

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3964456号
(P3964456)

(45) 発行日 平成19年8月22日(2007.8.22)

(24) 登録日 平成19年6月1日(2007.6.1)

(51) Int. Cl. F I
H O 4 B 3/46 (2006.01) H O 4 B 3/46 Z

請求項の数 10 (全 19 頁)

(21) 出願番号	特願平6-502156	(73) 特許権者	390028587
(86) (22) 出願日	平成5年6月24日(1993.6.24)		ブリティッシュ・テレコミュニケーションズ・パブリック・リミテッド・カンパニー
(65) 公表番号	特表平8-501910		BRITISH TELECOMMUNICATIONS PUBLIC LIMITED COMPANY
(43) 公表日	平成8年2月27日(1996.2.27)		イギリス国, イーシー1エー・7エー ジェイ, ロンドン, ニューゲート・ストリート 81
(86) 国際出願番号	PCT/GB1993/001322	(74) 代理人	100058479
(87) 国際公開番号	W01994/000922		弁理士 鈴江 武彦
(87) 国際公開日	平成6年1月6日(1994.1.6)	(74) 代理人	100084618
審査請求日	平成12年6月23日(2000.6.23)		弁理士 村松 貞男
審査番号	不服2004-17407(P2004-17407/J1)	(74) 代理人	100092196
審査請求日	平成16年8月23日(2004.8.23)		弁理士 橋本 良郎
(31) 優先権主張番号	9213459.2		
(32) 優先日	平成4年6月24日(1992.6.24)		
(33) 優先権主張国	英国 (GB)		
(31) 優先権主張番号	9312758.7		
(32) 優先日	平成5年6月21日(1993.6.21)		
(33) 優先権主張国	英国 (GB)		

最終頁に続く

(54) 【発明の名称】 電気通信装置の客観的音声品質測定の方法および装置

(57) 【特許請求の範囲】

【請求項 1】

試験信号を供給する手段と、試験される装置により歪みを受けた前記試験信号に対応する歪んだ信号を受信し、そして典型的な人間の聴者が知覚し得るであろう前記信号の歪みの程度を示す歪み知覚評価値を生成するように前記歪んだ信号を分析する分析手段とを含む、音声信号を処理する装置を試験する試験装置であって、

前記分析手段は、

元のおよび歪んだ信号から周波数により変化する帯域幅を有する複数の帯域制限された時間域区分信号を生成するフィルター手段と、

それぞれのかかる区分信号に関し、各振幅値が前記信号のそれぞれの時間部分に関連する複数の平均信号振幅値を生成する手段と、

各それぞれの時間部分および周波数帯域に関し、元のおよび歪んだ信号の振幅値の間の差を計算する手段と、そして、

そこから、試験される装置による信号について適用される、歪の知覚される程度の評価値を計算する手段と、

を含む

ことを特徴とする試験装置。

【請求項 2】

区分信号の帯域幅は、それらが典型的な聴者によって知覚されるのと等しいマスキングの程度を生成するように選択されることを特徴とする請求項 1 記載の試験装置。

10

20

【請求項3】

さらに、各帯域制限された時間域区分信号の複数の連続する時間部分のそれぞれに関し、前記歪んだ信号と対応して計算された試験信号成分間の差を形成することにより前記歪んだ信号の時間的マスキングを計算する手段と、前記区分信号の複数の前記時間部分にわたって前記差を評価することにより人間の聴者が知覚し得るであろう前記信号の歪みの程度を示す歪み知覚評価値 (E_A) を生成する手段を含むことを特徴とする請求項1または請求項2記載の試験装置。

【請求項4】

スペクトル区分帯域はピッチにほぼ等しい間隔であり、そして歪み知覚評価手段が前記時間部分およびスペクトル区分にわたって前記差の統計的分布の評価値 (E_E) を生成することを特徴とする請求項3記載の試験装置。

10

【請求項5】

各部分により示される時間間隔はより高い周波数区分信号よりもより低い周波数区分信号に関するほうが長いことを特徴とする請求項1乃至4のいずれか1項記載の試験装置。

【請求項6】

音声信号処理装置に試験信号を供給するステップと、前記音声信号処理装置により歪んだ試験信号を分析するステップと、典型的な人間の聴者が知覚し得るであろう前記信号の歪みの程度を決定するステップと、前記決定された程度を示す歪み知覚評価値を生成するステップとを有する音声信号処理装置を試験する方法であって、

20

前記試験信号は人間の音声に類似するスペクトルを有し、そして前記分析するステップは

元のそして前記歪んだ試験信号をフィルタリングし、周波数に関して変化する帯域幅を有する複数の帯域制限された時間域区分信号を生成するステップと、

各帯域制限された時間域区分信号について、各かかる区分信号に関する複数の平均信号振幅値を生成するステップと、 なお、各振幅値は前記信号のそれぞれの時間部分に関し、各それぞれの時間部分および周波数帯域に関し、元のおよび歪んだ信号の振幅値の間の差を計算するステップと、そして、

そこから、試験される装置による信号について適用される、歪の知覚される程度の評価値を計算するステップを含む

30

ことを特徴とする試験方法。

【請求項7】

それらが典型的な聴者によって知覚されるのと等しいマスキングの程度を生成するように複数の帯域制限された区分を選択することにより、前記試験信号が典型的な人間の聴者に知覚され得るであろう程度を分析するステップと、

前記試験信号の知覚と歪んだ信号の知覚間の差の評価値を導くステップと、そして前記差に従って前記歪み知覚評価値 (E_A) を導くステップ

を含むことを特徴とする請求項6記載の試験方法。

【請求項8】

各区分信号の前記時間部分の各々に関し、前記歪んだ信号の時間マスキングを計算し、前記歪んだ信号と前記試験信号の対応する部分および区分間の差を形成するステップと、

40

前記時間部分および前記区分信号にわたって前記差の評価値 (E_A) を導くステップを含むことを特徴とする請求項6または請求項7記載の試験方法。

【請求項9】

前記帯域制限された区分信号はピッチにほぼ等しい間隔で配され、歪み知覚評価手段が前記時間部分および区分信号にわたって前記差の統計的分布の評価値 (E_E) を生成することを特徴とする請求項8記載の試験方法。

【請求項10】

所定の時間間隔で区分信号レベルを示す連続する値の時間シーケンスを生成するステップを含み、前記時間間隔はより低周波区分信号に関する方がより高周波区分信号に関するも

50

のよりも長いことを特徴とする請求項6乃至9のいずれか1項記載の試験方法。

【発明の詳細な説明】

本発明は電気通信装置の客観的音声品質測定の方法および装置に関するものである。電気通信装置（例えば、電話回線、電話ネットワーク、または符号機のような通信装置）の試験において、試験信号が試験装置の入力に伝えられ、装置の出力結果についての何らかの試験が行われる。自動的処理装置により、計算可能な信号とノイズの比のような“客観的”試験の測定値を導くことが知られている。また、人間である聴者が電気通信装置の出力を聴取し出力の品質について評価を与える、“主観的”試験が行われることが知られている。

電気通信システムのある構成要素は線形性である。従って、離散的周波数の正弦波、掃引された正弦波信号、チャープ（chirp）信号、ランダムまたは疑似ランダムノイズ信号、またはインパルスのような、簡単な人工的な試験信号を用いることが可能である。そして出力信号は、例えば、高速フーリエ変換（FFT）または何らかの他のスペクトル分析技術を使用して分析することができる。リニアシステムの動作の特性を表すには、1つまたはそれ以上のかかる簡単な試験信号で十分である。

他方、最新の電気通信システムは非線形性の、そして/または時間的に変動する多数の構成要素を備えている。例えば、最新の電話システムの一部を形成する最新の低ビット伝送速度デジタル音声符号機は非線形的応答を有し、自動利得制御（AGC）、音声活性度検出器（VAD）および結合された音声スイッチそしてバーストエラーは、それらが一部を形成する電気通信システムの時間的偏差の一因となる。従って、電気通信装置の歪みまたは許容性についての客観的測定を行うために、リニアシステムのために開発された簡単な試験方法を使用することは益々不可能となっている。

他方、人間である聴者を利用する主観的試験は費用がかかり時間が浪費され実行が困難であり首尾一貫しない。しかしながら、これらの問題にもかかわらず、システムの性能または歪みの客観的測定と、システムの利用者である人間の主観的反応との間の相関関係が低い場合、結局かかる主観的試験が電気通信装置の試験の最良の方法となっている。

最近、オーディオ技術学会によって、AES前刷りに予稿3070（L-8）として出版された“音響装置の品質の測定” John G. Beerends および Jan A. Stemerdink, 90th AES Convention, 1991 February 19-22, Paris, の論文において、テスト信号として実際の記録された音声のデータベースを使用し、符号機の対応する出力を、人間の耳の中で起こっていると教えられている処理の状況に対応するように設計された知覚分析方法を用いて分析する、デジタル移動無線のための音声符号機の品質測定が提案された。

また、（例えば、“低ビット伝送速度音声符号化の音声品質を評価するための客観的測定方法” Irii, Kurashima, Kitawaki, および Ito, NTTレビュー, Vol 3, No. 5 September 1991）、人工的音声信号（即ち、スペクトル感度が人間の声に類似するが、何らの情報も伝達しない信号）を、ケプストラル間隔（CD）測定のような伝統的な歪み分析測定と組み合わせて使用することが、電気通信装置の測定のために提案された。

人間の音声を符号化するために設計された符号機のような装置を試験する場合、そして人間の耳に基づいた分析方法を採用する場合、上記論文において Beerends および Stemerdink により提案されたように、実際の人間の音声を使用することが明らかにわかりきったことであろう。実際には、しかしながら、かかる試験システムの性能は特別に良いものではない。

従って、本願発明の目的は改善された電気通信試験装置とその方法を提供することである。本願発明の他の目的は、主観的な人間のシステムの動作による知覚に合った、電気通信システムの動作の測定を提供し得る電気通信試験装置を提供することである。

本願発明は、人間の音声に類似するスペクトルを有するが、1人より多くの話者には相当せず、同じ長さの了解極限よりも多くの音声内容を有する試験信号を供給するための信号発生器（7）、および試験により電気通信装置（1）によって歪みを受けた場合に、前記

10

20

30

40

50

信号の歪みが典型的な人間の聴者に知覚されるであろう程度を示す歪み知覚評価値を発生させるために前記歪み信号を分析するために、前記試験信号に対応する歪み信号を受信する、分析手段(8)を含む電気通信試験装置を提供する。

本発明の他の特性および好ましい実施例は以下の記述と特許請求の範囲により明白になるであろう。

本発明は、添付の図面を単に例示として参照することとして描かれ、ここで：

図1は発明の実施例に用いられている配置を示すブロックダイアグラムであり；

図2は発明の実施例の構成要素のより詳細を示すブロックダイアグラムであり；

図3は図2の実施例の一部を形成する試験信号発生器のより詳細を示すブロックダイアグラムであり；

10

図4は試験信号の時間に関する構造を概略的に示し；

図5aは、マスキングノイズの異なるレベルに対する、臨界帯域率(Bark)単位によるピッチ(例えば、概略対数表示周波数)軸に対するマスクされたノイズのレベル(dBS)のグラフであり；

図5bは、7つの与えられた周波数におけるマスキングノイズに関する、臨界帯域率(Bark)単位によるピッチ(概略対数表示周波数)軸に対する励起閾値の変化を示す図形であり；

図6は図2の実施例の一部を形成する分析構成部のより詳細を示すブロックダイアグラムであり；

図7aおよび図7bは、図6の実施例における分析構成部の動作を概略的に示す工程図であり；

20

図8aは、この励起振幅の実施例において、予め定められた音声に似た信号によって人間の耳の中に形成されるであろう時間とピッチの関数として形成された評価を概略的に示し；

図8bは、2つの間隔を置いたクリック(click)によって形成されるであろう励起を示すプロットに対応し；

図9は、この実施例において知覚される音の大きさに振幅を変換するために周波数に対する重み付けをした値のプロットであり；

図10は、連続する時間セグメントに対する、分析手段によって図7に従って計算された、エラー音量値の典型的なプロットであり；

30

図11は、発明の後続する実施例における図7bの一部の修正部分に対応し；

図12aは、図8aに記載された音声信号の低振幅非線形歪みを表示する、ピッチおよび時間軸に対する歪み振幅の図形であり；そして

図12bは、図12aから導かれ、図10に形状が対応する知覚されたエラー音量のプロットであり；

図13aは、高振幅非線形歪みであるが、図12aに対応し；そして

図13bは、図12bに同様に対応し；

図14aは、MNRU歪みの代用であるが、図12aに対応し；そして

図14bは、時間に対するエラー音量のプロットに対応し；

図15aは、クロスオーバー歪みの代用であるが、図12aに対応し；

40

図15bは、時間に対するエラー音量のプロットに対応し；

図16aは、音声活性度検出器に起因するクリッピング歪みの代用であるが、図12aに対応し；

図16bは、時間に対するエラー音量のプロットに対応する。

装置の概要

第1図を参照して、電気通信装置1は入力端子2および出力端子3を具備する。試験装置4は試験において電気通信装置の入力端子2と結合するための出力端子5、および試験において電気通信装置の出力端子3と結合するための入力端子6を具備する。

第2図を参照して、試験装置4は音声に類似する試験信号を供給するために出力端子5に結合された試験信号発生器7、および電気通信装置1から受け取った信号を分析するた

50

に入力端子 6 に結合された信号分析ユニット 8 を具備する。下記においてより詳細に論述されるように、分析器 8 はまた試験信号発生器 7 が発生した試験信号の分析を利用し、このことはこの実施例において出力端子 5 から入力端子 6 に通じる経路 9 によって示される。

また、分析ユニット 8 から測定信号出力端子 10 が提供され、ここで電気通信装置の受容性についての何らかの測定（例えば、歪み）を示す信号がその後の処理のためか、または図示されていない可視表示ユニット（VDU）に表示するために供給される。

実施例 1

音声信号の発生

その最も簡単な形態は、人工的音声発生器が、単に再生可能な音声信号から記録されたデジタルデータを有するデジタル記録器 71（例えば、ハードディスクまたはデジタルオーディオテープ）を具備する事であろう。記録されたデータは個々にデジタル化された音声サンプルであり、これらは連続して記録器 71 から出力端子 5 に結合された信号再変換手段 72（例えば、デジタル - アナログ変換器（DAC））に供給される。記録器 71 に蓄積されたサンプルデータは 1 またはそれ以上の長さが数秒間継続する（例えば、およそ 10 秒程度）音声発生を含む。

代わりに、記録器 71 は、音声データを、例えば LPC 音声シンセサイザーを駆動するためのフィルター係数の形式で、または再生手段を含む音素シンセサイザーを駆動するために高レベルデータ（例えば、音素、ピッチおよび強度データ）の形式で記録することが可能である。

制御回路 73（例えば、マイクロプロセッサ）は出力である特別の試験信号を選択するために記録ユニット 71 の動作を制御する。

図 4 を参照して、記録器 71 に記録された試験信号データは複数のセグメント $t_0, t_1, t_2 \dots t_n$ を含む試験信号を形成するために再生される。

セグメント $t_0 - t_n$ の各々は典型的に異なった音声音（例えば、異なった音素）または無音に対応する。1 つの知られている人工的音声試験信号は CCITT の勧告 P50（Recommendation on Artificial Voices, Vol. Rec P50, Melbourne 1988, published by CCITT）に開示されている。P50 試験信号において、各セグメントは 60 ms 継続する。

セグメントは、それぞれが不規則に選択された 16 の予め定められたスペクトルパターンの連続を有するパターンに分類され、

$$S_i(f) = \frac{1}{12} \sum_{j=1}^{12} A_{ij} + 2 \sum_{j=1}^{12} A_{ij} [\cos(2\pi i f)] \quad , i=1, 2, \dots, 16$$

に等しい音圧スペクトル $S_i(f)$ に関する勧告によって定義される。

各パターンにおける異なるセグメント間の伝送は滑らかであるように調整される。パターンに関し、13 が有声音に対応し、残りの 3 が無声音に対応する。連続する音声は記録媒体に記録され再生されるか、または、例えば、上記参照の Irii の論文に記載されているようにボコーダを用いて記録データから発生させられるかのいずれかである。

P50 信号は、約 10 秒間の平均を取るとき、長期間および短期間の音声へのスペクトル類似性を有する。従って、むしろ、図 4 に示す音声の連続は少なくともこの長さで持続する。

歪み

試験において電気通信装置 1 から出る信号は入力端子 2 に供給される試験信号とは相違する。第 1 に、信号の時不変性線形歪みがあり、その結果全体にわたって振幅の変化をもたらし、そしてそのスペクトル形状を変化させるように信号を濾波する。第 2 に、信号に一

10

20

30

40

50

定のノイズ源（熱雑音のような）および不連続ノイズ源（ノイズバースト、ダイヤルノイズ、干渉性スパイク、回線の交差のような）を含む種々のノイズ源からノイズが追加されるであろう。第3に、コーデックのような非線形素子およびエコーキャンセラおよび閾値回路のような時変素子に起因する信号の非線形性および時変歪みが存在するであろう。非線形歪みの存在はノイズと信号間の相互変調歪みの原因となり得、そして出力端子3における歪みはそれ故に信号および装置1ばかりでなくノイズにもまた依存する。さらに、時変歪みの存在は信号の何らかの所定の一部分の時間に歪みが生ずることを意味し、信号およびノイズの先行する特定の時間の一部分に依存する。例えば、音素の開始の前に、もし高いレベルのノイズが存在する場合、音声活性度検出器は少しも音素を抽出できないかも知れず、ところが、もし音素の前が無音の場合には、音声活性度検出器は実質的歪みを発生させる音素の開始を十分に抽出するであろう。

10

分析器8

本願発明に係る分析は、現在解明されているような人間の耳の反応に類似するような、試験信号の歪みに依存する許容信号出力を提供することを目的とする。

これらの現象を生じさせる物理的または生物学的機構を考慮するまでもなく、音に対する人間の知覚は幾つかの要因の影響を受けることが良く知られている。第1は、1つの音の存在が同様のスペクトル（周波数）領域の他の音を“マスク”（即ち、知覚を抑制する）することである。他の音がマスクされる程度は、第1にピッチが第1の音にいかに接近しているか、第2に第1の音の振幅に依存する。

このように、音に関する人間のエラーまたは歪みの知覚は音それ自身に依存し、音自身としての同じスペクトル領域での低振幅のエラーはマスクされ得るそしてその結果聞き取れない（例えば、サブバンドの符号化において量子化エラーが生ずるように）。

20

第2に、マスキング現象はある時間依存性を有する。音はその音が消えた後短い時間他の音をマスクし続け、マスクされるであろう後の音の振幅は最初の音が消えた後急速に減衰する。このように、エラーまたは歪みは、現在の信号によるばかりでなく知覚された信号の部分によってもまたマスクされるであろう（程度は小さい）。これは“順向性マスキング”と言われている。別の場合には聞くことが可能であるような低レベルの音のすぐ後に高レベルの音を適用すると最初の音が聞き取れないようにマスクすることがまた見つけられている。これは“逆向性マスキング”と言われている。

第3は、人間の耳は周波数に直接には反応しないが、周波数軸の非線形的ゆがみに対応する、音の“ピッチ”として知覚された現象に反応する。

30

第4は、信号がマスクされない時でさえ、人間の耳は振幅に直接に反応するのではなく、振幅に対し非線形関数である音の大きさとして知覚された現象に反応することである。

従って、この実施例において、分析器8は電気通信装置1から受けた信号を処理し、上記既知の人間の耳の性質に一致させて、試験信号に形成された歪みが人間の聞き手にいかに意味があるかまたは不快であるかを決定するために調整される。特に、分析器ユニット8は試験信号発生器7によって発生した試験信号に人間の耳がどう反応するかを決定するために調整され、それから電気通信装置の出力3からの信号を、歪みが知覚できる程度を決定することによって、元の試験信号から知覚として相違する程度を決定するために、同様の処理をするために調整される。

40

図5aは、固定された周波数において狭い帯域のノイズに対するスペクトルのマスキングの閾値（それより上では第2の音が第1の音によって覆い隠される閾値）の変化を概略的に示す。5つの曲線は累進的に高くなるマスキングノイズのレベルに関するものであり、マスキングノイズのレベルが増加するとその効果は、マスキングノイズ周波数においてマスキング閾値をおおよそ直線的に増加させる原因となっており、またノイズ周波数から離れたところでは閾値の形を変化させている（高い周波数の方が大きく）ことが分かるであろう。マスキング効果はそれ故にマスキングノイズの振幅に関して振幅は非線形性である。

与えられたマスキングノイズレベルについて、マスクされたスペクトル帯域の幅（例えば、マスキング周波数の中央から3dB下の点で測定された）はマスキングノイズの周波数

50

により変化する。このマスクされた帯域の幅の変化は人間の聴覚の周波数識別力によるフィルターの形状の特性に関係し、それ故にピッチの人間の知覚に係る。

従って、図5bに示すように、ピッチの尺度は、周波数よりもむしろ、マスクング帯域の幅が一定である新しい尺度を創造するように周波数尺度を歪ませた周波数尺度から生成させる。図5bは、-3dBの点で横切る異なる周波数において音声をマスクングする狭い幅の集合を考慮することにより導かれる、臨界帯域レートまたはバルク(Bark)尺度を示す。この尺度は、例えば、“オーディオ工学と音響心理学：最終的受話器、人間の聴覚システムに信号を調和させる”、J. Audio Eng. Soc. Vol. 39, March 1991, Zwicker and Zwickerに記載されている。なお成分信号の帯域幅はそれらが各帯域の中央の信号に等しいマスクング振幅に相当するよう

10

に選択することができる。
図5bの臨界帯域は線形的周波数尺度表示された場合、500ヘルツ以下で形状(周波数軸において)が類似する。500ヘルツ以上では、それらは対数的周波数尺度で見た場合形状が類似する。典型的な電話方式の帯域幅は300から3150ヘルツであり、電気通信装置は多くの場合これらの限界の間に限定された帯域であり、ピッチ尺度への変換はこの実施例においては正確性について単に少し妥協することにより500ヘルツ以下の線形的領域は無視される。

図6を参照して、分析ユニット8は入力端子6から信号を受信し対応するデジタル信号パルス列を生成するよう調整されたアナログ-デジタル変換器(ADC)81、ADC81のデジタル出力を受け取るために結合された計算プロセッサ82(例えば、インテ

20

ル社の80486プロセッサのようなマイクロプロセッサ、またはウエスタンエレクトリック社のDSP32C、またはテキサスインスツルメント社のTMS C30装置のようなデジタル信号処理装置)、プロセッサ82のための命令シーケンスを記録し計算結果を記録するためのワーキングメモリを提供する記録装置83、出力10に接続されたプロセッサ82からの出力回線84を含む。

図7を参照して、この実施例においてプロセッサ82により実行される工程について記述する。

第1に、ステップ100において試験信号発生器7から供給される試験信号が、電気通信装置1を通過することなく入力端子6に直接入力される。
次のステップ101において、ADC81からの信号が耳の外側の部分と内耳の間の変換

30

関数に相当するフィルターによって濾波される。濾波は記録器83に記録されている濾波のデータに従ってデジタル濾波動作を実行するによって典型的に達成することができる。フィルターは“音響システムにおいてエラー(error)を評価するための音響心理学モデル”J. R. Stuart, Procs. IOA, vol. 13, part 7, 1991に記載された形の伝達関数によって特徴付けることができる。
実際、内耳への伝達関数は音が耳に密着して(例えば、ヘッドホンを通して)、またはより離れて(例えば、拡声器から)接続されているかどうかに従ってわずかに変化するであろう。それ故に、プロセッサ82および記録器83は、試験される電気通信装置1の型式に係る異なった音声配置に対応して幾つかの異なった伝達関数の特性を記録し、電気通信装置の型式を条件として指定する使用者の入力に応じて適切なフィルターを選択す

40

るために調整される。ステップ101の実行の後、濾波された信号は内耳において受けとられたであろう信号に相当すると言える。
次に、ステップ102において、信号は周波数からピッチへの転換をもたらすように、周波数に関し対数的に変化する帯域幅を有する複数のスペクトル帯域に分割される。この実施例では、国際標準ISO 532Bに従って、信号はそれぞれが1オクターブの1/3の帯域幅の100ヘルツから8キロヘルツの20の帯域に帯域濾波される。ISO帯域フィルターは対数的周波数軸で見た場合には形状が類似し、良く知られかつ文書に記載されている。各20の帯域の平均の信号振幅は各々4ミリ秒算定され、濾波された後の信号は、それぞれが20の周波数帯域振幅値を有する連続する時間セグメントを有する。この帯域濾波は試験信号のすべの音価について実行される(これは数秒程度、例えば10秒継

50

続する)。

各フィルター帯域の中で比較的幅の広いフィルターがマスキングについて考慮され、幅の広い、部分的に重なり合ったフィルターの縁が隣接する周波数に起因するスペクトルのマスキングを確実にすることが考慮される。

次に、ステップ103で、国際標準ISO 226に規定された周波数依存聴覚閾値が帯域出力のそれぞれに適用される。これは図5aに示された最小聴力閾値の効果をシミュレートする。

次に、ステップ104において、帯域通過信号の振幅はホン(phon)または人間の聴覚システムにより知覚されるラウドネスと等価である感覚レベルに変換される。変換は非線形であり、信号の振幅と周波数の双方に依存する。従って、変換を行うため、国際標準ISO 226に規定された等しいラウドネス変調曲線が帯域出力のそれぞれに適用される。これら等しいラウドネス変調曲線およびステップ103で使用された閾値の両者は記録器83に記録される。

10

次に、ステップ105において、十分大きな振幅値の後に指数関数的減衰を用意することによって時間的マスキング、特に順向性マスキング、が実行される。実際には、マスキング効果の減衰の度合いはマスキング音の適用時間に依存し、適用時間が長ければ短時間の場合より減衰時間は長い。しかしながら、この実施例では、

$$y = 56.5 * 10^{(-0.01x)}$$

(ここで、yはレベルを意味し、xは時間を表す)により決定される固定された指数関数的に重み付けされた減衰を適用することで十分であることが分かった。これは実際に生ずる最大減衰(200ミリ秒の間隔に相当する)と最小減衰(5ミリ秒の間隔に相当する)の間で低下する。

20

順向性マスキングを適用して、各帯域通過濾波振幅に関する各時間セグメントにおいて、次ぎに続く3つの時間セグメントの対応する帯域通過のマスキング値が、上記指数的減衰を用いて計算される。3つの値はこれらの帯域の実際の振幅と比較され、実際の振幅より高い場合には実際の振幅の代わりに置き換えられる。

上記に注目して、音に関し、前に生じた音をマスクすること(“逆向性マスキング”と言われてもまた可能である。好ましくは、この実施例において、順向性マスキング工程は逆向性マスキングを形成するために複製され、定数の数値は異なるが同じ形の指数的減衰が用いられる(換言すれば、各時間セグメントに関して、前に生じた時間セグメントに関するマスキングの値が計算され、これらの帯域において実際の振幅より高い場合には、実際の振幅の代わりに置き換えられる)。

30

このように、ステップ105の後、計算された信号データは、それぞれ20の帯域通過信号の振幅を有する連続する時間セグメントのデータ、ある振幅は零とするための閾値、および与えられた時間セグメントにおける与えられた帯域の、順向性および逆向性マスキング工程に起因する過去と未来の時間セグメントの帯域の相当する振幅に影響される振幅を含む。

これは信号のピッチと時間の軸に関する面状表示の、試験信号が電気通信装置1を通さず直接に人間の耳に適用された場合のマスキング効果に相当する。

図8を参照して、図8aおよび図8bは上記工程によって発生させた励起面を示す。図8aは無声音が続く有声音を含む音声の発生に対応し、第1の音と、第2の音の広帯域特性のフォルマント構造をただちに識別できる。図8bは2つのカチカチという音に対応する面を示し、図7の順向性マスキング工程105の、指数的に減衰する効果が明白に観察できる。

40

次に、ステップ106において、試験信号発生器7は試験信号を再生するが、今度はそれは電気通信装置1の入力端子2に供給され、出力端子3は試験装置4の入力端子6に接続される。計算工程101-105が電気通信装置1から受信した信号に対する対応する面を計算するためにここで繰り返される。

元の試験信号および電気通信装置からの出力(歪んだ試験信号)の耳(励起)における効果が計算され、2つが耳を刺激する大きさの差は、人間の聴覚システムによって知覚され

50

るような試験信号の歪みのレベルに相当する。従って、対応する帯域通過振幅間の比をとることにより、各セグメントに対する電気通信装置の振幅変換機能が計算される（または、図8 aまたは図8 bに示すように、dBで振幅間の差をとることにより帯域通過振幅はdB尺度で表示される）。

工程107において、変換関数の中で電気通信装置によって形成された知覚歪みには無関係な全ての利得期間を無効化するために、各帯域通過期間は試験信号シーケンスにおける全帯域、全時間セグメントの帯域通過フィルターの出力の平均振幅によって割算することにより（dB表示の場合は減算）規格化される。

もし最初の試験信号および電気通信装置1の出力が等しく、全体的レベルに差がなかったならば（換言すれば、もし電気通信設備1が歪みを発生させないならば）、2つの信号の各帯域通過フィルターの出力間の比は一定であり、そして振幅におけるdBによる対数の差は零である。それ故に、図8 aまたは図8 bに対応する差のプロットは全ての時間および全てのピッチ帯域において完全に平坦となるであろう。何らかの偏差は電気通信装置1の歪みに起因する。歪みの無い平均レベルに対して、付加的な歪みエラーはピークとして表れるであろうし、単一の欠損は凹部として表れるであろう。

これらのエラーによって与えられる知覚的重要性はそれらの振幅に直接には依存せず、むしろ振幅の非線形関数であり周波数の関数であるラウドネスに依存する。知覚されるラウドネスの計算は国際標準ISO 532 Bに与えられている。しかしながら、この規格は両耳の音および通常の音（電気通信応用装置において一般に生ずる）に適用され、確立された片耳の電話方式の知覚のラウドネスに関するCCITT勧告P79（Blue Book Volume V, Melbourne 1988, CCITT）により与えられた重み付けに基づき、ラウドネスのより簡単な計算に使用することが可能である。このエラーラウドネス評価方法は、ある周波数でのエラーはより容易に知覚されるという事実を考慮し、そしてこのことからこの周波数は他の周波数のよりも大きな重み付けを与えられる。信号シーケンスにおける各時間セグメントについて、この実施例では、エラーの大きさを次のように計算され：

14

$$Err Loud_t = 0.8 \sum_{n=1} E r_n * 10^{(-0.0175 * W_{s_n})}$$

ここで：

$Err Loud_t$ 時間tでのエラーラウドネス（+veおよび-ve部分は分離して計算される）

n 200 Hzから4 kHzの1/3オクターブ帯域のn番目

$E R_n$ dBによるエラー振幅

W_{s_n} n番目の周波数のSLR重み付け

であり200 Hzから4 kHz間にわたるエラーの狭い帯域モデルに対するものであり、ここでP79勧告から導かれる重み付け係数は図9に示される。

全20帯域通過出力を利用する広帯域電話方式モデルにおいて、対応するエラーラウドネスは次のような計算され、

21

$$Err Loud_t = 1.28 \sum_{n=1} E r_n * 10^{(-0.0175 * W_{s_n})}$$

この場合、nの値は100 Hzから8 kHz間の全ての20帯域をカバーする。

正および負の小計を求めるため、付加的エラー（正エラー）および欠損的エラー（負のエラー値）は分離して累加される。

図10に示すように、計算工程109の結果は時間セグメントのエラーラウドネス値の時

10

20

30

40

50

間的経過である。この実施例ではステップ110において、電気通信装置の受容性その他が、図9に示されたデータにより、例えばエラーラウドネスのピーク値そして/またはエラーラウドネスの平均値を得ることにより、直接に知りうる。次ぎにステップ111で、これらの特徴の一方または両方が電気通信装置1の歪みの測定として出力端子9に出力される。

実施例2

実施例2において、分析ユニット8は実施例1におけるものと同じか同様のものである。しかしながら、試験信号発生ユニット7はP50試験信号を使用せず、その代わり人工的で音声に似た試験信号の異なったタイプを発生する。

P50試験信号は多くの目的に適用可能であるにもかかわらず、摩擦音の領域が全て欠けていると認められる。さらに、それはどちらかという規則的であり単調な長時間の構成であり、その音はむしろ、母音-子音-母音-子音...の連続のようである。しかしながら、上記で論じたように、多くの電気通信システムは、自動利得制御器や音声スイッチのような時間に独立の構成要素を含むので、与えられた試験信号の部分に適用される歪みは試験信号の先行する部分に部分的に従属する。換言すれば、全体としての信号の時間的継続の中での音声信号の部分の前後関係に。

従って、本実施例では、小さな、典型的な、音声セグメントの部分集合(10の既知の音素より選択された)が使用され、試験信号は異なった前後関係のシーケンスで集められたこれらの音から構成される。歪みは測定されつつあるから、試験シーケンスは、比較的小互いに似ていない音か、またはより一般的には、一方が他方に続いたとき歪みの原因となるように比較的似ている音の連続を含むべきであることが重要である。本実施例のより簡単な形において、試験信号は、高い、低い、またはゼロのレベルから選択された調整部分が前置された選択されたセグメントのそれぞれを含み得、それによって試験信号は各典型的な音声セグメント(音素)が異なったレベルの前置音に続いて試験されることが可能となる。前置信号の長さは試験されるシステムの時定数よりも長くなるように選択され、例えば、コーデック適合器および能動的利得制御器は数秒の桁数であり、ところが音声変換器の過渡応答は数ミリ秒の桁数である。

さらに本実施例の詳細は我々の先行出願である英国特許出願No. 9312758.7(代理人参照:A24613), 1993年6月21日出願の名称“音声同様の試験の励振”に見いだすことができ、その内容は全てを参照することによりここに組み込まれている。本実施例の試験信号はまた従来の分析手段において利用し得る。

実施例3

本発明の実施例3において試験信号発生器7は実施例1および2と同じ方法で動作する。しかしながら、分析ユニット8の動作はステップ102から110において相違する。たとえば、実施例1の対数的に区分されたフィルターが人間の耳のピッチの尺度に妥当な近似であると分かっても、同等のより良い性能がバルク尺度上(前述したように)で均等に置かれたフィルターの使用によって与えられることが見出される。従って、ステップ102において、20の帯域通過フィルターが、ピッチ尺度上で1バルク間隔で離れて置かれた曲線的、指数関数的(roex)フィルターである。曲線的指数関数は“聴覚-フィルター帯域幅および励起パターンの計算に関する規格の提案”(J. Acoust. Soc. Am. 74, 750-753 1983), B. C. J. MooreおよびM. R. Glasburgに記載されている。

各帯域4ミリ秒毎の平均信号振幅を計算するよりもむしろ、この実施例においては、介在帯域に関する介在平均時間を含む、最も高いピッチ帯域に関する2ミリ秒および最も低い帯域に関する48ミリ秒にわたって平均し、異なった帯域に関する異なった平均時間にわたって信号振幅が計算される。低周波数においてより長い間隔にわたって分析するように、ピッチ(または、一般的には、周波数)に従って時間的歪みを変化させることは本質的に作業を増加することが分かる。

その後の工程において、前のように、各2ミリ秒の時間セグメントに関し、帯域通過フィルターの出力値の配列を発生させる。最も高いピッチより低い帯域に関して、介在時間セ

10

20

30

40

50

グメントのために評価は1回以上繰り返される(例えば、最も低いピッチ帯域に関し、各48ミリ秒の平均振幅値間の2ミリ秒の時間セグメントに関し各評価は24回繰り返される)。勿論、単にそれらを繰り返すよりもむしろ、次ぎの値との間で数値補間法を実行することは可能であろう。

ステップ103 - 106は、実施例1と同じである(異なったフィルター応答を反映するために、数値定数の調整を伴う)。

この実施例においては、歪みのラウドネスを計算するよりもむしろ、主観的“聴音作用”の基準 Y_{LE} により密接に関係する異なった試験基準が導かれる。

帯域通過可聴励起評価の組のシーケンス(時間軸とピッチ軸に応じた面に相当する)は長さ96ミリ秒(即ち、48の連続する2ミリ秒のセグメント)の隣接する領域に、最小のピッチ帯域にとって少なくとも2つの異なった評価を示すように、分割される。エラーまたはエラー活性度の合計値は

エラー活性度、

$$48 \quad 20$$

$$E_A = 10 \log \sum_{i=1} \sum_{j=1} |c(i, j)|$$

として計算され:

ここで $c(i, j)$ は分析されたエラー面区域の第 i 時間セグメントおよび第 j ピッチ帯域におけるエラー評価値である。

これは存在する歪みの完全な合計の表示を与える。

次に、時間およびピッチに関するエラーの分布(またはむしろ歪みのエントロピー、それはエネルギーが分布する広がり相互関係に相当する)が次の式で計算され:

エラーエントロピー、

$$E_E = - \sum_{i=1} \sum_{j=1} a(i, j) * \ln(a(i, j))$$

ここで、

$$a(i, j) = \frac{|c(i, j)|}{E_A}$$

上記表現における対数項はエネルギーの影響の範囲であるエントロピー E_E を統制し、非線形圧縮関数として作用する。

エラー活性度およびエラーエントロピー基準は共に、ピッチと時間にわたって分布しているよりもむしろ短い時間間に単一のピッチに集中した場合に、聴者はエラーを高いレベルで極めてより顕著に発見するように、歪みの主観的に知覚されたレベルに良く一致することが分かる。従って、本実施例において、図12に示すように、図7bのステップ109でラウドネスを計算するよりもむしろ、歪みの合計と分布(活性度とエントロピー)を計算するステップ119が実行される。

ステップ110において、2つの尺度が閾値として単独に提示されるか、またはそれらは結合されそして結合された尺度が閾値とされる。例えば、それらは後のステップにおいて適切な重み付けと共に合計されまたは掛け合わされるであろう。

実施例4

この実施例において、実施例1あるいは実施例2のいずれかに従って音声信号を発生することが可能である。しかしながら、分析ユニット8は上記マスキングの計算を実行するよりもむしろ、例えば“基底膜のデジタルフィルターシュミレーション”、電算機の音声

10

20

30

40

50

および言語、No. 3 1989, Ambika ir a j h, B l a c k, 及び L i n g g a r d (参照により総てここに組み込まれている) に記載されているように直接に人間の耳をシュミレートする。かかる、モデルは A D C 8 1 から入力として信号を受けとり、電気通信装置 1 からの歪んだ信号についての人間の聴取り構造部分に関する効果に相当する各時間セグメントについての連続する出力を発生する。モデルの出力は次に、信号における歪みの知覚の重要性を示す信号を供給するための実際の聞き手の反応を有する経験的に導かれた相互関係に基づく、適切な工程および判断論理 (例えば、ニューラルネットワークまたはファジー論理制御器) により結合される。

この実施例の分析方法の特徴は、また他の試験信号 (例えば、実際の人間の音声) と共に使用し得る事である。

10

発明の効果

図 1 2 から 1 6 を参照すると、発明の実施例 1 および実施例 2 による図 8 a の試験信号の電気通信装置の歪みの種々の型の表現がここに描れている。

図 1 2 a は、信号に低振幅の 2 次および 3 次の次数の項目を追加することによる瞬間的振幅歪みにより作成されたエラー励起面を示す。歪みは人間の聞き手による “ かるうじて聴取できる ” 特徴を示した。図 1 2 b に示すようにエラーラウドネス特性は小さくそして大部分が正であるように見える。

図 1 3 a は、同種の完全に聞き取れる非線形歪みに関するものであるが、高い 2 次および 3 次の次数の項目を有する対応するエラー振幅面を示す。エラーおよびエラーラウドネス (図 1 3 b) は両方ともより大きい。その上、歪みラウドネスの大部分は、これがその高調波が十分知覚される低周波ホルマント音調を含むことから、図 8 a の試験信号の有声音の部分に一致することが分かるであろう。

20

図 1 4 a および図 1 4 b を参照すると、変調ノイズ照合ユニット (M N R U) 歪みの効果が示されている。M N R U 歪みは C C I T T 勧告 P 8 1 の別紙 A に記載されており、それは信号の A l a w P C M s t a g e (電気通信システムにおいて広く使用されている種類のものである) により導入された歪みに関し、理論的に等価に設計される。歪みのレベルが人間の聴者に完全に聞き取れるものとして計算された。重ねて、図 1 4 a から知覚される歪みは試験信号の有声音の中のホルマントと主に結合することが分かるであろう。

図 1 5 a および図 1 5 b を参照すると、クロスオーバー歪みが与えられた場合 (即ち、 x が 0 より大の場合 $y = m x + c$ そして x が 0 より小の場合 $y = m x - c$)、低振幅信号は伝送されず、そして試験信号の第 2 の部分のより低いエネルギーの無声音は著しく減衰する。それ故、図 1 5 a および図 1 5 b はこの種類の歪みの非常に重要な主観的衝撃を示唆し、これは人間の聴者の反応に対応する。

30

最後に図 1 6 a および図 1 6 b は 5 0 ミリ秒の立上がり時間を有する音声活性度検出器の効果を明らかにする。信号の最初の部分において、信号が欠損したための大きな負のエラーラウドネスが存在する。それに続く正のエラーラウドネスは行き過ぎ量または整定に起因する。エラーラウドネスの値は知覚される歪みの高いレベルを示し、人間の聴者の反応に一致する。

他の変更または修正

40

本発明の動作原理を変更することなしに、前に記載した実施例の多くの変形が成され得ることが前述より明白であろう。例えば、もし電気通信装置がデジタル入力を受信するために設置された場合、D A C 7 1 を不要にすることができる。出力端子 5 からの信号は電気通信装置の入力端子 2 にデジタル形式で供給され、同様に A D C 8 1 も不要とされるかもしれない。代替案として、出力端子 5 に電気 - 機械変換器を準備することができ、信号は音声の信号として供給される。後者の場合に関し、双方とも基準としてここに組み込まれる、人口の耳および人工の口に関する C C I T T P . 5 1 勧告、V o l u m e 5 , R e c P . 5 1 , M e l b o u r n e 1 9 8 8 , および先行する英国特許出願 G B 2 2 1 8 2 9 9 (8 7 3 0 3 4 6) で論述されているように、試験信号は人工の口を経由して供給することができる。同様に、歪んだ音声信号は、基準としてここに組み込ま

50

れる上記 C C I T T 勧告および我々の先行する英国特許出願 G B 2 2 1 8 3 0 0 (8 7 3 0 3 4 7) において論述されているような、人工の耳の聴覚構造を經由して受信することができる。これはステップ 1 0 1 において必要とするフィルターを減少させるであろう。たとえ、前に記述した実施例において、一時的なマスキングによる単一の減衰形状が記載されているとしても、発明の代わりの実施例において順向性マスキング（および逆向性マスキング）に関する複数の（例えば 2 つ）減衰率を準備し、マスキング音の持続時間に従って必要とする減衰率を選択するのが好ましいであろう（即ち、通過帯域の 1 つにおいてその振幅が予め定められたレベルの限界を超える時間セグメントの数）。例えば、最大および最小減衰（それぞれ 2 0 0 ミリ秒および 5 ミリ秒間隔）は

$$y = 58.4039 * 10^{(-0.0059x)}$$

$$y = 55.5955 * 10^{(-0.0163x)}$$

によって決定することができる。

たとえ、実際の電気通信装置への接続がここに記載されていたとしても、多くのかかる歪みは比較的特徴付けることが容易であるから（例えば、これらは V A D またはコーデックに従う）、電気通信装置によって導入された歪みをシミュレートするために電算機装置をプログラムすることは同等に可能であろう。従って、発明は信号がかかるシミュレーション装置に適用される実施例と同様に拡張され、電気通信装置のシミュレートされた歪んだ信号が処理される。この方法で、多くの複雑なそして非線形通信装置の結合の人間の聴者への受容性が、かかる装置の現場での組み立てまたは接続以前にモデル化されるであろう。

たとえ、分析ユニット 8 および試験信号発生器 7 が分離された機器で記載されていたとしても、実際にはそれらは単一の適切に処理されたデジタルプロセッサにより実現化可能であり、同様に、上記実施例において好ましい電気通信装置シミュレータが同じプロセッサにより準備し得る。

たとえ、前に記載の実施例において、分析ユニット 8 が本文の信号発生器 7 から試験信号を受けとり分析していても、実際には分析ユニット 8 は先行する分析によりその又はそれぞれ各々の試験シーケンスのために以前に導かれた励起データを記録することができる。このように、かかる実施例における分析ユニットはそれ自身、歪みのない試験信号を分析するために配置される必要はない。

前に記載した実施例において、他に、エラーラウドネスより信号歪みを測定し、計算されたデータから図 1 1 a , 1 3 a , 1 4 a , 1 5 a および 1 6 a に相当するエラー活性度またはエラーエントロピーが容易に導かれる。実際、歪みのラウドネスは人間の聴者にとって単にその効果の評価の 1 つであり、他の評価は聴者の疲労および聴取するための努力である。例えば、前に記載した実施例に従って計算された歪みまたはエラーデータは、本物の人間の聴者との比較試験によって経験的に導かれたパラメータに従って動作する統計的分類機、ニューラルネットワーク、またはファジー論理機器への入力として採用することができる。

通常の使用において“音素 (p h o n e m e) ” は音声の前後関係によって変更される 1 つの音を表示するにも拘らず、この文書においては、用語“音素”は便宜的に単一の、繰り返し可能な、人間の音声を示すために使用される。

逆のことが指示されまたは明白でなければ、上記実施例の特徴はここで明瞭に記載された方法以外の方法で結合することができる。

たとえ、前に記載された実施例が電気通信装置の試験に関するものであっても、この発明の新規な概念の試験または分析への応用を排除するものではない。

従って、ここに開示した新規事項または新規事項の組み合わせについて、かかる事項または変形が下記請求項の範囲にあると専門家である読者にとって明白であるそれらの変形と共に保護が要求される。

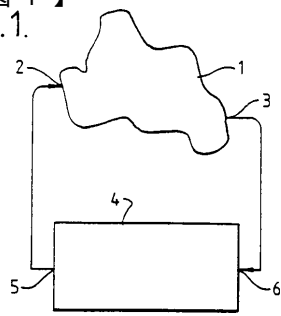
10

20

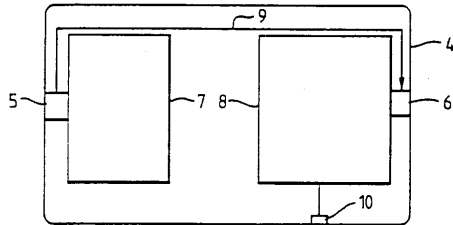
30

40

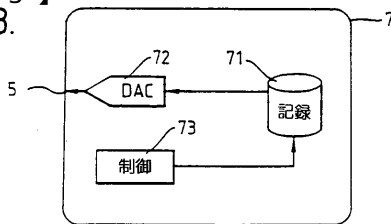
【図1】
Fig.1.



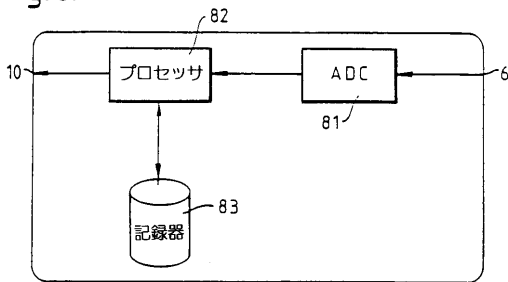
【図2】
Fig.2.



【図3】
Fig.3.



【図6】
Fig.6.



【図7b】

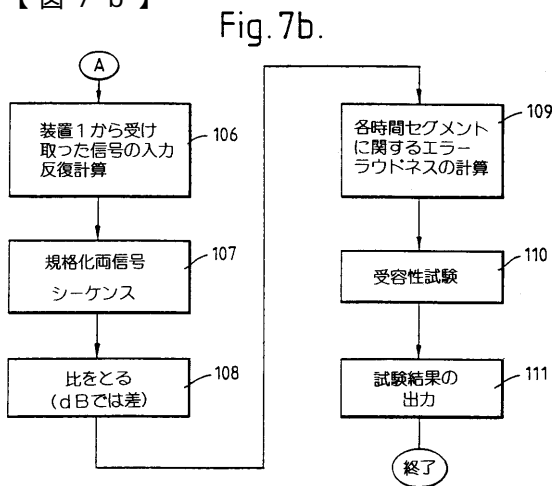
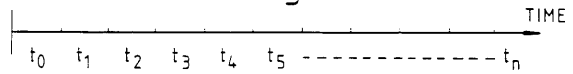


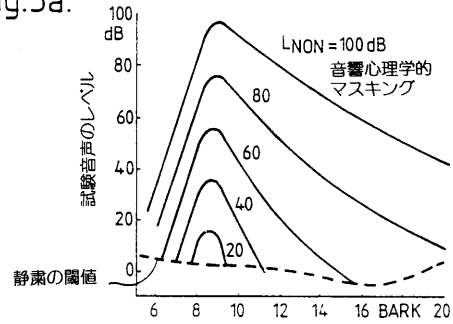
Fig.7b.

【図4】

Fig.4.

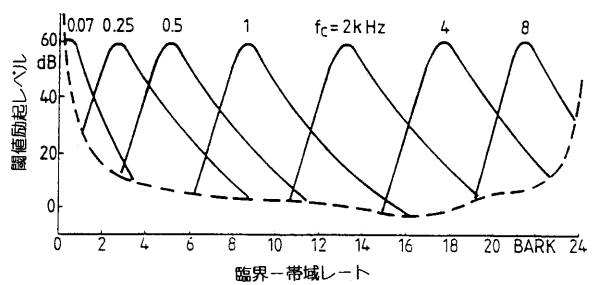


【図5a】
Fig.5a.

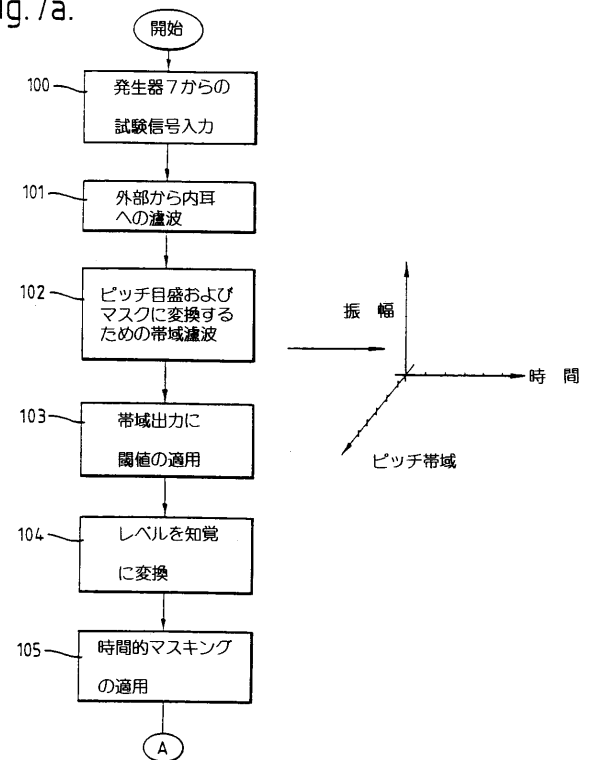


【図5b】

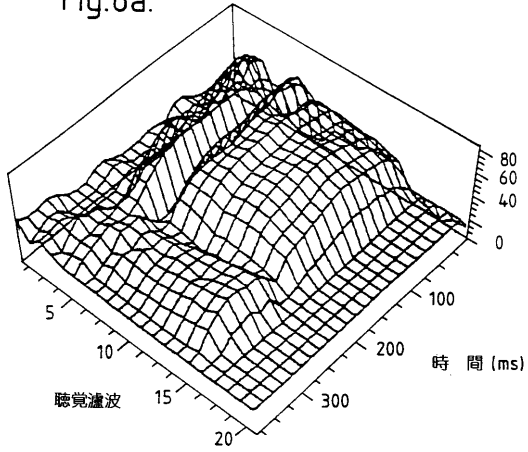
Fig.5b.



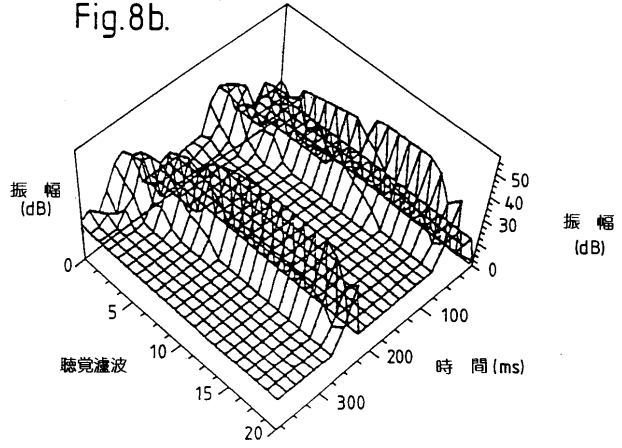
【図7a】
Fig.7a.



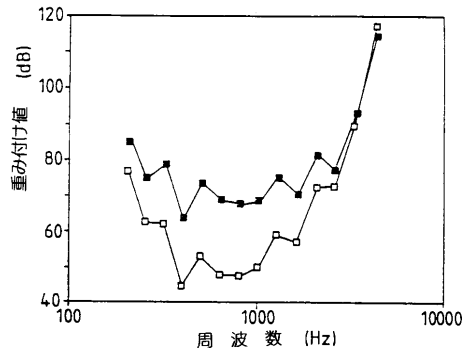
【図 8 a】
Fig.8a.



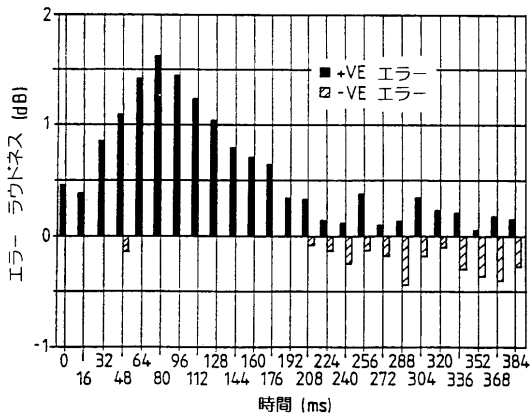
【図 8 b】
Fig.8b.



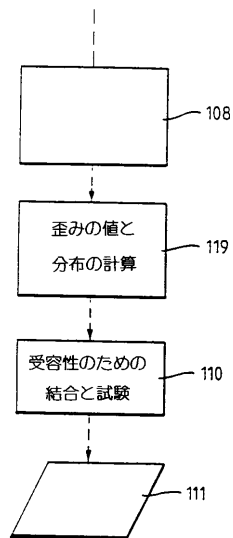
【図 9】
Fig.9.



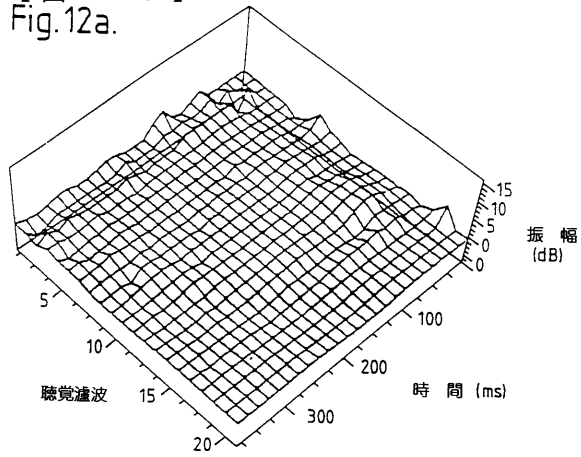
【図 10】
Fig.10.



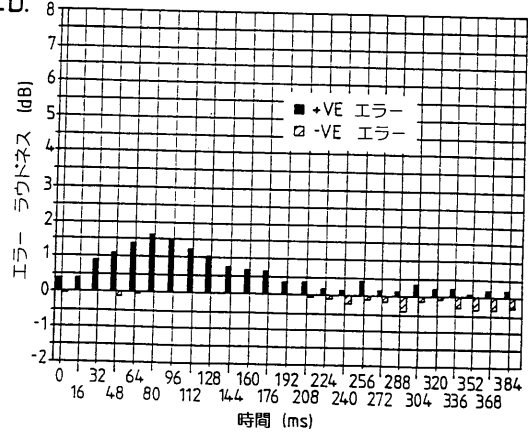
【図 11】
Fig.11.



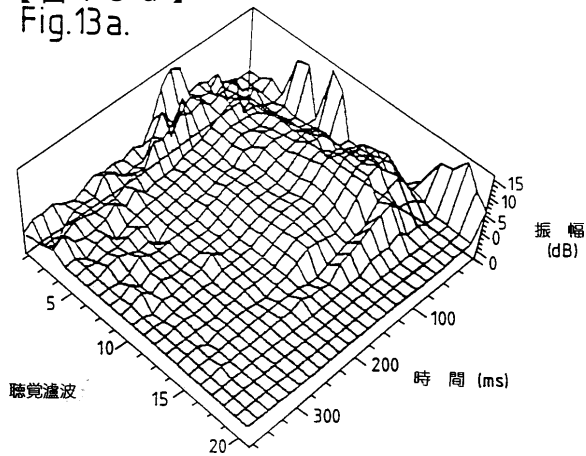
【図 1 2 a】
Fig.12a.



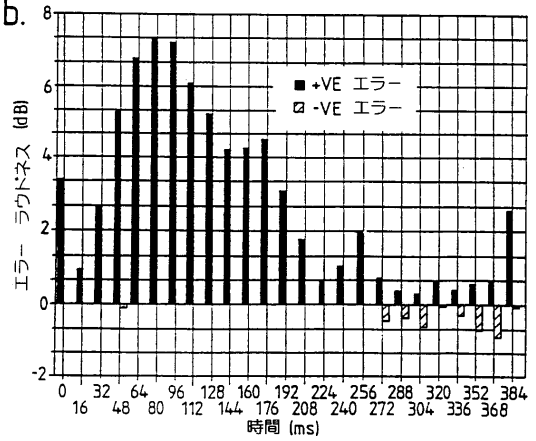
【図 1 2 b】
Fig.12b.



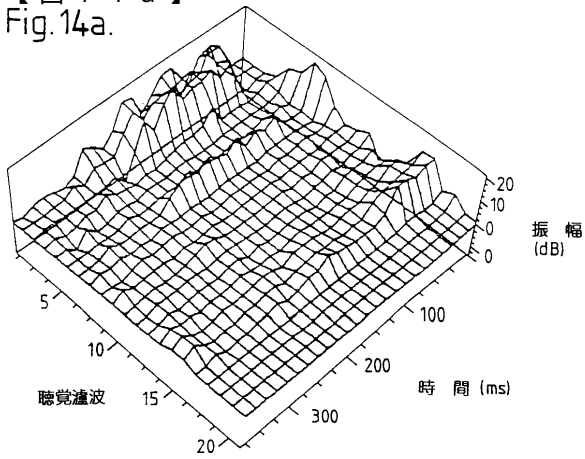
【図 1 3 a】
Fig.13a.



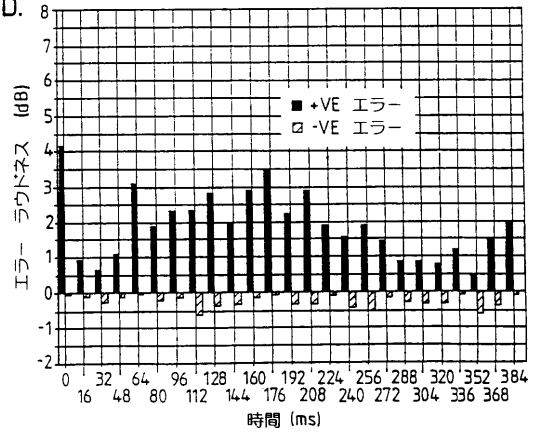
【図 1 3 b】
Fig.13b.



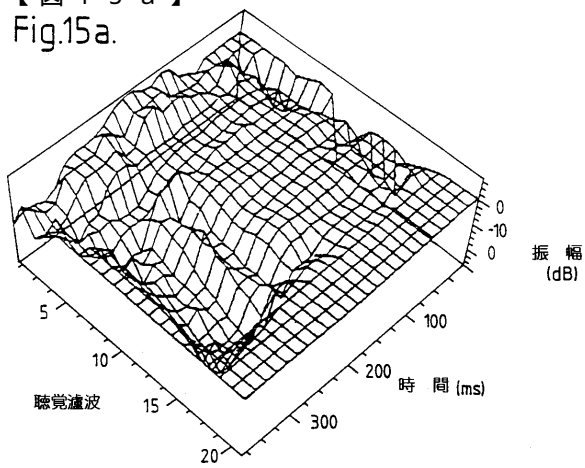
【図14a】
Fig.14a.



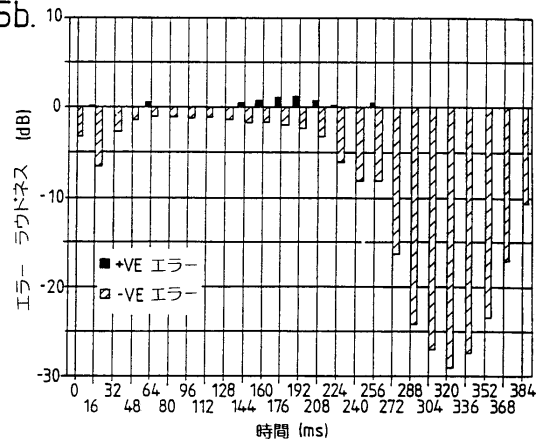
【図14b】
Fig.14b.



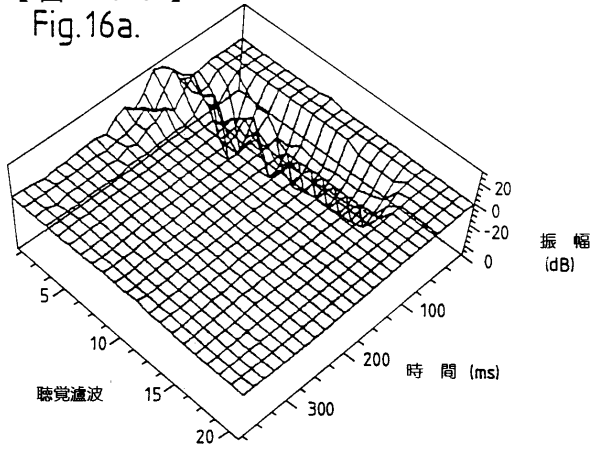
【図15a】
Fig.15a.



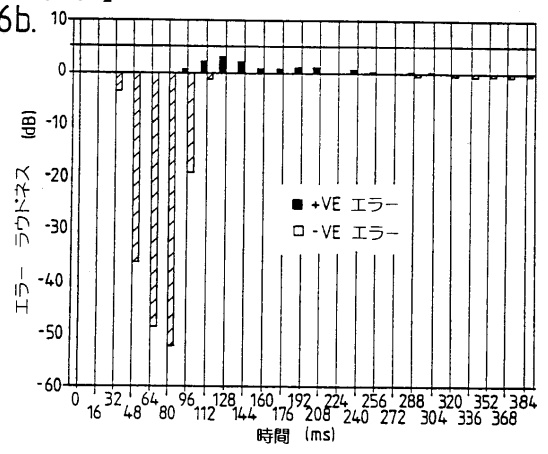
【図15b】
Fig.15b.



【図16a】
Fig.16a.



【図16b】
Fig.16b.



フロントページの続き

(72)発明者 ホーリア、マイケル・ピーター
イギリス国、アイピー４・２アールジェイ、サフォーク、イプスウィッチ、ノーティッジ・ロード
24

合議体

審判長 井関 守三

審判官 小田 浩

審判官 桑江 晃

(56)参考文献 入井寛ほか3名著「電話網の総合的な通話品質を推定するプログラムを開発」NTT技術ジャーナル, 74頁~77頁(1990.11)
長淵裕実ほか2名著「通信の品質[II]-音声品質の客観評価-」電子情報通信学会誌Vol. 75, No. 3, PP. 270~PP. 275(1992年3月)
J. G. BEERENDS ほか1名著「MEASURING THE QUALITY OF AUDIO DEVICES」AES-90th CONVENTION Paris, 19-22 February 1991, NEW YORK(US), Preprint 3070(L-8)、P. 1~10

(58)調査した分野(Int.Cl., DB名)

H04B 17/00

G01H 17/00