

(12) **Österreichische Patentanmeldung**

(21) Anmeldenummer: **A 1547/2006**

(51) Int. Cl.⁸: **G10L 21/02** (2006.01)

(22) Anmeldetag: **15.09.2006**

(43) Veröffentlicht am: **15.03.2008**

(73) Patentanmelder:

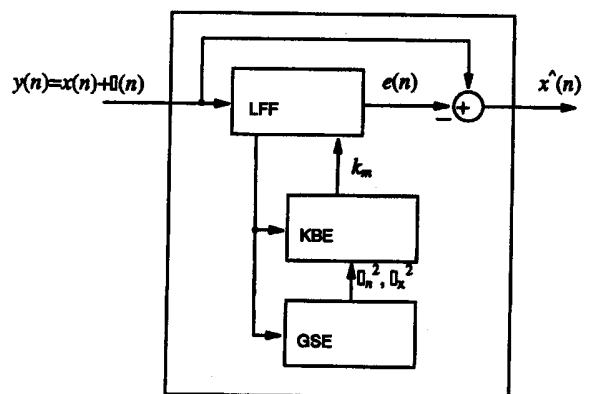
TECHNISCHE UNIVERSITÄT GRAZ
A-8010 GRAZ (AT)

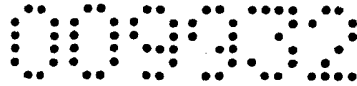
(72) Erfinder:

RANK ERHARD DIPL.ING. DR.TECHN.
WIEN (AT)
KUBIN GERNOT DIPL.ING. DR.TECHN.
GRAZ (AT)

(54) **VORRICHTUNG ZUR GERÄUSCHUNTERDRÜCKUNG BEI EINEM AUDIOSIGNAL**

(57) Eine Vorrichtung zur Geräuschunterdrückung mit einer linearen Prädiktions-Analysierschaltung mit einem LP-Fehler-Filter (LFF), das auf Basis eines ersten, von Geräusch überlagerten Sprachsignals $y(n) = x(n) + \varepsilon(n)$ ein LP-Fehler-Filter Ausgangssignal $e(n)$ erzeugt, einer Koeffizientenberechnungseinheit, welche die Koeffizienten des LP-Fehler-Filters auf Basis der internen Signale (inklusive des Ein- und Ausgangssignals $y(n)$ und $e(n)$) des LP-Fehler-Filters aktualisiert, und einer Subtrahiereinheit, welche das LP-Fehler-Filter Ausgangssignal $e(n)$ von dem ersten Sprachsignal $y(n)$ in einem Subtrahierer subtrahiert und nach der Subtraktion den Rest als zweites Sprachsignal $x(n) = y(n) - e(n)$ ausgibt, in welchem das Geräusch unterdrückt ist, wobei eine Geräuschschätzungseinheit (GSE) vorgesehen ist, welche auf Basis der internen Signale des LP-Fehler-Filters ein ein Geräuschleistungssignal σ_n^2 und ein Sprachleistungssignal σ_x^2 erzeugt, das der Koeffizientenberechnungseinheit (KBE) zugeführt ist und von dieser im Sinne einer Optimierung der Geräuschunterdrückung berücksichtigt wird.

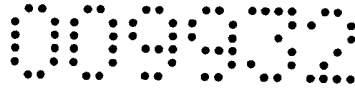




ZUSAMMENFASSUNG

Eine Vorrichtung zur Geräuschunterdrückung mit einer linearen Prädiktions-Analysierschaltung mit einem LP-Fehler-Filter (LFF), das auf Basis eines ersten, von Geräusch überlagerten Sprachsignals $y(n) = x(n) + e(n)$ ein LP-Fehler-Filter Ausgangssignal $e(n)$ erzeugt, einer Koeffizientenberechnungseinheit, welche die Koeffizienten des LP-Fehler-Filters auf Basis der internen Signale (inklusive des Ein- und Ausgangssignals $y(n)$ und $e(n)$) des LP-Fehler-Filters aktualisiert, und einer Subtrahiereinheit, welche das LP-Fehler-Filter Ausgangssignal $e(n)$ von dem ersten Sprachsignal $y(n)$ in einem Subtrahierer subtrahiert und nach der Subtraktion den Rest als zweites Sprachsignal $\hat{x}(n) = y(n) - e(n)$ ausgibt, in welchem das Geräusch unterdrückt ist, wobei eine Geräuschschätzungseinheit (GSE) vorgesehen ist, welche auf Basis der internen Signale des LP-Fehler-Filters ein ein Geräuschleistungssignal σ_n^2 und ein Sprachleistungssignal σ_x^2 erzeugt, das der Koeffizientenberechnungseinheit (KBE) zugeführt ist und von dieser im Sinne einer Optimierung der Geräuschunterdrückung berücksichtigt wird.

Fig. 6



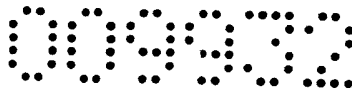
VORRICHTUNG ZUR GERÄUSCHUNTERDRÜCKUNG BEI EINEM AUDIOSIGNAL

Die Erfindung bezieht sich auf eine Vorrichtung zur Geräuschunterdrückung mit einer linearen Prädiktions-Analysierschaltung mit einem LP-Fehler-Filter (LFF), das auf Basis eines ersten, von Geräusch überlagerten Sprachsignals $y(n) = x(n) + \varepsilon(n)$ ein LP-Fehler-Filter Ausgangssignal $e(n)$ erzeugt, einer Koeffizientenberechnungseinheit, welche die Koeffizienten des LP-Fehler-Filters auf Basis der internen Signale (inklusive des Ein- und Ausgangssignals $y(n)$ und $e(n)$) des LP-Fehler-Filters aktualisiert, und einer Subtrahiereinheit, welche das LP-Fehler-Filter Ausgangssignal $e(n)$ von dem ersten Sprachsignal $y(n)$ in einem Subtrahierer subtrahiert und nach der Subtraktion den Rest als zweites Sprachsignal $x^{\wedge}(n) = y(n) - e(n)$ ausgibt, in welchem das Geräusch unterdrückt ist.

Die Geräuschunterdrückung bei Audiosignalen, insbesondere auch bei Sprachsignalen gewinnt immer mehr an Bedeutung, so in der Telephonie, der automatischen Spracherkennung oder, als nur eines von vielen weiteren Beispielen, bei digitalen Hörgeräten. Geräusche, die es zu unterdrücken gilt, sind in erster Linie nicht-weißes Rauschen, nichtstationäre Geräusche und stark impulsartige Geräusche.

Verfahren zur Geräuschunterdrückung für Audiosignale sind in einer großen Vielfalt bekannt geworden und als Beispiele seien genannt: Verfahren, bei welchen leise Audiosignale zunächst angehoben und später wieder abgesenkt werden, wie Pre-/Deemphase für den Rundfunk oder die Dolby-Rauschunterdrückungsverfahren für Tonträgeraufzeichnungen. Weiters Verfahren der spektralen Subtraktion, bei welchen z.B. in Gesprächspausen das Geräusch geschätzt und dann von dem Eingangssignal abgezogen wird. Zu den letztgenannten Verfahren zählen auch Wiener-Filter sowie Ephraim-Malah Filter mit adaptiver Verstärkung für in mehrere Transformationskanäle aufgespaltete Signale. Die bekannten Verfahren sind zum Teil nicht sehr wirkungsvoll, da sie von einem sehr vereinfachten Modell des Geräuschsignals ausgehen, oder sie führen auf Grund einer blockweisen Verarbeitung des Eingangssignals zu Artefakten, die sich als unangenehme Nebengeräusche bemerkbar machen, als so genannte Musical Tones, die im Signal nach der Geräuschreduktion verbleiben. Viele Verfahren führen auch zu einer relativ großen Verzögerung des Ausgangssignals.

Der Stand der Technik, von welchem die vorliegende Erfindung ausgeht, sind lineare Prädiktionsfilter (LP Filter), in direkter oder in Lattice-Form (Kreuzglied Kettenfilter), bei welchen Eigenschaften des gesamten Eingangssignals zur Einstellung der Filterkoeffizienten herangezogen werden. Eine entsprechende Vorrichtung, wie z.B. in der US 2001/0005822



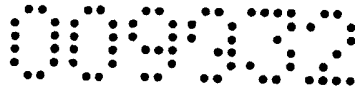
Als geoffenbart, besteht aus einem Lattice Filter, dem ein Eingangssignal $y(n)$ zugeführt ist, welches eine Sprach/Audiokomponente sowie eine Geräuschkomponente aufweist. Für die Einstellung der Komponenten ist eine Koeffizientenberechnungseinheit KBE vorgesehen, welcher die Vorwärts- und Rückwärtsfehlersignale, diese beinhalten in der ersten Stufe auch das Eingangssignal, zugeführt werden. Die Koeffizientenberechnungseinheit liefert dem Filter sodann immer im Sinne einer Minimierung des Prädiktionsfehlers aktualisierte Filterkoeffizienten. Eine Geräuschreduktion unter Verwendung linearer Prädiktionsfilter ist unter anderem auch in der GB 1 520 148 A oder in der US 4, 587, 620 geoffenbart. Die Verfahren bzw. Vorrichtungen nach dem Stand der Technik gehen immer von dem Eingangssignal aus, ohne dass die besonderen Eigenschaften des Sprachsignals einerseits und des Geräusches andererseits Berücksichtigung finden.

Als Publikationen auf diesem Gebiet seien weiters nachstehend genannt:

- [1] J. D. Markel and A. H. Gray, Jr., *Linear Prediction of Speech*. Berlin, Heidelberg, New York: Springer, 1976.
- [2] J. I. Makhoul and L.K. Cosell, "Adaptive lattice analysis of speech," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29, pp.654–659, June 1981.
- [3] M. L. Honig and D. G. Messerschmitt, *Adaptive Filters: Structures, Algorithms, and Applications*. Boston-The Hague-London-Lancaster: Kluwer Academic Publishers, 1984.
- [4] A. Kawamura, K. Fujii, Y. Itoh, and Y. Fukui, "A noise reduction method based on linear prediction analysis," *Electronics and Communications in Japan, Part 3*, vol. 86, no. 3, pp. 1-10, 2003.
- [5] M. H. Savoji, "Effective noise reduction of speech signals using adaptive lattice filtering, segmentation and soft decision," in *IEE Colloquium on New Directions in Adaptive Signal Processing*, pp. 7/1-7/5, Feb. 1993.

Eine Aufgabe der Erfindung liegt in der Schaffung einer Vorrichtung bzw. eines Verfahrens zur Geräuschunterdrückung für Audiosignale, insbesondere für Sprachsignale, welche ein praktisch unverzögertes Ausgangs-Audiosignal liefert, das auch keine störenden Artefakte mit sich führt.

Diese Aufgabe wird mit einer Vorrichtung der eingangs genannten Art gelöst, bei welcher erfindungsgemäß eine Geräuschschätzungseinheit (GSE) vorgesehen ist, welche auf Basis



der internen Signale des LP-Fehler-Filters ein Geräuschleistungssignal σ_n^2 und ein Sprachleistungssignal σ_x^2 , erzeugt, das der Koeffizientenberechnungseinheit (KBE) zugeführt ist und von dieser im Sinne einer Optimierung der Geräuschunterdrückung berücksichtigt wird und/oder eine Sprachaktivitätsschätzungseinheit (SAE) vorgesehen ist, welche auf Basis der internen Signale des LP-Fehler-Filters ein Sprachaktivitätssignal v erzeugt, das der Koeffizientenberechnungseinheit (KBE) zugeführt ist und von dieser im Sinne einer Optimierung der Geräuschunterdrückung berücksichtigt wird.

Bei einer zweckmäßigen Weiterbildung der Erfindung ist vorgesehen, dass die Sprachaktivitätsschätzungseinheit einen Geräuschunterdrückungsfaktor k_n bildet, welcher einem Eingang eines ersten Multiplizierers zugeführt ist und dessen anderem Eingang das Ausgangssignal des LP-Fehler-Filters zugeführt ist und der vor dem Subtrahierer liegt.

Es kann weiters vorteilhaft sein, wenn die Sprachaktivitätsschätzungseinheit einen Gesamtsignalfaktor k_g bildet, welcher einem Eingang eines zweiten Multiplizierers zugeführt ist dessen anderem Eingang das Ausgangssignal des Subtrahierers zugeführt ist.

Bei einer praxisgerechten Variante kann man vorsehen, dass das LP-Fehler-Filter aus einem Lattice Filter (Fig. 1) besteht, wobei die Vorwärts- und Rückwärtsfehlersignale die internen Signale des LP-Fehler-Filters darstellen.

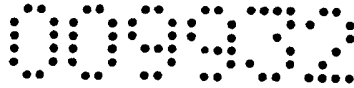
Andererseits ist es mit Vorteil auch möglich, dass das LP-Fehler-Filter als Filter in direkter Filterform aus dem Eingangssignal ein Prädiktionssignal an seinem Ausgang erzeugt, und ein Subtrahierer das Prädiktionssignal vom Eingangssignal abzieht und somit das Ausgangssignal des LP-Fehler-Filters $e(n)$ erzeugt; wobei die verzögerten Abtastwerte des Eingangssignals (vgl. Gleichung 1) sowie das Ausgangssignal des Subtrahierers $e(n)$ den internen Signalen des LP-Fehler-Filters entsprechen.

Bei einer anderen vorteilhaften Variante ist vorgesehen, dass die Koeffizientenberechnungseinheit (KBE) dazu eingerichtet ist, die korrigierte Fehlervarianz \hat{q}_0 gemäß

$$\hat{q}_0 = q_0 - \sigma_n^2$$

und den korrigierten Reflexionskoeffizienten \hat{k} gemäß

$$\hat{k}_1 = -\frac{\hat{r}_0}{\hat{q}_0} = -\frac{r_0}{q_0 - \sigma_n^2}$$



zu ermitteln.

Wieder eine weitere zweckmäßige Variante zeichnet sich dadurch aus, dass die Koeffizientenberechnungseinheit (KBE) dazu eingerichtet ist, die Fehlerkorrelation gemäß

$$\tilde{r}_m(n) = \lambda_r \tilde{r}_m(n-1) + f_m(n) b_m(n-1)$$

und die Fehlervarianz gemäß

$$\tilde{q}_m(n) = \lambda_q \tilde{q}_m(n-1) + \frac{1}{2}(f_m^2(n) + b_m^2(n-1))$$

zu ermitteln.

Bei einer günstigen Ausführung kann für die Leistungsschätzung von $\tilde{q}(n)$ ein einpoliger Tiefpass und ein für die Korrelationsschätzung $\tilde{r}(n)$ zweipoliger Tiefpass vorgesehen sein.

Zur Verbesserung der Geräuschunterdrückung bei entsprechend erhöhtem Aufwand sieht eine weitere Variante der Erfindung eine Kaskade vor, die aus zumindest zwei hintereinander geschalteten Vorrichtungen besteht, welche entsprechen der oben genannten Erfindungsmerkmale ausgebildet sind.

Die Erfindung samt weitere Vorteile ist im folgenden an Hand beispielsweise Ausführungsformen näher erläutert, die in der Zeichnung veranschaulicht sind. In dieser zeigen

Fig. 1 ein Lattice-Filter nach dem Stand der Technik,

Fig. 2 aus dem geräuschbehafteten Signal ohne Korrektur berechnete Werte des Reflexionskoeffizienten in einem Diagramm,

Fig. 3 a und b den Frequenzgang von im Rahmen der Erfindung verwendeten Tiefpässen,

Fig. 4 an Hand des zeitlichen Verlaufes sowie des Spektrogramms eines geräuschbehafteten Eingangssignal (oben) und dessen Geräuschbefreiung (unten) nach der Erfindung,

Fig. 5 das Blockschaltbild einer prinzipiellen Vorrichtung zur Geräuschbefreiung nach dem Stand der Technik

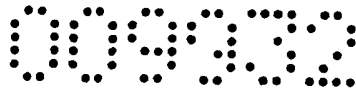


Fig. 6 das Blockschaltbild einer ersten Ausführungsform einer Vorrichtung nach der Erfindung,

Fig. 7 das Blockschaltbild einer zweiten Ausführungsform einer Vorrichtung nach der Erfindung und

Fig. 8 das Blockschaltbild einer dritten Ausführungsform einer Vorrichtung nach der Erfindung.

Lineare Prädiktion wird üblicherweise auf ein Sprachsignal $x(n)$ angewendet, beispielsweise um die Varianz eines Sprachsignals für dessen Übertragung zu reduzieren. Um einen Signalwert vorher zu sagen, werden langsam zeitveränderliche FIR-Filter (finite impulse response filter, Filter mit endlicher Impulsantwort) niedriger Ordnung verwendet:

$$\hat{x}(n) = \sum_{i=1}^M b_i(n)x(n-i). \quad (1)$$

Hier bedeuten M die Ordnung des LP-Filters und $b_i(n)$ die Filterkoeffizienten, die auf Basis der Signaleigenschaften geschätzt und rahmenweise, z.B. alle 10 ms, aufgefrischt werden. Algorithmen welche die Koeffizienten $b_i(n)$ für das Filter direkt ergeben sind das so genannte „Autokorrelationsverfahren“ oder das „Kovarianzverfahren“. Die zur Modellierung der spektralen Hüllkurve üblicherweise verwendete Filterordnung liegt, abhängig von der Abtastrate, bei $M = 10 \dots 20$.

Ein dem eben beschriebenen direkten FIR-Filter, auf welches sich z.B. Patentanspruch 7 bezieht, ebenbürtiges ist das bereits genannte Lattice-Filter, welches einen unmittelbaren Bezug zum menschlichen Stimmtrakt besitzt [1].

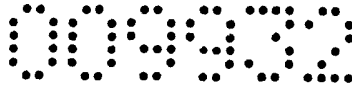
Charakterisiert ist ein solches Filter durch die Gleichungen

$$f_0(n) = b_0(n) = x(n),$$

$$f_m(n) = f_{m-1}(n) + k_m(n)b_{m-1}(n-1), \quad (2)$$

$$b_m(n) = b_{m-1}(n-1) + k_m(n)f_{m-1}(n). \quad (3)$$

welche zu jedem Zeitpunkt n für alle Gliederstufen $m = 1 \dots M$ berechnete werden. Dabei bedeuten $f_m(n)$ und $b_m(n)$ den Vorwärts- bzw. Rückwärts-Fehler in der Stufe m zur Zeit n und



$k_m(n)$ die Reflexionskoeffizienten des Filters. Hier sei angemerkt, dass bei einer verallgemeinerten Darstellung die Reflexionskoeffizienten in (2) und (3) unterschiedlich sind, hier jedoch gleiche Reflexionskoeffizienten für vorwärts und rückwärts verwendet werden. In Fig. 1 ist ein Lattice-Filter wie eben beschrieben schematisch dargestellt. Auf ein solches Filter bezieht sich beispielsweise Patentanspruch 6.

Der Vorwärtsfehler in der Stufe M ist das Vorhersage-Fehlersignal des LP-Filters:

$$f_M(n) = e(n) = x(n) - \hat{x}(n). \quad (4)$$

Optimale Reflexionskoeffizienten k_m zur Minimierung des quadratischen Mittelwertes des Vorhersagefehlers eines unverzerrten Signals ergeben sich als:

$$k_m(n) = -\frac{r_{m-1}(n)}{q_{m-1}(n)}. \quad (5)$$

mit den erwarteten Werten für Vorwärts- und Rückwärts- Fehlerkorrelation bzw. Leistung zu:

$$r_m(n) = E\{f_m(n)b_m(n-1)\}, \quad (6)$$

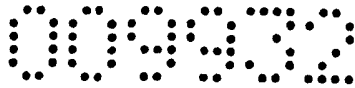
$$q_m(n) = \frac{1}{2}E\{f_m^2(n) + b_m^2(n-1)\}. \quad (7)$$

Üblicherweise werden die Erwartungswertoperatoren E in (6) und (7) unter Verwendung Tiefpass-gefilterter Augenblickswerte von $f_m(n)b_m(n-1)$ bzw. $f_m^2(n) + b_m^2(n-1)$ ausgewertet, beispielsweise mit Hilfe einpoliger rekursiver Tiefpassfilter („lossy integration“, verlustbehaftete Integration, siehe weiter unten).

Somit erfolgt das Adaptieren des Lattice-Filters zur Formung des sich langsam ändernden Eingangs-Sprachsignals durch Berechnung der Gleichungen (6), (7) und (5) für jeden Zeitpunkt n nach dem Filtern - im Gegensatz zur dem rahmenweisen Aktualisieren der Koeffizienten bei der direkten Form des LP-Filters entsprechend Gleichung (1).

Was die Geräuschverminderung betrifft, sei angenommen, dass ein beobachtetes Signal $y(n)$ eines additiven linearen Geräuschmodells vorliege:

$$y(n) = x(n) + \varepsilon(n), \quad (8)$$



wobei $x(n)$ die Sprachkomponente und $\varepsilon(n)$ eine additive Hintergrund-Geräuschkomponente sei. Gegenstand der Geräuschreduktion ist es, eine gute Schätzung für die Sprachsignal-Komponente $x(n)$ zu liefern. Für den vorliegenden einkanaligen Fall beruht diese Schätzung ausschließlich auf der Beobachtung des geräuschbehafteten Signals $y(n)$ d.h. es wird keine zusätzliche Information verwendet, wie beispielsweise ein zweites Signal eines Mikrophons, welches lediglich die Hintergrundgeräusche aufnimmt.

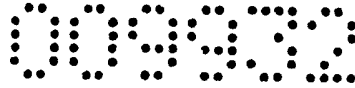
Eine Verringerung des Hintergrundgeräusches in Sprachsignalen unter Verwendung linearer Prädiktionsfilterung kann auf der Annahme basieren, dass die Sprachsignalkomponente gut vorhersagbar ist, wogegen die Geräuschkomponente diese Eigenschaft nicht besitzt. Somit kann das vorher gesagte Signal $\hat{x}(n)$ als Schätzung für die Sprachkomponente genommen werden. Während bei den Prädiktionsfiltern in der direkten Form in Gleichung (1) der Ausgang unmittelbar das vorhergesagte Signal ist, wird bei den Lattice-Filtern $\hat{x}(n)$ wirksam als die Differenz zwischen dem Eingangssignal und dem Ausgang des Vorwärts-Prädiktionspfades berechnet:

$$\hat{x}(n) = y(n) - e(n) \quad (9)$$

$\hat{x}(n)$ ist die Schätzung der Sprachkomponente mit $e(n) = f_M(n)$. Vergleiche Gleichung (4).

Wenn man ein Lattice-LP-Filter für die Geräuschverminderung nach A. Kawamura et al. verwendet, wo ein Filter hoher Ordnung ($N = 256$) benutzt wird, um die spektrale Feinstruktur der Sprache zu modellieren, kann man beobachten, dass $r_m(n)$ eine große Varianz aufgrund der Geräuschsignalkomponente in den höheren Filterstufen aufweist. Es wird vorgeschlagen, die entsprechende Varianz der Reflektionskoeffizienten dadurch zu reduzieren, dass man einen festen (großen) Wert für die Leistungsschätzungen $q_m(n) = G$ verwendet. Die Verringerung des Betrags der Reflektionskoeffizienten oder die Verringerung der Radien der Nullstellen der LP Filterübertragungsfunktion wurde auch für andere Zwecke vorgeschlagen, z.B. für die Modellierung der spektralen Hüllkurve oder für eine genauere Schätzung der Formanten.

Eine Verringerung des Betrags der Reflektionskoeffizienten kann durch die Herleitung von Schätzern für r_m und q_m erfolgen, die den mittleren quadratischen Schätzfehler minimieren. Zur Berechnung der Reflektionskoeffizienten für das Lattice Filter oder der partiellen Korrelationen, welche gleich $-k_m$ sind und auf den Schätzungen von einem Geräuschsignal beruhen, sei für das zusätzliche Geräuschsignal weißes Rauschen mit der Varianz σ_n^2 angenommen, das auch mit $x(n)$ nicht korreliert sein soll. Dies stellt das am wenigsten informierte Modell (maximale Entropie) dar.



Es ist zwar diese Annahme für das Umgebungsgeräusch im Allgemeinen nicht realistisch, doch beweist sie die Notwendigkeit einer Korrektur der Reflektionskoeffizienten.

Die Schätzungen für die Korrelation in Gleichung (6) und die Leistung in Gleichung (7) aus der Berechnung der Reflektionskoeffizienten in Gleichung (5) beruhen nun auf dem geräuschbehafteten beobachteten Signal $y(n)$ und man kann die Notwendigkeit eines Korrekturterms zeigen, um die Schätzungen für die Reflektionskoeffizienten \hat{k}_m bezogen auf das geräuschfreie Signal $x(n)$ zu gewinnen.

In besonderen erhält man für die Schätzung der Reflektionskoeffizienten in der ersten Filterstufe $m = 1$ den folgenden Erwartungswert für r_0 :

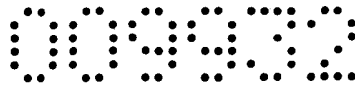
$$\begin{aligned}
 r_0 &= E\{f_0(n) b_0(n-1)\} \\
 &= E\{y(n) y(n-1)\} \\
 &= E\{(x(n) + \varepsilon(n))(x(n-1) + \varepsilon(n-1))\} \\
 &= E\{x(n) x(n-1)\}
 \end{aligned} \tag{10}$$

Für die Fehlervarianz q_0 in der ersten Filterstufe erhält man

$$\begin{aligned}
 q_0 &= \frac{1}{2} E\{f_0^2(n) + b_0^2(n-1)\} \\
 &= \frac{1}{2} E\{y^2(n) + y^2(n-1)\} \\
 &= \frac{1}{2} E\{(x(n) + \varepsilon(n))^2 + (x(n-1) + \varepsilon(n-1))^2\} \\
 &= \frac{1}{2} E\{x^2(n) + x^2(n-1)\} + \sigma_n^2
 \end{aligned} \tag{11}$$

Der resultierende Fehler in den Werten der Reflektionsfaktoren ist in Fig. 2 dargestellt. Genauer gesagt sind hier aus dem geräuschbehafteten Signal ohne Korrektur berechnete Werte des Reflexionskoeffizienten k_1 als Funktion eines a priori Signal/Geräusch-Abstandes für unterschiedliche Werte der Autokorrelation $\rho_{xx}(1)$ des ungestörten Signals $x(n)$ veranschaulicht.

Was die auf das geräuschfreie Signal $x(n)$ bezogenen Reflektionskoeffizienten betrifft, kann die Korrelationsschätzung aus der geräuschbehafteten Beobachtung unverändert verwendet werden, das heißt $\hat{r}_0 = r_0$ wogegen die berechnete Fehlerleistungsschätzung zu korrigieren ist als



$$\hat{q}_0 = q_0 - \sigma_n^2 \quad (12)$$

und der korrigierte Reflektionskoeffizient berechnet wird als

$$\hat{k}_1 = -\frac{\hat{r}_0}{\hat{q}_0} = -\frac{r_0}{q_0 - \sigma_n^2}. \quad (13)$$

Wenn man einführt: $\gamma = \frac{E\{y^2\}}{\sigma_n^2}$ wobei $\gamma - 1 = \frac{E\{y^2\} - \sigma_n^2}{\sigma_n^2}$ der a posteriori bestimmte Signalstörabstand ist und beachtet wird dass $q_0 = \frac{1}{2}E\{f_0^2(n) + b_0^2(n-1)\}$ kann man diese Gleichung erneut anschreiben als

$$\hat{k}_1 = -\frac{1}{1 - \frac{\sigma_n^2}{q_0}} \frac{r_0}{q_0} = \frac{1}{1 - \frac{1}{\gamma}} k_1. \quad (14)$$

Das bedeutet eine Skalierung des Reflektionskoeffizienten k_1 , der ursprünglich für das gestörte Signal $y(n)$ berechnet wurde unter Verwendung der Gleichungen (5), (6) und (7) mit einem Faktor $\frac{1}{1 - \frac{1}{\gamma}}$

Eine Schätzung der Geräuschleistung σ_n^2 kann erfolgen anhand der Leistung des Ausgangssignals $e(n)$ des LP-Fehler-Filters,

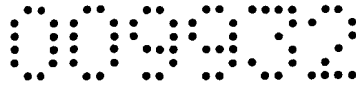
$$\sigma_n^2 = E\{e(n)\},$$

für das Lattice Filter ist eine möglich Auswertung des Erwartungswerts gegeben durch die Leistungsschätzung in der letzten Stufe des Lattice Filters $q_{M-1}(n)$:

$$\sigma_n^2 = q_{M-1}(n),$$

oder, bei Verwendung der Sprachaktivitätsschätzung (siehe unten) anhand der Leistungsschätzung des Gesamteingangssignales bei fehlender Sprachaktivität:

$$\sigma_n^2 = q_0(n), \text{ wenn } v \approx 0$$



Für die höheren Lattice Stufen $m = 2, 3, \dots$ kann die Gleichung (14) verallgemeinert werden, wodurch sich eine Korrektur der übrigen Reflektionskoeffizienten \hat{k}_m ergibt.

Dessen ungeachtet kann man aus dem obigen schließen, dass eine Verminderung der Größe der Reflektionskoeffizienten, das heißt eine Verringerung des Verhältnisses zwischen Korrelation und Leistungsschätzung für die Vorhersage eines Signals $x(n)$ bei Beobachtung eines Signal $y(n)$, welches zusätzliches Geräusch enthält, zweckmäßig ist. Das Auffinden der Korrekturgrößen setzt eine verlässliche Schätzung der Geräuschleistung σ_n^2 voraus. Weiters berücksichtigt das Modell bis jetzt nicht irgendwelche Kenntnisse bezüglich der Eigenschaften der Sprache und des zu erwartenden Geräuschsignals.

Die Erfindung schafft nun ein Verfahren bzw. eine Vorrichtung, mit welcher man, basierend auf einfachen Annahmen über die Änderung der Korrelation und der Leistung der Sprach- und Geräuschsignale nach der Zeit eine Korrektur der Reflektionsfaktoren erhält.

Wie oben ausgeführt, beruhen die Schätzungen der Fehlerkorrelation (6) und der Fehlervarianz (7) üblicherweise auf einer Tiefpassfilterung der Augenblickswerte. Oft wird auch eine einpolige Tiefpassfilterung (verlustbehaftet Integration) verwendet:

$$\tilde{r}_m(n) = \lambda_r \tilde{r}_m(n-1) + f_m(n) b_m(n-1), \quad (15)$$

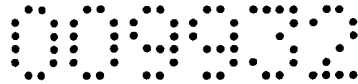
$$\tilde{q}_m(n) = \lambda_q \tilde{q}_m(n-1) + \frac{1}{2} (f_m^2(n) + b_m^2(n-1)), \quad (16)$$

mit den gleichen Polen bzw. Integrationsfaktoren $\lambda_r = \lambda_q$ sowohl für die Korrelations- als auch die Leistungsschätzung.

In Einklang mit der Erfindung sind verschiedene Pollagen $\lambda_q \geq \lambda_r$ erlaubt. Die resultierenden Filterfunktionen

$$H_r(z) = \frac{1}{1 - \lambda_r z^{-1}}, \quad H_q(z) = \frac{1}{1 - \lambda_q z^{-1}}, \quad (17)$$

für $\lambda_r = 0,99608$ und $\lambda_q = 0,99843$ und eine Abtastrate von 16 kHz sind in Fig.3a dargestellt. Man kann sehen, dass das Verhältnis von $\tilde{r}_m(n)$ und $\tilde{q}_m(n)$ bei geringeren Frequenzen beeinflusst wird, das heißt für langsam sich ändernde Korrelation und Leistung, wogegen



für raschere Änderungen (über $\approx 10\text{Hz}$) das Verhältnis verglichen mit den Schätzungen mit $\lambda_r = \lambda_q$ ungeändert bleibt. Unter der Annahme, dass diese Parameter sich für das Sprachsignal rascher ändern (unter der Annahme beispielsweise einer Phonemrate von 10 pro Sekunde) als für das Geräuschsignal (stationäres oder langsam sich mit der Zeit änderndes Geräusch) wird das sich ergebende Lattice Prädiktionsfilter die Sprachsignalkomponente gut vorhersagen, wogegen die Geräuschkomponente unterdrückt wird.

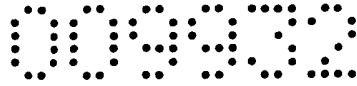
Was impulsartige Geräusche betrifft, kann man vorsehen, das Verhältnis zwischen Korrelation und Leistungsschätzung auch für hohe Frequenzen zu reduzieren, was beispielsweise durch Verwendung eines zweiten Pols in dem Tiefpassfilter für die Korrelation $H_r(z)$ erfolgen kann. Eine entsprechende Übertragungsfunktion ist in Fig.3b dargestellt.

Im Einzelnen zeigen die Figuren 3a und b die Frequenzgänge eines Tiefpassfilters für eine Fehlerkorrelation $H_r(z)$ (ausgezogene Linien) und die Varianz $H_q(z)$ (punktierte Linien) für in Fig. a zwei einpolige Tiefpassfilter mit $\lambda_r = 0,99608$ und $\lambda_q = 0,99843$ bzw. Figur 3b einen einpoligen Tiefpass für die Leistungsschätzung von $\tilde{q}(n)$ mit $\lambda_q = 0,99843$ und einen zweipoligen Tiefpass für die Korrelationsschätzung $\tilde{r}(n)$ mit $\lambda_{r1} = 0,99608$ und $\lambda_{r2} = 0,9$. Je größer der Abstand zwischen den beiden Übertragungsfunktionen umso stärker ist die Geräuschunterdrückung.

Um eine gute Reduzierung des Geräusches zu erreichen, kann unter diesen Umständen die Ordnung M des LP Filters überraschend niedrig gewählt werden, sogar niedriger als die üblicherweise zur Modellierung der spektralen Hüllkurve von Sprachsignalen verwendete Ordnung. Beispielsweise wurde in dem in Fig. 4a gezeigten Beispiel ein Prädiktor mit der Ordnung $M = 10$ für ein Signal mit einer Abtastrate von 16 kHz verwendet. Dieses Beispiel enthält ein mehrmaliges Auftreten von starken, nicht stationären Geräuschbursts, die dank der Erfindung gut beseitigt werden. Das gezeigte Geräusch stammt aus einer Fabrikhallen-Umgebung, das heißt einer äußerst ungünstigen akustischen Umgebung.

Die Wirksamkeit der Geräuschunterdrückung kann gesteuert werden durch die Einstellung unterschiedlicher Werte für λ_r (bzw. λ_{r1} und λ_{r2}) und λ_q . Wobei diese in Abhängigkeit von der Signalleistung und der Geräuschleistung gewählt werden:

$$(\lambda_r, \lambda_q) = f(\sigma_x^2, \sigma_n^2), \text{ bzw. } (\lambda_{r1}, \lambda_{r2}, \lambda_q) = g(\sigma_x^2, \sigma_n^2). \quad (18)$$



Weiters ist es zweckmäßig, die Wirksamkeit der Geräuschunterdrückung anhand einer Schätzung der Sprachaktivität zu steuern. Bei der Verwendung eines LP-Fehler-Filters bietet sich (vgl. auch [5]) die Schätzung der wahrscheinlichen Sprachaktivität als reeller Zahl im Wertebereich von 0 bis 1 anhand der Leistungen des Filtereingangssignals und des Filterausgangssignals an:

$$v = \frac{E\{y^2(n)\} - E\{e^2(n)\}}{E\{e^2(n)\}}$$

für ein Lattice Filter ist eine mögliche Auswertung der Erwartungswerte gegeben durch

$$v = \frac{q_0(n) - q_{M-1}(n)}{q_{M-1}(n)}$$

Zur Steuerung der Geräuschunterdrückung kann dann ein Faktor

$$k_n = 1 - v ,$$

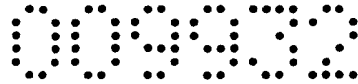
für das Ausgangssignal des Lattice Filters $e(n)$, und/oder ein Faktor

$$k_g = v ,$$

für das Ausgangssignal verwendet werden.

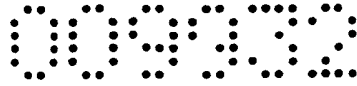
Das LP-Fehler-Filter kann als Filter in direkter Filterform (DFF) ausgebildet sein, welches aus dem Eingangssignal ein Prädiktionssignal an seinem Ausgang erzeugt, wobei ein Subtrahierer das Prädiktionssignal vom Eingangssignal abzieht und somit das Ausgangssignal des LP-Fehler-Filters $e(n)$ erzeugt. Die verzögerten Abtastwerte des Eingangssignals (vgl. Gleichung 1) sowie das Ausgangssignal des Subtrahierers $e(n)$ entsprechen den internen Signalen des LP-Fehler-Filters.

Ein wichtiges Merkmal der erfindungsgemäßen Geräuschunterdrückung besteht in der an die Eigenschaften des Sprach- und des Geräuschsignales angepassten Auswertung der Erwartungswertoperatoren und damit der optimalen Einstellung der Filterkoeffizienten für das lineare Prädiktionsfilter, sowie in der Sprachaktivitätsschätzung und deren Verwendung bei der Schätzung des Geräuschsignals, und zur Steuerung der Wirksamkeit der Geräuschunterdrückung und der Amplitude des Ausgangssignales.



Wenngleich es klar sein sollte, dass der Rechenaufwand mit der gewählten Filterordnung steigt und daher der Rechenaufwand unter Umständen größer sein kann, als jener bei Verwendung einer schnellen Fouriertransformation, liegt ein wesentlicher Vorteil der Erfindung darin, dass sie eine Geräuschreduktion ohne Verzögerung des Sprachsignals ermöglicht, was ein besonderer Vorteil ist, vor allem beim Einsatz in Hörgeräten.

Wien, den 15. September 2006



ANSPRÜCHE

1. Vorrichtung zur Geräuschunterdrückung mit einer linearen Prädiktions-Analysierschaltung mit einem LP-Fehler-Filter (LFF), das auf Basis eines ersten, von Geräusch überlagerten Sprachsignals $y(n) = x(n) + \varepsilon(n)$ ein LP-Fehler-Filter Ausgangssignal $e(n)$ erzeugt, einer Koeffizientenberechnungseinheit, welche die Koeffizienten des LP-Fehler-Filters auf Basis der internen Signale (inklusive des Ein- und Ausgangssignals $y(n)$ und $e(n)$) des LP-Fehler-Filters aktualisiert, und einer Subtrahiereinheit, welche das LP-Fehler-Filter Ausgangssignal $e(n)$ von dem ersten Sprachsignal $y(n)$ in einem Subtrahierer subtrahiert und nach der Subtraktion den Rest als zweites Sprachsignal $\hat{x}(n) = y(n) - e(n)$ ausgibt, in welchem das Geräusch unterdrückt ist,

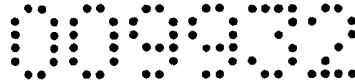
dadurch gekennzeichnet, dass

eine Geräuschschätzungseinheit (GSE) vorgesehen ist, welche auf Basis der internen Signale des LP-Fehler-Filters ein ein Geräuschleistungssignal σ_n^2 und ein Sprachleistungssignal σ_x^2 , erzeugt, das der Koeffizientenberechnungseinheit (KBE) zugeführt ist und von dieser im Sinne einer Optimierung der Geräuschunterdrückung berücksichtigt wird.

2. Vorrichtung zur Geräuschunterdrückung mit einer linearen Prädiktions-Analysierschaltung mit einem LP-Fehler-Filter, das auf Basis eines ersten, von Geräusch überlagerten Sprachsignals $y(n) = x(n) + \varepsilon(n)$ ein LP-Fehler-Filter Ausgangssignal $e(n)$ erzeugt, einer Koeffizientenberechnungseinheit, welche die Koeffizienten des LP-Fehler-Filters auf Basis der internen Signale des LP-Fehler-Filters aktualisiert, und einer Subtrahiereinheit, welche das LP-Fehler-Filter Ausgangssignal von dem ersten Sprachsignal subtrahiert und nach der Subtraktion den Rest als zweites Sprachsignalausgibt, $\hat{x}(n) = y(n) - e(n)$ in welchem das Geräusch unterdrückt ist,

dadurch gekennzeichnet, dass

eine Sprachaktivitätsschätzungseinheit (SAE) vorgesehen ist, welche auf Basis der internen Signale des LP-Fehler-Filters ein Sprachaktivitätssignal v erzeugt, das der Koeffizientenberechnungseinheit (KBE) zugeführt ist und von dieser im Sinne einer Optimierung der Geräuschunterdrückung berücksichtigt wird.



3. Vorrichtung zur Geräuschunterdrückung mit einer linearen Prädiktions-Analysierschaltung mit einem LP-Fehler-Filter, das auf Basis eines ersten, von Geräusch überlagerten Sprachsignals ein LP-Fehler-Filter Ausgangssignal erzeugt, einer Koeffizientenberechnungseinheit, welche die Koeffizienten des LP-Fehler-Filters auf Basis der internen Signale des LP-Fehler-Filters aktualisiert, und einer Subtrahiereinheit, welche das LP-Fehler-Filter Ausgangssignal von dem ersten Sprachsignal subtrahiert und nach der Subtraktion den Rest als zweites Sprachsignal ausgibt, in welchem das Geräusch unterdrückt ist,

dadurch gekennzeichnet, dass

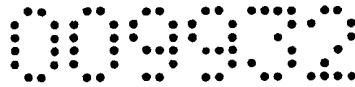
eine Geräuschschätzungseinheit (GSE) sowie eine Sprachaktivitätsschätzungseinheit (SAE) vorgesehen sind, beiden Einheiten die internen Signale des LP-Fehler-Filters zugeführt sind und die Einheiten auf deren Basis ein Geräuschleistungssignal σ_n^2 , ein Sprachleistungssignal σ_s^2 , und ein Sprachaktivitätssignal v erzeugen, welche der Koeffizientenberechnungseinheit (KBE) zugeführt sind und von dieser im Sinne einer Optimierung der Geräuschunterdrückung berücksichtigt werden.

4. Vorrichtung nach Anspruch 2 oder 3, **dadurch gekennzeichnet, dass** die Sprachaktivitätsschätzungseinheit (SAE) einen Geräuschunterdrückungsfaktor (k_n) bildet, welcher einem Eingang eines ersten Multiplizierers (MU1) zugeführt ist und dessen anderem Eingang das Ausgangssignal des LP-Fehler-Filters zugeführt ist und der vor dem Subtrahierer (SUB) liegt.

5. Vorrichtung nach einem der Ansprüche 2 bis 4, **dadurch gekennzeichnet, dass** die Sprachaktivitätsschätzungseinheit (SAE) einen Gesamtsignalfaktor (k_g) bildet, welcher einem Eingang eines zweiten Multiplizierers (MU2) zugeführt ist dessen anderem Eingang das Ausgangssignal des Subtrahierers (SUB) zugeführt ist.

6. Vorrichtung nach Anspruch 1 bis 5, **dadurch gekennzeichnet, dass** das LP-Fehler-Filter aus einem Lattice Filter (Fig. 1) besteht; wobei die Vorwärts- und Rückwärtsfehlersignale die internen Signale des LP-Fehler-Filters darstellen.

7. Vorrichtung nach Anspruch 1 bis 5, **dadurch gekennzeichnet, dass** das LP-Fehler-Filter als Filter in direkter Filterform (DFF) aus dem Eingangssignal ein Prädiktionssignal an seinem Ausgang erzeugt, und ein Subtrahierer das Prädiktionssignal vom Eingangssignal abzieht und somit das Ausgangssignal des LP-Fehler-Filters $e(n)$ erzeugt; wobei die verzögerten



Abtastwerte des Eingangssignals (vgl. Gleichung 1) sowie das Ausgangssignal des Subtrahierers $e(n)$ den internen Signalen des LP-Fehler-Filters entsprechen.

8. Vorrichtung nach einem der Ansprüche 1 bis 7, **dadurch gekennzeichnet**, dass die Koeffizientenberechnungseinheit (KBE) dazu eingerichtet ist, die korrigierte Fehlervarianz \hat{q}_0 gemäß

$$\hat{q}_0 = q_0 - \sigma_n^2$$

und den korrigierten Reflexionskoeffizienten \hat{k} gemäß

$$\hat{k}_1 = -\frac{\hat{r}_0}{\hat{q}_0} = -\frac{r_0}{q_0 - \sigma_n^2}$$

zu ermitteln.

9. Vorrichtung nach einem der Ansprüche 1 bis 7, **dadurch gekennzeichnet**, dass die Koeffizientenberechnungseinheit (KBE) dazu eingerichtet ist, die Fehlerkorrelation gemäß

$$\tilde{r}_m(n) = \lambda_r \tilde{r}_m(n-1) + f_m(n) b_m(n-1)$$

und die Fehlervarianz gemäß

$$\tilde{q}_m(n) = \lambda_q \tilde{q}_m(n-1) + \frac{1}{2}(f_m^2(n) + b_m^2(n-1))$$

zu ermitteln.

10. Vorrichtung nach einem der Ansprüche 1 bis 9, **dadurch gekennzeichnet**, dass für die Leistungsschätzung von $\tilde{q}(n)$ ein einpoliger Tiefpass und ein für die Korrelationsschätzung $\tilde{r}(n)$ zweipoliger Tiefpass vorgesehen ist.

11. Kaskade, bestehend aus zumindest zwei hintereinander geschalteten Vorrichtungen nach einem der Ansprüche 1 bis 9.

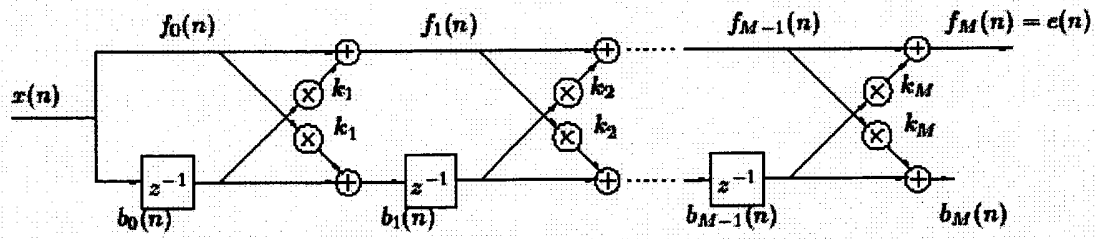


Fig. 1

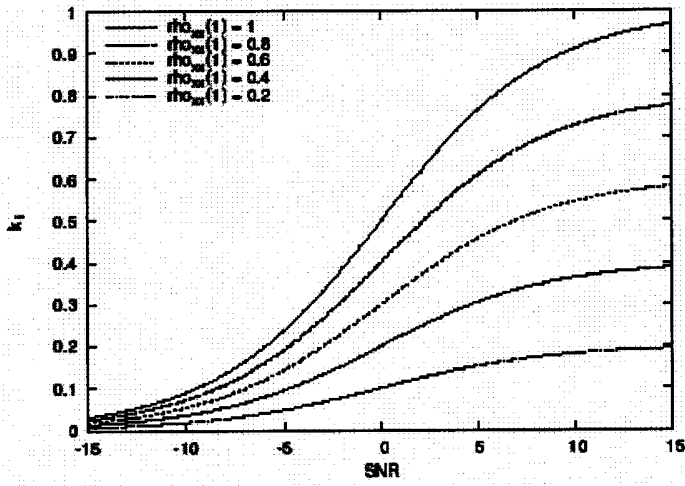


Fig. 2

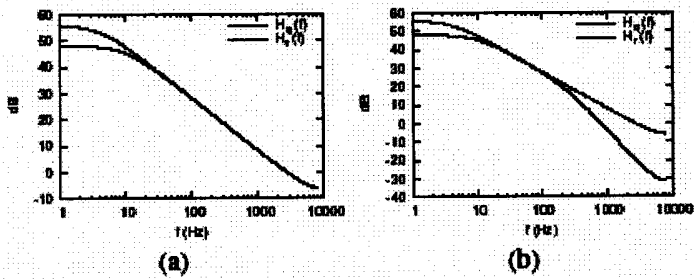


Fig. 3

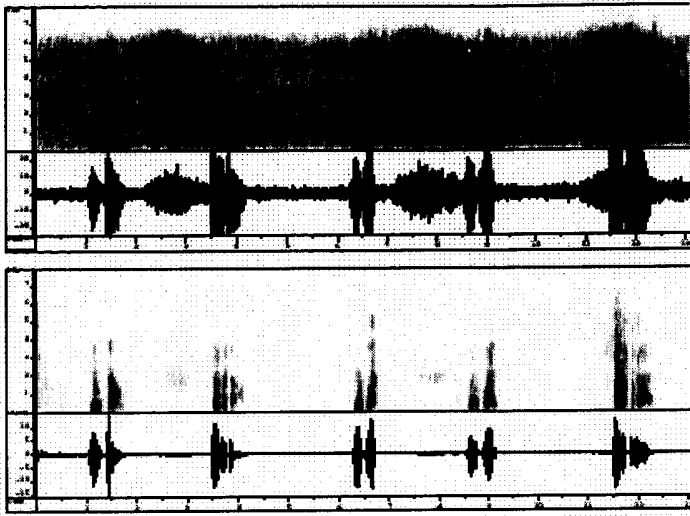
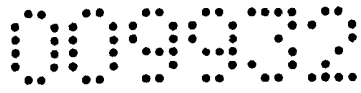


Fig. 4

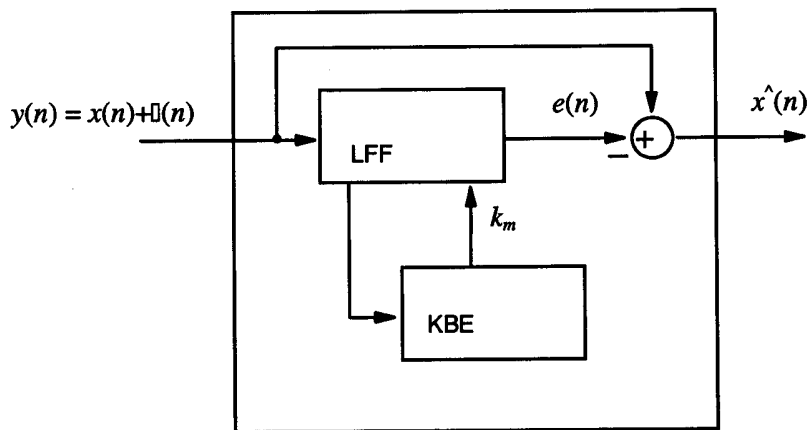


Fig. 5 (Stand der Technik)

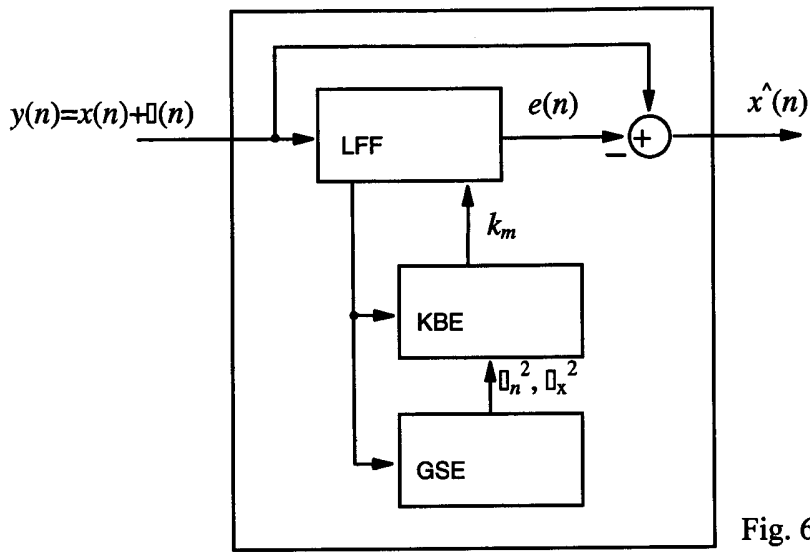
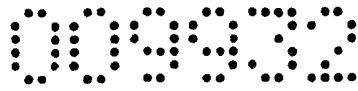


Fig. 6

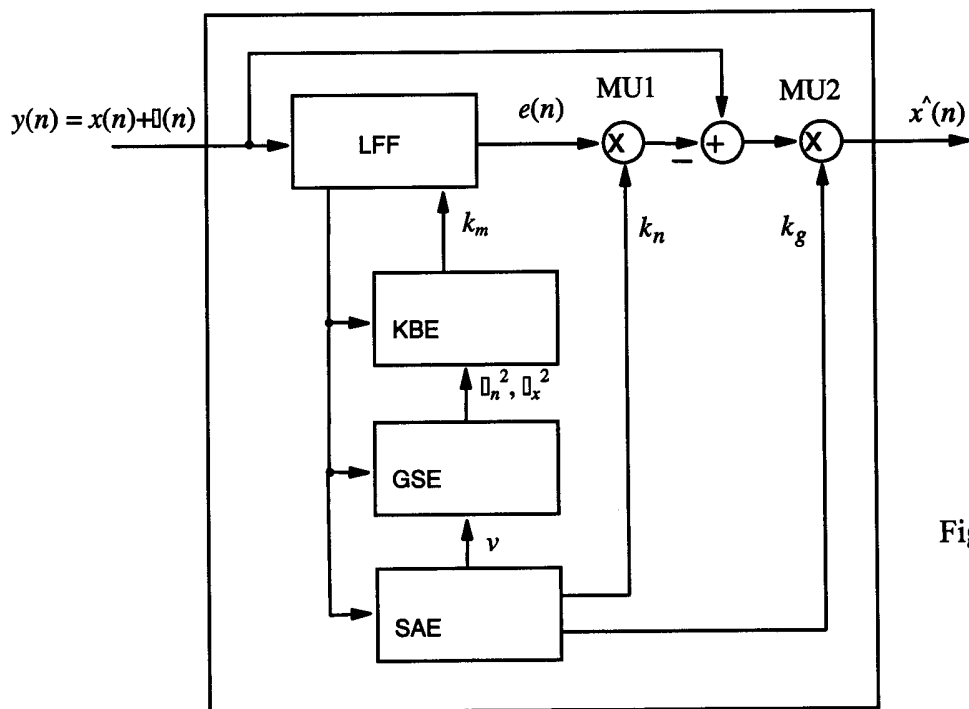


Fig. 7

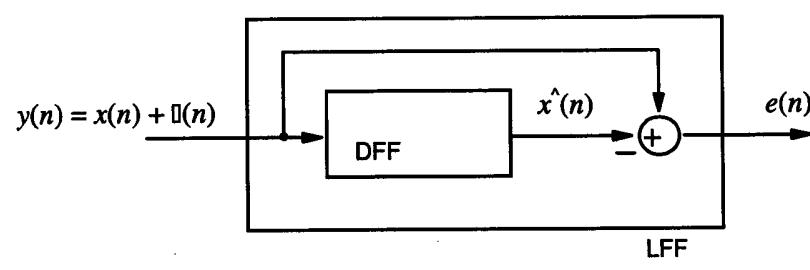


Fig. 8

Klassifikation des Anmeldegegenstands gemäß IPC ⁸ : G10L 21/02 (2006.01)
Klassifikation des Anmeldegegenstands gemäß ECLA: G10L 21/02A1
Recherchierter Prüfstoff (Klassifikation): G10L
Konsultierte Online-Datenbank: EPODOC WPI INSPEC
Dieser Recherchenbericht wurde zu den am 15. September 2006 eingereichten Ansprüchen 1-11 erstellt.

Kategorie ¹⁾	Bezeichnung der Veröffentlichung: Ländercode, Veröffentlichungsnummer, Dokumentart (Anmelder), Veröffentlichungsdatum, Textstelle oder Figur soweit erforderlich	Betreffend Anspruch
X	US 7065486 B1 (Thyssen) 30. Juni 2006 (30.06.2006) <i>Fig 2 und 3 mit Beschreibung</i> --	1-11
X	US 20010005822 A1 (Fujii et al.) 28. Juni 2001 (28.06.2001) <i>Fig 3 und beschreibung</i> ----	2, 4-11

Datum der Beendigung der Recherche:
11. Mai 2007

Fortsetzung siehe Folgeblatt

Prüfer(in):
Dipl.-Ing. SCHLECHTER

¹⁾ **Kategorien** der angeführten Dokumente:

- X** Veröffentlichung von **besonderer Bedeutung**: der Anmeldegegenstand kann allein aufgrund dieser Druckschrift nicht als neu bzw. auf erfinderischer Tätigkeit beruhend betrachtet werden.
- Y** Veröffentlichung von **Bedeutung**: der Anmeldegegenstand kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren weiteren Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese **Verbindung für einen Fachmann naheliegend** ist.

- A** Veröffentlichung, die den **allgemeinen Stand der Technik** definiert.
- P** Dokument, das **von Bedeutung** ist (Kategorien X oder Y), jedoch **nach dem Prioritätstag** der Anmeldung veröffentlicht wurde.
- E** Dokument, das **von besonderer Bedeutung** ist (Kategorie X), aus dem ein **älteres Recht** hervorgehen könnte (früheres Anmeldedatum, jedoch nachveröffentlicht, Schutz ist in Österreich möglich, würde Neuheit in Frage stellen).
- &** Veröffentlichung, die Mitglied der selben **Patentfamilie** ist.