

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2001 年 12 月 13 日 (13.12.2001)

PCT

(10) 国際公開番号
WO 01/95314 A1

(51) 国際特許分類: G10L 21/02, 15/20, B25J 13/00

(21) 国際出願番号: PCT/JP01/04858

(22) 国際出願日: 2001 年 6 月 8 日 (08.06.2001)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:
特願2000-173915 2000 年 6 月 9 日 (09.06.2000) JP

(71) 出願人 (米国を除く全ての指定国について): 科学技術
振興事業団 (JAPAN SCIENCE AND TECHNOLOGY
CORPORATION) [JP/JP]; 〒332-0012 埼玉県川口市本
町四丁目1番8号 Saitama (JP).

(NAKADAI, Kazuhiro) [JP/JP]; 〒285-0863 千葉県佐
倉市臼井86 Chiba (JP). 奥乃 博 (OKUNO, Hiroshi)
[JP/JP]; 〒151-0066 東京都渋谷区西原2-10-9 Tokyo
(JP). 北野宏明 (KITANO, Hiroaki) [JP/JP]; 〒350-0035
埼玉県川越市西小仙波町2-18-3 Saitama (JP).

(74) 代理人: 平山一幸 (HIRAYAMA, Kazuyuki); 〒160-
0022 東京都新宿区新宿2-3-10 新宿御苑ビル6階 Tokyo
(JP).

(81) 指定国 (国内): JP, US.

(84) 指定国 (広域): ヨーロッパ特許 (AT, BE, CH, CY, DE,
DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).

添付公開書類:
— 国際調査報告書

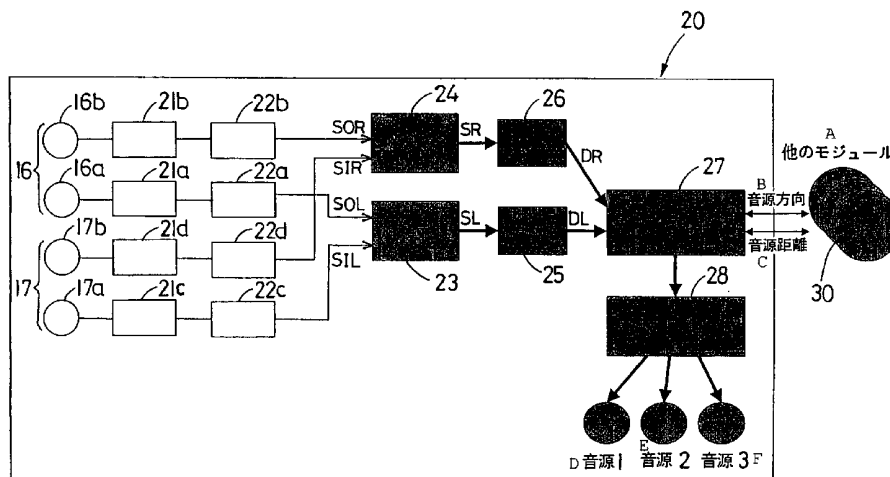
(72) 発明者; および

(75) 発明者/出願人 (米国についてのみ): 中臺一博

2 文字コード及び他の略語については、定期発行される
各PCTガゼットの巻頭に掲載されている「コードと略語
のガイダンスノート」を参照。

(54) Title: ROBOT ACOUSTIC DEVICE AND ROBOT ACOUSTIC SYSTEM

(54) 発明の名称: ロボット聴覚装置及びロボット聴覚システム



A...ANOTHER MODULE

B...DIRECTION OF SOUND SOURCE

C...DISTANCE TO SOUND SOURCE

D...SOUND SOURCE 1

E...SOUND SOURCE 2

F...SOUND SOURCE 3

(57) Abstract: A robot acoustic device and a robot acoustic system, which are capable of active perception by collecting sound from an outside target without being influenced by noise produced inside a robot such as a driving mechanism, comprising at least two outside microphones (16, 16) for mainly collecting outside sound outside a sound-proof armor, at least one inside microphone (17) for mainly collecting inside noise inside the armor, processing sections (23, 24) for canceling noise signals (SIR, SIL) from the inside microphone from acoustic signals (SOR, SOL) from the outside microphones, and a directional information extracting section (27) for deciding the direction of sound on the basis of right and left acoustic signals from the processing sections, wherein the processing sections detect inside burst noise and remove signal components in a band including the burst noise.

[続葉有]

WO 01/95314 A1



(57) 要約:

駆動機構等のロボット内部で発生するノイズに影響されることなく、外部の目標からの音を集音して能動知覚を行うことができるロボット聴覚装置及びロボット聴覚システムであって、防音性の外装の外側において主として外部の音を集音する少なくとも二つの外部マイク（１６，１６）と、外装の内側において主として内部のノイズを集音する少なくとも一つの内部マイク（１７）と、外部マイクからの音響信号（ＳＯＲ，ＳＯＬ）から内部マイクからのノイズ信号（ＳＩＲ，ＳＩＬ）をキャンセルする処理部（２３，２４）と、処理部からの左右の音響信号から、音の方向付けを行なう方向情報抽出部（２７）と、を含んでおり、処理部が、内部のバーストノイズを検出して、このバーストノイズを含む帯域の信号成分を除去する。

明 細 書

ロボット聴覚装置及びロボット聴覚システム

技術分野

本発明はロボット、特に人型または動物型ロボットにおける聴覚装置に関するものである。

技術背景

近年、人型または動物型ロボットにおいては、視覚、聴覚の能動知覚が注目されてきている。能動知覚とは、ロボット視覚やロボット聴覚等の知覚を担当する知覚装置を、知覚すべき目標に追従するように、これらの知覚装置を支持する例えば頭部を駆動機構により姿勢制御するものである。

ここで、能動視覚に関しては、少なくとも知覚装置であるカメラが、駆動機構による姿勢制御によってその光軸方向が目標に向かって保持され、目標に対して自動的にフォーカシングやズームイン、ズームアウト等が行なわれることにより、目標がカメラによって撮像されるようになっており、種々の研究が行なわれている。

これに対して、能動聴覚に関しては、少なくとも知覚装置であるマイクが、駆動機構による姿勢制御によってその指向性が目標に向くようにして保持され、目標からの音がマイクによって集音される。このとき、能動聴覚の不利な点として、駆動機構が作動している間は、マイクが駆動機構の作動音、特にバーストノイズを拾ってしまうため、目標からの音に大きなノイズが混入してしまい、目標からの音を正確に認識できなくなってしまうことがある。

しかしながら、駆動機構を停止させた状態での聴覚研究では、特に目標が移動しているような場合、目標の移動に追従しながらの所謂能動聴覚を行なうことはできない。

さらに、上述した駆動機構だけでなく、ロボット内部で発生する各種動作音及び定常的に発生するノイズも、聴覚装置としてのマイクが集音してしまうので、

同様に完全な能動聴覚を得ることは困難であった。

ところで、ノイズキャンセルの方法としては、所謂アクティブノイズコントロール（ANC）という方法が知られている。

このANC法は、騒音源の近傍にマイクを設けて、騒音源からのノイズを集音し、騒音源からのノイズをキャンセルしたい場所におけるノイズを、IIR（無限インパルス応答）フィルタやFIR（有限インパルス応答）フィルタ等の適応フィルタによって予測して、ノイズをキャンセルしたい場所にて予想した騒音と逆位相の音をスピーカから出力してノイズを打ち消すことにより、ノイズを消滅させる方法である。

しかしながら、このようなANC法においては、ノイズを過去のデータに基づく予想によってキャンセルしているため、所謂バーストノイズに対応することは困難である。また、適応フィルタを利用してノイズをキャンセルしていることから、左右のチャンネル間の位相差情報が歪められ、あるいは消滅してしまうので、音の方向付けができなくなってしまう。

さらに、騒音源からのノイズを集音するマイクは、できるだけノイズのみを集音することが望ましいが、ロボット聴覚装置の場合にはノイズのみを集音することが困難である。

また、ノイズをキャンセルしたい場所におけるノイズを予想するための計算時間が必要であることから、騒音源とスピーカとの距離はある程度離れていることが前提条件となるが、ロボット聴覚装置の場合には、外部音を集音する外部マイクと内部の騒音を集音する内部マイクが比較的近くに設置されることから、計算時間が短くなるため、実質的にANC法を採用することが困難である。

従って、上述したANC法をロボット内部で発生するノイズをキャンセルするために採用することは不適である。

この発明は、以上の点にかんがみて、駆動機構等のロボット内部で発生するノイズに影響されることなく、外部の目標からの音を集音して能動知覚を行なうことができるようにした、ロボット聴覚装置及びロボット聴覚システムを提供することを目的としている。

発明の開示

上記目的は、本発明の第一の構成によれば、内部にノイズ発生源を備えたロボットにおいて、少なくともロボットの一部を覆う防音性の外装と、上記外装の外側に設けられ主として外部の音を集音する少なくとも二つの外部マイクと、上記外装の内側に設けられ主として内部のノイズ発生源からのノイズを集音する少なくとも一つの内部マイクと、上記外部マイク及び内部マイクからの信号に基づいてそれぞれ外部マイクからの音響信号から内部のノイズ発生源からのノイズ信号をキャンセルする処理部と、上記処理部からの左右の音響信号から音の方向付けを行なう方向情報抽出部と、を含んでおり、上記処理部が、内部マイクからの信号からノイズ発生源によるバーストノイズを検出して、このバーストノイズを含む帯域の信号部分を除去することを特徴とするロボット聴覚装置により、達成される。

本発明によるロボット聴覚装置は、防音性の外装が自己認識用に構成されるのが望ましい。

本発明によるロボット聴覚装置は、内外マイクの強度差がテンプレートの駆動機構のノイズの強度差と近く、内外マイクの入力音のスペクトルの強度とパターンとがテンプレートの駆動機構のノイズ周波数応答に近く、さらに駆動機構が動作しているときにノイズをバーストノイズとして、上記処理部が、好ましくはこの帯域の信号部分を除去する。

本発明によるロボット聴覚装置は、好ましくは、上記方向情報抽出部が、聴覚用のエピソード幾何に基づいた方法により、音の方向情報を計算して音の方向付け（音源定位）を行い、調波構造を有する音に関しては、調波構造を利用した音源分離と、音の強度差情報とを利用してロバスト（頑健）な音源定位を行う。

また、上記目的は、本発明の第二の構成によれば、内部にノイズ発生源を備えたロボットにおいて、少なくともロボットの一部を覆う、好ましくはロボットの自己認識用の防音性の外装と、上記外装の外側に設けられ主として外部の音を集音する少なくとも二つの外部マイクと、上記外装の内側に設けられ主として内部のノイズ発生源からのノイズを集音する少なくとも一つの内部マイクと、上記外部マイク及び内部マイクからの信号に基づいて、それぞれ外部マイクからの音響

信号から内部のノイズ発生源からのノイズ信号をキャンセルする処理部と、上記処理部からの左右の音響信号からそれぞれ周波数解析を行なって当該音響データが示すピッチを伴う調波構造から時間、周波数及びパワーに関する音響データを取り出すピッチ抽出部と、上記ピッチ抽出部で取り出された左右の音響データに基づいて音の方向付けを行なう左右チャンネル対応部と、音の調波構造又は上記左右チャンネル対応部で取り出された音の方向付け情報に基づいて上記音響データから音源毎の音響データに分離する音源分離部と、を含んでおり、上記処理部が、内部マイクからの信号からノイズ発生源によるバーストノイズを検出して、このバーストノイズを含む帯域の信号部分を除去することを特徴とするロボット聴覚システムにより達成される。

さらに、上記目的は、本発明の第三の構成によれば、内部に駆動機構等のノイズ発生源を備えた人型または動物型のロボットにおいて、少なくともロボットの頭部を覆う、好ましくはロボットの自己認識用の防音性の外装と、上記外装の外側にて両側の耳に対応する両耳位置に設けられ主として外部の音を集音する少なくとも一對の外部マイクと、上記外装の内側に設けられ主として内部のノイズ発生源からのノイズを集音する少なくとも一つの内部マイクと、上記外部マイク及び内部マイクからの信号に基づいて外部マイクからの音響信号から内部のノイズ発生源からのノイズ信号をキャンセルする処理部と、上記処理部からの左右の音響信号からそれぞれ周波数解析を行なって当該音響データが示すピッチを伴う調波構造から時間、周波数及びパワーに関する音響データを取り出すピッチ抽出部と、上記ピッチ抽出部で取り出された左右の音響データに基づいて音の方向付けを行なう左右チャンネル対応部と、音の調波構造又は上記左右チャンネル対応部で取り出された音の方向付け情報に基づいて上記音響データから音源毎の音響データに分離する音源分離部とを含んでおり、上記処理部が、内部マイクからの信号からノイズ発生源によるバーストノイズを検出して、このバーストノイズを含む帯域の信号部分を除去することを特徴とするロボット聴覚システムにより達成される。

この発明によるロボット聴覚システムは、好ましくは、ロボットが視覚、触覚等の知覚システムをさらに備えており、上記左右チャンネル対応部が、これらの

知覚システムからの情報及び駆動機構の制御信号を参照して音の方向付け及び画像との対応付けを行なう。

この発明によるロボット聴覚システムは、好ましくは、上記左右チャンネル対応部が、上記知覚システムに対して音の方向付けに関する情報を出力する。

本発明によるロボット聴覚システムは、内外マイクの強度差がテンプレートの駆動機構のノイズの強度差と近く、内外マイクの入力音のスペクトルの強度とパターンとがテンプレートの駆動機構のノイズ周波数応答に近く、さらに駆動機構が動作しているときに、ノイズをバーストノイズとして、好ましくは上記処理部がこの帯域の信号部分を除去する。

本発明によるロボット聴覚システムは、内部マイクからの音響信号のパワーが外部マイクからの音響信号のパワーより十分大きく、所定周波数幅の複数のサブバンドにて所定値以上のパワーを有しており、駆動機構の制御信号により駆動機構の連続動作を検出したとき、好ましくは、上記処理部がバーストノイズとしてこの帯域の信号部分を除去する。

本発明によるロボット聴覚システムは、外部マイク及び内部マイクからの音響信号のスペクトルパワー差のパターンが前もって測定された駆動機構のノイズのスペクトルパワー差のパターンとほぼ同じであり、スペクトルの音圧及びパターンが前もって測定された駆動機構のノイズの周波数応答とほぼ同じであって、駆動機構の制御信号により駆動機構の連続動作を検出したとき、好ましくは、上記処理部がバーストノイズとしてこの帯域の信号部分を除去する。

本発明によるロボット聴覚システムは、好ましくは、上記左右チャンネル対応部が、聴覚用のエピポーラ幾何に基づいた方法により音の方向情報を計算して音の方向付け（音源定位）を行い、調波構造を有する音に関しては、調波構造を利用した音源分離と音の強度差情報とを利用してロバスト（頑健）な音源定位を行う。

上記構成によれば、外部マイクが主として外部の目標からの音を集音し、内部マイクが主としてロボット内部の駆動機構等のノイズ発生源からのノイズを集音する。その際、外部マイクが集音した音響信号には、ロボット内部のノイズ発生源からのノイズ信号が混入しているが、この混入したノイズ信号は、処理部にお

ける演算処理によって、内部マイクが集音したノイズ信号によりキャンセルされて著しく低減される。その際、処理部が、内部マイクからの信号からノイズ発生源によるバーストノイズを検出して、外部マイクからの信号からこのバーストノイズを含む帯域の信号部分を除去することによって、方向情報抽出部または左右チャンネル対応部における音の方向付けをバーストノイズに影響されることなくより正確に行なうことができる。

そして、ノイズがキャンセルされた音響信号から、ピッチ抽出部により周波数解析によって音響信号を取り出して、この音響信号から左右チャンネル対応部により音の方向付けを行なった後、この音の方向付けに基づいて音源分離部によって各音源毎の音響データに分離する。

したがって、外部マイクからの音響信号は、処理部での演算処理によってロボット内部の駆動機構等のノイズ発生源からのノイズが容易に著しく低減されると共に、特にバーストノイズを含む帯域の信号部分が除去されることによりそのS/N比が大幅に改善されるので、音源毎の音響データの分離をより一層良好に行なうことができる。

さらに、ロボットが視覚、触覚等の知覚システムを備えており、上記左右チャンネル対応部が、これらの知覚システムからの情報を参照して音の方向付けを行なう場合には、例えば視覚装置からの目標に関する視覚情報に基づいて、左右チャンネル対応部がより明確な方向付けを行なうことができる。

上記左右チャンネル対応部が、上記知覚システムに対して、音の方向付けに関する情報を出力する場合には、例えば視覚装置に対して目標に関する聴覚による方向付けの情報を出力するので、視覚装置がより一層正確な方向付けを行なうことができる。

処理部が、内外マイクの強度差がテンプレートの駆動機構のノイズの強度差と近く、内外マイクの入力音のスペクトルの強度とパターンとがテンプレートの駆動機構のノイズ周波数応答に近く、さらに駆動機構が動作しているときに、ノイズをバーストノイズとして、この帯域の信号部分を除去する場合、または内部マイクからの音響信号のパワーが外部マイクからの音響信号のパワーより十分大きく、所定周波数幅の複数個のサブバンドにて所定値以上のパワーを有しており、

駆動機構の制御信号により駆動機構の連続動作を検出したとき、上記処理部がバーストノイズとしてこの帯域の信号部分を除去する場合には、バーストノイズの除去を容易に行なうことができる。

外部マイク及び内部マイクからの音響信号のスペクトルパワー差のパターンが前もって測定された駆動機構のノイズのスペクトルパワー差のパターンとほぼ同じであり、スペクトルの音圧及びパターンが前もって測定された駆動機構のノイズの周波数応答とほぼ同じであって、駆動機構の制御信号により駆動機構の連続動作を検出したとき、上記処理部がバーストノイズとしてこの帯域の信号部分を除去する場合には、バーストノイズの除去をより正確に行なうことができる。

上記方向情報抽出部が、エピポーラ幾何に基づいた方法により、音の方向情報を計算して音の方向付け（音源定位）を行い、調波構造を有する音に関しては、調波構造を利用した音源分離と、音の強度差情報とを利用してロバスト（頑健）な音源定位を行う場合には、従来の視覚システムで行なわれているエピポーラ幾何による計算方法を聴覚システムに応用することによって、ロボットの外装や音響環境による影響を受けずに、音の方向付けを一層正確に行なうことができる。

ここで、従来の両聴システムにおいて一般的である頭部伝達関数（HRTF）を利用することは、本発明では必要としない。頭部伝達関数は音環境の変化に弱いことが知られているが、本発明では、音環境が変化したとしても頭部伝達関数の再計算や調整を行なう必要がないので、汎用性の高い聴覚システムを構成することができる。

図面の簡単な説明

本発明は、以下の詳細な説明及び本発明の実施例を示す添付図面に基づいて、より良く理解されるものとなろう。なお、添付図面に示す種々の実施例は本発明を特定又は限定することを意図するものではなく、単に本発明の説明及び理解を容易とするためだけのものである。

図中、

図1はこの発明によるロボット聴覚装置の第一の実施形態を組み込んだ人型ロボットの外観を示す正面図である。

図 2 は図 1 の人型ロボットの側面図である。

図 3 は図 1 の人型ロボットにおける頭部の構成を示す概略拡大図である。

図 4 は図 1 の人型ロボットにおけるロボット聴覚システムの電氣的構成を示すブロック図である。

図 5 は図 4 のロボット聴覚システムにおける要部を示すブロック図である。

図 6 (A) は視覚における、また、図 6 (B) は聴覚におけるエピソード幾何による方向付けを示す概略図である。

図 7 及び図 8 は、それぞれ音源定位と音源分離の処理を示す概念図である。

図 9 は図 4 のロボット聴覚システムの実験例を示す概略図である。

図 10 は図 9 の実験におけるロボット頭部の (A) 速い動作及び (B) 遅い動作における入力信号のスペクトログラムである。

図 11 (A) は図 9 の実験におけるバーストノイズを除去しない場合の速い動作における方向情報を示すグラフであり、図 11 (B) は遅い動作における方向情報を示すグラフである。

図 12 (A) は図 9 の実験における弱いバーストノイズを除去した場合の速い動作における方向情報を示すグラフであり、図 12 (B) は遅い動作における方向情報を示すグラフである。

図 13 (A) は図 9 の実験における強いバーストノイズを除去した場合の速い動作における方向情報を示すグラフであり、図 12 (B) は遅い動作における方向情報を示すグラフである。

図 14 (A) は図 13 (A) に対応するスペクトログラムを、図 14 (B) は図 13 (B) に対応するスペクトログラムで、いずれもノイズに比べて信号が強い場合を示している。

図 15 (A) は駆動手段のノイズの内部マイクによる周波数応答を示すグラフであり、図 15 (B) は外部マイクによる周波数応答を示すグラフである。

図 16 (A) は図 15 の周波数応答における駆動手段のノイズのグラフ、図 16 (B) は外部音のスペクトルパワー差のパターンを示すグラフである。

図 17 は、ロボット頭部の遅い動作における入力信号のスペクトログラムである。

図18はバーストノイズを除去しない場合の方向情報を示すグラフである。

図19は図9の実験と同様の第一のバーストノイズ除去方法による方向情報を示すグラフである。

図20は第二のバーストノイズ除去方法による方向情報を示すグラフである。

発明を実施するための最良の形態

以下、本発明のロボット聴覚装置及びロボット聴覚システムにおける実施の形態を図面を用いて詳細に説明する。

図1乃至図2はこの発明によるロボット聴覚システムの一実施形態を備えた実験用の人型ロボットの全体構成を示している。

図1において、人型ロボット10は、4DOF（自由度）のロボットとして構成されており、ベース11と、ベース11上にて一軸（垂直軸）周りに回動可能に支持された胴体部12と、胴体部12上にて三軸方向（垂直軸、左右方向の水平軸及び前後方向の水平軸）の周りに揺動可能に支持された頭部13と、を含んでいる。

上記ベース11は固定配置されていてもよく、脚部として動作可能としてもよく、或いは、移動可能な台車等の上に載置されていてもよい。

上記胴体部12は、ベース11に対して垂直軸の周りに、図1にて矢印Aで示すように回動可能に支持されており、図示しない駆動手段によって回転駆動されると共に、図示の場合、防音性の外装によって覆われている。

上記頭部13は胴体部12に対して連結部材13aを介して支持されており、この連結部材13aに対して前後方向の水平軸の周りに、図1にて矢印Bで示すように揺動可能に、また左右方向の水平軸の周りに、図2にて矢印Cで示すように揺動可能に支持されていると共に、上記連結部材13aが、胴体部12に対してさらに前後方向の水平軸の周りに、図1にて矢印Dで示すように揺動可能に支持されており、それぞれ図示しない駆動手段によって各矢印A、B、C、D方向に回転駆動される。

ここで、上記頭部13は、図3に示すように全体が防音性の外装14により覆われていると共に、前側にロボット視覚を担当する視覚装置としてのカメラ15

を、また両側にロボット聴覚を担当する聴覚装置としての一对の外部マイク 1 6 (1 6 a, 1 6 b) を備えている。

さらに、上記頭部 1 3 は、図 3 に示すように、外装 1 4 の内側にて左右に離隔して配設された一对の内部マイク 1 7 (1 7 a, 1 7 b) を備えている。

上記外装 1 4 は、例えばウレタン樹脂等の吸音性の合成樹脂から構成されており、頭部 1 3 の内部をほぼ完全に密閉することにより、頭部 1 3 の内部の遮音を行なうように構成されている。なお、胴体部 1 2 の外装も、同様にして吸音性の合成樹脂から構成されている。なお、この外装 1 4 は、ロボットが自己を認識するためのものでもあり、自己認識用にロボット内外から発せられる音を仕切る役割をするようになっている。ここで、自己認識とは、ロボットが、自分の内部の駆動手段から発したノイズやロボットの口（図示省略）から発せられた声などの音声とロボット外部からの音とを区別することを意味する。本発明では上記外装 1 4 は、ロボットの内部と外部の音が区別できる程度にロボット内部が密閉されるように構成されている。

上記カメラ 1 5 は公知の構成であって、例えば所謂パン、チルト、ズームの 3 D O F (自由度) を有する市販のカメラが適用され得る。

上記外部マイク 1 6 は、それぞれ頭部 1 3 の側面において、前方に向かって指向性を有するように取り付けられている。

ここで、外部マイク 1 6 の左右の各外部マイク 1 6 a, 1 6 b は、それぞれ図 1 及び図 2 に示すように、外装 1 4 の両側にて前方に向いた段部 1 4 a, 1 4 b にて内側に取り付けられ、段部 1 4 a, 1 4 b に設けられた貫通穴を通して前方の音を集音すると共に、外装 1 4 の内部の音をできるだけ拾わないように、音反射材や吸音材などを配した適宜の手段により遮音されている。これにより、外部マイク 1 6 a, 1 6 b は所謂バイノーラルマイクとして構成されている。なお、外部マイク 1 6 a, 1 6 b の取付位置の近傍において、段部 1 4 a, 1 4 b は人間の外耳形状或いは碗状に形成されていてもよい。

上記内部マイク 1 7 は、外装 1 4 の内側において、上述した各外部マイク 1 6 a, 1 6 b の近傍に、図示の場合には、前記カメラ 1 5 の両端付近の上方にそれぞれ左右一对で配設されている。すなわち、一方の外部マイク 1 6 a の近傍には

一方の内部マイク 17 a が、また、他方の外部マイク 16 b の近傍には他方の内部マイク 17 b が配設されている。尤も、上記内部マイク 17 は、上記した位置に限定されることなく外装 14 の内側で任意の位置に備えられていてよい。

図 4 は、上記外部マイク 16 及び内部マイク 17 を含む音響処理のための聴覚システムの電氣的構成を示している。図 4 において、聴覚システム 20 は、各外部マイク 16 a, 16 b 及び内部マイク 17 a, 17 b からの音響信号をそれぞれ増幅するアンプ 21 a, 21 b, 21 c, 21 d と、これらの増幅信号を AD 変換によりデジタル音響信号 SOL, SOR, SIL, SIR に変換する AD コンバータ 22 a, 22 b, 22 c, 22 d と、これらのデジタル音響信号が入力される処理部としての左右のノイズ除去回路 23, 24 と、ノイズ除去回路 23, 24 からのデジタル音響信号 SR, SL が入力されるピッチ抽出部 25, 26 と、ピッチ抽出部 25, 26 からの音響データが入力される左右チャンネル対応部 27 と、左右チャンネル対応部 27 からのデータが入力される音源分離部 28 と、から構成されている。

上記 AD コンバータ 22 a 乃至 22 d は、例えば量子化ビット数 16 または 24 で、48 kHz サンプリングした信号を取り出すように構成されている。

そして、左側の外部マイク 16 a からのデジタル音響信号 SOL 及び左側の内部マイク 17 a からのデジタル音響信号 SIL はノイズ除去回路 23 に入力され、また右側の外部マイク 16 b からのデジタル音響信号 SOR 及び左側の内部マイク 17 b からのデジタル音響信号 SIR はノイズ除去回路 24 に入力される。これらのノイズ除去回路 23, 24 は同一の構成であり、それぞれ外部マイク 16 からの音響信号から内部マイク 17 のノイズ信号によりノイズキャンセルするように構成されている。即ち、ノイズ除去回路 23 は、外部マイク 16 a からのデジタル音響信号 SOL を、内部マイク 17 a が集音したロボット内部のノイズ発生源からのノイズ信号 SIL に基づいて、例えば最も簡便には、外部マイク 16 a からのデジタル音響信号 SOL から内部マイク 17 a からの音響信号 SIL を減算する等の適宜の演算処理によるノイズキャンセル処理によって、外部マイク 16 a からの音響信号 SOL に混入したロボット内部の各駆動機構（駆動手段）等のノイズ発生源からのノイズを除去して左の音響信号 SL を生成する。また

、ノイズ除去回路 24 は、外部マイク 16b からのデジタル音響信号 SOR を、内部マイク 17b が集音したロボット内部のノイズ発生源からのノイズ信号 SIR に基づいて、例えば最も簡便には、外部マイク 16b からのデジタル音響信号 SOR から内部マイク 17b からの音響信号 SIR を減算する等の適宜の演算処理によるノイズキャンセル処理によって、外部マイク 16b からの音響信号 SIR に混入したロボット内部の各駆動機構等のノイズ発生源からのノイズを除去して右の音響信号 SR を生成する。

ここで、上記ノイズ除去回路 23, 24 は、内部マイク 17a, 17b からの音響信号 SIL, SIR から所謂バーストノイズを検出して、外部マイク 16a, 16b からの音響信号 SOL, SOR からこのバーストノイズの帯域に対応する信号部分を除去することにより、バーストノイズの混入による音の方向付けの精度を高めるようになっている。このバーストノイズの除去は、ノイズ除去回路 23, 24 内にて、以下のようにして行なわれる。

先ず、ノイズ除去回路 23, 24 は、第一のバーストノイズ除去方法として、内部マイク 17a, 17b からの音響信号 SIL, SIR と外部マイク 16a, 16b からの音響信号 SOL, SOR とを比較して、音響信号 SIL, SIR のパワーが音響信号 SOL, SOR のパワーより十分大きく、また音響信号 SIL, SIR のパワーが所定値（例えば 30 dB）以上のピークが所定周波数幅、例えば 47 Hz のサブバンドに関して、一定数（例えば 20）以上のサブバンドに亘って連続していて、さらに駆動手段が作動を続けている場合に、バーストノイズがあると判断して、音響信号 SOL, SOR の当該サブバンドに対応する信号部分を除去する。このために、ノイズ除去回路 23, 24 には駆動機構の制御信号が入力されている。

なお、バーストノイズの除去やその検出の判断方法としては、好ましくは後述する第二のバーストノイズ除去方法を行うのが望ましい。

このようなバーストノイズの除去は、例えばアダプティブフィルタによって行なわれる。このアダプティブフィルタは、リニアフェーズフィルタであって、例えば 100 のオーダーの FIR フィルタにより構成される。各 FIR フィルタのパラメータは、適応アルゴリズムとしての最小二乗法により計算される。

かくして、ノイズ除去回路 2 3, 2 4 は、図 6 に示すように、それぞれバーストノイズ除去部として機能してバーストノイズの検出及び除去を行なう。

上記ピッチ抽出部 2 5, 2 6 は同一の構成であり、ノイズ除去回路 2 3, 2 4 からの左右の音響信号 S L, S R をそれぞれ周波数解析して、時間、周波数及びパワーの三軸から成る音響データを取り出すように構成されている。即ち、ピッチ抽出部 2 5 は、ノイズ除去回路 2 3 からの左の音響信号 S L を周波数解析することにより、時間とパワーで成る二軸の音響信号 S L から時間、周波数及びパワーの三軸から成るスペクトログラムと呼ばれる左の音響データ D L を取り出す。また、ピッチ抽出部 2 6 も同様にして、ノイズ除去回路 2 4 からの右の音響信号 S R を周波数解析することにより、時間とパワーから成る二軸の音響信号 S R から時間、周波数及びパワーの三軸で成る右の音響データ D R を取り出す。

ここで、上記周波数解析は、例えば 2 0 m 秒の窓長で、7. 5 m 秒づつずらしながら、F F T（高速フーリエ変換）を行なうことにより実行される。なお、このような周波数解析は、F F T に限らず一般的な種々の方法で行なわれるようにしてもよい。

このようにして得られた音響データ D L においては、音声や音楽における各音はスペクトログラム上にてピークの連なりとして示されると共に、一般的に調波構造を有しており、整数倍の周波数値に規則的にピークを有するようになっている。

ピークの抽出は次のように行われる。スペクトルは、例えば 4 8 K H z のサンプリング速度で 1 0 2 4 のサブバンドについてフーリエ変換することにより計算される。次に、スペクトルの内、閾値以上のパワーを持ったローカルピークを抽出する。閾値は周波数ごとに異なり、部屋の暗騒音を一定時間計測することによって自動的に求められる。この際、バンドパスフィルタを用いて、ノイズの大きい 9 0 H z 以下の周波数域とパワーの小さい 3 K H z 以上の高周波数域とは計算量低減のためにカットする。これにより十分高速なピーク抽出が実現できる。

上記左右チャンネル対応部 2 7 は、左右のピッチ抽出部 2 5, 2 6 からの音響データ D L, D R にてそれぞれピークにより示される調波構造のピッチに関して、その位相差、時間差等に基づいて同じ音から派生するピッチを左右チャンネル

で対応付けることにより音の方向付けを行なう。この音の方向付けは、例えばエピポーラ幾何に基づいた方法により、音の方向情報を計算して音の方向付け（音源定位）を行い、調波構造を有する音に関しては、調波構造を利用した音源分離と音の強度差情報とを利用してロバスト（頑健）な音源定位を行う。

ここで、視覚によるエピポーラ幾何の場合には、図6（A）に示すように、互いに光軸が平行であって画面が同一面に位置し、同じ焦点距離を有する二つのカメラから成る簡単なステレオカメラにおいて、点P（X，Y，Z）が各カメラの画面にて点P1（x1，y1）及び点P2（xr，yr）に投影されるとき、以下の関係式

$$X = \frac{b(x_1 + x_r)}{2d}, \quad Y = \frac{b(y_1 + y_r)}{2d}, \quad Z = \frac{bf}{d}$$

が成立する。ここで、fは各カメラの焦点距離、bはベースライン、dは（x1－xr）で定義される。

このエピポーラ幾何の概念を聴覚に導入することにより、図6（B）に示すように、外部マイク16a，16bの中心から音源Pへの角度θについて、以下の方程式

$$\cos \theta = \frac{v}{2\pi f b} \Delta \phi$$

が成立する。ここで、vは音速、fは音の周波数である。

そして、左右の外部マイク16a，16bから音源までの距離差Δlに基づいて、外部マイクからの左右の音響信号SOL，SORの間に位相差IPD＝Δφが発生する。

音の方向付けは、例えば各サブバンドの帯域幅が47HzになるようなFFT（高速フーリエ変換）によりピークを抽出し、位相差IPDを計算することにより行なわれる。なお、ピークの抽出は、例えば48kHzのサンプリング速度で1024のサブバンドについてフーリエ変換を計算することにより、HRTFを利用した場合に比較して十分速く、正確に計算される。

これにより、頭部伝達関数（HRTF）によらない音の方向付け（音源定位）が実現できる。ピークの抽出においては、例えば48kHzのサンプリング速度で1024ポイントのFFTを使用して、スペクトルサブトラクションによる方法を用いる。これにより、実時間の処理が精度よく行える。なお、このスペクトルサブトラクションは、FFTの窓関数の性質をも考慮に入れたスペクトル補間を伴っている。

かくして、左右チャンネル対応部27は、図5に示すように、方向情報抽出部として機能して、方向情報を抽出する。図示の場合、左右チャンネル対応部27は、聴覚システム20以外にロボット10に備えられた図示しない他の知覚システム30から目標に関する情報、具体的には例えば視覚システムによる目標の位置、方向、形状、移動の有無等に関する情報や、触覚システムによる目標の柔軟度、振動の有無、触感等の情報が入力されることにより、上述した目標からの音の方向付けが正確に行なわれる。例えば、左右チャンネル対応部27は、上記（聴覚による）方向情報と、カメラ15からの視覚情報に基づく（視覚による）方向情報と比較して、これらの整合性をチェックし関連付ける。

さらに、左右チャンネル対応部27は、人型ロボット10の各駆動手段の制御信号に基づいて、頭部13の方向情報（ロボット座標）を取得することにより、目標との相対位置を計算することができる。これにより、人型ロボット10が移動している場合であっても、目標からの音の方向付けが一層正確に行なわれる。

上記音源分離部28は、公知の構成により、左右チャンネル対応部27からの方向付けの情報及び音響データDL、DRが入力されることにより、ディレクションパスフィルタを使用することによって、方向付けの情報に基づいて音響データDL、DRから音源を同定すると共に、各音源毎に音響データを分離するように構成されている。

このディレクションパスフィルタは、例えば以下のようにしてサブバンドを収集する。即ち、特定の方向 θ を、各サブバンド（47Hz）に関して $\Delta\phi$ に変換した後、ピークを抽出して、位相差（IPD）及び $\Delta\phi'$ を計算する。そして、位相差が $\Delta\phi' = \Delta\phi$ であるとき、当該サブバンドを収集する。このようにして全てのサブバンドについて上記処理をして、収集されたサブバンドから成る波形

を構成する。

ここで、同時のFFTにより得られる左右チャンネルのスペクトルを $S_p^{(l)}$ 及び $S_p^{(r)}$ とし、 $S_p(f_p)$ のピーク周波数を f_p としたとき、左右のチャンネルのスペクトル $S_p^{(l)}(f_p)$ 及び $S_p^{(r)}(f_p)$ は、それぞれ実数部 $R[S_p^{(r)}(f_p)]$ 、 $R[S_p^{(l)}(f_p)]$ 及び虚数部 $I[S_p^{(r)}(f_p)]$ 、 $I[S_p^{(l)}(f_p)]$ で表わされる。

従って、上記 $\Delta\phi$ は、

$$\Delta\phi = \tan^{-1}\left(\frac{I[S_p^{(r)}(f_p)]}{R[S_p^{(r)}(f_p)]}\right) - \tan^{-1}\left(\frac{I[S_p^{(l)}(f_p)]}{R[S_p^{(l)}(f_p)]}\right)$$

により求められる。

かくして、視覚（カメラ15）によるエピポーラ平面から聴覚（外部マイク16）によるエピポーラ平面への変換は図6に示すように容易に行なわれるので、目標の方向（ θ ）は、聴覚によるエピポーラ幾何に基づいて、上記の方程式〔数2〕により、 $f = f_p$ として容易に求められる。

このように、左右チャンネル対応部27で音源定位が行われ、その後音源分離部28で音源が分離される。なお、図7はそれらの処理の概念図である。

また、音の方向付けと音源定位については、調波構造抽出によって音源分離を実現する方法により、調波構造を有する音に対してロバストな音源定位は行える。すなわち、これは、図4に示すモジュールの内、左右チャンネル対応部27と音源分離部28とを入れ替えて、音源分離部28からのデータが左右チャンネル対応部27へ入力されることにより実現できる。

ここで、調波構造を有する音に対する音源分離と音源定位について説明する。図8に示すように、まず音源分離においては、ピーク抽出で抽出されたピークは、周波数が低いものから順番に取りだされる。その周波数 F_0 と一定以内の誤差（例えば心理学的実験によって得られた6%）で整数倍とみなせる周波数 F_n とを持つローカルピークが、周波数 F_0 の倍音としてクラスタリングを行われる。このクラスタリングによって集められた最終的なピークの集合が一つの音と見なされる。これによって、音源分離が行われる。

次に、音源定位について説明する。一般に両耳聴における音源定位には、頭部伝達関数（HRTF）から求められる両耳間位相差（IPD）と両耳間強度差（IID）とが使用される。しかし、HRTFは頭部の形状や環境に大きく依存し、環境が変わる都度、計測が必要であるため実環境アプリケーションには不向きである。

そこで、本発明は、HRTFに依らないIPDを利用した音源定位法として、ステレオ視におけるエピポーラ幾何の概念を聴覚に拡張した聴覚エピポーラ幾何に基づく方法を適用している。

この際、（１）音の倍音構造の利用、（２）IPDを用いた聴覚エピポーラ幾何による定位結果とIIDを用いた定位結果とのデンプスターシェイファー理論を用いた統合、（３）モータ動作中でも正確な音源定位を可能とするアクティブオーディションの導入、によって音源定位のロバスト性を向上させている。

図８に示すように、この音源定位は、音源分離によって分離された調波構造を有した各音に対して行う。ロボットでは、左右のマイクのベースラインから1.5KHz以下の周波数域に対してはIPD、それ以上の周波数域ではIID、による音源定位が有効である。このため、入力音のうち1.5KHz以上の倍音成分と1.5KHz以下の倍音成分との２つに分けて処理を行う。先ず、入力音のうち1.5KHz以下の周波数 f_k を持った各倍音成分に対して、聴覚エピポーラ幾何を使用して、ロボット正面に対して $\pm 90^\circ$ の範囲で 5° おきにIPD仮説（ $P_h(\theta, f_k)$ ）を生成する。

次に、次式に示す距離関数により、入力各倍音におけるIPD（ $P_s(f_k)$ ）と各仮説間の距離（ $d(\theta)$ ）とを計算する。ここで、 $n_{f < 1.5 \text{ KHz}}$ は周波数が1.5KHz以下である倍音数である。

$$d(\theta) = \frac{1}{n_{f < 1.5 \text{ KHz}}} \sum_{k=0}^{n_{f < 1.5 \text{ KHz}} - 1} \frac{(P_h(\theta, f_k) - P_s(f_k))^2}{f_k}$$

次に、得られた距離に対し、次式によって定義される確率密度関数を適用し、

距離を I P D を用いた場合の音源方向を支持する確信度 $B F_{I P D}$ に変換する。ここで、 m と s とは、それぞれ $d(\theta)$ の平均と分散とであり、 n は d の個数である。

$$B F_{I P D}(\theta) = \frac{d(\theta) - m}{\sqrt{\frac{s}{n}}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

入力音のうち 1.5 KHz 以上の周波数を持った倍音に関しては、I I D の総和の正負に応じて下記表 1 に示す値を I I D を用いた場合の音源方向を支持する確信度 $B F_{I I D}$ として与える。

表 1. I I D 確信度 ($B F_{I I D}(\theta)$) を示す表

θ		$90^\circ \sim 35^\circ$	$30^\circ \sim -30^\circ$	$-35^\circ \sim 90^\circ$
I I D の総和	+	0.35	0.5	0.65
	-	0.65	0.5	0.35

I P D, I I D それぞれの処理によって得られた音源方向を支持する値から、これらを次式で示されるデンプスター-シェイファー理論によって統合し、I P D と I I D との両方から音源方向を支持する新しい確信度を生成する。

$$B F_{I P D + I I D}(\theta) = B F_{I P D}(\theta) B F_{I I D}(\theta) + (1 - B F_{I P D}(\theta)) B F_{I I D}(\theta) + B F_{I P D}(\theta) (1 - B F_{I I D}(\theta))$$

このような確信度 $B F_{I P D + I I D}$ を各角度ごとに生成し、それらのうち最大の確信度値を持った方向が音源の方向と成る。

本発明の実施形態による人型ロボット 10 は以上のように構成されており、外部マイク 16 a, 16 b が集音しようとする目標からの音は、以下のようにして集音され、ノイズキャンセルされることにより音源として知覚される。

先ず、外部マイク 16 a, 16 b が、主として目標からの外部の音を集音してそれぞれアナログ音響信号を出力する。ここで、外部マイク 16 a, 16 b はロボット内部からのノイズも集音するが、外装 14 自体が頭部 13 の内部を密閉しており、また外部マイク 16 a, 16 b が頭部 13 の内部に対して遮音されてい

ることにより、混入するノイズが比較的低いレベルに抑えられている。

これに対して、内部マイク 17 a, 17 b は、主としてロボット内部からのノイズ、例えば前述した各駆動機構の作動音や冷却ファンの作動音等のノイズ発生源からのノイズを集音する。ここで、内部マイク 17 a, 17 b は外部からの音も集音するが、外装 14 が内部を密閉していることからそのレベルは比較的強く抑えられている。

このようにして集音された外部マイク 16 a, 16 b からのアナログ音響信号及び内部マイク 17 a, 17 b からのアナログ音響信号は、それぞれアンプ 21 a 乃至 21 d により増幅された後、ADコンバータ 22 a 乃至 22 d によりデジタル音響信号 SOL, SOR, SIL, SIR にAD変換されて、ノイズ除去回路 23 及び 24 に入力される。

ノイズ除去回路 23, 24 は、それぞれ外部マイク 16 a, 16 b からの音響信号 SOL, SOR から内部マイク 17 a, 17 b からの音響信号 SIL, SIR を減算する等の演算処理を行なうことにより、外部マイク 16 a, 16 b からの音響信号 SOL, SOR からロボット内部のノイズ発生源からのノイズ信号を除去すると共に、バーストノイズを検出して、外部マイク 16 a, 16 b からの音響信号 SOL, SOR からバーストノイズを含むサブバンドの信号部分を除去することにより、それぞれノイズ特にバーストノイズを除去した真の音響信号 SL, SR を外部へ出力する。

そして、ピッチ抽出部 25, 26 が、それぞれ音響信号 SL, SR に基づいて周波数解析により、音響信号 SL, SR に含まれる全ての音に関して当該音に関するピッチを抽出し、このピッチに対応する当該音の調波構造、開始時、終了時と共に、音響データ DL, DR として左右チャンネル対応部 27 に出力する。

続いて、左右チャンネル対応部 27 が、これらの音響データ DL, DR に基づいて音源毎の音の方向付けを行なう。

この場合、左右チャンネル対応部 27 は、例えば各ピッチ抽出部 25, 26 で抽出された音響データ DL, DR に基づいて左右のチャンネルの調波構造を比較して、最も近いピッチを対応付ける。その際、左右のチャンネルのピッチを一对一で比較するだけでなく、一方のチャンネルの複数のピッチを他方のチャンネル

の一つのピッチと比較する等により、より正確な対応付けを行なうことが好ましい。

そして、左右チャンネル対応部 27 は、対応付けられたピッチの位相を比較すると共に、前述したエピソード幾何に基づいた方法により、音の方向情報を計算することにより、音の方向付けを行なう。

これにより、音源分離部 28 が、左右チャンネル対応部 27 からの音の方向付け情報に基づいて、音響データ DL, DR から、音源毎の音に関する音響データを抽出して、音源毎の音に分離する。かくして、聴覚システム 20 は音源毎の音の分離によって音響認識を行ない、能動聴覚を行なうことができる。

このようにして、本発明実施形態による人型ロボット 10 によれば、ノイズ除去回路 23, 24 により、外部マイク 16a, 16b からの音響信号 SOL, SOR から、内部マイク 17a, 17b からの音響信号 SIL, SIR に基づいてノイズキャンセルを行なうと共に、外部マイク 16a, 16b からの音響信号 SOL, SOR から、バーストノイズを含むサブバンドの信号成分を除去することにより、各駆動機構によって外部マイク 16a, 16b の指向性の方向を目標に向けながら、バーストノイズの影響を受けることなく、音の方向付けを行なうことができると共に、方向付けが、従来のように HRTF を使用することなく、エピソード幾何に基づいた方法により計算することによって行なわれ得るので、音環境の変化により HRTF の調整または再測定を行なう必要がなく、計算時間が短縮され得ると共に、未知の音環境においても、各音源からの音を分離して、より正確な音声認識を行なうことができる。

従って、例えば目標が移動している場合であっても、各駆動機構によって、外部マイク 16a, 16b の指向性の方向を常に目標に対して追従させながら、目標の音響認識を行なうことができる。このとき、左右チャンネル対応部 27 が、例えば他の知覚システム 30 として、視覚システムからの視覚による目標の方向付け情報を参照して音の方向付けを行なうことにより、より正確な音の方向付けを行なうことができる。

また、他の知覚システム 30 として、視覚システムが使用される場合、左右チャンネル対応部 27 が視覚システムに対して音の方向付け情報を出力するように

してもよい。この場合、視覚システムが画像認識により目標の方向付けを行なう際に、聴覚システム 20 の音に関する方向付け情報を参照することによって、例えば目標が移動して障害物の裏に隠れてしまった場合でも、目標からの音を参照して、目標の方向付けをより一層正確に行なうことができる。

以下に、具体的な実験例を示す。

図 9 に示すように、上記人型ロボット 10 が、10 平方 m の居住用の部屋 40 内にて、二つの音源としてのスピーカ 41, 42 に対向している。ここで、人型ロボット 10 は、その頭部 13 を 53 度（右を 0 度とし、左回りの角度）の方向（回転前の方向）に向けている。

一方のスピーカ 41 は、500 Hz の単調音を再生し、人型ロボット 10 の前方 5 度左（58 度）の位置に在る。これに対して、他方のスピーカ 42 は 600 Hz の単調音を再生し、人型ロボット 10 から見てスピーカ 41 の 69 度左（127 度）の位置に在る。人型ロボット 10 から各スピーカ 41, 42 までの距離は、約 210 cm である。

ここで、人型ロボット 10 のカメラ 15 の視野は、水平方向に関して約 45 度であるから、人型ロボット 10 は、スピーカ 42 をカメラ 15 により見るができない。

この状態で、スピーカ 41 が音を再生し、スピーカ 42 が約 3 秒遅れて音を再生する場合について、人型ロボット 10 が聴覚によりスピーカ 42 の音の方向付けを行なって、スピーカ 42 の方向に向けて頭部 13 を回転し、音源としてのスピーカ 42 と視覚対象物としてのスピーカ 42 とを関連付ける、という実験を行なった。なお、回転後の頭部 13 の方向は 131 度の方向になる。

実験は、人型ロボット 10 の頭部 13 の回転速度について、速い動作（68.8 度/秒）及び遅い動作（14.9 度/秒）と、S/N 比 0 dB の弱いノイズ（内部の待機状態の音と同様のパワー）と S/N 比約 50 dB の強いノイズ（バーストノイズ）の条件で行なわれ、以下のような結果が得られた。

図 10 は、人型ロボット 10 の内部で発生する内部音（ノイズ）のスペクトログラムで、（A）は速い動作の場合、（B）は遅い動作の場合を示している。これらのスペクトログラムによれば、駆動モータによるバーストノイズが明白に示

されている。

従来のノイズ除去による方向情報は、図11(A)または(B)に示すように、頭部13が回転している間(5~6秒の間)、ノイズにより大きな影響を受けており、人型ロボット10が音源を追跡するために頭部13を回転駆動している間は、聴覚がほぼ無効となるようなノイズを発生していることが分かる。

これに対して、本発明によるバーストノイズ除去による方向情報は、図12に示す弱いノイズ及び図13に示す強いノイズの場合に、それぞれ頭部13の回転駆動の間でも、バーストノイズの影響を受けずに、方向情報が正確に得られることが分かる。なお、図14(A)は図13(A)に対応するスペクトログラムを表し、また、図14(B)は図13(B)に対応するスペクトログラムを表すもので、いずれもノイズに比べて信号が強い場合を示している。

上記ノイズ除去回路23, 24は、上述したように、各サブバンドについて、音響信号SIL, SIRのパワーに基づいて、バーストノイズの有無の判定を行なってバーストノイズを除去しているが、バーストノイズの除去は、外装14の音響特性に基づいて、以下のようにして行なうようにしてもよい。

第二のバーストノイズ除去方法においては、マイクに入力されるノイズは、ある時刻に、下記の3要件((1)~(3))を具備するとき、バーストノイズとして取り扱われる。

(1) 内外のマイク16a, 16b, 17a, 17bの強度差がテンプレートのモータなどの駆動手段のノイズの強度差と近い。

(2) 内外マイクの入力音のスペクトル強度とパターンがテンプレートのモータノイズ周波数応答に近い。

(3) モータなどの駆動手段が動作している。

即ち、第二のバーストノイズ除去方法においては、先ず、ノイズ除去回路23, 24は、前以てロボット10の各駆動手段の動作時における音響測定データ(後述の図15(A)及び(B)や図16(A)及び(B))、即ち外部マイク16及び内部マイク17からの音響信号データを測定し、テンプレートとして記憶しておく。

次に、ノイズ除去回路23, 24は各サブバンド毎に、内部マイク17a, 1

7 bからの音響信号S I L, S I Rと外部マイク1 6 a, 1 6 bからの音響信号S O L, S O Rに関して、上記記憶した音響測定データをテンプレートとしてバーストノイズの判定を行なう。即ち、ノイズ除去回路2 3, 2 4は、外部マイクと内部マイクのスペクトルパワー差（又は音圧差）のパターンが、測定された音響測定データにおける駆動手段のノイズのスペクトルパワー差のパターンとほぼ同じであり、スペクトルの音圧及びパターンが測定された駆動手段のノイズの周波数応答とほぼ同じであって、さらに駆動手段が作動を続けている場合に、バーストノイズがあると判断して当該サブバンドに対応する信号部分を除去する。

このようなバーストノイズの判定は、以下の理由に基づいている。外装1 4の音響特性は無響室において測定される。その際、測定される音響特性の項目は次の通りである。カバーロボット1 0の各駆動手段、即ち頭部1 3を前後方向に揺動させる第一のモータ（モータ1）、頭部1 3を左右方向に揺動させる第二のモータ（モータ2）、頭部1 3を垂直軸の周りに回動させる第三のモータ（モータ3）及び胴体部1 2を垂直軸の周りに回動させる第四のモータ（モータ4）のノイズに関する内部マイク1 7及び外部マイク1 6による周波数対応は、図1 5（A）、（B）に示すようになっている。また、内部マイク1 7及び外部マイク1 6のスペクトルパワー差のパターンは、図1 6（A）に示すようになり、外部マイクの周波数応答から内部マイクの周波数応答を減算することにより得られる。同様にして、外部音のスペクトルパワー差のパターンは、図1 6（B）に示すようになる。これはインパルス応答により得られる。インパルス応答は、水平方向及び垂直方向のマトリックス要素、即ち水平方向がロボット中心から0度、±45度、±90度及び±180度、そして垂直方向が0度及び30度である12点で測定される。

これらの図から、以下のことが観察される。即ち、

1. 駆動手段（モータ）のノイズは、広帯域であり、図1 5（A）、（B）に示すように、外部マイクによる信号よりも内部マイクによる信号が10 dB程度大きい。

2. 駆動手段（モータ）のノイズは、図1 6（A）に示すように、2.5 kHz以上の周波数に関して、内部マイクより外部マイクの方がやや大きいほぼ同

等である。これは、外装14で外部音が遮断されることにより、内部マイクがより駆動手段のノイズを拾いやすいことを示している。

3. 駆動手段（モータ）のノイズは、2 kHz以下の周波数では、外部マイクより内部マイクの方がやや大きく、この傾向は、図16（B）に示すように、特に700 Hz以下の周波数で顕著である。これは、外装14内での共鳴を示しており、外装14の直径が約18 cmであることから、周波数500 Hzにおける $\lambda/4$ に対応している。また、図16（A）においても、同様の共鳴が発生している。

4. 内部音は、図15（A），（B）を比較して、平均して外部音より約10 dB大きい。従って、外装14の内部音及び外部音の分離効率は約10 dBである。

このようにして、前以て外部マイクと内部マイクのスペクトルパワー差のパターン、そして共鳴によるピークを含むスペクトルの音圧及びパターンを記憶させておくことにより、駆動手段（モータ）のノイズに関する測定データを保持しておき、ノイズ除去回路23，24が、各サブバンド毎に上述したバーストノイズの判定を行なって、バーストノイズがあると判定されたサブバンドに対応する信号部分を除去することにより、バーストノイズの影響を排除することができる。

前述したと同様の実験例を示す。

この場合、前述の実験例と同じ条件で、遅い動作（14.9度/秒）でのみ実験が行なわれ、以下のような結果が得られた。

図17は、人型ロボット10の内部で発生する内部音（ノイズ）のスペクトログラムを示している。このスペクトログラムによれば、駆動モータによるバーストノイズが明白に示されている。

ノイズ除去のない場合による方向情報は、図18に示すように、頭部13が回転している間（5～6秒の間）、ノイズにより影響を受けており、人型ロボット10が音源を追跡するために頭部13を回転駆動している間は、聴覚がほぼ無効となるようなノイズを発生していることが分かる。

また、前述した第一のバーストノイズ除去方法による方向情報は、図19に示すように、頭部13の回転駆動の間でもバーストノイズの影響による揺れがやや

少なく、方向情報が比較的正確に得られる。

これに対して、上述した第二のバーストノイズ除去方法による方向情報は、図 20 に示すように、頭部 13 の回転駆動の間でもバーストノイズの影響による揺れが極めて少なくなり、方向情報がより一層正確に得られることが分かる。

なお、上記の実験と並行して、前述した ANC 法（適応フィルタとして FIR フィルタを使用）によるノイズキャンセルも実験してみたが、バーストノイズを有効にキャンセルすることはできなかった。

上述した実施形態において、人型ロボット 10 は、4 DOF（自由度）を有するように構成されているが、これに限らず、任意の動作を行なうように構成されたロボットに本発明によるロボット聴覚システムを組み込むことも可能である。

また、上述した実施形態においては、本発明によるロボット聴覚システムを人型ロボット 10 に組み込んだ場合について説明したが、これに限らず、犬型等の各種動物型ロボットや、その他の形式のロボットに組み込むことも可能であることは明らかである。

さらに、上述した実施形態においては、内部マイク 17 は一対のマイク 17 a, 17 b から構成されているが、一つ以上のマイクから構成されていてもよい。また、外部マイク 16 は一対のマイク 16 a, 16 b から構成されているが、二対以上のマイクから構成されていてもよい。

従来技術の ANC は、位相に影響を与えるフィルタリング処理を行うことで位相のずれが生じるため、音源定位を正確に行う場合には適していなかった。これに対して、本願発明によれば、位相情報に影響を与えるフィルタリングを行わないため、即ちノイズが混入している部分のデータを使わないようにしていることで、位相のずれは生じず、それゆえ、音源定位に適している。

産業上の利用可能性

以上のように、本発明によれば、駆動機構等のロボット内部で発生するノイズに影響されることなく、外部の目標からの音を集音して能動知覚を行なうことができるようにした、極めて優れたロボット聴覚装置及びロボット聴覚システムを提供できる。

請 求 の 範 囲

1. 内部にノイズ発生源を備えたロボットにおいて、
少なくともロボットの一部を覆う防音性の外装と、
上記外装の外側に設けられ、主として外部の音を集音する少なくとも二つの外部マイクと、
上記外装の内側に設けられ、主として内部のノイズ発生源からのノイズを集音する少なくとも一つの内部マイクと、
上記外部マイク及び内部マイクからの信号に基づいて、それぞれ外部マイクからの音響信号から、内部のノイズ発生源からのノイズ信号をキャンセルする処理部と、
上記処理部からの左右の音響信号から、音の方向付けを行なう方向情報抽出部と、を含んでおり、
上記処理部が、内部マイクからの信号からノイズ発生源によるバーストノイズを検出して、このバーストノイズを含む帯域の信号部分を除去することを特徴とするロボット聴覚装置。
2. 内部にノイズ発生源を備えたロボットにおいて、
少なくともロボットの一部を覆う自己認識用の防音性の外装と、
上記外装の外側に設けられ、主として外部の音を集音する少なくとも二つの外部マイクと、
上記外装の内側に設けられ、主として内部のノイズ発生源からのノイズを集音する少なくとも一つの内部マイクと、
上記外部マイク及び内部マイクからの信号に基づいて、それぞれ外部マイクからの音響信号から、内部のノイズ発生源からのノイズ信号をキャンセルする処理部と、
上記処理部からの左右の音響信号から、音の方向付けを行なう方向情報抽出部と、を含んでおり、
上記処理部が、内部マイクからの信号からノイズ発生源によるバーストノイ

ズを検出して、このバーストノイズを含む帯域の信号部分を除去することを特徴とするロボット聴覚装置。

3. 前記処理部が、内部マイクからの音響信号のパワーが外部マイクからの音響信号のパワーより十分大きく、所定値以上のピークが所定値以上の帯域に亘って検出されたとき、バーストノイズとしてこの帯域の信号部分を除去することを特徴とする、請求の範囲第1項又は第2項に記載のロボット聴覚装置。
4. 前記方向情報抽出部が、聴覚用のエピポーラ幾何により音の方向情報を計算して、音の方向付けを行なうことを特徴とする、請求の範囲第1項又は第2項に記載のロボット聴覚装置。
5. 前記方向情報抽出部が、聴覚用のエピポーラ幾何に基づいた方法により、音の方向情報を計算して音の方向付けを行い、
調波構造を有する音に関しては、調波構造を利用した音源分離と音の強度差情報とを利用して音の方向付けを行なうことを特徴とする、請求の範囲第1項又は第2項に記載のロボット聴覚装置。
6. 内部にノイズ発生源を備えたロボットにおいて、
少なくともロボットの一部を覆う防音性の外装と、
上記外装の外側に設けられ、主として外部の音を集音する少なくとも二つの外部マイクと、
上記外装の内側に設けられ、主として内部のノイズ発生源からのノイズを集音する少なくとも一つの内部マイクと、
上記外部マイク及び内部マイクからの信号に基づいて、外部マイクからの音響信号から内部のノイズ発生源からのノイズ信号をキャンセルする処理部と、
上記処理部からの左右の音響信号から、それぞれ周波数解析を行なって、当該音響データが示すピッチを伴う調波構造から、時間、周波数及びパワーに関する音響データを取り出すピッチ抽出部と、

上記ピッチ抽出部で取り出された左右の音響データに基づいて、音の方向付けを行なう左右チャンネル対応部と、

音の調波構造又は上記左右チャンネル対応部で取り出された音の方向付け情報に基づいて、上記音響データから音源毎の音響データに分離する音源分離部と、を含んでおり、

上記処理部が、内部マイクからの信号からノイズ発生源によるバーストノイズを検出して、このバーストノイズを含む帯域の信号部分を除去することを特徴とするロボット聴覚システム。

7. 内部にノイズ発生源を備えたロボットにおいて、

少なくともロボットの一部を覆う自己認識用の防音性の外装と、

上記外装の外側に設けられ、主として外部の音を集音する少なくとも二つの外部マイクと、

上記外装の内側に設けられ、主として内部のノイズ発生源からのノイズを集音する少なくとも一つの内部マイクと、

上記外部マイク及び内部マイクからの信号に基づいて、外部マイクからの音響信号から内部のノイズ発生源からのノイズ信号をキャンセルする処理部と、

上記処理部からの左右の音響信号からそれぞれ周波数解析を行なって、当該音響データが示すピッチを伴う調波構造から時間、周波数及びパワーに関する音響データを取り出すピッチ抽出部と、

上記ピッチ抽出部で取り出された左右の音響データに基づいて、音の方向付けを行なう左右チャンネル対応部と、

音の調波構造又は上記左右チャンネル対応部で取り出された音の方向付け情報に基づいて、上記音響データから音源毎の音響データに分離する音源分離部と、を含んでおり、

上記処理部が、内部マイクからの信号からノイズ発生源によるバーストノイズを検出して、このバーストノイズを含む帯域の信号部分を除去することを特徴とするロボット聴覚システム。

8. 内部に駆動機構等のノイズ発生源を備えた人型または動物型のロボットにおいて、

少なくともロボットの頭部を覆う防音性の外装と、

上記外装の外側にて両側の耳に対応する両耳位置に設けられ、主として外部の音を集音する少なくとも一対の外部マイクと、

上記外装の内側に設けられ、主として内部のノイズ発生源からのノイズを集音する少なくとも一つの内部マイクと、

上記外部マイク及び内部マイクからの信号に基づいて、外部マイクからの音響信号から内部のノイズ発生源からのノイズ信号をキャンセルする処理部と、

上記処理部からの左右の音響信号から、それぞれ周波数解析を行なって、当該音響データが示すピッチを伴う調波構造から、時間、周波数及びパワーに関する音響データを取り出すピッチ抽出部と、

上記ピッチ抽出部で取り出された左右の音響データに基づいて、音の方向付けを行なう左右チャンネル対応部と、

音の調波構造又は上記左右チャンネル対応部で取り出された音の方向付け情報に基づいて、上記音響データから音源毎の音響データに分離する音源分離部と、を含んでおり、

上記処理部が、内部マイクからの信号からノイズ発生源によるバーストノイズを検出して、このバーストノイズを含む帯域の信号部分を除去することを特徴とするロボット聴覚システム。

9. 内部に駆動機構等のノイズ発生源を備えた人型または動物型のロボットにおいて、

少なくともロボットの頭部を覆う自己認識用の防音性の外装と、

上記外装の外側にて両側の耳に対応する両耳位置に設けられ、主として外部の音を集音する少なくとも一対の外部マイクと、

上記外装の内側に設けられ、主として内部のノイズ発生源からのノイズを集音する少なくとも一つの内部マイクと、

上記外部マイク及び内部マイクからの信号に基づいて、外部マイクからの音

響信号から内部のノイズ発生源からのノイズ信号をキャンセルする処理部と、
上記処理部からの左右の音響信号からそれぞれ周波数解析を行なって、当該音響データが示すピッチを伴う調波構造から時間、周波数及びパワーに関する音響データを取り出すピッチ抽出部と、

上記ピッチ抽出部で取り出された左右の音響データに基づいて、音の方向付けを行なう左右チャンネル対応部と、

音の調波構造又は上記左右チャンネル対応部で取り出された音の方向付け情報に基づいて、上記音響データから音源毎の音響データに分離する音源分離部と、を含んでおり、

上記処理部が、内部マイクからの信号からノイズ発生源によるバーストノイズを検出して、このバーストノイズを含む帯域の信号部分を除去することを特徴とするロボット聴覚システム。

10. 視覚、触覚等の知覚システムをさらに備えており、前記左右チャンネル対応部が、これらの知覚システムからの情報及び駆動機構の制御信号を参照して音の方向付け及び画像との対応付けを行なうことを特徴とする、請求の範囲第6項から第9項のいずれかに記載のロボット聴覚システム。

11. 視覚、触覚等の知覚システムをさらに備えており、前記左右チャンネル対応部が、これらの知覚システムからの情報及び駆動機構の制御信号を参照して音の方向付け及び画像との対応付けを行い、

前記左右チャンネル対応部が、前記知覚システムに対して音の方向付けに関する情報を出力することを特徴とする、請求の範囲第6項から第9項のいずれかに記載のロボット聴覚システム。

12. 前記処理部が、内外マイクの強度差がテンプレートの駆動機構のノイズの強度差と近く、内外マイクの入力音のスペクトルの強度とパターンとがテンプレートの駆動機構のノイズ周波数応答に近く、さらに駆動機構が動作しているときに、ノイズをバーストノイズとしてこの帯域の信号部分を除去することを

特徴とする、請求の範囲第 6 項から第 9 項のいずれかに記載のロボット聴覚システム。

- 1 3. 視覚、触覚等の知覚システムをさらに備えており、前記左右チャンネル対応部が、これらの知覚システムからの情報及び駆動機構の制御信号を参照して音の方向付け及び画像との対応付けを行ない、

前記処理部が、内外マイクの強度差がテンプレートの駆動機構のノイズの強度差と近く、内外マイクの入力音のスペクトルの強度とパターンとがテンプレートの駆動機構のノイズ周波数応答に近く、さらに駆動機構が動作しているときに、ノイズをバーストノイズとしてこの帯域の信号部分を除去することを特徴とする、請求の範囲第 6 項から第 9 項のいずれかに記載のロボット聴覚システム。

- 1 4. 視覚、触覚等の知覚システムをさらに備えており、前記左右チャンネル対応部が、これらの知覚システムからの情報及び駆動機構の制御信号を参照して音の方向付け及び画像との対応付けを行ない、

前記左右チャンネル対応部が、前記知覚システムに対して音の方向付けに関する情報を出力し、

前記処理部が、内外マイクの強度差がテンプレートの駆動機構のノイズの強度差と近く、内外マイクの入力音のスペクトルの強度とパターンとがテンプレートの駆動機構のノイズ周波数応答に近く、さらに駆動機構が動作しているときに、ノイズをバーストノイズとしてこの帯域の信号部分を除去することを特徴とする、請求の範囲第 6 項から第 9 項のいずれかに記載のロボット聴覚システム。

- 1 5. 前記処理部が、外部マイク及び内部マイクからの音響信号のスペクトルパワー差のパターンが前もって測定された駆動機構のノイズのスペクトルパワー差のパターンとほぼ同じであり、スペクトルの音圧及びパターンが前もって測定された駆動機構のノイズの周波数応答とほぼ同じであって、駆動機構の制御

信号により駆動機構の連続動作を検出したとき、バーストノイズとしてこの帯域の信号部分を除去することを特徴とする、請求の範囲第 8 項又は第 9 項に記載のロボット聴覚システム。

1 6. 前記左右チャンネル対応部が、聴覚用のエピソード幾何により音の方向情報を計算して音の方向付けを行なうことを特徴とする、請求の範囲第 6 項から第 9 項のいずれか 1 項に記載のロボット聴覚システム。

1 7. 視覚、触覚等の知覚システムをさらに備えており、前記左右チャンネル対応部が、これらの知覚システムからの情報及び駆動機構の制御信号を参照して音の方向付け及び画像との対応付けを行ない、

前記左右チャンネル対応部が、聴覚用のエピソード幾何により音の方向情報を計算して音の方向付けを行なうことを特徴とする、請求の範囲第 6 項から第 9 項のいずれかに記載のロボット聴覚システム。

1 8. 視覚、触覚等の知覚システムをさらに備えており、前記左右チャンネル対応部が、これらの知覚システムからの情報及び駆動機構の制御信号を参照して音の方向付け及び画像との対応付けを行ない、

前記左右チャンネル対応部が、前記知覚システムに対して音の方向付けに関する情報を出力し、

前記左右チャンネル対応部が、聴覚用のエピソード幾何により音の方向情報を計算して音の方向付けを行なうことを特徴とする、請求の範囲第 6 項から第 9 項のいずれかに記載のロボット聴覚システム。

1 9. 視覚、触覚等の知覚システムをさらに備えており、前記左右チャンネル対応部が、これらの知覚システムからの情報及び駆動機構の制御信号を参照して音の方向付け及び画像との対応付けを行ない、

前記左右チャンネル対応部が、前記知覚システムに対して音の方向付けに関する情報を出力し、

前記処理部が、内外マイクの強度差がテンプレートの駆動機構のノイズの強度差と近く、内外マイクの入力音のスペクトルの強度とパターンとがテンプレートの駆動機構のノイズ周波数応答に近く、さらに駆動機構が動作しているときに、ノイズをバーストノイズとしてこの帯域の信号部分を除去し、

前記左右チャンネル対応部が、聴覚用のエピポーラ幾何により音の方向情報を計算して音の方向付けを行なうことを特徴とする、請求の範囲第6項から第9項のいずれかに記載のロボット聴覚システム。

20. 前記処理部が、外部マイク及び内部マイクからの音響信号のスペクトルパワー差のパターンが前もって測定された駆動機構のノイズのスペクトルパワー差のパターンとほぼ同じであり、スペクトルの音圧及びパターンが前もって測定された駆動機構のノイズの周波数応答とほぼ同じであって、駆動機構の制御信号により駆動機構の連続動作を検出したとき、バーストノイズとしてこの帯域の信号部分を除去し、

前記左右チャンネル対応部が、聴覚用のエピポーラ幾何により音の方向情報を計算して音の方向付けを行なうことを特徴とする、請求の範囲第8項又は第9項に記載のロボット聴覚システム。

21. 前記左右チャンネル対応部が、聴覚用のエピポーラ幾何に基づいた方法により音の方向情報を計算して音の方向付けを行い、

調波構造を有する音に関しては、調波構造を利用した音源分離と音の強度差情報とを利用して音の方向付けを行なうことを特徴とする、請求の範囲第6項から第9項のいずれかに記載のロボット聴覚システム。

22. 視覚、触覚等の知覚システムをさらに備えており、前記左右チャンネル対応部が、これらの知覚システムからの情報及び駆動機構の制御信号を参照して音の方向付け及び画像との対応付けを行ない、

前記左右チャンネル対応部が、聴覚用のエピポーラ幾何に基づいた方法により音の方向情報を計算して音の方向付けを行い、調波構造を有する音に関して

は、調波構造を利用した音源分離と音の強度差情報とを利用して音の方向付けを行なうことを特徴とする、請求の範囲第 6 項から第 9 項のいずれかに記載のロボット聴覚システム。

- 2 3. 視覚、触覚等の知覚システムをさらに備えており、前記左右チャンネル対応部が、これらの知覚システムからの情報及び駆動機構の制御信号を参照して音の方向付け及び画像との対応付けを行ない、

前記左右チャンネル対応部が、前記知覚システムに対して音の方向付けに関する情報を出力し、

前記左右チャンネル対応部が、聴覚用のエピポーラ幾何に基づいた方法により音の方向情報を計算して音の方向付けを行い、調波構造を有する音に関しては、調波構造を利用した音源分離と音の強度差情報とを利用して音の方向付けを行なうことを特徴とする、請求の範囲第 6 項から第 9 項のいずれかに記載のロボット聴覚システム。

- 2 4. 視覚、触覚等の知覚システムをさらに備えており、前記左右チャンネル対応部が、これらの知覚システムからの情報及び駆動機構の制御信号を参照して音の方向付け及び画像との対応付けを行ない、

前記左右チャンネル対応部が、前記知覚システムに対して音の方向付けに関する情報を出力し、

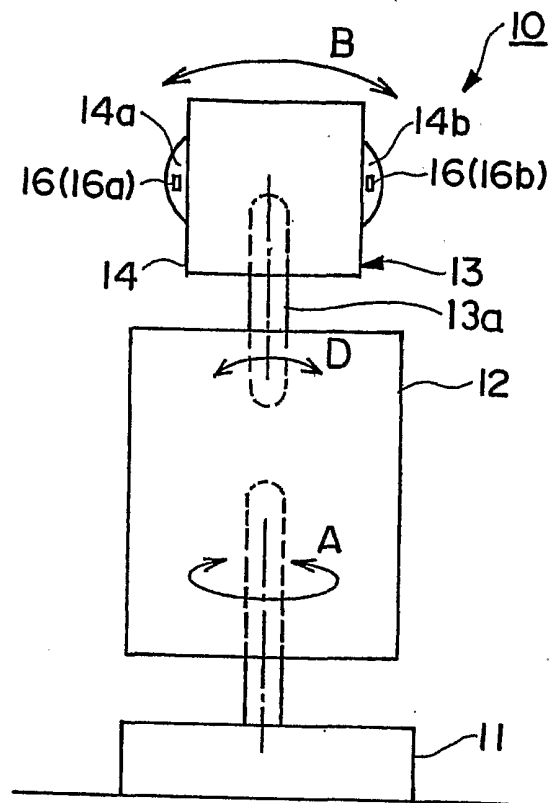
前記処理部が、内外マイクの強度差がテンプレートの駆動機構のノイズの強度差と近く、内外マイクの入力音のスペクトルの強度とパターンとがテンプレートの駆動機構のノイズ周波数応答に近く、さらに駆動機構が動作しているときに、ノイズをバーストノイズとしてこの帯域の信号部分を除去し、

前記左右チャンネル対応部が、聴覚用のエピポーラ幾何に基づいた方法により音の方向情報を計算して音の方向付けを行い、調波構造を有する音に関しては、調波構造を利用した音源分離と音の強度差情報とを利用して音の方向付けを行なうことを特徴とする、請求の範囲第 6 項から第 9 項のいずれかに記載のロボット聴覚システム。

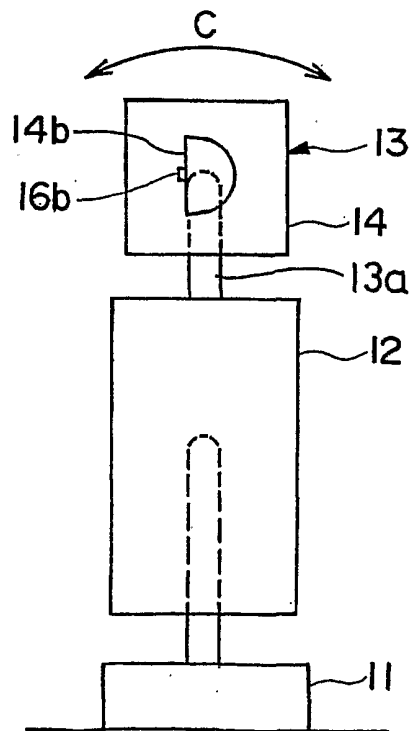
25. 前記処理部が、外部マイク及び内部マイクからの音響信号のスペクトルパワー差のパターンが前もって測定された駆動機構のノイズのスペクトルパワー差のパターンとほぼ同じであり、スペクトルの音圧及びパターンが前もって測定された駆動機構のノイズの周波数応答とほぼ同じであって、駆動機構の制御信号により駆動機構の連続動作を検出したとき、バーストノイズとしてこの帯域の信号部分を除去し、

前記左右チャンネル対応部が、聴覚用のエピソード幾何に基づいた方法により音の方向情報を計算して音の方向付けを行い、調波構造を有する音に関しては、調波構造を利用した音源分離と音の強度差情報とを利用して音の方向付けを行なうことを特徴とする、請求の範囲第8項又は第9項に記載のロボット聴覚システム。

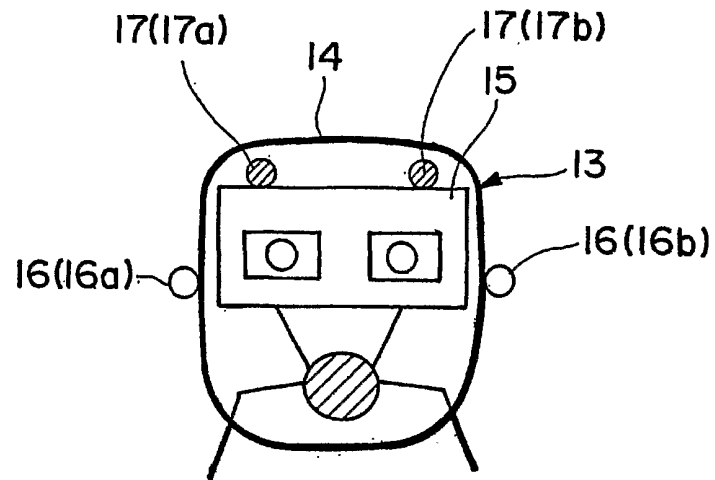
第 1 図



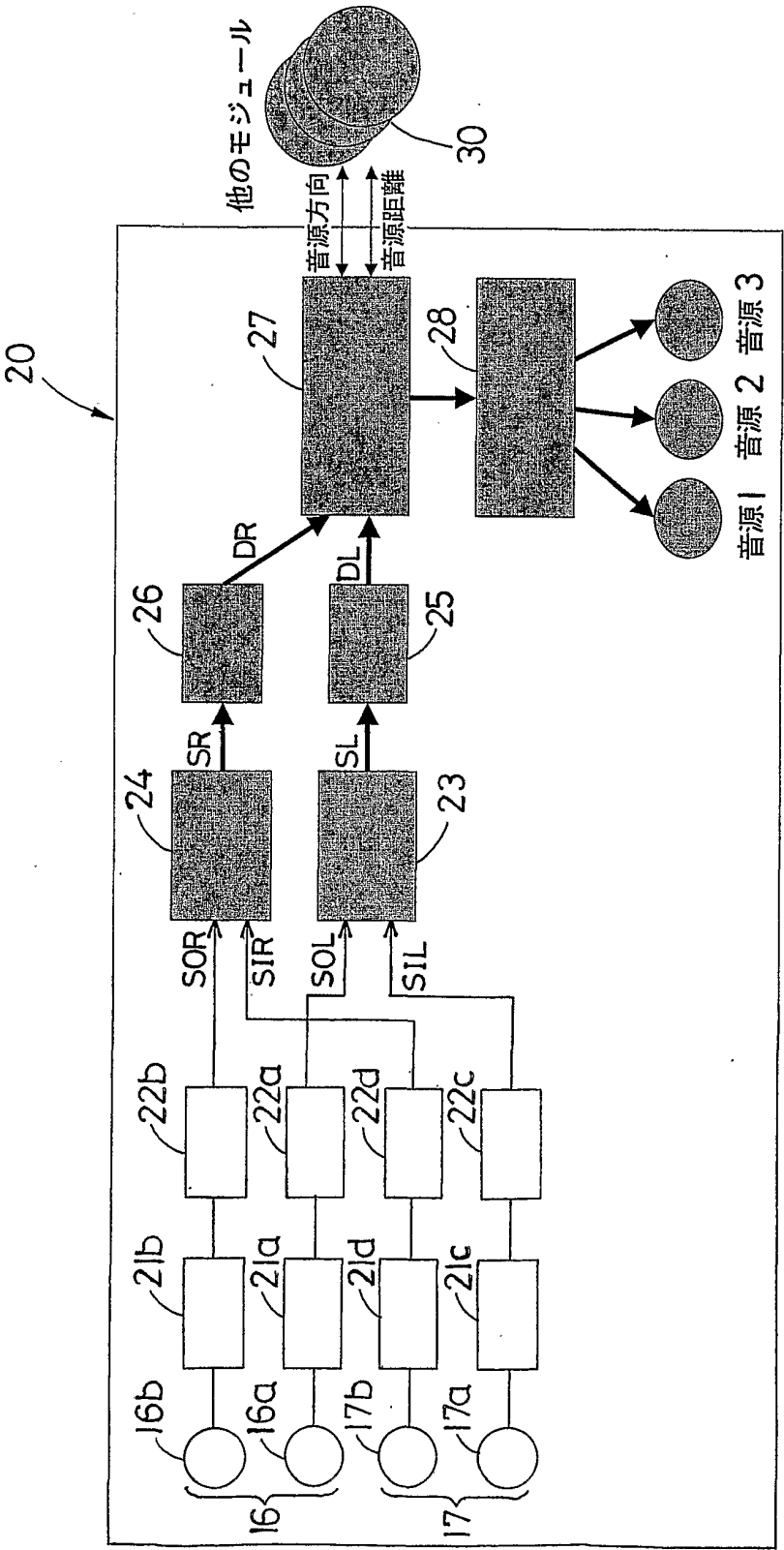
第 2 図



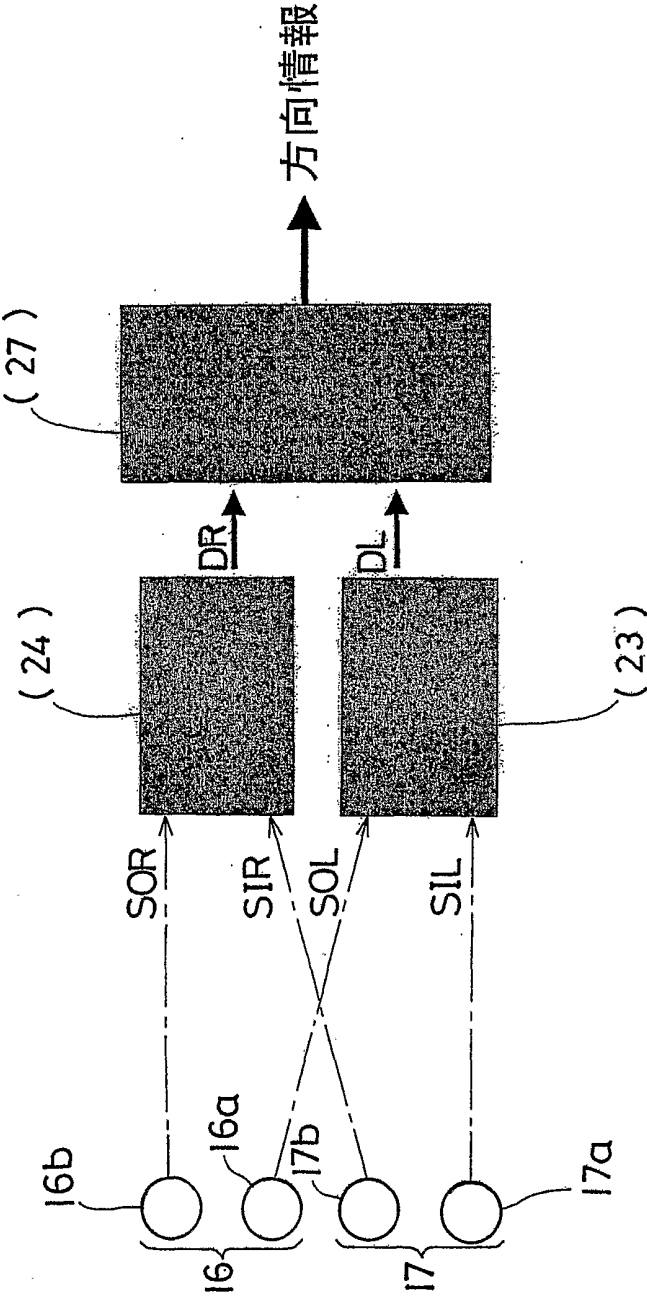
第 3 図



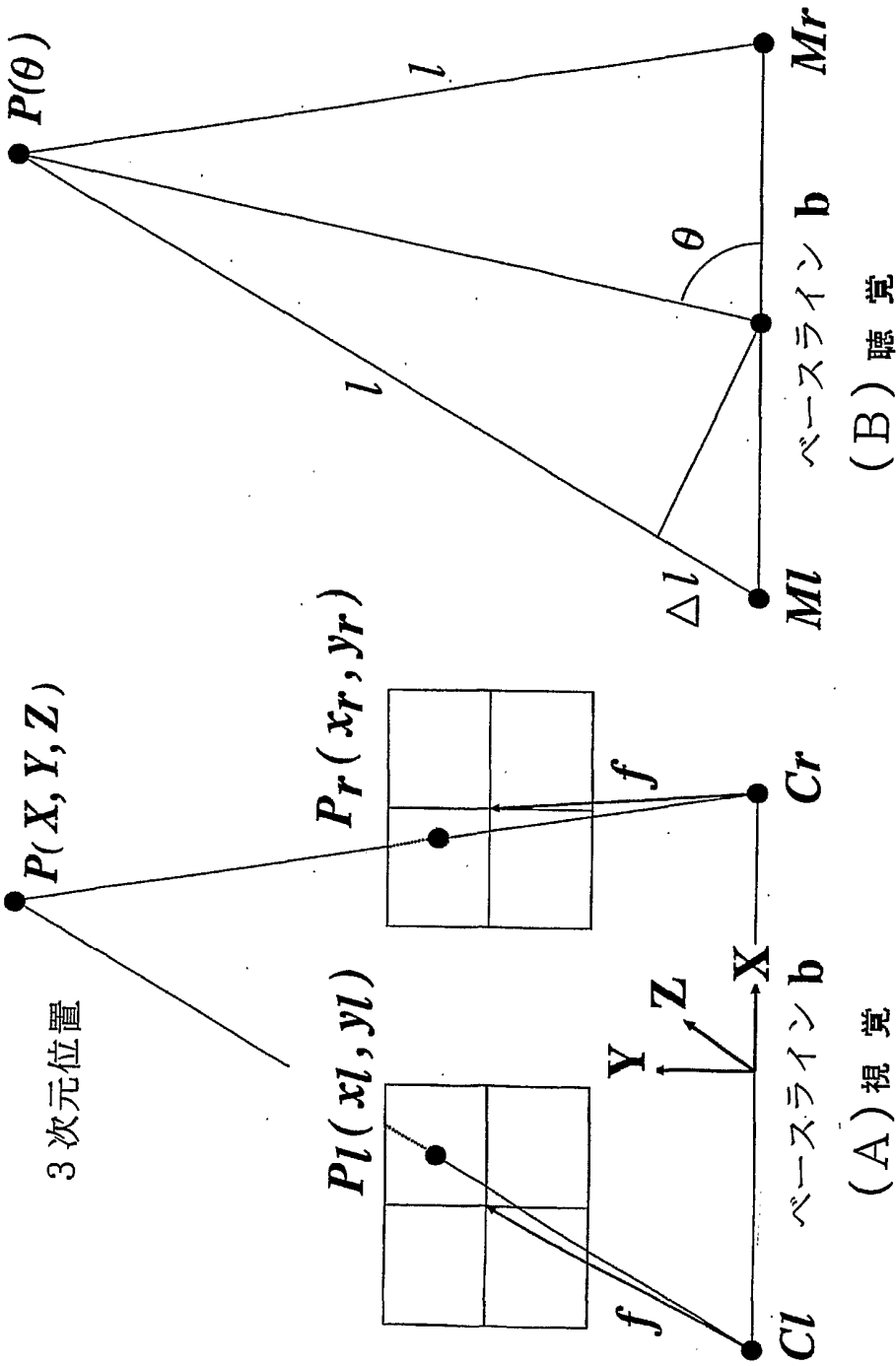
第 4 図



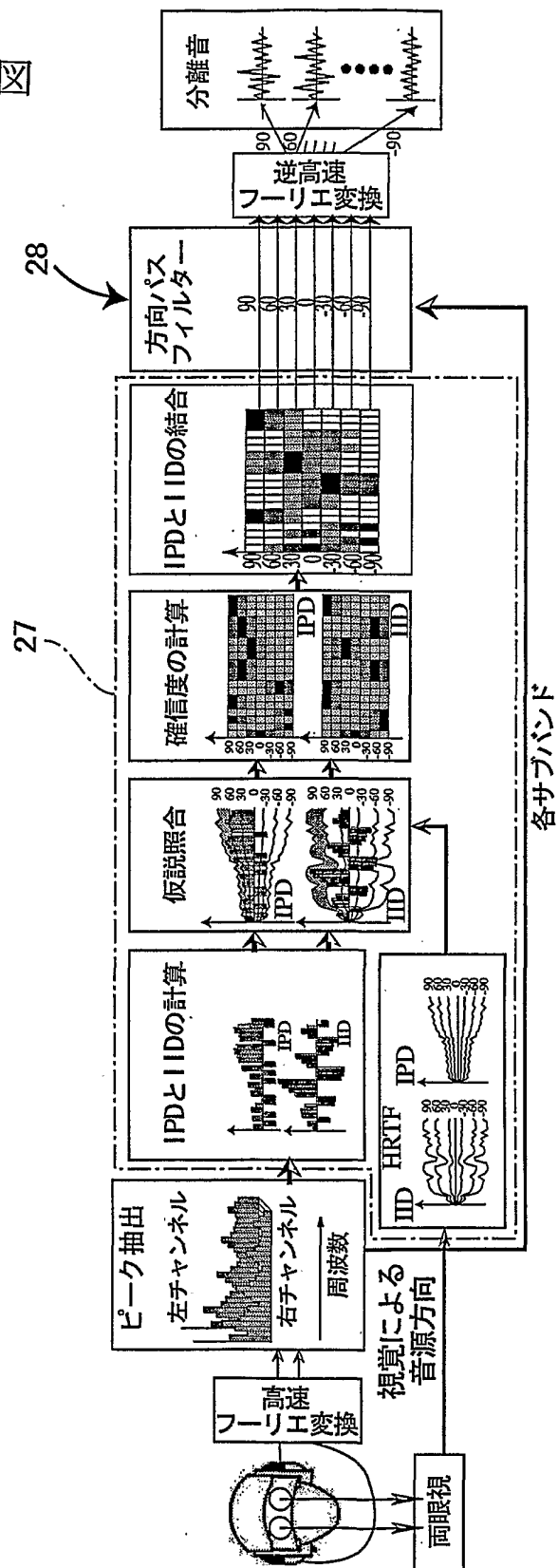
第 5 図



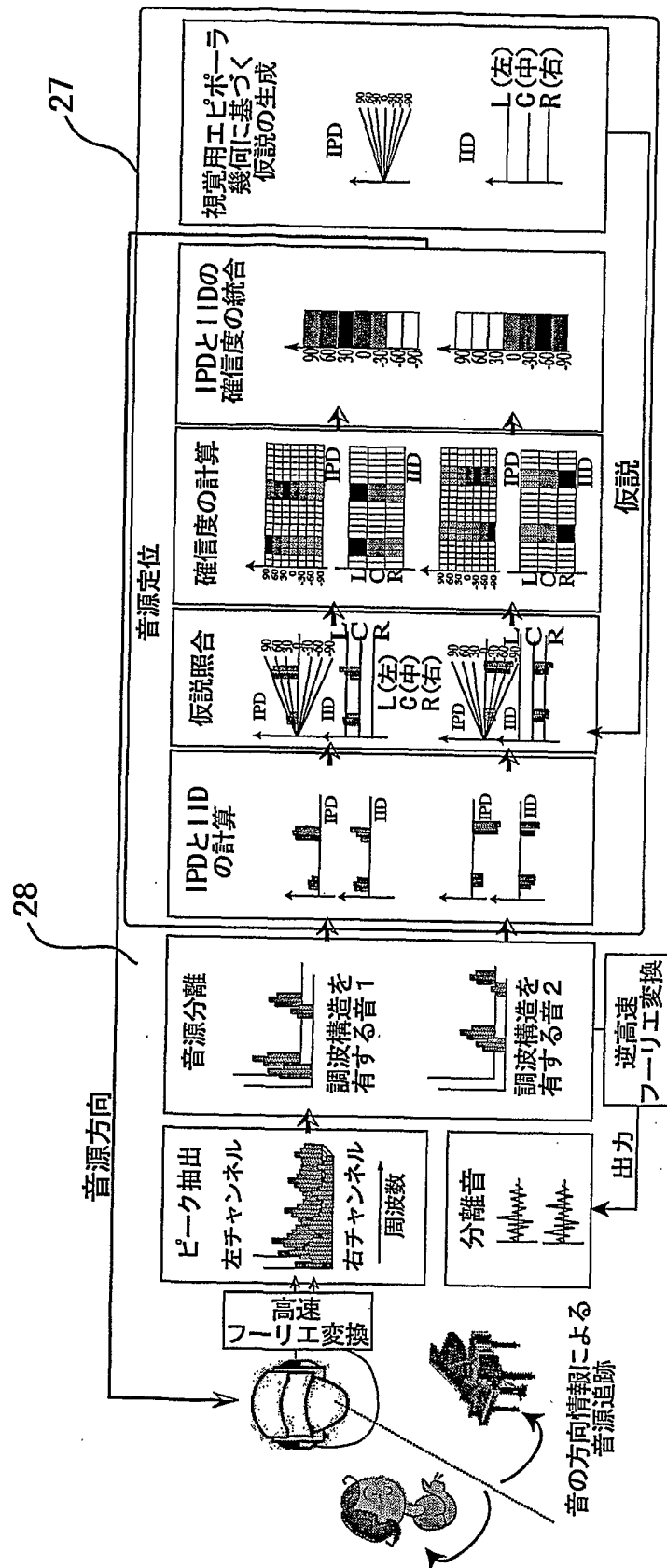
第 6 図



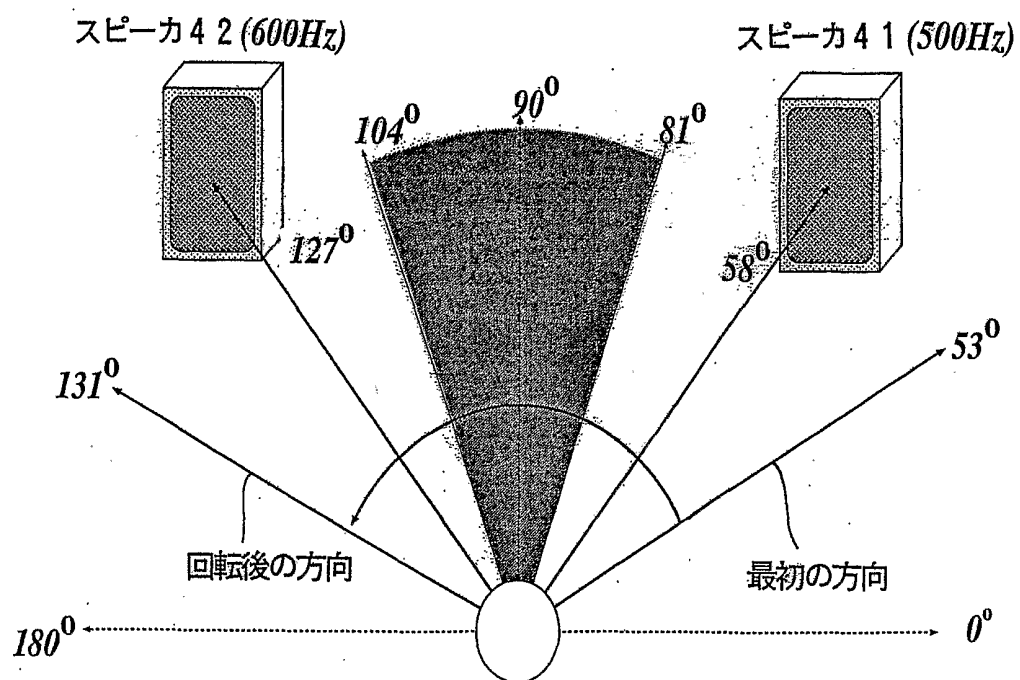
第7図



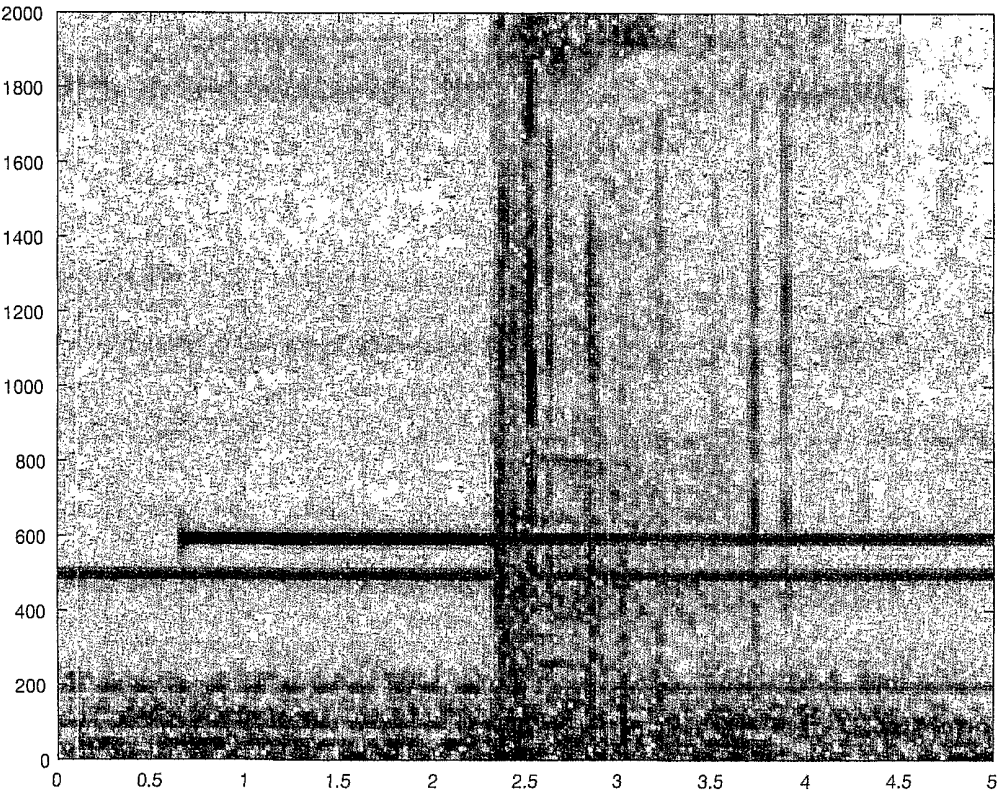
第 8 図



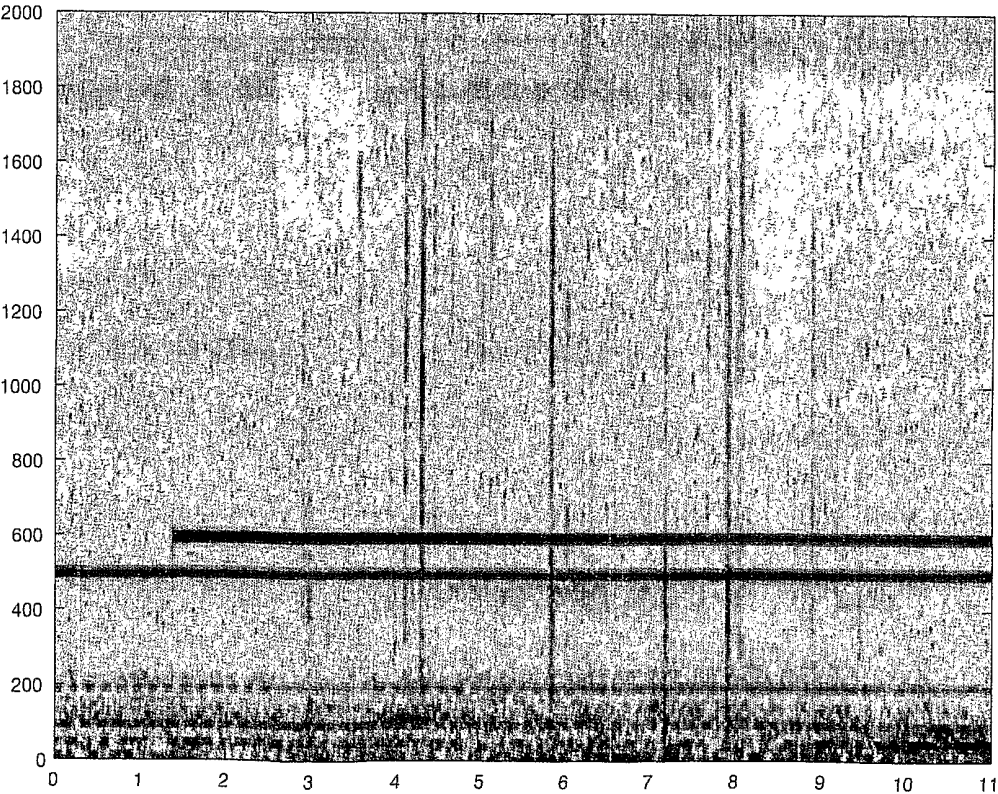
第 9 図



第 1 0 図 (A)

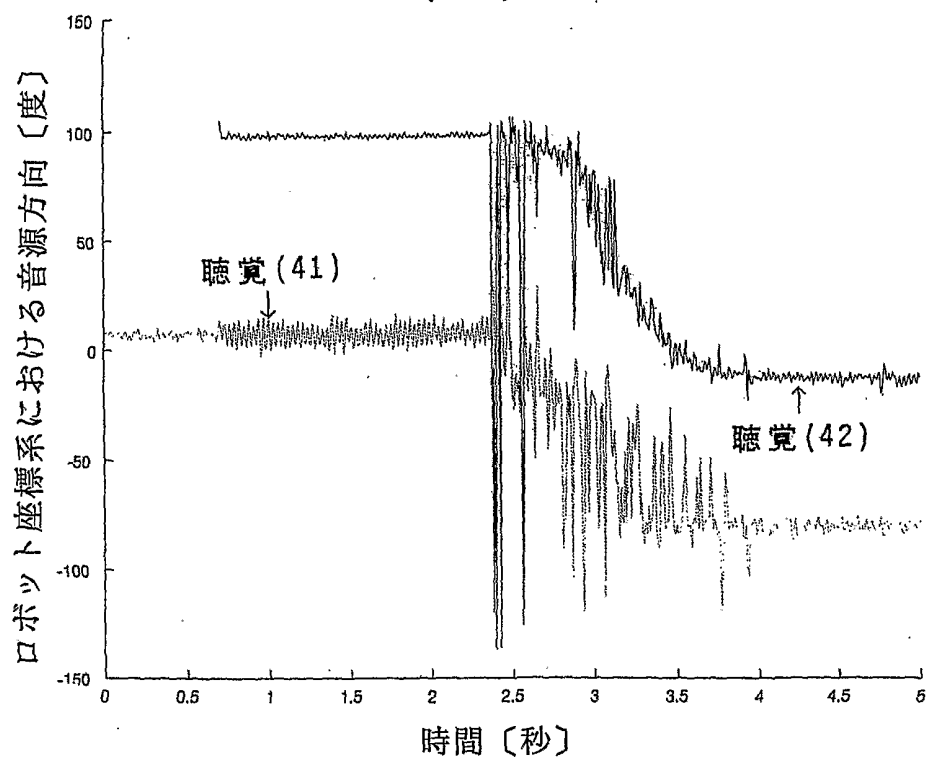


(B)

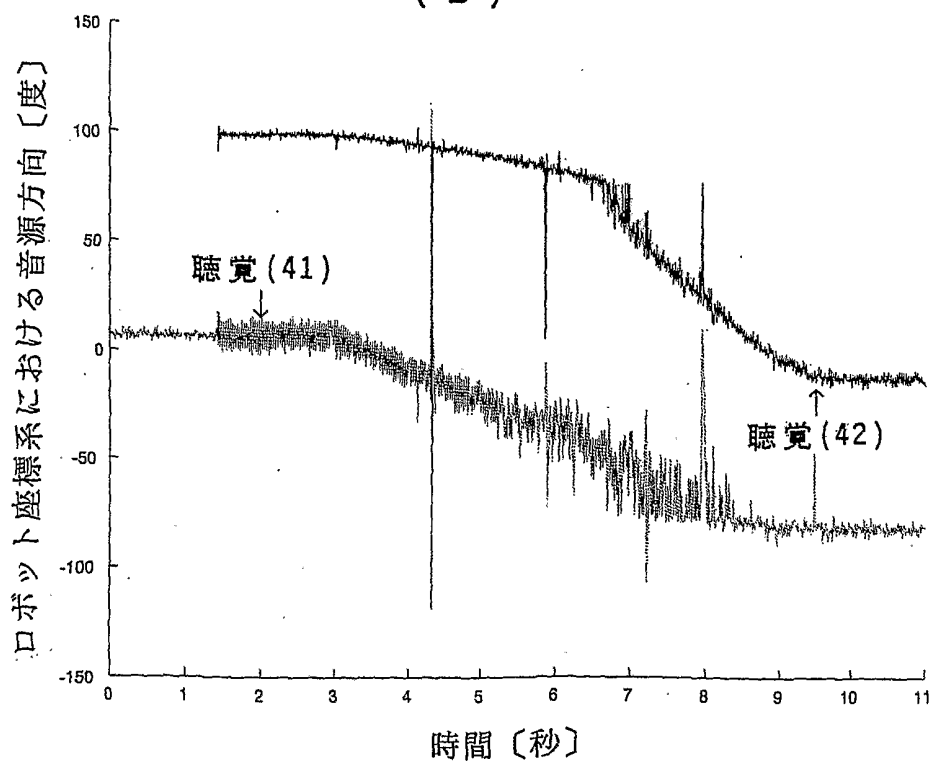


第 1 1 図

(A)

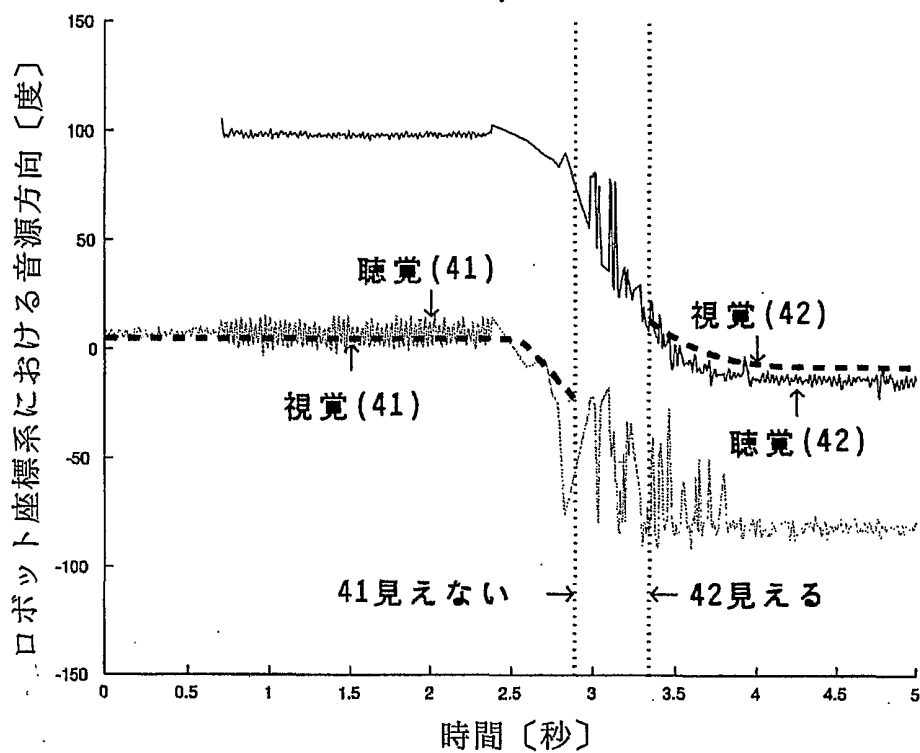


(B)

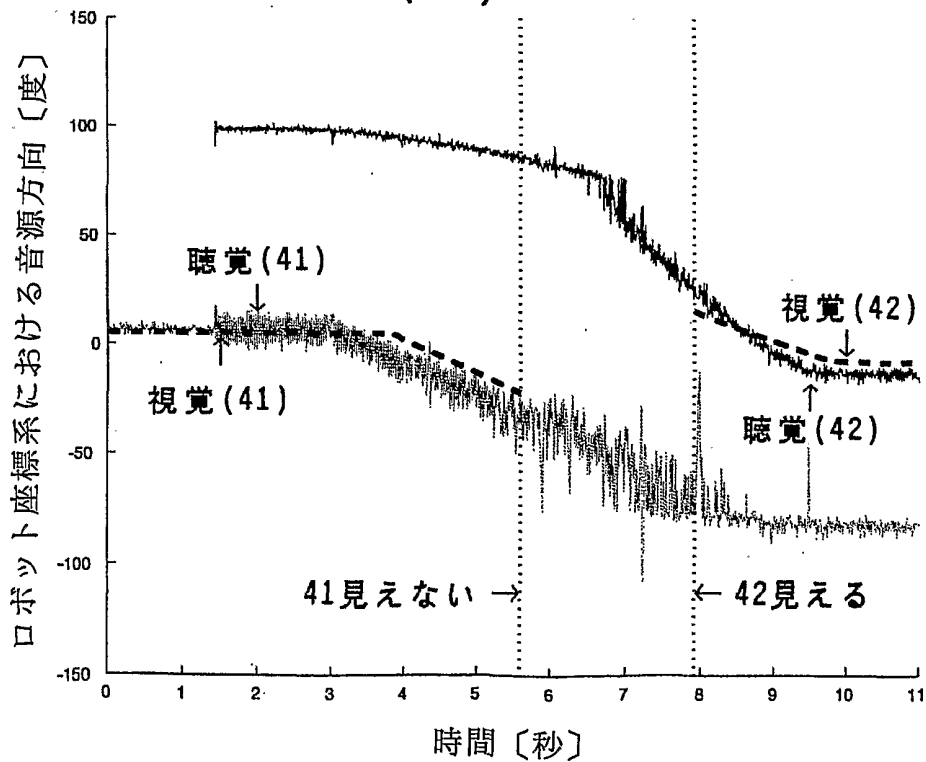


第 1 2 図

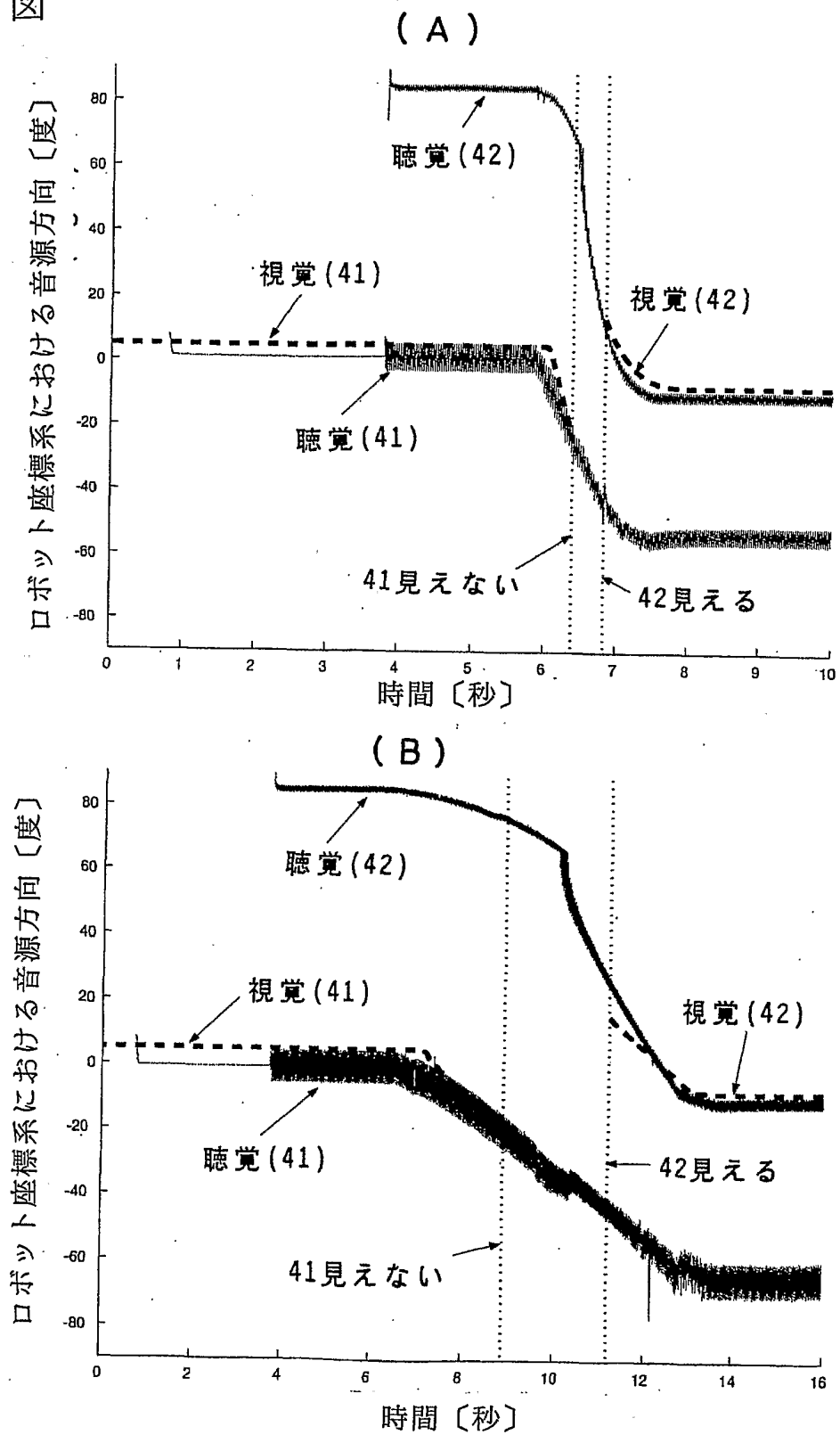
(A)



(B)

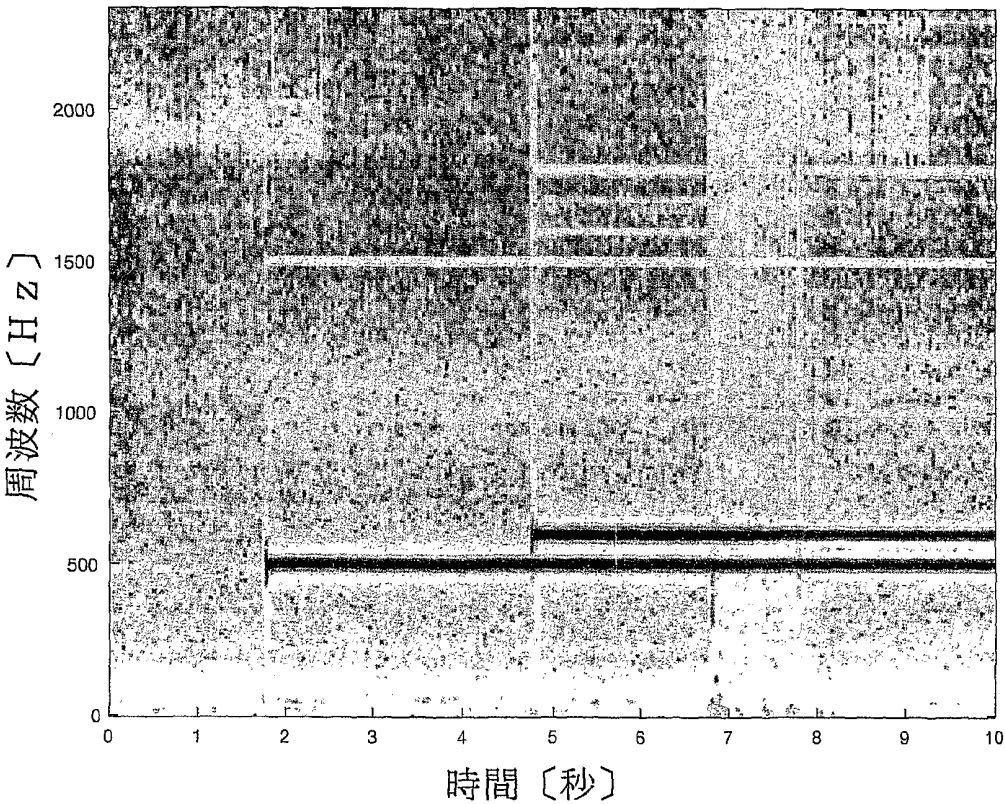


第 1 3 図

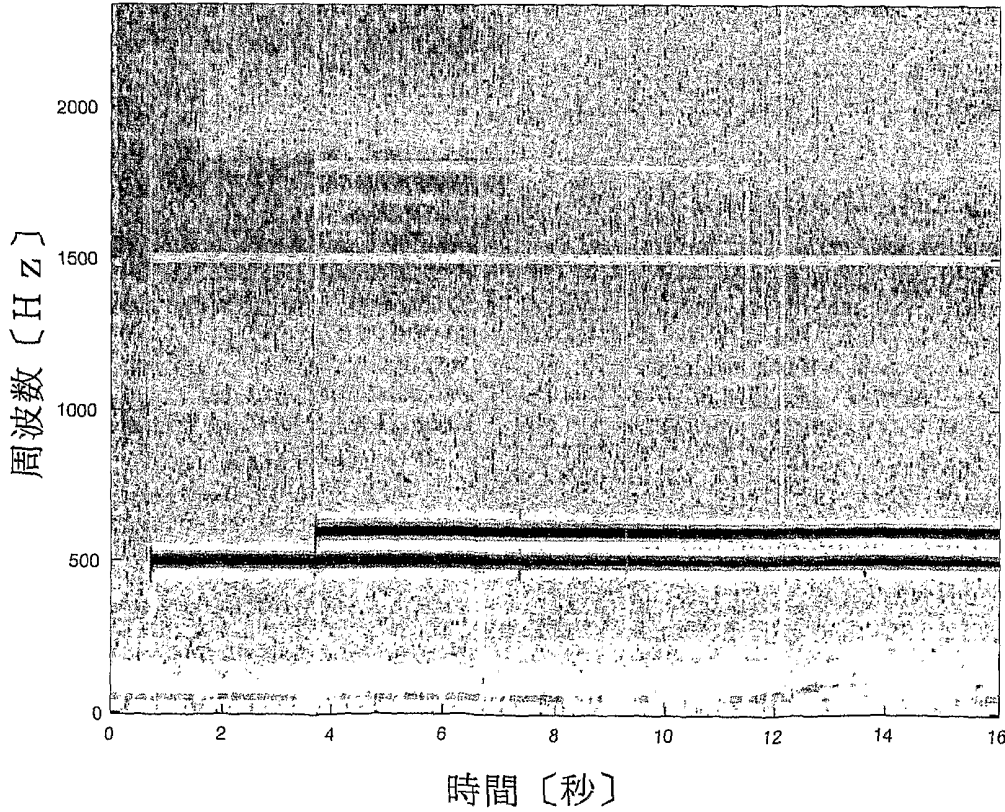


第 1 4 図

(A)

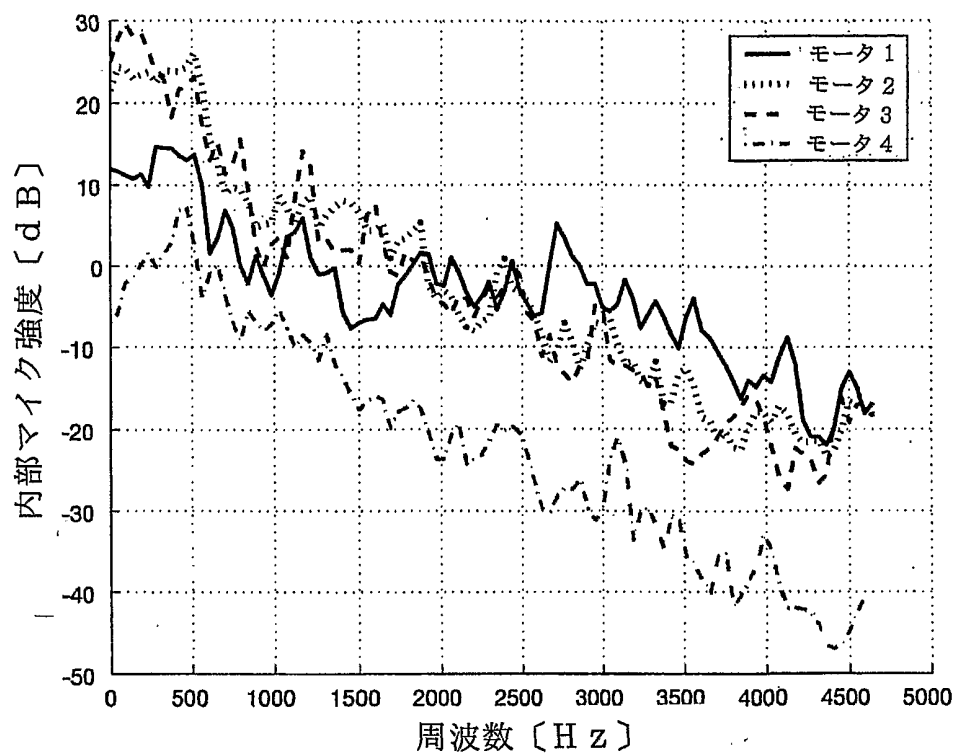


(B)

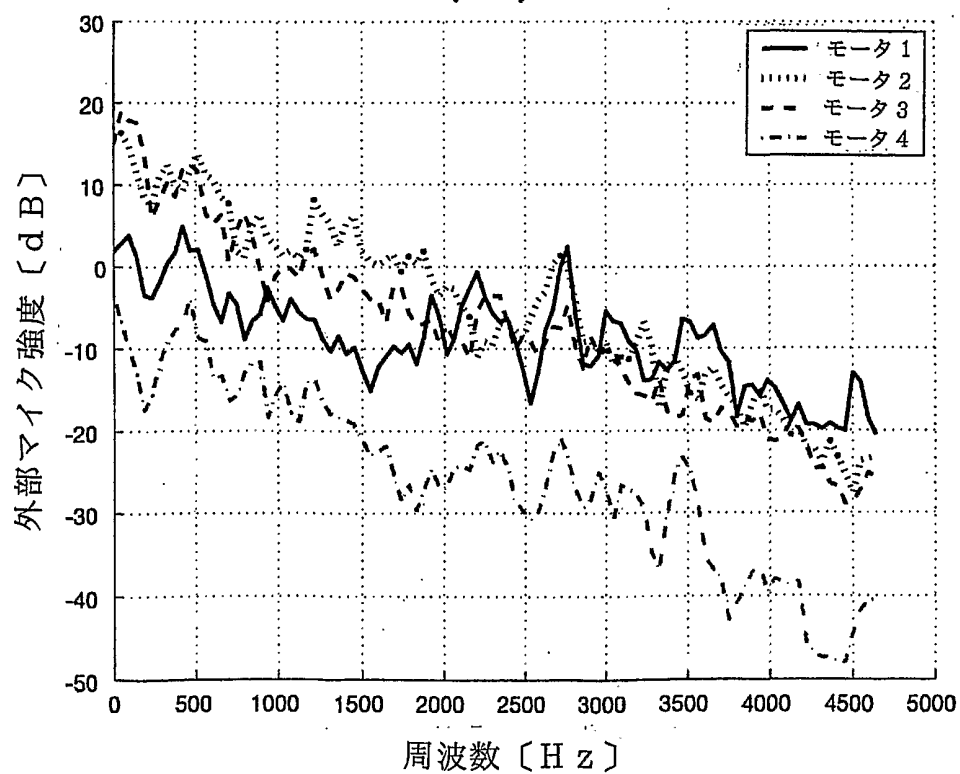


第 1 5 図

(A)

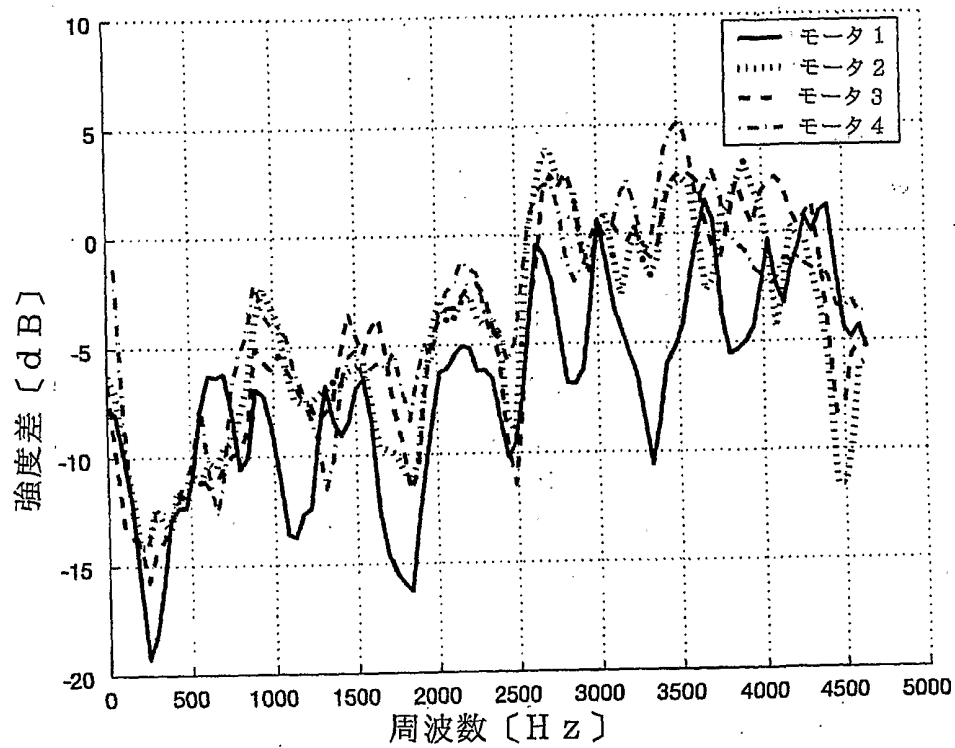


(B)

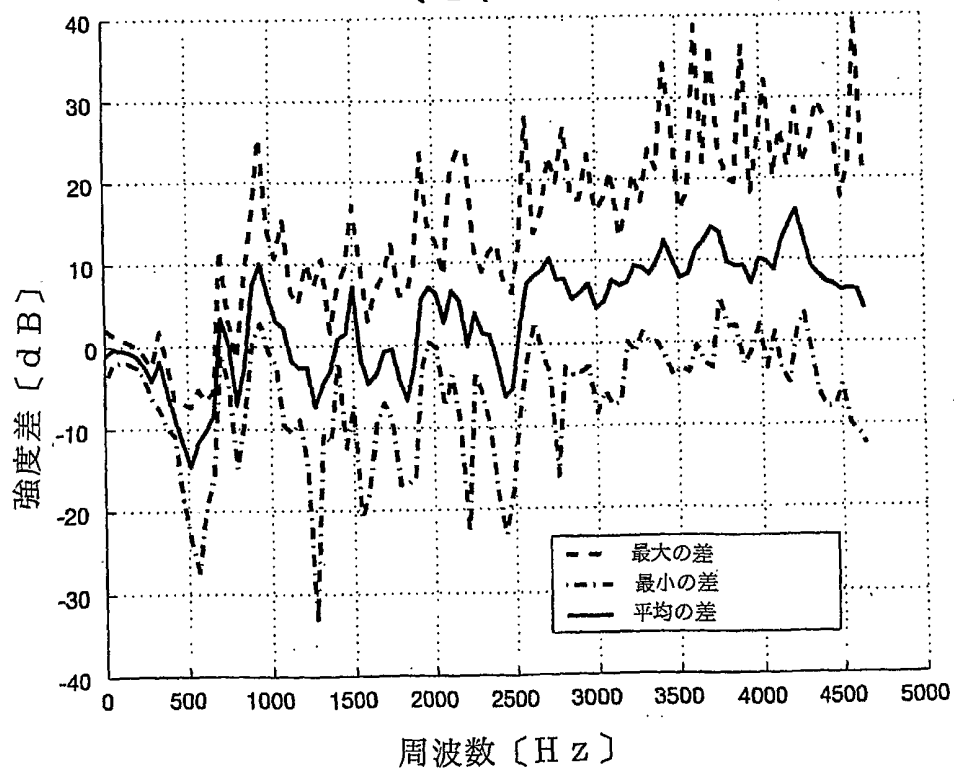


第 16 図

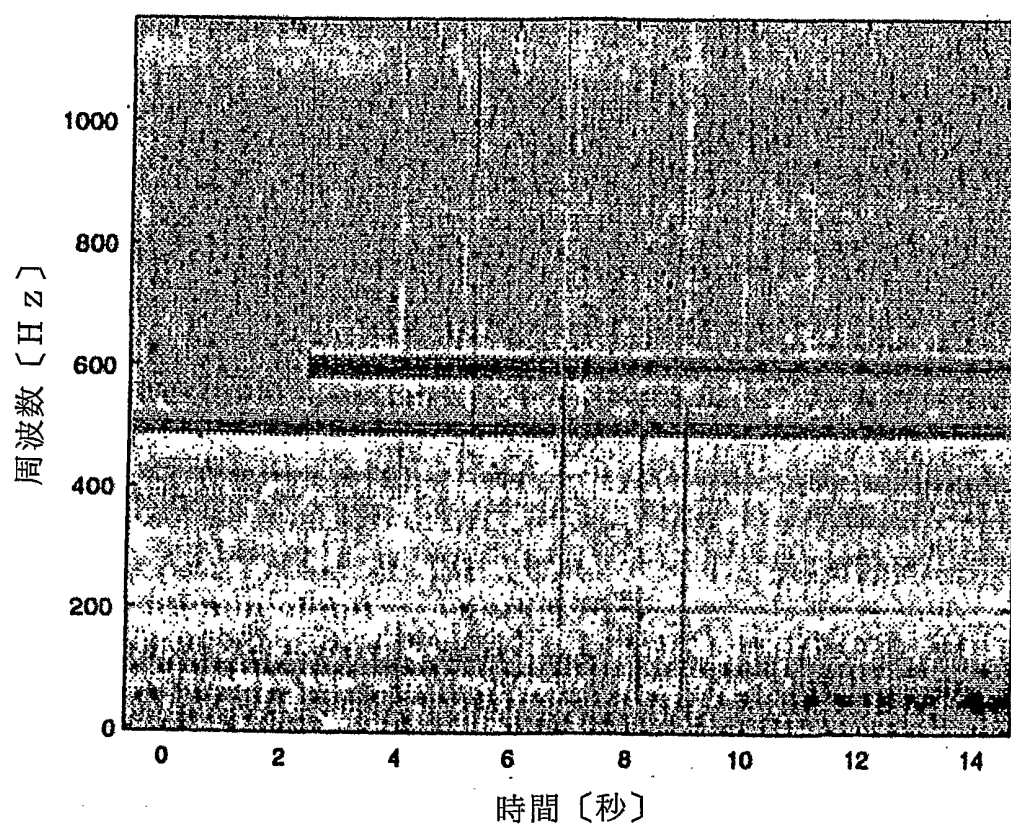
(A)



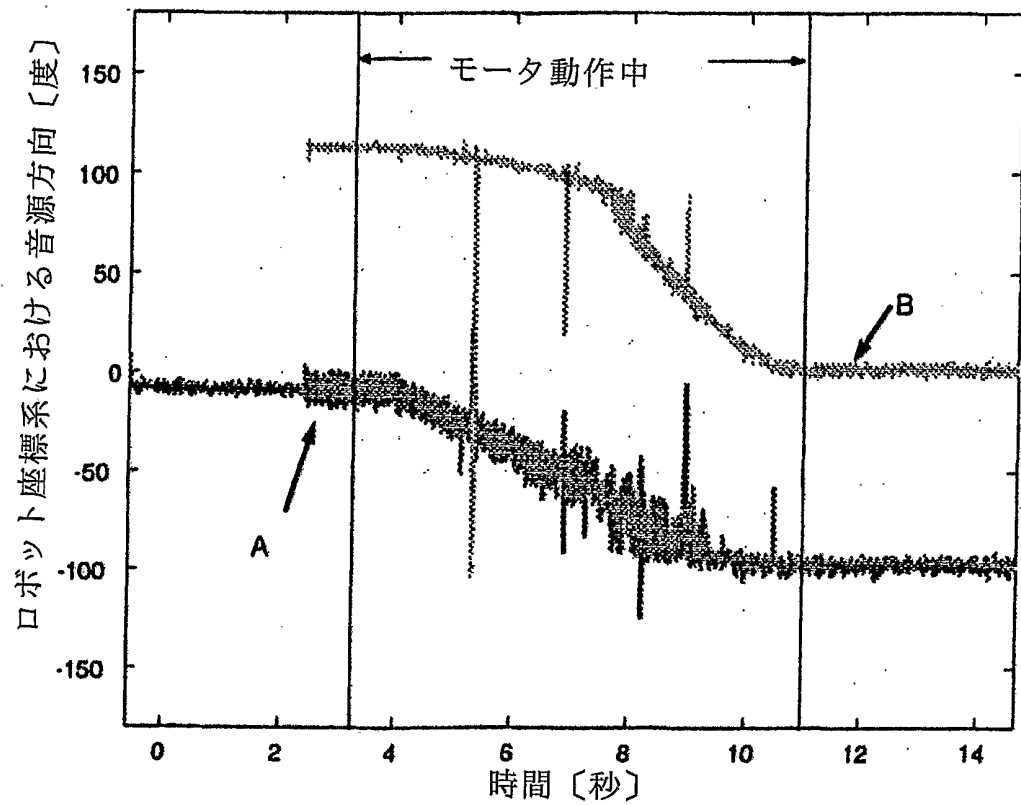
(B)



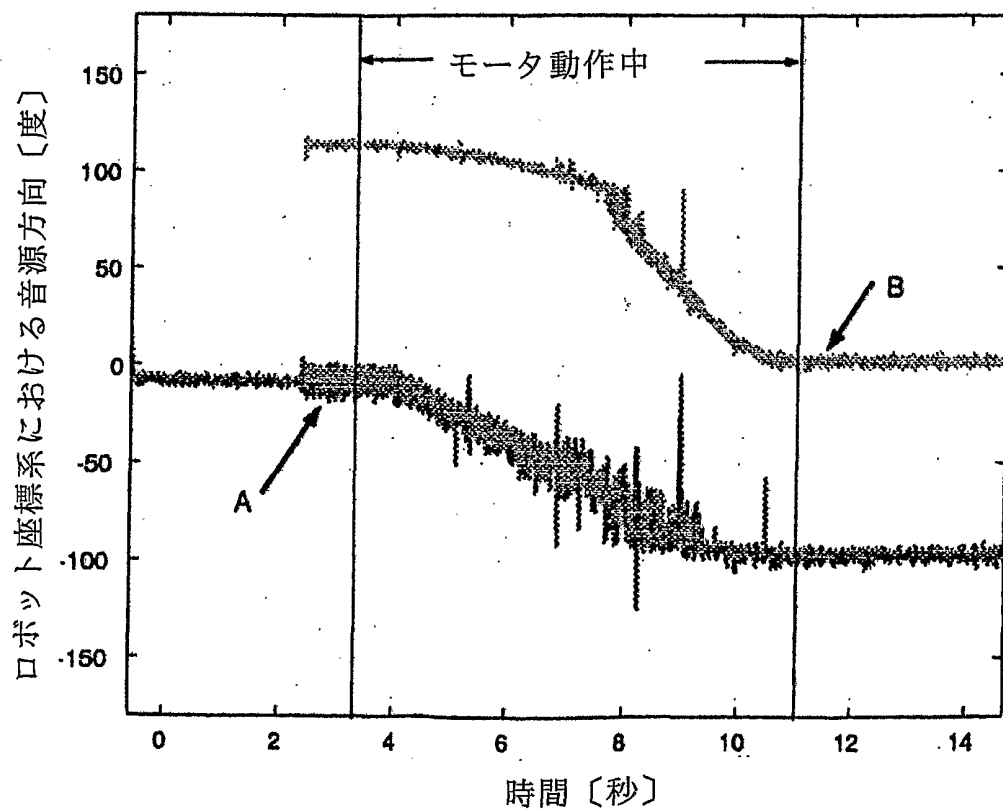
第 1 7 図



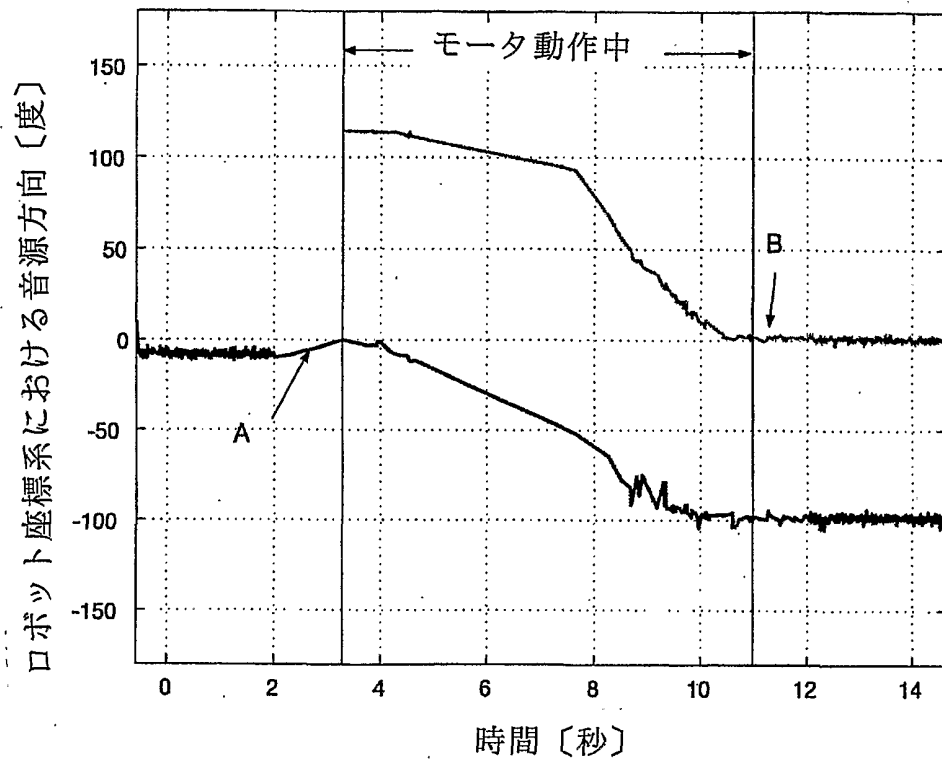
第 1 8 図



第 1 9 図



第 2 0 図



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP01/04858

A. CLASSIFICATION OF SUBJECT MATTER Int.Cl ⁷ G10L21/02, 15/20, B25J13/00		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) Int.Cl ⁷ G10L21/02, 15/20		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Jitsuyo Shinan Koho 1922-1996 Toroku Jitsuyo Shinan Koho 1994-2001 Kokai Jitsuyo Shinan Koho 1971-2001 Jitsuyo Shinan Toroku Koho 1996-2001		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) JICST FILE on Science and Technology (JOIS)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JSAI Technical Report, Proceedings of the Seventh Meeting of Special Interest Group on AI Challenges, SIG-Challenge-9907-10, H. G. OKUNO et al., "Research Issues of Humanoid Audition", pages 61 to 65, 02 November, 1999 (02.11.99), Japanese Society for Artificial Intelligence	1-2, 6-12
Y	IEICE Technical Report, Vol.98, No.534, DSP98-164, T. KIKUCHI, et al., "Microphone Array System with DOA Estimation by Using Harmonic Structure of Speech Signals", pages 23 to 28, 22 January, 1999 (22.01.99), The Institute of Electronics, Information and Communications Engineers	1-2, 6-12
Y	JP 11-41577 A (Fujitsu Limited), 12 February, 1999 (12.02.99), Full text; all drawings (Family: none)	10-11
A	The Heisei-7 Spring Meeting of the Acoustical Society of Japan, Vol.1, 1-5-8, S. NAKAMURA, et al., "Speech Recognition with Source Localization by Microphone Array", pages 15 to 16, 14 March, 1995 (14.03.95), The Acoustical Society of Japan	1-25
☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 24 July, 2001 (24.07.01)		Date of mailing of the international search report 07 August, 2001 (07.08.01)
Name and mailing address of the ISA/ Japanese Patent Office		Authorized officer
Facsimile No.		Telephone No.

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl⁷ G10L21/02, 15/20, B25J13/00

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl⁷ G10L21/02, 15/20

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922~1996年

日本国公開実用新案公報 1971~2001年

日本国登録実用新案公報 1994~2001年

日本国実用新案登録公報 1996~2001年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

JICST科学技術文献ファイル (JOIS)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	人工知能学会研究会資料 (JSAT Technical Report), AIチャレンジ研究会 (第7回) (Proceedings of the Seventh Meeting of Special Interest Group on AI Challenges), SIG-Challenge-9907-10, H.G.Okuno et al., "Research Issues of Humanoid Audition", p.61-65, 2.11月.1999 (02.11.99), 社団法人 人人工知能学会 (Japanese Society for Artificial Intelligence)	1-2, 6-12

☐ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」 特に関連のある文献ではなく、一般的技術水準を示すもの

「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)

「O」 口頭による開示、使用、展示等に言及する文献

「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&」 同一パテントファミリー文献

国際調査を完了した日

24.07.01

国際調査報告の発送日

07.08.01

国際調査機関の名称及びあて先

日本国特許庁 (ISA/JP)

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

榎本 剛

5C

9379

電話番号 03-3581-1101 内線 3541

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	電子情報通信学会技術研究報告("IEICE Technical Report") [ディジタル信号処理], Vol. 98, No. 534, DSP98-164, 菊池隆文外("T. Kikuchi et al."), 「音声の調波構造を方向推定に適用したマイクロフォンアレイシステム」 ("Microphone Array System with DOA Estimation by Using Harmonic Structure of Speech Signals"), p. 23-28, 22. 1月. 1999 (22. 01. 99), 社団法人電子情報通信学会(The Institute of Electronics, Information and Communications Engineers)	1-2, 6-12
Y	JP11-41577 A (富士通株式会社) 12. 2月. 1999 (12. 02. 99) 全文, 全図 (ファミリーなし)	10-11
A	日本音響学会平成7年度春季研究発表会講演論文集 I ("The Heisei-7 Spring Meeting of the Acoustical Society of Japan, Vol. 1"), 1-5-8, 中村哲外("S. Nakamura et al."), 「マイクロホンアレーを用いた音源方向検出による音声認識」 ("Speech Recognition with Source Localization by Microphone Array"), p. 15-16, 14. 3月. 1995 (14. 03. 95), 社団法人日本音響学会("The Acoustical Society of Japan")	1-25