(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Veröffentlichungsnummer: **0 308 817**
**A2**

# (12) EUROPÄISCHE PATENTANMELDUNG

(21) Anmeldenummer: **88115139.3**

(51) Int. Cl.4: **G10L 9/14 , G10L 7/04**

(22) Anmeldetag: **15.09.88**

(30) Priorität: **23.09.87 DE 3732047**

(43) Veröffentlichungstag der Anmeldung:
**29.03.89 Patentblatt 89/13**

(84) Benannte Vertragsstaaten:
**DE FR GB IT NL**

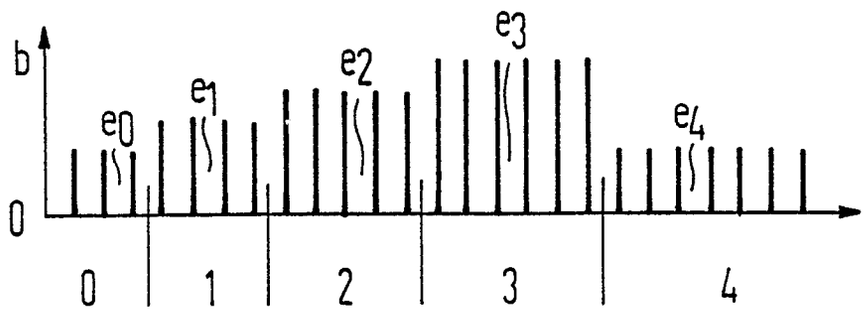
(71) Anmelder: **Siemens Aktiengesellschaft Berlin und München**
**Wittelsbacherplatz 2**
**D-8000 München 2(DE)**

(72) Erfinder: **Brandl, Hans, Dipl.-Ing.**
**Ottweilerstrasse 15**
**D-8000 München 83(DE)**

## (54) Verfahren zur Umcodierung von Kanalvocoder-Parameter in LPC-Vocoder-Parameter.

(57) Die Erfindung betrifft ein Verfahren zur Umcodierung von Kanalvocoder-Parametern in LPC-Vocoder-Parameter, wobei aus dem geglätteten Leistungsspektrum der Kanalvocoder-Parameter durch eine inverse diskrete Fouriertransformation die LPC-Vocoder-Parameter errechnet werden. Erfindungsgemäß werden bei vorgegebener Kanalzahl des Kanal-Vocoders und bei vorgegebener Parameterzahl des LPC-Vocoders aus den hierbei konstanten Größen Matrixelemente errechnet und die LPC-Vocoder-Parameter durch Matrixmultiplikationen aus den Kanalvocoder-Parametern berechnet.

FIG 1




EP 0 308 817 A2

# Verfahren zur Umcodierung von Kanalvocoder-Parameter in LPC-Vocoder-Parameter

Die Erfindung betrifft ein Verfahren gemäß dem Oberbegriff des Patentanspruchs 1.

Derzeit werden digitale Schmalband-Kommunikationsnetze mit niedrigen Datenübertragungsraten (1-2 kbit/s) geplant. Die hierbei angewandten Codierungsverfahren bauen entweder auf dem Prinzip des Kanalvocoders oder der linearen Prädiktion (LPC-Vocoder) auf. Eine Kommunikation zwischen den Vocodern ist nur möglich, falls an ihrer Schnittstelle eine geeignete Datenumcodierung erfolgt.

Der hierzu benötigte Umsetzer soll möglichst aufwandgünstig gestaltet sein und die Sprachqualität möglichst nicht verschlechtern.

Eine Möglichkeit, einen Umsetzer aufzubauen, besteht in der Rücktransformation der Sprachdaten in das Sprachsignal und dessen Neucodierung.

Dieses Verfahren ist sehr aufwendig, da zwei Analyseeinheiten und zwei Syntheseeinheiten benötigt werden. Durch Analysefehler verschlechtert sich außerdem die Sprachqualität. Die Verschlechterung der Sprachqualität läßt sich durch direkte Umcodierung der Daten der verschiedenen Vocoder umgehen. Diese Möglichkeit ergibt sich aus dem sehr ähnlichen Syntheseprinzip, das bei dem Kanalvocoder und dem LPC-Vocoder angewandt wird.

Das Sprachsignal wird hierbei durch ein Anregungssignal, welches durch ein variables Filter gefiltert wird, erzeugt. Das Anregungssignal besteht bei stimmhaften Lauten aus einer Pulsfolge und bei stimmlosen Lauten aus weißem Rauschen. Mit den Anregungsparametern wird die Pulsfrequenz und der Anregungsmodus - stimmhaft oder stimmlos - festgelegt. Das variable Übertragungsverhalten des Filters entspricht dem variablen Resonanzverhalten des menschlichen Vokaltraktes. Dieses ändert sich langsam und wird durch Filterparameter alle 10 bis 20 ms neu eingestellt. Aufgabe der Sprachsignal-Analyse eines Vocoders ist es, aus einem Sprachsignal die Anregungsparameter und die Filterparameter zu gewinnen. Der LPC-Vocoder und der Kanalvocoder unterscheiden sich im wesentlichen in der Struktur des Filters. LPC geht von einem Allpolfilter und der Kanalvocoder von einer Filterbank aus. Damit unterscheiden sich die Analyseverfahren zur Bestimmung der entsprechenden Filterparameter und es ergeben sich andere Filterparameter, die in den verschiedenen Netzen übertragen werden. Dagegen sind die Anregungsparameter im Prinzip die gleichen.

Es wird also ein Umcodierverfahren gesucht, welches die Filterparameter einer Filterbank eines Kanalvocoders in die Filterparameter eines Allpolfilters eines LPC-Vocoders umwandelt.

Die Kanalvocoder-Parameter (oder Koeffizienten) stellen nachrichtentheoretisch meist ein nicht-äquidistant abgetastetes Spektrum dar. Aus dem Amplitudenspektrum wird nun das Leistungsspektrum berechnet und mit Hilfe der Fouriertransformation in die Autokorrelationsfunktion (AKF) transformiert. Aus der AKF kann nun in bekannter Weise mit Hilfe der üblichen Verfahren (z.B. Levinson-Rekursion) der entsprechende LPC-Vocoder-Parametersatz berechnet werden (siehe H. Hermansky, B. Hanson, H. Witka; "Perceptually based Predictive Analysis of Speech" on ICASSP 85, S. 13.10 Tagungsband).

Die direkte Transformation ist mit hohem technischen Aufwand verbunden. Es werden leistungsfähige Real-time-Prozessoren zur Berechnung von Spektren und Korrelationsfunktionen benötigt.

Der Erfindung liegt die Aufgabe zugrunde, ein Verfahren zur Umcodierung von Kanalvocoder-Parameter in LPC-Vocoder-Parameter anzugeben, das bei hoher Genauigkeit relativ wenige Rechenoperationen benötigt.

Diese Aufgabe wird erfindungsgemäß durch die im Patentanspruch 1 angegebenen Merkmale gelöst.

Im folgenden wird ein bekanntes Verfahren zur Umcodierung anhand der mathematischen Methoden erläutert.

Ausgangspunkt sind die Kanalvocoder-Parameter, die beispielsweise als Leistungsspektrum vorliegen (siehe FIG 1). Dieses Leistungsspektrum liegt bei einem Kanalvocoder nur in einer abschnittsweisen konstanten Form $b_k$ mit Sprüngen an den Übergangsstellen von $b_k$ nach $b_{k+1}$ vor. In FIG 1 sind als diese Parameter $b_k$ Energiewerte $e_j$ dargestellt, wobei der Wert $e_j$ der Energie im Kanal mit der Nummer j entspricht. Hierbei entspricht in allgemein bekannter Weise die Kanalenergie der Leistung in einem 20 ms-Intervall (dies ist das Intervall, nach dem jeweils neue Filterparameter eingestellt werden). Dieses Intervall ist auch gleichzeitig das Transformationsintervall.

Aus diesem "rohen" Spektrum $b_k$ wird durch Faltung mit einer Glättungsfunktion g (i, s) ein geglättetes Spektrum $a_k$ (siehe FIG 2) gebildet. Die Glättungsfunktion g ist eine gerade Funktion, g (i, s) = g (-i, s), mit i als Argument und mit s als Streuung, durch die die Breite der Glättungsfunktion g gegeben ist.

Für diese Glättungsfunktion g eignen sich beispielsweise Gauß'sche Glockenkurven oder ähnliche Funktionen. Als Beispiel für die Gauß'sche Glockenkurve wird folgende Funktion angegeben:

Weitere mögliche Glättungsfunktionen g sind die aus der Filtertheorie und der digitalen Signalverarbeitung bekannten Tiefpaßfunktionen. In diesen Fällen definiert die Streuung s die Eckfrequenzen des jeweiligen Tiefpasses.

Für den Spezialfall eines Diracimpulses

$$g(i, s) = \begin{cases} 1 & \text{für } i = 0 \\ 0 & \text{" } \quad i \neq 0 \end{cases}$$

würde $b_k$ unverändert auf das geglättete Spektrum $a_k$ abgebiltet werden.

Bei der Glättung eines realen Sprachspektrums ($b_k$) kann die Streuung s eine Funktion der aktuellen Spektrallinie sein. In diesem Fall wird bei höheren Frequenzen und damit breiteren Kanälen in $b_k$ eine größere Streuung s für die Glättungsfunktion g (i, s) gewählt als bei tieferen Frequenzen. Damit ist eine Anpassung der Glättung an die Tonheitsempfindung (Bark -Skala) des menschlichen Ohres möglich. Über die Wahl des oder der Streuungen s ist der "Wohlklang" bei der Sprachsynthese empirisch wählbar.

Für die Berechnung des geglätteten Spektrums $a_k$ aus dem "Roh"-Spektrum $b_k$ ergibt sich somit folgende Formel:

$$(1) \quad a_k = \sum_{l=0}^{N-1} b_l \cdot g[(l-k), (u \cdot k)]$$

mit g : Glättungsfunktion
u : Glättungsbreite (Normierung)
u . k : Streuung
$a_k$: K-ter Koeffizient des geglätteten Leistungsspektrums
N : Anzahl der Spektralkoeffizienten
$b_l$: l-ter Koeffizient des Rohspektrums.

Die LPC-Koeffizienten werden i. a. aus der Kurzzeit-Autokorrelationsfunktion (ca. 20 ms), kurz AKF genannt, des Sprachsignals errechnet. Diese AKF, d.h. deren Korrelationskoeffizienten $r_i$ lassen sich auch aus dem Leistungsspektrum des Sprachsignals durch die inverse, diskrete Fouriertransformation bestimmen.

Für die M Korrelationskoeffizienten $r_i$ ergeben sich dann folgende Gleichungen:

$$(2) \quad r_i = \sum_{k=0}^{N-1} a_k \cdot \cos\left(i \cdot k \cdot \frac{2\pi}{N}\right)$$

i = 0,1 ... M, Anzahl der Korrelationskoeffizienten (sonst wie in Formel (1)).

Formel (1) in Formel (2) eingesetzt ergibt nach Anwendung des Kommutativgesetzes:

$$(3) \quad r_i = \sum_{l=0}^{N-1} b_l \cdot \sum_{k=0}^{N-1} g[(l-k), (u \cdot k)] \cdot \cos\left(i \cdot k \cdot \frac{2\pi}{N}\right)$$

Die N Spektrallinien $b_l$ des Rohspektrums lassen sich von den Kanalenergiewerten $e_j$ ableiten (siehe FIG 1)

Bei realen Vocodern liegen die Kanalzahlen und damit auch die Anzahl der Kanalenergiewerte $e_j$ bei etwa 16-18. Für die Anzahl der Spektralkoeffizienten N im Bereich von etwa 256 lassen sich die Koeffizienten $b_k$ des "rohen" Leistungsspektrums folgendermaßen darstellen:

(4) $b_l = e_i$ für $l = m_j....(m_{j+1}-1)$

$m_j$ : Index der ersten Spektrallinie des Kanals j

$m_{j+1}-1$ : Index der letzten Spektrallinie des Kanals j

Formel (4) eingesetzt in Formel (3) ergibt folgende allgemeine Gleichung zur Berechnung der AKF aus den Vocoder-Kanalenergiewerten

$e_j$ mit j = l-P

(5)

$$r_i = \sum_{j=1}^{P} e_j \cdot \sum_{l=m_j}^{m_{j+1}-1} \sum_{k=0}^{N-1} g[(l-k),(m-k)] \cdot \cos\left(i \cdot k \cdot \frac{2\pi}{N}\right)$$

m = l erste Spektrallinien des ersten Kanals

$m_p$ = N letzte Spektrallinie des letzten Kanals

Im folgenden wird das erfindungsgemäße Verfahren zur Umcodierung erläutert.

Alle Elemente nach den Vocoder-Kanalenergiewerten $e_j$ sind Konstante.

Für ein vorgegebenes Frequenz- und Zeitraster, hinsichtlich der Kanalvocoder- und der LPC-Vocoder-Parameter, läßt sich die Formel (5) in eine Matrixmultiplikation umschreiben:

(6)

$$r_i = \sum_{j=1}^{P} c_{ij} \cdot e_j$$

i = 0...M: Koeffizienten der AKF

P : Kanalzahl

mit:

(7)

$$c_{ij} = \sum_{l=m_j}^{m_{j+1}-1} \sum_{k=0}^{N} g[(l-k),(m-k)] \cdot \cos\left(i \cdot k \cdot \frac{2\pi}{N}\right)$$

oder in Matrix-Schreibweise

$\vec{R} = C \times \vec{E}$

mit $\vec{R}$ = AKF-Vektor

C : Matrix mit den Elementen aus Formel (7)

$\vec{E}$ : Kanalvocoder-Energievektor

Zur Umcodierung werden beim erfindungsgemäßen Verfahren nur einmal die Elemente der Matrix C für eine bestimmte Vocoder-Kombination berechnet. Anschließend sind zur Umcodierung der jeweiligen Sprach-Parameter nur noch Matrixmultiplikationen zwischen den Energievektoren E (der die Parameter enthält) und der Matrix C auszuführen.

Für einen praktischen Fall mit beispielsweise P = 18 Kanälen eines Kanalvocoders und einer gewünschten Zahl von 11 Autokorrelationswerten für LPC-10 sind somit nur noch ca 200 Multiplikationen und etwa ebensoviele Additionen nötig. Bei konventionellen Verfahren werden ca. 4000 Rechenoperationen benötigt.

Im folgenden wird anhand von FIG 3 eine Schaltungsanordnung zur Durchführung der vorstehend beschriebenen Matrixmultiplikation erläutert.

An einem Eingang 1 eines ersten Speichers 2 liegen die geglätteten Kanalvocoder-Parameter $a_p$ an. Es wird beispielsweise jeweils ein Satz dieser Parameter, bei 18 Kanälen also 18 Werte, in den ersten Speicher 2 eingeschrieben.

Es soll folgende Rechenoperation ausgeführt werden:

$$(8)\quad l_i = \sum_{p=0}^{17} c_{ip} \cdot a_p$$

mit $l_i$ : LPC-Vocoder-Parameter (diese entsprechen den Autokorrelationskoeffizienten $r_i$ in Formel (6))
$c_{ip}$: Transformationskoeffizienten (Matrixelemente), berechnet nach Formel (7)
$a_p$: Kanalvocoder-Parameter

Hierbei werden für eine Umcodierung der Parameter eines vorgegegebenen Kanalvocoders in Parameter eines vorgegebenen LPC-Vocoders die Transformationskoeffizienten $c_{ip}$ der Matrix C berechnet und in einem Koeffizientenspeicher 3 abgelegt.

Zur Durchführung der Matrixmultiplikation werden die Kanalvocoder-Parameter $a_p$ im ersten Speicher 2 von einem ersten Zähler 4 nacheinander adressiert. Analog dazu werden die Koeffizienten $c_{ip}$ im Koeffizientenspeicher 3 nach ihrem Index p adressiert.

In einem Multiplizierer 5 werden die adressierten Kanalvocoder-Parameter $a_p$ und die adressierten Koeffizienten $c_{ip}$ multipliziert, und in einem nachgeschalteten Addierer 6 aufsummiert. Hierbei wird der Index i der Koeffizienten $c_{ip}$ solange konstant gehalten, bis der Index i seinen größten Wert, in Formel 8 beispielsweise 17, erreicht hat. Die gebildete Summe wird als LPC-Parameter $l_i$ in einen zweiten Speicher 7 eingeschrieben. Danach wird von einem zweiten Zähler 8 der Index i um eins erhöht, und der nächste LPC-Parameter $l_{i+1}$ berechnet. Hierzu adressiert der zweite Zähler 8 zum einen die Koeffizienten $c_{ip}$ im Koeffizientenspeicher 3 nach ihrem Index i, und zum anderen die LPC-Vocoder-Parameter im zweiten Speicher 7.

Die beiden Zähler 4 und 8 werden von einer Taktsteuerung 9 getaktet.

An einem Ausgang 10 des zweiten Speichers 7 ist dann jeweils ein transformierter bzw. umcodierter Satz von LPC-Vocoder-Parametern abnehmbar.

## Ansprüche

1. Verfahren zur Umcodierung von digitalen Kanalvocoder-Parametern, die im Analyseteil des Kanalvocoders aus einem natürlichen Sprachsignal gewonnen wurden, in digitale LPC-Vocoder-Parameter, die im Syntheseteil des LPC-Vocoders zu einem synthetischen Sprachsignal verarbeitet werden, wobei die Kanalvocoder-Parameter als Leistungsspektrum vorliegen, wobei die LPC-Vocoder-Parameter aus der Kurzzeit-Autokorrelationsfunktion berechnet werden, wobei das Leistungsspektrum mit einer Glättungsfunktion (g) geglättet wird,und wobei aus dem geglätteten Leistungsspektrum durch eine inverse, diskrete Fouriertransformation die Korrelationskoeffizienten der Autokorrelationsfunktion errechnet werden, **dadurch gekennzeichnet,** daß bei vorgegebener Kanalzahl des Kanal-Vocoders und bei vorgegebener Parameterzahl des LPC-Vocoders bei einem vorgegebenen Frequenz- und Zeitraster aus den hierbei konstanten Größen Matrixelemente ($c_{ij}$) errechnet und in einem Koeffizientenspeicher (3) abgelegt werden, so daß die LPC-Vocoder-Parameter durch Matrixmultiplikationen aus den Kanalvocoder-Parametern ableitbar sind, wobei jeweils die Parameter eines der Vocoder einen Vektor bilden.

2. Verfahren nach Anspruch 1, **dadurch gekennzeichnet,** daß die Glättungsfunktion (g = g (i, s)) eine Streuung (s) beinhaltet, durch die die Breite der Glättungsfunktion gegeben ist.

3. Verfahren nach Anspruch 1 oder Anspruch 2, **dadurch gekennzeichnet,** daß die Breite (s) der Glättungsfunktion (g) eine Funktion der Parameter des Kanal-Vocoders ist.

FIG 1

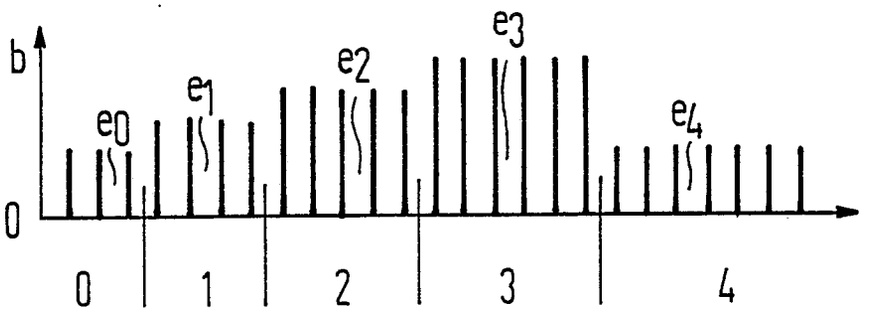



FIG 2

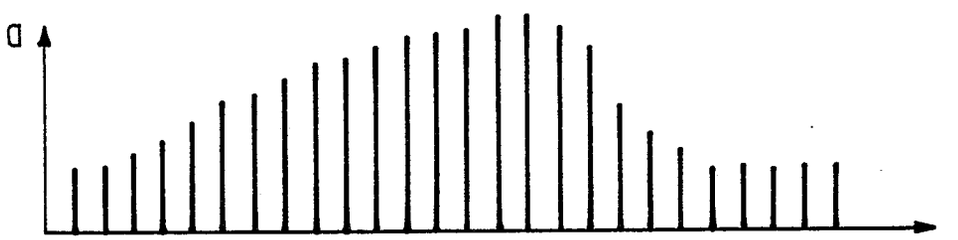


FIG 3

