

19



Europäisches Patentamt  
European Patent Office  
Office européen des brevets



11 Veröffentlichungsnummer: **0 272 723 B1**

12

## EUROPÄISCHE PATENTSCHRIFT

45

Veröffentlichungstag der Patentschrift: **21.10.92**

51

Int. Cl.<sup>5</sup>: **G10L 3/00**

21

Anmeldenummer: **87202298.3**

22

Anmeldetag: **24.11.87**

54

**Verfahren zur Bestimmung des zeitlichen Verlaufs eines Sprachparameters und Anordnung zur Durchführung des Verfahrens.**

30

Priorität: **26.11.86 DE 3640355**

43

Veröffentlichungstag der Anmeldung:  
**29.06.88 Patentblatt 88/26**

45

Bekanntmachung des Hinweises auf die  
Patenterteilung:  
**21.10.92 Patentblatt 92/43**

84

Benannte Vertragsstaaten:  
**DE FR GB IT SE**

56

Entgegenhaltungen:

**PROCEEDINGS OF THE 6th INTERNATIONAL  
CONFERENCE - PATTERN RECOGNITION,  
München, 19.-20. Oktober 1982, Seiten  
1119-1124, IEEE, New York, US; H. NEY:  
"Dynamic programming as a technique for  
pattern recognition"**

**IEEE TRANSACTIONS ON SYSTEMS, MAN,  
AND CYBERNETICS, Band SMC-12, Nr. 3,  
Mai/Juni 1982, Seiten 383-388, IEEE, New  
York, US; H. NEY: "A time warping approach  
to fundamental period estimation"**

73

Patentinhaber: **Philips Patentverwaltung  
GmbH  
Wendenstrasse 35 Postfach 10 51 49  
W-2000 Hamburg 1(DE)**

84

Benannte Vertragsstaaten:  
**DE**

73

Patentinhaber: **N.V. Philips' Gloeilampenfa-  
brieken  
Groenewoudseweg 1  
NL-5621 BA Eindhoven(NL)**

84

Benannte Vertragsstaaten:  
**FR GB IT SE**

72

Erfinder: **Ney, Hermann  
Gandersheimer Weg 6  
W-2000 Hamburg 61(DE)**

74

Vertreter: **Poddig, Dieter, Dipl.-Ing. et al  
Philips Patentverwaltung GmbH Wenden-  
strasse 35 Postfach 10 51 49  
W-2000 Hamburg 1(DE)**

**EP 0 272 723 B1**

Anmerkung: Innerhalb von neun Monaten nach der Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents kann jedermann beim Europäischen Patentamt gegen das erteilte europäische Patent Einspruch einlegen. Der Einspruch ist schriftlich einzureichen und zu begründen. Er gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist (Art. 99(1) Europäisches Patentübereinkommen).

## Beschreibung

Die Erfindung betrifft ein Verfahren zur Bestimmung des zeitlichen Verlaufs eines Sprachparameters eines Sprachsignals, wobei zu diskreten Zeitpunkten für jeden Wert eines vorgegebenen Wertebereichs des Sprachparameters ein Einzelwert vorhanden ist und der zeitliche Verlauf die Folge der benachbarten, einschließlich diagonal benachbarten Sprachsignalparameterwerte darstellt, deren Einzelwert zumindest nahe den Extremwerten der Einzelwerte für die einzelnen Zeitpunkte liegen, wobei die Summe der Einzelwerte dieser Folge einen Summen-Extremwert gegenüber anderen Folgen bildet.

Ein Sprachparameter kann beispielsweise die Grundfrequenz oder eine Formante eines zu untersuchenden Sprachsignals sein. Andere Sprachparameter sind beispielsweise LPC-Koeffizienten (Linearprädiktion-Koeffizient).

Die Einzelwerte beispielsweise der Sprachgrundfrequenz können mit Hilfe AMDF (Average Magnitude Difference Function) bestimmt werden. Dafür wird das Sprachsignal beispielsweise mit einer Abtastperiode von 10 kHz abgetastet, und eine bestimmte Anzahl aufeinanderfolgender Abtastwerte, die also insgesamt einen Sprachsignalabschnitt darstellen, werden gegenüber dem Sprachsignal schrittweise um eine Anzahl Abtastpunkte verschoben, und die Differenz der Abtastwerte des unverschobenen und des verschobenen Signals werden für die einzelnen Verschiebungsschritte aufsummiert. Die Verschiebung, die den kleinsten Summenwert ergibt, bezeichnet im allgemeinen die Periode der Sprachgrundfrequenz. Diese Werte sind jedoch nicht immer ganz eindeutig, es können bei Perioden von Oberwellen bzw. Formanten kleine Summenwerte auftreten, und es gibt noch weitere Einflüsse, die die korrekte Ermittlung der Sprachgrundfrequenz verfälschen. Die AMDF ergibt jedoch für jeden der aufeinanderfolgenden Sprachabschnitte, die gegen das Sprachsignal verschoben werden, bei verschiedenen Verschiebungen jeweils einen Wert, der eine gewisse Wahrscheinlichkeit bzw. genauer Unwahrscheinlichkeit dafür angibt, daß die betreffende Verschiebung die Grundperiode des Sprachsignals angibt.

Für viele Untersuchungen, beispielsweise für die Spracherkennung oder auch die Sprechererkennung, ist es erforderlich, einen zusammenhängenden Verlauf eines oder mehrerer Sprachparameter wie die Sprachgrundfrequenz über die Zeit zu erhalten, um diesen Verlauf mit Musterverläufen vergleichen zu können. Die ermittelten AMDF-Werte müssen also von Ausreißern befreit und im übrigen geglättet werden. Eine Glättung durch ein beispielsweise lineares Filter ergibt jedoch häufig zu große Abweichungen bzw. Verfälschungen von ei-

nem optimalen bzw. wahrscheinlichsten Verlauf. Ein solcher Verlauf ist am besten durch diejenige Folge der Sprachparameterwerte, deren Einzelwert zumindest nahe an dem Minimum bei dem betreffenden Zeitpunkt liegen, charakterisiert, wenn die Summe dieser Einzelwerte insgesamt ein Minimum gegenüber anderen Folgen bilden.

In IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, Band SMC-12, Nr. 3, Mai/Juni 1982, Seiten 383-388 ist die AMDF erläutert und davon ausgehend beschrieben, wie zur Bestimmung der Sprachgrundfrequenz bei der Autokorrelation gegeneinander gegeneinander verschobener Sprachsignalabschnitte das eine zu korrelierende Sprachsignal zeitverzerrt werden kann, um eine optimale Übereinstimmung, d.h. minimale Werte der AMDF zu erzielen. Die Grundfrequenz wird für jeden Sprachabschnitt für das Minimum des AMDF-Wertes angenommen. Aufeinanderfolgende Sprachsignalabschnitte ergeben dabei jedoch Grundfrequenzwerte, die nicht immer aneinander anschließen, d.h. der zeitliche Verlauf der so gewonnenen Sprachgrundfrequenz stellt keine glatte Kurve dar.

Das Glätten des Verlaufs der Sprachgrundfrequenz, unabhängig davon, wie die einzelnen Werte gewonnen werden, ist in PROCEEDINGS OF THE 6th INTERNATIONAL CONFERENCE - PATTERN RECOGNITION, München, 19.-20. Oktober 1982, Seiten 1119-1124, IEEE, New York, US beschrieben. Mit dem dort angegebenen Verfahren werden jedoch nur Ausreißer in dem Kurvenverlauf eliminiert, so daß dann Lücken in der Kurve übrig bleiben. Auch damit ist es also nicht ohne weiteres möglich, einen optimal geglätteten zeitlichen Verlauf der Folge von Sprachparametern zu erzielen.

Aufgabe der Erfindung ist es daher, ein Verfahren der eingangs genannten Art anzugeben, das diesen wahrscheinlichsten Verlauf eines Sprachparameters möglichst gut annähert.

Diese Aufgabe wird erfindungsgemäß dadurch gelöst, daß bezogen auf das Minimum als Extremwert der Einzelwerte und als Summen-Extremwert zu jedem Zeitpunkt  $i$  in einem ersten Durchlauf nacheinander für alle in der einen Richtung aufeinanderfolgenden Sprachparameterwerte  $k$  ein erster Richtungswert  $D'(k,i)$  als die Summe aus dem betreffenden Einzelwert und dem Minimum der folgenden Werte

$$\begin{aligned} D(k,i-1) \\ D(l,i-1) + a(d(k,i) + A) \\ D'(l,i) + b(d(k,i) + A) \end{aligned}$$

sowie ein Hinweiswert  $h'(k,i)$  auf den das Minimum liefernden Einzelwert gebildet und abgespeichert wird, wobei

$D(k,i-1)$  bzw.  $D(l,i-1)$  einen beim jeweils vorherge-

henden Zeitpunkt  $i-1$  und denselben Sprachparameterwert  $k$  bzw. dem vorhergehenden Sprachparameterwert  $l$  erzeugten und gespeicherten Summenwert,

$D(l,i)$  der beim unmittelbar vorhergehenden Sprachparameterwert gebildete Richtungswert und  $a, b, A$  vorgegebene feste Größen sind, wobei  $a$  und  $b$  vorzugsweise zwischen  $0,5$  und  $2,0$  liegen und  $A$  den Einzelwerten in stimmlosen Sprachsignalsegmenten und in Sprachpausen entspricht, daß danach in einem zweiten Durchlauf für den gleichen Zeitpunkt  $i$  für alle in der anderen Richtung aufeinanderfolgenden Sprachparameterwerte  $k$  ein zweiter Richtungswert  $D''(k,i)$  mit einem Hinweiswert  $h''(k,i)$  in entsprechender Weise wie der erste Richtungswert gebildet wird, daß für jeden Sprachparameterwert  $k$  das Minimum der beiden Richtungswerte  $D', D''$  und der zu diesem Richtungswert gehörende Hinweiswert  $h', h''$  als jeweils neuer Summenwert  $D$  bzw. als Gesamthinweiswert  $H$  gespeichert wird, und daß spätestens am Ende des Sprachsignals aus dem Gesamthinweiswert  $H(k,l)$ , der zu dem Sprachparameterwert  $k$  mit dem Summen-Extremwert beim letzten Zeitpunkt  $l$  gehört, der unmittelbar vorhergehende Sprachparameterwert ermittelt und der zugehörige Gesamthinweiswert ausgelesen wird usw., wobei die dabei entstehende Folge der Sprachparameterwerte ausgegeben bzw. gespeichert wird.

Durch das erfindungsgemäße Verfahren werden also mit Mitteln der sogenannten dynamischen Programmierung Folgen von Sprachparameterwerten ermittelt, die am Ende des Sprachsignals verschiedene Werte der Summen der Einzelwerte ergeben, und die Folge, die am Ende des Sprachsignals das Minimum des Summenwerts ergibt, wird als wahrscheinlichste Folge angesehen. Durch die Speicherung der Gesamthinweiswerte für jeden Sprachparameter zu jedem Zeitpunkt kann diese Folge dann zurückverfolgt werden.

Bei der Bildung der Summenwerte für die einzelnen Folgen von Sprachparameterwerten werden bestimmte vorgegebene feste Größen verwendet, von denen die Größen  $a$  und  $b$  überwiegend die Glätte des Verlaufs beeinflussen, indem ein Verlauf in diagonalen Richtung und in senkrechter Richtung erschwert wird, und dafür haben diese beiden Größen  $a$  und  $b$  zweckmäßig Werte zwischen  $0,5$  und  $2,0$ . Die Größe  $A$  und entspricht etwa dem Wert, den die AMDF in stimmlosen Sprachabschnitten und Pausenabschnitten aufweist, und erzwingt einen im wesentlichen waagerechten Verlauf des Sprachparameters in diesen Abschnitten.

Die mindestens zwei Durchläufe für die getrennte Ermittlung der beiden Richtungswerte, d.h. einmal in Aufwärtsrichtung und einmal in Abwärtsrichtung oder umgekehrt, sind erforderlich, da für jeden Richtungswert für einen möglichen Verlauf in

senkrechter Richtung auch der in der betreffenden Richtung unmittelbar vorhergehende Richtungswert berücksichtigt werden muß und somit vorher ermittelt sein muß.

Die bei dem erfindungsgemäßen Verfahren gebildeten verschiedenen Werte müssen gespeichert werden, da sie anschließend noch weiter verwendet werden. Eine genaue Überprüfung des Ablaufs des erfindungsgemäßen Verfahrens ergibt jedoch, daß verschiedene Werte nur während eines relativ kurzen Zeitabschnittes benötigt werden. Eine Ausgestaltung der Erfindung ist daher dadurch gekennzeichnet, daß die Richtungswerte  $D', D''$  und die Hinweiswerte  $h', h''$  und die neuen Summenwerte für jeweils nur einen Zeitpunkt gespeichert und danach wieder überschrieben werden. Für alle Sprachparameter aller Zeitpunkte müssen also nur die Gesamthinweiswerte gespeichert werden, da nur diese für das Zurückverfolgen der am Schluß als optimal ermittelten Folge erforderlich sind, während die anderen Werte jeweils nur für einen Zeitpunkt gespeichert werden, da sie danach nicht mehr benötigt werden. Insbesondere ist bei der Ermittlung der Summenwert aller Sprachparameter nur für den vorhergehenden Zeitpunkt erforderlich, während die ermittelten Richtungswerte nur für den momentanen Zeitpunkt benötigt werden. Dadurch kann sehr viel Speicherplatz eingespart werden.

Die Bestimmung des jeweils neuen Summenwerts aus dem Minimum der beiden Richtungswerte kann in einem weiteren Durchlauf erfolgen, nachdem beide Richtungswerte für jeden Sprachparameter eines Zeitpunktes ermittelt worden sind. In vielen Fällen ist es jedoch auch möglich, den jeweils neuen Summenwert unmittelbar nach der Ermittlung des zweiten Richtungswert für jeweils einen Sprachparameter zu ermitteln und damit den alten Summenwert zu überschreiben, da dieser dann nicht mehr benötigt wird. Eine weitere Ausgestaltung der Erfindung ist daher dadurch gekennzeichnet, daß die Richtungswerte  $D', D''$  und die Hinweiswerte  $h', h''$  nur für die in der einen Richtung aufeinanderfolgenden Sprachparameterwerte  $k$  gespeichert werden und daß nach der Bildung jedes Richtungswerts  $D'', D'$  für die in der anderen Richtung aufeinanderfolgenden Sprachparameterwerte  $k$  unmittelbar der neue Summenwert  $D$  und der Gesamthinweiswert  $H$  gebildet gespeichert werden. Auf diese Weise kann der erforderliche Speicherbedarf noch weiter verringert werden.

Eine Anordnung zur Durchführung des erfindungsgemäßen Verfahrens ist gekennzeichnet durch

- einen ersten Speicher für die Einzelwerte  $d(k,i)$  aller Sprachparameterwerte  $k$  mindestens jeweils eines Zeitpunktes  $i$ ,
- einen zweiten Speicher für jeweils einen Summenwert  $D(k,i)$  für jeden Sprachparameterwert  $k$  zu jedem Zeitpunkt  $i$ .

terwert k mindestens jeweils eines Zeitpunktes i,

- einen dritten Speicher für jeweils einen ersten Richtungswert  $D'(k,i)$  und einen Hinweiswert  $h'(k,i)$  für jeden Sprachparameterwert k mindestens jeweils eines Zeitpunktes i,
- einen vierten Speicher für einen zweiten Richtungswert  $D''(k,i)$  und einen zugehörigen Hinweiswert  $h''(k,i)$  mindestens eines Sprachparameterwertes k für denselben Zeitpunkt i wie die Werte im dritten Speicher,
- einen fünften Speicher für die Gesamthinweiswerte  $H(k,i)$  für alle Sprachparameterwerte k und alle Zeitpunkte i,
- eine Verarbeitungsanordnung mit Eingängen, die mit einem Datenausgang des ersten, des zweiten und des dritten bzw. vierten Speichers gekoppelt sind, und mit einem Ausgang zum Abgeben jeweils eines Richtungswertes  $D'(k,i), D''(k,i)$  und eines zugehörigen Hinweiswertes  $h'(k,i), h''(k,i)$ , der mit dem Dateneingang des dritten Speichers (30) gekoppelt ist,
- einen Vergleicher mit zwei Eingängen, von denen der eine vom Datenausgang des dritten Speichers jeweils einen ersten Richtungswert  $D'(k,i)$  und der andere Eingang den zugehörigen zweiten Richtungswert  $D''(k,i)$  empfängt, und mit einem Ausgang, der einen Umschalter steuert, der den kleineren der beiden Richtungswerte dem Dateneingang des zweiten Speichers zuführt, und
- eine Steueranordnung mit einem Adressengenerator, der zyklisch die Adressen aller Sprachparameterwerte in der einen Richtung der Adressenfolge und danach in der anderen Richtung der Adressenfolge für mindestens den ersten bis dritten und den fünften Speicher erzeugt und der außerdem die Adressenauswahl für den Zeitpunkt mindestens für den fünften Speicher erzeugt, und mit einer Ablaufsteuerung, die das Einschreiben und Auslesen mindestens des zweiten, dritten und fünften Speichers steuert. Auf diese Weise läßt sich das erfindungsgemäße

Verfahren mit relativ geringem Aufwand realisieren.

Ausführungsbeispiele der Erfindung werden nachstehend anhand der Zeichnung erläutert. Es zeigen:

- Fig. 1 ein Beispiel für einen möglichen Verlauf eines Sprachparameters über die aufeinanderfolgenden Zeitpunkte eines Sprachsignals,
- Fig. 2 ein Diagramm zur Erläuterung, wie die einzelnen Richtungswerte bestimmt werden,
- Fig. 3 ein Schema für den Ablauf des erfindungsgemäßen Verfahrens,

Fig. 4 eine Schaltungsanordnung zur Durchführung des erfindungsgemäßen Verfahrens,

Fig. 5 eine mögliche Ausführung der Verarbeitungsanordnung darin.

In Fig. 1 ist in einer zweidimensionalen Darstellung der Verlauf eines Sprachparameters über die Zeit schematisch dargestellt. Als Beispiel werde angenommen, daß es sich bei dem Sprachparameter um die Stimmgrundfrequenz handelt. Die über die Ordinate aufgetragenen Sprachparameterwerte k, die von 1 bis K laufen, stellen also verschiedene diskrete Frequenzwerte dar, während über die Abszisse die Zeit in Form von diskreten Zeitpunkten i aufgetragen ist, die von 1 bis I am Ende des Sprachsignals reichen. Es ist klar, daß die Anzahl der Werte in beiden Koordinatenrichtungen in der Praxis tatsächlich wesentlich größer ist.

Zu jedem Zeitpunkt i gibt es für jeden Sprachparameterwert k, d.h. für jeden durch einen kleinen Kreis angegebenen Kreuzungspunkt der beiden Koordinaten, einen Einzelwert  $d(k,i)$ . Dieser kann beispielsweise dadurch gewonnen, daß das Sprachsignal mit einer hohen Frequenz, beispielsweise 10 kHz abgetastet wird. Jeweils eine Anzahl, beispielsweise 100 bis 200, Abtastwerte ergeben ein Sprachsegment, das also eine Dauer von 10 bis 20 msec hat. Die Abtastwerte des Sprachsignals werden dann mit  $s(i,j)$  bezeichnet, wobei der Index i die Sprachsegmente und der Index j die Abtastwerte in einem Sprachsegment angibt. rch Anwendung der AMDF (Average Magnitude Difference Function) ergeben sich dann die Einzelwerte:

$$d(k,i) = \sum_j |s(i, j+k) - s(i, j)|$$

d.h. das Sprachsegment wird um eine Anzahl k Abtastwerte gegenüber dem Sprachsignal verschoben und die Differenzen zwischen korrespondierenden Abtastwerten innerhalb des Sprachsignals werden betragsmäßig aufsummiert. Jeder Verschiebung k entspricht damit eine bestimmte Frequenz. Auf diese Weise entsteht für jeden Wert von k innerhalb eines vorgegebenen Wertebereichs, der den praktisch vorkommenden Sprachgrundfrequenzen entspricht, ein Einzelwert  $d(k,i)$ . Diese Einzelwerte können als eine Art von Wahrscheinlichkeit bzw. Unwahrscheinlichkeit aufgefaßt werden, daß die durch den Sprachparameterwert k angegebene Frequenz tatsächlich die Grundfrequenz des Sprachsignals in diesem Sprachsegment ist, so daß der Einzelwert für den Sprachparameterwert k entsprechend der tatsächlichen Grundfrequenz ein Minimum ist.

In Fig. 1 sind bestimmte, jeweils einem Einzel-

wert entsprechende Kreuzungspunkte der beiden Koordinaten  $k$  und  $i$  durch eine Linie miteinander verbunden, um ein Beispiel für einen möglichen zeitlichen Verlauf der Stimmgrundfrequenz zu zeigen. Die Linie verbindet eine derartige Folge von Einzelwerten, so daß die Summe dieser so verbundenen Einzelwerte ein Minimum gegenüber jeder anderen Verbindung ergibt, wobei diese Verbindung nur waagrecht, diagonal oder senkrecht verlaufen kann. Ein senkrechter Verlauf der Sprachgrundfrequenz kann in der Praxis zwar nicht auftreten, jedoch handelt es sich hier um zeitdiskrete Werte, so daß eine schnelle Änderung der Sprachgrundfrequenz zwischen zwei aufeinanderfolgenden Zeitpunkt  $i$  in dem diskreten Modell der Fig. 1 durch einen senkrechten Verlauf angenähert werden muß. Die Einschränkung, daß nur ein waagerechter, diagonal oder senkrechter Verlauf zugelassen wird, d.h. daß die Linie nur unmittelbar benachbarte Einzelwerte miteinander verbindet, ergibt eine glattere Kurve, jedoch kann es dabei vorkommen, daß die Linie nicht genau die Minima der Einzelwerte jedes Zeitpunktes  $i$  miteinander verbindet, sondern es wird eine Verbindung gebildet, die die Abweichungen von den minimalen Einzelwerten, d.h. die Fehler minimiert. Dabei werden Ausreißer von Einzelwerten an einzelnen Zeitpunkten, die durch die komplexe Form des Sprachsignals und durch andere Einflüsse möglich sind, eliminiert.

Die Ermittlung desjenigen Verlaufs, der am Ende die kleinste Gesamtsumme der Einzelwerte ergibt, wird anhand der Fig. 2 näher erläutert. Da ein senkrechter Verlauf sowohl von oben nach unten als auch von unten nach oben möglich ist, wie aus Fig. 1 hervorgeht, wird der jeweilige Summenwert der Einzelwerte für beide Richtungen getrennt bestimmt, und danach wird mit dem kleineren der sich dabei ergebenden Summenwerte, die im folgenden mit Richtungswerte bezeichnet werden, weitergegangen.

Um dies leichter erläutern zu können, ist in Fig. 2 ein momentan bearbeiteter Punkt  $k,i$ , d.h. der Sprachparameterwert  $k$  zum Zeitpunkt  $i$ , doppelt dargestellt. Die waagerechte Verbindung zum gleichen Sprachparameterwert  $k$  des vorhergehenden Zeitpunktes  $i-1$  verläuft dadurch lediglich aus zeichnerischen Gründen etwas schräg.

Zunächst werde der Richtungswert  $D^+(k,i)$  für die aufsteigende Richtung bestimmt, wofür von jedem der drei benachbarten Punkte der dafür gebildete Summenwert bzw. Richtungswert verwendet und um den Einzelwert  $d(k,i)$  an der betreffenden Stelle erhöht wird. Da ein waagerechter Verlauf wahrscheinlicher ist als ein schräger Verlauf oder insbesondere ein senkrechter Verlauf, werden die Werte  $D(k-1,i-1)$  des diagonal benachbarten Punktes und  $D^+(k-1,i)$  des senkrecht darunterliegenden Punktes nicht direkt verwendet, sondern um einen

bestimmten Wert erhöht, wie später im einzelnen erläutert wird. Der Wert  $D^+(k-1,i)$  des senkrecht darunterliegenden Punktes muß vorher ermittelt worden sein, und für diesen muß wiederum der Wert des darunterliegenden Punktes ermittelt sein, usw. bis zum untersten Punkt, so daß die Richtungswerte  $D^+(k,i)$  ausgehend von  $k=1$  ermittelt werden müssen. Entsprechendes gilt jedoch auch für den Richtungswert  $D^-(k,i)$  für die Richtung nach unten, für den der Summenwert  $D(k+1,i-1)$  des diagonal darüberliegenden Punktes und der Richtungswert  $D^-(k+1,i)$  des senkrecht darüberliegenden Punktes benötigt wird, wobei letzterer vorher ermittelt sein muß, so daß in dieser Richtung beim Wert  $k=K$  begonnen werden muß. Daraus ergibt sich, daß es zweckmäßig ist, zuerst alle Richtungswerte der einen Richtung, beispielsweise die Richtungsweise  $D^+$  zu bestimmen und zwischenspeichern, und dann erst die Richtungswerte für die andere Richtung, also in diesem Falle die Richtungswerte  $D^-$ , und dann erst das Minimum der beiden Richtungswerte als neuen Summenwert  $D$  zu bestimmen. Eine Möglichkeit, daß nur alle Richtungswerte, z.B.  $D^+$ , der einen Richtung ermittelt und zwischengespeichert werden müssen, wird später erläutert.

In der anhand der Fig. 2 erläuterten Weise werden also jeweils für einen Zeitpunkt  $i$  die Richtungswerte  $D^+$  und die Richtungswerte  $D^-$  beider Richtungen für alle Sprachparameterwerte bestimmt und daraus jeweils das Minimum als neuer Summenwert  $D$  ermittelt und gespeichert. Beim letzten Zeitpunkt  $I$  gibt dann der Sprachparameterwert  $k$ , bei dem dieser Summenwert  $D$  von allen anderen Summenwerten zu diesem Zeitpunkt am kleinsten ist, das Ende der Folge an. Damit von diesem Ende ausgehend der Verlauf der Folge zurückverfolgt werden kann, muß bei jedem Richtungswert und dann beim Summenwert gespeichert werden, von welchem vorhergehenden Punkt ausgehend dieser Richtungswert bzw. Summenwert erreicht worden ist. Dadurch kann beim minimalen Summenwert  $D$  beim letzten Zeitpunkt  $I$  festgestellt werden, von welchem Punkt dieser erreicht worden ist, d.h. welches der vorhergehende Punkt ist, und aus dem für diesen Punkt gespeicherten Wert kann der diesem vorhergehende Punkt ermittelt werden, usw., so daß sich schließlich die gesamte Folge der Sprachparameterwerte ergibt, die den minimalen Summenwert  $D$  beim letzten Zeitpunkt  $I$  ergeben hat. Die Werte, die auf die Richtung hinweisen, in der der Punkt liegt, aus dem jeweils der folgende Punkt erreicht worden ist, werden mit Hinweiswert bzw. Gesamthinweiswert bezeichnet und müssen selbstverständlich für jeden Sprachparameterwert  $k$  zu jedem Zeitpunkt  $i$  gespeichert werden. Da es jedoch nur fünf Richtungen gibt, sind dafür lediglich drei Bit erforderlich. Alle anderen Werte müs-

sen lediglich höchstens für alle Sprachparameterwerte  $k$  eines Zeitpunktes  $i$  bzw.  $i-1$  gespeichert werden.

Übrigens ist im jeden der beiden Richtungs-  
werte  $D^+$  und  $D^-$  die waagerechte Richtung enthal-  
ten, so daß, wenn diese Richtung den kleinsten  
Richtungswert ergibt, beide Richtungswerte gleich  
sind. Für den einen Richtungswert könnte also die  
waagerechte Richtung weggelassen werden.

In Fig. 3 ist der gesamte Ablauf der einzelnen  
Verarbeitungsschritte zur Ermittlung des zeitlichen  
Verlaufs der Sprachparameterwerte schematisch  
angegeben. Der Block 101 bezeichnet übliche Ein-  
stellungen von Anfangszuständen wie Rückstellen  
von Zählern und Löschen von Speicherbereichen,  
die nicht gesondert angegeben sind. Lediglich das  
Füllen der Speicherplätze für die Summenwerte  $D$ -  
( $k,1$ ) für den ersten Zeitpunkt  $i=1$  mit den entspre-  
chenden Einzelwerten  $d(k,1)$  ist angegeben.

Danach wird im Block 102 ein Zähler, der den  
momentanen Zeitpunkt  $i$  angibt, auf den nächsten  
Zeitpunkt  $i+1$  geschaltet. Danach werden im Block  
103 die aufsteigenden Richtungswerte  $D^+(k)$  in fol-  
gender Weise bestimmt:

$$D^+(k,i) = d(k,i) + \min\{D(k,i-1), D(k-1,i-1) + a [d(k,i) + A], D^+(k-1,i) + b [d(k,i) + A]\} \quad (1)$$

Diese Gleichung bedeutet gemäß Fig. 2, daß  
der Einzelwert  $d(k,i)$  addiert wird zu dem Minimum  
der Summenwerte der drei benachbarten Punkte,  
wobei für die waagerechte Richtung der Summen-  
wert  $D(k,i-1)$  direkt verwendet wird, während bei  
den anderen Werten noch ein Term addiert wird,  
der von dem Einzelwert des gerade betrachteten  
Punktes und von festen Größen  $a$  bzw.  $b$  und  $A$   
abhängt. Die Größe  $A$  entspricht den Einzelwerten  
in stimmlosen Sprachsignalsegmenten und in Pau-  
sensegmenten und führt dazu, daß in diesen Berei-  
chen der Verlauf praktisch stets waagerecht ist. Mit  
den Größen  $a$  und  $b$  wird die Glätte des Verlaufs  
beeinflusst, d.h. je größer  $a$  und  $b$  ist, desto stärker  
werden die Diagonale und die senkrechte Richtung  
benachteiligt. Es sind daher empirisch gewonnene  
Größen, die im allgemeinen bei Sprachsignalen  
zwischen 0,5 und 2,0 liegen. Im Block 103 wird  
außerdem noch der Hinweiswert  $h^+(k,i)$  ermittelt,  
der angibt, von welchem vorhergehenden Punkt  
der momentane Punkt erreicht worden ist, d.h. wel-  
cher der drei Terme, von denen beim Richtungs-  
wert  $D^+(k,i)$  das Minimum bestimmt wird, dieses  
Minimum ergeben hat. Außerdem wird ein Zähler  
für den Sprachparameterwert  $k$  um 1 auf  $k+1$   
erhöht.

Der Block 103 wird also nacheinander für alle  
Sprachparameterwerte  $k$  eines bestimmten Zeit-  
punktes  $i$  durchgeführt.

Wenn für alle Werte  $k$  der Richtungswert und

der Hinweiswert bestimmt und zwischengespei-  
chert ist, d.h.  $k=K$ , wird gemäß Block 104 der  
andere Richtungswert  $D^-(k,i)$  für die andere Rich-  
tung in folgender Weise bestimmt:

$$D^-(k,i) = d(k,i) + \min\{D(k,i-1), D(k+1,i-1) + a [d(k,i) + A], D^-(k+1,i) + a [d(k,i) + A]\} \quad (2)$$

Hierfür gilt das gleiche wie für die Berechnun-  
gen des Blocks 103, es werden lediglich die in der  
anderen Richtung vorangehenden Punkte berück-  
sichtigt, d.h.  $k-1$  ist ersetzt durch  $k+1$ . In entspre-  
chender Weise wird auch der Hinweiswert  $h^-(k,1)$   
bestimmt. Da dies die Richtungswerte für die ab-  
wärts zeigende Richtung sind, wobei die Berechnung  
mit dem höchsten Sprachparameterwert  $K$  beginnt,  
wird hier jeweils nach der Bestimmung  
des Richtungswerts und des zugehörigen Hinweis-  
werts der Zähler für die Sprachparameter  $k$  um 1  
auf  $k-1$  verringert.

Wenn alle Richtungswerte und Hinweiswerte  
auch der anderen Richtung ermittelt sind, d.h. auch  
für den ersten Sprachparameterwert  $k=1$ , wird ge-  
mäß Block 105 von den zwei Richtungswerten  $D^+$ -  
( $k,i$ ) und  $D^-(k,i)$  aller Sprachparameter  $k$  jeweils das  
Minimum ermittelt und als Summenwert  $D(k,i)$  ge-  
speichert, und außerdem wird jeweils der zugehörige  
Hinweiswert  $h^+(k,i)$  bzw.  $h^-(k,i)$  als Gesamthin-  
weiswert  $H(k,i)$  gespeichert. Wenn dies für alle  
Sprachparameterwerte  $k$  erfolgt ist, d.h.  $k=K$ , wird  
auf den Block 102 zurückgegangen, wo der Zähler  
für den Zeitpunkt  $i$  auf den nächsten Wert  $i+1$   
eingestellt wird, und der beschriebene Vorgang  
wird wiederholt.

Die in den Blocks 104 und 105 getrennt ange-  
gebenen Vorgänge können auch gemeinsam  
durchgeführt werden, d.h. nach jeder Ermittlung  
eines Richtungswerts  $D^-(k,i)$  für einen Sprachpara-  
meterwert  $k$  kann unmittelbar danach festgestellt  
werden, ob dieser oder der im vorhergehenden  
Durchlauf ermittelte Richtungswert  $D^+(k,i)$  kleiner  
ist, und der kleinere von beiden wird als neuer  
Summenwert  $D(k,i)$  gespeichert. Dabei muß aller-  
dings der bisherige Summenwert  $D(k,i-1)$  zwischen-  
gespeichert werden, da er für den folgenden Rich-  
tungswert  $D^-(k-1,i)$  noch benötigt werden. Außer-  
dem sind dann für jeden Sprachparameterwert  $k$   
eine größere Anzahl Schritte durchzuführen.

Wenn für alle Zeitpunkte  $i$  die Summenwerte  
 $D(k,i)$  und insbesondere die Hinweiswerte  $H(k,i)$  be-  
stimmt worden sind, d.h. es ist der letzte Zeitpunkt  
 $i=I$  erreicht, werden die Vorgänge gemäß Block  
106 ausgelöst. Danach wird zunächst der Sum-  
menwert  $D(m,I)$  bestimmt, der das Minimum aller  
Summenwerte  $D(k,I)$  beim letzten Zeitpunkt  $I$  dar-  
stellt. Danach wird der zu diesem minimalen Sum-  
menwert  $D(m,I)$  gehörende Gesamthinweiswert  $H$ -  
( $m,I$ ) ausgelesen und daraus der dem Punkt  $m,I$

vorhergehende Punkt  $k_1, i_1$  bestimmt. Der an dieser Stelle gespeicherte Hinweiswert  $H(k_1, i_1)$  wird dann ausgelesen und der diesem wiederum vorhergehende Punkt ermittelt, usw., bis der Anfang des auf diese Weise bestimmten Verlaufs des Sprachparameters erreicht ist. Die bei diesem Zurückverfolgen entstehende Folge von Punkten in Form von deren Koordinaten, d.h. dem Zeitpunkt  $i$  und dem Sprachparameter  $k$ , stellt die gesuchte Folge dar.

In Fig. 4 ist das Blockschaltbild einer Anordnung dargestellt, die die in Fig. 3 angegebenen Bearbeitungsschritte ausführt. Der Block 10 bezeichnet vorzugsweise einen Speicher, der alle Einzelwerte  $d(k, i)$  enthält und der von den Sprachparameterwerten  $k$  und den Werten entsprechend den Zeitpunkten  $i$  adressiert wird. Der Block 10 muß zumindest die Einzelwerte für alle Sprachparameter  $k$  eines Zeitpunktes  $i$  enthalten, da diese jeweils zweimal, nämlich für beide Richtungswerte  $D^+$  und  $D^-$  benötigt werden. Der Block 10 kann auch die Anordnung zur Erzeugung der Einzelwerte einschließen, die jedoch nicht Bestandteil der Erfindung ist und daher nicht näher ausgeführt ist.

Der Block 20 stellt einen Speicher dar, der die Summenwerte  $D(k, i-1)$  des jeweils vorhergehenden Zeitpunktes  $i-1$  zu Beginn eines neuen Zeitpunktes  $i$  enthält und der am Ende dieses neuen Zeitpunktes die neuen Summenwerte  $D(k, i)$  enthält. Das Erzeugen und Einschreiben dieser Summenwerte wird später erläutert. Der Speicher 20 wird von den Sprachparameterwerten  $k$  adressiert, und über einen Eingang, der ein Signal  $d$  erhält, wird er auf das Einschreiben der über die Verbindung 35 zugeführten Werte umgeschaltet. Der Datenausgang 21 des Speichers 20 ist mit der Reihenschaltung von zwei Registern 22 und 24 verbunden, von denen das Register 22 den Wert vom Ausgang 21 übernimmt und seinen früheren Inhalt gleichzeitig in das andere Register 24 überträgt. Damit enthält das Register 22 jeweils den Summenwert  $D(k, i-1)$  und das Register 24 den vorhergehenden Summenwert  $D(k-1, i-1)$  bzw.  $D(k+1, i-1)$ , je nachdem welche Richtungswerte gerade berechnet werden. Ein weiteres Register 26 ist vorgesehen, das den auf der Verbindung 13 vorhandenen, gerade ermittelten Richtungswert  $D^+(k-1, i)$  bzw.  $D^-(k+1, i)$  aufnimmt und während der Ermittlung des jeweils folgenden Richtungswerts am Ausgang 27 bereitstellt.

Die Ausgänge 11 des Speichers 10 sowie 23, 25 und 27 der Register 22, 24 und 26 führen auf eine Verarbeitungsanordnung 12, die die in den Blöcken 103 und 104 in Fig. 3 angegebenen Berechnungen durchführt und am Ausgang 13 wie erwähnt den jeweils neuen Richtungswert  $D^+(k, i)$  bzw.  $D^-(k, i)$  und am Ausgang 19 den zugehörigen Hinweiswert  $h^-(k, i)$  bzw.  $h^+(k, i)$  erzeugt. Diese Werte werden einem Speicher 30 zugeführt, der eben-

falls von den Sprachparameterwerten  $k$  adressiert wird und der in einem ersten Durchlauf zu Beginn jedes neuen Zeitpunktes, bei dem beispielsweise die einen Richtungswerte  $D^+(k, i)$  und die zugehörigen Hinweiswerte  $h^+(k, i)$  für alle Sprachparameterwerte  $k$  erzeugt werden, durch ein Signal  $c$  an einem zusätzlichen Eingang auf Schreiben geschaltet ist und somit alle Richtungswerte und Hinweiswerte in diesem Durchlauf aufnimmt.

Wenn alle Richtungswerte und Hinweiswerte der einen Richtung erzeugt sind, werden die Speicher 10 und 20 in der umgekehrten Reihenfolge der Sprachparameter  $k$  adressiert, d.h. also ausgehend vom maximalen Wert  $k=K$  bis zum minimalen Wert  $k=1$ , wenn vorher die Adressierung vom minimalen Wert beginnend bis zum maximalen Wert gelaufen war. Die bei diesem zweiten Durchlauf von der Verarbeitungsanordnung 12 auf der Verbindung 13 erzeugten Richtungswerte  $D^-(k, i)$  und auf der Verbindung 19 erzeugten Hinweiswerte  $h^-(k, i)$  werden einem weiteren Speicher 40 zugeführt, der ebenfalls durch die Sprachparameterwerte  $k$  adressiert wird und durch ein Signal  $d$  an einem weiteren Eingang auf Schreiben eingestellt ist.

Nach dem Erzeugen beider Richtungswerte und Hinweiswerte für alle Sprachparameterwerte  $k$  eines Zeitpunktes und dem Einschreiben dieser Werte in die Speicher 30 bzw. 40 werden nochmals nacheinander alle Sprachparameterwerte  $k$  erzeugt und für jeden Wert  $k$  die Inhalte dieser Speicher 30 und 40 ausgelesen. Über die Leitungen 29 und 39 werden die beiden Richtungswerte  $D^+(k, i)$  und  $D^-(k, i)$  einem Vergleicher 14 zugeführt, der abhängig davon, welcher der beiden Richtungswerte kleiner ist, an seinem Ausgang 15 ein entsprechendes Signal abgibt. Wenn der Richtungswert auf der Verbindung 29 kleiner ist als der Richtungswert auf der Verbindung 39, erzeugt der Vergleicher 14 auf der Leitung 15 ein Signal, das die beiden Umschalter 32 und 34 in die linke Stellung umschaltet, so daß also der kleinere der beiden Richtungswerte über den Umschalter 34 und die Leitung 35 als neuer Summenwert dem Speicher 20 zugeführt und dort eingeschrieben wird, und gleichzeitig wird der auf der Verbindung 31 abgegebene zugehörige Hinweiswert als Gesamthinweiswert über den Schalter 32 einem Speicher 50 zugeführt und dort gespeichert. Die Speicheradresse für den Speicher 50 wird über den Umschalter 36 zugeführt, der zunächst durch ein Signal  $e$  in der unteren Stellung gehalten wird, so daß die Adresse aus dem Sprachparameterwert  $k$  und dem Wert  $i$  für den jeweiligen Zeitpunkt gebildet wird.

Wenn der Richtungswert auf der Verbindung 39 kleiner ist als der Richtungswert auf der Verbindung 29, erzeugt der Vergleicher 14 am Ausgang

15 ein Signal, das die beiden Umschalter 32 und 34 in die entgegengesetzte Lage umschaltet, so daß dann wieder der kleinere Richtungswert über die Verbindung 35 dem Speicher 20 und der zugehörige, auf der Verbindung 41 abgegebene Gesamthinweiswert über den Umschalter 32 dem Speicher 50 zugeführt wird. Im Speicher 20 werden durch die auf der Verbindung 35 zugeführten Minimum-Richtungswerte, die den neuen Summenwert  $D(k,i)$  darstellen, die entsprechenden früher gespeicherten Summenwerte desselben Sprachparameterwerts  $k$  überschrieben, so daß der Speicher 20 nur eine Kapazität von  $K$  Worten haben muß. Entsprechendes gilt auch für die Speicher 30 und 40. Lediglich der Speicher 50 muß für jeden Sprachparameterwert  $k$  jedes Zeitpunktes  $i$  einen Speicherplatz für einen Gesamthinweiswert aufweisen, der wegen der Tatsache, daß nur fünf verschiedene Richtungen möglich sind, nur drei Bit groß sein muß.

Wenn auf diese Weise das letzte Sprachsegment verarbeitet und damit der Zeitpunkt  $I$  erreicht ist, wird eine Anordnung aus einem Vergleicher 38 und einem Register 48 eingeschaltet, wodurch der minimale Summenwert  $D(m,I)$  bei diesem letzten Zeitpunkt  $I$  und der zugehörige Sprachparameterwert  $k = m$  dieses Minimums festgestellt wird. Dazu wird beispielsweise der Inhalt des Registers 48 auf seinem Maximalwert festgehalten und erst mit dem letzten Zeitpunkt  $I$  freigegeben, was der Übersichtlichkeit halber nicht näher dargestellt ist.

Der erste Summenwert  $D(k,I)$ , der zum Zeitpunkt  $I$  am Ausgang des Umschalters 34 auf der Verbindung 35 erscheint, wird dem einen Eingang des Vergleichers 38 sowie dem Dateneingang des Registers 48 zugeführt, dessen Ausgang 49 mit dem anderen Eingang des Vergleichers 38 verbunden ist. Da der auf der Verbindung 35 auftretende Summenwert kleiner sein wird als der Maximalwert, auf den das Register 38 vorher festgehalten wurde und der noch auf der Verbindung 49 vorhanden ist, erzeugt der Vergleicher 38 auf seiner Ausgangsleitung 37 ein Signal, das den momentanen Summenwert auf der Verbindung 35 in das Register 48 einschreibt, und außerdem wird in ein weiteres Register 42 der momentane Sprachparameterwert  $k$  und der Wert  $i$  des momentanen Zeitpunktes eingeschrieben. Wenn der nächste, am Ausgang des Schalters 34 auftretende Summenwert kleiner ist als der vorhergehende Wert, erzeugt der Vergleicher 38 wieder ein Signal auf der Ausgangsleitung 37, so daß dieser kleinere Summenwert in das Register 48 und die entsprechenden Werte  $k$  und  $I$  in das Register 42 eingeschrieben werden. Dies geschieht so lange, bis der am Ausgang des Schalters 34 erscheinende Summenwert nicht mehr kleiner ist als der vorhergehende Summenwert, so daß das Register 48 stets den kleinsten Summenwert

und das Register 42 die zugehörigen Werte  $k$  und  $i$  enthält. Dies wird fortgesetzt bis zum letzten Summenwert zum Zeitpunkt  $I$ , so daß dann der Speicher 50 alle Gesamthinweiswerte  $H(k,i)$  und das Register 48 den kleinsten aller Summenwerte zum Zeitpunkt  $I$  und das Register 42 die entsprechenden Werte  $k$  und  $I$  enthält. Damit ist das Ende der gesuchten Folge identifiziert.

Nun wird durch das Signal  $e$  der Umschalter 36 umgeschaltet, so daß der Speicher 50 von dem Ausgang 45 eines Adreßrechners 44 adressiert wird. Dieser erhält vom Register 42 über die Verbindung 43 die darin enthaltenen Werte  $m$  und  $I$  entsprechend dem minimalen Summenwert, und der an dieser Adresse im Speicher 50 enthaltene Gesamthinweiswert  $H(m,I)$  wird ausgelesen und über die Verbindung 51 dem Adressenrechner 44 zugeführt. Dieser modifiziert mit diesem Hinweiswert die über die Verbindung 43 zugeführten Werte, so daß am Ausgang 45 die Werte  $k_1$  und  $i_1$  des vorhergehenden Punktes erscheinen. Diese werden in das Register 42 eingeschrieben und adressieren gleichzeitig den Speicher 50, so daß der zugehörige Hinweiswert  $H(k_1,i_1)$  ausgelesen und über die Verbindung 51 dem Adressenrechner 44 zugeführt wird. Dieser erzeugt damit wieder die Werte des nächsten vorhergehenden Punktes auf der Verbindung 45, die wieder im Register 42 abgespeichert werden und gleichzeitig den Speicher 50 adressieren usw., so daß am Ausgang 46 schließlich die gesuchte Folge der Werte  $k$  und  $i$  in umgekehrter Reihenfolge erscheint. Diese können zwischengespeichert oder auch direkt verarbeitet werden, was nicht näher dargestellt ist, da es nicht mehr zur Erfindung gehört.

Es ist auch möglich, den Speicher 40 einzusparen, wobei dann die Verbindung 13 direkt mit der Verbindung 39 und die Verbindung 19 direkt mit der Verbindung 41 gekoppelt wird, entsprechend der Erläuterung bei Fig. 3, wenn die Blöcke 104 und 105 zusammengefaßt werden. Dann wird also, wenn die Richtungswerte für eine Richtung, beispielsweise die Richtungswerte  $D^+(k,i)$  und die zugehörigen Hinweiswerte  $h^+(k,i)$  für alle Sprachparameterwerte  $k$  erzeugt und im Speicher 30 gespeichert sind, jeder von der Verarbeitungsanordnung 12 auf der Verbindung 13 erzeugte Richtungswert  $D^-(k,i)$  direkt dem Vergleicher 14 zugeführt, der dann gleichzeitig vom Speicher 30 über die Verbindung 29 den anderen Richtungswert  $D^+(k,i)$  erhält, und der kleinere dieser beiden Richtungswerte wird dann unmittelbar über den Umschalter 34 und die Verbindung 35 dem Speicher 20 zugeführt und darin eingeschrieben. Gleichzeitig wird der zugehörige Hinweiswert über den Schalter 32 als Gesamthinweiswert  $H(k,i)$  in den Speicher 50 eingeschrieben. Dabei sind dann für jeden Zeitpunkt  $i$  nur zwei Durchläufe aller Sprachparameterwerte  $k$  erforder-

lich.

Eine Steueranordnung, die die dafür erforderlichen Werte  $k$  und  $i$  sowie die Steuersignale  $c, d$  und  $e$  erzeugt, ist im gestrichelten Block 16 dargestellt. Darin ist ein Taktgenerator 52 vorhanden, der einen Zähler 54 ansteuert, von dem jede mögliche Stellung einen anderen Wert des Sprachparameters  $k$  darstellt. Dieser Zähler 54 möge von 1 zu zählen beginnen, und wenn er den letzten Sprachparameterwert  $k=K$  erreicht hat, gibt er auf der Leitung 55 ein Signal ab, daß eine bistabile Kippstufe 56 umschaltet, so daß das Signal  $c$  verschwindet, das das Einschreiben in den Speicher 30 gesteuert hat, und das Signal  $d$  beginnt, mit dem die Summenwerte  $D(k,i)$  in den Speicher 20 und die Gesamthinweiswerte  $H(k,i)$  in den Speicher 50 eingeschrieben werden. Gleichzeitig wird die Zählrichtung des Zählers 54 umgekehrt, der nun wieder rückwärts zählt, so daß die Werte  $k$  an seinem Ausgang von  $K$  bis 1 laufen. Dies ist der zweite Durchlauf, bei dem die anderen Richtungswerte und die neuen Summenwerte und die Gesamthinweiswerte erzeugt und gespeichert werden.

Wenn der Zähler 54 wieder bei seiner Anfangsstellung angelangt ist, wird wieder auf der Leitung 55 ein Signal erzeugt, so daß die bistabile Kippstufe 56 wieder umschaltet und das Signal  $d$  verschwindet und das Signal  $c$  wieder beginnt, und gleichzeitig wird ein Zähler 58 um eine Stellung weitergeschaltet, dessen Ausgänge die Werte  $i$  liefern.

Wenn auf diese Weise der Zähler 58 schließlich alle Stellungen durchlaufen hat und den letzten Zeitpunkt  $I$  erreicht hat, erzeugt er anschließend, beispielsweise an einem Übertragsausgang, das Signal  $e$ , das den Umschalter 36 umschaltet, wie vorstehend beschrieben wurde, so daß das Zurückverfolgen der ermittelten Folge beginnen kann.

Die für die Speicher 20, 30 usw. sowie für die Register 22, 24 usw. und die Verarbeitungsanordnung 12 gegebenenfalls erforderliche Taktsteuerung, die in Fig. 4 der Übersichtlichkeit halber nicht dargestellt ist, kann ebenfalls von dem Taktgenerator 52 abgeleitet werden.

In Fig. 5 ist eine mögliche Ausführungsform der Verarbeitungsanordnung 12 in Fig. 4 dargestellt. Der über die Verbindung 11 zugeführte Einzelwert  $d(k,i)$  wird dem einen Eingang eines Addierers 60 zugeführt, dessen anderer Eingang den konstanten Wert  $A$  empfängt, der beispielsweise durch feste Verdrahtung vorgegeben ist. Die darin gebildete, auf der Verbindung 61 erscheinende Summe wird einem Multiplizierer 62 zugeführt, wo diese Summe mit dem festen Wert  $a$  multipliziert wird. Da der genaue Wert der Größe  $a$  nicht sehr kritisch ist, kann sie aus einer geringen Anzahl von einzelnen Summanden gebildet werden, die jeweils eine ganze negative Potenz von 2 sind, so daß der

Multiplizierer 62 aus einer kleinen Anzahl kaskadierter Addierer aufgebaut sein kann.

Das auf der Verbindung 63 entstehende Produkt wird einem Addierer 64 zugeführt, wo es zu dem Wert auf der Verbindung 25 des Registers 24 addiert wird, und dies ist der Summenwert  $D(k-1,i-1)$ , wenn die Sprachparameterwerte  $k$  ansteigend aufeinander folgen.

Der im Addierer 64 gebildete, auf der Verbindung 65 gelieferte Wert wird dem einen Eingang eines Vergleichers 66 zugeführt, wo er mit dem über die Verbindung 23 vom Register 22 zugeführten Wert verglichen wird, d.h. mit dem Summenwert  $D(k,i-1)$ . Der Vergleichsergebnis einen Umschalter 68, der den kleineren, dem Vergleichsergebnis zugeführten Wert über die Verbindung 69 dem einen Eingang eines weiteren Vergleichers 76 zuführt.

Der andere Eingang des Vergleichers 76 ist mit der Ausgangsverbindung 75 eines weiteren Addierers 74 verbunden, der den auf der Verbindung 27 vom Register 26 gelieferten Wert zu der im Multiplizierer 72 mit der Größe  $b$  multiplizierten Summe auf der Verbindung 61 addiert. Der Vergleichsergebnis steuert wieder einen Umschalter 78 derart, daß der kleinere der dem Vergleichsergebnis zugeführten Werte weitergeleitet und dem einen Eingang eines weiteren Addierers 70 zugeführt wird, der am anderen Eingang den über die Verbindung 11 zugeführten Einzelwert  $d(k,i)$  erhält. Die am Ausgang 13 des Addierers 70 entstehende Summe ist der Richtungswert  $D^+(k,i)$ . In entsprechender Weise entstehen am Ausgang 13 die Richtungswerte  $D^-(k,i)$  für die entgegengesetzte Richtung, wenn die Sprachparameter  $k$  von großen nach kleinen Werten laufen.

Parallel zum Umschalter 68 wird vom Ausgangssignal des Vergleichers 66 ein Umschalter 82 betätigt, der einer Leitung 83 entweder den logischen Wert "0" oder "1" zuführt, wobei letzterer Wert bedeutet, daß der vorhergehende Punkt zumindest schräg darunter liegt, d.h. in dem Adressenrechner 44 in Fig. 1 muß für die Adressierung des nächsten Wertes beim Sprachparameter  $k$  eine Einheit abgezogen werden.

Die Leitung 83 bildet den einen Anschluß eines weiteren Umschalters 84, der vom Ausgangssignal des Vergleichers 76 parallel zum Umschalter 78 gesteuert wird und dessen anderer Eingang fest den logischen Wert "1" erhält. Parallel dazu wird ein weiterer Umschalter 86 gesteuert, der der Ausgangsleitung 87 in der linken Stellung fest den logischen Wert "1" und in der rechten Stellung den logischen Wert "0" zuführt, wobei der letztere Wert angibt, daß der vorhergehende Punkt zum selben Zeitpunkt gehört, d.h. senkrecht darunter liegt, so daß der Adressenrechner 44 in Fig. 1 in diesem Falle den gleichen Adressenteil  $i$  für die Adressie-

rung des Speichers 50 liefert wie bei der jeweils  
 vorhergehenden Adresse. Die beiden Ausgangslei-  
 tungen 85 und 87 der Umschalter 84 und 86 bilden  
 zusammen die Verbindung 19, die in Fig. 4 auf die  
 Speicher 30 und 40 bzw. den Umschalter 32 führt.  
 Dieser Umschalter kann beispielsweise über ein  
 drittes Bit, dessen Wert umgeschaltet wird, ange-  
 ben, ob der vorhergehende Punkt oberhalb oder  
 unterhalb des gerade betrachteten Punktes liegt  
 und ob im Adressenrechner 44 dann zu dem mo-  
 mentanen Wert k eine Einheit addiert oder davon  
 eine Einheit subtrahiert werden muß.

Die in den Fig. 4 und 5 dargestellte Anordnung  
 ist nur beispielhaft zu betrachten, insbesondere  
 können einige oder auch alle Teile durch einen  
 entsprechend programmierten Mikroprozessor rea-  
 lisiert werden.

### Patentansprüche

- Verfahren zur Bestimmung des zeitlichen Ver-  
 laufs eines Sprachparameters eines Sprachsi-  
 gnals, wobei zu diskreten Zeitpunkten für je-  
 den Wert eines vorgegebenen Wertebereichs  
 des Sprachparameters ein Einzelwert vorhan-  
 den ist und der zeitliche Verlauf die Folge der  
 benachbarten, einschließlich diagonal benach-  
 barten Sprachsignalparameterwerte darstellt,  
 deren Einzelwert zumindest nahe den Extrem-  
 werten der Einzelwerte für die einzelnen Zeit-  
 punkte liegen, wobei die Summe der Einzel-  
 werte dieser Folge einen Summen-Extremwert  
 gegenüber anderen Folgen bildet,  
 dadurch gekennzeichnet, daß, bezogen auf das  
 Minimum als Extremwert der Einzelwerte und  
 als Summen-Extremwert, zu jedem Zeitpunkt  
 (i) in einem ersten Durchlauf nacheinander für  
 alle in der einen Richtung aufeinanderfolgen-  
 den Sprachparameterwerte (k) ein erster Rich-  
 tungswert  $D'(k,i)$  als die Summe aus dem be-  
 treffenden Einzelwert  $d(k,i)$  und dem Minimum  
 der folgenden Werte

$$\begin{aligned}
 &D(k,i-1) \\
 &D(l,i-1) + a(d(k,i) + A) \\
 &D'(l,i) + b(d(k,i) + A)
 \end{aligned}$$

sowie ein Hinweiswert  $h'(k,i)$  auf den das Mini-  
 mum liefernden Einzelwert gebildet und abge-  
 speichert wird, wobei  $D(k,i-1)$  bzw.  $D(l,i-1)$  ei-  
 nen beim jeweils vorhergehenden Zeitpunkt i-1  
 und denselben Sprachparameterwert k bzw.  
 dem vorhergehenden Sprachparameterwert l  
 erzeugten und gespeicherten Summenwert,  
 $D(l,i)$  der beim unmittelbar vorhergehenden  
 Sprachparameterwert l gebildete Richtungswert  
 und  
 a, b, A vorgegebene feste Größen sind, wobei

a und b vorzugsweise zwischen 0,5 und 2,0  
 liegen und A den Einzelwerten in stimmlosen  
 Sprachsignalsegmenten und in Sprachpausen  
 entspricht,

daß danach in einem zweiten Durchlauf für den  
 gleichen Zeitpunkt i für alle in der anderen  
 Richtung aufeinanderfolgenden Sprachparame-  
 terwerte k ein zweiter Richtungswert  $D''(k,i)$  mit  
 einem Hinweiswert  $h''(k,i)$  in entsprechender  
 Weise wie im ersten Durchlauf gebildet wird,  
 daß für jeden Sprachparameterwert k das Mini-  
 mum der beiden Richtungswerte  $D',D''$  und der  
 zu diesem Richtungswert gehörende Hinweis-  
 wert  $h',h''$  als jeweils neuer Summenwert D  
 bzw. als Gesamthinweiswert H gespeichert  
 wird, und daß spätestens am Ende des  
 Sprachsignals aus dem Gesamthinweiswert H-  
 (k,l), der zu dem Sprachparameterwert k mit  
 dem Summen-Extremwert beim letzten Zeit-  
 punkt l gehört, der unmittelbar vorhergehende  
 Sprachparameterwert ermittelt und der zuge-  
 hörige Gesamthinweiswert ausgelesen wird  
 usw., wobei die dabei entstehende Folge der  
 Sprachparameterwerte ausgegeben bzw. ge-  
 speichert wird.

- Verfahren nach Anspruch 1,  
 dadurch gekennzeichnet, daß die Richtungs-  
 werte  $D',D''$  die Hinweiswerte  $h',h''$  und die neu-  
 en Summenwerte D für jeweils nur einen Zeit-  
 punkt i gespeichert und danach wieder über-  
 schrieben werden.
- Verfahren nach Anspruch 2,  
 dadurch gekennzeichnet, daß die Richtungs-  
 werte  $D',D''$  und die Hinweiswerte  $h',h''$  nur für  
 die in der einen Richtung aufeinanderfolgen-  
 den Sprachparameterwerte k gespeichert wer-  
 den und daß nach der Bildung jedes Rich-  
 tungswerts  $D'',D'$  für die in der anderen Rich-  
 tung aufeinanderfolgenden Sprachparameter-  
 werte k unmittelbar der neue Summenwert D  
 und der Gesamthinweiswert H gebildet und  
 gespeichert wird.
- Anordnung zur Durchführung des Verfahrens  
 nach einem der Ansprüche 1 bis 3,  
 gekennzeichnet durch
  - einen ersten Speicher (10) für die Einzel-  
 werte  $d(k,i)$  aller Sprachparameterwerte k  
 mindestens jeweils eines Zeitpunktes i,
  - einen zweiten Speicher (20) für jeweils  
 einen Summenwert  $D(k,i)$  für jeden  
 Sprachparameterwert k mindestens je-  
 weils eines Zeitpunktes i,
  - einen dritten Speicher (30) für jeweils

- einen ersten Richtungswert  $D'(k,i)$  und einen Hinweiswert  $h'(k,i)$  für jeden Sprachparameterwert  $k$  mindestens jeweils eines Zeitpunktes  $i$ ,
- einen vierten Speicher (26; 40) für einen zweiten Richtungswert  $D''(k,i)$  und einen zugehörigen Hinweiswert  $h''(k,i)$  mindestens eines Sprachparameterwertes  $k$  für denselben Zeitpunkt  $i$  wie die Werte im dritten Speicher (30),
  - einen fünften Speicher (50) für die Gesamthinweiswerte  $H(k,i)$  für alle Sprachparameterwerte  $k$  und alle Zeitpunkte  $i$ ,
  - eine Verarbeitungsanordnung (12) mit Eingängen, die mit einem Datenausgang des ersten, des zweiten und des dritten bzw. vierten Speichers (10, 20, 30, 26, 40) gekoppelt sind, und mit einem Ausgang zum Abgeben jeweils eines Richtungswertes  $D'(k,i), D''(k,i)$  und eines zugehörigen Hinweiswertes  $h'(k,i), h''(k,i)$ , der mit dem Dateneingang des dritten Speichers (30) gekoppelt ist,
  - einen Vergleicher (14) mit zwei Eingängen, von denen der eine vom Datenausgang des dritten Speichers (30) jeweils einen ersten Richtungswert  $D'(k,i)$  und der andere Eingang den zugehörigen zweiten Richtungswert  $D''(k,i)$  empfängt, und mit einem Ausgang, der einen Umschalter (32) steuert, der den kleineren der beiden Richtungswerte dem Dateneingang des zweiten Speichers (20) zuführt, und
  - eine Steueranordnung (16) mit einem Adressengenerator (54, 58), der zyklisch die Adressen aller Sprachparameterwerte  $k$  in der einen Richtung der Adressenfolge und danach in der anderen Richtung der Adressenfolge für mindestens den ersten bis dritten und den fünften Speicher (10, 20, 30, 50) erzeugt und der außerdem die Adressenauswahl für den Zeitpunkt  $i$  mindestens für den fünften Speicher (50) erzeugt, und mit einer Ablaufsteuerung (56), die das Einschreiben und Auslesen mindestens des zweiten, dritten und fünften Speichers (20, 30, 50) steuert.

## Claims

1. A method for determining the variation with time of a speech parameter of a speech signal, where an individual value exists at discrete instants for each value of a predetermined value range of the speech parameter and the variation with time represents the sequence of

adjacent, including diagonally adjacent, speech signal parameter values whose individual values are at least close to the extreme values of the individual values for the individual instants, the sum of the individual values of this sequence forming an extreme value sum in relation to other sequences, characterized in that referred to the minimum as extreme value of the individual values and as extreme value sum at each instant (i), a first direction value  $D'(k,i)$  is formed, in a first pass successively for all speech parameter values ( $k$ ) succeeding one another in the one direction, as the sum of the relevant individual value  $d(k,i)$  and the minimum of the following values

$$\begin{aligned} & d(k,i-1) \\ & D(l,i-1) + (a(d(k,i), + A) \\ & D'(l,i) + b)d(k,i) + A) \end{aligned}$$

as well as a pointing value  $h'(k,i)$ , pointing to the individual value supplying the minimum, being formed and stored, where  $D(k,i-1)$  or  $D(l,i-1)$  is a sum value generated and stored at the respective preceding instant  $i-1$  and the same speech parameter value  $k$  or the preceding speech parameter value  $l$ , respectively,  $D(l,i)$  is the direction value formed at the immediately preceding speech parameter value  $l$ , and  $a, b, A$  are predetermined, fixed quantities where  $a$  and  $b$  are preferably between 0.5 and 2.0 and  $A$  corresponds to the individual values in non-voiced speech signal segments and in speech intervals, that thereafter a second direction value  $D''(k,i)$  with a pointing value  $h''(k,i)$  is formed, in the manner corresponding to the first pass, in a second pass for the same instant  $i$  for all speech parameter values  $k$  succeeding one another in the other direction, that for each speech parameter value  $k$  the minimum of the two direction values  $D', D''$  and the pointing value  $h', h''$  associated with this direction value is stored as the respective new sum value  $D$  or as the total pointing value  $H$ , and that the immediately preceding speech parameter value is determined, at the latest at the end of the speech signal, from the total pointing value  $H(k,l)$  associated with the speech parameter value  $k$  with the extreme value sum at the last instant  $l$ , and the associated total pointing value is read out and so forth, the sequence of speech parameter values has produced being output and stored.

2. A method as claimed in Claim 1, characterized in that the direction values  $D', D''$ , the pointing values  $h', h''$  and the new sum values  $D$  are stored each time for instant  $i$  only, after which

they are overwritten again.

3. A method as claimed in Claim 2, characterized in that the direction values  $D', D''$  and the pointing values  $h', h''$  are stored only for the speech parameter values  $k$  succeeding one another in the one direction and that, after each direction value  $D'', D'$  has been formed for the speech parameter values  $k$  succeeding one another in the other direction, the new sum value  $D$  and the total pointing value  $H$  are directly formed and stored. 5 10
4. An arrangement for carrying out the method claimed in any one of Claims 1 to 3, characterized by 15
- a first memory (10) for the individual values  $d(k, i)$  of all speech parameter values  $k$  of each time at least one instant  $i$ ,
  - a second memory (20) for each time one sum value  $D(k, i)$  for each speech parameter value  $k$  of each time at least one instant  $i$ , 20
  - a third memory (30) for each time one first direction value  $D'(k, i)$  and one pointing value  $h'(k, i)$  for each speech parameter value  $k$  of each time at least one instant  $i$ , 25
  - a fourth memory (26; 40) for a second direction value  $D''(k, i)$  and an associated pointing value  $h''(k, i)$  of at least one speech parameter value  $k$  for the same instant  $i$  as the values in the third memory (30), 30
  - a fifth memory (50) for the total pointing values  $H(k, i)$  for all speech parameter values  $k$  and all instants  $i$ , 35
  - a processing arrangement (12) with inputs which are coupled to a data output of the first, the second, and the third or the fourth memory (10, 20, 30, 26, 40), and with an output, for supplying each time one direction value  $D'(k, i)$ ,  $D''(k, i)$  and an associated pointing value  $h'(k, i)$ ,  $h''(k, i)$ , which is coupled to the data input of the third memory (30), 40 45
  - a comparator (14) with two inputs, one of which receives from the data output of the third memory (30) each time one first direction value  $D'(k, i)$ , while the other input receives the associated second direction value  $D''(k, i)$ , and with an output which controls a switch (32) which applies the smaller one of the two direction values to the data input of the second memory (20), and 50 55
  - a control arrangement (16) with an address generator (54, 58) which cyclically

generates the addresses of all speech parameter values  $k$  in the one direction of the address sequence and subsequently, in the other direction, of the address sequence for at least the first to the third and the fifth memory (10, 20, 30, 50) and which, moreover, generates the address selection for the instant  $i$  at least for the fifth memory (50), and with a sequence controller (56) which controls the writing and reading of at least the second, the third and the fifth memory (20, 30, 50).

## Revendications

1. Procédé pour déterminer l'évolution temporelle d'un paramètre de parole d'un signal de parole, selon lequel, à des moments discrets, pour chaque valeur d'un domaine de valeurs prédéfini du paramètre de parole, une valeur individuelle est présente et l'évolution temporelle représente la séquence des valeurs de paramètre de signal de parole voisines, y compris diagonalement voisines, dont les valeurs individuelles sont situées au moins à proximité des valeurs extrêmes des valeurs individuelles pour les moments individuels, la somme des valeurs individuelles de cette séquence formant une valeur extrême de somme par rapport à d'autres séquences, caractérisé en ce que, sur base du minimum comme valeur extrême des valeurs individuelles et comme valeur extrême de somme, à chaque moment ( $i$ ), dans un premier parcours, successivement pour toutes les valeurs de paramètre de parole ( $k$ ) qui se suivent dans un sens, une première valeur de sens  $D'(k, i)$ , en tant que somme de la valeur individuelle en question  $d(k, i)$  et du minimum des valeurs suivantes :

$$D(k, i-1) \\ D(l, i-1) + a(d(k, i) + A) \\ D'(l, i) + b(d(k, i) + A)$$

ainsi qu'une valeur d'indication  $h'(k, i)$  vers la valeur individuelle fournissant le minimum sont formées et stockées, où  $D(k, i-1)$  ou  $D(l, i-1)$  est une valeur de somme produite et stockée chaque fois au moment précédent  $i-1$  pour la même valeur de paramètre de parole  $k$  ou la valeur de paramètre de parole précédente  $l$ ,  $D(l, i)$  est la valeur de sens formée lors de la valeur de paramètre de parole directement précédente, et  $a, b, A$  sont des grandeurs fixes prédéfinies,  $a$  et  $b$  étant de préférence compris entre 0,5 et

- 2,0 et A correspondant aux valeurs individuelles dans des segments de signal de parole sourds et dans des interruptions de parole, que, par la suite, dans un second parcours pour le même moment  $i$  pour toutes les valeurs de paramètre de parole  $k$  se suivant dans l'autre sens, une seconde valeur de sens  $D''(k,i)$  avec une valeur d'indication  $h''(k,i)$  est formée de la même manière que dans le premier parcours,
- que, pour chaque valeur de paramètre de parole  $k$ , le minimum des deux valeurs de sens  $D', D''$  et de la valeur d'indication  $h', h''$  appartenant à cette valeur de sens est stocké chaque fois en tant que nouvelle valeur de somme  $D$  ou valeur d'indication d'ensemble  $H$ , et qu'au plus tard à la fin du signal de parole, à partir de la valeur d'indication d'ensemble  $H(k,i)$ , qui appartient à la valeur de paramètre de parole  $k$  de la valeur de somme extrême au dernier moment  $i$ , la valeur de paramètre de parole immédiatement précédente est déterminée et la valeur d'indication d'ensemble associée est lue, etc., la séquence des valeurs de paramètre de parole ainsi formée étant soit sortie, soit stockée.
2. Procédé suivant la revendication 1, caractérisé en ce que les valeurs de sens  $D', D''$ , les valeurs d'indication  $h', h''$  et les nouvelles valeurs de somme  $D$  pour un seul moment  $i$  sont chaque fois stockées, puis à nouveau recouvertes.
3. Procédé suivant la revendication 2, caractérisé en ce que les valeurs de sens  $D', D''$  et les valeurs d'indication  $h', h''$  ne sont stockées que pour les valeurs de paramètre de parole  $k$  qui se suivent dans un des sens et que, après la formation de chaque valeur de sens  $D'', D'$  pour les valeurs paramétriques de parole  $k$  qui se suivent dans l'autre sens, la nouvelle valeur de somme  $D$  et la valeur d'indication d'ensemble  $H$  sont immédiatement formées et stockées.
4. Dispositif pour l'exécution du procédé suivant l'une quelconque des revendications 1 à 3, caractérisé par
- une première mémoire (10) pour les valeurs individuelles  $d(k,i)$  de toutes les valeurs de paramètre de parole  $k$  d'au moins chaque fois un moment  $i$ ,
  - une deuxième mémoire (20) pour chaque fois une valeur de somme  $D(k,i)$  pour chaque valeur de paramètre de parole  $k$  d'au moins chaque fois un moment  $i$ ,
  - une troisième mémoire (30) pour chaque fois une première valeur de sens  $D'(k,i)$
- et une valeur d'indication  $h'(k,i)$  pour chaque valeur de paramètre de parole  $k$  d'au moins chaque fois un moment  $i$ ,
- une quatrième mémoire (26; 40) pour une seconde valeur de sens  $D''(k,i)$  et une valeur d'indication  $h''(k,i)$  associée d'au moins une valeur de paramètre de parole  $k$  pour le même moment  $i$  que les valeurs dans la troisième mémoire (30),
  - une cinquième mémoire (50) pour les valeurs d'indication d'ensemble  $H(k,i)$  pour toutes les valeurs de paramètre de parole  $k$  de tous les moments  $i$ ,
  - un dispositif de traitement (12) comportant des entrées qui sont connectées à une sortie de données de la première, de la deuxième et de la troisième ou de la quatrième mémoire (10, 20, 30, 26, 40), et comportant une sortie pour fournir chaque fois une valeur de sens  $D'(k,i)$ ,  $D''(k,i)$  et une valeur d'indication  $h'(k,i)$ ,  $h''(k,i)$  associée, qui est connectée à l'entrée de données de la troisième mémoire (30),
  - un comparateur (14) à deux entrées, dont l'une reçoit de la sortie de données de la troisième mémoire (30) chaque fois une première valeur de sens  $D'(k,i)$  et l'autre, la seconde valeur de sens  $D'(k,i)$  associée, et une sortie, qui commande un commutateur (32) qui applique la plus petite des deux valeurs de sens à l'entrée de données de la deuxième mémoire (20), et
  - un dispositif de commande (16) comportant un générateur d'adresses (54, 58), qui produit cycliquement les adresses de toutes les valeurs de paramètre de parole  $k$  dans l'un des sens de la séquence d'adresses, puis dans l'autre sens de la séquence d'adresses pour au moins les trois premières mémoires et pour la cinquième (10, 20, 30, 50) et qui, en outre, produit la sélection d'adresse pour le moment  $i$  au moins pour la cinquième mémoire (50), et comportant une commande séquentielle (56) qui commande l'écriture et la lecture d'au moins la deuxième, la troisième et la cinquième mémoire (20, 30, 50).

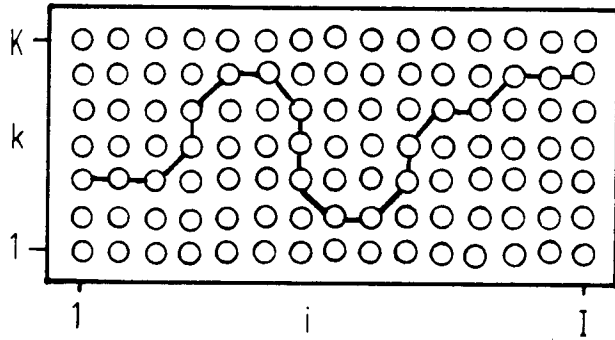


Fig. 1

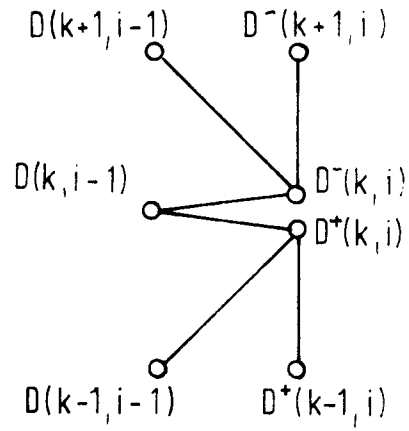


Fig. 2

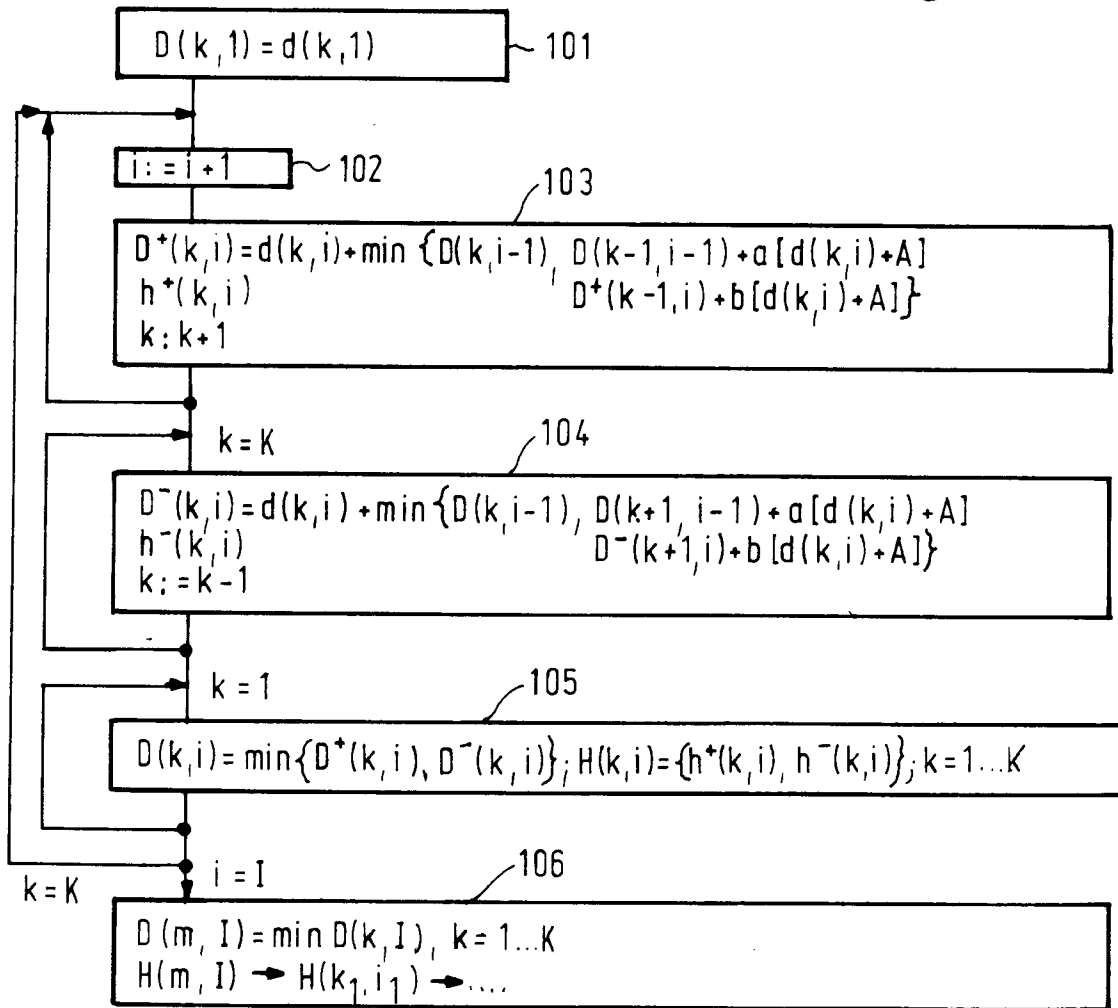


Fig. 3

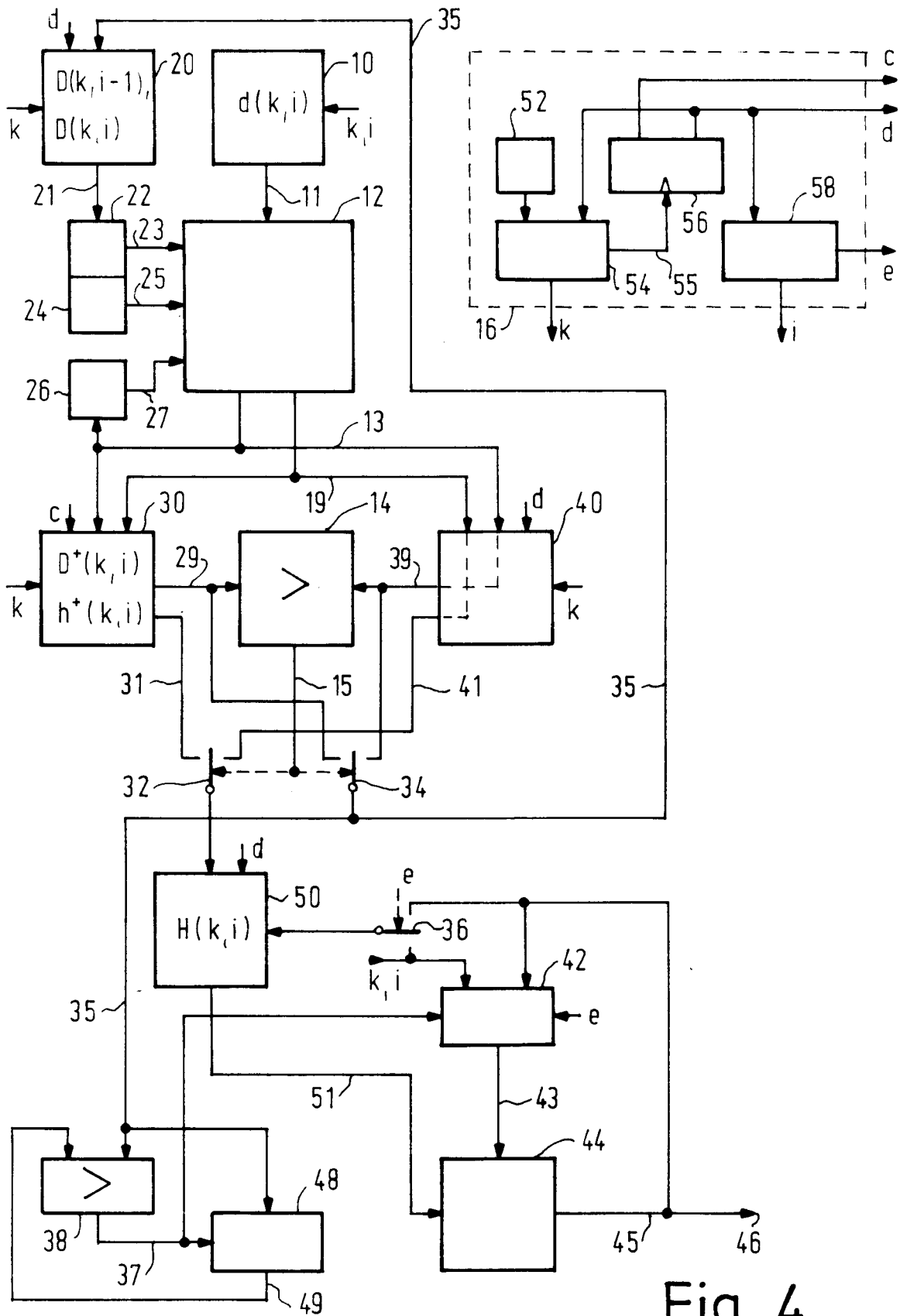


Fig. 4

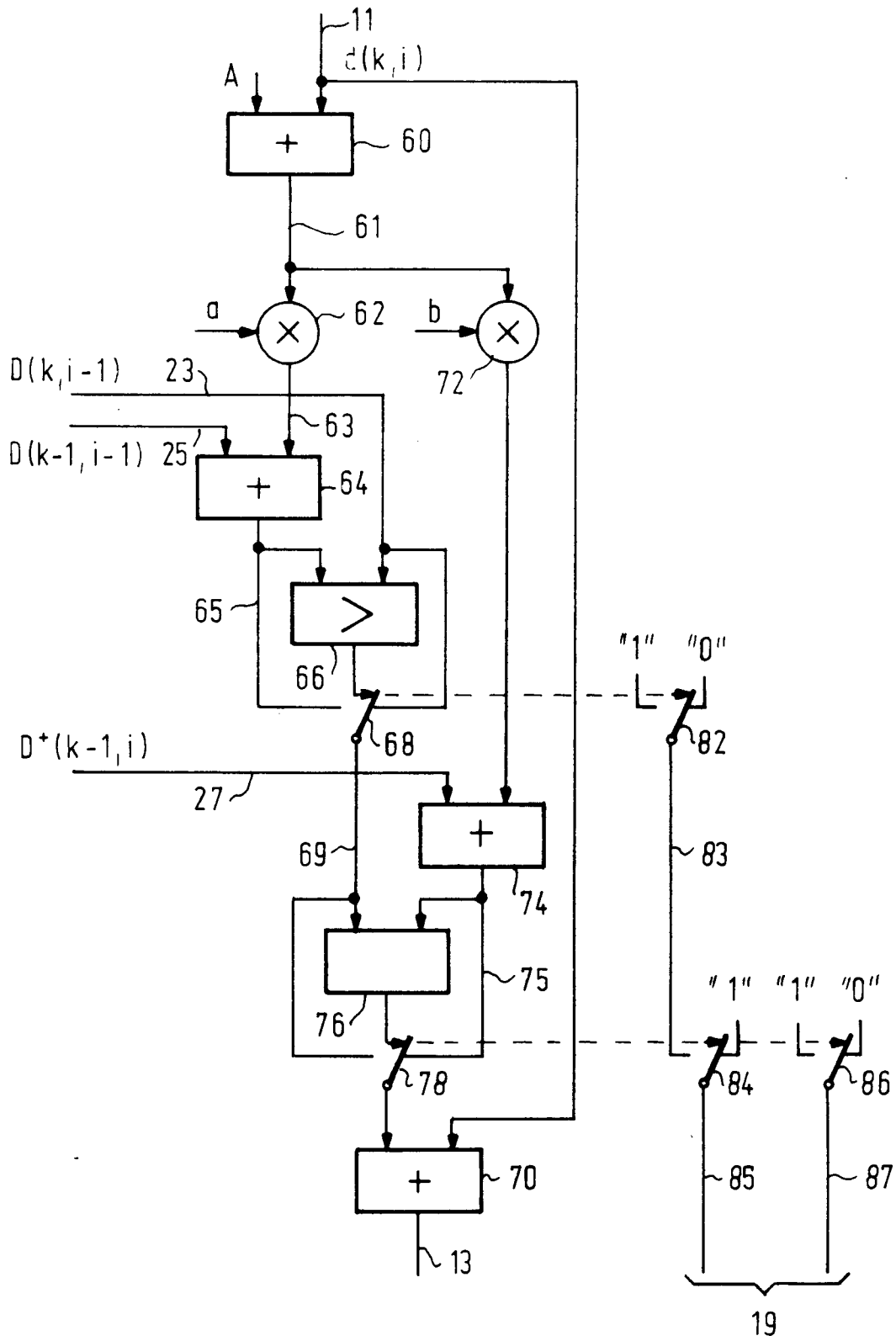


Fig. 5