



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11) **EP 1 382 034 B1**

(12) **EUROPÄISCHE PATENTSCHRIFT**

(45) Veröffentlichungstag und Bekanntmachung des Hinweises auf die Patenterteilung:
16.02.2005 Patentblatt 2005/07

(21) Anmeldenummer: **02727282.2**

(22) Anmeldetag: **03.04.2002**

(51) Int Cl.7: **G10L 19/00, G10L 11/02**

(86) Internationale Anmeldenummer:
PCT/DE2002/001200

(87) Internationale Veröffentlichungsnummer:
WO 2002/084644 (24.10.2002 Gazette 2002/43)

(54) **VERFAHREN ZUR BESTIMMUNG VON INTENSITÄTSKENNWERTEN VON HINTERGRUNDGERÄUSCHEN IN SPRACHPAUSEN VON SPRACHSIGNALEN**

METHOD FOR DETERMINING INTENSITY PARAMETERS OF BACKGROUND NOISE IN SPEECH PAUSES OF VOICE SIGNALS

PROCEDE DE DETERMINATION DE VALEURS CARACTERISTIQUES D'INTENSITE DE BRUITS DE FOND DANS DES PAUSES DE VOIX DE SIGNAUX VOCAUX

(84) Benannte Vertragsstaaten:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE TR

(30) Priorität: **18.04.2001 DE 10120168**

(43) Veröffentlichungstag der Anmeldung:
21.01.2004 Patentblatt 2004/04

(73) Patentinhaber: **Deutsche Telekom AG**
53113 Bonn (DE)

(72) Erfinder: **BERGER, Jens**
10405 Berlin (DE)

(56) Entgegenhaltungen:
WO-A-00/52683 **US-A- 4 811 404**
US-A- 5 598 466 **US-A- 6 044 342**

EP 1 382 034 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 bezieht sich auf ein Verfahren zur Bewertung von Hintergrundgeräuschen in Sprachpausen von aufgezeichneten oder übertragenen Sprachsignalen wie in Anspruch 1 definiert.

[0002] Die empfundene Sprachqualität, z.B. in Telefonverbindungen oder Rundfunkübertragungen, wird hauptsächlich von sprachsimultanen Störungen, also von Störungen während der Sprachaktivität, bestimmt. Aber auch Geräusche in den Sprachpausen gehen in das Qualitätsurteil ein, insbesondere bei hochqualitativer Sprachwiedergabe.

[0003] Die Intensität des Hintergrundgeräusches in den Sprachpausen kann als ergänzender Kennwert zur Bestimmung der Sprachqualität (Sprachgüte) verwendet werden.

[0004] Sprachqualitätsbestimmungen von Sprachsignalen werden in der Regel mittels auditiver ("subjektiver") Untersuchungen mit Versuchspersonen vorgenommen.

[0005] Das Ziel von instrumentellen ("objektiven") Verfahren zur Sprachqualitätsbestimmung ist es dagegen, aus Eigenschaften des zu bewertenden Sprachsignals mittels geeigneter Rechenverfahren Kennwerte zu ermitteln, die die Sprachqualität des Sprachsignals beschreiben, ohne auf Urteile von Versuchspersonen zurückgreifen zu müssen.

[0006] Eine sichere Qualitätsbestimmung liefern instrumentelle Verfahren, die auf einem Vergleich von ungestörtem Referenzsprachsignal (Quellsprachsignal) und dem gestörten Sprachsignal am Ende der Übertragungskette beruhen. Es existieren viele solcher Verfahren, die meist in sogenannten Probeverbindungs-systemen eingesetzt werden. Dabei wird an der Quelle das ungestörte Quellsprachsignal eingespeist und nach der Übertragung wieder aufgezeichnet.

Stand der Technik und Nachteile bekannter Verfahren

[0007] Bekannte Verfahren zur Bestimmung der Intensität von Hintergrundgeräuschen gehen meist vom gestörten Signal selbst aus und nutzen eine festgelegte Intensitätsschwelle zur Unterscheidung von aktiver Sprache und Sprachpausen (Fig. 1). Diese Schwelle ist im einfachsten Fall konstant im Verfahren eingestellt, kann aber auch anhand des Signalverlaufs adaptiert werden (z.B. festgelegter Abstand zum Signal-Spitzenwert). Das Ziel ist eine sichere Unterscheidung zwischen Sprache und Sprachpause. Gelingt die Unterscheidung, können die gesuchten Intensitätskennwerte des Hintergrundgeräuschs aus den als Sprachpause detektierten Signalabschnitten bestimmt werden. Dazu werden im Allgemeinen die als Sprachpause detektierten Signalabschnitte nochmals in kürzere Segmente (typisch sind 8...40ms) unterteilt und für diese die Intensitätsberechnungen (z.B. Effektivwert oder Lautheit) vorgenommen. Aus den Ergebnissen können dann Intensitätskennwerte bestimmt werden.

[0008] Die Verfahren liefern bei geringen Geräuschintensitäten in Sprachpausen und gleichzeitig hoher Intensität der Sprache (großes Sprach-Geräusch-Verhältnis) sichere Meßwerte, da die Unterscheidung zwischen Sprache und Sprachpause sicher erfolgen kann (Fig.1).

[0009] Bei steigenden Geräuschintensitäten in Sprachpausen (abnehmendes Sprach-Geräusch-Verhältnis) treten zunehmend Unsicherheiten in der Unterscheidung zwischen Sprache und Sprachpausen auf. Hier ist es schwierig den Schwellenwert so festzulegen, dass zum einen keine Geräuschabschnitte mit höheren Intensitäten als Sprache detektiert werden (Schwelle zu niedrig) und zum anderen keine Sprachabschnitte geringerer Intensität als Sprachpause gewertet werden (Schwelle zu hoch) (Fig. 2).

[0010] Erreicht die Intensität des Geräusches in den Sprachpausen die Intensität der aktiven Sprache oder übersteigt diese sogar, ist keine Intensitätsschwelle zu finden, die eine Unterscheidung zwischen Sprache und Sprachpause ermöglicht.

[0011] Lösungen für die beschriebenen Probleme sind möglich, wenn z.B. unterschiedliche spektrale Charakteristika von Sprache und Hintergrundgeräuschen vorliegen. Hier kann durch geeignete, Vorfilterung des Signals bzw. durch eine spektrale Analyse und Auswertung von ausgewählten Frequenzbändern ein höheres Verhältnis von Sprache zu Hintergrundgeräusch in den betrachteten Frequenzbereichen erreicht werden, so dass wieder eine sichere Unterscheidung zwischen aktiver Sprache und Sprachpause möglich ist.

[0012] Andere Lösungen bedienen sich bestimmter Parameter, die bei Sprachcodierung ermittelt werden und nutzen diese zur Unterscheidung zwischen Sprache und Abschnitten mit Hintergrundgeräuschen. Dabei ist es das Ziel, aus den Parametern abzuleiten, ob das betrachtete Signalsegment typische Eigenschaften von Sprache (z.B. stimmhafte Anteile) aufweist. Ein Beispiel hierfür ist "Voice-Activity Detector" (ETSI Recommendation GSM 06.92, Valboune, 1989).

[0013] Diese Verfahren arbeiten bei geringen Sprach-Geräusch-Verhältnissen robuster und werden vorrangig zur Unterdrückung der Übertragung von Sprachpausen z.B. im Mobilfunk eingesetzt. Die Verfahren zeigen jedoch Unsicherheiten, wenn das Hintergrundgeräusch selbst Sprache beinhaltet oder sprachähnlich ist. Solche Abschnitte werden dann als Sprache klassifiziert, obwohl sie von einem Zuhörer als störendes Hintergrundgeräusch empfunden werden.

[0014] Instrumentelle Sprachqualitätsmessverfahren basieren meist auf dem Prinzip des Signalvergleichs von ungestörtem Referenzsprachsignal und gestörtem und zu bewertenden Signal. Beispiele hierfür sind die Veröffentlichungen:

"A perceptual speech-quality measure based on a psychacoustic sound representation" (Beerends, J. G.; Stemerdink, J. A., J. Audio Eng. Soc. 42(1994)3, S. 115-123)

"Auditory distortion measure for speech coding" (Wang, S; Sekey, A.; Gersho, A.: IEEE Proc. Int. Conf. acoust., speech and signalprocessing (1991), S.493-496).

[0015] Der derzeit gültige ITU-T Standard P.861 beschreibt ebenfalls ein derartiges Verfahren: "Objective quality measurement of telephone-band speech codecs" (ITU-T Rec. P.861, Genf 1996).

[0016] Solche Messverfahren werden in sogenannten Probeverbindungs-systemen eingesetzt, bei denen ein bekanntes Referenzsprachsignal (Quellsprachsignal) an der Quelle eingespeist, über z. B. eine Telefonverbindung übertragen und an der Senke aufgezeichnet wird. Nach der Aufzeichnung des Sprachsignals werden zur Bewertung der Sprachqualität des möglicherweise gestörten Signals dessen Eigenschaften mit denen des ungestörten Quellsprachsignals verglichen.

[0017] Steht für die Bestimmung des Hintergrundgeräuschs in Sprachpausen das ungestörte Quellsprachsignal zur Verfügung, dann kann dieses zur Festlegung der Übergangszeitpunkte von Sprache zur Sprachpause bzw. von Sprachpause zur Sprache benutzt werden. Dazu wird z.B. ein Verfahren mit Schwellwertbestimmung - wie oben beschrieben - auf das Quellsprachsignal angewandt. Das Verfahren liefert sichere Unterscheidungen zwischen Sprache und Sprachpause, da das Sprach-Geräusch-Verhältnis im ungestörten Quellsprachsignal ausreichend hoch ist (Fig. 3a). Die Zeitpunkte der Schwellpassage, d.h. Beginn bzw. Ende der Sprachaktivität, können nun auf das gestörte Sprachsignal übertragen werden (Fig. 3b).

[0018] Unproblematisch kann ein solches Verfahren modifiziert werden, wenn zwischen Quellsprachsignal und gestörtem Signal eine konstante Zeitdifferenz (z.B. Verzögerung durch Signalübertragung) eintritt. Bedingung ist aber, dass diese Zeitdifferenz vorab sicher bestimmt werden kann und dann zur Korrektur der Zeitpunkte Ende bzw. Beginn der Sprachaktivität genutzt wird. Das ist meist bei zeit-invarianten Systemen möglich, da diese eine konstante Verzögerung besitzen (Fig. 3c).

Prinzipiell funktioniert ein solches Verfahren auch, wenn der Zeitversatz zwischen beiden Signalen nicht für die gesamte Signallänge konstant ist, sondern variabel verläuft. Zu diesen zeit-invarianten Systemen zählen insbesondere paketbasierte Übertragungssysteme, bei denen durch unterschiedliche Paketlaufzeiten und entsprechendes Management im Empfänger deutliche Schwankungen in der Systemverzögerung auftreten können. Um Verlusten durch verspätet eintreffende Pakete vorzubeugen, werden teilweise Sprachpausen im Empfänger verlängert und spätere wieder verkürzt. Eine Übertragung der Zeitpunkte von Beginn bzw. Ende der Sprachaktivität ist nur noch bei Kenntnis der aktuellen Verzögerung an diesen Punkten möglich. Die adaptive Bestimmung des Zeitversatzes ist rechenzeitintensiv und gelingt insbesondere bei verringerten Sprach-Geräusch-Verhältnissen oft nur unzureichend. Wenn die adaptive Bestimmung des Zeitversatzes nicht sicher gelingt, können Anfang und Ende von Sprachpausen nicht exakt oder gar nicht ermittelt werden. Dadurch ist keine oder nur eine unsichere Bestimmung der Intensitätskennwerte von Pausengeräuschen möglich.

[0019] Beispiele von der Bestimmung von Hintergrundgeräuschen gemäß dem Stand der Technik sind von US6044342A, US5598466A, WO0052683A und US4811404A bekannt.

Aufgabe

[0020] Wie beschrieben, ist die Bestimmung von Hintergrundgeräuschen in Sprachpausen auch bei Kenntnis des ungestörten Quellsprachsignals schwierig oder teilweise unmöglich, insbesondere wenn

- ein geringes Verhältnis von Sprache zu Hintergrundgeräusch vorliegt,
- das Hintergrundgeräusch Sprache beinhaltet oder selbst sprachähnlich ist,
- der Zeitversatz zwischen ungestörtem Quellsprachsignal und gestörtem Sprachsignal nicht konstant über die gesamte Signallänge ist.

[0021] Es soll ein Verfahren vorgestellt werden, mit dem auch unter den genannten Bedingungen eine sichere und schnelle Bestimmung von Intensitätskennwerten des Hintergrundgeräuschs in Sprachpausen gewährleistet wird. Bedingung ist, dass sowohl Quellsprachsignal als auch gestörtes Sprachsignal vollständig aufgezeichnet zur Verfügung stehen.

Lösungsprinzip

[0022] Die bekannten Verfahren gehen davon aus, den Zeitpunkt von Beginn und Ende einer Sprachpause möglichst exakt zu ermitteln. Im Ergebnis steht dann das Signal von den Pausenabschnitten zur weiteren Auswertung zur Ver-

fügung. Aus diesen separierten Pausenabschnitten des Signals werden die Intensitätskennwerte ermittelt.

[0023] Mit dem vorliegenden Verfahren können Intensitätskennwerte von Hintergrundgeräuschen in Sprachpausen von Sprachsignalen bestimmt werden, ohne dass die exakten Zeitpunkte von Beginn und Ende eines Pausenabschnitts ermitteln werden müssen. Auch ist eine Separierung des Sprachpausensignals für die Auswertung nicht erforderlich.

[0024] Basis für das hier beschriebene Verfahren zur Bestimmung von Intensitätskennwerten von Hintergrundgeräuschen in Sprachpausen von Sprachsignalen ist die kumulative Häufigkeitsverteilung der Intensitätswerte von den Signalsegmenten, in die das Sprachsignal zuvor unterteilt wird. Diese Kurzzeit-Signalintensitäten beziehen sich auf Signalsegmente mit einer Dauer von z.B. 8ms oder 16ms. Die Häufigkeitsverteilung gibt an, wie hoch der Anteil an Kurzzeit-Intensitäten unterhalb eines definierten Schwellwertes ist.

[0025] Für die Berechnung der Häufigkeitsverteilung wird das zu analysierende Sprachsignal in kurze aufeinanderfolgende Signalsegmente unterteilt und von jedem Signalsegment der Intensitätswert (z.B. Lautheit oder Effektivwert) bestimmt.

[0026] Fig. 4 zeigt einen typischen Kurvenverlauf für Sprachsignale mit stationärem Hintergrundgeräusch (Sprach-Geräusch-Abstand ca. 10dB). Die kumulative Häufigkeitsverteilung ist am Beispiel von Kurzzeit-Lautheiten (Lautheiten berechnet nach ISO532) dargestellt. Ausgewertet wurden 2000 Segmente von 16ms Länge. Es ist zu erkennen, dass keines der Segmente einen geringeren Wert als 30 sone aufweist ($P = 0 \%$) und auch kein Segment eine höhere Lautheit als 80 sone erreicht, da hier schon der Wert $P=100 \%$ erreicht wird. Der steile Anstieg der Funktion bei ca. 30 sone lässt auf eine geringe Fluktuation der Signalintensität in großen Bereichen (fast 70%) des Signals schließen. Als Signal wurde hier ein Sprachsignal mit additiven weißen Rauschen benutzt.

[0027] Eine solche Verteilungsfunktion soll nun dazu benutzt werden, Intensitätskennwerte von Hintergrundgeräuschen in den Sprachpausen zu ermitteln. Dazu ist es erforderlich, den Anteil an Sprachpausen im Gesamtsignal zu kennen. Dieser Anteil kann aus dem ungestörten Quellsprachsignal bestimmt werden (Fig. 3a).

$$\text{Gesamtlänge der Sprachpausen} = (t_1 - t_0) + (t_3 - t_2)$$

$$\text{Gesamtlänge des Signalabschnitts} = (t_4 - t_0)$$

$$\text{Sprachpausenanteil} = \frac{\text{Gesamtlänge der Sprachpausen}}{\text{Gesamtlänge des Signalabschnitts}}$$

[0028] Wird davon ausgegangen, dass das Verhältnis von aktiver Sprache zu Sprachpausen während der Übertragung weitgehend konstant bleibt, kann dieser Wert auch auf das gestörte Signal übertragen werden.

[0029] Ist der Anteil an Sprachpausen am gesamten Sprachsignal bekannt und wird dieser Anteil als Häufigkeitsschwelle definiert, so kann aus der Häufigkeitsverteilung der Kurzzeit-Intensitäten der der Häufigkeitsschwelle entsprechende Intensitätsschwellwert ermittelt werden.

[0030] In Fig. 4 ist als Beispiel ein Anteil an Sprachpausen von 58 % eingetragen. Dieser Häufigkeitsschwelle $P_z = 0.58$ entspricht ein Intensitätsschwellwert von $N = 34.5$ sone, das bedeutet, dass von 58 % der Signalsegmente der Intensitätswert (Lautheit) von 34,5 sone nicht überschritten wird.

[0031] Der Bereich unterhalb des Intensitätsschwellwertes zeigt die Häufigkeitsverteilung für Intensitätswerte von Signalsegmenten in den Sprachpausen und kann für die Ermittlung von Intensitätskennwerten von den Hintergrundgeräuschen in den Sprachpausen benutzt werden.

[0032] Es wird davon ausgegangen, dass kein Sprachpausensegment einen höheren Intensitätswert als ein Sprachsegment besitzt, so dass der Intensitätsschwellwert als Maximalwert für das Hintergrundgeräusch in Sprachpausen angesehen werden kann.

Ermittlung des arithmetischen Mittelwertes von Intensitäten

[0033] Aus der kumulativen Verteilungsfunktion lässt sich auch der arithmetische Mittelwert aller Segmente ableiten, deren Intensitäten sich unter einer vorher ermittelten Häufigkeitsschwelle befinden. Dazu ist zunächst eine Differenzierung der kumulativen Verteilungsfunktion $P(x)$ in eine Verteilungsdichtefunktion $p(x)$ vorzunehmen.

Das arithmetische Mittel aller ausgewerteten Intensitäten X des Gesamtsignals berechnet sich wie bekannt aus dem Integral der Verteilungsdichtefunktion $p(x)$:

$$\bar{X} = \int_{-\infty}^{\infty} x p(x) dx$$

Gl. 1

5

[0034] Eine Begrenzung der Integration bei einem bestimmten Wert x_G ermöglicht die Ermittlung des arithmetischen Mittelwertes über alle Werte X , die unter diesem Grenzwert liegen. Dabei ist jedoch das Ergebnis mit der Häufigkeit $P(x_G)$ zu wichten. Diese Häufigkeit entspricht dem Integral über $p(x)$ bis zum Wert x_G .

10

$$\bar{X} = \int_{-\infty}^{x_G} x p(x) dx \Big/ \int_{-\infty}^{x_G} p(x) dx = \int_{-\infty}^{x_G} x p(x) dx \Big/ P(x_G)$$

Gl.2

15

[0035] Der Intensitätsschwellwert x_G kann aus der Verteilungsfunktion $P(x)$ abgeleitet werden. Im Beispiel nach Fig. 4 ist der Häufigkeitsschwellwert $P(x_G)$ der Anteil von Sprachpausen im Gesamtsignal $P_z = 0.58$, dem der Intensitätsschwellwert $x_G = 34.5$ *sone* zugeordnet ist. Das arithmetische Mittel aller Segmente mit einer Intensität, die geringer als x_G ist, berechnet sich nach Gl. 2, wobei $x_G = 34.5$ *sone* gilt. Die Häufigkeit von 58% entspricht hier dem Wichtungswert $P(x_G=34.5) = 0.58$. Grafisch ist dieses Vorgehen in Fig. 5 dargestellt.

20

[0036] Wird nun wieder davon ausgegangen, dass die Intensitäten von Segmenten in Sprachpausen, die Intensitäten von Sprachsegmenten nicht übersteigen oder das Hintergrundgeräusch nur schwache zeitliche Fluktuationen aufweist, kann der berechnete arithmetische Mittelwert als Mittelwert der Intensität in Sprachpausen betrachtet werden.

25

Vereinfachtes Verfahren zur Bestimmung des arithmetischen Mittelwertes

30

[0037] Ein vereinfachtes Verfahren zur Bestimmung des Mittelwertes über alle X geht von der Annahme aus, dass die relative Häufigkeitsverteilung der Intensitätswerte der Signalsegmente im Bereich $P(x) = 0$ bis zum Häufigkeitsschwellwert von Sprachpausen P_z durch eine gewichteten Normalverteilung $G(x, \mu, \sigma^2)$ angenähert werden kann. Der Wert für die Verteilungsfunktion $G(x, \mu, \sigma^2)$ für $x \rightarrow \infty$ ist 1. Wie bekannt, entspricht der Wert x , bei dem gilt $G(x, \mu, \sigma^2) = 0.5$, dem arithmetischen Mittel über alle Einzelwerte X .

35

[0038] Gelingt eine Näherung der relativen Häufigkeitsverteilung $P(x)$ im Bereich von $P(x) = 0$ bis P_z mit einer gewichteten Normalverteilung $\kappa P_z G(x, \mu, \sigma^2)$, dann entspricht der arithmetische Mittelwert über X für die gewichtete Normalverteilung dem Wert x für den gilt: $G(x, \mu, \sigma^2) = 0.5 \kappa P_z$. Durch die Annahme, dass $\kappa P_z G(x, \mu, \sigma^2)$ die Verteilung $P(x)$ im Bereich von $P(x) = 0$ bis P_z gut annähert und $\kappa \geq 1$ ist, entspricht der gesuchte arithmetische Mittelwert dem Wert x_A , für den gilt $P(x_A) = 0.5 \kappa P_z$.

40

[0039] Für den hier betrachteten Anwendungsfall von Sprache mit additivem Hintergrundgeräusch zeigen Werte für $\kappa = 1 \dots 1.3$ gute Approximationsergebnisse. In Fig. 6 ist ein Beispiel für die Annäherung durch gewichtete Normalverteilungen gezeigt. Dabei wurde ein Wert $\kappa = 1.1$ gewählt. Das Diagramm zeigt Sprache als Hintergrundgeräusch und hat einen Sprachpausenanteil von 58 %. Die starke zeitliche Fluktuation des Sprachhintergrundes lässt sich deutlich als flachere Steigung im Bereich $N = 0 \dots 40$ *sone* erkennen. Der arithmetische Mittelwert, der aus der Normalverteilungsfunktion mit $P(x_A) = 0.5 \kappa P_z = 0.32$ abgeleitet wird, beträgt 20 *sone*.

45

[0040] Der Vorteil dieses vereinfachten Verfahrens ist die geringere Rechenintensität, da auf die Berechnung der Verteilungsdichte und deren Integration verzichtet werden kann. Es ist ebenfalls nicht notwendig, die Normalverteilungsfunktion $\kappa P_z G(x, \mu, \sigma^2)$ exakt zu bestimmen, es genügt bereits die Festlegung von κ . Da P_z bekannt ist, wird der Mittelwert über alle $X < x_G$ als Wert x_A bestimmt, bei dem gilt $P(x_A) = 0.5 \kappa P_z$. Der arithmetische Mittelwert über alle X bis x_G entspricht somit dem Intensitätswert, der einem Häufigkeitswert von $0.5 \cdot \kappa \cdot \text{Anteil der Sprachpausen}$ am Gesamtsignal entspricht, d.h. der Intensität, die von einem Anteil von Segmenten von $0.5 \cdot \kappa \cdot \text{Anteil der Sprachpausen}$ nicht überschritten wird.

50

Bestimmung weiterer statistischer Kennwerte

55

[0041] Auch andere statistische Intensitätskennwerte können mit diesem Verfahren ermittelt werden. In Fig. 7 ist am Beispiel aus Fig. 4 demonstriert, wie aus der Funktion der Intensitätswert ermittelt werden kann, der von nur 20% der

Sprachpausensegmente überschritten wird (20%-Perzentil-Lautheit).

[0042] Im angeführten Beispiel wird der Intensitätswert gesucht, der von 80% der Segmente in Sprachpausen unterschritten wird, d.h. gesucht wird der Abszissenwert, der für den Ordinatenwert $P = 0.58 * 0.8 = 0.46$ gilt. Der Wert ist aufgrund des im Beispiel gewählten wenig schwankenden Störgeräusches nur wenig geringer als der Maximalwert.

Ausführungsbeispiel für die Bestimmung des arithmetischen Mittelwertes aus der Verteilungsdichtefunktion

[0043] Das hier vorgestellte Ausführungsbeispiel des Verfahrens zur Intensitätsbestimmung von Hintergrundgeräuschen ermittelt den arithmetischen Mittelwert aller Lautheiten der Segmente, die unter einer bestimmten Häufigkeitsschwelle liegen. Diese Häufigkeitsschwelle entspricht dem Anteil an Sprachpausen im Signal und der errechnete arithmetische Mittelwert wird als mittlere Lautheit in Sprachpausen betrachtet. Dazu wird in diesem Ausführungsbeispiel die Verteilungsdichtefunktion benutzt.

[0044] Vorbedingung ist, dass beide Signale, d.h. das ungestörte Quellsprachsignal und das gestörte zu bewertende Signal, vollständig aufgezeichnet vorliegen.

[0045] Zunächst wird mittels einer geeigneten Schwelle anhand des Quellsprachsignals der Anteil an Sprachpausen P_z in diesem Signal bestimmt.

[0046] Der zweite Schritt ist die Berechnung der gewünschten Intensitätswerte für aufeinanderfolgende kurze Signalsegmente des zu bewertenden Sprachsignals. In diesem Ausführungsbeispiel werden die Lautheiten nach ISO532 in aufeinanderfolgenden Signalabschnitten von 16ms Länge berechnet. Die Verteilungsfunktion wird durch eine Reihe von Einzelwerten (diskrete relative Häufigkeitsverteilung) angenähert. Diese Einzelwerte werden durch aufeinanderfolgende Indexe m bezeichnet. Die Reihe von Einzelwerten ist bei einem Maximalwert M begrenzt (z.B.: $P_0 \dots P_{200}$). In der Auswertung wird jeder Einzelwert P_m - dessen Index die ermittelte Intensität X des ausgewerteten Signalsegments übersteigt - um den Zähler 1 erhöht. Nach Auswertung des gesamten Signals werden alle Einzelwerte durch die Anzahl aller ausgewerteten Signal-segmente dividiert. Jeder Einzelwert P_m enthält dann die relative Häufigkeit der Signal-segmente, die eine Lautheit kleiner als der Wert des Indexes aufweisen.

[0047] Anhand des vorher ermittelten Anteils an Sprachpausen P_z , wird derjenige Häufigkeitwert P_s ermittelt, welcher die geringste absolute Differenz zu P_z besitzt. Der Index S dieses Einzelwertes P_s gibt die entsprechende Lautheit an, d.h. der Lautheit, die von einem Anteil P_s aller Segmente nicht überschritten wird. Zur Bestimmung des arithmetischen Mittels der Lautheiten aller Segmente, deren Lautheiten sich unter der vorgegebenen Häufigkeitsschwelle P_s befinden, ist als nächstes die Umwandlung der diskreten Häufigkeitsverteilung $P_0 \dots P_M$, in eine diskrete Häufigkeitsdichte (Streifenhäufigkeit) $p_0 \dots p_{M-1}$ vorzunehmen. Dazu werden die Differenzen zweier aufeinanderfolgender Einzelwerte gebildet und als Wertefolge $p_0 \dots p_{N-1}$ abgelegt:

$$p_m = p_{m+1} - p_m \quad \text{für alle } m = 0 \dots M-1 \quad \text{Gl.3}$$

[0048] Der Wert p_m enthält dann die relative Häufigkeit der Segmente, deren Lautheit sich zwischen m und $m+1$ befindet. Der gesuchte arithmetische Mittelwert entspricht der gewichteten Summe über die Streifenhäufigkeit P_m bis $m = S$, d.h. der Lautheit, die von einem Anteil P_s aller Segmente nicht überschritten wird:

$$\tilde{N}_{av} = \frac{\sum_{m=0}^S (m + \frac{1}{2}) p_m}{\sum_{m=0}^S p_m} = \frac{\sum_{m=0}^S (m + \frac{1}{2}) p_m}{P_s} \quad \text{Gl. 4}$$

[0049] Der Korrekturwert $\frac{1}{2}$ entspricht dem halben Abstand zweier aufeinanderfolgender Indexe. Der Wert p_m enthält die relative Häufigkeit von Segmenten, deren Lautheiten sich zwischen m und $m+1$ befinden. Der Erwartungswert aller hier erfassten Lautheiten ist, bei angenommener Gleichverteilung der Lautheiten von $m \dots m+1$, daher $m+0.5$.

[0050] Das Verfahren liefert wie im Anwendungsfall beschrieben, eine diskrete Häufigkeitsverteilung mit einer Auflösung / *some*, da der Index m ganzzahlig ist und die Lautheitswerte direkt den entsprechenden Indexen zugeordnet werden. Um gegebenenfalls andere höhere oder verringerte Auflösungen zu erzielen ist der Lautheitswert vor Berechnung der relativen Häufigkeitsverteilung mit entsprechenden Faktoren zu multiplizieren.

[0051] Zur Demonstration der Messsicherheit des vorgestellten Verfahrens sind in Tabelle 1 Messwerte für verschiedene Signale und Hintergrundgeräusche aufgeführt. Es wurde Sprachsignale von 32 s Länge und verschiedenem Anteil an Sprachpausen (35%, 58% und 91%) jeweils mit verschiedenen Geräuschen gemischt. Als Geräusche wurde

EP 1 382 034 B1

zunächst weißes Rauschen mit verschiedenen Sprach-Geräusch-Abständen benutzt. Des weiteren wurde auch kontinuierlich gesprochene Sprache sowie zwei Geräusche aus realen akustischen Umgebungen (Straße und Büro) eingesetzt.

[0052] Vor Berechnung der Häufigkeitsverteilung wird eine Multiplikation aller Lauheitswerte mit dem Faktor 2 durchgeführt, um die Auflösung der Darstellung bei Benutzung ganzzahliger Indexe zu erhöhen. Dies entspricht dann einer Lautheitsstufung bei ganzzahligen Indexen von *0.5 sone*. Mit einer Begrenzung der Häufigkeitsverteilungsfunktion bei P_{200} , können so Lautheiten von *0 ... 100 sone* in Schritten von *0.5 sone* abgebildet werden. Es ist aber zu beachten, dass dieser Faktor als Divisor zur Korrektur auf alle Ergebnisse angewendet wird muss. Im hier gewählten Ausführungsbeispiel bedeutet dies, dass der errechnete arithmetische Mittelwert durch 2 zu teilen ist.

[0053] Erläuterungen zu Tabelle 1: Der Sprach-Geräusch-Abstand dient lediglich zur Information; Grundlage bildet der Abstand des mittleren Effektivpegels bei Sprachaktivität zum mittleren Effektivpegel des Hintergrundgeräusches. Der mittlere Lautheitswert (Zielwert) wurde in einer Referenzmessung bestimmt, bei der die Sprachpausen manuell markiert und in Segmenten zu 16 ms ausgewertet wurden. Die berechneten Standardabweichungen beziehen sich auf die derart gemessenen Referenz-Lautheiten und geben Information über die Stärke der auftretenden Fluktuationen. Die Messwerte in Spalte 5 wurden mit dem in diesem Ausführungsbeispiel beschriebenen Verfahren ermittelt.

Tabelle 1

Geräusch	SNR	mittlere Lautheit (sone) Zielwert	Standardabweichung der Segment-Lautheiten	mittlere Lautheit (sone) gemessen mit beschriebenen Verfahren	Abweichung (Messfehler) abs. / rel.
Pausenanteil des Sprachsignals 91%					
weißes Rauschen	6 dB	41.4	1.55	42.0	0.6 / 1.4%
weißes Rauschen	10 dB	32.3	1.22	32.6	0.3/0.9%
weißes Rauschen	16 dB	22.2	0.87	22.3	0.1 / 0.4%
Sprache	6 dB	21.3	11.7	20.6	-0.7/-3.3%
Sprache	10 dB	16.5	9.16	16.2	-0.3/-1.8%
Sprache	16 dB	11.2	6.21	11.3	0.1/0.9%
Straßengeräusch	10 dB	26.0	3.22	26.2	0.2/0.8%
Bürogeräusch	10 dB	26.3	2.78	26.6	0.3/1.1%
Pausenanteil des Sprachsignals 58%					
weißes Rauschen	6 dB	41.3	1.55	44.8	3.5/8.5%
weißes Rauschen	10 dB	32.3	1.22	34.2	1.9/6.0%
weißes Rauschen	16 dB	22.1	0.87	22.6	0.5/2.2%
Sprache	6 dB	20.7	11.7	19.0	-1.7/-8.2%
Sprache	10 dB	16.0	9.16	15.4	-0.6 /-3.8%
Sprache	16 dB	10.7	6.21	10.8	0.1 / 0.9%
Straßengeräusch	10 dB	26.1	3.22	27.0	0.9/3.4%
Bürogeräusch	10 dB	26.3	2.78	27.3	1.0 / 3.8%
Pausenanteil des Sprachsignals 35%					
weißes Rauschen	6 dB	41.3	1.55	46.1	4.8 / 11.6%
weißes Rauschen	10 dB	32.3	1.22	35.6	3.3 / 10.2%

EP 1 382 034 B1

Tabelle 1 (fortgesetzt)

Geräusch	SNR	mittlere Lautheit (sone) Zielwert	Standardabweichung der Segment-Lautheiten	mittlere Lautheit (sone) gemessen mit beschriebenen Verfahren	Abweichung (Messfehler) abs. / rel.
Pausenanteil des Sprachsignals 35%					
weißes Rauschen	16 dB	22.1	0.87	23.3	1.2/5.4%
Sprache	6 dB	20.0	11.22	17.6	-2.4 / -12%
Sprache	10 dB	15.6	8.7	15.0	-0.6 / -3.8%
Sprache	16 dB	10.9	5.93	11.8	0.9/8.3%
Straßengeräusch	10 dB	26.1	3.22	27.3	1.2/4.6%
Bürogeräusch	10 dB	26.3	1.78	27.9	1.6/6.1%

[0054] Zunächst ist festzustellen, dass die Messsicherheit mit zunehmenden Pausenanteil im zu bewertenden Signal zunimmt. Eine Zunahme der Messsicherheit ist ebenfalls bei sinkender Geräuschintensität sowie geringerer zeitlicher Fluktuation des Hintergrundgeräusches festzustellen. Ausgehend von einem typischen Anteil an Sprachpausen in einer Telefonkommunikation von $P_z > 50\%$ sind die mit dem vorgestellten Verfahren erreichten Messwerte selbst bei stärkeren Fluktuationen im Hintergrundgeräusch (z.B. Sprache) zufriedenstellend.

Ausführungsbeispiel für die Bestimmung des arithmetischen Mittelwertes mit vereinfachtem Verfahren

[0055] Dieses spezielle Ausführungsbeispiel zeigt eine Anwendung des beschriebenen vereinfachten Verfahren zur Bestimmung des arithmetischen Mittels unter Nutzung einer gewichteten Normalverteilung.

[0056] Das vereinfachte Verfahren verzichtet auf die Berechnung der Streifenhäufigkeit und leitet einen Schätzwert für das arithmetische Mittel der Lautheiten aller Segmente, deren Lautheiten sich unter der vorgegebenen Häufigkeitsschwelle P_z befinden, direkt aus der relativen Häufigkeitsverteilung P_m ab. Wie beschrieben muss lediglich der Wert κ für die Schätzung festgelegt werden.

[0057] in diesem Ausführungsbeispiel wird mit $\kappa = 1.1$ definiert. Der Schätzwert entspricht dann dem Lautheitswert, der von einem Anteil von $0.5 * 1.1 * P_z$ aller ausgewerteten Segmente nicht überschritten wird. Im Ausführungsbeispiel entspricht dieser Schätzwert des arithmetischen Mittels der Lautheiten, dem Index m des Häufigkeitswertes, welcher die geringste absolute Differenz zu $0.55 P_z$ besitzt. In Tabelle 2 sind die Messwerte aufgeführt, die mit diesem vereinfachten Verfahren gewonnen worden. Auch hier wurden zur Erhöhung der Auflösung auf 0.5 sone alle Lautheitswerte vor Berechnung der Häufigkeitsverteilung mit dem Faktor 2 multipliziert und die Ergebnisse entsprechend korrigiert.

Tabelle 2

Geräusch	SNR	mittlere Lautheit (sone) Zielwert	Standardabweichung der Segment-Lautheiten	mittlere Lautheit (sone) gemessen mit vereinfachtem Verfahren	Abweichung (Messfehler) abs. / rel.
Pausenanteil des Sprachsignals 91%					
weißes Rauschen	6 dB	41.4	1.55	41.5	0.1/0.2%
weißes Rauschen	10 dB	32.3	1.22	32.5	0.2/0.6%
weißes Rauschen	16 dB	22.2	0.87	22.5	0.3/ 1.3%
Sprache	6 dB	21.3	11.7	20.5	-0.8/-3.8%
Sprache	10 dB	16.5	9.76	16.5	0.0/0.0%
Sprache	16 dB	11.2	6.21	11.0	-0.2/1.8%

EP 1 382 034 B1

Tabelle 2 (fortgesetzt)

Geräusch	SNR	mittlere Lautheit (sone) Zielwert	Standardabweichung der Segment-Lautheiten	mittlere Lautheit (sone) gemessen mit vereinfachtem Verfahren	Abweichung (Messfehler) abs. / rel.
Pausenanteil des Sprachsignals 91%					
Straßengeräusch	10 dB	26.0	3.22	26.0	0.0 / 0.0%
Bürogeräusch	10 dB	26.3	2.78	26.5	0.2/0.6%
Pausenanteil des Sprachsignals 58%					
weißes Rauschen	6 dB	41.3	1.55	41.50	0.2/0.5%
weißes Rauschen	10 dB	32.3	1.22	32.5	0.2/0.6%
weißes Rauschen	16 dB	22.1	0.87	22.5	0.4 / 1.8%
Sprache	6 dB	20.7	11.7	20.0	-0.7/-3.4%
Sprache	10 dB	16.0	9.16	16.0	0.0/0.0%
Sprache	16 dB	10.7	6.21	11.0	0.3/2.8%
Straßengeräusch	10 dB	26.1	3.22	26.0	-0.1 /-0.4%
Bürogeräusch	10 dB	26.3	2.78	26.5	0.2/0.8%
Pausenanteil des Sprachsignals 35%					
weißes Rauschen	6 dB	41.3	1.55	41.0	-0.3/0.7%
weißes Rauschen	10 dB	32.3	1.22	32.5	0.2/0.6%
weißes Rauschen	16 dB	22.1	0.87	22.5	0.4/ 1.8%
Sprache	6 dB	20.0	11.12	19.0	-1.0/-5%
Sprache	10 dB	15.6	8.7	15.5	-0.1 / -0.6%
Sprache	16 dB	10.9	5.93	11.5	0.6/5.5%
Straßengeräusch	10 dB	26.1	3.22	25.5	-0.6/-1.4%
Bürogeräusch	10 dB	26.3	2.78	26.5	0.2 / 0.8%

[0058] Das vereinfachte Verfahren spart nicht nur Rechenzeit sondern liefert in den ausgewerteten Beispielen Messwerte mit einer deutlich höheren Genauigkeit im Vergleich zu den Werten aus Tabelle 1. Da als Schätzwert direkt der Index m benutzt wird, ist die Genauigkeit der Schätzung auf die Auflösung der relativen diskreten Häufigkeitsverteilung (hier: 0.5 sone) begrenzt.

[0059] Mit dem beschriebenen vereinfachten Messverfahren werden auch bei Geräuschen mit stärkerer Fluktuation gute Messwerte erzielt. Bei den gewählten Sprach-Geräusch-Abständen von 6dB kann auch nicht mehr davon ausgegangen werden, dass alle Lautheiten in Sprachpausen eine geringere Lautheit als Sprachsegmente aufweisen. Trotzdem sind die Messwerte kaum verfälscht wurden. Das beschriebene vereinfachte Verfahren eignet sich zudem auch für Signale mit geringerem Pausenanteil.

Ausführungsbeispiel für die Bestimmung von Perzentil-Lautheiten aus der relativen Häufigkeitsverteilung

[0060] Die Perzentil-Lautheit aller Segmente, die unter einer bestimmten Häufigkeitsschwelle P_z liegen, kann durch Multiplikation dieser relativen Häufigkeit P_z mit einem Wert $1 - \text{Perzentilwert}$ erfolgen (z.B. 10%-Perzentil-Lautheit: $P_{Z10\%} = 0.9 * P_z$). Der ganzzahlige Index m des Häufigkeitwertes P_m , welcher die geringste absolute Differenz zu $P_{S10\%}$ besitzt, liefert den gesuchten Perzentil-Lautheitswert.

[0061] In Tabelle 3 sind für die bereits in den Tabellen 1 und 2 aufgeführten Beispiele die 10%-Perzentil-Lautheiten aufgeführt und werden mit einem manuell bestimmten Referenzwert verglichen.

Tabelle 3

Geräusch	SNR	10%-Perzentil-Lautheit (sone) Zielwert	Standardabweichung der Segment-Lautheiten	10%-Perzentil-Lautheit (sone) gemessen über Häufigkeitsverteilung	Abweichung (Messfehler) abs. rel.
Pausenanteil des Sprachsignals 91%					
weißes Rauschen	6 dB	42.5	1.55	43.0	0.5/1.2%
weißes Rauschen	10 dB	33.0	1.22	34.0	1.0/3.0%
weißes Rauschen	16 dB	22.5	0.87	23.5	1.0/4.4%
Sprache	6 dB	37.0	11.7	34.5	-2.5/-6.8%
Sprache	10 dB	28.5	9.16	27.5	-1.0/-3.5%
Sprache	16 dB	19.0	6.21	19.5	0.5/2.6%
Straßengeräusch	10 dB	29.5	3.22	30.0	0.5/1.7%
Bürogeräusch	10 dB	29.0	2.78	29.5	0.5/1.7%
Pausenanteil des Sprachsignals 58%					
weißes Rauschen	6 dB	42.5	1.55	42.5	0.0/0.0%
weißes Rauschen	10 dB	33.0	1.22	33.5	0.5/1.5%
weißes Rauschen	16 dB	22.5	0.87	23.0	0.5/2.2%
Sprache	6 dB	36.0	11.7	29.0	-7.0/-19%
Sprache	10 dB	28.5	9.16	24.5	-4.0/-14%
Sprache	16 dB	19.0	6.21	18.0	-1.0/-5.3%
Straßengeräusch	10 dB	30.0	3.22	29.0	-1.0/-3.3%
Bürogeräusch	10 dB	29.0	2.78	28.5	-0.5/-1.6%
Pausenanteil des Sprachsignals 35%					
weißes Rauschen	6 dB	42.5	1.55	42.5	0.0/0.0%
weißes Rauschen	10 dB	33.0	1.22	33.5	0.5/ 1.5%
weißes Rauschen	16 dB	22.5	0.87	23.5	1.0/2.2%
Sprache	6 dB	35.5	11.21	24.0	-11.5/-33%

Tabelle 3 (fortgesetzt)

Geräusch	SNR	10%-Perzentil-Lautheit (sone) Zielwert	Standardabweichung der Segment-Lautheiten	10%-Perzentil-Lautheit (sone) gemessen über Häufigkeitsverteilung	Abweichung (Messfehler) abs. rel.
Pausenanteil des Sprachsignals 35%					
Sprache	10 dB	27.5	8.7	21.0	-6.5 / -24%
Sprache	16 dB	19.0	5.93	17.5	-1.5/-7.9%
Straßengeräusch	10 dB	29.5	3.22	28.0	-1.5/-4.8%
Bürogeräusch	10 dB	29.0	1.78	28.5	-0.5 / -1.6%

[0062] Die Messwerte zeigen eine gute Abschätzung der Perzentil-Lautheit für Hintergrundgeräusche mit schwacher Fluktuation, für Sprache werden - vor allem bei geringem Pausenanteil - nur unzureichende Genauigkeiten erzielt. Lediglich bei höheren Sprach-Geräusch-Abständen sind die Ergebnisse brauchbar bis gut.

Patentansprüche

- Verfahren zur Bestimmung von Intensitätskennwerten von Hintergrundgeräuschen in Sprachpausen von Sprachsignalen, von denen das ungestörte Quellsprachsignal und das gestörte Sprachsignal aufgezeichnet zur Verfügung stehen und aus dem ungestörten Quellsprachsignal der Anteil Sprachpausen im Gesamtsignal nach bekannten Methoden ermittelt wird und das gestörte Sprachsignal in kurze aufeinanderfolgende Signalelemente unterteilt wird und für jedes Signalelement ein Intensitätswert bestimmt wird, wobei aus den Intensitätswerten der einzelnen Signalelemente des gestörten Sprachsignals die kumulative relative Häufigkeitsverteilung (1) gebildet wird, der ermittelte Anteil an Sprachpausen im Quellsprachsignal als Häufigkeitsschwelle definiert wird und die Häufigkeitsschwelle auf das gestörte Sprachsignal angewendet wird, aus der Häufigkeitsverteilung der Intensitätswerte der Signalsegmente der definierten Häufigkeitsschwelle (2) entsprechende Intensitätsschwellwert (3) ermittelt wird, alle Signalsegmente mit einem geringeren Intensitätswert als dem des Intensitätsschwellwertes zu den Sprachpausen gehörend bewertet werden, die Verteilungsfunktion für die Intensitätswerte der Signalsegmente in dem Bereich unterhalb des Intensitätsschwellwertes die Häufigkeitsverteilung für die Intensitätswerte in den Sprachpausen (4) darstellt, und dieser Bereich der Verteilungsfunktion für die Ermittlung von Intensitätskennwerten von den Hintergrundgeräuschen in den Sprachpausen genutzt werden kann.
- Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** als Intensitätskennwert von den Hintergrundgeräuschen in den Sprachpausen der arithmetische Mittelwert der Intensitätswerte der Signalelemente in den Sprachpausen bestimmt wird, und dass der arithmetische Mittelwert berechnet wird, indem aus der Häufigkeitsverteilung die Verteilungsdichte abgeleitet wird und durch eine nachfolgende Integration über die Verteilungsdichte im Bereich unterhalb des Intensitätsschwellwertes der arithmetische Mittelwert der Intensitätswerte in den Sprachpausen ermittelt wird.
- Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** als Intensitätskennwert von den Hintergrundgeräuschen in den Sprachpausen der arithmetische Mittelwert der Intensitätswerte der Signalelemente in den Sprachpausen bestimmt wird, und dass der arithmetische Mittelwert aus der Häufigkeitsverteilung ermittelt wird, indem die Intensitätsverteilung im Bereich unterhalb des Intensitätsschwellwertes durch eine mit einem Faktor gewichtete Normalverteilung angenähert wird und für die Berechnung des arithmetischen Mittelwertes der Intensitätsschwellwert mit 0,5 und dem Wichtungsfaktor multipliziert wird.
- Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** als Intensitätskennwerte von den Hintergrundgeräuschen in den Sprachpausen Perzentil-Kennwerte bestimmt werden können, dass die Perzentil-Kennwerte aus

der Häufigkeitsverteilung ermittelt werden können, indem der vorgegebene Perzentilwert von 100 Prozent subtrahiert wird, die Differenz mit dem Häufigkeitsschwellwert multipliziert wird und für den sich ergebenden Häufigkeitswert der diesem Wert entsprechende Intensitätswert als Perzentil-Kennwert aus der Verteilfunktion bestimmt wird.

5

Claims

1. Method for the determination of intensity characteristic values of background noises in speech pauses of speech signals of which the undisturbed source speech signal and the disturbed speech signal are available in recorded form, wherein, from the undisturbed source speech signal, the proportion of speech pauses in the total signal is determined according to known methods and the disturbed speech signal is broken down into short consecutive signal elements and an intensity value is determined for each signal element, wherein the cumulative relative frequency distribution (1) is formed from the intensity values of the individual signal elements of the disturbed speech signal;
- 15 the determined proportion of speech pauses in the source speech signal is defined as a frequency threshold and the frequency threshold is applied to the disturbed speech signal; the intensity threshold value (3) corresponding to the defined frequency threshold (2) is determined from the frequency distribution of the intensity values of the signal segments; all the signal segments with an intensity value lower than that of the intensity threshold value are assessed as belonging to the speech pauses;
- 20 the distribution function for the intensity values of the signal segments in the region below the intensity threshold value represents the frequency distribution for the intensity values in the speech pauses (4), and this region of the distribution function can be used for the determination of intensity characteristic values of the background noises in the speech pauses.
- 25 2. Method according to claim 1, **characterized in that** the arithmetic mean value of the intensity values of the signal elements in the speech pauses is determined as the intensity characteristic value of the background noises in the speech pauses, and **in that** the arithmetic mean value is calculated **in that** the distribution density is derived from the frequency distribution and the arithmetic mean value of the intensity values in the speech pauses is determined by subsequent integration over the distribution density in the region below the intensity threshold value.
- 30 3. Method according to claim 1, **characterized in that** the arithmetic mean value of the intensity values of the signal elements in the speech pauses is determined as the intensity characteristic value of the background noises in the speech pauses, and **in that** the arithmetic mean value is determined from the frequency distribution **in that** the intensity distribution in the region below the intensity threshold value is approximated by a normal distribution weighted with a factor and, for calculation of the arithmetic mean value, the intensity threshold value is multiplied by 0.5 and the weighting factor.
- 35 4. Method according to claim 1, **characterized in that** percentile characteristic values can be determined as intensity characteristic values of the background noises in the speech pauses, **in that** the percentile characteristic values can be determined from the frequency distribution, **in that** the specified percentile value is subtracted from 100 percent, the difference is multiplied by the frequency threshold value and, for the resulting frequency value, the intensity value corresponding to said value is determined as a percentile characteristic value from the distribution function.

45

Revendications

1. Procédé de détermination de paramètres d'intensité de bruits de fond dans les pauses à l'intérieur de signaux vocaux, dont le signal vocal d'origine, non perturbé, et le signal vocal perturbé sont disponibles sous formes d'enregistrement, la part des pauses dans le signal global étant déterminée selon des méthodes connues à partir du signal vocal d'origine non perturbé et le signal vocal perturbé étant découpé en de courts éléments successifs avec détermination d'une valeur d'intensité pour chaque élément du signal, dans lequel la répartition de fréquence relative cumulative (1) est établie à partir des valeurs d'intensité des différents éléments du signal vocal perturbé, la part mesurée des pauses dans le signal vocal d'origine est définie comme seuil de fréquence et le seuil de fréquence est appliqué au signal vocal perturbé,
- 50 la valeur seuil d'intensité (3) correspondant au seuil de fréquence défini (2) est déterminée à partir de la répartition de fréquence des valeurs d'intensité des éléments de signal,
- 55 tous les éléments de signal sont évalués à une valeur d'intensité inférieure à la valeur seuil d'intensité correspon-

5 dant aux pauses,

la fonction de répartition des valeurs d'intensité des éléments de signal située en dessous de la valeur seuil d'intensité représente la répartition de fréquence des valeurs d'intensité pendant les pauses (4), cette gamme inférieure de la fonction de répartition pouvant être utilisée pour déterminer les paramètres d'intensité de bruits de fond dans les pauses.

10 2. Procédé selon la revendication 1, **caractérisé en ce que** l'on retient comme paramètre d'intensité de bruits de fond pendant les pauses la moyenne arithmétique des valeurs d'intensité des éléments de signal pendant les pauses, la moyenne arithmétique étant calculée en déduisant la densité de répartition de la répartition de fréquence et en procédant ensuite à une intégration au niveau de la densité de répartition pour les valeurs situées en dessous de la valeur seuil d'intensité.

15 3. Procédé selon la revendication 1, **caractérisé en ce que** l'on retient comme paramètre d'intensité de bruits de fond pendant les pauses la moyenne arithmétique des valeurs d'intensité des éléments de signal pendant les pauses, la moyenne arithmétique étant calculée à partir de la répartition de fréquence en rapprochant la répartition d'intensité en dessous de la valeur seuil d'intensité d'une répartition normale pondérée par un facteur et en multipliant la valeur seuil d'intensité par 0,5 et par le facteur de pondération.

20 4. Procédé selon la revendication 1, **caractérisé en ce que** l'on retient comme paramètres d'intensité de bruits de fond pendant les pauses des percentiles, lesquels sont calculés à partir de la répartition de fréquence en soustrayant le percentile de 100 % prédéfini, en multipliant la différence avec la valeur seuil de fréquence et en déterminant la valeur d'intensité correspondant à la valeur de fréquence ainsi obtenue en tant que percentile résultant de la fonction de répartition.

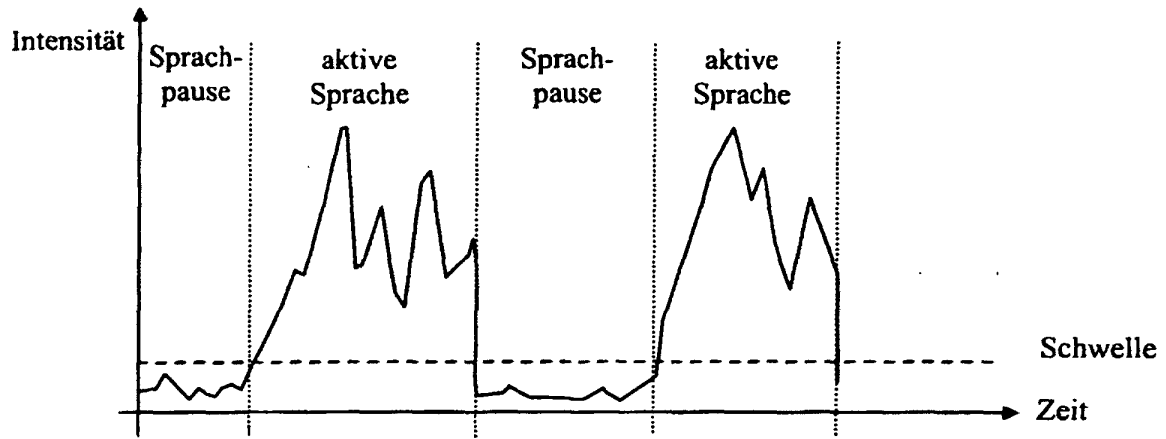


Fig. 1

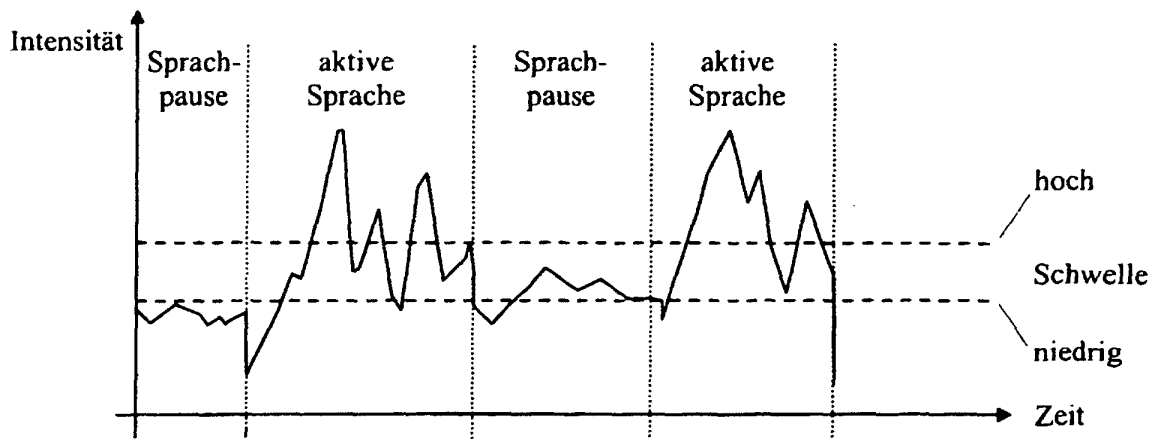


Fig. 2

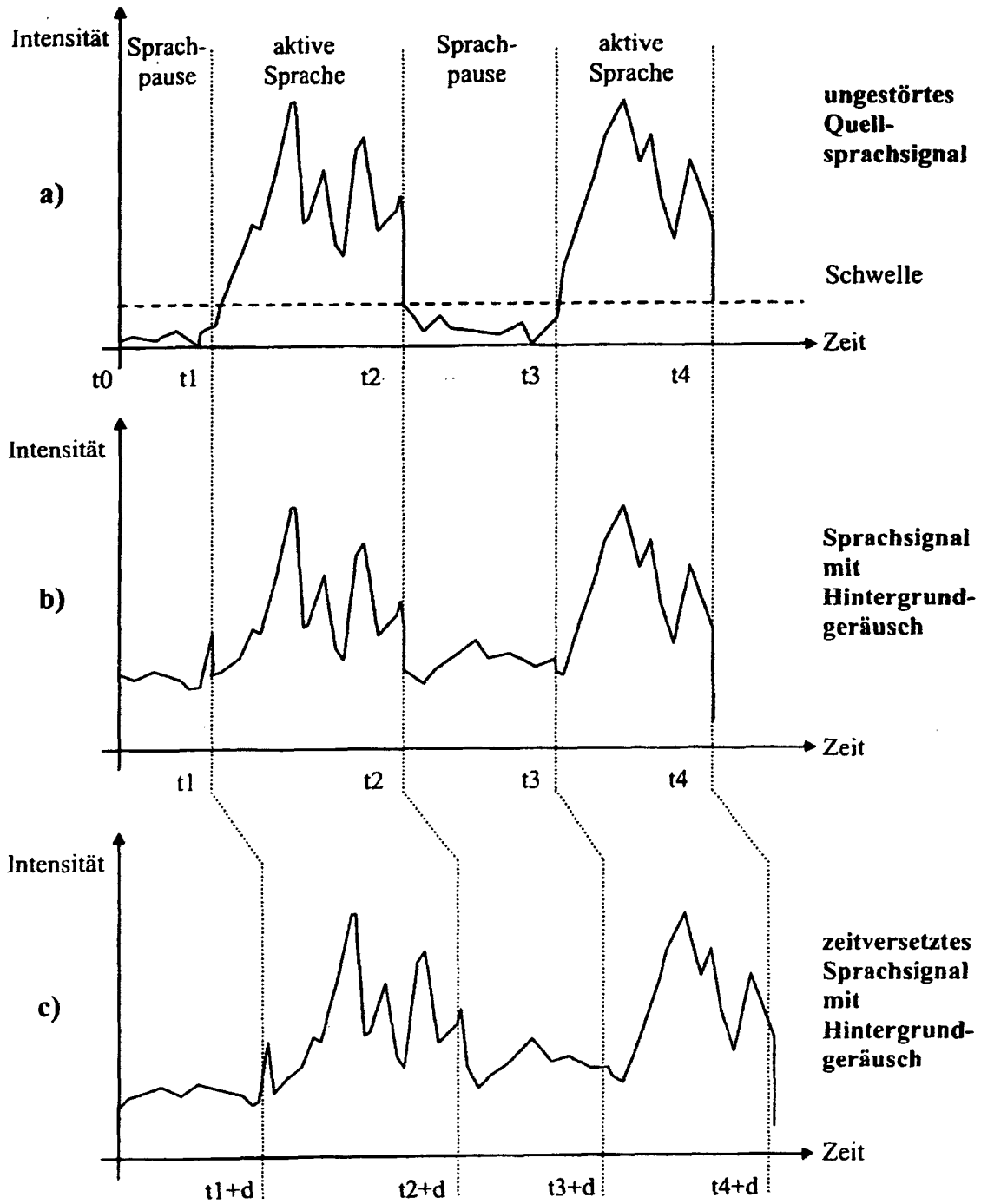


Fig. 3

Relative Häufigkeitsverteilung der Kurzzeit-Intensitäten
des gestörten Sprachsignals

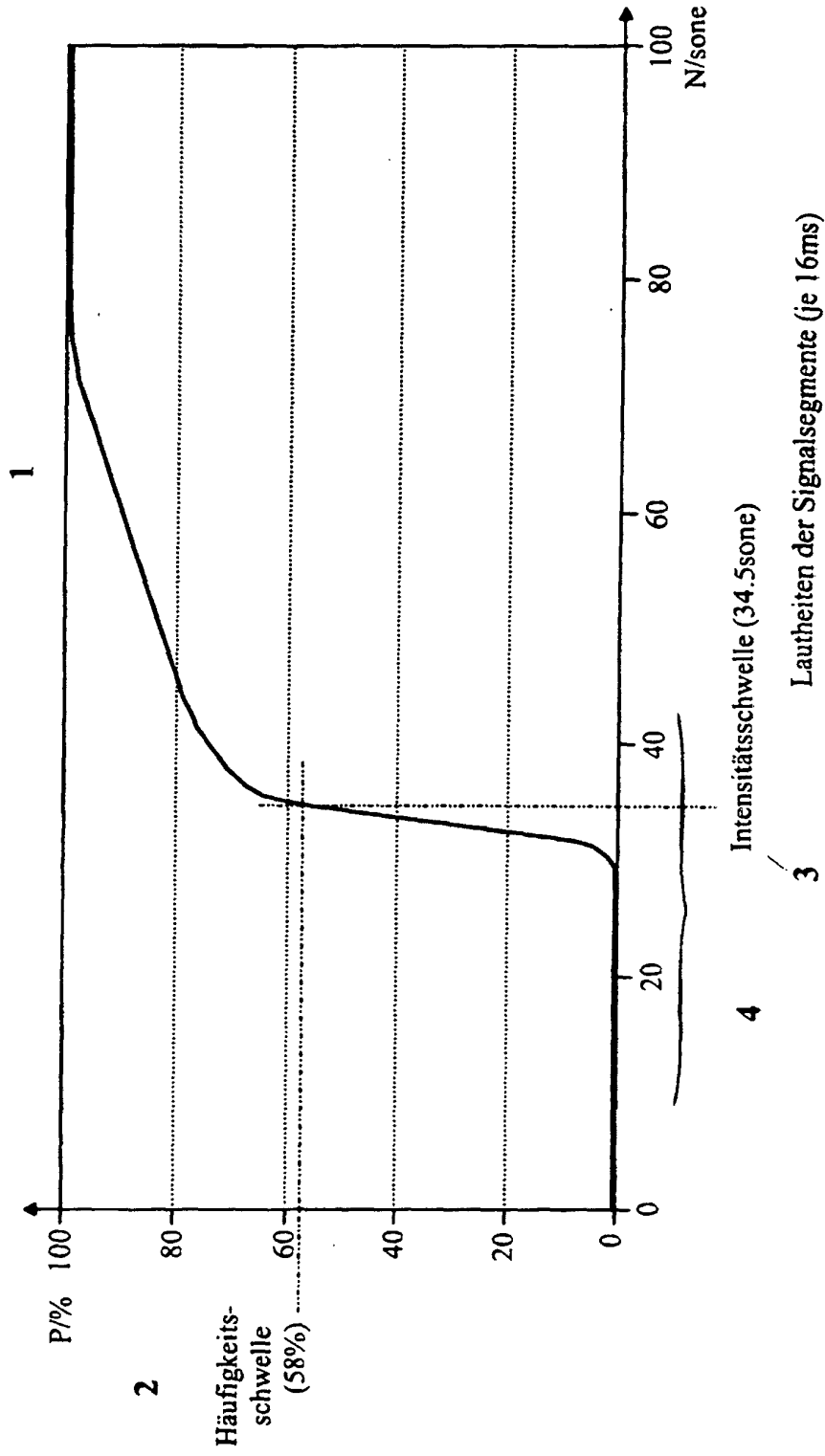
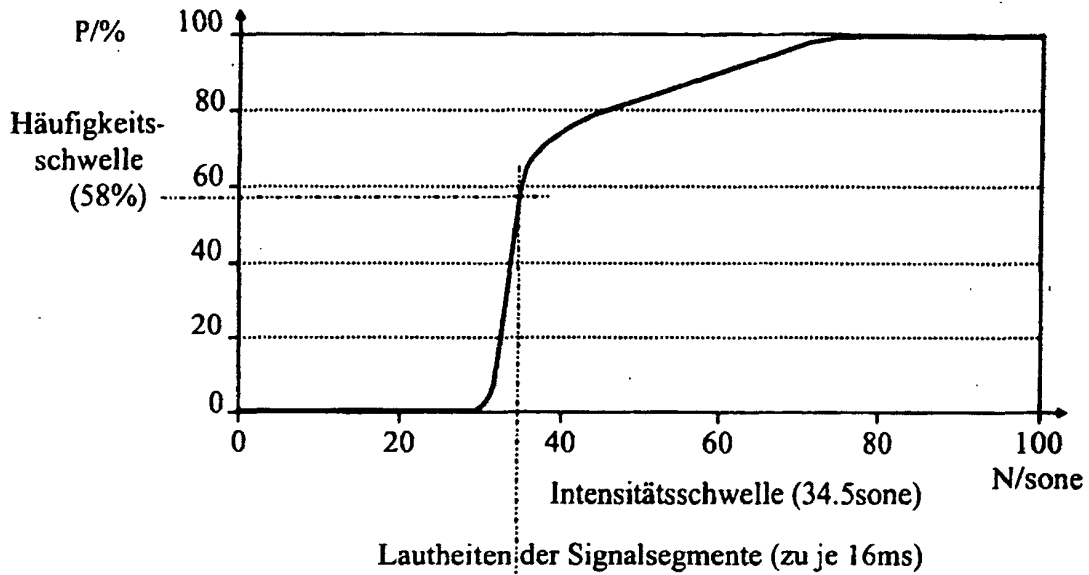


Fig. 4

Häufigkeitsverteilung von Lautheiten in Signalsegmenten



Verteilungsdichte von Lautheiten in Signalsegmenten

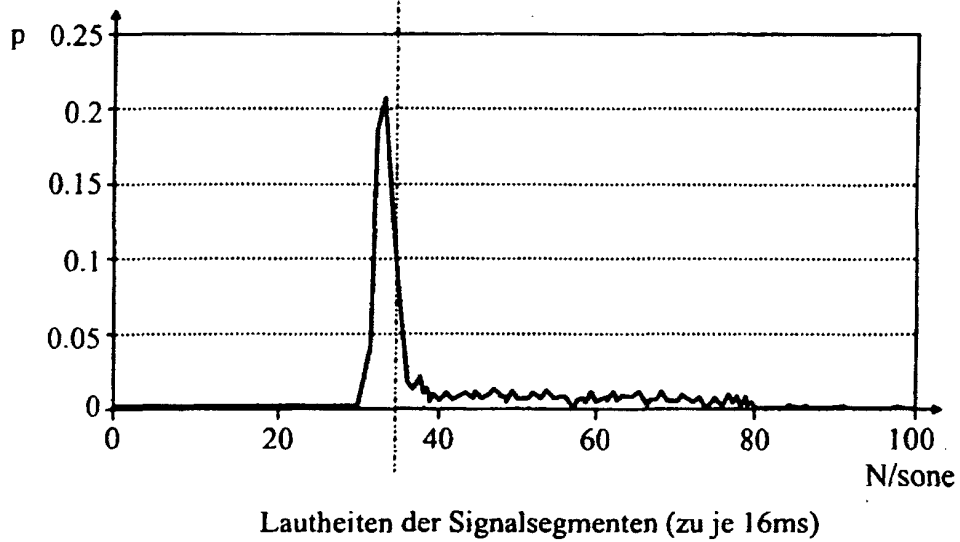


Fig.5

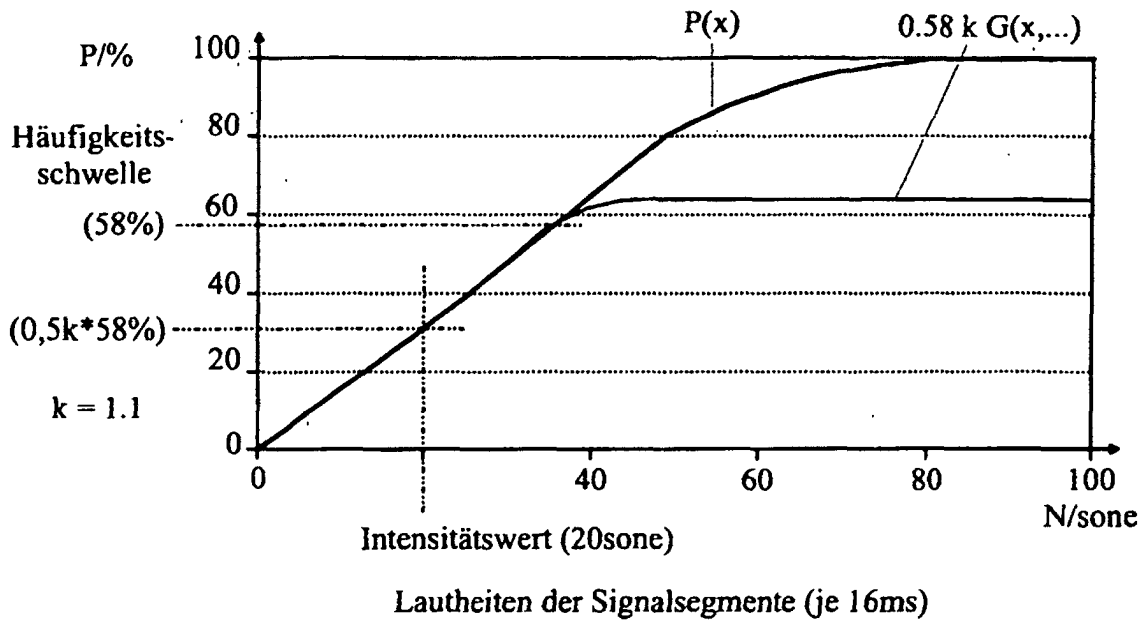


Fig. 6

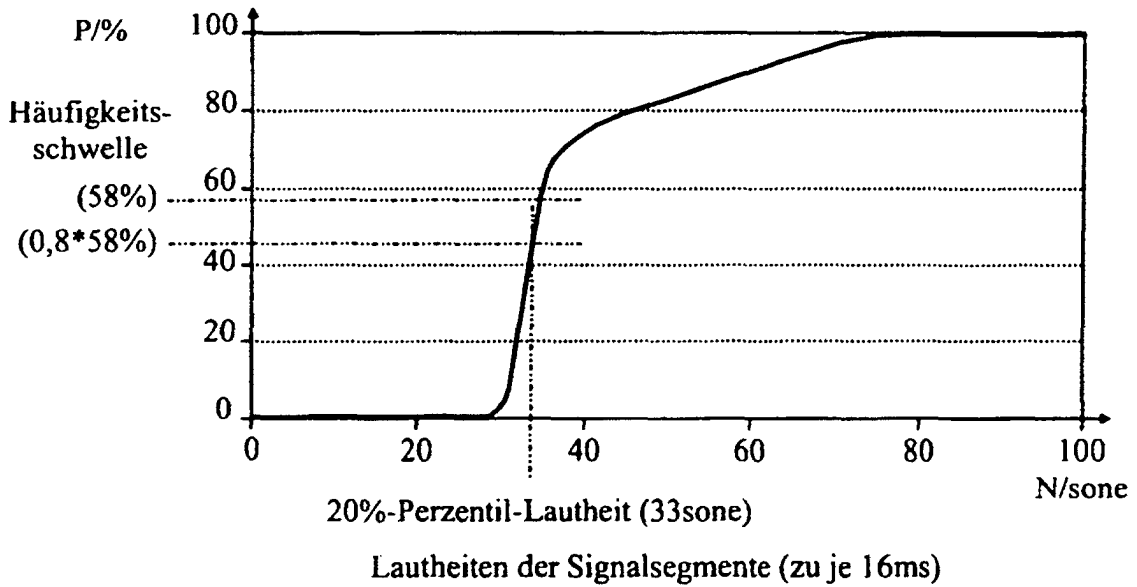


Fig.7