

(19)



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11)

EP 0 612 059 B1

(12)

EUROPÄISCHE PATENTSCHRIFT

(45) Veröffentlichungstag und Bekanntmachung des Hinweises auf die Patenterteilung:
17.03.1999 Patentblatt 1999/11

(51) Int Cl.⁶: **G10L 3/00, G10L 3/02**

(21) Anmeldenummer: **93120010.9**

(22) Anmeldetag: **11.12.1993**

(54) Verfahren zur Laufzeitschätzung an gestörten Sprachkanälen

Method for estimating the propagation time in noisy speech channels

Méthode pour estimer le temps de propagation dans des canaux vocaux perturbés

(84) Benannte Vertragsstaaten:
DE FR GB

(30) Priorität: **23.12.1992 DE 4243831**

(43) Veröffentlichungstag der Anmeldung:
24.08.1994 Patentblatt 1994/34

(73) Patentinhaber: **Daimler-Benz Aktiengesellschaft
70546 Stuttgart (DE)**

(72) Erfinder: **Linhard, Klaus, Dr.-
D-89231 Neu-Ulm (DE)**

(56) Entgegenhaltungen:

**US-A- 4 112 430 US-A- 4 254 417
US-A- 4 333 170**

- **ITG-FACHTAGUNG: DIGITALE SPRACHVERARBEITUNG, 26.August 1988 - 28.August 1988 BAD NAUHEIM, Seiten 69-73, SCHLANG 'Ein Verfahren zur automatischen Ermittlung der Sprecherposition bei Freisprechen'**
- **INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING 1988, Bd. 5, 11.April 1988 - 14.April 1988 NEW YORK, NY, US, Seiten 2578-2581, ZELINSKI 'A microphone array with adaptive post-filtering for noise reduction in reverberant rooms'**

EP 0 612 059 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

[0001] Die Erfindung betrifft ein Verfahren nach dem Oberbegriff des Patentanspruchs 1.

[0002] Ein derartiges Verfahren findet Verwendung bei automatischen Spracherkennungssystemen oder für Freisprechanlagen z.B. in Büroräumen, Kraftfahrzeugen etc..

[0003] Gestörte Sprache ist besser erfaßbar, wenn sie mit zwei oder mehreren Kanälen aufgezeichnet wird. Der Mensch benutzt zwei Kanäle, seine beiden Ohren. Durch eine psychoakustische Nachverarbeitung wird bei ihm die Richtung des Sprechers ermittelt und die Hintergrundstörung ausgeblendet. Bei technischen Geräten können zwei oder mehrere Kanäle zur Aufzeichnung verwendet werden. Diese Signale können dann mit einer digitalen Signalverarbeitung aufbereitet werden.

[0004] Ein wesentlicher Aspekt der mehrkanaligen Verarbeitung ist die Schätzung des Laufzeitunterschiedes der einzelnen Kanäle. Ist der Laufzeitunterschied bekannt, kann die Richtung des Schallereignisses (Sprecher) ermittelt werden. Die Signale der einzelnen Kanäle können entsprechend laufzeitkorrigiert und weiterverarbeitet werden. Werden z.B. nicht korrigierte Signale zu einem Summensignal zusammengefaßt, können sich einzelne spektrale Anteile des Signals durch Interferenz verstärken, dämpfen oder auslöschen.

[0005] Ein Verfahren zur automatischen Ermittlung der Laufzeitunterschiede zweier Mikrofone ist aus einer Veröffentlichung von M. Schlang, ITG-Fachtagung 1988, Bad Nauheim S. 69-73 bekannt. Es arbeitet im Zeitbereich. Jedoch ist dieses Verfahren bei starken Störungen nicht anwendbar.

[0006] Der Erfindung liegt deshalb die Aufgabe zugrunde ein Verfahren zur Laufzeitschätzung für ein Spracherkennungssystem anzugeben, das auch bei starken Hintergrundgeräuschen anwendbar ist, für ein Mehrkanalübertragungssystem geeignet ist und zeit- und kostensparend arbeitet.

[0007] Die Aufgabe wird gelöst durch die im kennzeichnenden Teil des Patentanspruchs 1 angegebenen Merkmale. Vorteilhafte Ausgestaltungen und/oder Weiterbildungen sind den Unteransprüchen zu entnehmen.

[0008] Die Erfindung wird anhand eines Ausführungsbeispiels beschrieben unter Bezugnahme auf schematische Zeichnungen.

[0009] In FIG. 1 wird anhand eines Blockschaltbilds die Phasenschätzung erläutert.

[0010] FIG. 2 gibt für ein Fahrgeräusch von 140km/h eine Darstellung der Größen S_B , S_I , S_N und g in Abhängigkeit von der Zeit an.

[0011] In der vorliegenden Erfindung wird ein 2-kanaliger Laufzeitausgleich vorgestellt. Die Erweiterung auf mehrere Kanäle ist mit dem entsprechenden Mehraufwand leicht möglich. Der Laufzeitausgleich ist ein Teil der Signalvorverarbeitung einer mehrkanaligen Geräuschreduktion, die z.B. für einen Spracherkennung im Fahrzeug verwendet werden kann.

[0012] Die Laufzeit wird im Frequenzbereich ermittelt. Dies ermöglicht eine einfache Laufzeitkorrektur durch die Multiplikation des Spektrums mit der neuen Phase und führt zu einem geringen Rechenaufwand.

[0013] Die Sprach- und Geräuschaufnahmen zur Entwicklung und Bewertung des vorliegenden Verfahrens wurden in einem Fahrzeug mit zwei Mikrofonen durchgeführt. Die Störung ist das Fahrgeräusch bei verschiedenen Fahrsituationen.

[0014] Mit dem erfindungsgemäßen Verfahren werden im Frequenzbereich die Phasen an einer Anzahl von Maxima der Kreuzkorrelation bestimmt. Die Hintergrundstörung und das Einschwingverhalten des Raumes werden ständig mitgeschätzt. Die einzelnen Phasenwerte werden nur zu Beginn eines Einschwingvorgangs verarbeitet und wenn das Hintergrundgeräusch um einen gewissen Faktor überschritten wird. Bei der Weiterverarbeitung der Phasenwerte wird eine lineare Phasenbeziehung vorausgesetzt und die Varianz der Schätzung wird bei der Glättung der Werte mitberücksichtigt. Die Berücksichtigung des Einschwingvorgangs des Raumes führt dazu, daß nur bei starken Energieanstiegen der Sprache eine Phasenschätzung stattfindet. Sofort zu Beginn des Wortes steht ein neuer Phasenschätzwert zur Verfügung. Der Einfluß von Reflexionen wird vermindert. Durch die Berücksichtigung des Hintergrundgeräuschs ist das Verfahren für den praktischen Einsatz z.B. im Fahrzeug gut geeignet. Anhand eines Blockschaltbildes in FIG. 1 wird der Verfahrensablauf der Phasenschätzung näher erläutert.

[0015] Die Mikrofonensignale x und y werden in den Frequenzbereich transformiert (FFT, Fast Fourier Transformation). Die Transformationslänge wird zu $N = 256$ gewählt. Es ergaben sich die transformierten Segmente $X_i(i)$ und $Y_i(i)$. i bezeichnet den Blockindex der Segmente, i die diskrete Frequenz ($i = 0, 1, 2, \dots, N-1$). Die Segmente sind halb überlappt und werden mit einem Hamming Fenster gewichtet. (Die Abtastrate der Signale x und y beträgt 12 kHz.)

[0016] Im Frequenzbereich wird der Langzeitmittelwert des Betragsspektrums subtrahiert (SPS, spektrale Subtraktion). Die Phase der Signale wird nicht verändert. Das Störgeräusch wird reduziert. Es ergeben sich die Schätzwerte \hat{X} und \hat{Y} . Die SPS ist ein Standardverfahren und kann hier in einer einfachen Version eingesetzt werden. Sind nur geringe Störungen vorhanden, kann auf die SPS ganz verzichtet werden.

[0017] Mit der Glättungskonstante β wird das Störspektrum $S_{nn}(i)$ geschätzt. Das Störspektrum wird normiert und subtrahiert. i bezeichnet den Blockindex, i die diskrete Frequenz. Als Glättungskonstante wird z.B. $\beta_i = 0.03$ verwendet.

EP 0 612 059 B1

$$\hat{S}_{nn,l}(i) = (1 - \beta_l) \hat{S}_{nn,l-1}(i) + \beta_l |X_l(i)|^2 \quad (1)$$

5

$$|\hat{X}_l(i)| = |X_l(i)| - \frac{\hat{S}_{nn,l}(i)}{|X_l(i)|} \quad (2)$$

10

$$\hat{X}_l(i) = \left[1 - \frac{\hat{S}_{nn,l}(i)}{|X_l(i)|^2} \right] X_l(i) \quad (3)$$

Für den zweiten Kanal Y gelten die entsprechenden Gleichungen.

15 **[0018]** Aus den geschätzten Werten \hat{X} und \hat{Y} wird der Betrag der Kreuzleistungsdichte $B_{XY,l}$ berechnet. Der Bereich (N_u, N_o) liegt z.B. zwischen 300 und 1500Hz ($N_u = 6, N_o = 31$, bei $N = 256$). Dabei gilt

$$S_{xy,l}(i) = (1 - \alpha) S_{xy,l-1}(i) + \alpha \hat{X}_l(i) \hat{Y}_l^*(i); N_u \leq i \leq N_o \quad (4)$$

20

$$B_{xy,l}(i) = |S_{xy,l}(i)| \quad (5)$$

Als Glättungskonstante α wird z.B. $\alpha = 1$ gewählt. Werte $\alpha \ll 1$ sind nicht sinnvoll.

25 **[0019]** Mit einer Präemphase können höhere Frequenzen angehoben werden. Dies ist dann vorteilhaft, wenn das Sprachsignal und das Störsignal bei höheren Frequenzen eine geringere Leistung aufweisen. Die Werte der Kreuzleistung $B_{xy}(i)$ können z.B. im Bereich 300 bis 1500 Hz um 10dB linear ansteigend angehoben werden. Die Präemphase kann aber auch schon durch die Mikrofoncharakteristik vorgegeben sein.

[0020] Aus den Werten $B_{xy}(i)$ werden M Maxima bestimmt und summiert. Es können z.B. $M = 8$ verwendet werden. Es wird ein aktueller Schätzwert

30

$$S_{B,l} = \frac{1}{M} \sum_{i=1}^M B_{xy,l}(i) \quad (6)$$

35 bestimmt.

[0021] Über einen Impulsmonitor wird eine "simulierte Impulsantwort" S_l berechnet. Das Einschwingverhalten des umgebenden Raumes auf plötzliche energiestarke Schallereignisse (Sprache) wird hiermit grob simuliert (z.B. wird $\gamma = 0.1$ gewählt). Die Glättung des Phasenwerts "vom Wortanfang in das Wort hinein" ist mit γ einstellbar.

40

$$S_{l,i} = (1 - \gamma) S_{l,i-1} + \gamma S_{B,i} \quad (7)$$

45 Außerdem wird über einen Geräuschmonitor eine adaptive Glättungskonstante h berechnet. Mit dieser Glättungskonstanten ergibt sich ein Schätzwert S_N für die Störung. Wurde zuvor eine spektrale Subtraktion (SPS) durchgeführt, ist S_N ein Schätzwert für die Reststörung. Für die Glättungskonstante h_o gilt z.B. $h_o = 0.03$

50

$$h_l = h_o \frac{2S_{N,l-1}}{S_{N,l-1} + S_{B,l}} \quad (8)$$

55

$$S_{N,l} = (1 - h_l) S_{N,l-1} + h_l S_{B,l} \quad (9)$$

Die Phase der gestörten Signale wird aus den Real- und Imaginärteilen von S_{xy} berechnet. Die Phase wird nur an den M zuvor bestimmten Maxima berechnet.

EP 0 612 059 B1

$$\varphi_1(i) = \arctan \frac{\operatorname{Im}[S_{xy,i}(i)]}{\operatorname{Re}[S_{xy,i}(i)]}; \quad \text{für } \operatorname{Re} > 0 \quad (10)$$

5 und

$$\varphi_1(i) = \pi - \arctan \frac{-\operatorname{Im}[S_{xy,i}(i)]}{\operatorname{Re}[S_{xy,i}(i)]}; \quad \text{sonst} \quad (11)$$

10

Daraus ergibt sich der Phasenanstieg:

$$\varphi'_1(i) = \frac{\varphi_1(i)}{i} \quad (12)$$

15

Mit der Länge der Fouriertransformation N und der max. zulässigen Verschiebung um n Taps ergibt sich (N = 256):

$$|\varphi'|_{\max} = \ln \left| \frac{2\pi}{N} \right| \quad (13)$$

20

Übersteigt der Phasenanstieg $|\varphi'|$ an einem der Maxima $|\varphi'|_{\max}$, so wird dieser Wert φ' nicht weiterverwendet. Es wird eine adaptive Glättungskonstante g berechnet:

25

$$g_i = \frac{g_o (s_{B,i} - S_{i,i})}{S_{i,i}} \quad (14)$$

30

$$g_i \leq g_{\max} \quad (15)$$

$$g_o = 0,25 \quad (16)$$

35

Der aktuelle Wert S_B muß um den Faktor c größer sein als die simulierte Impulsantwort S_i

$$S_{B,i} \geq c S_{i,i}; \quad c = 2 \quad (17)$$

40

sonst gilt:

$$g_i = 0 \quad (18)$$

45

Der aktuelle Wert S_B muß um den Faktor d größer sein als das Restrauschen S_N

$$S_{B,i} \geq d S_{N,i}; \quad d = 3 \quad (19)$$

50

sonst gilt ebenfalls

$$g_i = 0 \quad (20)$$

55

Ist Gl. (17) oder Gl. (19) nicht erfüllt, d.h. gilt $g = 0$, so kann die Phasenschätzung abgebrochen werden. Es gilt der alte Phasenschätzwert.

[0022] Für alle

$$|\varphi'_l(i)| \leq |\varphi|_{\max} \quad (21)$$

gilt:

$$m_{\varphi', l} = \frac{1}{M'} \sum_{i=1}^{M'} \varphi'_{l}(i) \quad (22)$$

$$s^2_{\varphi', l} = \frac{1}{M'} \sum_{i=1}^{M'} (\varphi'_{l}(i))^2 \quad (23)$$

Von den ursprünglichen M Maxima werden wegen Gl. (21) nur M' für die Gl. (22, 23) verwendet. Ist die Anzahl M' der für die Summen gültigen Werte ϕ kleiner als M_{\min} , gilt der geschätzte Phasenanstieg als zu unsicher oder außerhalb des Nutzbereichs (z.B. $M_{\min} = 6$, bei $M = 8$). Die Phasenschätzung wird dann nicht aktualisiert und das Verfahren hier abgebrochen. Es gilt der alte Phasenschätzwert.

[0023] Es wird die Varianz der Schätzung berechnet:

$$\sigma^2_{\varphi', l} = s^2_{\varphi', l} - m^2_{\varphi', l} \quad (24)$$

Als maximale Varianz wird

$$\sigma^2_{\max} = |\varphi'|_{\max}^2 \quad (25)$$

verwendet.

[0024] Entsprechend der Varianz wird die Glättungskonstante g gewichtet. Bei einer großen Streuung gilt:

$$g_j := 0.09 * g_j; \quad \text{für } 0,2\sigma^2_{\max} < \sigma^2_{\varphi', l} < \sigma^2_{\max} \quad (26)$$

Bei einer mittleren Streuung gilt:

$$g_j := 0.3 * g_j; \quad \text{für } 0.02\sigma^2_{\max} \leq \sigma^2_{\varphi', l} \leq 0,2\sigma^2_{\max} \quad (27)$$

Bei sehr geringer Streuung gilt:

$$g_j := g_j; \quad \text{für } \sigma^2_{\varphi', l} < 0.02\sigma^2_{\max} \quad (28)$$

Entsprechend den Gl. 17 - 20 wird g in der Regel nur am Wortanfang größer Null sein. Dabei muß die Energie des Wortes größer sein als die Energie des Restgeräusches und der simulierten Impulsantwort. Mit der Variablen j wird die aufeinanderfolgende Anzahl für $g > 0$ gezählt. Entsprechend gilt für die Glättung:

$$j=1: \quad \varphi'_l = m_{\varphi', l} \quad (29)$$

$$j=2: \quad m_{\varphi', l} = \frac{(m_{\varphi', l} + m_{\varphi', l-1})}{2} \quad (30)$$

$$\tilde{\varphi}'_1 = (1 - 1,5g_1) \tilde{\varphi}'_{1-1} + 1,5 g_1 m_{\varphi'_1} \quad (31)$$

5

j=3, 4 . . . :

$$\tilde{\varphi}'_j = (1 - g_1) \tilde{\varphi}'_{j-1} + g_1 m_{\varphi'_j} \quad (32)$$

10

Wird z.B. infolge einer Störung die Bedingung $g > 0$ nur einmal in Folge erfüllt, wird die Phasenschätzung nicht aktualisiert. Eine Aktualisierung der Phasenschätzung erfolgt nur dann, wenn $g > 0$ mindestens 2-mal in Folge erfüllt wird.

[0025] Ein Beispiel für die Zwischengrößen S_B , S_I , S_N' und g und die daraus abgeleitete Phasenschätzung zeigt, FIG. 2. Dabei wird das Wort "Senderwahl" gesprochen und das Fahrgeräusch bei 140km/h addiert. Das Verfahren wird, wie oben angegeben, verwendet. Der Phasenschätzwert ist in Abtastwerten n angegeben. Mit der Größe S_I wird der "Sprachimpuls" teilweise verdeckt und so nur bei starken Energieanstiegen eine Schätzung erlaubt (S_B muß S_I um den Faktor 2 übersteigen). Die Schätzung der Reststörung S_N ermöglicht eine größere Robustheit gegenüber Geräuschen (S_B muß S_N um den Faktor 3 übersteigen).

20

Patentansprüche

1. Verfahren zur Laufzeitschätzung bei dem Laufzeitunterschiede von geräuschgestörten Signalen von zumindest zwei Sprachkanälen mittels einer Kreuzkorrelation bestimmt werden, dadurch gekennzeichnet,

25

- daß im Frequenzbereich die Phasenwerte von zumindest zwei Signalen über eine bestimmte Anzahl von Maxima der Kreuzleistungsdichte ermittelt werden und deren Phasenverschiebung bestimmt wird, und
- daß der erforderliche Phasenausgleich ebenfalls im Frequenzbereich durchgeführt wird.

30

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß Hintergrundstörungen und das Einschwingverhalten des Raumes bei der Bestimmung der Phasenwerte ständig mitgeschätzt werden.

3. Verfahren nach Anspruch 2, dadurch gekennzeichnet, daß das Hintergrundgeräusch über einen Geräuschmonitor geschätzt wird, und daß ein neuer Phasenwert lediglich dann ermittelt wird, wenn der Schätzwert des Hintergrundgeräusches um einen bestimmten Faktor überschritten wird.

35

4. Verfahren nach Anspruch 2, dadurch gekennzeichnet, daß das Einschwingverhalten des umgebenden Raumes über einen Impulsmonitor derart geschätzt wird, daß lediglich bei starkem Energieanstieg in den Signalen ein neuer Phasenschätzwert ermittelt wird.

40

5. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß eine lineare Laufzeitverschiebung der Signale angenommen wird.

6. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß eine Glättung des Phasenwertes vom Wortanfang in das gesprochene Wort hinein durchgeführt wird, und daß die Varianz der Schätzung bei der Glättung der Phasenwerte mitberücksichtigt wird.

45

7. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet,

50

- daß zumindest zwei Mikrofonsignale x , y mittels einer FFT (Fast Fourier Transformation) in den Frequenzbereich transformiert werden,
- daß durch spektrale Subtraktion aus den transformierten Signalen die Schätzwerte \hat{X} , \hat{Y} bestimmt werden,
- daß aus den geschätzten Werten \hat{X} , \hat{Y} der Betrag der Kreuzleistungsdichte B_{xy} bestimmt wird,
- daß die Maxima der Kreuzleistungsdichte bestimmt werden, und daß aus einer bestimmten Anzahl Maxima

55

der Kreuzleistungsdichte B_{xy} ein aktueller Wert S_B für die gestörten Signale ermittelt wird, daß abhängig vom aktuellen Wert S_B die Phasen φ der gestörten Signale ermittelt werden und damit der Phasenanstieg φ' bestimmt wird,

- 5 - daß der Phasenanstieg φ' geglättet wird, indem über einen Impulsmonitor ein simulierter Sprachimpuls S_I mit dem aktuellen Wert S_B der gestörten Signale gekoppelt wird, derart, daß eine erneute Phasenschätzung lediglich dann durchgeführt wird, wenn ein starker Energieanstieg des Mikrofonsignals registriert wird, und
- 10 - daß mit einem Geräuschmonitor ein Schätzwert S_N für die Hintergrundgeräuschstörung ermittelt wird und mit dem aktuellen Wert S_B der gestörten Signale gekoppelt wird, derart, daß eine erneute Phasenschätzung lediglich dann durchgeführt wird, wenn vom Signal die Hintergrundstörung deutlich überschritten wird.
- 8. Verfahren nach Anspruch 7, dadurch gekennzeichnet, daß ein maximaler Phasenanstieg $|\varphi'|_{\max}$ für die Phase an den einzelnen Maxima vorgebar ist und eine erneute Phasenschätzung lediglich dann durchgeführt wird, wenn der Phasenanstieg um mindestens M' der M Maxima den maximalen Anstieg $|\varphi'|_{\max}$ nicht überschreitet.
- 15 9. Verfahren nach Anspruch 7, dadurch gekennzeichnet, daß die Varianz der Phasenanstiege an den einzelnen Maxima bei der zeitlichen Glättung des Phasenanstiegs berücksichtigt wird.
- 20 10. Verfahren nach den Ansprüchen 7 bis 9, dadurch gekennzeichnet, daß eine erneute Phasenschätzung lediglich dann durchgeführt wird, wenn die Bedingungen für einen gültigen Phasenanstieg zeitlich mehrfach in Folge auftreten.
- 25 11. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß die gestörte Sprache auf mehr als zwei Sprachkanälen aufgenommen wird und daß die Laufzeitunterschiede der einzelnen Kanäle geschätzt werden.

Claims

- 30 1. Method of estimating transit time, in which transit time differences of noise-disturbed signals of at least two speech channels are determined by means of a cross-correlation, characterised thereby
 - 35 - that in the frequency range the phase values of at least two signals are ascertained by way of a specific number of maxima of the cross power density and the phase displacement thereof is determined, and
 - that the required phase equalisation is similarly undertaken in the frequency range.
- 40 2. Method according to claim 1, characterised thereby that background disturbances and the transient behaviour of the space are constantly estimated together in the determination of the phase values.
- 3. Method according to claim 2, characterised thereby that the background noise is estimated by way of a noise monitor and that a new phase value is ascertained merely when the estimated value of the background noise is exceeded by a specific factor.
- 45 4. Method according to claim 2, characterised thereby that the transient behaviour of the surrounding space is estimated by way of an impulse monitor in such a manner that a new phase estimated value is ascertained merely in the case of strong increase in energy in the signals.
- 50 5. Method according to one of the preceding claims, characterised thereby that a linear transit time displacement of the signals is adopted.
- 6. Method according to one of the preceding claims, characterised thereby that a smoothing of the phase value is undertaken from the word start in the spoken word and that the variance of the estimation is taken into consideration in the smoothing of the phase values.
- 55 7. Method according to one of the preceding claims, characterised thereby that
 - that at least two microphone signals x,y are transformed into the frequency range by means of an FFT (Fast

Fourier Transform),

- that the estimated values \hat{X}, \hat{Y} are determined by spectral subtraction from the transformed signals,
- that the amount of the cross power density B_{xy} is determined from the estimated values \hat{X}, \hat{Y} ,
- that the maxima of the cross power density are determined, that an actual value S_B for the disturbed signals is ascertained from a specific number of maxima of the cross power density B_{xy} , and that in dependence on the actual value S_B the phases φ of the disturbed signals are ascertained and thus the phase increase φ' is determined,
- that the phase increase φ' is smoothed, in that by way of an impulse monitor a simulated speech pulse S_1 is coupled with the actual value S_B of the disturbed signals in such a manner that a renewed phase estimation is undertaken merely when a strong increase in energy of the microphone signal is registered, and
- that an estimated value S_N for the background noise disturbance is ascertained by a noise monitor and is coupled with the actual value S_B of the disturbed signals in such a manner that a renewed phase estimation is undertaken merely when the background disturbance is significantly exceeded by the signal.

8. Method according to claim 7, characterised thereby that a maximum phase increase $k\varphi'_{\max}$ for the phase of the individual maxima is settable and a renewed phase estimation is undertaken merely when the phase increase does not exceed the maximum increase $k\varphi'_{\max}$ by a least M' of the M maxima.

9. Method according to claim 7, characterised thereby that the variance of the phase increases of the individual maxima are taken into consideration in the smoothing over time of the phase increase.

10. Method according to claims 7 to 9, characterised thereby that a renewed phase estimation is undertaken merely when the conditions for a valid phase increase occur in time several times in succession.

11. Method according to one of the preceding claims, characterised thereby that the disturbed speech is received on more than two speech channels and that the transit time differences of the individual channels are estimated.

Revendications

1. Procédé pour l'évaluation du temps de propagation dans le cas duquel des différences de temps de propagation de signaux, perturbés par des bruits, provenant d'au moins deux canaux vocaux sont déterminées au moyen d'une intermodulation, caractérisé par le fait que:

- les valeurs de phase d'au moins deux signaux sont déterminées dans la dimension fréquence au moyen d'un nombre déterminé de maximums de la densité de puissance d'intermodulation, et leur décalage de phase est déterminé, et
- que l'équilibrage de phase nécessaire est également effectué dans la dimension fréquence.

2. Procédé selon la revendication 1, caractérisé par le fait que, lors de la détermination des valeurs de phase, on évalue simultanément et en permanence des perturbations provoquées par l'arrière-plan et la réaction de l'espace en régime transitoire.

3. Procédé selon la revendication 2, caractérisé par le fait que le bruit provoqué par l'arrière-plan est évalué au moyen d'un moniteur de bruit et qu'une valeur de phase n'est déterminée que si la valeur évaluée du bruit provoqué par l'arrière-plan est dépassée d'un facteur déterminé.

4. Procédé selon la revendication 2, caractérisé par le fait que la réaction de l'espace environnant en régime transitoire est évaluée, par un moniteur d'impulsion, de façon qu'une nouvelle valeur évaluée de phase ne soit déterminée qu'en présence d'une forte croissance d'énergie dans les signaux.

5. Procédé selon l'une des revendications précédentes, caractérisé par le fait que l'on admet un décalage linéaire du temps de propagation des signaux.

6. Procédé selon l'une des revendications précédentes, caractérisé par le fait qu'il est procédé à un lissage de la valeur de phase, du début du mot, à l'intérieur du mot prononcé, et que, lors du lissage des valeurs de phase, on tient compte de la variance de l'évaluation.

7. Procédé selon l'une des revendications précédentes, caractérisé par le fait:

- qu'au moins deux signaux de microphone x, y sont transformés au moyen d'une transformation rapide de Fourier FFT, dans la dimension fréquence,
- que les valeurs évaluées X, Y sont déterminées, par soustraction spectrale, à partir des signaux transformés,
- qu'à partir des valeurs évaluées X, Y, on détermine la valeur de la densité de puissance d'intermodulation B_{xy} ,
- que l'on détermine les maximums de la densité de puissance d'intermodulation et qu'à partir d'un nombre déterminé de maximums de la densité de puissance d'intermodulation B_{xy} , on détermine une valeur actuelle S_B pour les signaux perturbés, que les phases φ des signaux perturbés sont déterminées, et donc également la croissance de phase φ' , en fonction de la valeur actuelle S_B .
- que la croissance de phase φ' est lissée, en ce sens qu'au moyen d'un moniteur par impulsions, une réponse simulée en régime impulsif S_I est couplée à la valeur actuelle S_B des signaux perturbés, de façon qu'il ne soit procédé à une nouvelle évaluation de phase que si l'on enregistre une forte croissance d'énergie des signaux des microphones et
- qu'avec un moniteur de bruit, une valeur évaluée S_N pour la perturbation provoquée par le bruit d'arrière-plan est déterminée et couplée avec la valeur actuelle S_B des signaux perturbés de façon qu'il ne soit procédé à une nouvelle évaluation de phase que si la perturbation provoquée par l'arrière-plan est nettement dépassée par le signal.

8. Procédé selon la revendication 7, caractérisé par le fait qu'une croissance de phase maximale $|\varphi'|_{\max}$ peut être prescrite pour la phase aux différents maximums et qu'il n'est procédé à une nouvelle évaluation de phase que si la croissance de phase ne dépasse pas la croissance maximale $|\varphi'|_{\max}$ pour au moins M' des M maximums.

9. Procédé selon la revendication 7, caractérisé par le fait que lors du lissage de la croissance de phase dans le temps, on tient compte de la variance de la croissance de phase aux différents maximums.

10. Procédé selon les revendications 7 à 9, caractérisé par le fait que l'on ne procède à une nouvelle évaluation de phase que si les conditions pour une croissance de phase valable se présentent plusieurs fois successivement dans le temps.

11. Procédé selon l'une des revendications précédentes, caractérisé par le fait que la parole perturbée est recueillie sur plus de deux canaux vocaux, et que l'on évalue les différences de temps de propagation des différents canaux.

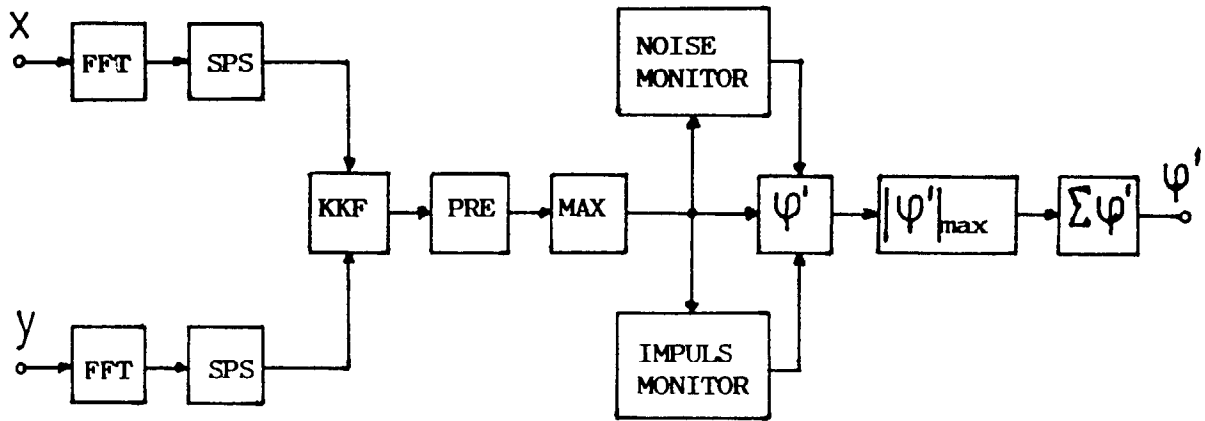


FIG.1

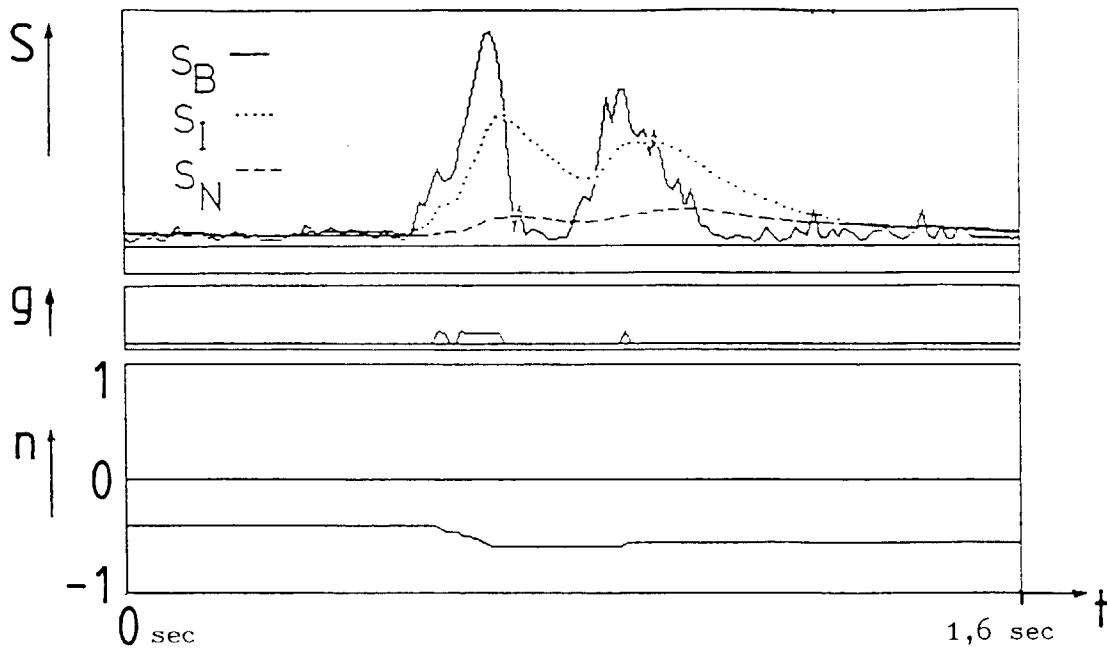


FIG. 2