(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2009 年4 月9 日 (09.04.2009)

PCT

(10) 国際公開番号
WO 2009/044509 A1

(51) 国際特許分類:
G01S 3/808 (2006.01)

(21) 国際出願番号: PCT/JP2008/002497

(22) 国際出願日: 2008 年9 月10 日 (10.09.2008)

(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2007-258084 2007 年10 月1 日 (01.10.2007) JP

(71) 出願人 (米国を除く全ての指定国について): パナソニック株式会社 (PANASONIC CORPORATION) [JP/JP]; 5718501 大阪府門真市大字門真１００６番地 Osaka (JP).

(72) 発明者; および
(75) 発明者/出願人 (米国についてのみ): 芳澤伸一 (YOSHIZAWA, Shinichi). 中藤良久 (NAKATOH, Yoshihisa).

(74) 代理人: 新居広守 (NII, Hiromori); 〒5320011 大阪府大阪市淀川区西中島５丁目３番１０号タナカ・イトーピア新大阪ビル６階新居国際特許事務所内 Osaka (JP).

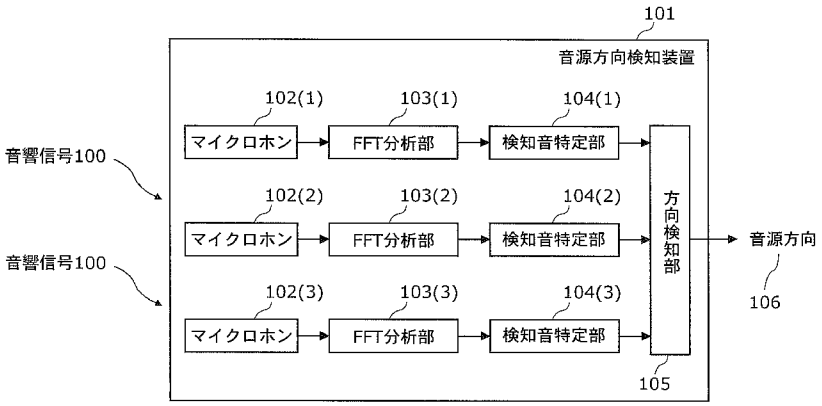
(81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE,

[続葉有]

(54) Title: SOUNND SOURCE DIRECTION DETECTOR

(54) 発明の名称: 音源方向検知装置

[図2]



101 SOUND SOURCE DIRECTION DETECTOR
100 ACOUSTIC SIGNAL
102(1) MICROPHONE
102(2) MICROPHONE
102(3) MICROPHONE
103(1) FFT ANALYSIS SECTION
103(2) FFT ANALYSIS SECTION
103(3) FFT ANALYSIS SECTION
104(1) DETECTION SOUND IDENTIFYING SECTION
104(2) DETECTION SOUND IDENTIFYING SECTION
104(3) DETECTION SOUND IDENTIFYING SECTION
105 DIRECTION DETECTING SECTION
106 SOUND SOURCE DIRECTION

(57) Abstract: A sound source direction detector comprises FFT analysis sections (103(1) to 103(3)) for generating a frequency spectrum in at least one frequency band of acoustic signals for each of the acoustic signals collected by two or more microphones arranged apart from one another, detection sound identifying sections (104(1) to 104(3)) for identifying a time portion of the frequency spectrum of a detection sound which obtains a sound source direction from the frequency spectrum in the frequency band, and a direction detecting section (105) for obtaining the difference between the times at which the detection sound reaches the microphones, obtaining the sound source direction from the time difference, the distance between the microphones, and the sound velocity , and outputting it depending on the degree of coincidence between the microphones of the frequency spectrum in the time portion identified by the detection sound identifying sections (104(1) to 104(3)) in a time interval which is the time unit to detect the sound source direction.

[続葉有]

WO 2009/044509 A1

SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:
— 国際調査報告書

(57) 要約: 互いに離間して配置された２つ以上のマイクロホンにて集音された音響信号の各々について、当該音響信号の少なくとも１つの周波数帯域における周波数スペクトルを生成するＦＦＴ分析部（１０３（１）～１０３（３））と、前記周波数帯域における周波数スペクトルから、音源方向を求める検知音の周波数スペクトルの時刻部分を特定する検知音特定部（１０４（１）～１０４（３））と、音源方向を検知する時間単位である時間区間の中で、前記検知音特定部（１０４（１）～１０４（３））が特定した時刻部分における周波数スペクトルの前記マイクロホン間での一致度合いにより、前記検知音が前記マイクロホンに到達する時間差を求めて、前記時間差と前記マイクロホンの距離と音速とから音源方向を求めて出力する方向検知部（１０５）とを備える。

明　細　書

音源方向検知装置

技術分野

[0001]　本発明は、雑音下において音源の方向を特定する音源方向検知装置に関し、特に、雑音の影響を軽減して、さらに、雑音除去の副作用により誤った方向を出力してしまうことを回避できる音源方向検知装置に関する。

背景技術

[0002]　従来の雑音除去方法として、検知音と雑音を含む音響信号の周波数スペクトルから、雑音の影響が少ない検知音の周波数スペクトルの時刻部分を特定するものがある（例えば、非特許文献１参照）。

[0003]　雑音の影響が少ない検知音の周波数スペクトルの時刻部分を、検知音（Ｓ）と雑音（Ｎ）との周波数スペクトルのパワーの比率であるＳＮ比が０ｄＢ以上の時刻部分（検知音の周波数スペクトルのパワー≧雑音の周波数スペクトルのパワー）により特定する。ここでは、検知音が含まれない時間区間から雑音の周波数スペクトルのパワーを求めて、検知音と雑音とが混合された周波数スペクトルのパワーから雑音の周波数スペクトルのパワーを減算することで検知音の周波数スペクトルのパワーを求めている。この後の処理として、検知音（音声）の認識を行っている。また、検知音の周波数スペクトルの時刻部分を特定する別の方法として、学習用データから入力音が検知音である確率分布と雑音である確率分布を求めて、ベイズ推定により検知音の時刻部分を特定するものもある。ここでは、確率分布の１つの変数として検知音（Ｓ）と雑音（Ｎ）との周波数スペクトルの比率であるＳＮ比や、周波数スペクトルの波形情報が用いられている。このため、ＳＮ比以外の情報も含めて精度良く検知音の周波数スペクトルの時刻部分を特定することができる。

[0004]　従来の音源方向を検知する方法として、互いに離間して配置された２つのマイクロホンにて集音された音響信号の各々について周波数帯域別の信号に

分解して（周波数スペクトルを求めて）、マイクロホン間での周波数帯域別の信号の相互相関（一致度合い）により各々のマイクロホンに到達する時間差を求めて、時間差とマイクロホンの距離とから音源方向を求めるものがある（例えば、特許文献１参照）。

非特許文献1：“Missing-Feature Approaches in Speech Recognition”, Bhiksha Raj and Richard M. Stern, IEEE SIGNAL PROCESSING MAGAZINE, pp.101-pp.116, 2005

特許文献1：特開２００２－６２３４８号公報（請求項１、図１）

発明の開示

発明が解決しようとする課題

[0005] しかしながら、非特許文献１に記載の方法を用いて求められた部分的に時刻部分が欠落した雑音除去後の検知音の周波数スペクトルから、検知音の音源方向を求めることを考える場合に、特許文献１に記載の方法において、相互相関の求め方で、雑音の影響により欠落した時刻部分の扱い方や、雑音の時刻部分を欠落させることで生じる音源方向の結果に与える影響についても新たに考える必要がでてくる。

[0006] このため、本発明は、雑音の影響を軽減して、さらに、雑音除去の副作用により誤った方向を出力してしまうことを回避できる音源方向検知装置を提供することを目的とする。

課題を解決するための手段

[0007] 本発明に係る音源方向検知装置は、互いに離間して配置された２つ以上のマイクロホンにて集音された音響信号の各々について、当該音響信号の少なくとも１つの周波数帯域における周波数スペクトルを生成する生成手段と、前記周波数帯域における周波数スペクトルから、音源方向を求める検知音の周波数スペクトルの時刻部分を特定する検知音特定手段と、音源方向を検知する時間単位である時間区間の中で、前記検知音特定手段が特定した時刻部分における周波数スペクトルの前記マイクロホン間での一致度合いにより、前記検知音が前記マイクロホンに到達する時間差を求めて、前記時間差と前

記マイクロホンの距離と音速とから音源方向を求めて出力する検知手段とを備え、前記検知手段は、前記時間差がゼロになるように前記２つ以上のマイクロホンからの周波数スペクトルの時間軸を調整したときに、前記音源方向を検知する時間単位である時間区間に占める、前記検知音特定手段により特定された時刻部分のうち前記マイクロホン間で共通の時刻に存在する部分の割合が、所定のしきい値より大きい場合に音源方向を出力することを特徴とする。

[0008] この構成によると、検知手段において、相互相関値の算出に用いる所定の時間区間に占める周波数スペクトルがマイクロホン間で共通に存在する時刻部分の割合が所定のしきい値よりも大きい部分のみにおいて音源方向を出力する。すなわち、検知音が２つ以上のマイクロホンに到達する時間差がゼロになるように周波数スペクトルの時間軸を調整したときに、２つ以上のマイクロホンの共通の時刻に存在する検知音の時刻部分における一致度合いにより音源方向を求めて、共通に存在する検知音の時刻部分が多い場合のみに音源方向を出力することで、雑音の影響により誤った方向を出力してしまうことを回避できる。誤った方向を求めてしてしまう原因としては、雑音により除去された時間区間が多くなると音源方向の信頼性がなくなり検知精度が悪くなるためである。

[0009] 好ましくは、前記検知手段は、さらに、前記２つ以上のマイクロホンからの周波数スペクトルの時間軸を調整したときの、前記音源方向を検知する時間単位である時間区間に占める、前記検知音特定手段により特定された時刻部分のうち前記マイクロホン間で共通の時刻に存在する部分の割合が、すべての音源方向の候補について、前記所定のしきい値より大きい場合にのみ音源方向を出力することを特徴とする。

[0010] この構成によると、全ての方向の候補に対して、２つ以上のマイクロホンに共通に存在する時刻部分が多い場合にのみ音源方向を出力することで、上記で、共通に存在する時刻部分が少ないため音源方向を出力できなかった方向に音源が存在した場合でも、他の方向を音源方向として誤って方向を出力

してしまうことを回避できる。

[0011] なお、本発明は、このような特徴的な手段を備える音源方向検知装置として実現することができるだけでなく、音源方向検知装置に含まれる特徴的な手段をステップとする音源方向検知方法として実現したり、音源方向検知方法に含まれる特徴的なステップをコンピュータに実行させるプログラムとして実現したりすることもできる。そして、そのようなプログラムは、ＣＤ－ＲＯＭ（Compact Disc-Read Only Memory）等の記録媒体やインターネット等の通信ネットワークを介して流通させることができるのは言うまでもない。

発明の効果

[0012] 本発明の音源方向検知装置によれば、検知音が２つ以上のマイクロホンに到達する時間差がゼロになるように周波数スペクトルの時間軸を調整したときに、２つ以上のマイクロホンの共通の時刻に存在する検知音の時刻部分における一致度合いにより音源方向を求めて、共通に存在する検知音の時刻部分が多い場合のみに音源方向を出力することで、雑音の影響により誤った方向を出力してしまうことを回避できる。誤った方向を求めてしてしまう原因としては、雑音により除去された時間区間が多くなると音源方向の信頼性がなくなり検知精度が悪くなるためである。さらに、－９０度から＋９０度までの全ての方向に対して、２つ以上のマイクロホンに共通に存在する時刻部分が多い場合にのみ音源方向を出力することで、上記で、共通に存在する時刻部分が少ないため音源方向を出力できなかった方向に音源が存在した場合でも、他の方向を音源方向として誤って方向を出力してしまうことを回避できる。

図面の簡単な説明

[0013] [図1]図１は、本発明の実施の形態１における音源方向検知装置の外観図である。

[図2]図２は、本発明の実施の形態１における音源方向検知装置の全体構成を示したブロック図である。

[図3]図３は、本発明の実施の形態１における音源方向検知装置の検知音特定

部１０４（ｋ）を示したブロック図である。

[図4]図４は、本発明の実施の形態１における音源方向検知装置の検知音特定部１０４（ｋ，ｊ）を示したブロック図である。

[図5]図５は、本発明の実施の形態１における音源方向検知装置の方向検知部１０５を示したブロック図である。

[図6]図６は、本発明の実施の形態１における音源方向検知装置の方向検知部１０５における方向検知部４００（１）を示したブロック図である。

[図7A]図７Ａは、バイク音（検知音）を周波数ｆで周波数分析した結果を模式的に示した図である。

[図7B]図７Ｂは、暗騒音（雑音）を周波数ｆで周波数分析した結果を模式的に示した図である。

[図8]図８は、本発明の実施の形態１における音源方向検知装置の動作手順を示すフローチャートである。

[図9]図９は、本発明の実施の形態１における音源方向検知装置の、検知音の時刻部分を特定するステップＳ６０１（ｋ）の動作手順を示すフローチャートである。

[図10]図１０は、本発明の実施の形態１における音源方向検知装置の、検知音の時刻部分を特定するステップＳ６０１（ｋ）における、検知音の時刻部分を特定するステップＳ６０１（ｋ，ｊ）の動作手順を示すフローチャートである。

[図11]図１１は、本発明の実施の形態１における音源方向検知装置の、音源の方向を特定するステップＳ６０２の動作手順を示すフローチャートである。

[図12]図１２は、本発明の実施の形態１における音源方向検知装置の、音源の方向を特定するステップＳ９００（１）における、音源の方向を特定するステップＳ９００（１，ｊ）の動作手順を示すフローチャートである。

[図13]図１３は、本発明の実施の形態１における音源方向検知装置の、マイクロホンの配置と音響信号１００との位置関係を示した図である。

[図14]図１４は、音響信号１００のスペクトログラムの一例を示した図である。

[図15]図１５は、音響信号１００を作成するときに用いた音声１１００（検知音）のスペクトログラムの一例を示した図である。

[図16]図１６は、周波数スペクトルを特定する方法の一例を説明する図である。

[図17A]図１７Ａは、周波数スペクトルを特定する方法の他の一例を説明する図である。

[図17B]図１７Ｂは、周波数スペクトルを特定する方法の他の一例を説明する図である。

[図18]図１８は、類似度の総和の求め方の一例を説明する図である。

[図19]図１９は、音響信号１００から抽出された音声のスペクトログラムの一例を示した図である。

[図20A]図２０Ａは、類似度の総和を求める時間範囲（所定の時間幅）における、音響信号の周波数スペクトルの位相を模式的に示した図である。

[図20B]図２０Ｂは、類似度の総和を求める時間範囲（所定の時間幅）における、音響信号の周波数スペクトルの位相を模式的に示した図である。

[図21A]図２１Ａは、「時間ずれに起因する位相ずれ」について説明するための図である。

[図21B]図２１Ｂは、「時間ずれに起因する位相ずれ」について説明するための図である。

[図22A]図２２Ａは、位相の時間変化が反時計回りになる仕組みについて説明するための図である。

[図22B]図２２Ｂは、位相の時間変化が反時計回りになる仕組みについて説明するための図である。

[図22C]図２２Ｃは、位相の時間変化が反時計回りになる仕組みについて説明するための図である。

[図23A]図２３Ａは、「時間ずれに起因する位相ずれ」について説明するため

の図である。

[図23B]図２３Ｂは、「時間ずれに起因する位相ずれ」について説明するための図である。

[図24]図２４は、割合判定部５００（１，ｊ）の動作における割合を求める方法の一例を説明する図である。

[図25]図２５は、割合判定部５００（１，ｊ）の動作における割合を求める方法の一例を説明する図である。

[図26]図２６は、割合判定部５００（１，ｊ）の動作における割合を求める方法の一例を説明する図である。

[図27]図２７は、音源の方向に対する相互相関値の関係の一例を示した図である。

[図28]図２８は、音源方向を算出する方法を模式的に示した図である。

[図29]図２９は、本発明の実施の形態１における音源方向検知装置の、音源の方向を特定した結果の一例を示した図である。

[図30]図３０は、音声１１００だけの理想的な音源の方向を示した図である。

[図31]図３１は、割合判定部５００（１，ｊ）を用いない場合の、音源の方向を特定した結果の一例を示した図である。

[図32]図３２は、割合判定部５００（１，ｊ）において、全ての方向に対して割合の条件を満たしていなくても音源の方向を特定した場合の結果の一例を示した図である。

[図33]図３３は、検知音特定部３０２（ｋ，ｊ）が用いた第１のしきい値の値を最適化した場合の、音源の方向を特定した結果の一例を示した図である。

[図34]図３４は、マイクロホンの各組が特定した音源の方向から、音源の方向を特定する方法の一例を説明する図である。

[図35]図３５は、３本のマイクロホンから音源位置を求める方法の一例を示した図である。

[図36]図３６は、３本のマイクロホンから割合を算出する方法の一例を示した図である。

[図37]図３７は、周波数スペクトルの位相成分のヒストグラムを作成する方法の一例について説明するための図である。

[図38A]図３８Ａは、周波数スペクトル特定部３００（ｋ，ｊ）が特定した周波数スペクトルの一例を示す図である。

[図38B]図３８Ｂは、周波数スペクトル特定部３００（ｋ，ｊ）が特定した周波数スペクトルの位相成分のヒストグラムの一例を示す図である。

[図39]図３９は、類似度の空間から類似度がしきい値以上となる周波数スペクトルのかたまりを特定する方法の一例を示す図である。

[図40]図４０は、本発明の実施の形態２における音源方向検知装置の全体構成を示したブロック図である。

[図41]図４１は、本発明の実施の形態２における音源方向検知装置の、位相補正部２８０１（ｋ）と検知音特定部２８０２（ｋ）とを示したブロック図である。

[図42]図４２は、本発明の実施の形態２における音源方向検知装置の検知音特定部２８０２（ｋ，ｊ）を示したブロック図である。

[図43]図４３は、本発明の実施の形態２における音源方向検知装置の動作手順を示すフローチャートを示した図である。

[図44]図４４は、本発明の実施の形態２における音源方向検知装置の、位相ずれを補正するステップＳ３１００（ｋ，ｊ）と、検知音の時刻部分を特定するステップＳ３１０１（ｋ，ｊ）との動作手順を示すフローチャートである。

[図45]図４５は、本発明の実施の形態２における音源方向検知装置の、検知音の時刻部分を特定するステップＳ３１０１（ｋ，ｊ）の動作手順を示すフローチャートである。

[図46]図４６は、位相が等速度でかつ１／ｆの時間間隔の間に３６０度回転するという条件を説明する図である。

[図47]図４７は、時間ずれに起因する位相ずれを補正する方法の一例について説明する図である。

[図48]図４８は、時間ずれに起因する位相ずれを補正する方法の一例について説明する図である。

[図49A]図４９Ａは、時間ずれに起因する位相ずれを補正する方法の一例について説明する図である。

[図49B]図４９Ｂは、時間ずれに起因する位相ずれを補正する方法の一例について説明する図である。

[図50]図５０は、時間ずれに起因する位相ずれを補正する方法の一例について説明する図である。

[図51]図５１は、周波数スペクトルの位相成分のヒストグラムを作成する方法の一例について説明するための図である。

[図52]図５２は、本発明の実施の形態１における他の雑音除去装置の全体構成を示したブロック図である。

[図53]図５３は、音響信号２４０１の２００Ｈｚにおける周波数スペクトルの時間波形を示した図である。

[図54]図５４は、音響信号２４０１を作成するときに用いた２００Ｈｚの正弦波における周波数スペクトルの時間波形を示した図である。

[図55]図５５は、本発明の実施の形態１における雑音除去装置の、雑音を除去するステップＳ３０１（ｊ）の動作手順を示すフローチャートを示した図である。

[図56]図５６は、音響信号２４０１から抽出された２００Ｈｚにおける周波数スペクトルの時間波形を示した図である。

[図57]図５７は、音響信号における分析結果についての考察を行うための図である。

[図58]図５８は、音響信号における分析結果についての考察を行うための図である。

[図59]図５９は、バイク音を分析した結果を示す図である。

[図60A]図６０Ａは、所定の時間幅における音響信号の位相を模式的に示した図である。

[図60B]図６０Ｂは、所定の時間幅における音響信号の位相を模式的に示した図である。

符号の説明

[0014] １００、２４０１　音響信号

１０１ａ　コンピュータ

１０１、２８００　音源方向検知装置

１０２（ｋ）（ｋ＝１～３）、４１０１（１）、４１０１（２）　マイクロホン

１０３（ｋ）（ｋ＝１～３）　ＦＦＴ分析部

１０４（ｋ）（ｋ＝１～３）、１０４（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）、１１３（ｊ）、３０２（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）、３００２（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）、２８０２（ｋ）（ｋ＝１～３）、２８０２（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）　検知音特定部

１０５、４００（１）、４００（２）、４００（３）、４０１、４００（１，ｊ）（ｊ＝１～Ｍ）　方向検知部

１０６　音源方向

１１０　雑音除去装置

３００（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）、３０００（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）　周波数スペクトル特定部

３０１（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）、３００１（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）　類似度算出部

５００（１，ｊ）（ｊ＝１～Ｍ）　割合判定部

５０１（１，ｊ）（ｊ＝１～Ｍ）　相関算出部

５０２（１，ｊ）（ｊ＝１～Ｍ）　方向特定部

５０３（１，ｊ）（ｊ＝１～Ｍ）　時間区間特定部

１１００　音声

１１０１～１１０３　白色雑音

１１０６　ＤＦＴ分析部

２４０８　出力スペクトル

２８０１（ｋ）（ｋ＝１～３）、２８０１（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）　位相補正部

４１００　音源

発明を実施するための最良の形態

[0015]　以下本発明の実施の形態について、図面を参照しながら説明する。

[0016]　（実施の形態１）

図１は、本発明の実施の形態１における音源方向検知装置の外観図である。

[0017]　音源方向検知装置は、３個のマイクロホン１０２（ｋ）（ｋ＝１～３）が接続されたコンピュータ１０１ａより構成される。

[0018]　図２～図６は、本発明の実施の形態１における音源方向検知装置の構成を示すブロック図である。

[0019]　図２において、音源方向検知装置１０１は、３個のマイクロホン１０２（ｋ）（ｋ＝１～３）と、各マイクロホンに対応した３個のＦＦＴ分析部１０３（ｋ）（ｋ＝１～３）（分析手段に対応）と、各マイクロホンに対応した３個の検知音特定部１０４（ｋ）（ｋ＝１～３）（検知音特定手段に対応）と、方向検知部１０５（検知手段に対応）とから構成される。ＦＦＴ分析部１０３（ｋ）（ｋ＝１～３）と、検知音特定部１０４（ｋ）（ｋ＝１～３）と、方向検知部１０５は、コンピュータ１０１ａのプロセッサ上で各処理部の機能を実現するためのプログラムを実行することにより実現される。また、各種データはコンピュータ１０１ａのメモリに記憶される。

[0020]　マイクロホン１０２（ｋ）（ｋ＝１～３）は、音響信号１００を入力として受ける。

[0021]　ＦＦＴ分析部１０３（ｋ）（ｋ＝１～３）は、対応するマイクロホン（ｋの値が同じもの）が入力として受けた音響信号１００に対して高速フーリエ

変換処理を施し、音響信号１００の位相成分を含む周波数スペクトル（周波数信号に対応）を求める。以下では、ＦＦＴ分析部１０３（ｋ）（ｋ＝１～３）が求めた周波数スペクトルの周波数帯域の個数をＭとして、それらの周波数帯域を指定する番号を記号ｊ（ｊ＝１～Ｍ）で表すこととする。

[0022] 図３に示すように、検知音特定部１０４（ｋ）（ｋ＝１～３）は、Ｍ個の検知音特定部１０４（ｋ，ｊ）（ｊ＝１～Ｍ）から構成される。検知音特定部１０４（ｋ）（ｋ＝１～３）は、対応するＦＦＴ分析部１０３（ｋ）（ｋの値が同じもの）が求めた周波数スペクトルに対して、周波数帯域ｊ（ｊ＝１～Ｍ）ごとに検知音特定部１０４（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）を用いて検知音の時刻部分を特定する。以下では、ｊ番目の周波数帯域がｆ（例えば、中心周波数がｆ）である場合について説明を行う。

[0023] 検知音特定部１０４（ｋ，ｊ）は、１／ｆの時間間隔の時刻から選択された複数の時刻における周波数スペクトルを用いて、上記複数の時刻のうち分析の対象とする時刻の周波数スペクトルと、上記複数の時刻のうち少なくとも前記時刻とは異なる複数の時刻における周波数スペクトルとの類似度の総和を求める。また、検知音特定部１０４（ｋ，ｊ）は、当該類似度の総和が第１のしきい値以上となる上記分析の対象とする時刻を特定することで、検知音の時刻部分を特定する。

[0024] 方向検知部１０５は、音源方向を求める時間区間（この時間区間で時間的に平均化された音源の方向が求まる）を特定して、この特定された時間区間における検知音特定部１０４（ｋ）（ｋ＝１～３）が特定した時刻部分の周波数スペクトルを用いて、３個のマイクロホンから得られるマイクロホンの組（マイクロホン１０２（１）とマイクロホン１０２（２）、マイクロホン１０２（２）とマイクロホン１０２（３）、マイクロホン１０２（１）とマイクロホン１０２（３））において、それぞれのマイクロホンに対応した特定された周波数スペクトルが共通に存在する部分の周波数スペクトル同士の一致度合いを求める。ここでは一致度合いとして相互相関値を用いる。方向検知部１０５は、それぞれのマイクロホンの組み合わせにおいて、相互相関

値の極大値から到達時間差を求める。方向検知部１０５は、求められた到達時間差とマイクロホン同士の間隔とから音源方向を求める。方向検知部１０５は、全てのマイクロホンの組合せにおいて、到達時間差がゼロになるように２つのマイクロホンからの周波数スペクトルの時間軸を調整したときに、音源方向を特定する時間区間に占める、検知音特定部により特定された時刻部分のうち前記マイクロホン間で共通の時刻に存在する部分の割合が所定のしきい値より大きい場合に、音源方向１０６を示す情報を出力する。

[0025] この例では、一致度合いとして相互相関値を用いたが、一致度合いとして周波数スペクトル間の差分誤差を用いてもよい。また、この例では、相互相関値の極大値から到達時間差を求めたが、相互相関値の最大値や、しきい値以上の相互相関値の極大値または最大値や、差分誤差の極小値や、差分誤差の最小値や、しきい値以下の差分誤差の極小値または最小値から到達時間差を求めてもよい。

[0026] 図４は、検知音特定部１０４（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）の構成を示すブロック図である。検知音特定部１０４（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）は、周波数スペクトル特定部３００（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）と、類似度算出部３０１（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）と、検知音特定部３０２（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）とから構成される。

[0027] 図７Ａおよび図７Ｂは、この実施の形態における検知音の時刻部分を特定する方法のポイントを説明する概念図である。図７Ａは、バイク音（検知音に対応）を周波数ｆで周波数分析した結果を模式的に示した図である。図７Ｂは、暗騒音（雑音に対応）を周波数ｆで周波数分析した結果を模式的に示した図である。両図ともに横軸は時間軸であり縦軸は周波数軸である。図７Ａに示すように、バイク音の周波数変化などの影響により周波数波形の振幅の大きさは変化するものの、周波数波形の位相の時間変化は、分析した周波数ｆに対応する周期である１／ｆの間隔で等角速度で０度から３６０度まで変化し、このような変化が規則的に繰り返される。例えば、１００Ｈｚにお

ける周波数波形では位相が１０ｍｓ間隔の間に３６０度回転して、２００Ｈｚにおける周波数波形では位相が５ｍｓ間隔の間に３６０度回転する。一方、図７Ｂに示すように、暗騒音などの白色雑音における周波数波形の位相の時間変化は不規則になる。また、混合音が原因で歪んだ雑音部分においても位相の時間変化は乱れて不規則になる。このように、周波数波形の位相の時間変化が規則的でない、時間と周波数とで定義された領域を除去することで雑音を除去することができる。この処理は、周波数分析をする前の時間波形において、抽出する音が周期的な音である否かに関わらず同様に扱うことができる。

[0028] 周波数スペクトル特定部３００（ｋ，ｊ）は、類似度を求めるときに用いる周波数スペクトルを特定する。類似度算出部３０１（ｋ，ｊ）は、周波数スペクトル特定部３００（ｋ，ｊ）が特定した周波数スペクトルを用いて類似度の総和を算出する。検知音特定部３０２（ｋ，ｊ）は、類似度算出部３０１（ｋ，ｊ）が算出した類似度の総和が第１のしきい値以上になる周波数スペクトルの時刻部分を、検知音の時刻部分として特定する。

[0029] 図５は、方向検知部１０５の構成を示すブロック図である。方向検知部１０５は、方向検知部４００（１）と、方向検知部４００（２）と、方向検知部４００（３）と、方向検知部４０１とから構成される。

[0030] 方向検知部４００（１）は、検知音特定部１０４（１）が特定した時刻部分の周波数スペクトルと、検知音特定部１０４（２）が特定した時刻部分の周波数スペクトルとを用いて、マイクロホン１０２（１）とマイクロホン１０２（２）とに入力された音響信号の音源の方向を求める。

[0031] 方向検知部４００（２）は、検知音特定部１０４（２）が特定した時刻部分の周波数スペクトルと、検知音特定部１０４（３）が特定した時刻部分の周波数スペクトルとを用いて、マイクロホン１０２（２）とマイクロホン１０２（３）とに入力された音響信号の音源の方向を求める。

[0032] 方向検知部４００（３）は、検知音特定部１０４（１）が特定した時刻部分の周波数スペクトルと、検知音特定部１０４（３）が特定した時刻部分の

周波数スペクトルとを用いて、マイクロホン１０２（１）とマイクロホン１０２（３）とに入力された音響信号の音源の方向を求める。

[0033] 方向検知部４０１は、全てのマイクロホンの組合せにおいて、到達時間差がゼロになるように２つのマイクロホンからの周波数スペクトルの時間軸を調整したときに、音源方向を特定する時間区間に占める、検知音特定部により特定された時刻部分のうち前記マイクロホン間で共通の時刻に存在する部分の割合が所定のしきい値より大きい場合に、方向検知部４００（１）と、方向検知部４００（２）と、方向検知部４００（３）とで求められた３つの音源の方向から音源の方向を求めて、音源方向１０６を出力する。

[0034] 図６は、方向検知部４００（１）の構成を示すブロック図である。方向検知部４００（１）は、方向検知部４００（１，ｊ）（ｊ＝１～Ｍ）から構成されている。方向検知部４００（１，ｊ）は、時間区間特定部５０３（１，ｊ）と、割合判定部５００（１，ｊ）と、相関算出部５０１（１，ｊ）と、方向特定部５０２（１，ｊ）とから構成される。方向検知部４００（２）と方向検知部４００（３）とは、方向検知部４００（１）と同じように構成されるため説明を繰り返さない。

[0035] 時間区間特定部５０３（１，ｊ）は、音源方向を求める時間区間（この時間区間で時間的に平均化された音源の方向が求まる）を特定する。

[0036] 割合判定部５００（１，ｊ）は、検知音特定部１０４（１）が特定した時刻部分のｊ番目の周波数帯域の周波数スペクトルと、検知音特定部１０４（２）が特定した時刻部分のｊ番目の周波数帯域の周波数スペクトルとを用いて、相互相関値の算出に用いる時間区間特定部５０３（１，ｊ）が特定した時間区間における、それぞれのマイクロホンに対応した特定された周波数スペクトルが共通に存在する部分の割合が第２のしきい値よりも小さいか否かを判定する。

[0037] 相関算出部５０１（１，ｊ）は、割合判定部５００（１，ｊ）が前記割合が第２のしきい値以上であると判定したときに、検知音特定部１０４（１）が特定した時刻部分のｊ番目の周波数帯域の周波数スペクトルと、時間区間

特定部５０３（１，ｊ）が特定した時間区間において検知音特定部１０４（２）が特定した時刻部分のｊ番目の周波数帯域の周波数スペクトルとが共通に存在する部分の、周波数スペクトル同士の相互相関値を算出する。

[0038] 方向特定部５０２（１，ｊ）は、相関算出部５０１（１，ｊ）が算出した相互相関値の極大値から到達時間差を求め、求められた到達時間差とマイクロホン同士の間隔とから音源方向を求める。ただし、ここでは、割合判定部５００（１，ｊ）は、－９０度から＋９０度までの全ての音源の方向に対応する相互相関値に対して、前記割合が前記第２のしきい値以上のときにのみ音源の方向を特定するように、方向特定部５０２（１，ｊ）に方向特定信号を送る。方向特定部５０２（１，ｊ）は、割合判定部５００（１，ｊ）からの方向特定信号を受信したときに音源の方向を求める。

[0039] 次に、以上のように構成された音源方向検知装置１０１の動作について説明する。

[0040] 図８～図１２は、音源方向検知装置１０１の動作手順を示すフローチャートである。

[0041] 図１３には、本実施の形態で用いる３個のマイクロホンの配置と音響信号１００の方向とが示されている。ここでは、音響信号１００として、音声１１００（検知音に対応）と白色雑音（雑音に対応）（白色雑音１１０１、白色雑音１１０２、白色雑音１１０３）との混合音（コンピュータ上で混合して作成したもの）を用いた場合を一例として説明を行う。

[0042] ３個のマイクロホンの配置は、マイクロホン１０２（１）とマイクロホン１０２（２）との間隔をＬ＝１０ｃｍとし、マイクロホン１０２（２）とマイクロホン１０２（３）との間隔をＬ＝１０ｃｍとし、マイクロホン１０２（１）とマイクロホン１０２（３）との間隔をＬ＝１４ｃｍとしている。マイクロホン１０２（１）とマイクロホン１０２（２）とを結ぶ直線とマイクロホン１０２（２）とマイクロホン１０２（３）とを結ぶ直線とは垂直に交わっている。

[0043] 音声１１００（検知音）の音源方向は、マイクロホン１０２（１）とマイ

クロホン１０２（２）とを結ぶ直線と垂直に交わる直線から、マイクロホン１０２（１）側を正の角度として－３０度の方向である。また、音声１１００の音源方向は、マイクロホン１０２（２）とマイクロホン１０２（３）とを結ぶ直線と垂直に交わる直線から、マイクロホン１０２（２）側を正の角度として＋６０度の方向である。

[0044] 白色雑音１１０１（雑音）の音源方向は、マイクロホン１０２（１）とマイクロホン１０２（２）とを結ぶ直線と垂直に交わる直線から、マイクロホン１０２（１）側を正の角度として－５０度の方向である。白色雑音１１０２（雑音）の音源方向は、マイクロホン１０２（１）とマイクロホン１０２（２）とを結ぶ直線と垂直に交わる直線から、マイクロホン１０２（１）側を正の角度として＋３０度の方向である。白色雑音１１０３（雑音）の音源方向は、マイクロホン１０２（１）とマイクロホン１０２（２）とを結ぶ直線と垂直に交わる直線から、マイクロホン１０２（１）側を正の角度として＋５０度の方向である。

[0045] ここでは、音響信号１００から白色雑音１１０１と白色雑音１１０２と白色雑音１１０３とを除去して音声１１００の音源の方向を特定することを目的とする。

[0046] 図１４に、この例における、マイクロホン１０２（１）における音響信号１００のスペクトログラムの一例を示す。横軸は時間軸（秒）であり縦軸は周波数軸（Ｈｚ）である。色の濃度は周波数スペクトルの大きさを表しており、濃い色ほど周波数スペクトルの大きさが大きいことを示している。ここでは、５０Ｈｚ～１０００Ｈｚの周波数範囲の０秒～５秒のスペクトログラムが表示されている。ここでの表示には、周波数スペクトルの位相成分の表示は省略されている。

[0047] 図１５に、図１４に示した音響信号１００を作成するときに用いた音声１１００のスペクトログラムを示す。表示の方法は図１４と同様であるため説明を繰り返さない。

[0048] 図１４と図１５とから、音響信号１００において、音声１１００の周波数

スペクトルの大きさが大きい部分においてのみ音声１１００を観測することができる。また、白色雑音（白色雑音１１０１と白色雑音１１０２と白色雑音１１０３とを加えたもの）の周波数スペクトルの大きさが音声１１００の周波数スペクトルの大きさと比較して大きな値をもっていることがわかる。

[0049] 初めに、ＦＦＴ分析部１０３（ｋ）（ｋ＝１～３）は、対応するマイクロホン１０２（ｋ）（ｋの値が同じもの）が入力した音響信号１００に対して高速フーリエ変換処理を施し、音響信号１００の位相成分を含む周波数スペクトルを求める（ステップＳ６００（ｋ）（ｋ＝１～３））。この例では、高速フーリエ変換処理により複素空間上での周波数スペクトルを求めている。この例における高速フーリエ変換処理では、サンプリング周波数＝１６０００Ｈｚの音響信号１００に時間窓の幅ΔＴ＝６４ｍｓ（１０２４ｐｔ）のハニング窓を掛け、高速フーリエ変換処理を行なっている。また、時間軸方向には１ｐｔ（０．０６２５ｍｓ）の時間シフトを行いながら各時刻における周波数スペクトルを求めている。この処理結果のマイクロホン１０２（１）における周波数スペクトルの大きさのみを表示したものが図１４である。

[0050] 次に、検知音特定部１０４（ｋ）（ｋ＝１～３）は、対応するＦＦＴ分析部１０３（ｋ）（ｋの値が同じもの）が求めた周波数スペクトルに対して、周波数帯域ｊ（ｊ＝１～Ｍ）ごとに検知音特定部１０４（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）を用いて検知音の時刻部分を特定する（ステップＳ６０１（ｋ）、ステップＳ６０１（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ））。この後の説明にはｊ番目の周波数帯域に関してのみ行う。この例では、ｊ番目の周波数帯域の周波数はｆである。

[0051] 検知音特定部１０４（ｋ，ｊ）は、１／ｆの時間間隔の時刻から選択された複数の時刻における周波数スペクトルを用いて、上記複数の時刻のうち分析の対象とする時刻の周波数スペクトルと、上記複数の時刻のうち少なくとも前記時刻とは異なる複数の時刻における周波数スペクトルとの類似度の総和が第１のしきい値以上となる分析の対象とする時刻を特定する（ステップＳ６０１（ｋ，ｊ）（ｊ＝１～Ｍ））。ここでは、周波数ｆ＝５００Ｈｚの

場合を一例として説明を行う。

[0052] 図１６（ｂ）には、図１６（ａ）に示された音響信号１００における、周波数ｆ＝５００Ｈｚにおける位相成分を含む周波数スペクトルが模式的に示されている。図１６（ａ）は、図１４と同じものであり、図１６（ｂ）において、水平軸は時間軸（秒）であり垂直平面の２軸は周波数スペクトルの実部と虚部とを表している。この例では周波数ｆ＝５００Ｈｚであるので１／ｆ＝２ｍｓとなる。

[0053] 初めに、周波数スペクトル特定部３００（ｋ，ｊ）が、１／ｆの時間間隔の時刻から選択された複数の時刻の周波数スペクトルを特定する（ステップＳ８００（ｋ，ｊ））。

[0054] 図１６（ｂ）には、１／ｆの時間間隔の時刻から特定された複数の時刻の周波数スペクトルの位置が白丸印で示されている。ここでは、図１６（ｂ）に示すように、１／ｆ＝２ｍｓの時間間隔の時刻から選択された全ての時刻の周波数スペクトルが特定される。すなわち、１／ｆ＝２ｍｓの時間間隔の時刻の周波数スペクトルが特定される。

[0055] ここで、図１７Ａと図１７Ｂとに、他の周波数スペクトルの特定方法を示す。表示の方法は図１６（ｂ）と同じであるため説明を繰り返さない。図１７Ａには、１／ｆの時間間隔の時刻から、１／ｆ×Ｎ（Ｎ＝２）の時間間隔の時刻の周波数スペクトルを特定する一例が示されている。また、図１７Ｂには、１／ｆの時間間隔の時刻から、ランダムに選択した時刻の周波数スペクトルを特定する一例が示されている。

[0056] ここで、周波数スペクトル特定部３００（ｋ，ｊ）は、類似度算出部３０１（ｋ，ｊ）が類似度の総和の算出に用いる周波数スペクトルの時間範囲も特定するが、時間範囲の特定方法の説明については、類似度算出部３０１（ｋ，ｊ）の説明と合わせて以下で行う。

[0057] 次に、類似度算出部３０１（ｋ，ｊ）は、周波数スペクトル特定部３００（ｋ，ｊ）が特定した周波数スペクトルを用いて類似度の総和を算出する（ステップＳ８０１（ｋ，ｊ））。ここでは、類似度として相関値を用いる。

[0058] 図１８に、類似度の総和の求め方の一例を示す。図１８の表示の方法において、図１６（ｂ）と共通する部分の説明は繰り返さない。図１８において、分析の対象とする時刻の周波数スペクトルを黒丸印で示して、分析の対象とする時刻とは異なる複数の時刻における特定された周波数スペクトルを白丸印で示す。

[0059] この例では、分析の対象とする時刻（黒丸印の時刻）から±１０ｍｓ以内の時刻（時間幅は２０ｍｓ）に存在する１／ｆ（＝２ｍｓ）の時間間隔の時刻から分析の対象とする時刻を除いた時刻（白丸印の時刻）の周波数スペクトルを、分析の対象の周波数スペクトルとの類似度を求める周波数スペクトルに特定する。ここでは、類似度の総和の算出に用いる周波数スペクトルの時間範囲を、分析の対象とする時刻から±１０ｍｓ以内の時刻（時間幅は２０ｍｓ）としたが、この値は、ここでの抽出したい音響信号である音声の特徴から実験的に求めた値である。

[0060] 類似度の総和の算出方法を以下に説明する。以下では、周波数スペクトルの実部を

[0061] [数1]

$$x_k \quad (k = -K,\ldots,-2,-1,0,1,2,\ldots,K)$$

と表すこととして、周波数スペクトルの虚部を

[0062] [数2]

$$y_k \quad (k = -K,\ldots,-2,-1,0,1,2,\ldots,K)$$

と表すこととする。ここでの記号ｋは周波数スペクトルを指定する番号である。ｋ＝０の周波数スペクトルは、分析の対象とする時刻の周波数スペクトルを表している。ゼロ以外のｋ（ｋ＝－Ｋ，…，－２，－１，１，２，…Ｋ）の周波数スペクトルは、分析の対象とする時刻の周波数スペクトルとの類似度を求めるための周波数スペクトルを表している（図１８を参照）。類似度の総和Ｓは、

[0063] [数3]

$$S=\sum_{k=-K}^{k=-1}\left(x_0\times x_k+y_0\times y_k\right)+\sum_{k=1}^{k=K}\left(x_0\times x_k+y_0\times y_k\right)$$

を用いて算出する。

[0064] ここで、他の類似度の総和Ｓの算出方法を以下に示す。相関値の計算において、総和した類似度の数で正規化する方法である

[0065] [数4]

$$S=\frac{1}{2K}\left(\sum_{k=-K}^{k=-1}\left(x_0\times x_k+y_0\times y_k\right)+\sum_{k=1}^{k=K}\left(x_0\times x_k+y_0\times y_k\right)\right)$$

や、分析の対象とする時刻の周波数スペクトルとの類似度も加える方法である

[0066] [数5]

$$S=\sum_{k=-K}^{k=K}\left(x_0\times x_k+y_0\times y_k\right)$$

や、周波数スペクトルの大きさで正規化された相関値を用いる方法である

[0067] [数6]

$$S=\sum_{k=-K}^{k=-1}\frac{\left(x_0\times x_k+y_0\times y_k\right)}{\sqrt{x_0^2+y_0^2}\sqrt{x_k^2+y_k^2}}+\sum_{k=1}^{k=K}\frac{\left(x_0\times x_k+y_0\times y_k\right)}{\sqrt{x_0^2+y_0^2}\sqrt{x_k^2+y_k^2}}$$

や、周波数スペクトルの距離の逆数を用いる方法である

[0068] [数7]

$$S=1\Big/\left(\sum_{k=-K}^{k=K}\sqrt{\left(x_0-x_k\right)^2+\left(y_0-y_k\right)^2}+\alpha\right)$$

や、位相成分の距離の逆数を用いる方法である

[0069]

[数8]

$$S = 1\Big/\left(\sum_{k=-K}^{k=K}\sqrt{\left(\arctan\left(y_0/x_0\right)-\arctan\left(y_k/x_k\right)\right)^2}+\alpha\right)$$

などがある。ここで

[0070] [数9]

$$\alpha$$

はＳが無限大に発散しないための予め定められた小さな値である。

[0071] 最後に、検知音特定部３０２（ｋ，ｊ）が、類似度の総和Ｓが第１のしきい値以上となる分析の対象とする周波数スペクトルの時刻部分を、検知音の時刻部分として特定する（ステップＳ８０２（ｋ，ｊ））。

[0072] 第１のしきい値をＡと表すとすると、検知音特定部３０２（ｋ，ｊ）は、

[0073] [数10]

$$S \geq A$$

の条件を満たす時刻の周波数スペクトルを特定する。この例では、第１のしきい値Ａの値を

[0074] [数11]

$$A = 10^{10.0}$$

としている。この値は、音響信号１００の音圧レベルなどに基づいて実験的に求めた値である。しきい値Ａの設定方法として、抽出する音のダイナミックレンジと雑音のダイナミックレンジとに基づいて雑音だけが除去されるようにしきい値Ａを設定する方法が挙げられる。また、アプリケーションに応じて、抽出したい音（少し歪んだもの）も多少は除去されてしまうことを許して、雑音を完全に除去できる値にしきい値Ａを設定する方法や、逆に、雑

音は完全に除去できないことを許して、抽出したい音（少し歪んだものも含む）は完全に抽出できる値にしきい値Ａを設定する方法が挙げられる。この処理は、時間軸方向に１ｐｔ（０．０６２５ｍｓ）の時間シフトを行いながら求めた全ての時刻の周波数スペクトルに対して行う。

[0075] 図１９に、図１４に示した音響信号１００から抽出された音声１１００のスペクトログラムを示す。表示の方法は図１４と同様であるため説明を繰り返さない。図１９と、図１４及び図１５とを比較すると、音響信号１００から白色雑音（白色雑音１１０１、白色雑音１１０２、白色雑音１１０３）と混合音が原因で大きく歪んでいる雑音部分が除去されて、音声１１００が抽出されていることがわかる。

[0076] ここで、雑音として除去される音響信号の周波数スペクトルの位相について考察を加える。図２０Ａ及び図２０Ｂは、類似度の総和を求める時間範囲（所定の時間幅）における、音響信号の周波数スペクトルの位相を模式的に示したものである。横軸は時間であり縦軸は位相である。黒丸印は分析の対象とする周波数スペクトルの位相を示し、白丸印は分析の対象とする周波数スペクトルとの間で類似度を求める周波数スペクトルの位相を示す。ここでは１／ｆの時間間隔での周波数スペクトルの位相が示されている。図２０Ａに示すように、類似度の総和を求めることは、分析の対象とする周波数スペクトルの位相を通る時間に対して２πｆの傾きをもつ直線（１／ｆの時間間隔では水平な値になる）との類似度の総和を求めることと同じになる。図２０Ａでは、この直線の近傍に類似度を求める周波数スペクトルの位相が集まっているため、類似度の総和はしきい値Ａより大きくなり雑音として除去されることはない。また、図２０Ｂのように、分析の対象とする周波数スペクトルの位相を通る時間に対して２πｆの傾きをもつ直線の近傍に、類似度を求める周波数スペクトルがほとんど存在しない場合には、類似度の総和はしきい値Ａ以下になり雑音として除去される。

[0077] かかる構成によれば、検知音特定部１０４（ｋ）（ｋ＝１～３）において位相成分が歪んだ雑音部分（暗騒音などの白色雑音の部分、または混合音が

原因で歪んだ雑音の部分）を特定して、検知音の時刻部分を精度良く特定することができる。

[0078] また、時間ずれに起因する位相ずれが起こらない１／ｆの時間間隔の時刻の周波数スペクトルを用いて類似度の総和を求めるため、位相ずれを補正する必要がなく計算量を削減できるという効果もある。すなわち、ψ（ｔ）－（２πｆｔ）で定義される空間での位相の類似度をψ（ｔ）で定義される空間で求めることができる。このため、類似度の算出方法が簡単になり、位相の補正を行うための計算量を削減することができるという効果もある。ここで、「時間ずれに起因する位相ずれ」について説明する。図７Ａを用いて説明したように抽出対象とされる音響信号の周波数ｆの周波数スペクトルは、位相が等速度かつ１／ｆの時間間隔の間に３６０度回転するとの仮定を設けている。このため、時間が変化すると位相も変化する。

[0079] 図２１Ａに、周波数分析を行うときに音響信号に畳み込むＤＦＴ（Discrete Fourier Transform）の波形を示す。実部はコサイン波形で虚部はマイナスのサイン波形である。ここでは、周波数ｆの信号について分析を行う。音響信号が周波数ｆの正弦波であるとき、周波数分析を行ったときの周波数スペクトルの位相の時間変化は、図２１Ｂに示すように反時計回りになる。このとき、横軸は実部であり縦軸は虚部である。反時計回りの位相を正とすると、位相は１／ｆの時間で２π（ラジアン）増加する。また、位相は時間変化に対して２πｆの傾きで変化するとも言える。図２２Ａ～図２２Ｃを用いて、位相の時間変化が反時計回りになる仕組みについて説明する。図２２Ａに、音響信号（周波数ｆの正弦波）を示す。ここでは音響信号の振幅の大きさを１としている。図２２Ｂに、周波数分析を行うときに音響信号に畳み込むＤＦＴの波形（周波数ｆ）を示す。実線は実部のコサイン波形を破線は虚部のマイナスのサイン波形を示している。図２２Ｃに、図２２Ａの音響信号と図２２ＢのＤＦＴの波形を畳み込んだときの値の符号を示す。図２２Ｃより、時刻が（ｔ１～ｔ２）のとき図２１Ｂの第１象限に、時刻が（ｔ２～ｔ３）の時に図２１Ｂの第２象限に、時刻が（ｔ３～ｔ４）のとき図２１Ｂの第

３象限に、時刻が（ｔ４～ｔ５）のとき図２１Ｂの第４象限に位相が変化することがわかる。このことから、位相の時間変化が反時計回りになることがわかる。

[0080] ここで補足であるが、図２３Ａのように、横軸を虚部にして縦軸を実部にするという特殊なことをすると位相の増減が反転して、位相は時間変化に対して（－２πｆ）の傾きで変化することが起こるが、ここでは図２１Ｂの軸の取り方に補正されているとして説明を行う。また、図２３Ｂのように、周波数分析を行うときに畳み込む波形を、実部をコサイン波形に虚部をサイン波形にするという特殊なことをすると位相の増減が反転して、位相は時間変化に対して（－２πｆ）の傾きで変化することが起こるが、ここでは、図２１Ａの周波数分析の結果にあうように実部と虚部の符号が補正されていることを前提として説明を行う。

[0081] このような現象を「時間ずれに起因する位相ずれ」と呼んでいる。ただし、１／ｆの時間間隔の間に位相は一周するため、１／ｆの時間間隔の時刻の音響信号の周波数スペクトルは、同一の位相を有すると考えられるため、位相ずれを補正する必要がなくなる。

[0082] また、類似度として相関値を用いることにより、周波数スペクトルの大きさのずれには着目せずに位相のずれを判定することができる。このため、周波数スペクトルの大きさが小さい音響信号に対しても位相成分が歪んだ雑音部分を特定することができるという効果もある。

[0083] 次に、方向検知部１０５は、検知音特定部１０４（ｋ）（ｋ＝１～３）が特定した時刻部分の雑音が除去された周波数スペクトルを用いて、３個のマイクロホンから得られるマイクロホンの組（マイクロホン１０２（１）とマイクロホン１０２（２）、マイクロホン１０２（２）とマイクロホン１０２（３）、マイクロホン１０２（１）とマイクロホン１０２（３））において、それぞれのマイクロホンに対応した特定された周波数スペクトルが共通に存在する部分の周波数スペクトル同士の相互相関値を算出する。方向検知部１０５は、それぞれのマイクロホンの組合せにおいて、相互相関値の極大値

から到達時間差を求め、求められた到達時間差とマイクロホン同士の間隔とから音源方向を求めて、全てのマイクロホンの組合せにおいて、到達時間差がゼロになるように２つのマイクロホンからの周波数スペクトルの時間軸を調整したときに、音源方向を特定する時間区間に占める、検知音特定部により特定された時刻部分のうち前記マイクロホン間で共通の時刻に存在する部分の割合が所定のしきい値より大きい場合に、音源方向１０６を示す情報を出力する（ステップＳ６０２）。

[0084] 初めに、方向検知部１０５は、３個のマイクロホンから得られるマイクロホンの組ごとに音源の方向を求める（ステップＳ９００（１）、ステップＳ９００（２）、ステップＳ９００（３））。ここでは、マイクロホンの組として、マイクロホン１０２（１）とマイクロホン１０２（２）との組を選択した場合について説明を行う（ステップＳ９００（１））。他の組（ステップＳ９００（２）とステップＳ９００（３）に対応する）については、同様な処理を行うため説明を繰り返さない。

[0085] 初めに、時間区間特定部５０３（１，ｊ）は、音源方向を求める時間区間（この時間区間で時間的に平均化された音源の方向が求まる）を特定する（ステップＳ１００３（１，ｊ））。この時間区間は、音源方向を特定する音源の移動速度により決定することができる。このとき、移動速度が速い場合は時間区間を短く設定する必要がある。また、時間区間を大きくすると音源方向を求めるための周波数スペクトルのデータ量が増えるため安定した音源方向を求めることができる。このため、許容される範囲で時間区間を長くすることが好ましい。

[0086] 次に、割合判定部５００（１，ｊ）は、検知音特定部１０４（１）が特定した時刻部分のｊ番目の周波数帯域の周波数スペクトルと、検知音特定部１０４（２）が特定した時刻部分のｊ番目の周波数帯域の周波数スペクトルとを用いて、相互相関値の算出に用いる時間区間特定部５０３（１，ｊ）が特定した時間区間における、それぞれのマイクロホンに対応した特定された周波数スペクトルが共通に存在する部分の割合が第２のしきい値よりも小さい

か否かを判定する（ステップＳ１０００（１，ｊ））。

[0087] 図２４～図２６に、相互相関値の算出に用いる時間区間における、それぞれのマイクロホンに対応した特定された周波数スペクトルが共通に存在する部分の割合が第２のしきい値よりも小さいか否かを判定する方法の一例を示す。この例では、第２のしきい値＝１００％としている。

[0088] 図２４には、マイクロホン１０２（１）に対応する検知音特定部１０４（１）が特定した時刻部分のｊ番目の周波数帯域の周波数スペクトルと、マイクロホン１０２（２）に対応する検知音特定部１０４（２）が特定した時刻部分のｊ番目の周波数帯域の周波数スペクトルとが模式的に示されている。横軸は時間軸であり、縦軸には周波数スペクトルの実部又は虚部における振幅値が示されている。斜線が施された部分は、雑音部分を表している。ここでは、相互相関値の算出に用いる時間区間をΔＴＡとしている。また、マイクロホン同士の到達時間差をτとしている。

[0089] 図２５は、図２４におけるマイクロホン１０２（１）に対応する周波数スペクトルの時間軸と、マイクロホン１０２（２）に対応する周波数スペクトルの時間軸とを到達時間差τだけずらして表示したものである。表示の方法は、図２４と同様であるので説明を繰り返さない。

[0090] 図２６は、図２５において、マイクロホン１０２（１）に対応する周波数スペクトルと、マイクロホン１０２（２）に対応する周波数スペクトルとが共通に存在する部分の時間区間の総和をΔＴＰとして表示したものである。図２６における横線が施された部分の時間区間の総和がΔＴＰである。ΔＴＰは、到達時間差τだけずらして時間軸を調整したときに、周波数スペクトルが除去されずに、マイクロホン１０２（１）に対応する周波数スペクトルと、マイクロホン１０２（２）に対応する周波数スペクトルとが同じ時刻に共通に残っている部分の時間区間の総和である。

[0091] そして、割合判定部５００（１，ｊ）は、（ΔＴＰ/ΔＴＡ）×１００％が第２のしきい値（＝１００％）よりも小さいか否かを判定する。

[0092] 次に、相関算出部５０１（１，ｊ）は、割合判定部５００（１，ｊ）が（

ΔＴＰ/ΔＴＡ）×１００％が第２のしきい値（＝１００％）以上であると判定したときに、検知音特定部１０４（１）が特定した時刻部分のｊ番目の周波数帯域の周波数スペクトルと、検知音特定部１０４（２）が特定した時刻部分のｊ番目の周波数帯域の周波数スペクトルとの、共通に存在する部分の周波数スペクトル同士の相互相関値を算出する（ステップＳ１００１（１，ｊ））。なお、相互相関値は、到達時間差τを変化させながら、到達時間差τ毎に算出される。

[0093] ここで、共通に存在する部分の、マイクロホン１０２（１）に対応する周波数スペクトルの実部を

[0094] [数12]

$$x_k^1$$

と表すこととして、マイクロホン１０２（１）に対応する周波数スペクトルの虚部を

[0095] [数13]

$$y_k^1$$

と表すこととして、共通に存在する部分の、マイクロホン１０２（２）に対応する周波数スペクトルの実部を

[0096] [数14]

$$x_k^2$$

と表すこととして、マイクロホン１０２（２）に対応する周波数スペクトルの虚部を

[0097] [数15]

$$y_k^2$$

と表すこととする。ここでの記号ｋは周波数スペクトルの時刻を指定する番号である。そして、相互相関値Ｈを以下の式で算出する。

[0098] [数16]

$$H = \sum_k \left(x_k^1 \times x_k^2 + y_k^1 \times y_k^2 \right)$$

ここで、他の相互相関値Ｈの算出方法を以下に示す。総和した時刻の数Ｋで正規化する方法である

[0099] [数17]

$$H = 1/K \times \sum_k \left(x_k^1 \times x_k^2 + y_k^1 \times y_k^2 \right)$$

や、周波数スペクトルの大きさで正規化する方法である

[0100] [数18]

$$H = \sum_k \frac{\left(x_k^1 \times x_k^2 + y_k^1 \times y_k^2 \right)}{\sqrt{\left(x_k^1\right)^2 + \left(y_k^1\right)^2}\sqrt{\left(x_k^2\right)^2 + \left(y_k^2\right)^2}}$$

などがある。

[0101] 最後に、方向特定部５０２（１，ｊ）は、相関算出部５０１（１，ｊ）が算出した相互相関値に対応する音源の方向を音源方向の候補として、相互相関値の極大値から到達時間差を求め、求められた到達時間差とマイクロホン同士の間隔とから音源方向を求める。ただし、ここでは、割合判定部５００（１，ｊ）は、－９０度から＋９０度までの全ての音源の方向に対応する相互相関値に対して、前記割合が前記第２のしきい値以上のときにのみ音源の方向を求めるように、方向特定部５０２（１，ｊ）に方向特定信号を送る。方向特定部５０２（１，ｊ）は、割合判定部５００（１，ｊ）からの方向特定信号を受信したときに音源の方向を求める（ステップＳ１００２（１，ｊ））。

[0102] 図２７には、音源の方向が－９０度から＋９０度の範囲にあるとし、５度刻みで音源の方向を変化させた場合の相互相関値の一例が示されている。なお、上述したように、実際には、相互相関値は到達時間差τを変化させながら求められる。横軸は音源の方向（度）を示しており、縦軸は相互相関値の対数値が示されている。この例では、－３０度の方向に極大値が存在する。すなわち、相互相関値の極大値に対応する到達時間差とマイクロホン同士の間隔とから求めた音源方向は－３０度であることを意味している。このとき、－９０度から＋９０度の範囲での５度刻みの全ての方向に対して、割合が第２のしきい値以上であり、相互相関値が算出できていることを意味する。

[0103] 以下に、到達時間差τとマイクロホン同士の間隔Ｌとから音源方向Θを算出する方法を示す。図２８は、音源方向を算出する方法を模式的に示した図である。ここで、音速をＣと、音源方向をΘと表すこととすると、

[0104] [数19]

$$\Theta = \arcsin\left(\tau \times C/L\right)$$

の関係式が成り立ち、この関係を用いて到達時間差から音源４１００の方向を求めることができる。ここで、Ｌはマイクロホンの間隔（マイクロホン４１０１（１）とマイクロホン４１０１（２）との距離）であり、τはマイクロホン間での到達時間差である。

[0105] 図２９に、この例における音源の方向を特定した結果を示す。横軸は時間軸（秒）であり縦軸は周波数軸（Ｈｚ）である。色の濃度は音源の方向を表しており、薄い色は－９０度の方向であり濃い色は＋９０度の方向である。空白の部分は音源の方向を特定していない部分である。図２９の下方に音源の方向と濃度との関係を示す。ここでは、５０Ｈｚ～１０００Ｈｚの周波数範囲の０秒～５秒の音源の方向が表示されている。

[0106] 図３０に、音声１１００だけを音響信号１００として入力して処理した場合での理想的な音源の方向を示す。表示の方法は図２９と同じなので説明を繰り返さない。図３０と図１５とを比較すると、周波数スペクトルの大きさ

が小さい部分においても音源の方向が表示されていることがわかる。

[0107] 図３１に、方向検知部１０５において、割合判定部５００（１，ｊ）（ｊ＝１～Ｍ）を用いずに、前記割合と第２のしきい値との関係を考慮しなかった場合の音源の方向を示す。表示の方法は図２９と同じなので説明を繰り返さない。図３１と図３０とを比較すると、雑音部分の影響が非常に大きくでており音声１１００の音源の方向とは異なる方向が特定されていることがわかる。

[0108] 図２９と図３０と図３１とから、割合判定部５００（１，ｊ）（ｊ＝１～Ｍ）を用いて前記割合と第２のしきい値との関係を考慮することで、雑音部分の影響を大幅に低減して音声１１００の音源の方向を求めることができることがわかる。

[0109] 図３２に、方向検知部１０５の割合判定部５００（１，ｊ）（ｊ＝１～Ｍ）において、全ての音源の方向に対応する相互相関値に対して、前記割合が前記第２のしきい値以上のときにのみ音源の方向を求めるという条件を設けなかった場合の本発明による音源の方向を示す。表示の方法は図２９と同じなので説明を繰り返さない。図３２と図３１とを比較すると、割合判定部５００（１，ｊ）（ｊ＝１～Ｍ）を用いて前記割合と第２のしきい値との関係を考慮することで、雑音部分の影響を大幅に低減して音声１１００の音源の方向を求めることができることがわかる。

[0110] ここで、図２９と図３１とを比較すると、割合判定部５００（１，ｊ）において、全ての音源の方向に対応する相互相関値に対して、前記割合が前記第２のしきい値以上のときにのみ音源の方向を求めるという条件を設けることで、雑音部分の影響をさらに低減して音声１１００の音源の方向を求めることができることがわかる（図２９と図３１において丸印で囲まれた部分を参照）。

[0111] 図３３に、検知音特定部１０４（ｋ）における検知音特定部３０２（ｋ，ｊ）が用いた第１のしきい値の値を最適化して、

[0112]

[数20]

$$A = 10^{11.5}$$

とした場合における音源の方向を示す。表示の方法は図２９と同じなので説明を繰り返さない。図３３より、第１のしきい値の値が最適化できた場合には、雑音部分の影響をさらに大幅に低減して音声１１００の音源の方向を求めることができることがわかる。

[0113] 最後に、方向検知部４０１は、方向検知部４００（１）が求めた音源の方向と、方向検知部４００（２）が求めた音源の方向と、方向検知部４００（３）が求めた音源の方向とから音源の方向を求める（ステップＳ９０１）。このとき、方向検知部４００（１）と方向検知部４００（２）と方向検知部４００（３）の全ての方向検知部において、音源方向が求められた場合に、方向検知部４０１は音源方向を求めて出力する。なお、別の方法として、方向検知部４０１は、一部の方向検知部により求められた音源方向を用いて音源方向を求めてもよい。また、方向検知部４０１は、全ての方向検知部において音源方向が求められなかった場合にも一部の方向検知部により求められた音源方向を用いて音源方向を求めてもよい。

[0114] 図３４には、方向検知部４００（１）が求めた音源の方向と、方向検知部４００（２）が求めた音源の方向と、方向検知部４００（３）が求めた音源の方向とから音源の方向を求める方法の一例が示されている。特定した３個の音源の方向と、マイクロホン１０２（１）とマイクロホン１０２（２）とマイクロホン１０２（３）との配置とから音声１１００の音源の方向を求めることができる。

[0115] かかる構成によれば、方向検知部１０５が、時間軸調整後の各々のマイクロホンに対応する周波数信号において、到達時間差を求める時間区間における、各々のマイクロホンに対応する検知音特定手段が特定した周波数信号同士が共通に存在する時刻部分の割合が、所定のしきい値より小さくなる音源位置を検知音の音源位置の候補から除外することで、雑音の影響により誤っ

た音源位置を特定してしまうことを回避できる。

[0116] また、割合判定部５００（１，ｊ）が、全ての音源位置の候補について、一つでも前記割合が前記所定のしきい値より小さくなる場合、前記検知音の音源位置を特定しないことで、雑音の影響で到達時間差が求められない音源位置が検知音の音源位置であった場合でも、別の候補である誤った音源位置を特定してしまうことを回避できる。

[0117] また、検知音特定部１０４（ｋ）が、位相成分が歪んだ雑音部分を除去して音源の方向を特定している。このことで、暗騒音などの白色雑音および混合音が原因で歪んだ雑音部分を除去してから、正確な音源の方向を特定することができる。

[0118] また、検知音の時刻部分を特定するときに、時間ずれに起因する位相ずれが起こらない１／ｆの時間間隔の時刻の周波数スペクトルを用いて類似度の総和を求める。このため、位相ずれを補正する必要がなく計算量を削減できるという効果もある。ここで、「時間ずれに起因する位相ずれ」について説明する。図７Ａを用いて説明したように抽出対象とされる音響信号の周波数ｆの周波数スペクトルは、位相が等速度かつ１／ｆの時間間隔の間に３６０度回転するとの仮定を設けている。このため、時間が変化すると位相も変化する。このような現象を「時間ずれに起因する位相ずれ」と呼んでいる。ただし、１／ｆの時間間隔の間に位相は一周するため、１／ｆの時間間隔の時刻の音響信号の周波数スペクトルは、同一の位相を有すると考えられるため、位相ずれを補正する必要がなくなる。

[0119] また、検知音の時刻部分を特定するときに、類似度として相関値を用いる。このことで、周波数スペクトルの大きさのずれには着目せずに位相のずれを判定することができる。このため、周波数スペクトルの大きさが小さい音響信号に対しても位相成分が歪んだ雑音部分を特定することができるという効果がある。また、方向検知部における相互相関値との尺度における整合が取れる。

[0120] 本発明の雑音部分の特定に関する考え方は、暗騒音などの白色雑音または

混合音が原因で歪んだ雑音部分は、位相成分が歪んでおり、類似度の総和Ｓが第１のしきい値Ａよりも小さくなるという考え方に基づいている。

[0121] なお、検知音特定部において、音源方向を求める検知音の周波数スペクトルの時刻部分を特定する方法として、検知音（Ｓ）と雑音（Ｎ）との周波数スペクトルのパワーの比率であるＳＮ比が０ｄＢ以上の時刻部分（検知音の周波数スペクトルのパワー≧雑音の周波数スペクトルのパワー）を検知音の時刻部分に特定してもよい。また、学習用データから入力音が検知音である確率分布と雑音である確率分布を求めて、ベイズ推定により検知音の時刻部分を特定してもよい。

[0122] また、この例では、音源方向を求める検知音の周波数スペクトルの時刻部分を特定してから、音源方向を求める時間区間を特定したが、初めに音源方向を求める時間区間を特定して、この時間区間において音源方向を求める検知音の周波数スペクトルの時刻部分を特定してもよい。

[0123] また、本実施の形態において、３個のマイクロホンのそれぞれに対応した３個のＦＦＴ分析部と３個の検知音特定部とを設けたが、１つのＦＦＴ分析部と１つの検知音特定部とを共有して各マイクロホンが入力した音響信号を処理しても良い。

[0124] また、本実施の形態において、時間及び周波数ごとに音源の方向を特定したが、時間軸方向に音源の方向のヒストグラムをとり最大値をとる音源方向を音源の方向に特定しても良いし、時間軸方向に音源の方向を平均化して音源の方向を特定しても良い。

[0125] さらに、本実施の形態において、時間及び周波数ごとに音源の方向を特定したが、周波数軸方向に音源の方向のヒストグラムをとり最大値をとる音源方向を音源の方向に特定しても良いし、周波数方向に音源の方向を平均化して音源の方向を特定しても良い。

[0126] さらにまた、本実施の形態において、マイクロホンの数を３個としたが２個以上あれば良い。

[0127] また、ＦＦＴ分析部におけるＦＦＴの代わりに、ＤＦＴ（Discrete Fourie

r Transform）、コサイン変換、ウェーブレット変換、又は、バンドパスフィルタを用いても良い。

[0128] なお、検知音特定部１０４（ｋ，ｊ）（周波数ｆに対応した処理）において１／ｆの時間間隔の時刻から全ての時刻の周波数スペクトルを特定したが、１／ｆの時間間隔の時刻から任意の時刻を特定することで周波数スペクトルを特定しても良い。また、分析の対象とする時刻ごとに周波数スペクトルの特定方法を変えてもよい。

[0129] なお、検知音特定部１０４（ｋ）はＦＦＴ分析部１０３（ｋ）が求めた全て（Ｍ個）の周波数帯域の周波数スペクトルに対して雑音の除去を行い音源の方向を特定したが、一部の周波数帯域を選択し、選択した周波数帯域において雑音の除去を行い音源の方向を特定しても良い。

[0130] また、ＦＦＴ分析部は、ΔＴの幅の時間窓を掛けた前記音響信号から前記周波数スペクトルを求めて、前記検知音特定部は、分析の対象とする時刻の前記周波数スペクトルと、分析の対象とする時刻をはさみΔＴの時間間隔よりも離れた複数の時刻を含む複数の時刻の前記周波数スペクトルとの類似度の総和を用いて検知音の時刻部分を特定してもよい。本構成によって、分析の対象とする時刻をはさみΔＴの時間間隔（周波数スペクトルを求めるときの時間窓の幅）よりも離れた時刻を含む複数の時刻の前記周波数スペクトルとの類似度の総和を用いることで、時間分解能（ΔＴ）を細かくしたときの影響による周辺の周波数からの周波数もれによる雑音（歪み）を除去することができる。

[0131] なお、この例では、２本のマイクロホンを組みにして３組の組合せから求めた３つの音源方向から音源位置を求めたが、３本のマイクロホンを組にして音源位置を求めても良い。図３５に、３本のマイクロホンから音源位置を求める方法を示す。音源位置に依存して、音声１１００（検知音）が、マイクロホン１０２（１）に到達する時間Ｔ（１）とマイクロホン１０２（２）に到達する時間Ｔ（２）とマイクロホン１０２（３）に到達する時間Ｔ（３）が異なってくる。このため、Ｔ（１）とＴ（２）とＴ（３）との時間差（

到達時間差）とマイクロホンの位置の関係から音源位置を求めることができる。図３６に、この例の場合での割合の求め方を示す。表示方法は図２６と同様であるため説明を繰り返さない。ただし、３本のマイクロホンの到達時間差を同時に合わせた結果が示されている。また、周波数スペクトルの実部だけを表示している。図３６に示すように、音源位置に対応した到達時間差だけ時間軸をずらして割合を求める。そして、割合（ΔＴＰ／ΔＴＡ）×１００％が第２のしきい値より小さくなる音源位置を前記検知音の音源位置の候補から除外する。さらには、全ての音源位置の候補について、一つでも前記割合が前記所定のしきい値より小さくなる場合、前記検知音の音源位置を特定しないようにしても良い。

[0132] （実施の形態１の変形例）

次に、実施の形態１に示した音源方向検知装置の変形例について説明する。

[0133] 本変形例に係る音源方向検知装置は、図２～図６を参照して説明した実施の形態１に係る音源方向検知装置と同様の構成を有する。ただし、検知音特定部１０４（ｋ）（ｋ＝１～３）が実行する処理が異なる。

[0134] 検知音特定部１０４（ｋ）（ｋ＝１～３）は、対応するＦＦＴ分析部１０３（ｋ）（ｋの値が同じもの）が求めた周波数スペクトルに対して、周波数帯域ｊ（ｊ＝１～Ｍ）ごとに検知音特定部１０４（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）を用いて検知音の時刻部分を特定する。以下では、ｊ番目の周波数帯域がｆ（例えば、中心周波数がｆ）である場合について説明を行う。

[0135] 検知音特定部１０４（ｋ，ｊ）は、１／ｆの時間間隔の時刻から選択された複数の時刻における周波数スペクトルを用いて、類似度の空間を、上記周波数スペクトルの位相成分のヒストグラムを作成することにより求める。検知音特定部１０４（ｋ，ｊ）は、類似度がしきい値以上となる周波数スペクトルのかたまりを、検知音の周波数スペクトルと特定する。

[0136] 類似度算出部３０１（ｋ，ｊ）は、周波数スペクトル特定部３００（ｋ，ｊ）が特定した周波数スペクトルの位相成分のヒストグラムを作成すること

により類似度の空間を求める。検知音特定部３０２（ｋ，ｊ）は、類似度算出部３０１（ｋ，ｊ）が求めた類似度の空間から類似度がしきい値以上となる周波数スペクトルの時刻を、検知音の時刻部分と特定する。

[0137] 次に、以上のように構成された音源方向検知装置１０１の動作について説明する。音源方向検知装置１０１の動作手順を示すフローチャートは、実施の形態１と同様であり、図８～図１２に示されている。

[0138] 以下の説明では、音源方向検知装置１０１のうち検知音特定部１０４（ｋ）（ｋ＝１～３）が実行する処理についてのみ説明する。その他の処理部が実行する処理は実施の形態１と同様であるため、説明を繰り返さない。

[0139] 検知音特定部１０４（ｋ）（ｋ＝１～３）は、対応するＦＦＴ分析部１０３（ｋ）（ｋの値が同じもの）が求めた周波数スペクトルに対して、周波数帯域ｊ（ｊ＝１～Ｍ）ごとに検知音特定部１０４（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）を用いて検知音の時刻部分を特定する（ステップＳ６０１（ｋ）、ステップＳ６０１（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ））。この後の説明はｊ番目の周波数帯域に関してのみ行う。この例では、ｊ番目の周波数帯域の周波数はｆである。

[0140] 検知音特定部１０４（ｋ，ｊ）は、１／ｆの時間間隔の時刻から選択された複数の時刻における周波数スペクトルを用いて、類似度の空間を、上記周波数スペクトルの位相成分のヒストグラムを作成することにより求める。検知音特定部１０４（ｋ，ｊ）は、類似度がしきい値以上となる周波数スペクトルのかたまりを、検知音の周波数スペクトルと特定する（ステップＳ６０１（ｋ，ｊ））。

[0141] 類似度算出部３０１（ｋ，ｊ）は、周波数スペクトル特定部３００（ｋ，ｊ）が特定した周波数スペクトルを用いて、上記周波数スペクトルの位相成分のヒストグラムを作成することにより類似度の空間を求める（ステップＳ８０１（ｋ，ｊ））。以下、類似度の空間を求める方法について説明する。

[0142] 周波数スペクトル特定部３００（ｋ，ｊ）が特定した周波数スペクトルを、数１、数２で表すこととする。ここで、以下の式を用いて位相成分を抽出

する。

[0143] [数21]

$$\theta_k = \arctan(y_k / x_k) \quad \left(k = -K,\ldots,-2,-1,0,1,2,\ldots,K\right)$$

[0144] 図３７に、周波数スペクトルの位相成分のヒストグラムを作成する方法の一例を示す。ここでは、位相区間がΔψ（ｉ）（ｉ＝１～４）で、位相が時間に対して２πｆの傾きで変化する帯領域ごとの、所定の時間幅における周波数スペクトルの出現頻度を求めることでヒストグラムを作成する。図３７の斜線で示されている部分はΔψ（１）の領域である。ここでは位相を０～２πの間に制限して表現しているためにとびとびの領域になっている。ここで、Δψ（ｉ）（ｉ＝１～４）ごとにそれらの領域に含まれる周波数スペクトルの数をカウントすることでヒストグラムを作成することができる。

[0145] 図３８Ａ及び図３８Ｂに、周波数スペクトル特定部３００（ｋ，ｊ）が特定した周波数スペクトルの一例と、上記特定された周波数スペクトルの位相成分のヒストグラムの一例とをそれぞれ示す。

[0146] 図３８Ａに、特定された周波数スペクトルを示す。図３８Ａの表示の方法は、図１６（ｂ）と同じであるので説明を繰り返さない。この例では、特定された周波数スペクトルの中に音声Ａ（検知音に対応）と音声Ｂ（検知音に対応）と暗騒音（雑音に対応）との周波数スペクトルが含まれている。

[0147] 図３８Ｂに、周波数スペクトルの位相成分のヒストグラムの一例を模式的に示す。音声Ａの周波数スペクトルのかたまりは類似した位相（この例では１００度の近傍）をもち、音声Ｂの周波数スペクトルのかたまりは類似した位相（この例では１９０度の近傍）をもつ。このため、ヒストグラムの１００度の近傍と１９０度の近傍に山ができている。また、暗騒音の周波数スペクトルは異なる位相をもつため、ヒストグラムでは山ができていない。

[0148] 最後に、検知音特定部３０２（ｋ，ｊ）は、類似度算出部３０１（ｋ，ｊ）が求めた類似度の空間（ヒストグラム）から、類似度がしきい値以上となる周波数スペクトルのかたまりを、検知音の周波数スペクトルと特定する（

ステップＳ８０２（ｋ，ｊ））。

[0149] 図３９に、類似度の空間から類似度がしきい値以上となる周波数スペクトルのかたまりを、検知音の周波数スペクトルと特定する方法の一例を示す。図３９の表示の方法は、図３８Ｂと同じであるので説明を繰り返さない。図３９のヒストグラムにおいて、類似度がしきい値以上の周波数スペクトルのかたまり（この例では位相差が３０度以内の類似したもの）を、検知音の周波数スペクトルと特定する。この例では、１００度近傍の周波数スペクトルのかたまりと、１９０度近傍の周波数スペクトルのかたまりとが、検知音の周波数スペクトルと特定される。これらはそれぞれ音声Ａと音声Ｂに対応している。暗騒音の周波数スペクトルは異なる位相をもつため、類似度がしきい値より小さくなる。

[0150] このように、音声Ａと音声Ｂというように異なる位相に対応する周波数スペクトルが複数含まれる場合でも、暗騒音を除去して音声Ａと音声Ｂの周波数スペクトルを区別して抽出することができる。

[0151] かかる構成によれば、類似度の空間を、周波数スペクトルの位相成分のヒストグラムを作成することにより求めて、類似度がしきい値以上となる周波数スペクトルのかたまりを、検知音の周波数スペクトルと特定する。このことで、音声Ａと音声Ｂというように異なる位相に対応するものが複数含まれる場合でも、暗騒音を除去して音声Ａと音声Ｂの周波数スペクトルを区別して、検知音の周波数スペクトルと特定することができるという効果がある。つまり、実施の形態１で説明したように、例えば、図１８に示すような分析対象とされる２０ｍｓｅｃの時間幅内に、音声Ａと音声Ｂとの周波数スペクトルが混在するような場合には、類似度を計算すると類似度が低くなってしまい、雑音とみなされる場合がある。これは、局所的に雑音か否かの判断を行なっているからである。しかし、周波数スペクトルの位相成分のヒストグラムを用いる方法は、局所的に雑音か否かの判断を行なうものではない。このため、同じような状況においても音声Ａと音声Ｂの周波数スペクトルを、検知音の周波数スペクトルと特定することができる。よって、音源の方向を

正確に特定することができる。

[0152] （実施の形態２）

次に、実施の形態２に係る音源方向検知装置について説明する。実施の形態２に係る音源方向検知装置は、実施の形態１に係る音源方向検知装置と異なり、音響信号の周波数スペクトルの位相成分を補正し、補正後の周波数スペクトルを用いて音源の方向を特定する。

[0153] 図４０～図４２は、本発明の実施の形態２における音源方向検知装置の構成を示すブロック図である。図４０において、図２と同じ構成要素については同じ符号を用い、説明を繰り返さない。

[0154] 図４０において、音源方向検知装置２８００は、３個のマイクロホン１０２（ｋ）（ｋ＝１～３）と、各マイクロホンに対応した３個のＦＦＴ分析部１０３（ｋ）（ｋ＝１～３）（分析手段に対応）と、各マイクロホンに対応した３個の位相補正部２８０１（ｋ）（ｋ＝１～３）と、各マイクロホンに対応した３個の検知音特定部２８０２（ｋ）（ｋ＝１～３）（検知音特定手段に対応）と、方向検知部１０５（検知手段に対応）とから構成される。

[0155] マイクロホン１０２（ｋ）（ｋ＝１～３）は、音響信号１００を入力として受ける。

[0156] ＦＦＴ分析部１０３（ｋ）（ｋ＝１～３）は、対応するマイクロホン（ｋの値が同じもの）が入力として受けた音響信号１００に対して高速フーリエ変換処理を施し、音響信号１００の位相成分を含む周波数スペクトル（周波数信号に対応）を求める。以下では、ＦＦＴ分析部１０３（ｋ）（ｋ＝１～３）が求めた周波数スペクトルの周波数帯域の個数をＭとして、それらの周波数帯域を指定する番号を記号ｊ（ｊ＝１～Ｍ）で表すこととする。以下では、ｊ番目の周波数帯域がｆ（例えば、中心周波数がｆ）である場合について説明を行う。ここでの処理は実施の形態１と同じである。

[0157] 位相補正部２８０１（ｋ）（ｋ＝１～３）は、対応するＦＦＴ分析部１０３（ｋ）（ｋの値が同じもの）が求めた周波数スペクトルに対して、周波数帯域ｊ（ｊ＝１～Ｍ）ごとに位相補正部２８０１（ｋ，ｊ）（ｋ＝１～３、

ｊ＝１～Ｍ）を用いて、位相が等速度でかつ１／ｆの時間間隔の間に３６０度回転するという条件のもとで、複数の時刻における周波数スペクトルに対して、時間ずれに起因する位相ずれを補正する（図４１を参照）。すなわち、時刻ｔの周波数スペクトルの位相ψ（ｔ）（ラジアン）をψ（ｔ）－（２πｆｔ）に補正を行う。

[0158] 検知音特定部２８０２（ｋ）（ｋ＝１～３）は、対応するＦＦＴ分析部１０３（ｋ）（ｋの値が同じもの）が求めた周波数スペクトルと、対応する位相補正部２８０１（ｋ）（ｋの値が同じもの）が位相ずれを補正した周波数スペクトルとを用いて、周波数帯域ｊ（ｊ＝１～Ｍ）ごとに検知音特定部２８０２（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）を用いて、位相ずれが補正された分析の対象とする時刻の周波数スペクトルと、位相ずれが補正された、少なくとも分析の対象とする時刻とは異なる複数の時刻における周波数スペクトルとの類似度の総和を求める。また、検知音特定部２８０２（ｋ）（ｋ＝１～３）は、類似度の総和が第１のしきい値以上となる分析の対象とする時刻を求める。検知音特定部２８０２（ｋ）（ｋ＝１～３）は、第１のしきい値以上となる時刻のＦＦＴ分析部１０３（ｋ）が求めた周波数スペクトルを、検知音の周波数スペクトルと特定することにより雑音を除去する（図４１を参照）。

[0159] 方向検知部１０５は、検知音特定部２８０２（ｋ）（ｋ＝１～３）が特定した検知音の周波数スペクトルを用いて、３個のマイクロホンから得られるマイクロホンの組（マイクロホン１０２（１）とマイクロホン１０２（２）、マイクロホン１０２（２）とマイクロホン１０２（３）、マイクロホン１０２（１）とマイクロホン１０２（３））において、それぞれのマイクロホンに対応した特定された周波数スペクトルが共通に存在する部分の周波数スペクトル同士の一致度合い（相互相関値）を算出する。方向検知部１０５は、それぞれのマイクロホンの組合せにおいて、相互相関値の極大値から到達時間差を求め、求められた到達時間差とマイクロホン同士の間隔とから音源方向を求める。方向検知部１０５は、全てのマイクロホンの組合せにおいて

、到達時間差がゼロになるように２つのマイクロホンからの周波数スペクトルの時間軸を調整したときに、音源方向を特定する時間区間に占める、検知音特定部により特定された時刻部分のうち前記マイクロホン間で共通の時刻に存在する部分の割合が所定のしきい値より大きい場合に、音源方向１０６を示す情報を出力する（ステップＳ６０２）。なお、別の方法として、方向検知部１０５は、一部の方向検知部により求められた音源方向を用いて音源方向を求めてもよい。また、方向検知部１０５は、全ての方向検知部において音源方向が求められなかった場合にも一部の方向検知部により求められた音源方向を用いて音源方向を求めてもよい。

[0160] この例では、一致度合いとして相互相関値を用いたが、一致度合いとして周波数スペクトル間の差分誤差を用いても良い。また、この例では、相互相関値の極大値から到達時間差を求めたが、相互相関値の最大値や、しきい値以上の相互相関値の極大値又は最大値や、差分誤差の最小値や、しきい値以下の差分誤差の極小値又は最小値から到達時間差を求めても良い。

[0161] 図４２は、検知音特定部２８０２（ｋ，ｊ）（ｋ＝１～３，ｊ＝１～Ｍ）の構成を示すブロック図である。検知音特定部２８０２（ｋ，ｊ）は、周波数スペクトル特定部３０００（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）と、類似度算出部３００１（ｋ，ｊ）（ｋ＝１～３，ｊ＝１～Ｍ）と、検知音特定部３００２（ｋ，ｊ）（ｋ＝１～３，ｊ＝１～Ｍ）とから構成される。

[0162] 周波数スペクトル特定部３０００（ｋ，ｊ）は、位相補正部２８０１（ｋ，ｊ）が位相ずれを補正した周波数スペクトルから、類似度算出部３００１（ｋ，ｊ）が類似度の総和を算出するのに用いる周波数スペクトルを特定する。

[0163] 類似度算出部３００１（ｋ，ｊ）は、周波数スペクトル特定部３０００（ｋ，ｊ）が特定した位相ずれが補正された周波数スペクトルを用いて類似度の総和を算出する。

[0164] 検知音特定部３００２（ｋ，ｊ）は、類似度算出部３００１（ｋ，ｊ）が算出した、類似度の総和が第１のしきい値以上の時刻の周波数スペクトルを

、ＦＦＴ分析部１０３（ｋ）が求めた周波数スペクトルから特定する。

[0165] 次に、以上のように構成された音源方向検知装置２８００の動作について説明する。

[0166] 図４３～図４５は、音源方向検知装置２８００の動作手順を示すフローチャートである。

[0167] 初めに、ＦＦＴ分析部１０３（ｋ）（ｋ＝１～３）は、対応するマイクロホン１０２（ｋ）（ｋの値が同じもの）が入力した音響信号１００に対して高速フーリエ変換処理を施し、音響信号１００の位相成分を含む周波数スペクトルを求める（ステップＳ６００（ｋ）（ｋ＝１～３））。ここでは、実施の形態１と同様の方法に従い周波数スペクトルを求める。

[0168] 次に、位相補正部２８０１（ｋ）（ｋ＝１～３）は、対応するＦＦＴ分析部１０３（ｋ）（ｋの値が同じもの）が求めた周波数スペクトル（周波数はｆである）に対して、位相が等速度でかつ１／ｆの時間間隔の間に３６０度回転するという条件のもとで、複数の時刻における周波数スペクトルに対して、時間ずれに起因する位相ずれを補正する（ステップＳ３１００（ｋ）、ステップＳ３１００（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ））。

[0169] ここで、図４６を用いて、位相が等速度でかつ１／ｆの時間間隔の間に３６０度回転するという条件を説明する。横軸は時間軸（秒）を示しており縦軸は位相（度）を表す。図４６に示されているように、上記の条件とは、１／ｆ（秒）の時間間隔の間に位相が０度から３６０度まで直線的に変化して、かつ、１／ｆ（秒）の時間間隔ごとに上記の位相の変化（０度から３６０度までの直線的な変化）を繰り返すことである。ここでのｆの単位はＨｚである。

[0170] 図４７～図５０を用いて、時間ずれに起因する位相ずれを補正する方法の一例について説明する。図４７（ａ）には、ＦＦＴ分析部１０３（ｋ）が求めた周波数スペクトルが模式的に示されており、図４７（ｂ）には、図４７（ａ）から分離された周波数スペクトルの位相成分が模式的に示されており、図４７（ｃ）には、図４７（ａ）から分離された周波数スペクトルの大き

さの成分が模式的に示されている。図４７（ａ）、図４７（ｂ）、及び図４７（ｃ）の横軸は時間軸（秒）である。図４７（ａ）の表示の方法は図１６（ｂ）と同様であるため説明を繰り返さない。図４７（ｂ）の縦軸は周波数スペクトルの位相を表しており０度から３６０度の間の値で示される。図４７（ｃ）の縦軸は周波数スペクトルの大きさを表している。周波数スペクトルの位相Ｐ及び大きさＬは、周波数スペクトル実部を

[0171] [数22]

$$x_t$$

と表すこととして、周波数スペクトルの虚部を

[0172] [数23]

$$y_t$$

と表すこととすると、

[0173] [数24]

$$P_t = \arctan(y_t / x_t)$$

及び

[0174] [数25]

$$L_t = \sqrt{x_t^2 + y_t^2}$$

である。ここでの記号ｔは周波数スペクトルの時刻を表している。

[0175] ここで、図４７（ｂ）に示されている周波数スペクトルの位相成分を用いて、時間ずれに起因する位相ずれを補正する。

[0176] 初めに、基準の時刻を決定する。図４８（ａ）は、図４７（ｂ）と同じ内容のものであり、この例では、図４８（ａ）の黒丸印の時刻ｔ０を基準の時刻に決定している。

[0177] 次に、位相を補正する周波数スペクトルの複数の時刻を決定する。この例

では、図４８（ａ）の５個の白丸印の時刻（ｔ１、ｔ２、ｔ３、ｔ４、ｔ５）を、位相を補正する周波数スペクトルの時刻に決定している。

[0178]　ここで、基準の時刻における周波数スペクトルの位相を

[0179] [数26]

$$P_{t_0} = \arctan\left(y_{t_0} / x_{t_0}\right)$$

と表すこととして、位相を補正する５個の時刻における周波数スペクトルの位相を

[0180] [数27]

$$P_{t_i} = \arctan\left(y_{t_i} / x_{t_i}\right) \quad (i = 1,2,3,4,5)$$

と表すこととする。これらの補正する前の位相を図４８（ａ）において×印で示してある。また、対応する時刻の周波数スペクトルの大きさＬを

[0181] [数28]

$$L_{t_i} = \sqrt{x_{t_i}^2 + y_{t_i}^2} \quad (i = 0,1,2,3,4,5)$$

とする。

[0182]　次に、図４９Ａ及び図４９Ｂに、時刻ｔ２における周波数スペクトルの位相を補正する方法を示す。図４９Ａと図４８（ａ）とは同じ内容のものである。また、図４９Ｂと図４６とは同じ内容のものであり、位相が等速度でかつ１／ｆの時間間隔の間に３６０度回転するという条件を示している。ここで、補正したあとの位相を

[0183] [数29]

$$P'_{t_i} \quad (i = 0,1,2,3,4,5)$$

と表すこととする。図４９Ｂにおいて、上記条件における、基準の時刻であ

る時刻ｔ０と時刻ｔ２との位相ずれを比較すると、時刻ｔ２の位相は時刻ｔ０の位相よりΔＰだけ大きい。そこで、図４９Ａにおいて、基準の時刻である時刻ｔ０の位相Ｐｔ０との時間ずれに起因する位相ずれを補正するために、時刻ｔ２の位相Ｐｔ２からΔＰを差し引いてＰ’ ｔ２を求める。これが補正後の時刻ｔ２の位相である。また、時刻ｔ０の位相は基準の時刻における位相であるので補正後も同じ値となる。具体的には、補正後の位相を

[0184] [数30]

$$P'_{t_0} = P_{t_0}$$

[0185] [数31]

$$P'_{t_i} = P_{t_i} - \left(\left(t_i - t_0\right)\mathrm{mod}\left(1/f\right)\right)/360 \quad \left(i = 1,2,3,4,5\right)$$

により求める。ただし、位相が０度から３６０度の範囲に入るように調整して求めている。補正前の位相が１／ｆの時間間隔の間に３６０度回転していると仮定すると、補正後の位相はすべてＰｔ０になる。

[0186] 補正したあとの周波数スペクトルの位相を図４８（ｂ）に×印で示す。図４８（ｂ）の表示の方法は図４８（ａ）と同様であるため説明を繰り返さない。

[0187] 最後に、補正された周波数スペクトルの位相Ｐ’と補正前の周波数スペクトルの大きさＬとを用いて補正後の周波数スペクトルを求める。ここでは、補正後の周波数スペクトル実部を

[0188] [数32]

$$x'_{t_i} \quad \left(i = 0,1,2,3,4,5\right)$$

と表すこととして、補正後の周波数スペクトルの虚部を

[0189]

[数33]

$$y'_{t_i} \quad (i = 0,1,2,3,4,5)$$

と表すこととする。

[0190] 図５０に、補正後の周波数スペクトルを求める方法の一例を示す。図５０（ａ）は、図４８（ｂ）と同じ内容のものである。図５０（ｂ）は、図４７（ｃ）と同じ内容のものであり、時刻ｔ０、ｔ１、ｔ２、ｔ３、ｔ４、ｔ５における補正前の周波数スペクトルの大きさを×印で示している。補正された周波数スペクトルの位相Ｐ’と補正前の周波数スペクトルの大きさＬを用いて

[0191] [数34]

$$x'_{t_i} = L_{t_i} \cos\left(P'_{t_i}\right) \quad (i = 0,1,2,3,4,5)$$

[0192] [数35]

$$y'_{t_i} = L_{t_i} \sin\left(P'_{t_i}\right) \quad (i = 0,1,2,3,4,5)$$

により補正後の周波数スペクトルを求める。

[0193] 図５０（ｃ）に、補正後の周波数スペクトルを×印で示す。図５０（ｃ）の表示の方法は、図４７（ａ）と同じであるので説明を繰り返さない。補正後の周波数スペクトルは、位相が等速度でかつ１／ｆの時間間隔の間に３６０度回転するという条件のもとで、時間ずれに起因する位相ずれが補正された周波数スペクトルである。

[0194] 次に、検知音特定部２８０２（ｋ）（ｋ＝１～３）は、対応するＦＦＴ分析部１０３（ｋ）（ｋの値が同じもの）が求めた周波数スペクトルと、対応する位相補正部２８０１（ｋ）（ｋの値が同じもの）が位相ずれを補正した周波数スペクトルとを用いて、周波数帯域ｊ（ｊ＝１～Ｍ）ごとに検知音特定部２８０２（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）を用いて、位相ずれが

補正された分析の対象とする時刻の周波数スペクトルと、位相ずれが補正された、少なくとも分析の対象とする時刻とは異なる複数の時刻における周波数スペクトルとの類似度の総和を求める。検知音特定部２８０２（ｋ）（ｋ＝１～３）は、類似度の総和が第１のしきい値以上となる分析の対象とする時刻を求めて、第１のしきい値以上となる時刻のＦＦＴ分析部１０３（ｋ）が求めた周波数スペクトルを、検知音の時刻部分の周波数スペクトルと特定することにより雑音を除去する（ステップＳ３１０１（ｋ）、ステップＳ３１０１（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ））。

[0195] 初めに周波数スペクトル特定部３０００（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）は、位相補正部２８０１（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）が求めた位相ずれが補正された周波数スペクトルから、類似度算出部３００１（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）が類似度の総和の算出に用いる周波数スペクトルを特定する（ステップＳ３３００（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ））。ここでは、分析の対象とする時刻をｔ０として、分析の対象とする時刻の補正後の周波数スペクトルとの類似度の総和を求める補正後の周波数スペクトルの時刻をｔ１、ｔ２、ｔ３、ｔ４、ｔ５とする。ここでの類似度の総和を求めるために用いる補正後の周波数スペクトルの時刻および時間幅は、抽出する音の特徴に基づいて決定する。

[0196] 次に、類似度算出部３００１（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）が、周波数スペクトル特定部３０００（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）が特定した補正後の周波数スペクトルを用いて類似度の総和を算出する（ステップＳ３３０１（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ））。以下、類似度の総和の算出方法について説明する。この例では、分析の対象とする時刻をｔ０として、分析の対象とする時刻の補正後の周波数スペクトルとの類似度の総和を求める補正後の周波数スペクトルの時刻をｔ１、ｔ２、ｔ３、ｔ４、ｔ５としている。ここでは、類似度として相関値を用いることとする。類似度の総和Ｓは、

[0197]

[数36]

$$S=\sum_{i=1}^{i=5}\left(x'_{t_0}\times x'_{t_i}+y'_{t_0}\times y'_{t_i}\right)$$

を用いて算出される。

[0198] ここで、その他の類似度の総和Ｓの計算方法を以下に示す。相関値の計算において、総和した類似度の数で正規化する方法である

[0199] [数37]

$$S=\frac{1}{5}\left(\sum_{i=1}^{i=5}\left(x'_{t_0}\times x'_{t_i}+y'_{t_0}\times y'_{t_i}\right)\right)$$

や、分析の対象とする時刻の周波数スペクトルとの類似度も加える方法である

[0200] [数38]

$$S=\sum_{i=0}^{i=5}\left(x'_{t_0}\times x'_{t_i}+y'_{t_0}\times y'_{t_i}\right)$$

や、周波数スペクトルの大きさで正規化された相関値を用いる方法である

[0201] [数39]

$$S=\sum_{i=1}^{i=5}\frac{\left(x'_{t_0}\times x'_{t_i}+y'_{t_0}\times y'_{t_i}\right)}{\sqrt{x'^2_{t_0}+y'^2_{t_0}}\sqrt{x'^2_{t_i}+y'^2_{t_i}}}$$

や、周波数スペクトルの距離の逆数を用いる方法である

[0202] [数40]

$$S=1\Big/\left(\sum_{i=1}^{i=5}\sqrt{\left(x'_{t_0}-x'_{t_i}\right)^2+\left(y'_{t_0}-y'_{t_i}\right)^2}+\alpha\right)$$

や、位相成分の距離の逆数を用いる方法である

[0203] [数41]

$$S = 1\Big/\left(\sum_{i=1}^{i=5}\sqrt{\left(\arctan\left(y'_{t_0}/x'_{t_0}\right)-\arctan\left(y'_{t_i}/x'_{t_i}\right)\right)^2}+\alpha\right) = 1\Big/\left(\sum_{i=1}^{5}\sqrt{\left(P'_{t_0}-P'_{t_i}\right)^2}+\alpha\right)$$

などがある。ここで

[0204] [数42]

$$\alpha$$

はＳが無限大に発散しないための予め定められた小さな値である。

[0205] また、分析の対象とする時刻をｔ２として、分析の対象とする時刻の補正後の周波数スペクトルとの類似度の総和を求める補正後の周波数スペクトルの時刻をｔ０、ｔ１、ｔ３、ｔ４、ｔ５としたときの類似度の総和の一例を以下に示す。

[0206] [数43]

$$S = \sum_{i=0}^{i=1}\left(x'_{t_2}\times x'_{t_i}+y'_{t_2}\times y'_{t_i}\right)+\sum_{i=3}^{i=5}\left(x'_{t_2}\times x'_{t_i}+y'_{t_2}\times y'_{t_i}\right)$$

この例では、周波数スペクトル特定部３０００（ｋ，ｊ）が、位相補正部２８０１（ｋ，ｊ）が求めた位相ずれが補正された周波数スペクトルから、類似度算出部３００１（ｋ，ｊ）が類似度の総和の算出に用いる周波数スペクトルを特定したが、他の方法として、位相補正部２８０１（ｋ，ｊ）が位相ずれを補正する周波数スペクトルを、あらかじめ周波数スペクトル特定部３０００（ｋ，ｊ）が特定しておいて、位相補正部２８０１（ｋ，ｊ）によりすでに特定された位相ずれが補正された周波数スペクトルを用いて、類似度算出部３００１（ｋ，ｊ）が類似度の総和を求めるようにしてもよい。

[0207] 次に、検知音特定部３００２（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）は、類似度の総和Ｓが第１のしきい値以上となる時刻の、対応するＦＦＴ分析部１０３（ｋ）（ｋの値が同じもの）が求めた周波数スペクトルを、検知音の時刻部分の周波数スペクトルと特定する（ステップＳ３３０２（ｋ，ｊ）（

ｋ＝１～３、ｊ＝１～Ｍ））。

[0208] 第１のしきい値をＡと表すとすると、

[0209] [数44]

$$S \geq A$$

の条件を満たす時刻の周波数スペクトルを、対応するＦＦＴ分析部１０３（ｋ）（ｋの値が同じもの）が求めた周波数スペクトルから特定する。

[0210] 最後に、方向検知部１０５は、検知音特定部２８０２（ｋ）（ｋ＝１～３）が特定した検知音の周波数スペクトルを用いて、３個のマイクロホンから得られるマイクロホンの組（マイクロホン１０２（１）とマイクロホン１０２（２）、マイクロホン１０２（２）とマイクロホン１０２（３）、マイクロホン１０２（１）とマイクロホン１０２（３））において、それぞれのマイクロホンに対応した特定された周波数スペクトルが共通に存在する部分の周波数スペクトル同士の相互相関値を算出する。方向検知部１０５は、相互相関値の極大値から到達時間差を求め、求められた到達時間差とマイクロホン同士の間隔とから音源方向を特定して、音源方向１０６を示す情報を出力する（ステップＳ６０２）。この例では、相互相関値の極大値から到達時間差を求めたが、相互相関値の最大値や、しきい値以上の相互相関値の極大値や、差分誤差の極小値や、差分誤差の最小値や、しきい値以下の差分誤差の極小値から到達時間差を求めてもよい。

[0211] かかる構成によれば、割合判定部５００（１，ｊ）が、全ての音源の方向に対応する相互相関値に対して、前記割合が前記第２のしきい値以上のときにのみ音源方向を特定する。このため、雑音の影響により、実際の音源の方向に対応する相互相関値が算出できない場合においても、誤った音源の方向を特定することを回避することができる。

[0212] また、検知音特定部２８０２（ｋ）が、位相成分が歪んだ雑音部分（暗騒音などの白色雑音の部分、または混合音が原因で歪んだ雑音の部分）を特定して、検知音の時刻部分を精度良く特定できる。

[0213] また、検知音の時刻部分を特定するときに、位相補正部２８０１（ｋ）で時間ずれに起因する位相ずれを補正できる。このため、時間ずれに起因する位相ずれが起こらない１／ｆの時間間隔の時刻の周波数スペクトルを必ずしも用いる必要がない。このため、１／ｆの時間間隔よりも短い時間の音響信号に対しても雑音を特定できるという効果もある。

[0214] また、方向検知部が、相互相関値の算出に用いる時間区間における、それぞれのマイクロホンに対応した特定された時刻部分の周波数スペクトルが共通に存在する部分の割合が第２のしきい値よりも小さいときに、その相互相関値に対応する方向を音源方向の候補から除去する。このことで、誤差が大きくなる音源方向の候補を除去してから音源の方向を特定することができる。

[0215] また、検知音の時刻部分を特定するときに、類似度として相関値を用いる。このことで、周波数スペクトルの大きさのずれには着目せずに位相のずれを判定することができる。このため、周波数スペクトルの大きさが小さい音響信号に対しても位相成分が歪んだ雑音部分を特定することができるという効果がある。また、方向検知部における相互相関値との尺度における整合が取れる。

[0216] 本発明の雑音部分の特定に関する考え方は、暗騒音などの白色雑音または混合音が原因で歪んだ雑音部分は、位相成分が歪んでおり、類似度の総和Ｓが第１のしきい値Ａよりも小さくなるという考え方に基づいている。

[0217] なお、本実施の形態において、３個のマイクロホンのそれぞれに対応した３個のＦＦＴ分析部と３個の検知音特定部とを設けたが、１つのＦＦＴ分析部と１つの検知音特定部とを共有して各マイクロホンが入力した音響信号を処理しても良い。

[0218] なお、本実施の形態において、時間及び周波数ごとに音源の方向を特定したが、時間軸方向に音源の方向のヒストグラムをとり最大値をとる音源方向を音源の方向に特定しても良いし、時間軸方向に音源の方向を平均化して音源の方向を特定しても良い。

[0219] また、本実施の形態において、時間及び周波数ごとに音源の方向を特定したが、周波数軸方向に音源の方向のヒストグラムをとり最大値をとる音源方向を音源の方向に特定しても良いし、周波数方向に音源の方向を平均化して音源の方向を特定しても良い。

[0220] さらに、本実施の形態において、マイクロホンの数を３個としたが２個以上あれば良い。

[0221] さらにまた、ＦＦＴ分析部におけるＦＦＴの代わりに、ＤＦＴ、コサイン変換、ウェーブレット変換、又は、バンドパスフィルタを用いても良い。

[0222] なお、本実施の形態において、位相補正部２８０１（ｋ）と検知音特定部２８０２（ｋ）を用いて、ＦＦＴ分析部１０３（ｋ）が求めた全て（Ｍ個）の周波数帯域の周波数スペクトルに対して雑音の除去を行い音源の方向を特定したが、一部の周波数帯域を選択し、選択した周波数帯域において雑音の除去を行い音源の方向を特定しても良い。

[0223] また、ＦＦＴ分析部は、ΔＴの幅の時間窓を掛けた前記音響信号から前記周波数スペクトルを求めて、前記検知音特定部は、分析の対象とする時刻の前記周波数スペクトルと、分析の対象とする時刻をはさみΔＴの時間間隔よりも離れた複数の時刻を含む複数の時刻の前記周波数スペクトルとの類似度の総和を用いて検知音の時刻部分を特定してもよい。本構成によって、分析の対象とする時刻をはさみΔＴの時間間隔（周波数スペクトルを求めるときの時間窓の幅）よりも離れた時刻を含む複数の時刻の前記周波数スペクトルとの類似度の総和を用いることで、時間分解能（ΔＴ）を細かくしたときの影響による周辺の周波数からの周波数もれによる雑音（歪み）を除去することができる。

[0224] （実施の形態２の変形例）

次に、実施の形態２に示した音源方向検知装置の変形例について説明する。

[0225] 本変形例に係る音源方向検知装置は、図４０～図４２を参照して説明した実施の形態２に係る音源方向検知装置と同様の構成を有する。ただし、検知

音特定部２８０２（ｋ）（ｋ＝１～３）が実行する処理が異なる。

[0226] 検知音特定部２８０２（ｋ）（ｋ＝１～３）は、対応するＦＦＴ分析部１０３（ｋ）（ｋの値が同じもの）が求めた周波数スペクトルと、対応する位相補正部２８０１（ｋ）（ｋの値が同じもの）が位相ずれを補正した周波数スペクトルとを用いて、周波数帯域ｊ（ｊ＝１～Ｍ）ごとに検知音特定部２８０２（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ）を用いて、位相ずれが補正された周波数スペクトルを用いて、類似度の空間を、上記周波数スペクトルの位相成分のヒストグラムを作成することにより求める。検知音特定部２８０２（ｋ）（ｋ＝１～３）は、類似度がしきい値以上となる周波数スペクトルのかたまりを、検知音の周波数スペクトルと特定する。

[0227] 類似度算出部３００１（ｋ，ｊ）は、周波数スペクトル特定部３０００（ｋ，ｊ）が特定した位相ずれが補正された周波数スペクトルの位相成分のヒストグラムを作成することにより類似度の空間を求める。検知音特定部３００２（ｋ，ｊ）は、類似度算出部３００１（ｋ，ｊ）が求めた類似度の空間から類似度がしきい値以上となるＦＦＴ分析部１０３（ｋ）が求めた周波数スペクトルのかたまりを、検知音の周波数スペクトルと特定する。

[0228] 次に、以上のように構成された音源方向検知装置２８００の動作について説明する。音源方向検知装置２８００の処理手順を示すフローチャートは、実施の形態２と同様であり、図４３～図４５に示されている。

[0229] 以下の説明では、音源方向検知装置２８００のうち検知音特定部２８０２（ｋ）（ｋ＝１～３）が実行する処理についてのみ説明する。その他の処理部が実行する処理は実施の形態２と同様であるため、説明を繰り返さない。

[0230] 検知音特定部２８０２（ｋ，ｊ）は、位相補正部２８０１（ｋ，ｊ）が求めた位相ずれが補正された周波数スペクトルを用いて、類似度の空間を、上記周波数スペクトルの位相成分のヒストグラムを作成することにより求める。検知音特定部２８０２（ｋ，ｊ）は、類似度がしきい値以上となるＦＦＴ分析部１０３（ｋ）が求めた周波数スペクトルのかたまりを、検知音の周波数スペクトルと特定することにより雑音を除去する（ステップＳ３１０１（

ｋ）、ステップＳ３１０１（ｋ，ｊ）（ｋ＝１～３、ｊ＝１～Ｍ））。

[0231] 類似度算出部３００１（ｋ，ｊ）（ｊ＝１～Ｍ）は、周波数スペクトル特定部３０００（ｋ，ｊ）が特定した位相が補正された後の周波数スペクトルを用いて、上記周波数スペクトルの位相成分のヒストグラムを作成することにより類似度の空間を求める（ステップＳ３３０１（ｋ，ｊ）（ｊ＝１～Ｍ））。

[0232] 補正後の周波数スペクトルの位相成分を、数２６で表すこととする。

[0233] 図５１に、周波数スペクトルの位相成分のヒストグラムを作成する方法の一例を示す。ここでは、位相区間がΔψ（ｉ）（ｉ＝１～４）で、位相が時間に対して２πｆの傾きで変化する帯領域ごとの所定に時間幅における周波数スペクトルの出現頻度を求めることでヒストグラムを作成する。図５１の斜線で示されている部分はΔψ（１）の領域である。ここでは周波数スペクトルの位相は、時間に対して２πｆの傾きがゼロになるように補正されているため横軸と平行な領域になっている。ここで、Δψ（ｉ）（ｉ＝１～４）ごとにそれらの領域に含まれる周波数スペクトルの数をカウントすることでヒストグラムを作成することができる。

[0234] 以下の処理は、実施の形態１の変形例と同じであるので説明を繰り返さない（図３８Ｂ、図３９を参照）。

[0235] かかる構成によれば、類似度の空間を、周波数スペクトルの位相成分のヒストグラムを作成することにより求めて、類似度がしきい値以上となる周波数スペクトルのかたまりを、検知音の周波数スペクトルと特定することで、音声Ａと音声Ｂというように異なる位相に対応するものが複数含まれる場合でも、暗騒音（雑音に対応）を除去して音声Ａ（検知音に対応）と音声Ｂ（検知音に対応）の周波数スペクトルを区別して検知音の周波数スペクトルと特定することができるという効果がある。よって、音源の方向を正確に特定することができる。

[0236] なお、検知音特定手段（検知音特定部）として、以下に示す雑音除去装置を用いても良い。

[0237] 雑音除去装置は、音響信号の、少なくとも１つの周波数ｆにおける、位相成分を含む周波数スペクトルを生成するスペクトル生成手段と、所定の時間幅に含まれる前記周波数スペクトルにおいて、分析の対象とする周波数スペクトルと、前記所定の時間幅に含まれる周波数スペクトルから任意に選択される一定の数以上から構成される周波数スペクトルの組との類似度が、いずれの組に対しても一定の値より大きくならない前記分析の対象とする周波数スペクトルを除去する検知音特定手段とを備え、前記類似度は、時刻ｔの周波数スペクトルの位相をψ（ｔ）（ラジアン）とするときにψ（ｔ）－（２πｆｔ）で定義される空間での位相の類似度であることを特徴とする。

[0238] この構成によると、位相成分が歪んだ雑音部分（暗騒音などの白色雑音の部分、または混合音が原因で歪んだ雑音の部分）を特定して、検知音の時刻部分を精度良く特定することができる。

[0239] 好ましくは、前記スペクトル生成手段は、前記音響信号に所定の時間窓幅の窓関数を掛け合わせ、当該窓関数が掛け合わされた後の音響信号から前記周波数スペクトルを生成し、前記所定の時間窓幅は前記所定の時間幅より小さいことを特徴とする。

[0240] この構成によると、周波数スペクトルを求める際に用いられた所定の時間窓幅よりも離れた時刻を含む周波数スペクトルを用いて類似度を求めることができる。これにより、周波数スペクトルの時間分解能を細かくしたときの影響による周辺の周波数からの周波数もれによる雑音（歪み）を除去することができる。

[0241] このことについて説明を加える。ここでは、音響信号２４０１として、１００Ｈｚと２００Ｈｚと３００Ｈｚとの正弦波の混合音を用いた場合を一例として説明を行う。この例では、混合音中の２００Ｈｚの正弦波において、１００Ｈｚと３００Ｈｚとの正弦波からの周波数もれによる雑音（歪み）を除去することを目的とする。

[0242] 図５２は、この変形例に係る雑音除去装置の構成を示すブロック図である。

[0243] 雑音除去装置１１０は、ＤＦＴ分析部１１０６と、検知音特定部１１３とを含む。ＤＦＴ分析部１１０６および検知音特定部１１３は、コンピュータ上で各処理部の機能を実現するためのプログラムを実行することにより実現される。

[0244] ＤＦＴ分析部１１０６は、入力された音響信号２４０１に対して離散フーリエ変換処理を施し、音響信号２４０１の位相成分を含む周波数スペクトルを求める処理部である。以下では、ＤＦＴ分析部１１０６で求められた周波数スペクトルの周波数帯域の個数をＭとして、それらの周波数帯域を指定する番号を記号ｊで表すこととする。

[0245] 検知音特定部１１３は、検知音特定部１１３（ｊ）（ｊ＝１～Ｍ）を含む。検知音特定部１１３は、ＤＦＴ分析部１１０６が求めた周波数スペクトルに対して、周波数帯域ｊ（ｊ＝１～Ｍ）ごとに検知音特定部１１３（ｊ）（ｊ＝１～Ｍ）を用いて雑音の除去を行う処理部である。以下では、ｊ番目の周波数帯域がｆ（例えば、中心周波数がｆ）である場合について説明を行う。検知音特定部１１３（ｊ）は、１／ｆの時間間隔の時刻から選択される複数の時刻における周波数スペクトルを用いて、上記複数の時刻のうち分析の対象とする時刻の周波数スペクトルと、上記複数の時刻のうち少なくとも上記時刻とは異なる複数の時刻における周波数スペクトルとの類似度の総和を求める。また、検知音特定部１１３（ｊ）は、当該類似度の総和が一定のしきい値以下となる上記分析の対象とする時刻の周波数スペクトルを特定して、特定した周波数スペクトルを除去することにより雑音を除去する。そして、検知音特定部１１３は、雑音が除去されたＭ個の周波数帯域における周波数スペクトルをまとめることで出力スペクトル２４０８を作成する。

[0246] 検知音特定部１１３（ｊ）（ｊ＝１～Ｍ）は、図４２に示した検知音特定部２８０２（ｋ，ｊ）と同様の構成を有する。

[0247] 図５３に、１００Ｈｚと２００Ｈｚと３００Ｈｚとの正弦波の混合音を用いた場合の、２００Ｈｚにおける周波数スペクトルの時間波形の一例を示す。図５３（ａ）には２００Ｈｚにおける周波数スペクトルの実部の時間波形

が、図５３（ｂ）には２００Ｈｚにおける周波数スペクトルの虚部の時間波形が示されている。横軸は時間軸（秒）であり縦軸は周波数スペクトルの振幅を表す。ここでは５０ｍｓの時間長の時間波形が示されている。

[0248] 図５４に、図５３に示した音響信号２４０１を作成するときに用いた２００Ｈｚの正弦波の、２００Ｈｚにおける周波数スペクトルの時間波形を示す。表示の方法は図５３と同じであるため説明を繰り返さない。

[0249] 図５３と図５４とから、音響信号２４０１において、２００Ｈｚの正弦波が、１００Ｈｚと３００Ｈｚとの正弦波からの周波数もれの影響により歪んでいる部分（雑音部分）が存在することがわかる。

[0250] 図５５は、雑音除去装置１１０の動作手順を示すフローチャートである。

[0251] 初めに、ＤＦＴ分析部１１０６は、音響信号２４０１を入力として受け、入力された音響信号２４０１に対して離散フーリエ変換処理を施し、音響信号２４０１の位相成分を含む周波数スペクトルを求める（ステップＳ３００）。この例における離散フーリエ変換処理の条件としては、サンプリング周波数＝１６０００Ｈｚの音響信号２４０１に対して時間窓の幅ΔＴ＝５ｍｓ（８０ｐｔ）のハニング窓を用いることで処理している。また、時間軸方向には１ｐｔ（０．０６２５ｍｓ）の時間シフトを行いながら各時刻における周波数スペクトルを求めている。この処理結果における周波数スペクトルの時間波形を表示したものが図５３である。

[0252] 次に、検知音特定部１１３は、ＤＦＴ分析部１１０６が求めた周波数スペクトルに対して、周波数帯域ｊ（ｊ＝１～Ｍ）ごとに検知音特定部１１３（ｊ）（ｊ＝１～Ｍ）を用いて雑音の除去を行う（ステップＳ３０１（ｊ）（ｊ＝１～Ｍ））。この例では、Ｍ＝１であり、ｊ＝１番目の周波数帯域の周波数はｆ＝２００Ｈｚである。

[0253] 検知音特定部１１３（１）（ｊ＝１）では、１／ｆの時間間隔の時刻から選択された複数の時刻における周波数スペクトルを用いて、分析の対象とする時刻の周波数スペクトルと、少なくとも上記時刻とは異なる複数の時刻における周波数スペクトルとの類似度の総和が一定のしきい値以下となる分析

の対象とする時刻の周波数スペクトルを特定して、特定した周波数スペクトルを除去することにより雑音を除去する（ステップＳ３０１（１）（ｊ＝１））。

[0254] 初めに、検知音特定部１１３（１）（ｊ＝１）の周波数スペクトル特定部が、実施の形態１に示した例と同様にして、１／ｆの時間間隔の時刻から選択された複数の時刻における周波数スペクトルを特定する（ステップＳ４００（１）（ｊ＝１））。

[0255] ここで、実施の形態１に示した例と異なる部分は、検知音特定部１１３（１）（ｊ＝１）の類似度算出部が類似度の総和の算出に用いる周波数スペクトルの時間範囲である。実施の形態１に示した例では、時間範囲は２０ｍｓであり、周波数スペクトルを求めるときに用いた時間窓の幅ΔＴ（＝６４ｍｓ）よりも時間幅が短かった。この例では、時間範囲を１００ｍｓとしており、周波数スペクトルを求めるときに用いた時間窓の幅ΔＴ（＝５ｍｓ）よりも時間幅が長く設定されている。

[0256] 次に、検知音特定部１１３（１）（ｊ＝１）の類似度算出部が、検知音特定部１１３（１）（ｊ＝１）の周波数スペクトル特定部が特定した全ての周波数スペクトルを用いて類似度の総和を算出する（ステップＳ４０１（１）（ｊ＝１））。ここでの処理は実施の形態１に示した例の処理と同じであるので説明を繰り返さない。

[0257] 最後に、検知音特定部１１３（１）（ｊ＝１）の検知音特定部が、類似度の総和Ｓがしきい値以下となる分析の対象とする時刻の周波数スペクトルを特定して、特定した周波数スペクトルを除去することにより雑音を除去する（ステップＳ４０２（１）（ｊ＝１））。ここでの処理は実施の形態１に示した例の処理と同じであるので説明を繰り返さない。

[0258] 図５６に、図５３に示した音響信号２４０１から抽出された２００Ｈｚにおける周波数スペクトルの時間波形を示す。表示方法において図５３と共通する部分の説明は繰り返さない。図５６において、斜線部分の領域は雑音部分として除去された領域である。図５６と、図５３及び図５４とを比較する

と、音響信号２４０１から１００Ｈｚと３００Ｈｚとの正弦波からの周波数もれによる雑音（歪み）を除去されて、２００Ｈｚの正弦波が抽出されていることがわかる。

[0259] 実施の形態１および実施の形態１の変形例にかかる構成によれば、分析の対象とする時刻における周波数スペクトルと分析の対象とする時刻をはさみ、かつΔＴの時間間隔（周波数スペクトルを求めるときの時間窓の幅）よりも離れた時刻を含む複数の時刻の周波数スペクトルとの類似度の総和を用いることで、時間分解能（ΔＴ）を細かくしたときの影響による周辺の周波数からの周波数もれによる雑音（歪み）を除去することができるという効果がある。

[0260] 好ましくは、前記検知音特定手段は、前記所定の時間幅に含まれる周波数スペクトルから１／ｆの時間間隔の時刻における周波数スペクトルを特定し、前記特定された周波数スペクトルを用いて類似度を求めることを特徴とする。

[0261] この構成によると、ψ（ｔ）－（２πｆｔ）で定義される空間での位相の類似度をψ（ｔ）で定義される空間で求めることができるため、類似度の算出方法が簡単になり、位相の補正を行うための計算量を削減することができる。

[0262] 好ましくは、雑音除去装置は、さらに、時刻ｔの周波数スペクトルの位相ψ（ｔ）（ラジアン）をψ（ｔ）-（２πｆｔ）に補正する位相補正手段を備え、前記検知音特定手段は、前記位相が補正された前記周波数スペクトルを用いて前記類似度を求めることを特徴とする。

[0263] この構成によると、１／ｆの時間間隔よりも細かい時間間隔で音響信号を詳細に分析できる。これにより、１／ｆの時間間隔よりも短い時間の音響信号に対しても雑音を除去できる。

[0264] 好ましくは、前記類似度は、分析の対象とする周波数スペクトルと、前記周波数スペクトルの組を構成する各周波数スペクトルとの各々の類似度の総和であることを特徴とする。

[0265] この構成によると、周波数スペクトルの組を構成する各周波数スペクトルの１つ１つの類似度を詳細に求めることができる。

[0266] さらに好ましくは、前記類似度は時刻ｔの周波数スペクトルの位相ψ（ｔ）（ラジアン）とするときにψ（ｔ）－（２πｆｔ）で定義される空間での、位相区間ごとの前記所定の時間幅に含まれる周波数スペクトルの出現頻度であり、分析の対象とする周波数スペクトルと、前記出現頻度を求める全ての周波数スペクトルとの各々の類似度を、類似度を求める２つの周波数スペクトルが同一の位相区間に存在する場合は「１」、異なる位相区間に存在する場合は「０」として前記類似度の総和を求めることを特徴とする。

[0267] この構成によると、音響信号中の位相が異なる複数の音声（検知音）が含まれている場合でも、位相成分が歪んだ雑音部分を除去して複数の音声の周波数スペクトルを区別して抽出することができる。

[0268] ここで別の音響信号における分析結果についての考察を行う。図５７に、ｆ＝２００Ｈｚの周波数を分析した結果を示す。図５８に、ｆ＝１５０Ｈｚの周波数を分析した結果を示す。図５７、図５８ともに、音響信号として２００Ｈｚの正弦波と白色雑音とを用いて分析した結果である。

[0269] 図５７（ａ）は、音響信号として２００Ｈｚの正弦波を用いて位相を分析した結果である。ここでは位相の補正は行っていない。この結果より、所定の時間幅において、位相は時刻に対して２π×２００の傾きで変化している。ここでは、２００Ｈｚの周波数の分析を行っており、位相が時刻に対して２π×２００の傾きで変化する直線と音響信号の位相との所定の時間幅における類似度は大きくなる。そして、一定の数以上の音響信号（所定の時間幅における全ての音響信号）との類似度は一定の値より大きくなる。このため、音響信号である２００Ｈｚの正弦波は除去されずに抽出することができる。ここでの類似度は、例えば、所定の時間幅においてこの直線と音響信号の位相との誤差（位相は０度と３６０度が同じでトーラス上になっていることを考慮した誤差）を求めて逆数をとれば求めることができる。

[0270] 図５７（ｂ）は、音響信号として２００Ｈｚの正弦波を用いて位相を分析

した結果である。ここでは位相の補正を行っている。この結果より、所定の時間幅において、位相は時刻に対してΔψの幅（この例ではπ／２）で一定の値を持つことがわかる。これより、位相が時刻に対して２π×２００の傾きで変化する直線を位相補正した傾きゼロの直線と音響信号の位相補正した位相との所定の時間幅における類似度は大きくなる。そして、一定の数以上の音響信号（所定の時間幅における全ての音響信号）との類似度は一定の値より大きくなる。このため、音響信号である２００Ｈｚの正弦波は除去されずに抽出することができる。

[0271] 図５７（ｃ）は、音響信号として白色雑音を用いて位相を分析した結果である。分析する時間幅や類似度を求める周波数スペクトルの数の下限値や類似度のしきい値などの分析条件は図５７（ａ）と同じである。ここでは位相の補正は行っていない。この結果より、所定の時間幅において、位相は時刻に対して２π×２００の傾きで変化していないことがわかる。ここでは、２００Ｈｚの周波数の分析を行っており、位相が時刻に対して２π×２００の傾きで変化する直線と音響信号の位相との所定の時間幅における類似度は小さくなる。そして、一定の数以上の音響信号（所定の時間幅における全ての音響信号）との類似度は一定の値以下になる。このため、白色雑音を除去することができる。図５７（ａ）の結果と合わせると、２００Ｈｚの正弦波を抽出して白色雑音を除去することができる。

[0272] 図５７（ｄ）は、音響信号として白色雑音を用いて位相を分析した結果である。分析する時間幅や類似度を求める周波数スペクトルの数の下限値や類似度のしきい値などの分析条件は図５７（ｂ）と同じである。ここでは位相の補正を行っている。この結果より、所定の時間幅において、位相は時刻に対してΔψの幅（この例ではπ／２）で一定の値を持たないことがわかる。このため、一定の数以上の音響信号（所定の時間幅における全ての音響信号）との類似度は一定の値以下になり白色雑音を除去することができる。図５７（ｂ）の結果と合わせると、２００Ｈｚの正弦波を抽出して白色雑音を除去することができる。ただし、２００Ｈｚの正弦波を抽出する場合は、Δψ

の幅をπ／３やπ／４と小さくしても白色雑音を除去することができる。

[0273] 図５８（ａ）は、音響信号として２００Ｈｚの正弦波を用いて位相を分析した結果である。分析する時間幅や類似度を求める周波数スペクトルの数の下限値や類似度のしきい値などの分析条件は図５７（ａ）と同じである。ここでは位相の補正は行っていない。この結果より、所定の時間幅において、位相は時刻に対して２π×２００の傾きで変化している。ここでは、１５０Ｈｚの周波数の分析を行っており、位相が時刻に対して２π×１５０の傾きで変化する直線と音響信号の位相との所定の時間幅における類似度は小さくなる。そして、一定の数以上の音響信号（所定の時間幅における全ての音響信号）との類似度は一定の値以下になる。このため、音響信号である２００Ｈｚの正弦波からの１５０Ｈｚの周波数帯域への周波数もれを除去することができる。

[0274] また、別の効果として、ここでの分析は中心周波数を１５０Ｈｚとして周波数分析を行っているが２００Ｈｚを含む周波数分析になっている。このとき、中心周波数１５０Ｈｚの周波数スペクトルに対して、位相が時刻に対して２π×２００の傾きで変化する直線と音響信号の位相との所定の時間幅における類似度を求め、類似度が一定の値より大きい周波数スペクトルを抽出すれば２００Ｈｚの周波数スペクトルを抽出することができる。また、中心周波数１５０Ｈｚの周波数スペクトルから音響信号に２００Ｈｚの周波数スペクトルが存在するか否かも求めることができる。

[0275] 図５８（ｂ）は、音響信号として２００Ｈｚの正弦波を用いて位相を分析した結果である。分析する時間幅や類似度を求める周波数スペクトルの数の下限値や類似度のしきい値などの分析条件は図５７（ｂ）と同じである。ここでは位相の補正を行っている。ここでは、１５０Ｈｚの周波数の分析を行っている。この結果より、所定の時間幅において、位相は時刻に対してΔψの幅（この例ではπ／２）で一定の値を持たないことがわかる。このため、一定の数以上の音響信号（所定の時間幅における全ての音響信号）との類似度は一定の値以下になる。このため、音響信号である２００Ｈｚの正弦波か

らの１５０Ｈｚの周波数帯域への周波数もれを除去することができる。

[0276] 図５８（ｃ）は、音響信号として白色雑音を用いて位相を分析した結果である。分析する時間幅や類似度を求める周波数スペクトルの数の下限値や類似度のしきい値などの分析条件は図５７（ａ）と同じである。ここでは位相の補正は行っていない。この結果より、所定の時間幅において、位相は時刻に対して２π×１５０の傾きで変化していないことがわかる。このため、音響信号の位相同士の類似度は一定の値以下になり白色雑音を除去することができる。

[0277] 図５８（ｄ）は、音響信号として白色雑音を用いて位相を分析した結果である。ここでは位相の補正を行っている。この結果より、所定の時間幅において、位相は時間に対してΔψの幅（この例ではπ／２）で一定の値を持っていないことがわかる。このため、一定の数以上の音響信号（所定の時間幅における全ての音響信号）との類似度は一定の値以下になり白色雑音を除去することができる。

[0278] 図５９に、バイク音を分析した結果を示す。図５９（ａ）は、バイク音のスペクトログラムを示しており黒い部分がバイク音の部分である。バイクが通過したときのドップラーシフトが現れている。図５９（ｂ）、図５９（ｃ）、図５９（ｄ）ともに、位相補正を行った場合の位相の時間変化を示している。分析する時間幅や類似度を求める周波数スペクトルの数の下限値や類似度のしきい値などの分析条件は図５７（ｄ）、図５８（ｄ）と同じである。

[0279] 図５９（ｂ）は、バイク音の周波数スペクトルの部分である１２０Ｈｚの周波数帯域を分析した結果であり、位相は時間に対してΔψの幅（この例ではπ／２）で一定の値を持つことがわかる。これより、位相が時刻に対して２π×１２０の傾きで変化する直線を位相補正した傾きゼロの直線と音響信号の位相補正した位相との所定の時間幅における類似度は大きくなる。図５７（ｄ）、図５８（ｄ）の結果と合わせると、バイク音の周波数スペクトルを特定して白色雑音を除去することができる。

[0280] 図５９（ｃ）は、バイク音の別の周波数スペクトルの部分である１４０Ｈｚの周波数帯域を分析した結果であり、位相は時間に対してΔψの幅（この例ではπ／２）で一定の値を持つことがわかる。これより、位相が時刻に対して２π×１４０の傾きで変化する直線を位相補正した傾きゼロの直線と音響信号の位相補正した位相との所定の時間幅における類似度は大きくなる。図５７（ｄ）、図５８（ｄ）の結果と合わせると、バイク音の周波数スペクトルを特定して白色雑音を除去することができる。

[0281] 図５９（ｄ）は、雑音の周波数スペクトルの部分である８０Ｈｚの周波数帯域を分析した結果であり、位相は時間に対してΔψの幅（この例ではπ／２）で一定の値を持たないことがわかる。これより、位相が時刻に対して２π×８０の傾きで変化する直線を位相補正した傾きゼロの直線と音響信号の位相補正した位相との所定の時間幅における類似度は小さくなる。このため、雑音の周波数スペクトルを除去することができる。

[0282] ここでのΔψの幅は、抽出するバイク音の位相の時間特性から実験的に求めている。このΔψの幅は、この例では、ドップラーシフトのバイク音と別の部分のバイク音とで同じ値を用いておりπ／２とした。また、図５７（ｄ）、図５８（ｄ）の結果より、このΔψの幅を２π／３と大きくしても、白色雑音の位相は一定の値とならないため白色雑音を除去できることがわかる。

[0283] ここで雑音として除去される周波数スペクトルと、除去されずに抽出される周波数スペクトルの典型的なパターンについて説明を加える。図６０Ａ及び図６０Ｂは、所定の時間幅における音響信号の位相を模式的に示したものである。横軸は時間であり縦軸は位相である。丸印で位相補正された音響信号の周波数スペクトルの位相が示されている。実線で囲まれた周波数スペクトル同士は同じクラスタに属しており類似度が一定の値より大きくなる周波数スペクトルの集まりである。これらのクラスタは多変量解析を用いても求めることができる。同一のクラスタの中に一定の数以上の周波数スペクトルが存在するクラスタの周波数スペクトルは除去されずに抽出され、一定の数

より少ない周波数スペクトルしか存在しないクラスタの周波数スペクトルは、一定の数以上の周波数スペクトルとの類似度が一定の値以下になるため雑音として除去される。図６０Ａに示すように、所定の時間幅に一部分だけ雑音部分が含まれる場合に、その一部分の雑音のみを除去することができる。例えば、図６０Ｂに示すように、２つの抽出したい音が存在する場合にも、所定の時間幅に対して４０％以上（ここでは、７個以上）の周波数スペクトルとの類似度が一定の値より大きい周波数スペクトルを特定する（一定の値以下の周波数スペクトルを除去する）ことで２つの音を抽出することができる。

[0284] なお、実施の形態１および実施の形態２において、検知音特定部として、従来の雑音除去部を用いてもよい。例えば、非特許文献１に開示されている雑音除去部を用いてもよい。

[0285] 今回開示された実施の形態はすべての点で例示であって制限的なものではないと考えられるべきである。本発明の範囲は上記した説明ではなくて請求の範囲によって示され、請求の範囲と均等の意味および範囲内でのすべての変更が含まれることが意図される。

産業上の利用可能性

[0286] 本発明にかかる音源方向検知装置は、雑音の影響により誤った方向を出力してしまうことを回避できる。このため、本発明の音源方向検知装置を、例えば、車載用の死角車両（例えば交差点死角に隠れたバイク）を音で検知する装置に組み込めば、実環境である雑音下で死角車両の方向を特定することができて運転者に危険なバイクの存在を知らせることができる。また、本発明の音源方向検知装置を、例えば、テレビ会議装置に組み込めば、雑音下で話者の方向を特定して話者の方向にマイクを向けて音声を収録したり、カメラを向けることができる。また、本発明の音源方向検知装置を、例えば、セキュリティカメラ装置に組み込めば、雑音下で悲鳴や異常音の発生方向を特定してカメラを向けることができる。

請求の範囲

[1] 互いに離間して配置された２つ以上のマイクロホンにて集音された音響信号の各々について、当該音響信号の少なくとも１つの周波数帯域における周波数スペクトルを生成する生成手段と、

前記周波数帯域における周波数スペクトルから、音源方向を求める検知音の周波数スペクトルの時刻部分を特定する検知音特定手段と、

音源方向を検知する時間単位である時間区間の中で、前記検知音特定手段が特定した時刻部分における周波数スペクトルの前記マイクロホン間での一致度合いにより、前記検知音が前記マイクロホンに到達する時間差を求めて、前記時間差と前記マイクロホンの距離と音速とから音源方向を求めて出力する検知手段とを備え、

前記検知手段は、前記時間差がゼロになるように前記２つ以上のマイクロホンからの周波数スペクトルの時間軸を調整したときに、前記音源方向を検知する時間単位である時間区間に占める、前記検知音特定手段により特定された時刻部分のうち前記マイクロホン間で共通の時刻に存在する部分の割合が、所定のしきい値より大きい場合に音源方向を出力する

ことを特徴とする音源方向検知装置。

[2] 前記検知手段は、さらに、前記２つ以上のマイクロホンからの周波数スペクトルの時間軸を調整したときの、前記音源方向を検知する時間単位である時間区間に占める、前記検知音特定手段により特定された時刻部分のうち前記マイクロホン間で共通の時刻に存在する部分の割合が、すべての音源方向の候補について、前記所定のしきい値より大きい場合にのみ音源方向を出力する

ことを特徴とする請求項１に記載の音源方向検知装置。

[3] 互いに離間して配置された２つ以上のマイクロホンにて集音された音響信号の各々について、当該音響信号の少なくとも１つの周波数帯域における周波数スペクトルを生成する生成ステップと、

前記周波数帯域における周波数スペクトルから、音源方向を求める検知音

の周波数スペクトルの時刻部分を特定する検知音特定ステップと、

音源方向を検知する時間単位である時間区間の中で、前記検知音特定ステップにおいて特定された時刻部分における周波数スペクトルの前記マイクロホン間での一致度合いにより、前記検知音が前記マイクロホンに到達する時間差を求めて、前記時間差と前記マイクロホンの距離と音速とから音源方向を求めて出力する検知ステップとを含み、

前記検知ステップでは、前記時間差がゼロになるように前記２つ以上のマイクロホンからの周波数スペクトルの時間軸を調整したときに、前記音源方向を検知する時間単位である時間区間に占める、前記検知音特定ステップにおいて特定された時刻部分のうち前記マイクロホン間で共通の時刻に存在する部分の割合が、所定のしきい値より大きい場合に音源方向を出力する

ことを特徴とする音源方向検知方法。

[4] 前記検知ステップでは、さらに、前記２つ以上のマイクロホンからの周波数スペクトルの時間軸を調整したときの、前記音源方向を検知する時間単位である時間区間に占める、前記検知音特定ステップにおいて特定された時刻部分のうち前記マイクロホン間で共通の時刻に存在する部分の割合が、すべての音源方向の候補について、前記所定のしきい値より大きい場合にのみ音源方向を出力する

ことを特徴とする請求項３に記載の音源方向検知方法。

[5] 互いに離間して配置された２つ以上のマイクロホンにて集音された音響信号の各々について、当該音響信号の少なくとも１つの周波数帯域における周波数スペクトルを生成する生成ステップと、

前記周波数帯域における周波数スペクトルから、音源方向を求める検知音の周波数スペクトルの時刻部分を特定する検知音特定ステップと、

音源方向を検知する時間単位である時間区間の中で、前記検知音特定ステップにおいて特定された時刻部分における周波数スペクトルの前記マイクロホン間での一致度合いにより、前記検知音が前記マイクロホンに到達する時間差を求めて、前記時間差と前記マイクロホンの距離と音速とから音源方向

を求めて出力する検知ステップとをコンピュータに実行させるためのプログラムであって、

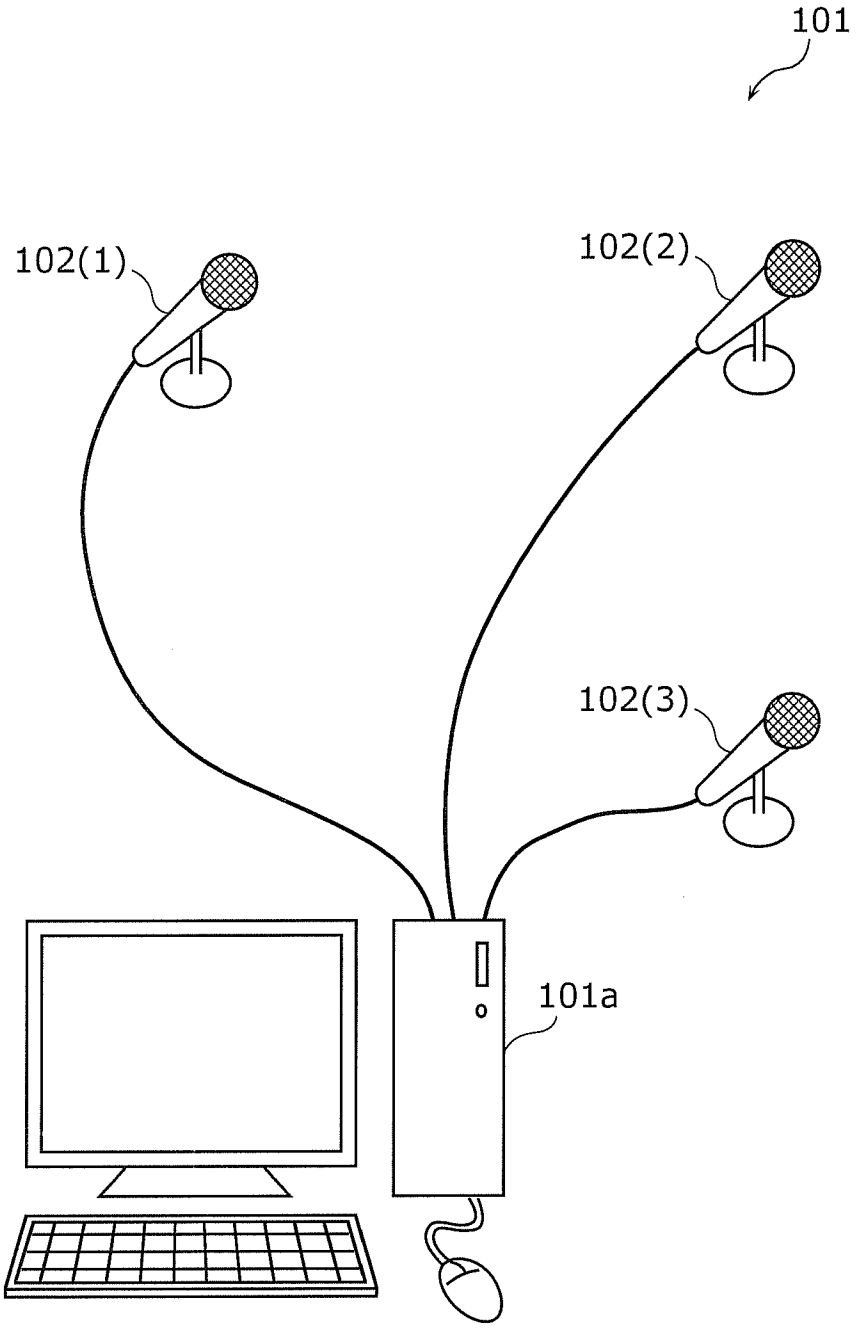
前記検知ステップでは、前記時間差がゼロになるように前記２つ以上のマイクロホンからの周波数スペクトルの時間軸を調整したときに、前記音源方向を検知する時間単位である時間区間に占める、前記検知音特定ステップにおいて特定された時刻部分のうち前記マイクロホン間で共通の時刻に存在する部分の割合が、所定のしきい値より大きい場合に音源方向を出力する

ことを特徴とするプログラム。

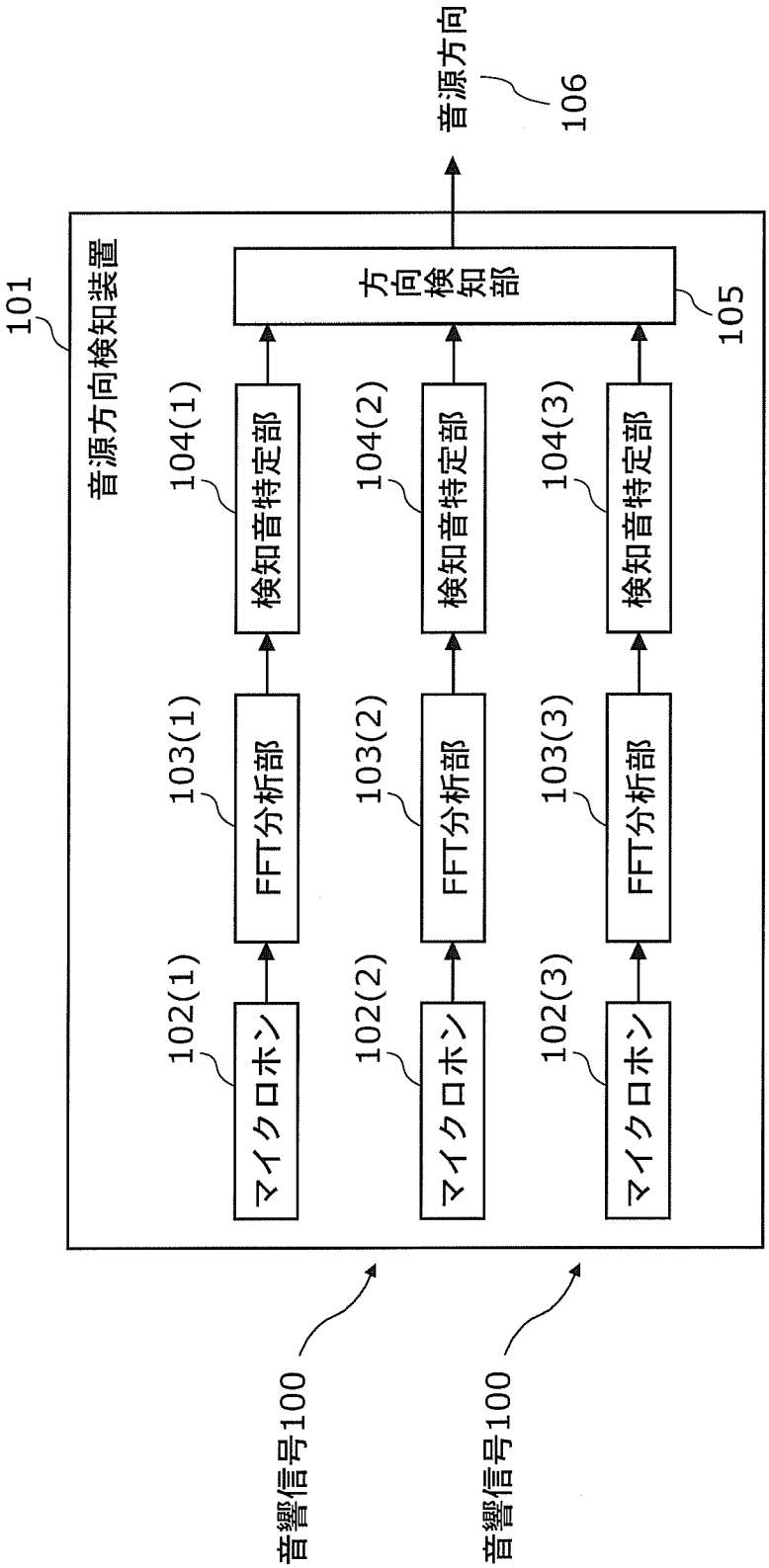
[6] 前記検知ステップでは、さらに、前記２つ以上のマイクロホンからの周波数スペクトルの時間軸を調整したときの、前記音源方向を検知する時間単位である時間区間に占める、前記検知音特定ステップにおいて特定された時刻部分のうち前記マイクロホン間で共通の時刻に存在する部分の割合が、すべての音源方向の候補について、前記所定のしきい値より大きい場合にのみ音源方向を出力する

ことを特徴とする請求項５に記載のプログラム。

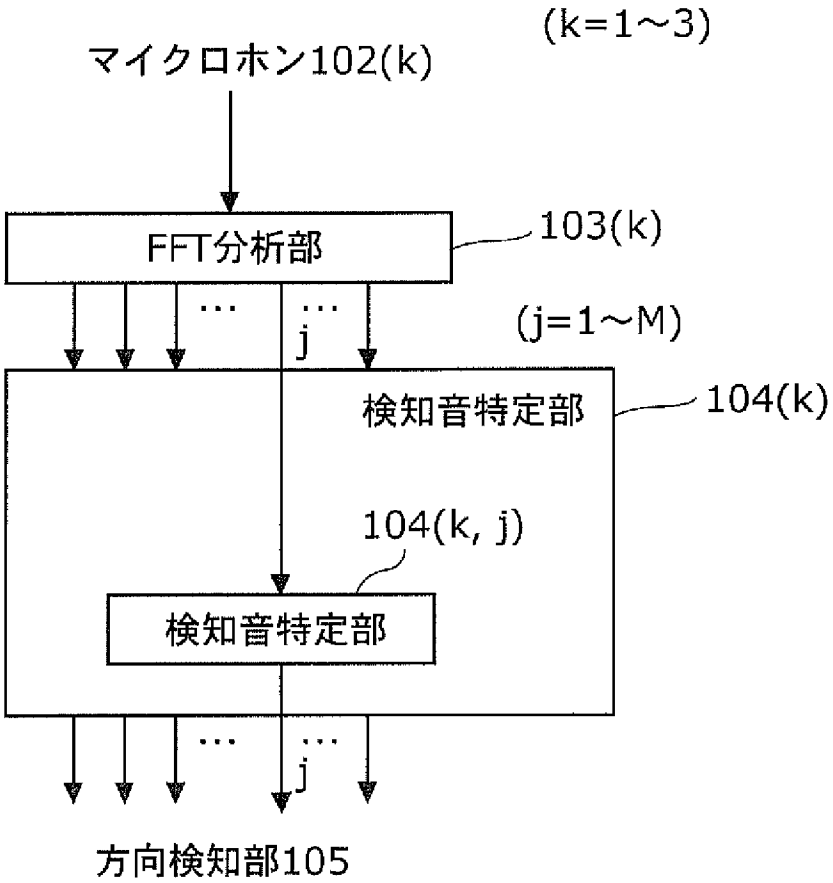
[図1]



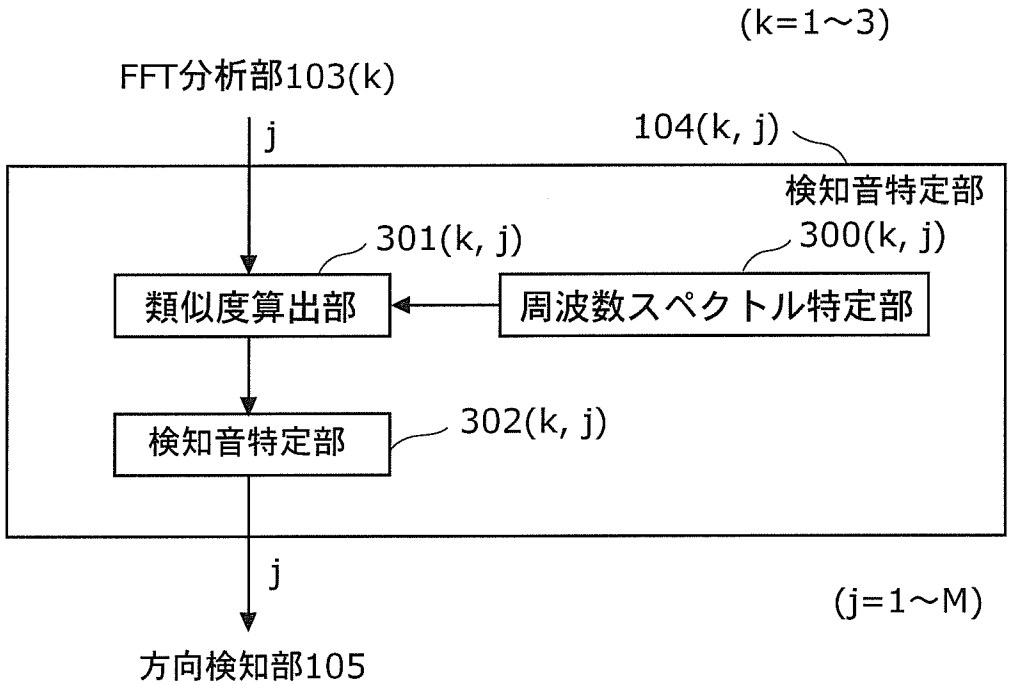
[図2]


音源方向検知装置
101
音響信号100
音響信号100
マイクロホン
102(1)
マイクロホン
102(2)
マイクロホン
102(3)
FFT分析部
103(1)
FFT分析部
103(2)
FFT分析部
103(3)
検知音特定部
104(1)
検知音特定部
104(2)
検知音特定部
104(3)
方向検知部
105
音源方向
106

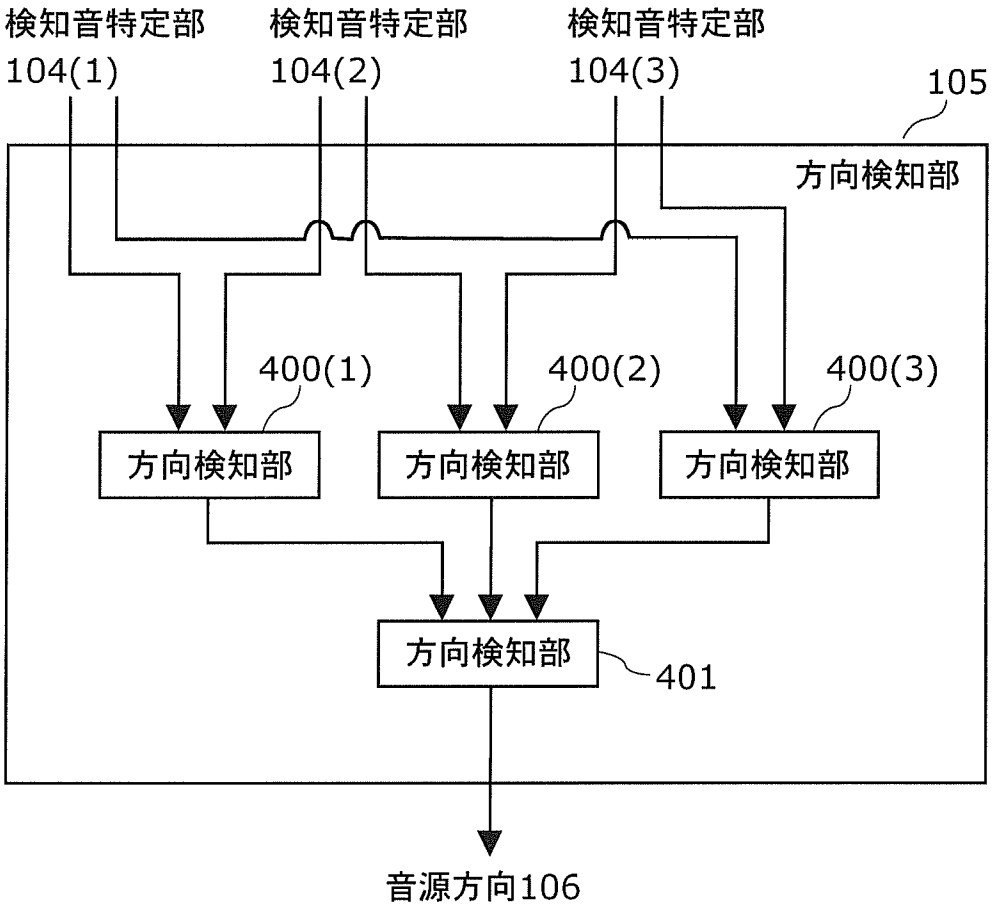
[図3]


(k=1～3)
マイクロホン102(k)
FFT分析部
103(k)
(j=1～M)
j
検知音特定部
104(k)
104(k, j)
検知音特定部
j
方向検知部105

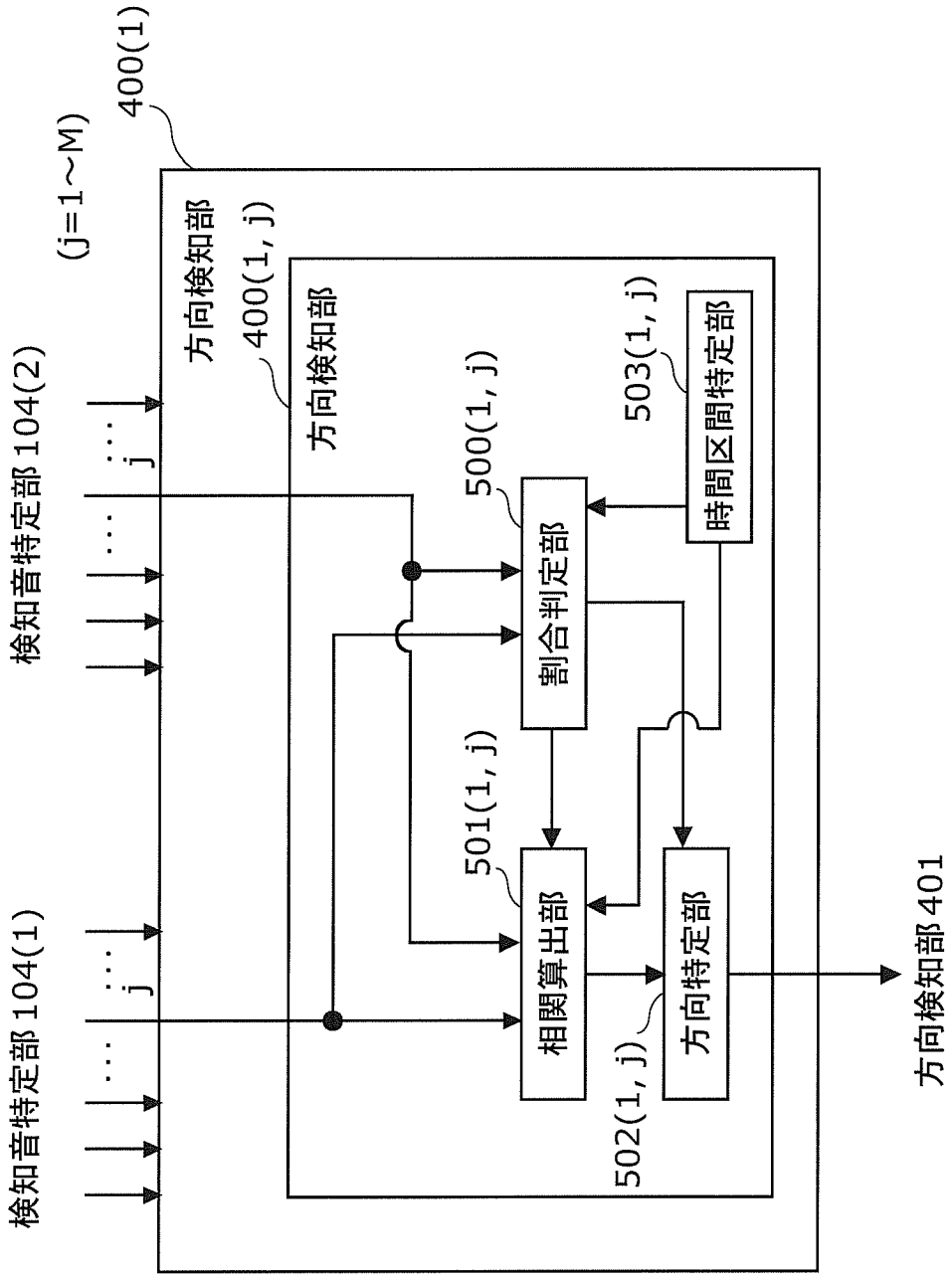
[図4]


(k=1～3)
FFT分析部103(k)
j
104(k, j)
検知音特定部
301(k, j)
300(k, j)
類似度算出部
周波数スペクトル特定部
検知音特定部
302(k, j)
j
(j=1～M)
方向検知部105

[図5]


検知音特定部
104(1)
検知音特定部
104(2)
検知音特定部
104(3)
105
方向検知部
400(1)
400(2)
400(3)
方向検知部
方向検知部
方向検知部
方向検知部
401
音源方向106

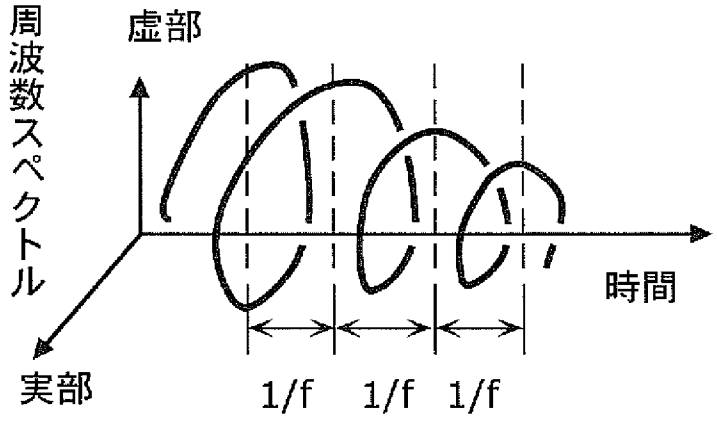
[図6]


400(1)
(j=1～M)
方向検知部
400(1, j)
方向検知部
検知音特定部104(2)
検知音特定部104(1)
j
j
500(1, j)
割合判定部
503(1, j)
時間区間特定部
501(1, j)
相関算出部
502(1, j)
方向特定部
方向検知部401

[図7A]

バイク音

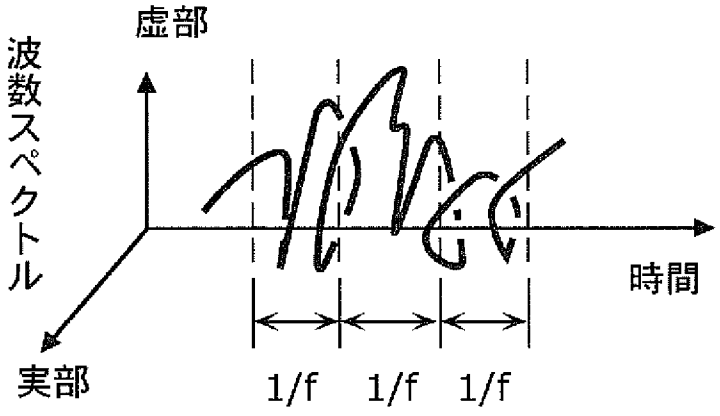
周波数： f



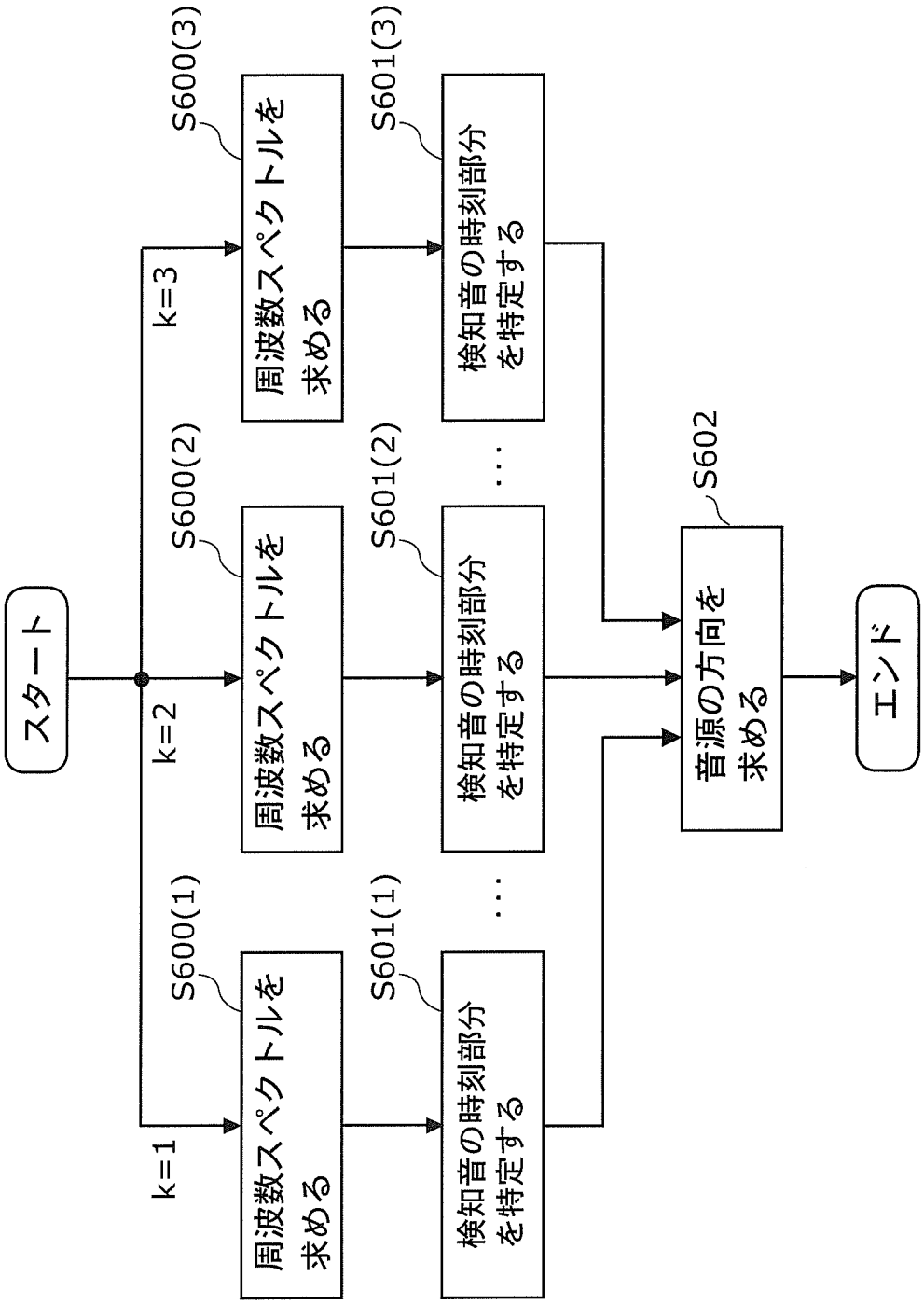
[図7B]

暗騒音

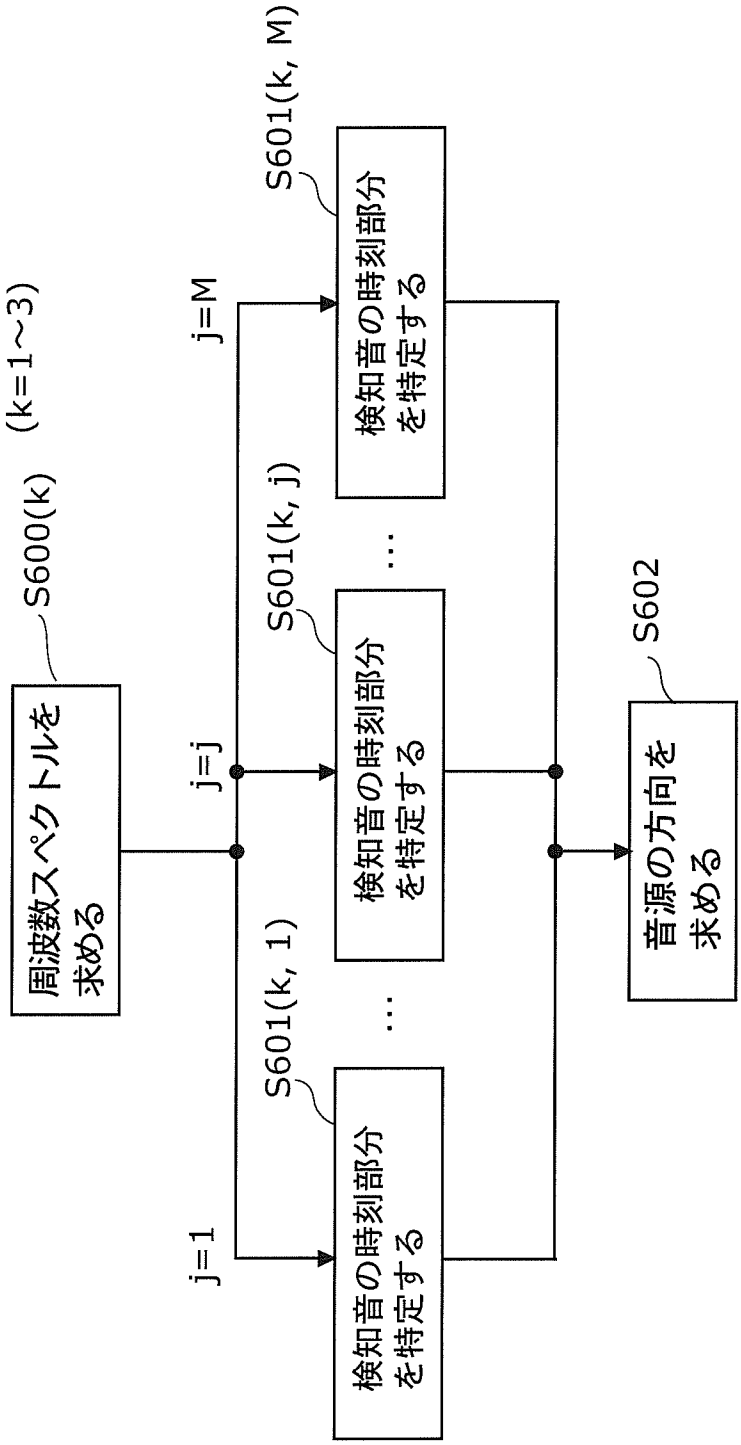
周波数： f



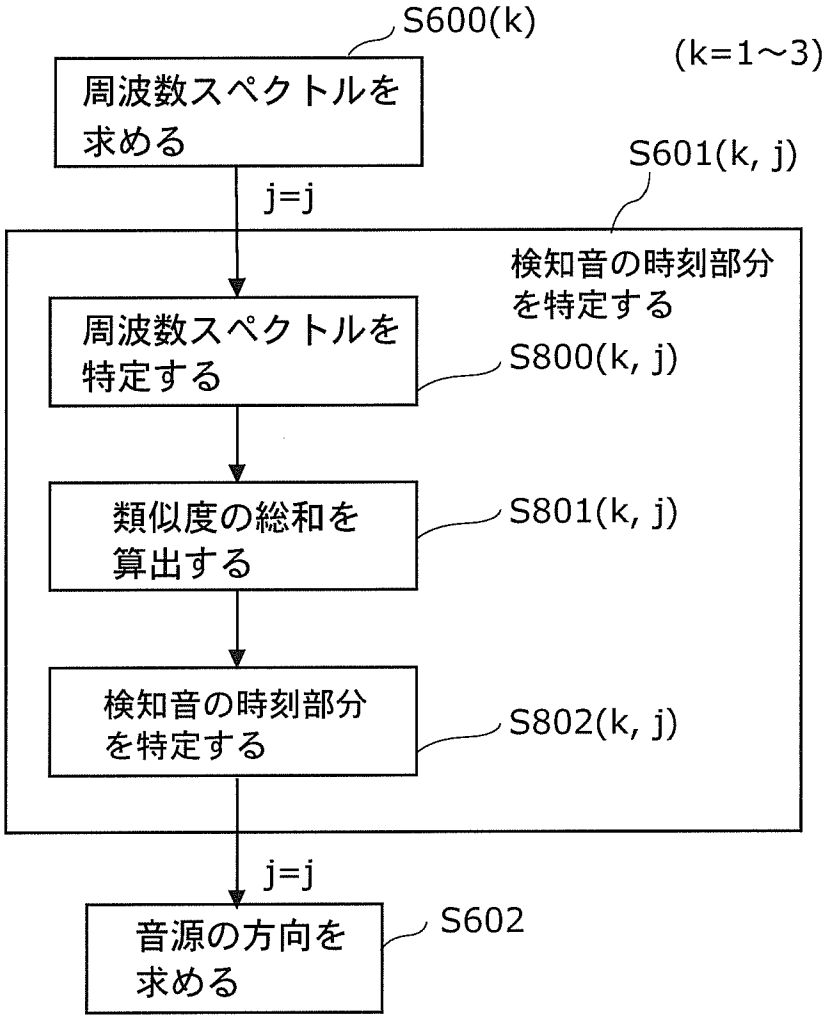
[図8]


スタート
k=1
k=2
k=3
S600(1)
S600(2)
S600(3)
周波数スペクトルを求める
周波数スペクトルを求める
周波数スペクトルを求める
S601(1)
S601(2)
S601(3)
検知音の時刻部分を特定する
検知音の時刻部分を特定する
検知音の時刻部分を特定する
S602
音源の方向を求める
エンド

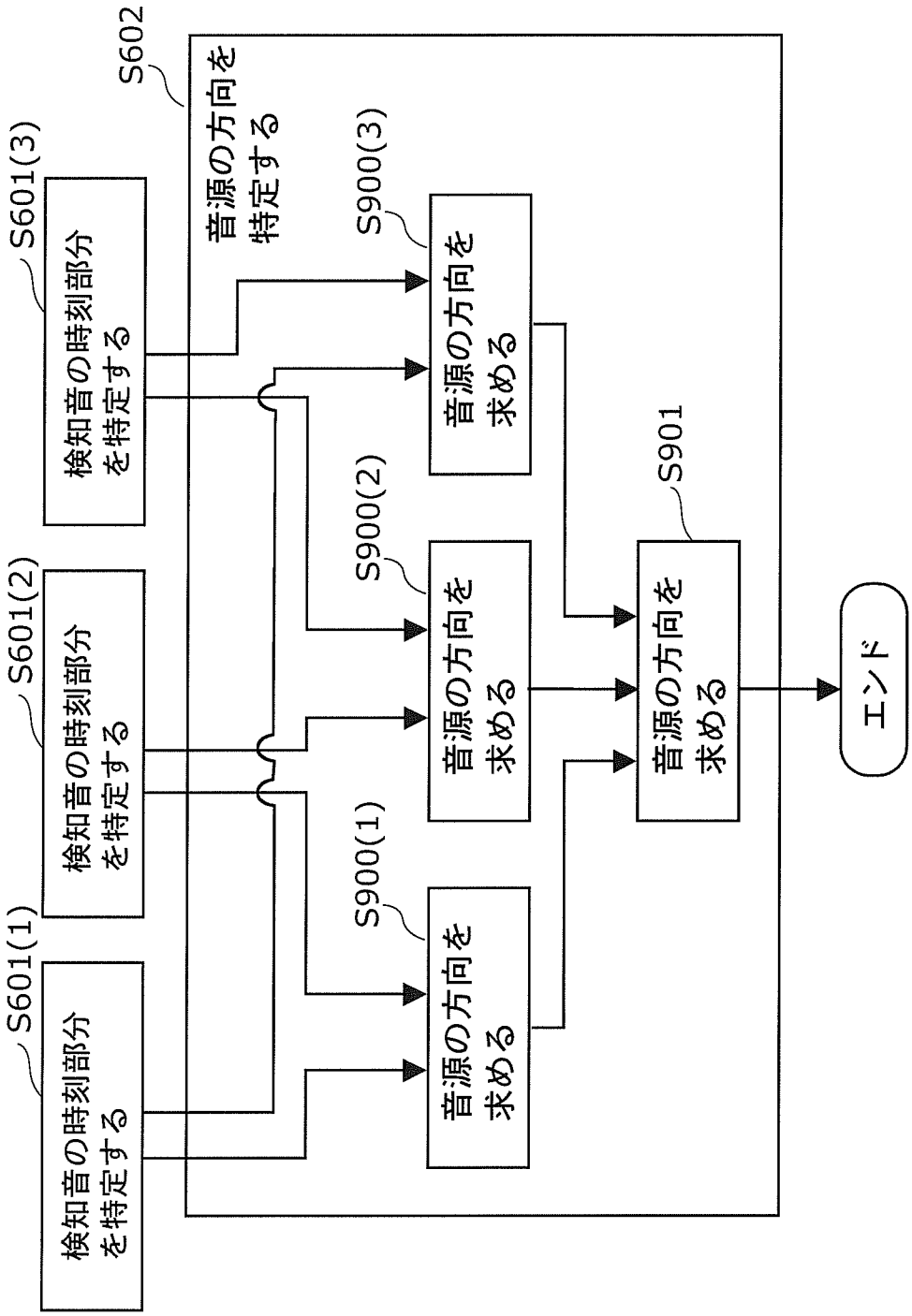
[図9]


S600(k)
(k=1～3)
周波数スペクトルを求める
j=1
j=j
j=M
S601(k, 1)
S601(k, j)
S601(k, M)
検知音の時刻部分を特定する
検知音の時刻部分を特定する
検知音の時刻部分を特定する
S602
音源の方向を求める

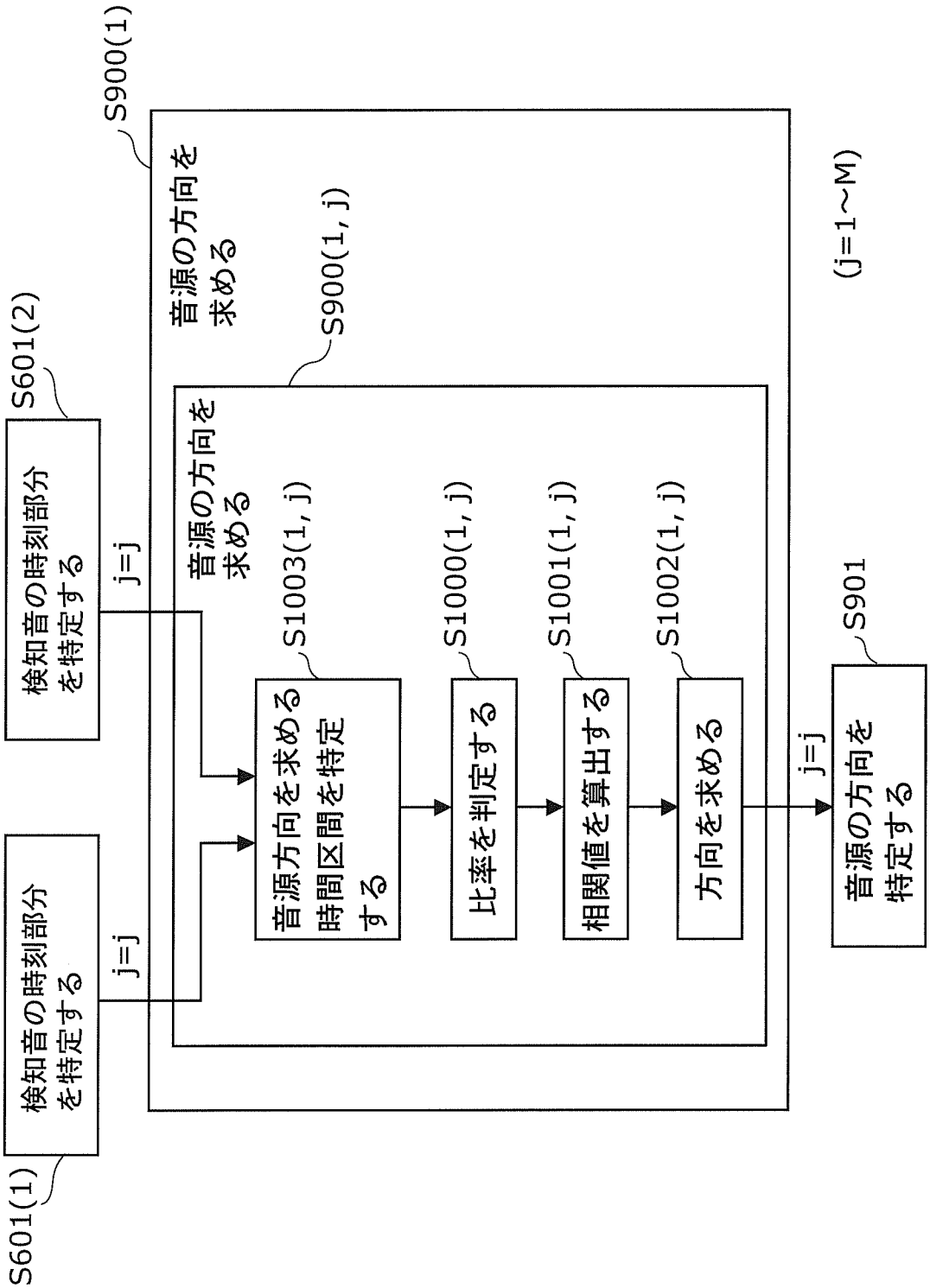
[図10]


S600(k)
(k=1～3)
周波数スペクトルを求める
S601(k, j)
j=j
検知音の時刻部分を特定する
周波数スペクトルを特定する
S800(k, j)
類似度の総和を算出する
S801(k, j)
検知音の時刻部分を特定する
S802(k, j)
j=j
音源の方向を求める
S602

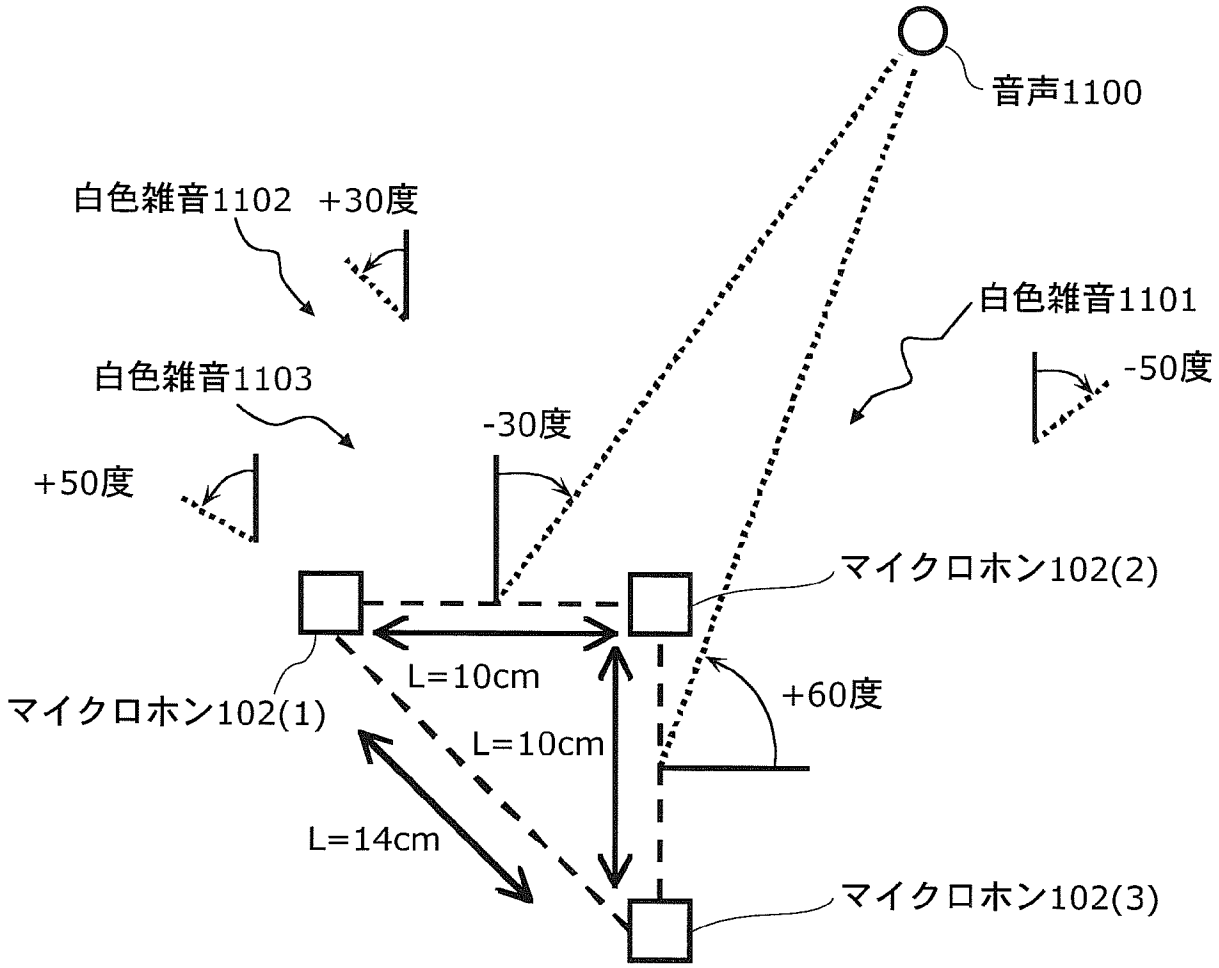
[図11]


S602
音源の方向を特定する
S601(3)
検知音の時刻部分を特定する
S601(2)
検知音の時刻部分を特定する
S601(1)
検知音の時刻部分を特定する
S900(3)
音源の方向を求める
S900(2)
音源の方向を求める
S900(1)
音源の方向を求める
S901
音源の方向を求める
エンド

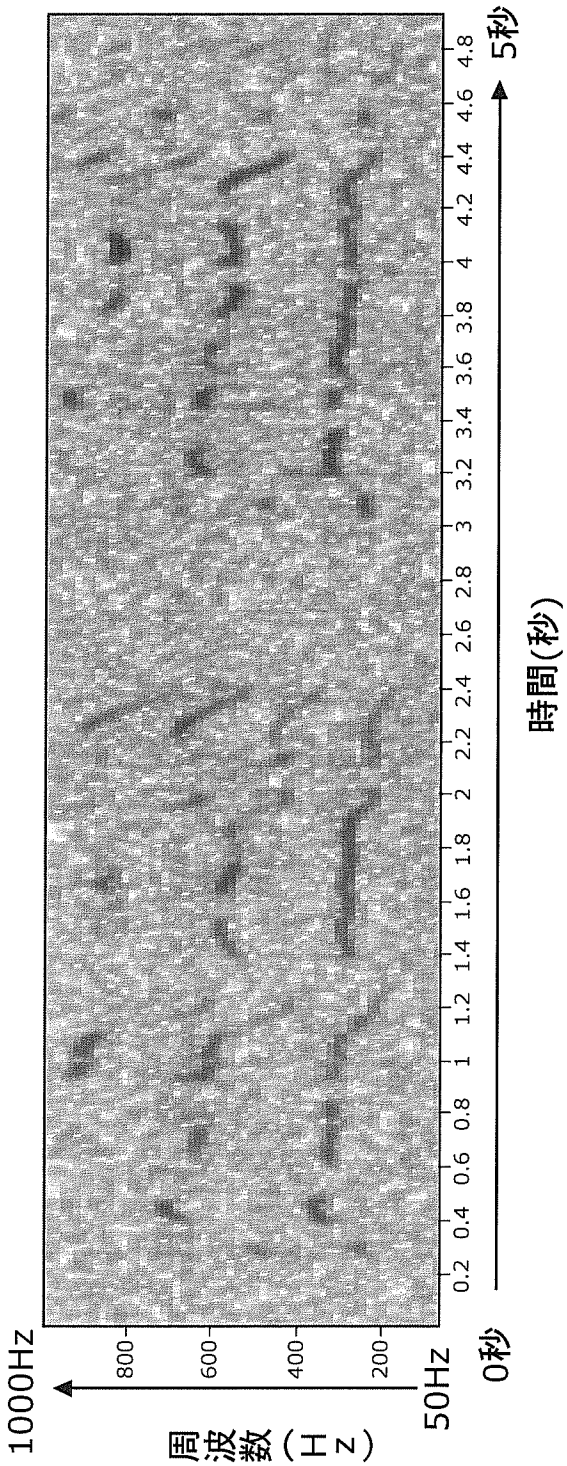
[図12]


S601(1)
検知音の時刻部分を特定する
S601(2)
検知音の時刻部分を特定する
j=j
j=j
S900(1)
音源の方向を求める
S900(1, j)
音源の方向を求める
S1003(1, j)
音源方向を求める時間区間を特定する
S1000(1, j)
比率を判定する
S1001(1, j)
相関値を算出する
S1002(1, j)
方向を求める
j=j
S901
音源の方向を特定する
(j=1～M)

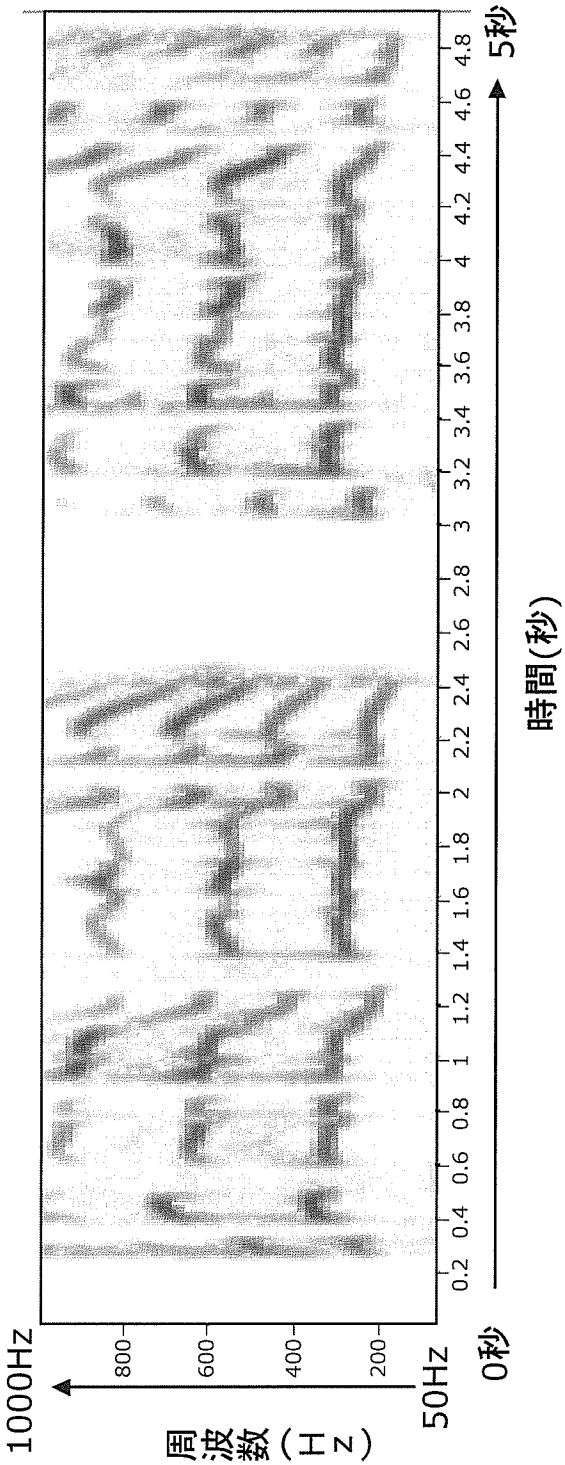
[図13]


音声1100
白色雑音1102
+30度
白色雑音1101
-50度
白色雑音1103
-30度
+50度
マイクロホン102(2)
L=10cm
+60度
マイクロホン102(1)
L=10cm
L=14cm
マイクロホン102(3)

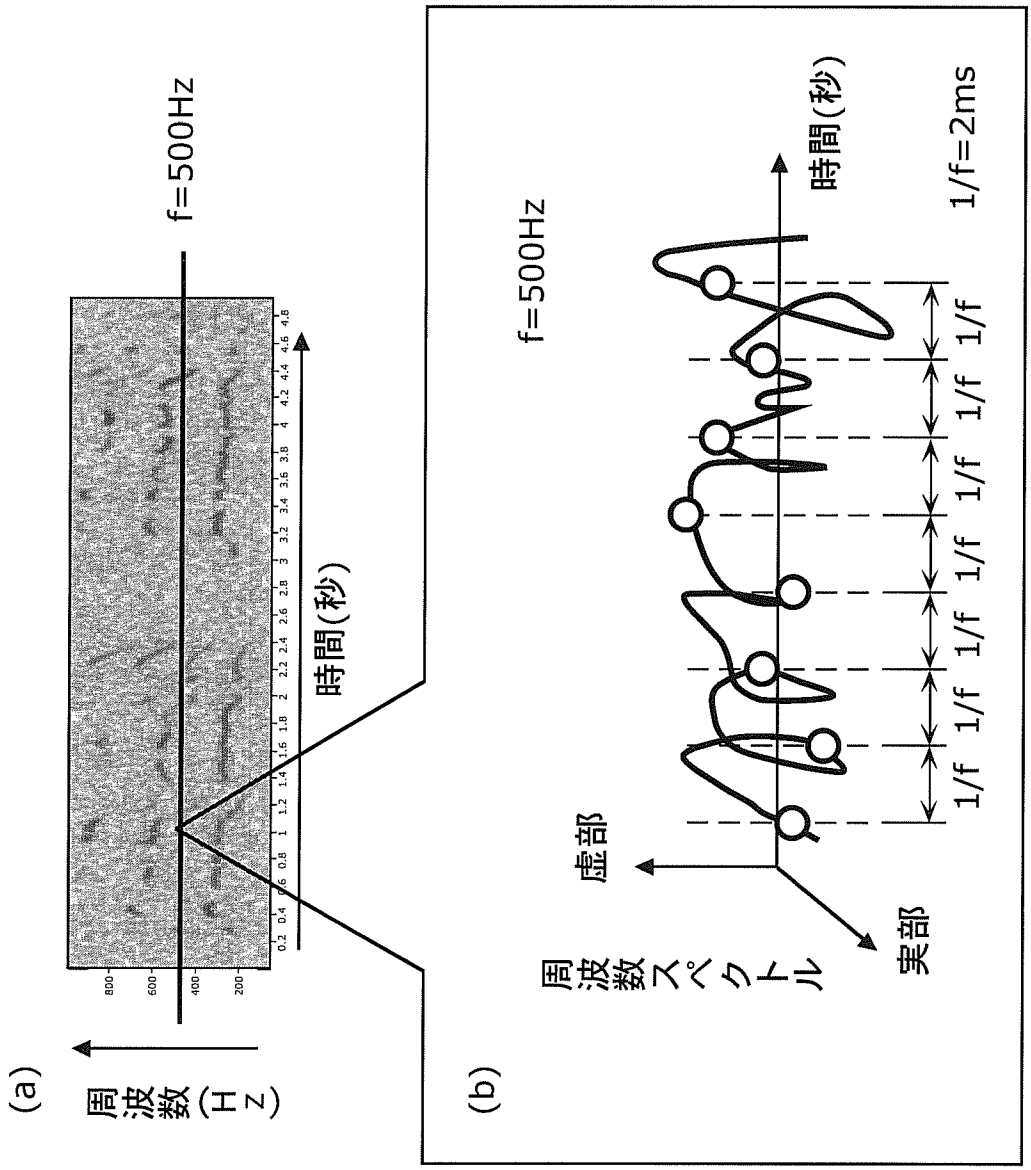
[図14]


1000Hz
800
600
400
200
50Hz
周波数(Hz)
0秒
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
2.2
2.4
2.6
2.8
3
3.2
3.4
3.6
3.8
4
4.2
4.4
4.6
4.8
5秒
時間(秒)

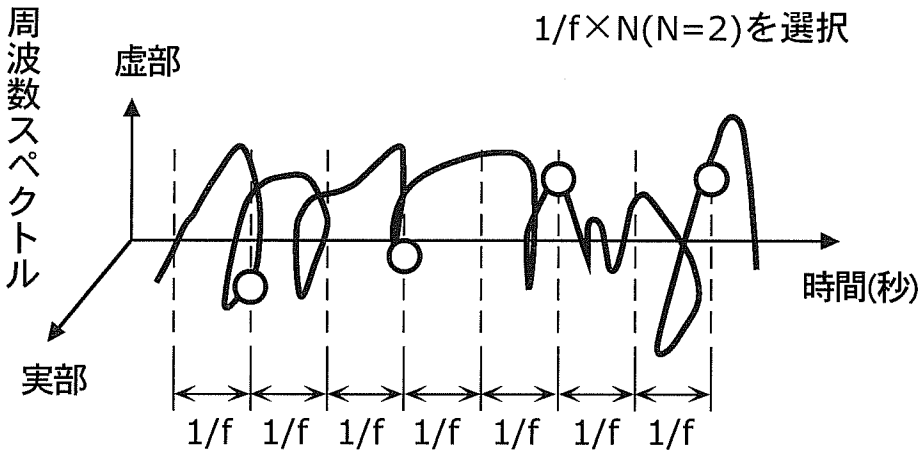
[図15]


1000Hz
800
600
400
200
50Hz
周波数(Hz)
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
2.2
2.4
2.6
2.8
3
3.2
3.4
3.6
3.8
4
4.2
4.4
4.6
4.8
0秒
5秒
時間(秒)

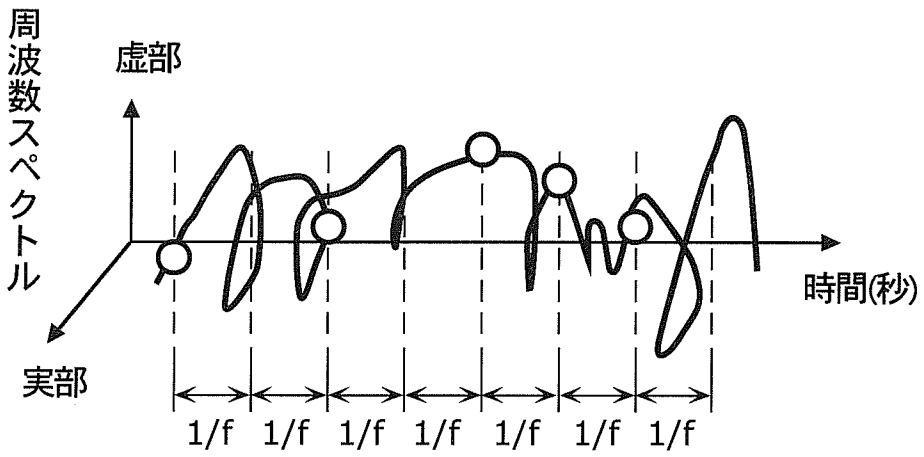
[図16]


(a)
周波数(HZ)
800
600
400
200
f=500Hz
時間(秒)
0.2 0.4 0.6 0.8 1 1.2 1.4 1.6 1.8 2 2.2 2.4 2.6 2.8 3 3.2 3.4 3.6 3.8 4 4.2 4.4 4.6 4.8
(b)
周波数スペクトル
虚部
実部
f=500Hz
時間(秒)
1/f=2ms
1/f 1/f 1/f 1/f 1/f 1/f 1/f 1/f

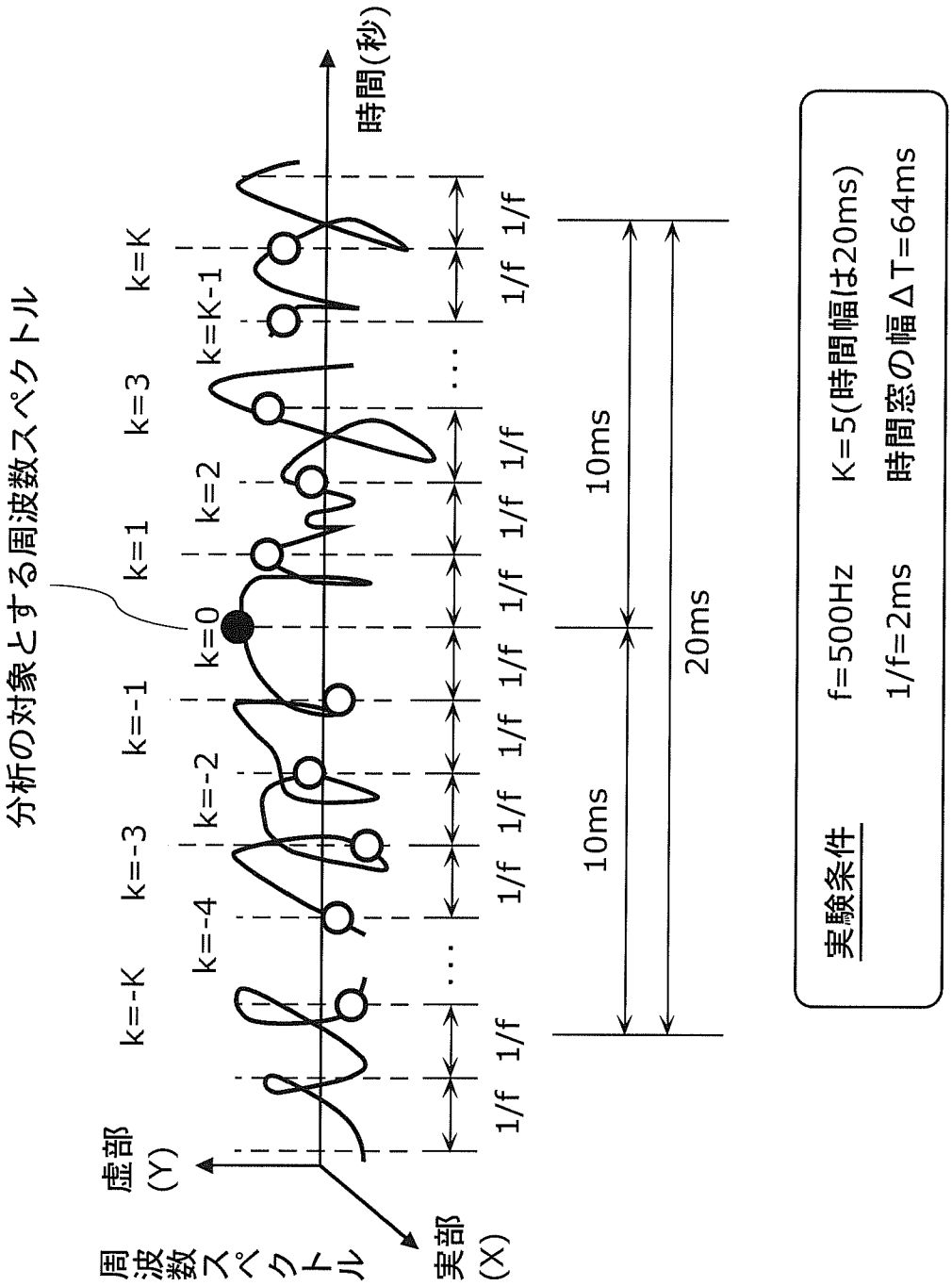
[図17A]


周波数スペクトル
虚部
実部
1/f×N(N=2)を選択
時間(秒)
1/f 1/f 1/f 1/f 1/f 1/f 1/f

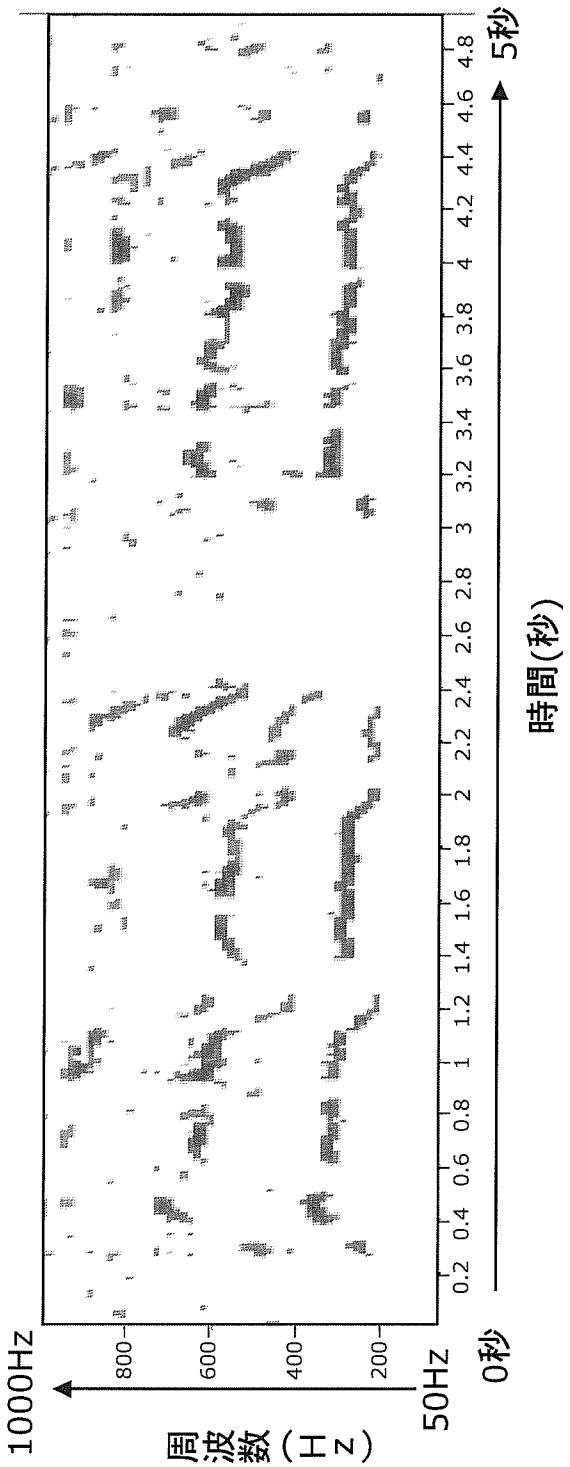
[図17B]



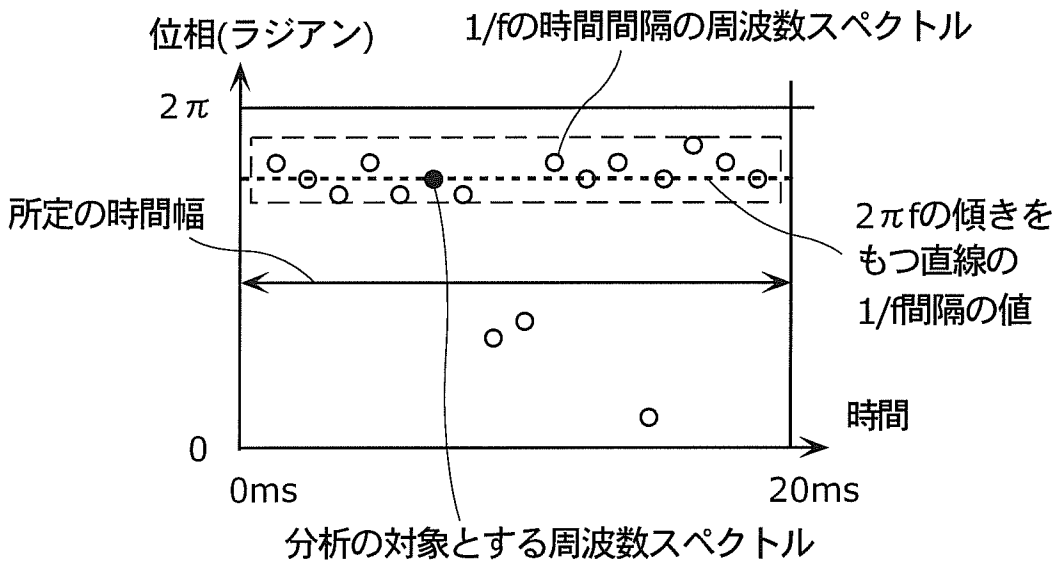
[図18]


分析の対象とする周波数スペクトル
時間(秒)
k=-K
k=-4
k=-3
k=-2
k=-1
k=0
k=1
k=2
k=3
k=K-1
k=K
1/f
10ms
20ms
周波数スペクトル
虚部(Y)
実部(X)
実験条件
f=500Hz
1/f=2ms
K=5(時間幅は20ms)
時間窓の幅ΔT=64ms

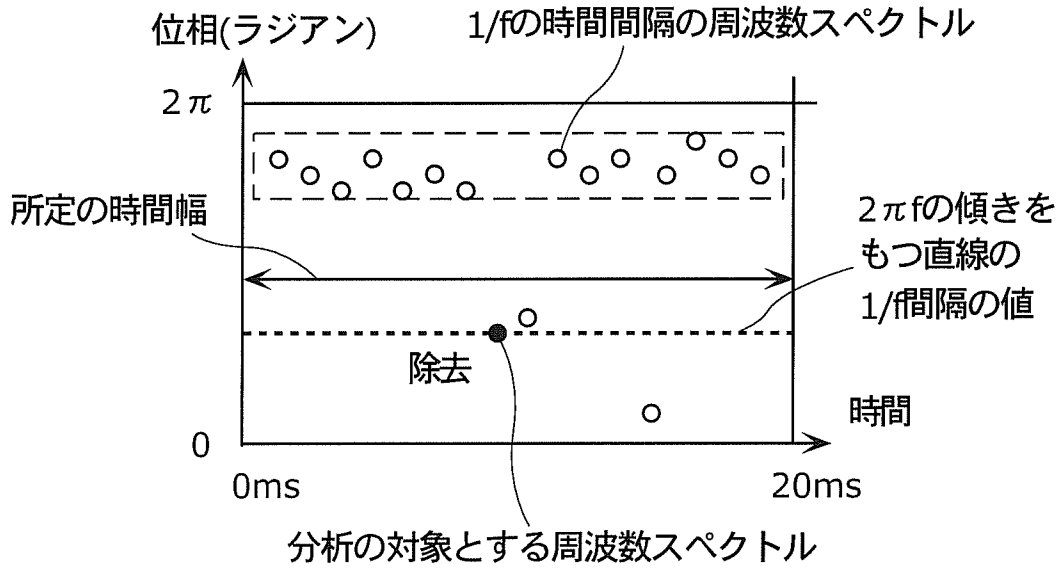
[図19]


1000Hz
800
600
400
200
50Hz
周波数(Hz)
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
2.2
2.4
2.6
2.8
3
3.2
3.4
3.6
3.8
4
4.2
4.4
4.6
4.8
0秒
5秒
時間(秒)

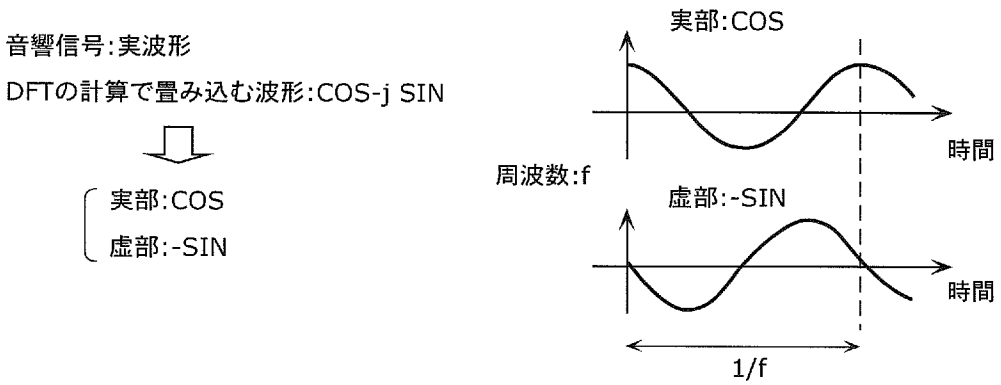
[図20A]



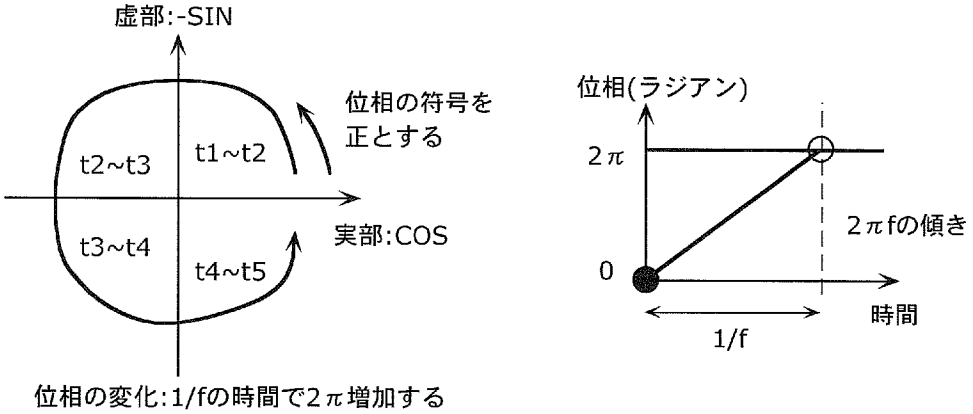
[図20B]



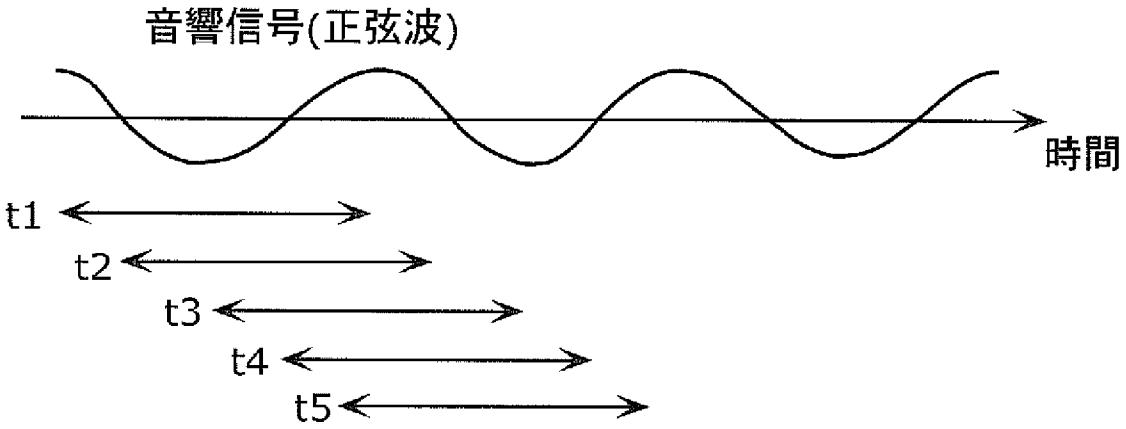
[図21A]



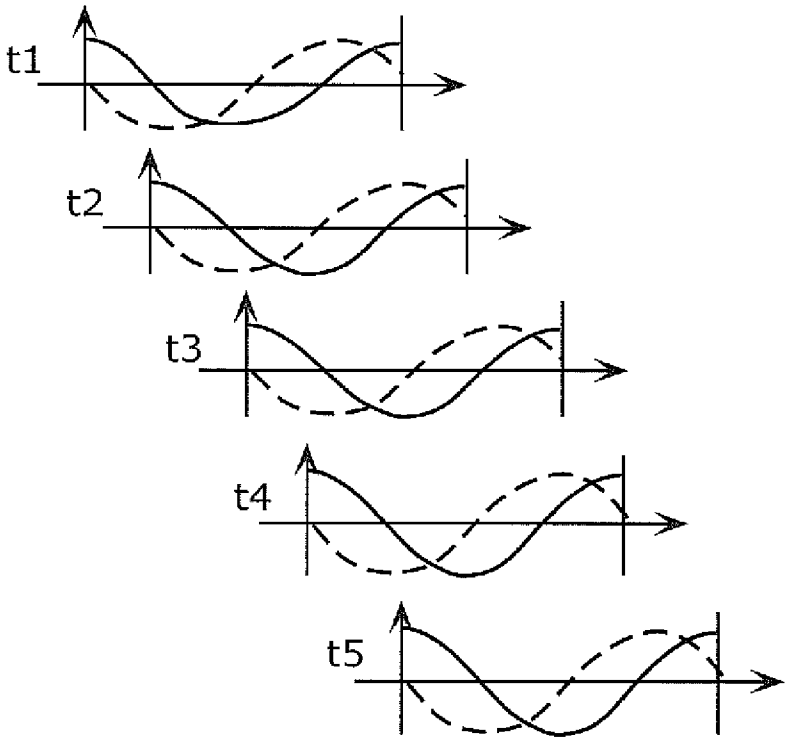
[図21B]



[図22A]



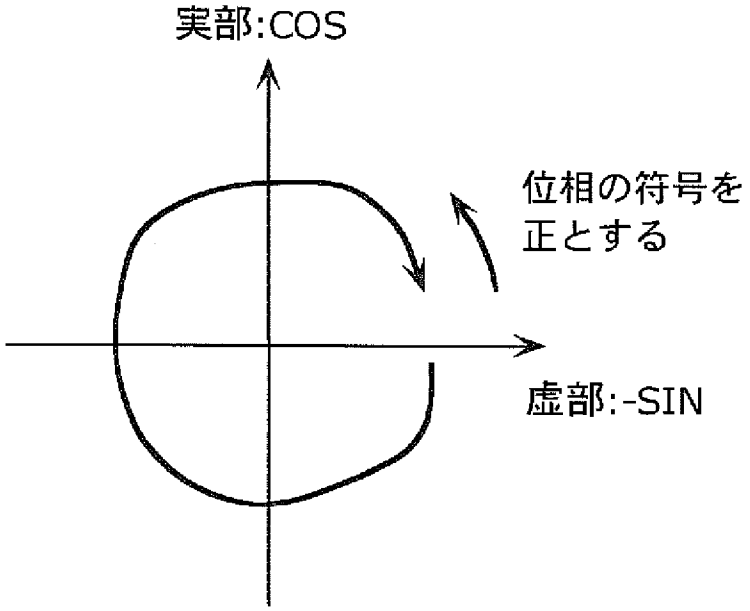
[図22B]



[図22C]

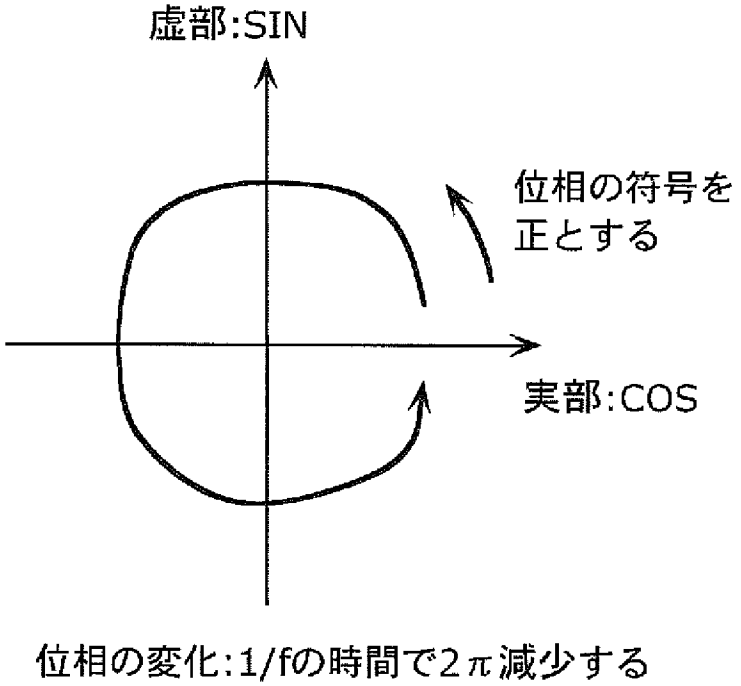
時間	実部:COS (実線)	虚部:-SIN (破線)
t1	+1	±0
t1~t2	+	+
t2	±0	+1
t2~t3	-	+
t3	-1	±0
t3~t4	-	-
t4	±0	-1
t4~t5	+	-
t5	+1	±0

[図23A]

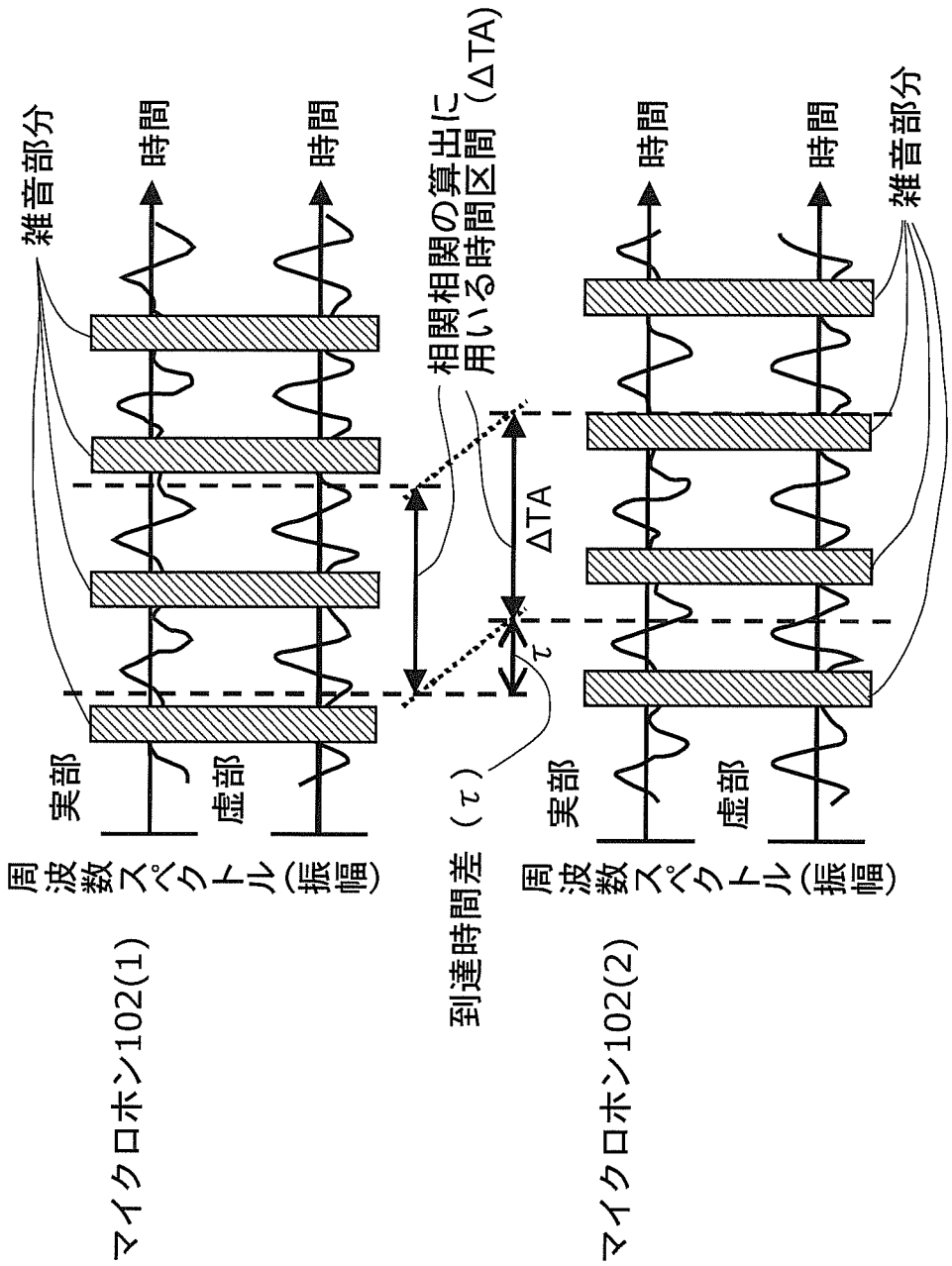


位相の変化:1/fの時間で2π減少する

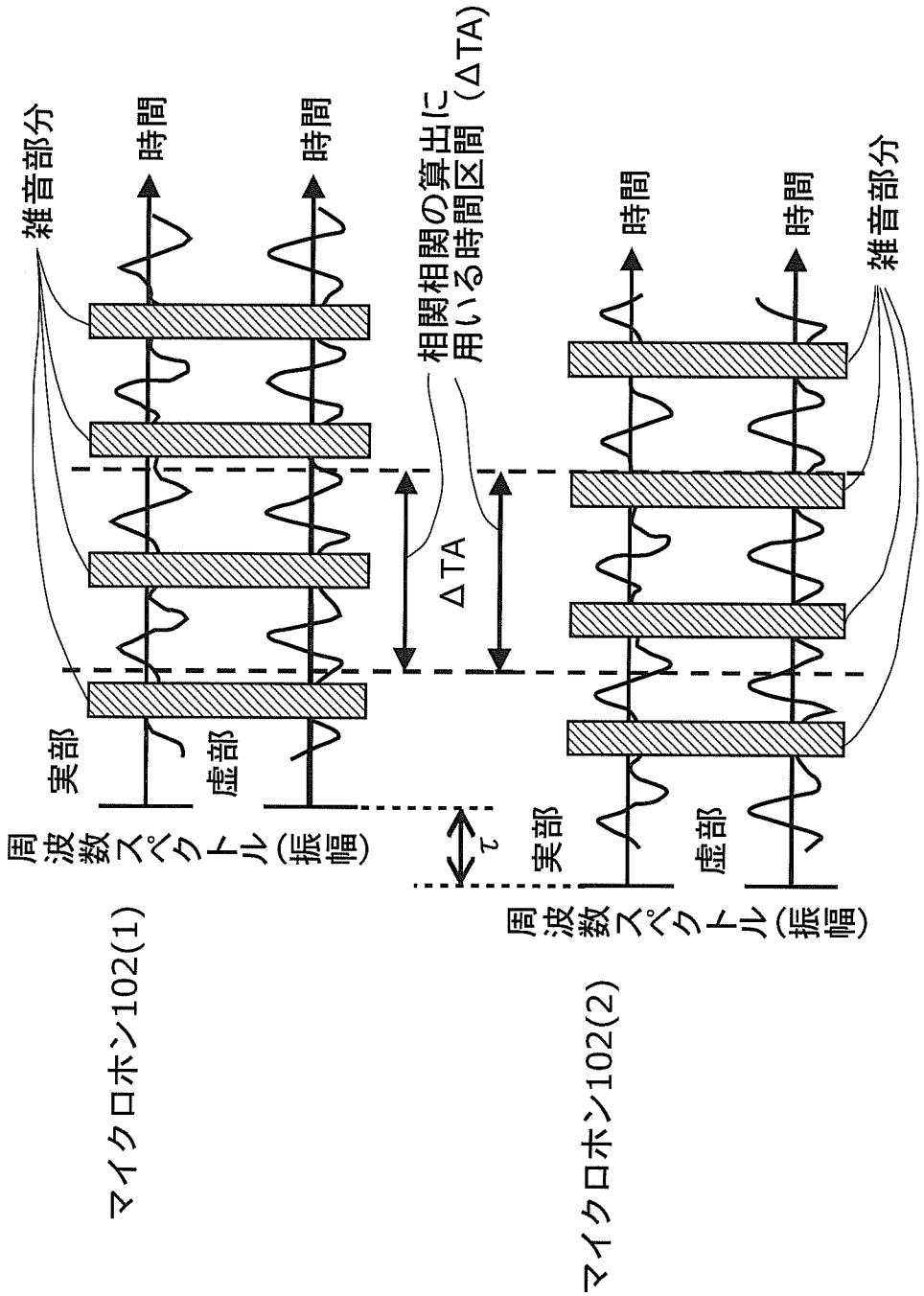
[図23B]



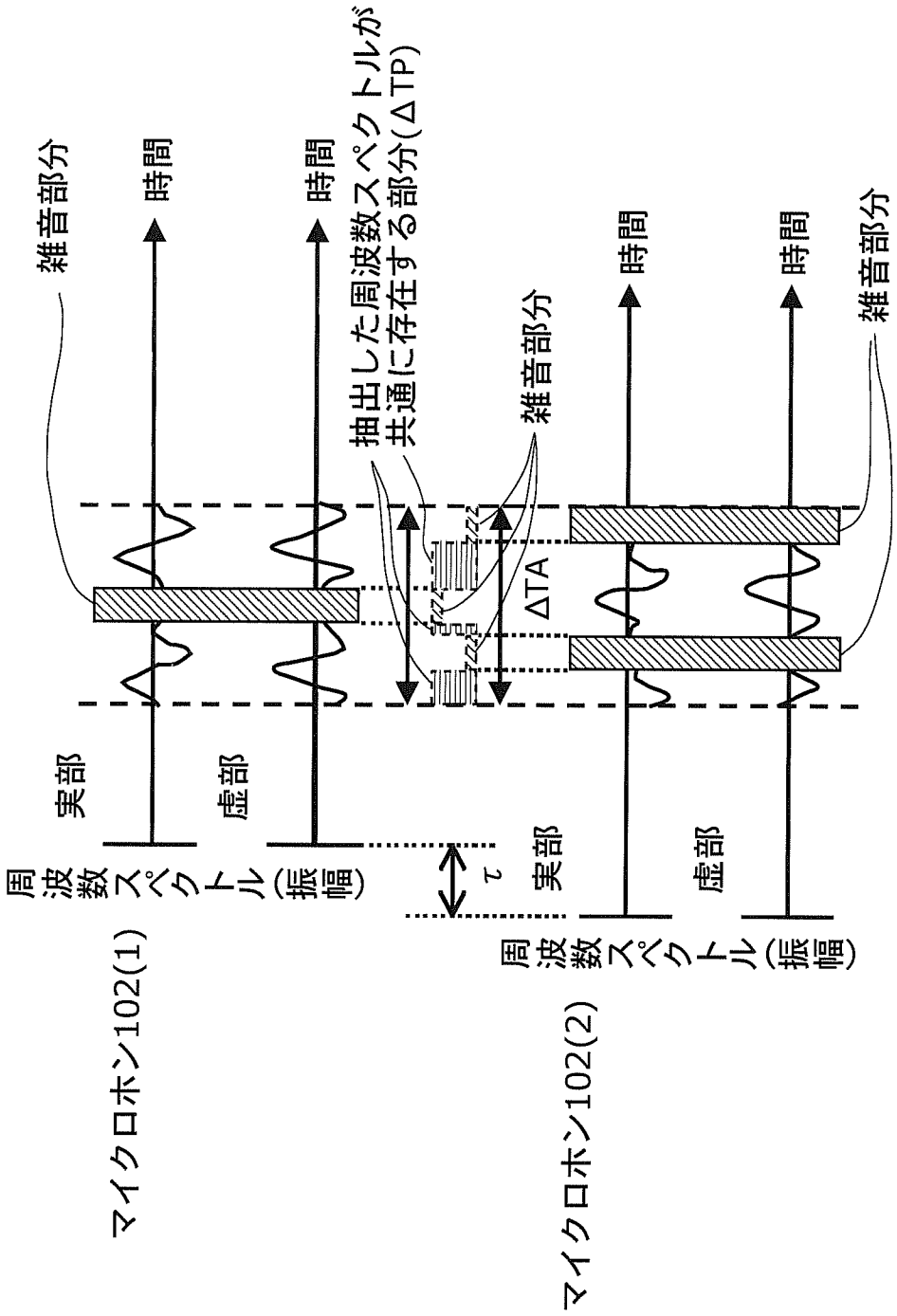
[図24]


マイクロホン102(1)
周波数スペクトル(振幅)
実部
虚部
雑音部分
時間
時間
到達時間差（τ）
τ
相関相関の算出に用いる時間区間（ΔTA）
ΔTA
マイクロホン102(2)
周波数スペクトル(振幅)
実部
虚部
雑音部分
時間
時間

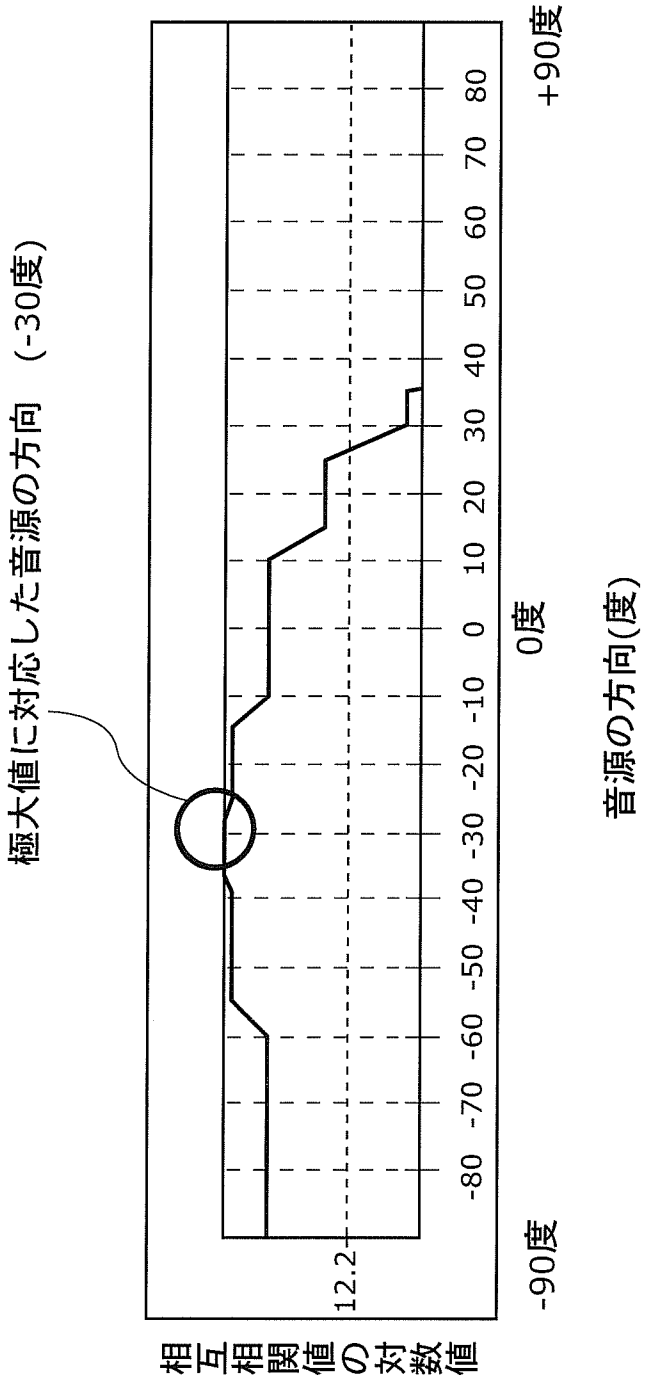
[図25]


雑音部分
時間
時間
相関相関の算出に用いる時間区間（ΔTA）
ΔTA
実部
虚部
周波数スペクトル（振幅）
τ
マイクロホン102(1)
時間
時間
雑音部分
実部
虚部
周波数スペクトル（振幅）
マイクロホン102(2)

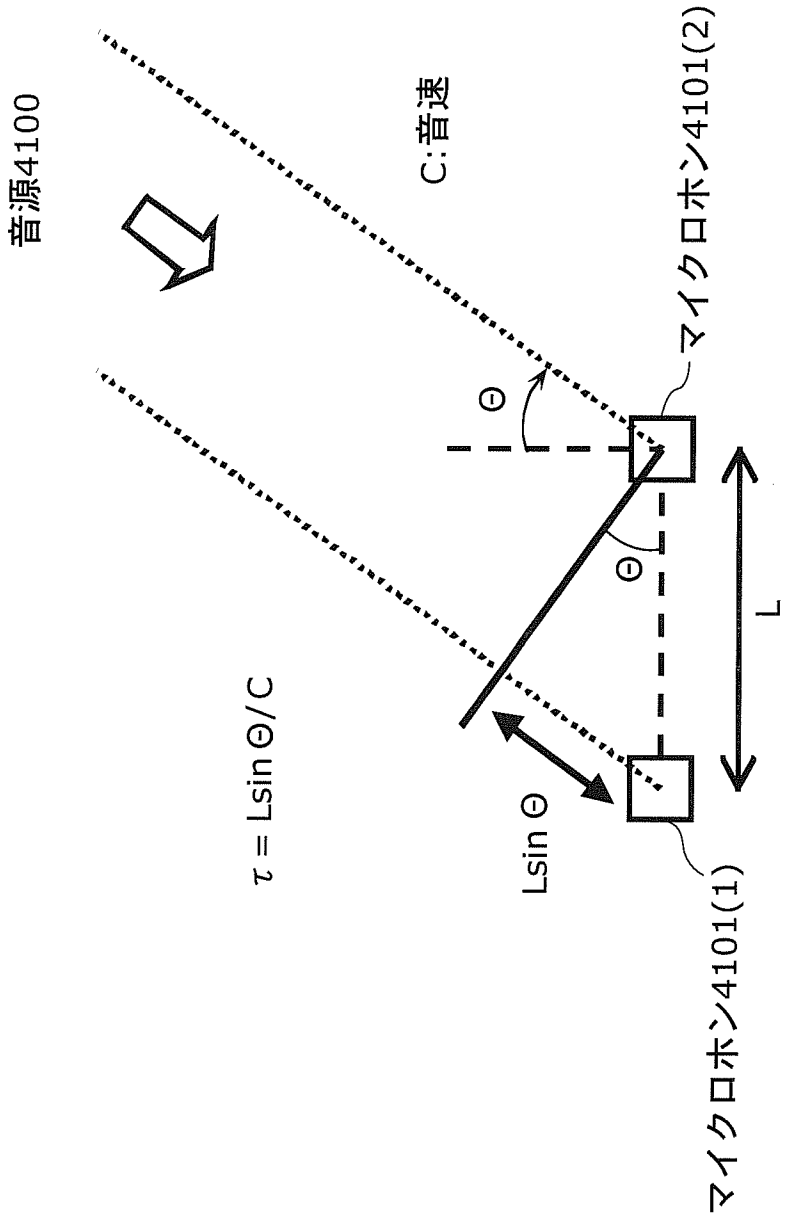
[図26]


マイクロホン102(1)
周波数スペクトル(振幅)
実部
虚部
τ
時間
雑音部分
時間
抽出した周波数スペクトルが共通に存在する部分(ΔTP)
雑音部分
ΔTA
マイクロホン102(2)
周波数スペクトル(振幅)
実部
虚部
時間
雑音部分
時間

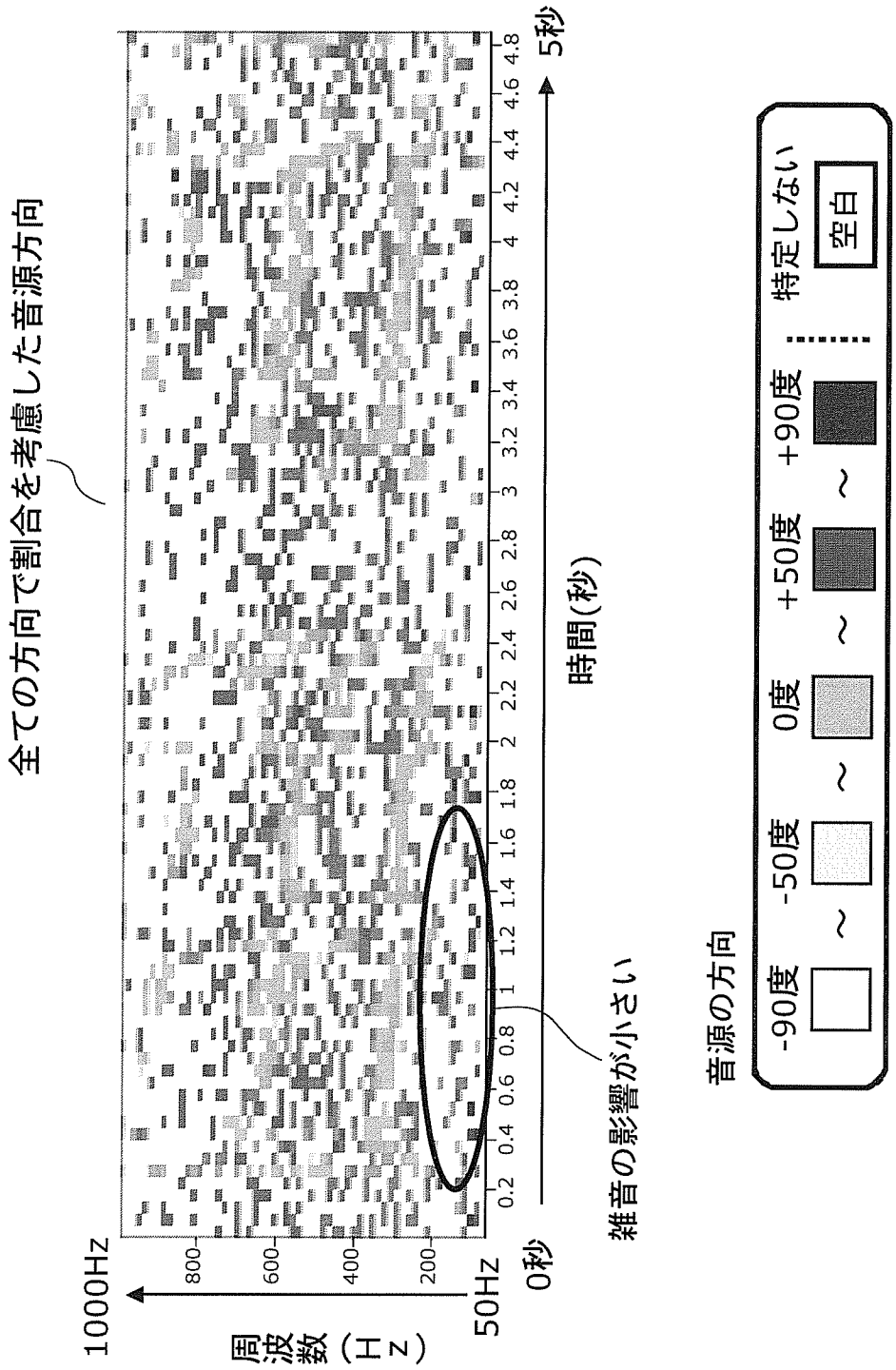
[図27]


極大値に対応した音源の方向 (-30度)
相互相関値の対数値
12.2
-90度
-80 -70 -60 -50 -40 -30 -20 -10 0 10 20 30 40 50 60 70 80
0度
+90度
音源の方向(度)

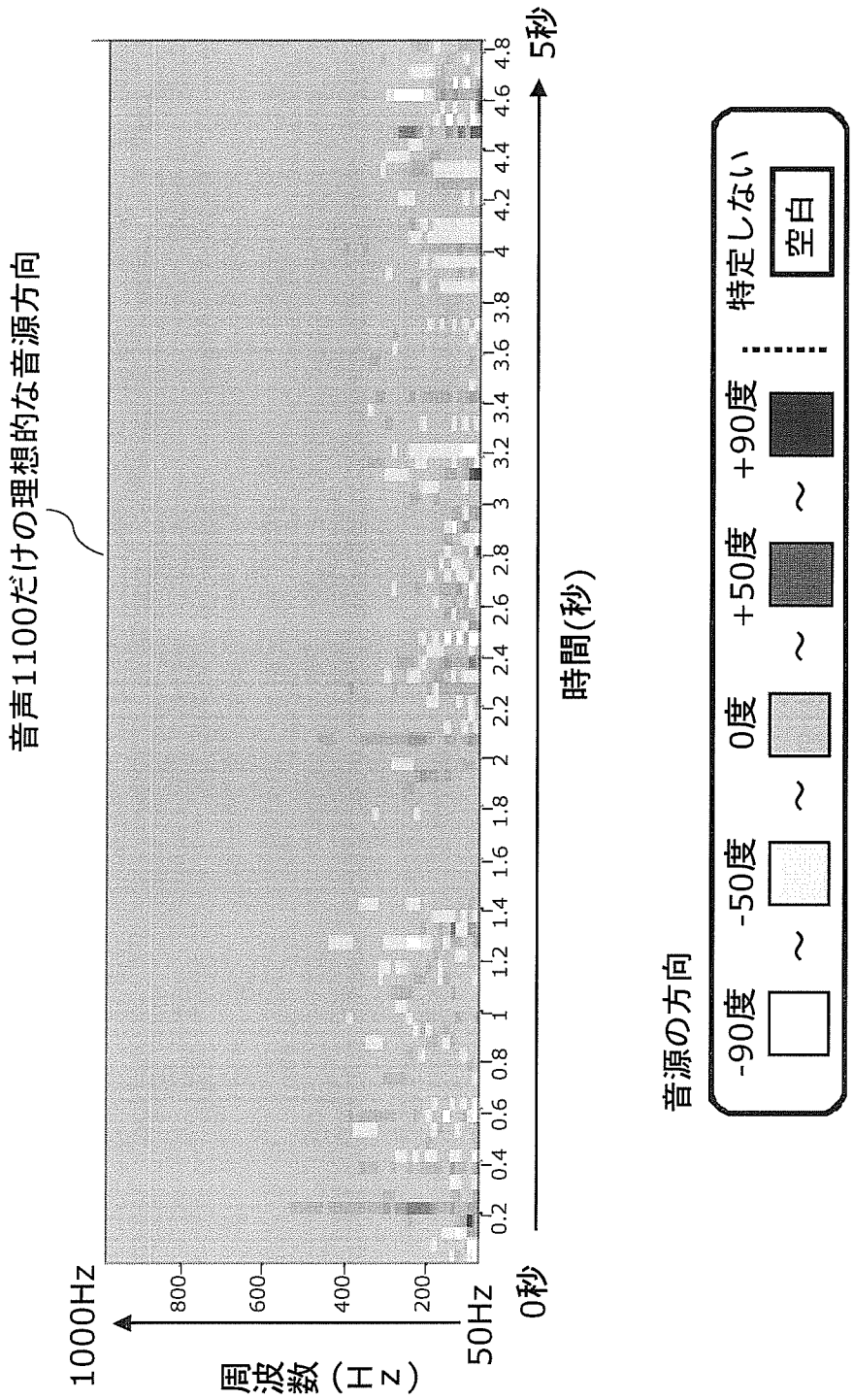
[図28]



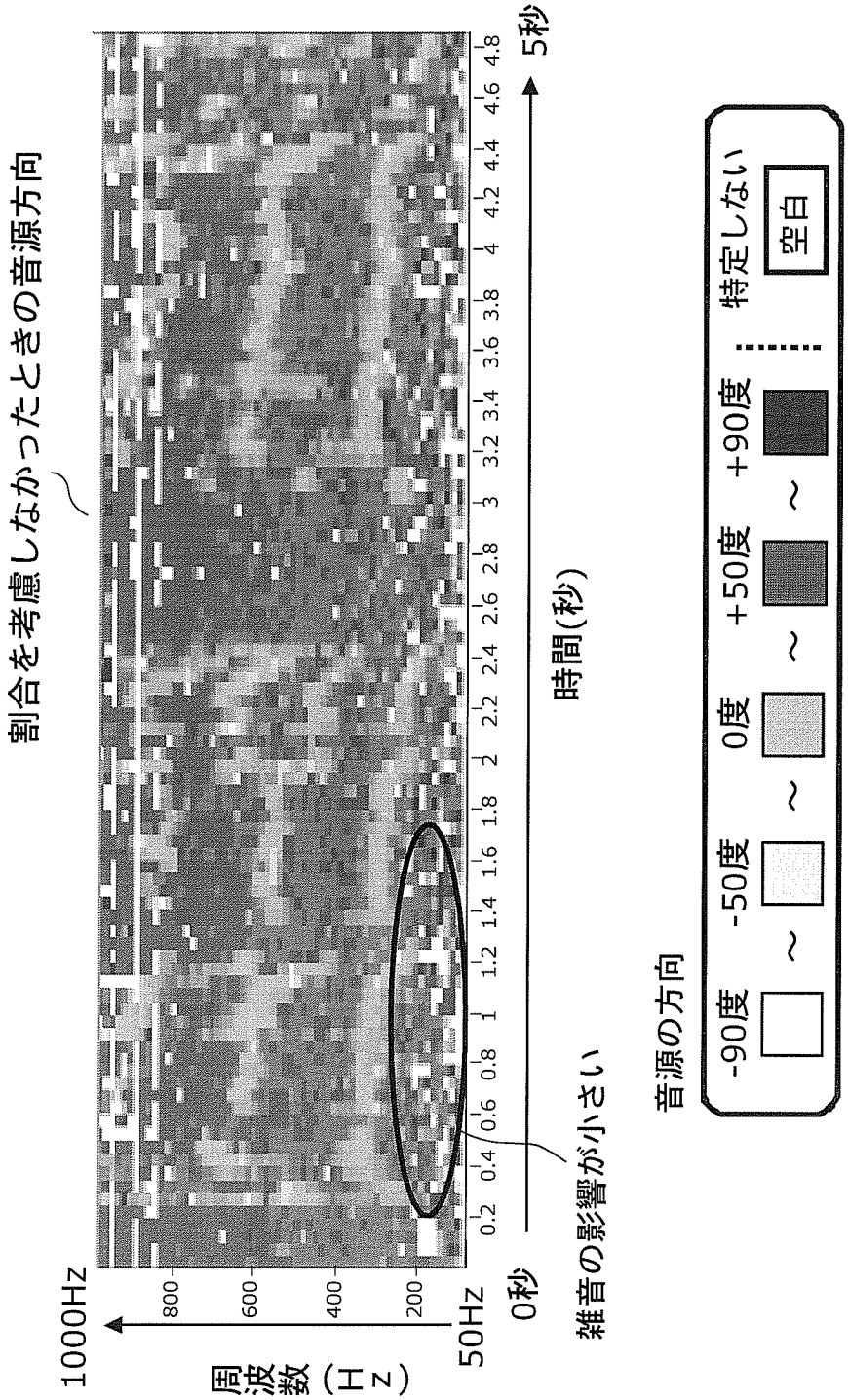
[図29]


全ての方向で割合を考慮した音源方向
1000Hz
800
600
400
200
50Hz
周波数（Hz）
0秒
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
2.2
2.4
2.6
2.8
3
3.2
3.4
3.6
3.8
4
4.2
4.4
4.6
4.8
5秒
時間(秒)
雑音の影響が小さい
音源の方向
-90度
~
-50度
~
0度
~
+50度
~
+90度
特定しない
空白

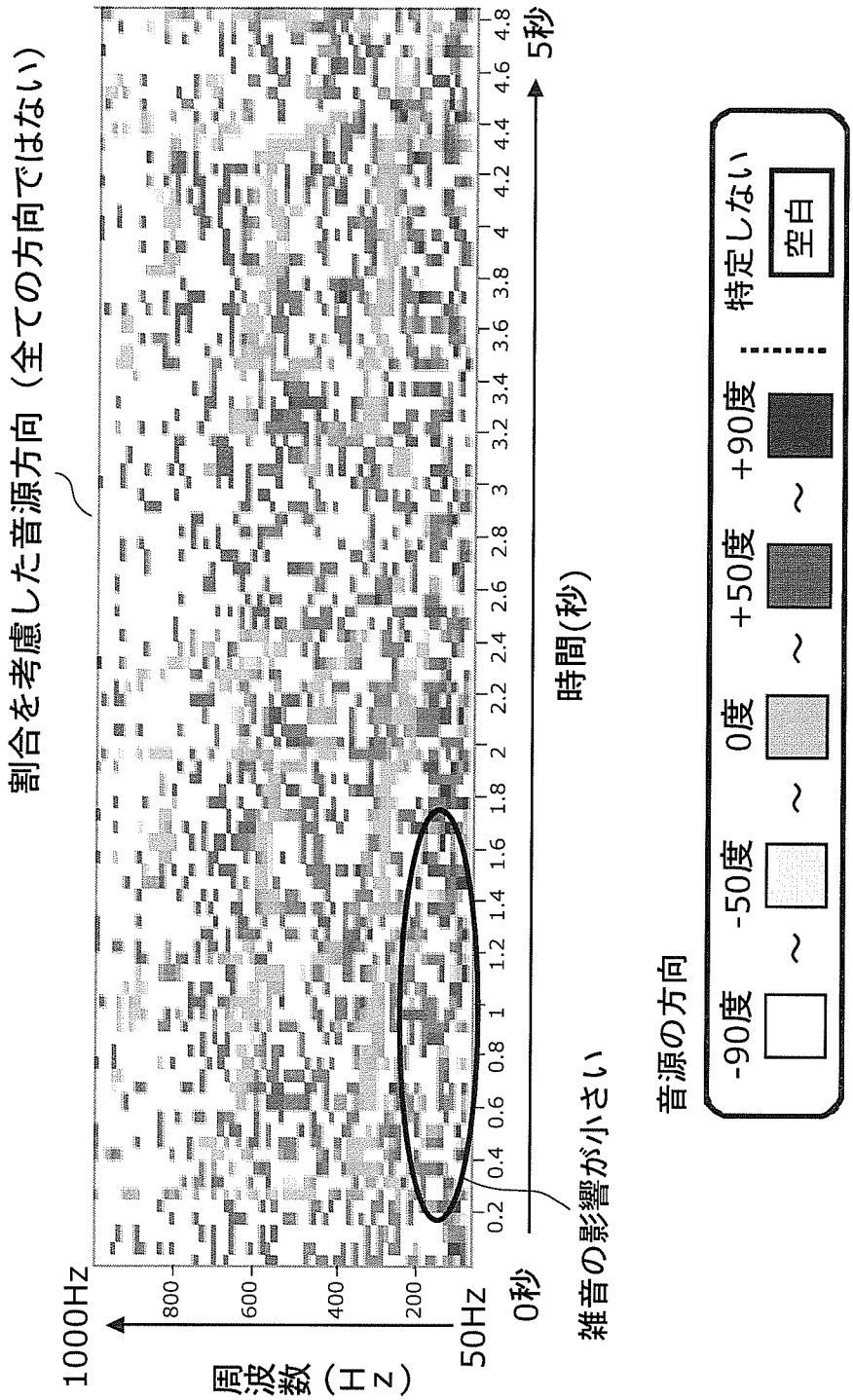
[図30]


音声1100だけの理想的な音源方向
1000Hz
800
600
400
200
50Hz
周波数（Hz）
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
2.2
2.4
2.6
2.8
3
3.2
3.4
3.6
3.8
4
4.2
4.4
4.6
4.8
0秒
5秒
時間(秒)
音源の方向
-90度
~
-50度
~
0度
~
+50度
~
+90度
特定しない
空白

[図31]

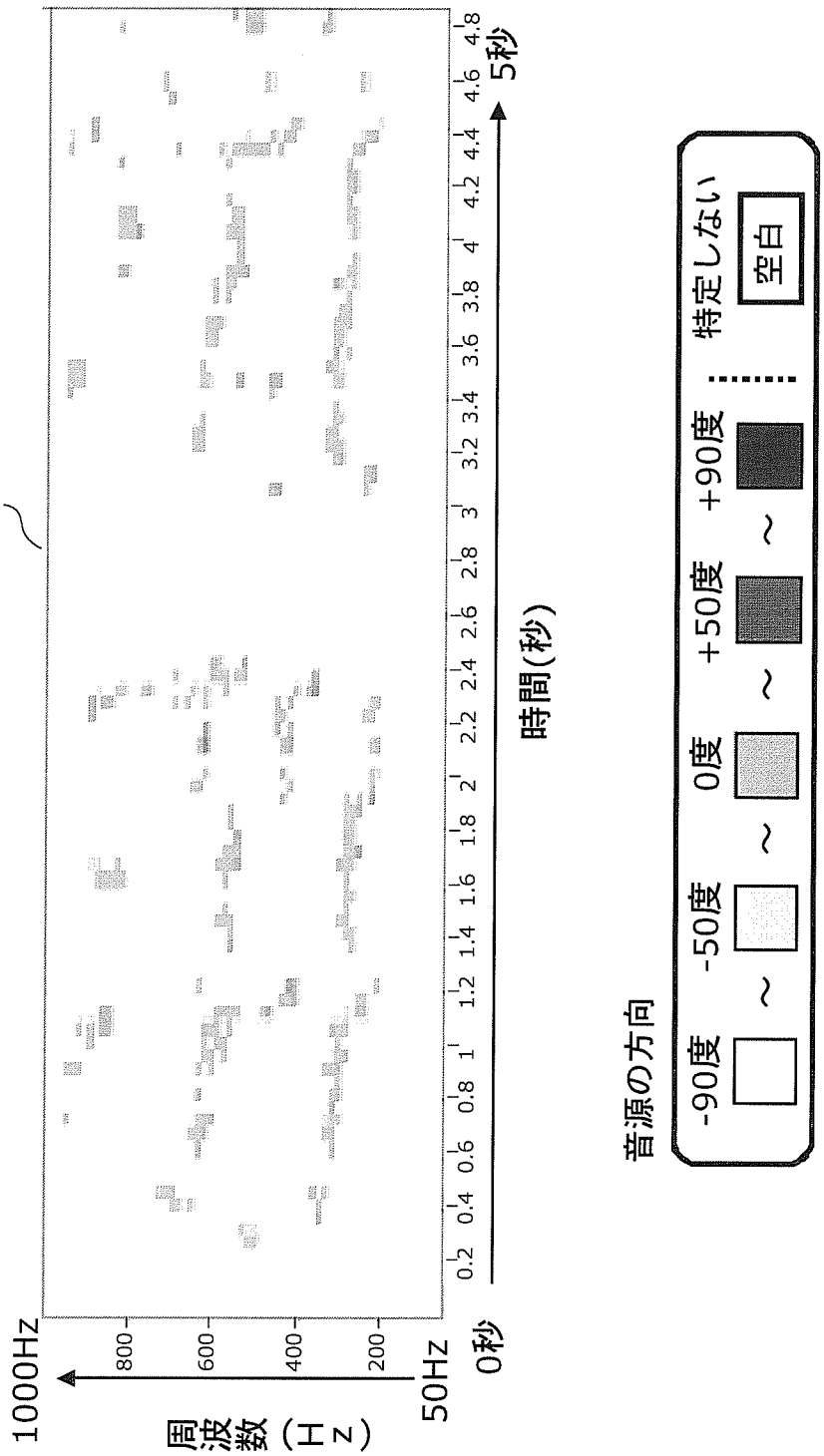

割合を考慮しなかったときの音源方向
1000Hz
周波数（Hz）
50Hz
800
600
400
200
0秒
5秒
時間(秒)
0.2 0.4 0.6 0.8 1 1.2 1.4 1.6 1.8 2 2.2 2.4 2.6 2.8 3 3.2 3.4 3.6 3.8 4 4.2 4.4 4.6 4.8
雑音の影響が小さい
音源の方向
-90度
~
-50度
~
0度
~
+50度
~
+90度
特定しない
空白

[図32]

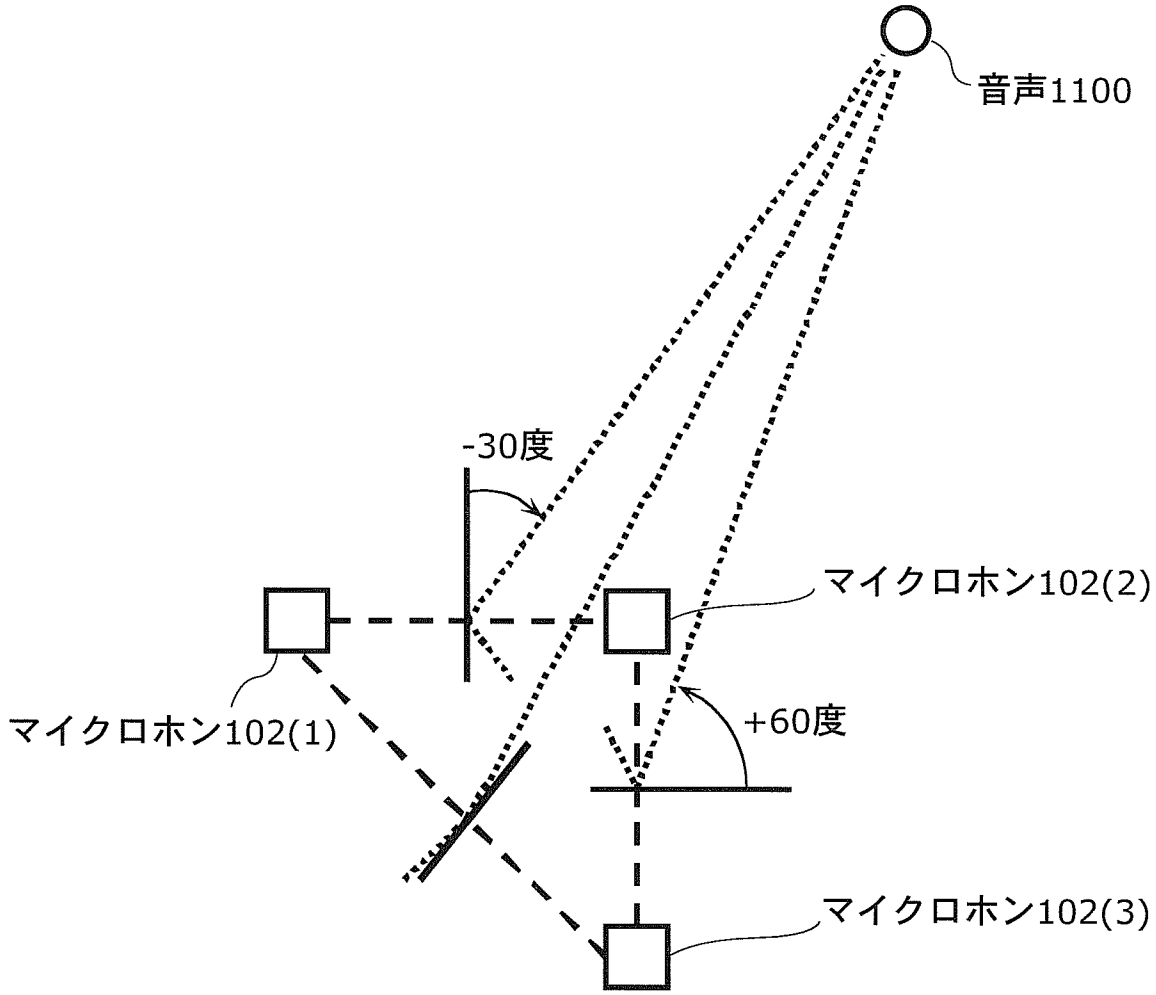

割合を考慮した音源方向（全ての方向ではない）
1000Hz
800
600
400
200
50Hz
周波数（Hz）
0秒
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
2.2
2.4
2.6
2.8
3
3.2
3.4
3.6
3.8
4
4.2
4.4
4.6
4.8
5秒
時間(秒)
雑音の影響が小さい
音源の方向
-90度
~
-50度
~
0度
~
+50度
~
+90度
特定しない
空白

[図33]

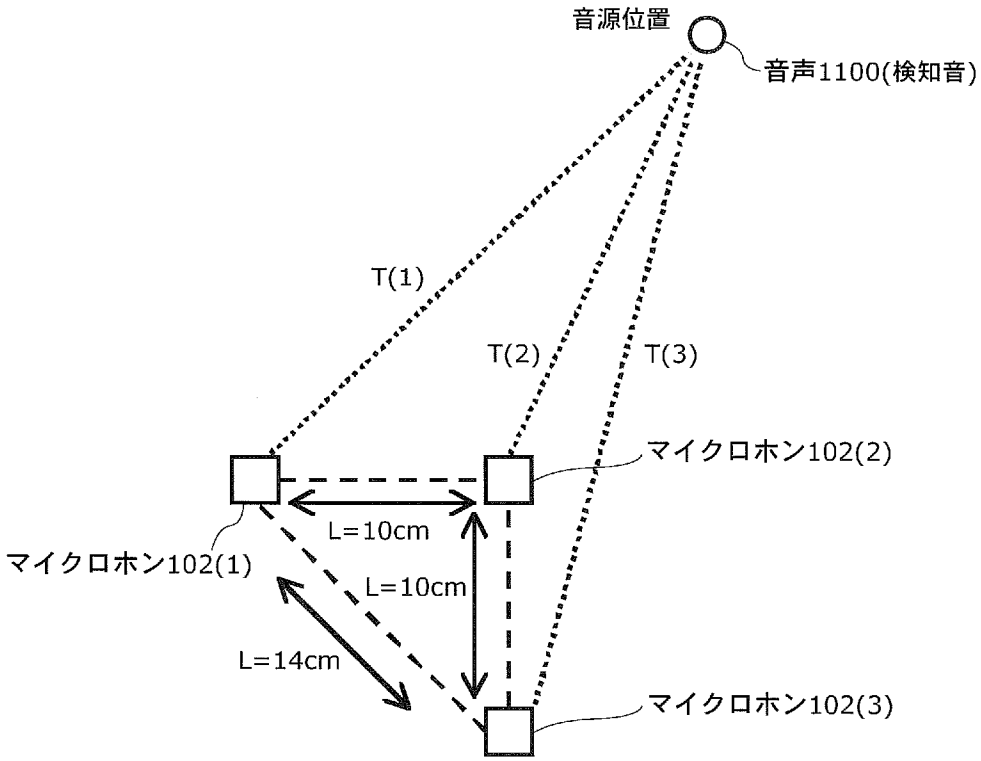
第１のしきい値を１１．５に設定したときの音源方向



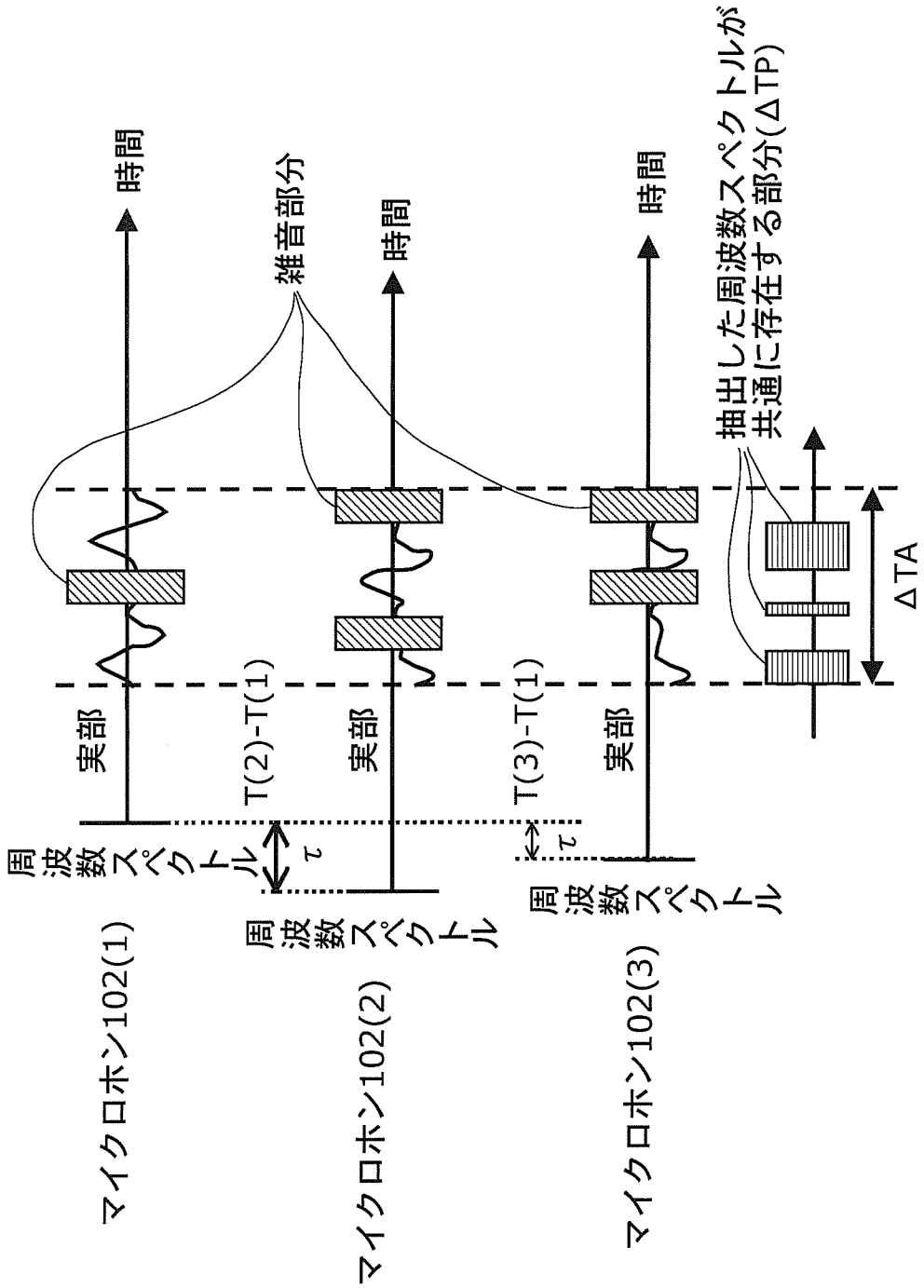
[図34]


音声1100
-30度
マイクロホン102(2)
マイクロホン102(1)
+60度
マイクロホン102(3)

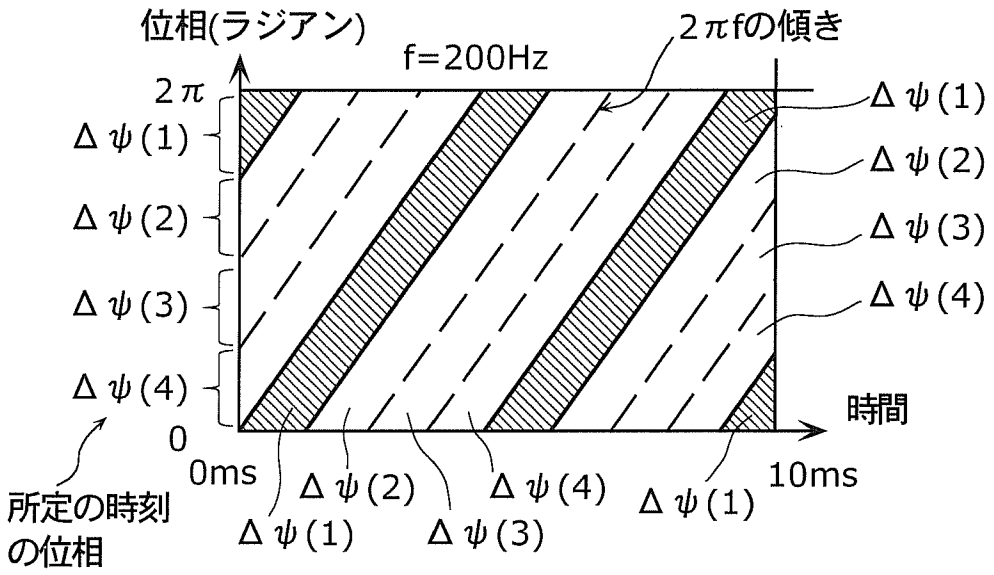
[図35]


音源位置
音声1100(検知音)
T(1)
T(2)
T(3)
マイクロホン102(2)
L=10cm
マイクロホン102(1)
L=10cm
L=14cm
マイクロホン102(3)

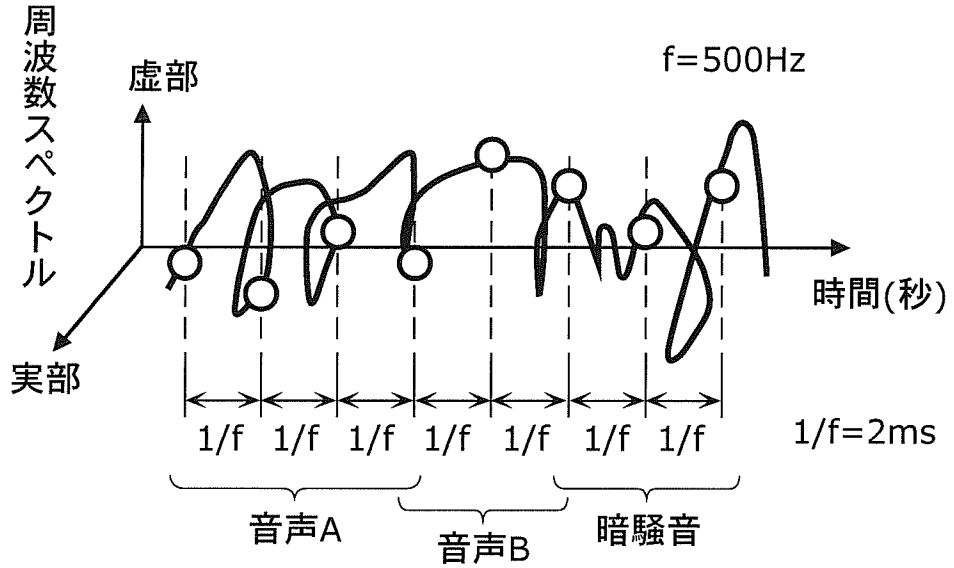
[図36]


マイクロホン102(1)
周波数スペクトル
実部
時間
τ
T(2)-T(1)
マイクロホン102(2)
周波数スペクトル
実部
雑音部分
時間
τ
T(3)-T(1)
マイクロホン102(3)
周波数スペクトル
実部
時間
抽出した周波数スペクトルが共通に存在する部分(ΔTP)
ΔTA

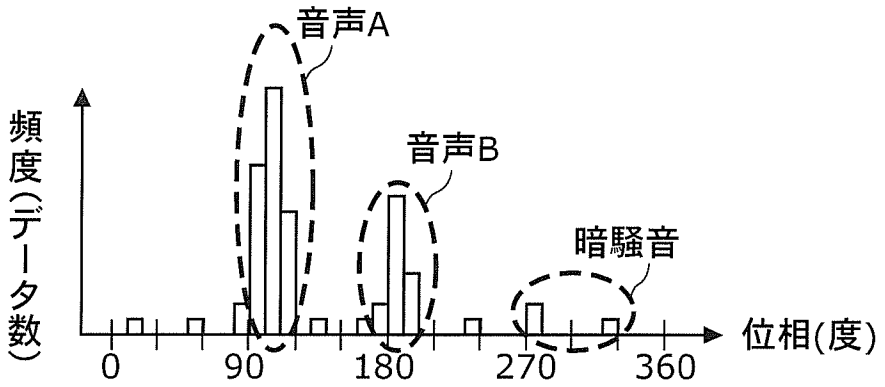
[図37]



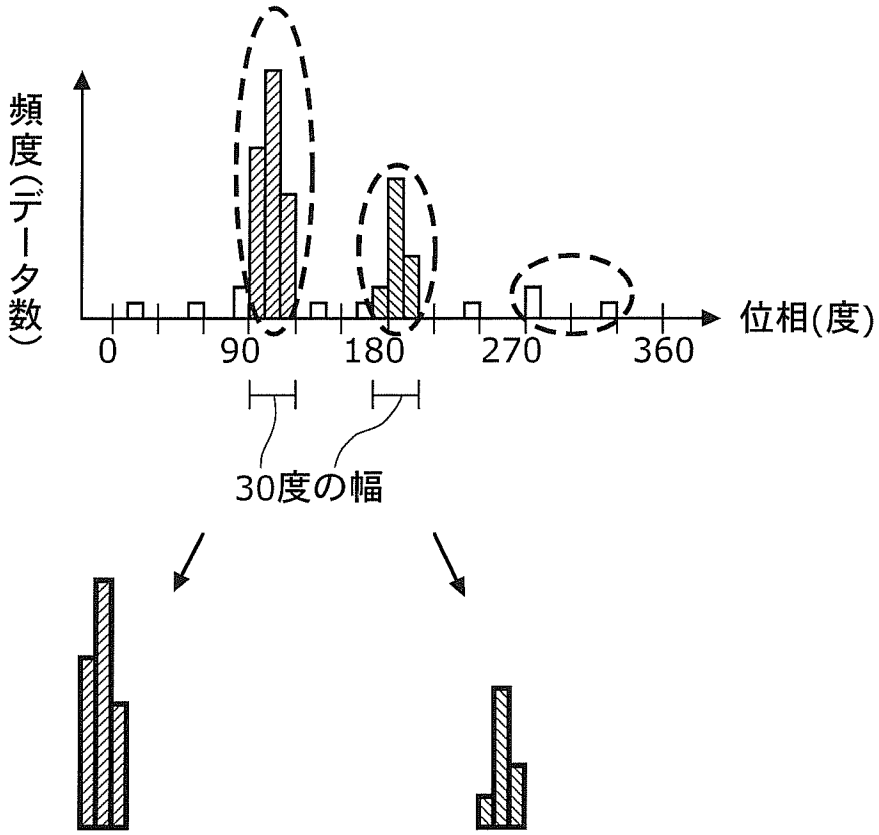
[図38A]



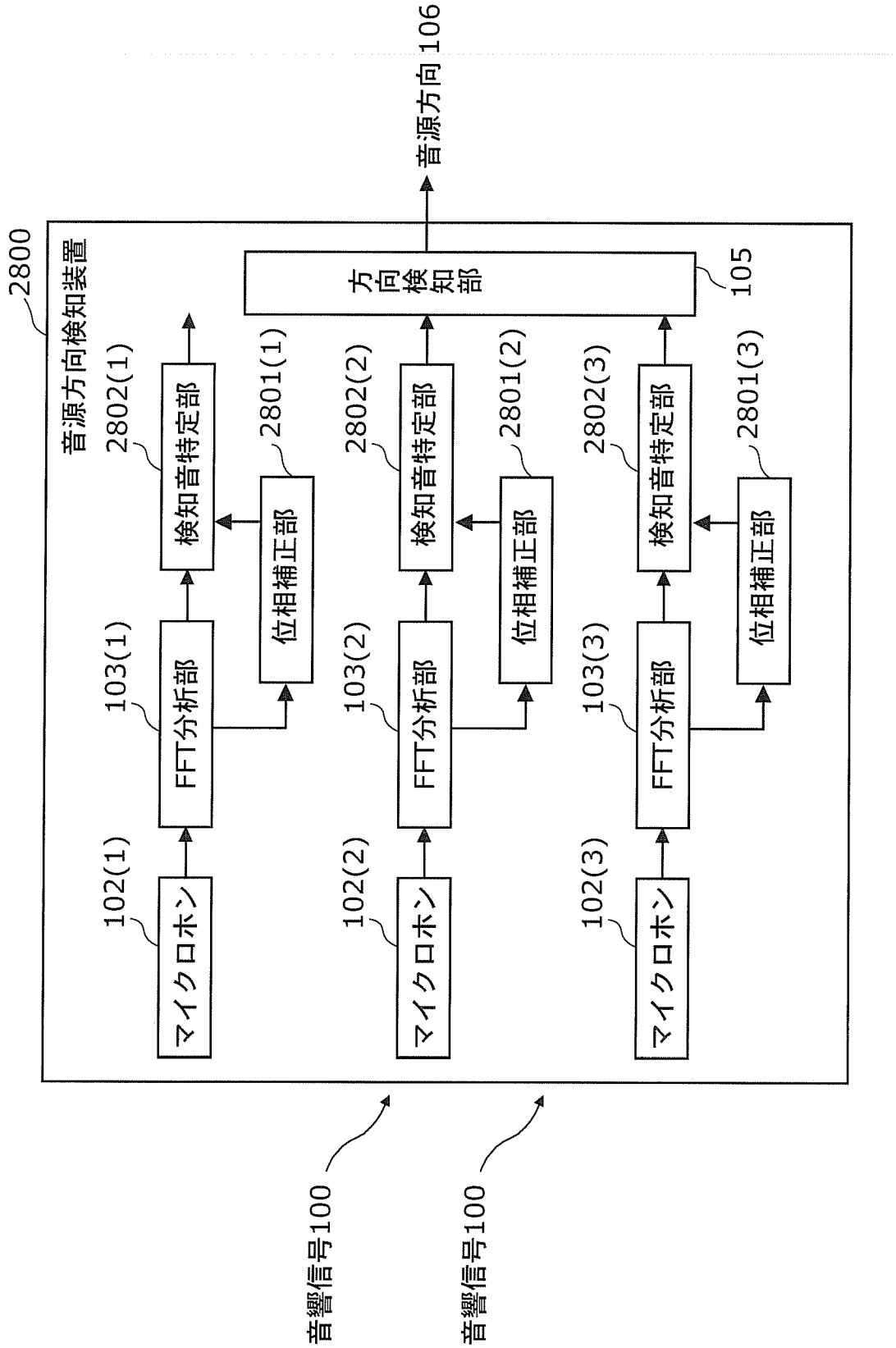
[図38B]



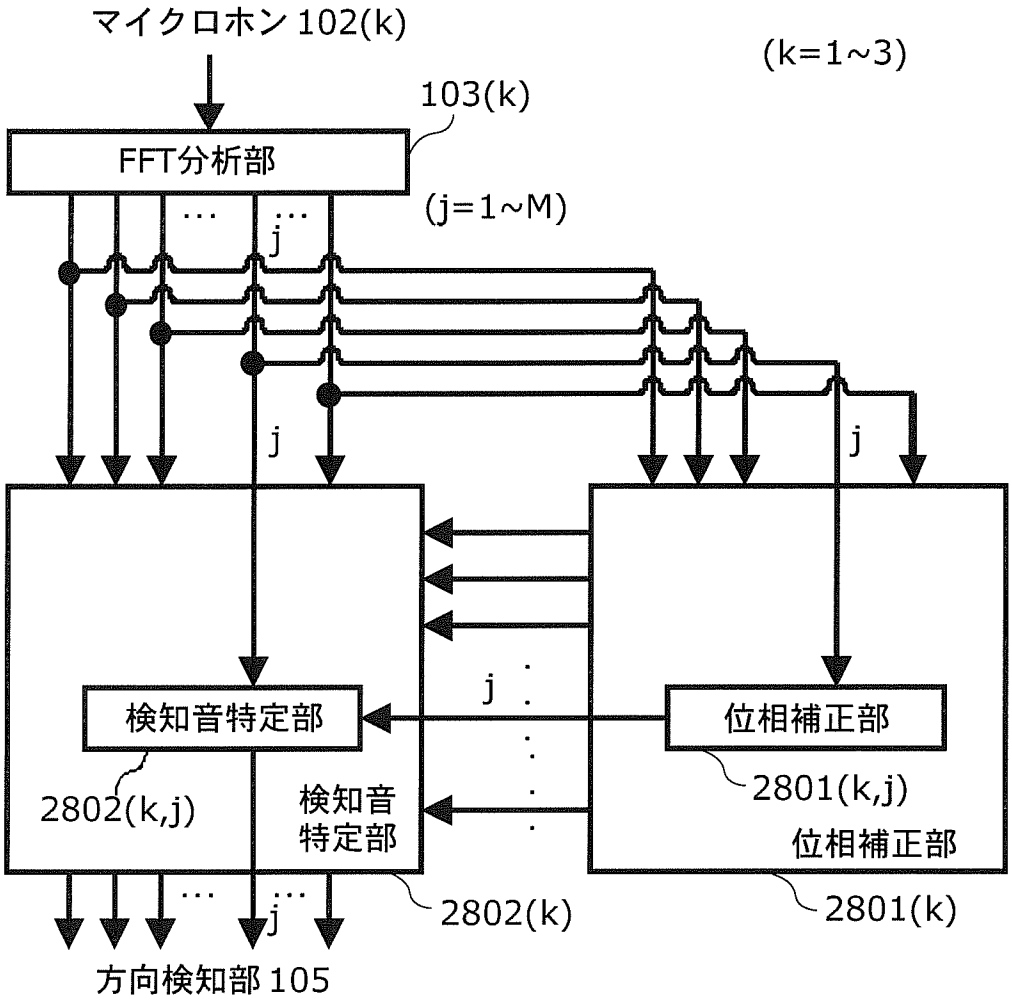
[図39]



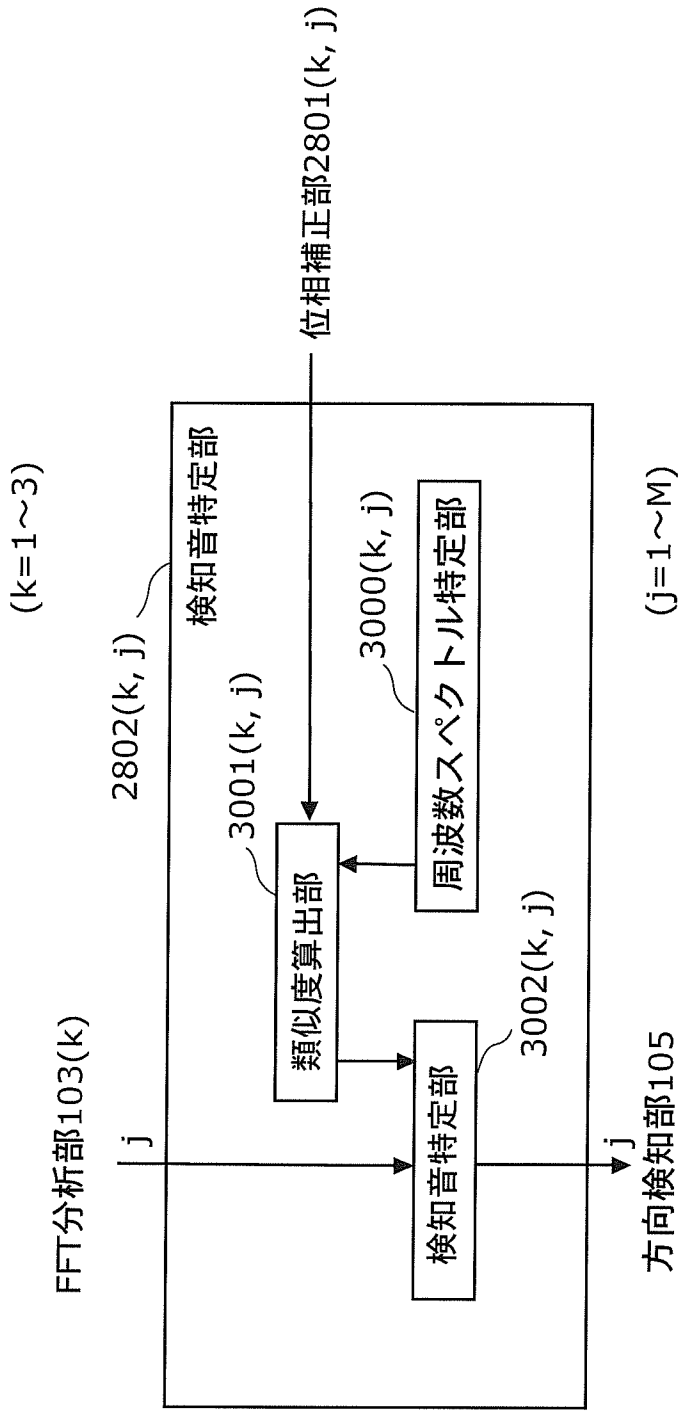
[図40]


音源方向 106
2800
音源方向検知装置
方向検知部
105
2802(1)
検知音特定部
2801(1)
位相補正部
103(1)
FFT分析部
102(1)
マイクロホン
2802(2)
検知音特定部
2801(2)
位相補正部
103(2)
FFT分析部
102(2)
マイクロホン
2802(3)
検知音特定部
2801(3)
位相補正部
103(3)
FFT分析部
102(3)
マイクロホン
音響信号100
音響信号100

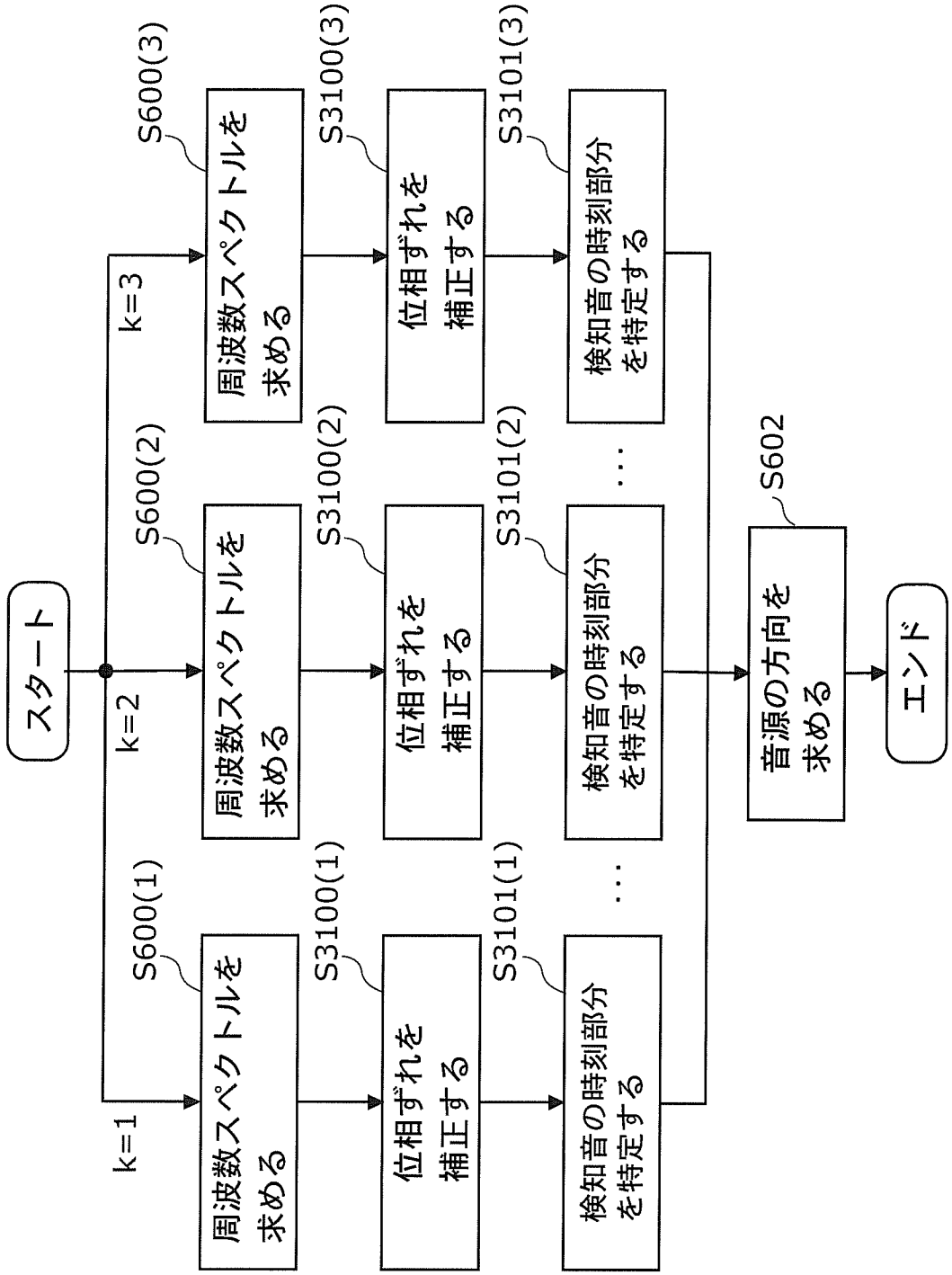
[図41]


マイクロホン 102(k)
(k=1~3)
103(k)
FFT分析部
(j=1~M)
j
j
j
j
検知音特定部
位相補正部
2802(k,j)
2801(k,j)
検知音
特定部
位相補正部
2802(k)
2801(k)
j
方向検知部 105

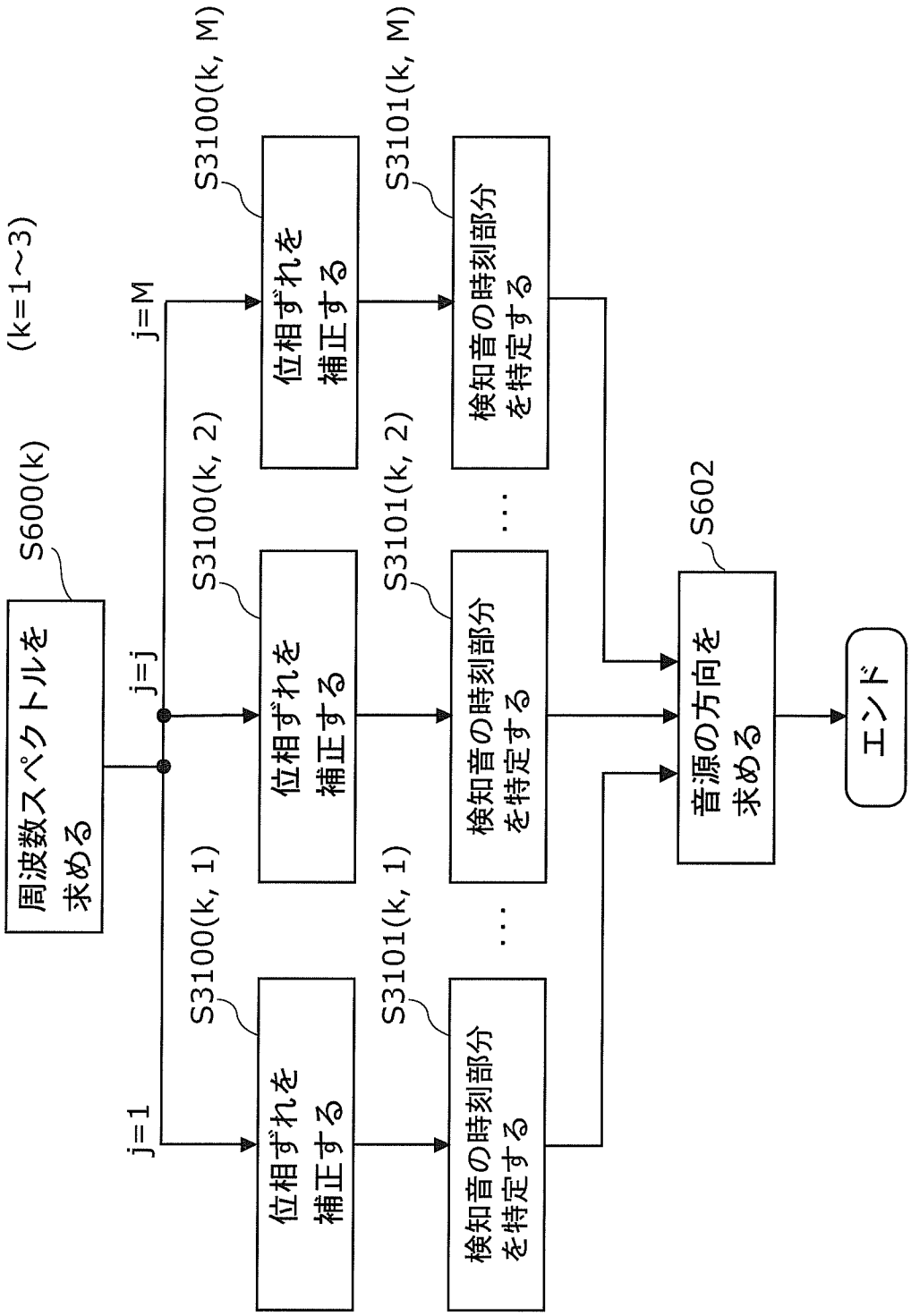
[図42]


位相補正部2801(k, j)
(k=1～3)
検知音特定部
2802(k, j)
3000(k, j)
周波数スペクトル特定部
(j=1～M)
3001(k, j)
類似度算出部
3002(k, j)
FFT分析部103(k)
j
検知音特定部
j
方向検知部105

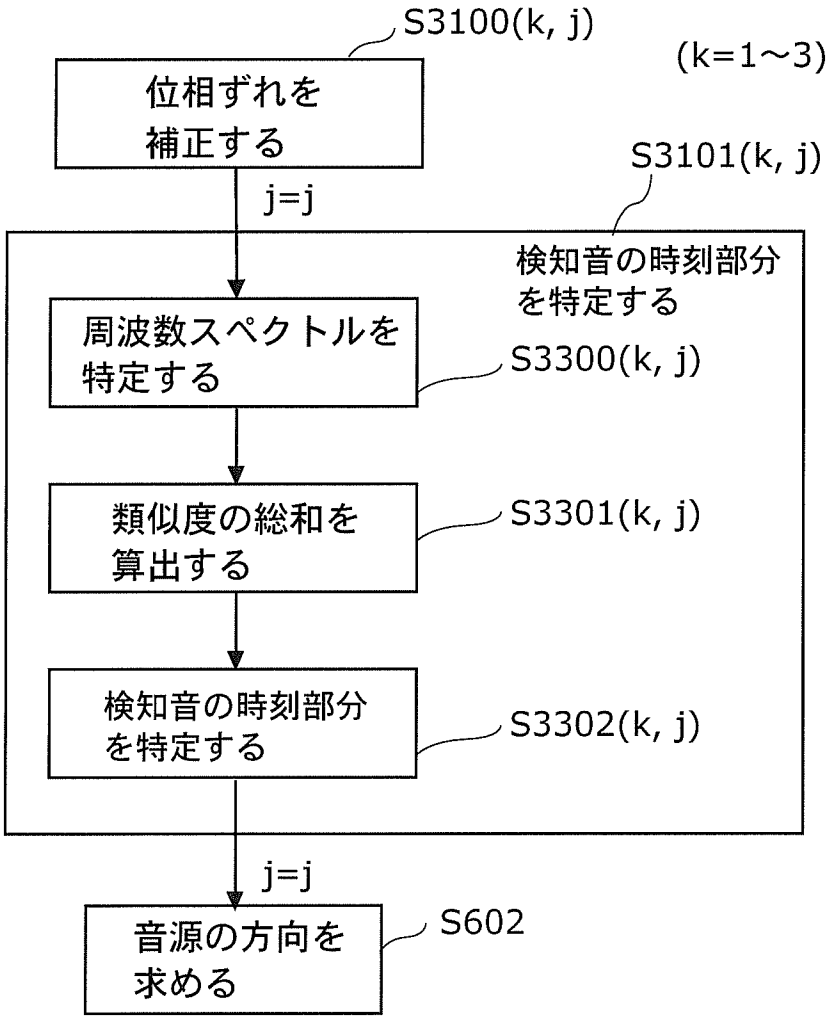
[図43]


スタート
k=1
k=2
k=3
S600(1)
S600(2)
S600(3)
周波数スペクトルを求める
周波数スペクトルを求める
周波数スペクトルを求める
S3100(1)
S3100(2)
S3100(3)
位相ずれを補正する
位相ずれを補正する
位相ずれを補正する
S3101(1)
S3101(2)
S3101(3)
検知音の時刻部分を特定する
検知音の時刻部分を特定する
検知音の時刻部分を特定する
S602
音源の方向を求める
エンド

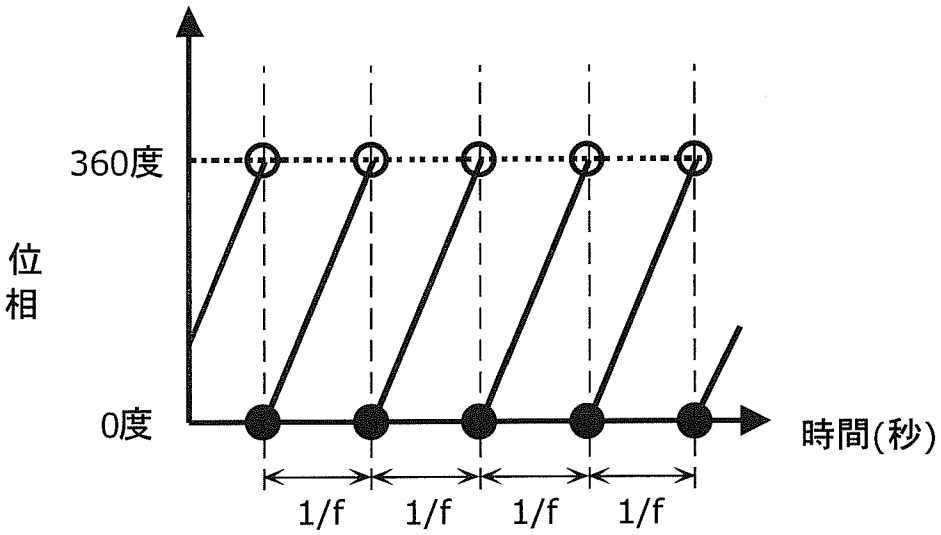
[図44]


(k=1～3)
S600(k)
周波数スペクトルを求める
j=1
j=j
j=M
S3100(k, 1)
S3100(k, 2)
S3100(k, M)
位相ずれを補正する
位相ずれを補正する
位相ずれを補正する
S3101(k, 1)
S3101(k, 2)
S3101(k, M)
検知音の時刻部分を特定する
検知音の時刻部分を特定する
検知音の時刻部分を特定する
S602
音源の方向を求める
エンド

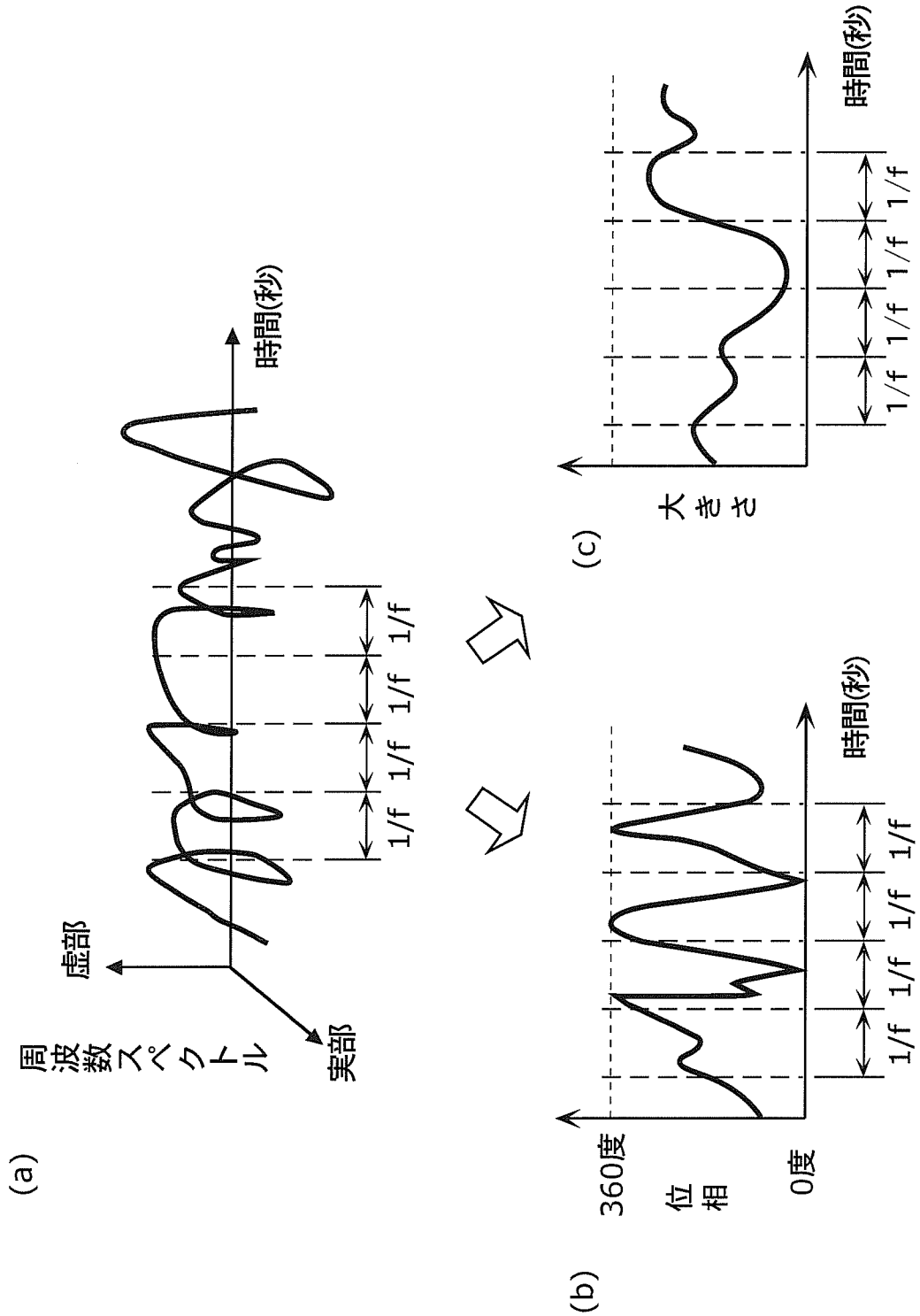
[図45]


S3100(k, j)
(k=1～3)
位相ずれを
補正する
S3101(k, j)
j=j
検知音の時刻部分
を特定する
周波数スペクトルを
特定する
S3300(k, j)
類似度の総和を
算出する
S3301(k, j)
検知音の時刻部分
を特定する
S3302(k, j)
j=j
音源の方向を
求める
S602

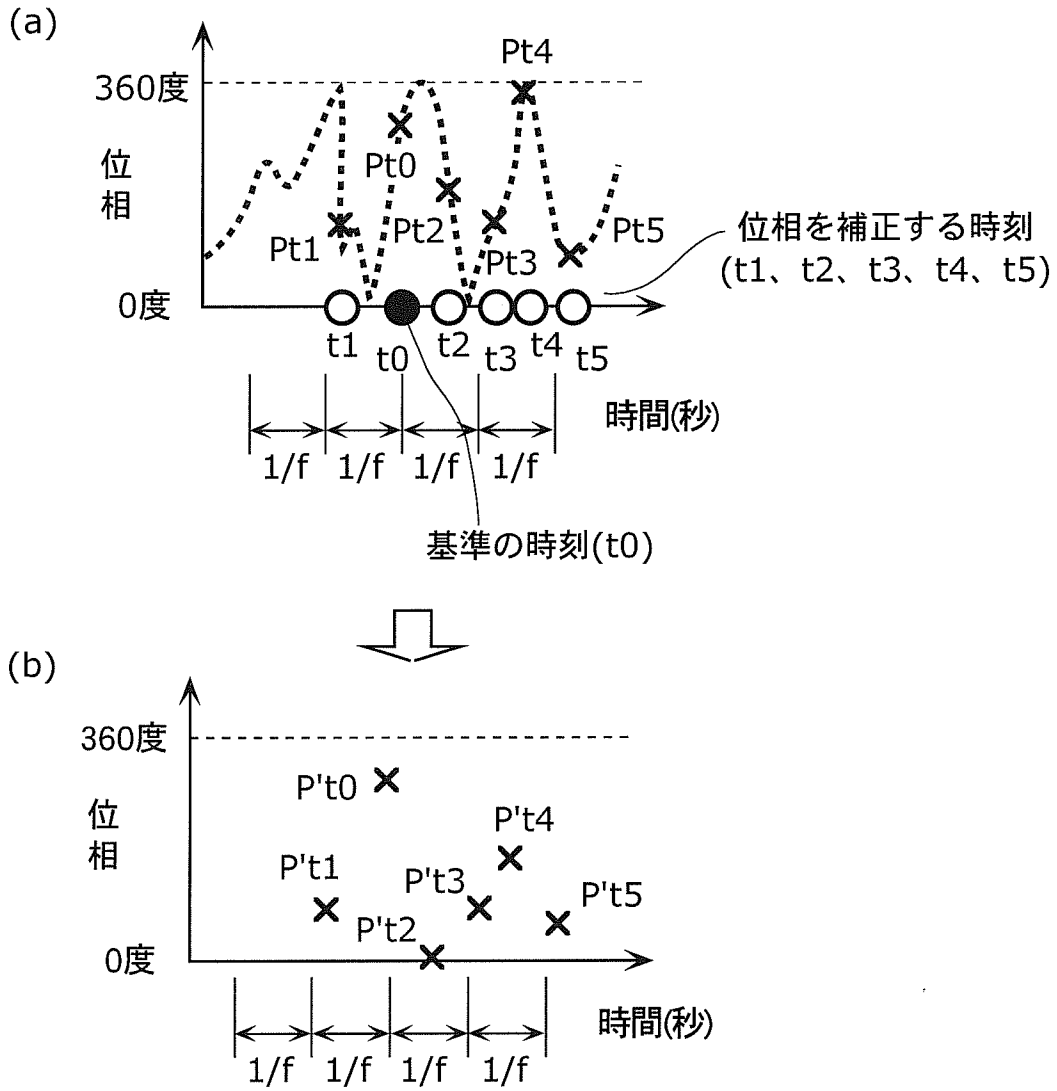
[図46]


360度
位
相
0度
時間(秒)
1/f
1/f
1/f
1/f

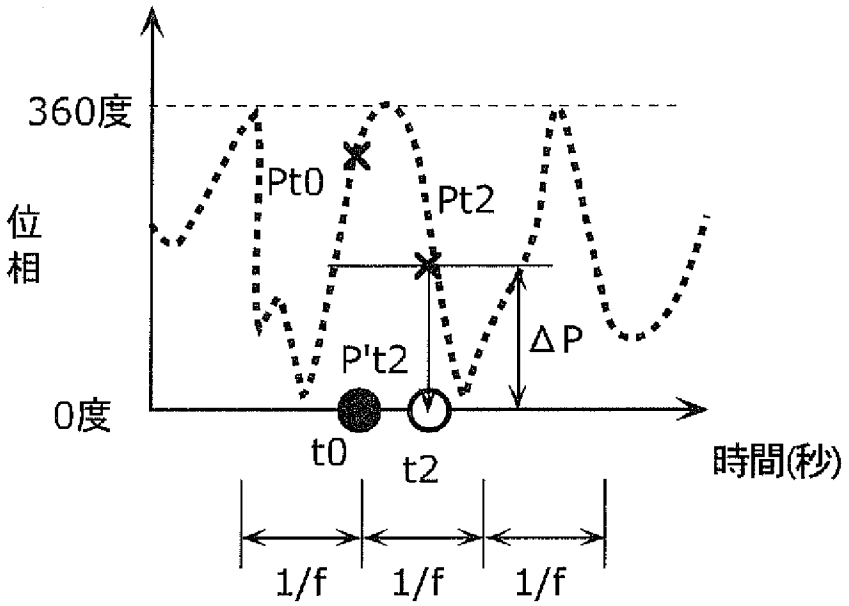
[図47]


(a)
周波数スペクトル
虚部
実部
時間(秒)
1/f
1/f
1/f
1/f
(b)
360度
位相
0度
時間(秒)
1/f
1/f
1/f
1/f
(c)
大きさ
時間(秒)
1/f
1/f
1/f
1/f

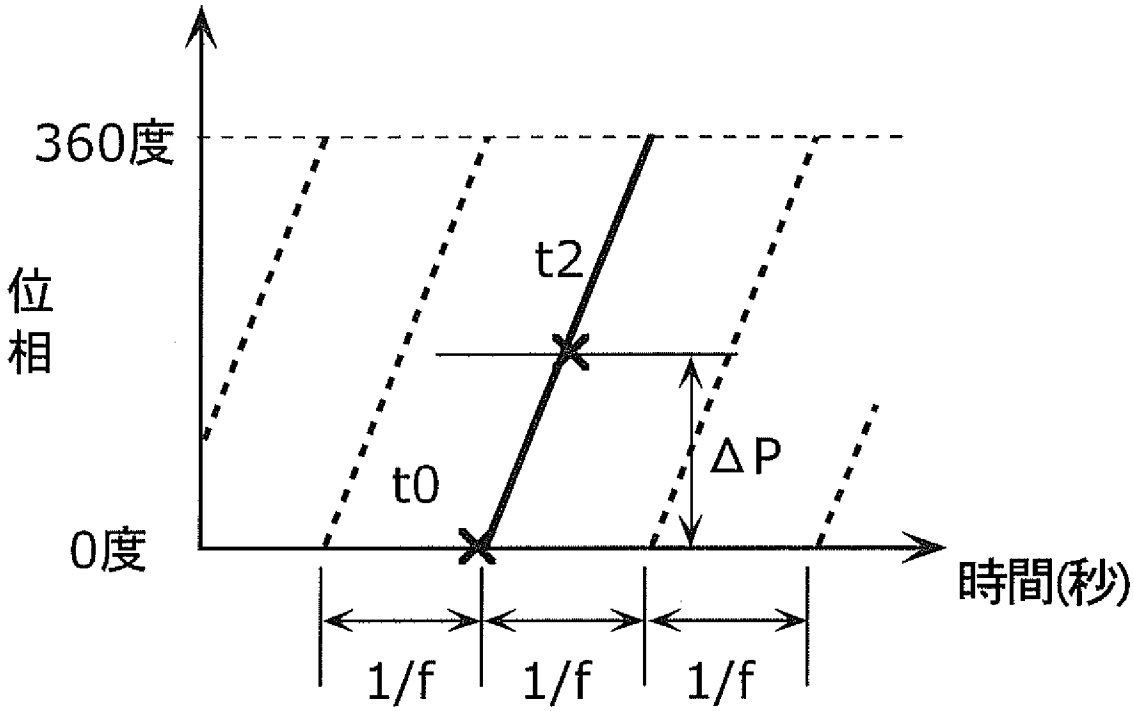
[図48]


(a)
360度
位相
0度
Pt4
Pt0
Pt1
Pt2
Pt3
Pt5
位相を補正する時刻
(t1、t2、t3、t4、t5)
t1
t0
t2
t3
t4
t5
時間(秒)
1/f
1/f
1/f
1/f
基準の時刻(t0)
(b)
360度
位相
0度
P't0
P't4
P't1
P't3
P't5
P't2
時間(秒)
1/f
1/f
1/f
1/f

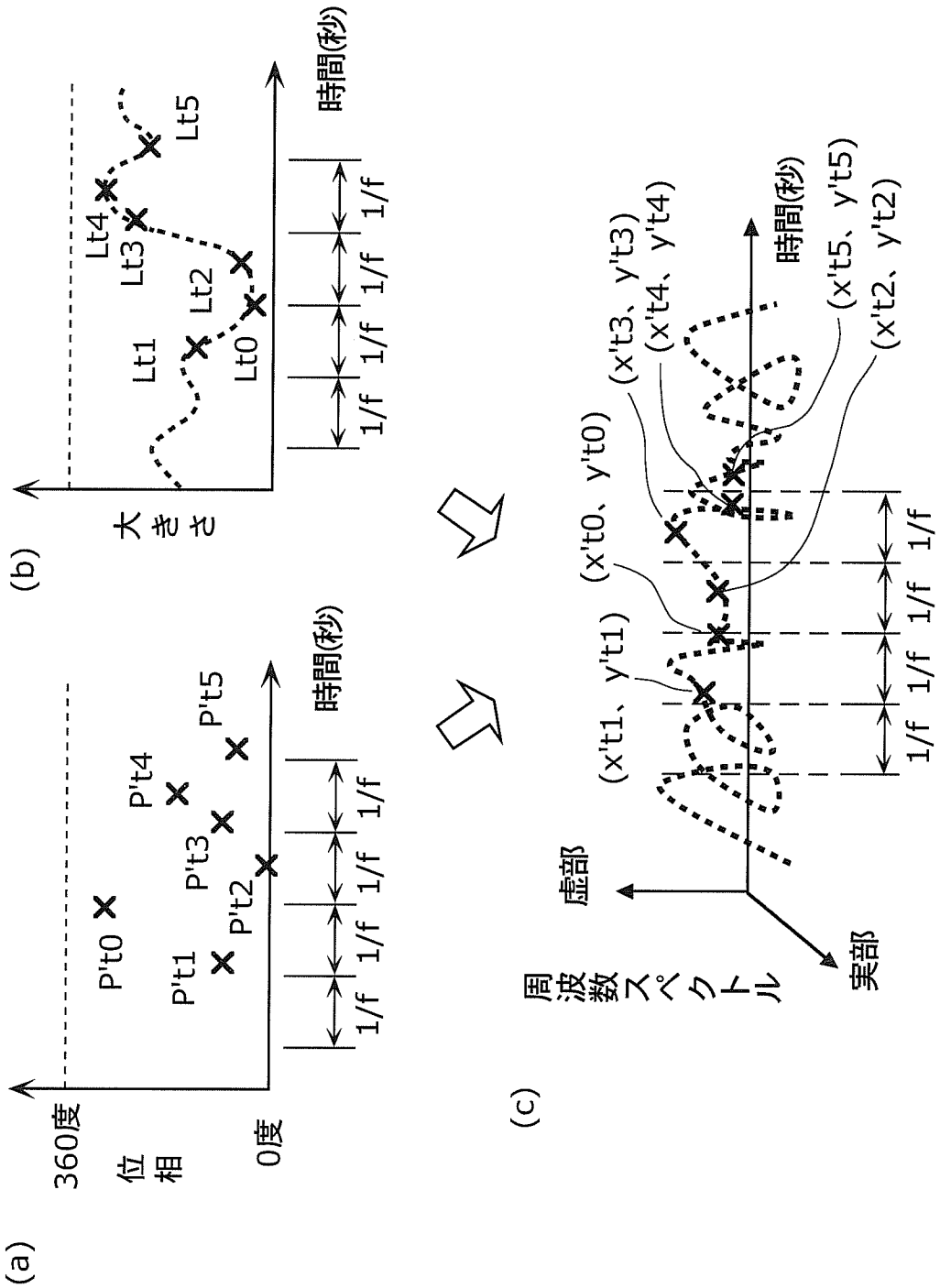
[図49A]


360度
位相
0度
Pt0
Pt2
P't2
ΔP
t0
t2
時間(秒)
1/f
1/f
1/f

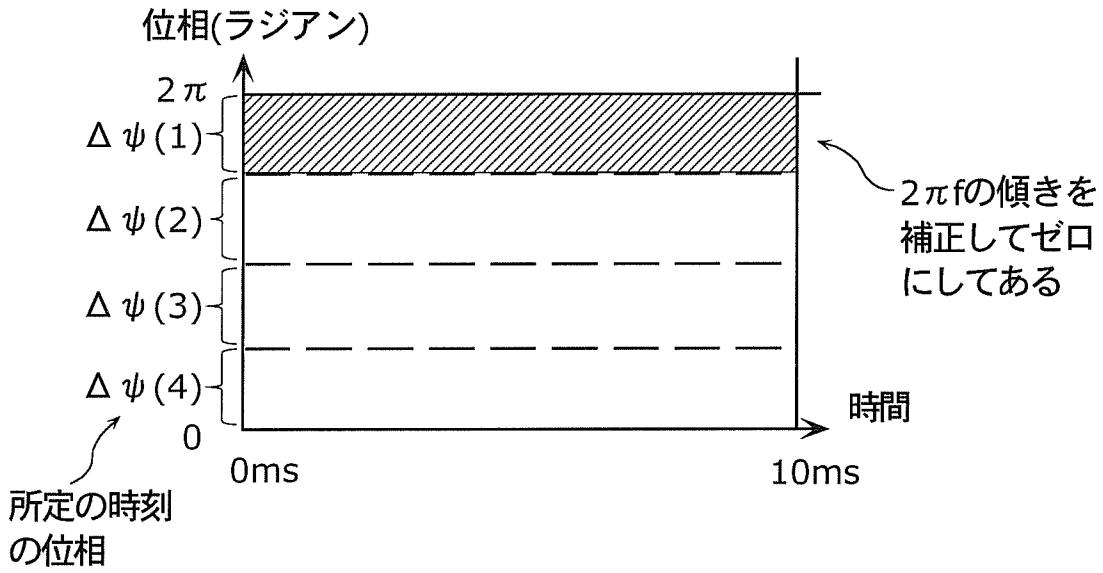
[図49B]



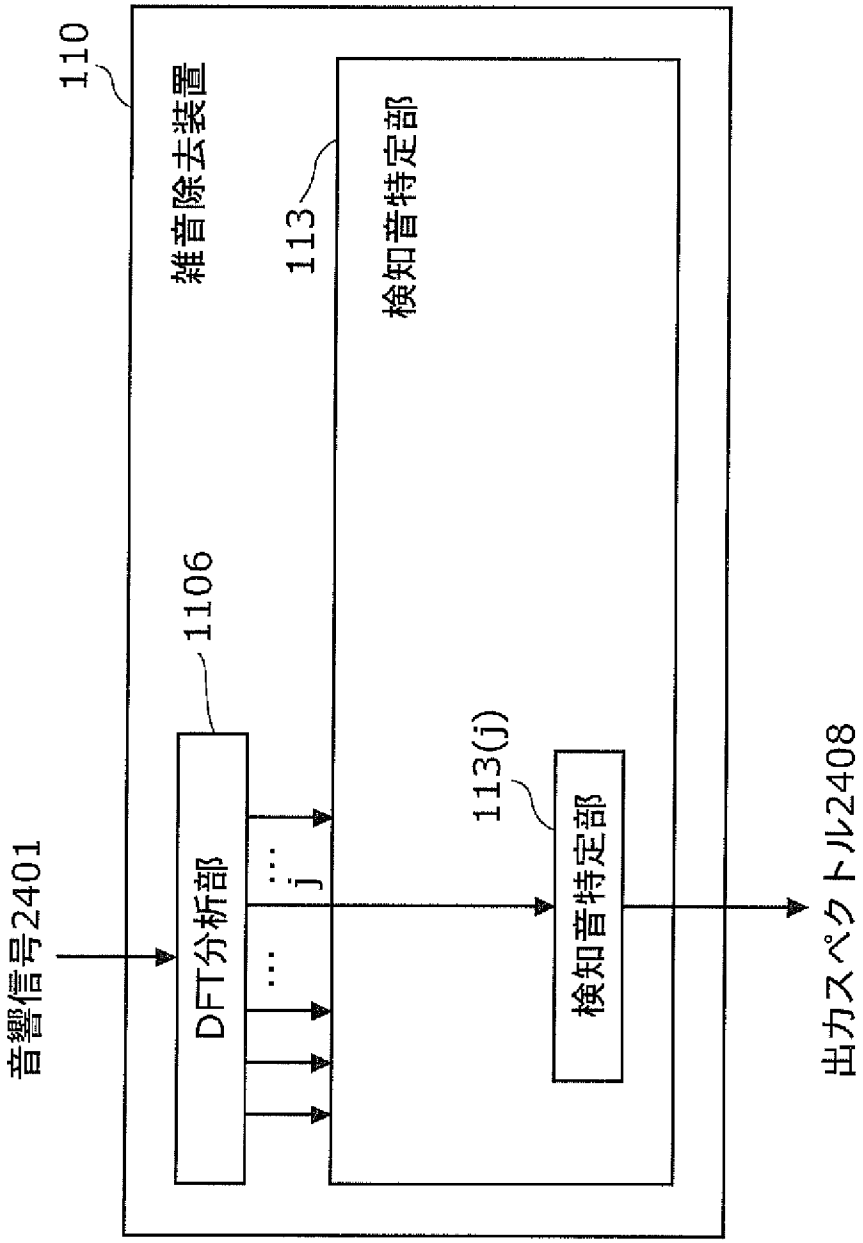
[図50]


(a)
360度
位相
0度
P't0
P't1
P't2
P't3
P't4
P't5
時間(秒)
1/f
(b)
大きさ
Lt0
Lt1
Lt2
Lt3
Lt4
Lt5
時間(秒)
1/f
(c)
周波数スペクトル
虚部
実部
時間(秒)
(x't0、y't0)
(x't1、y't1)
(x't2、y't2)
(x't3、y't3)
(x't4、y't4)
(x't5、y't5)
1/f

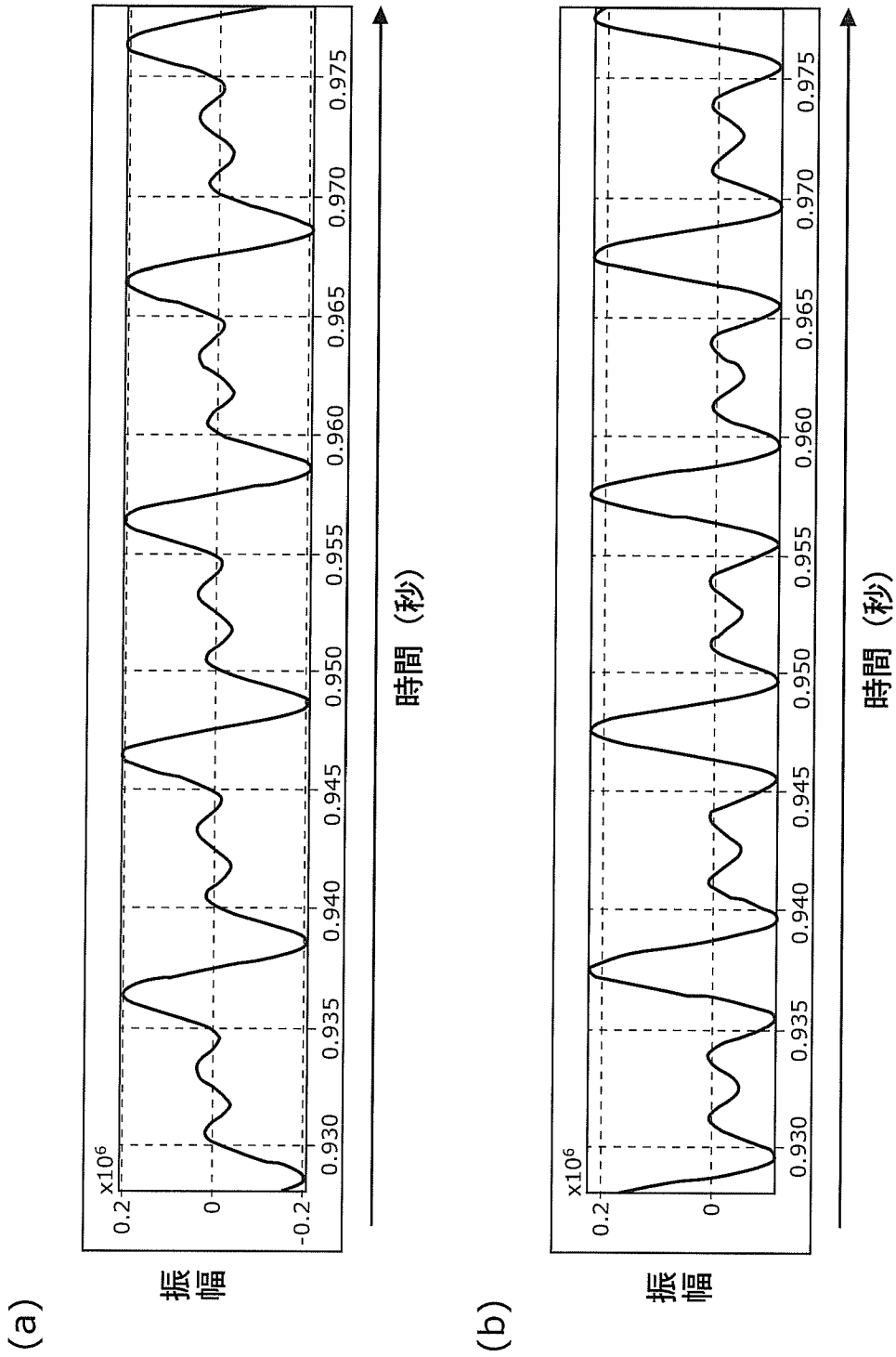
[図51]


位相(ラジアン)
2π
Δψ(1)
Δψ(2)
Δψ(3)
Δψ(4)
0
0ms
10ms
時間
2πfの傾きを補正してゼロにしてある
所定の時刻の位相

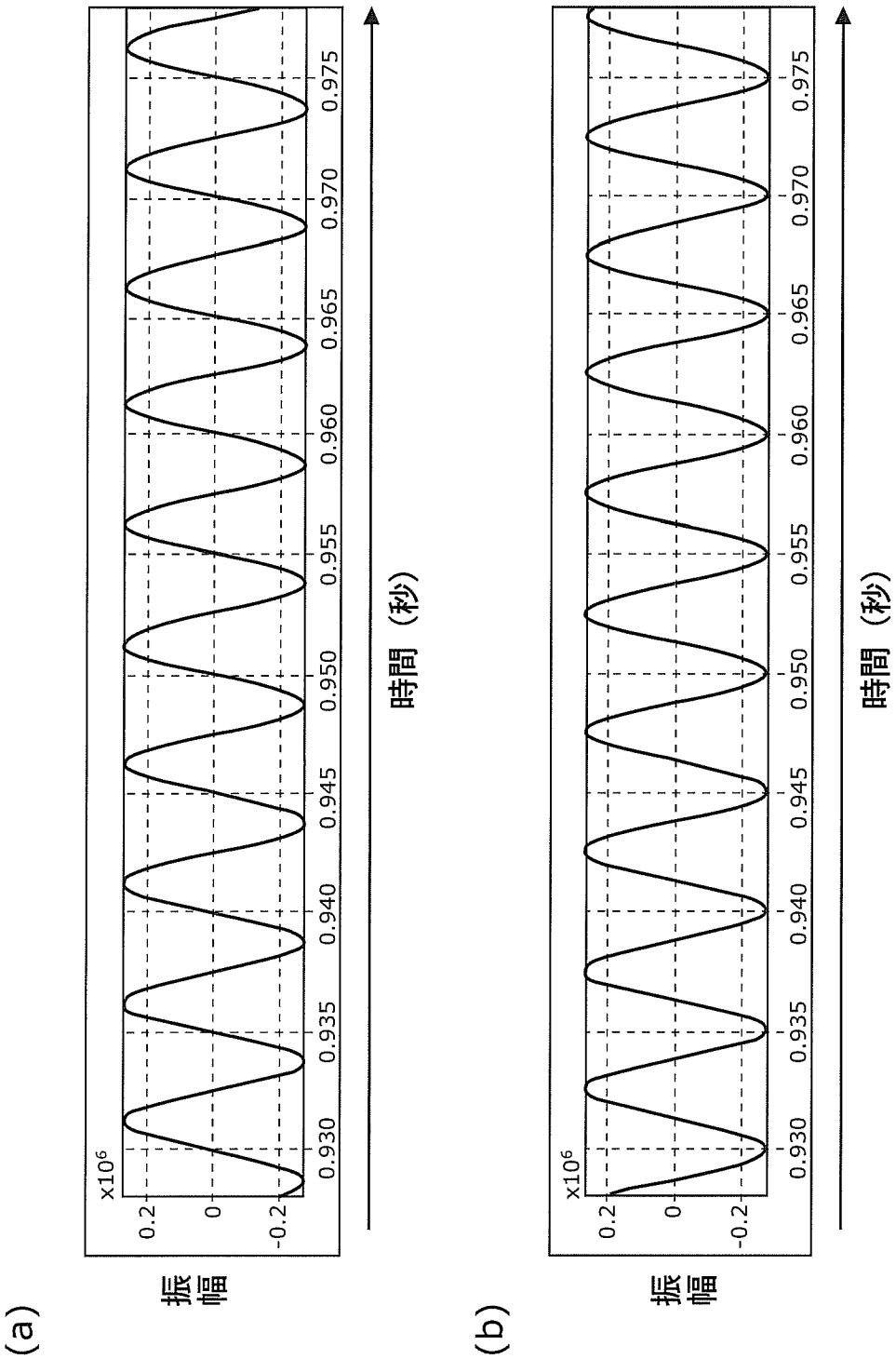
[図52]


110
雑音除去装置
音響信号2401
1106
DFT分析部
…
j
…
113
検知音特定部
113(j)
検知音特定部
出力スペクトル2408

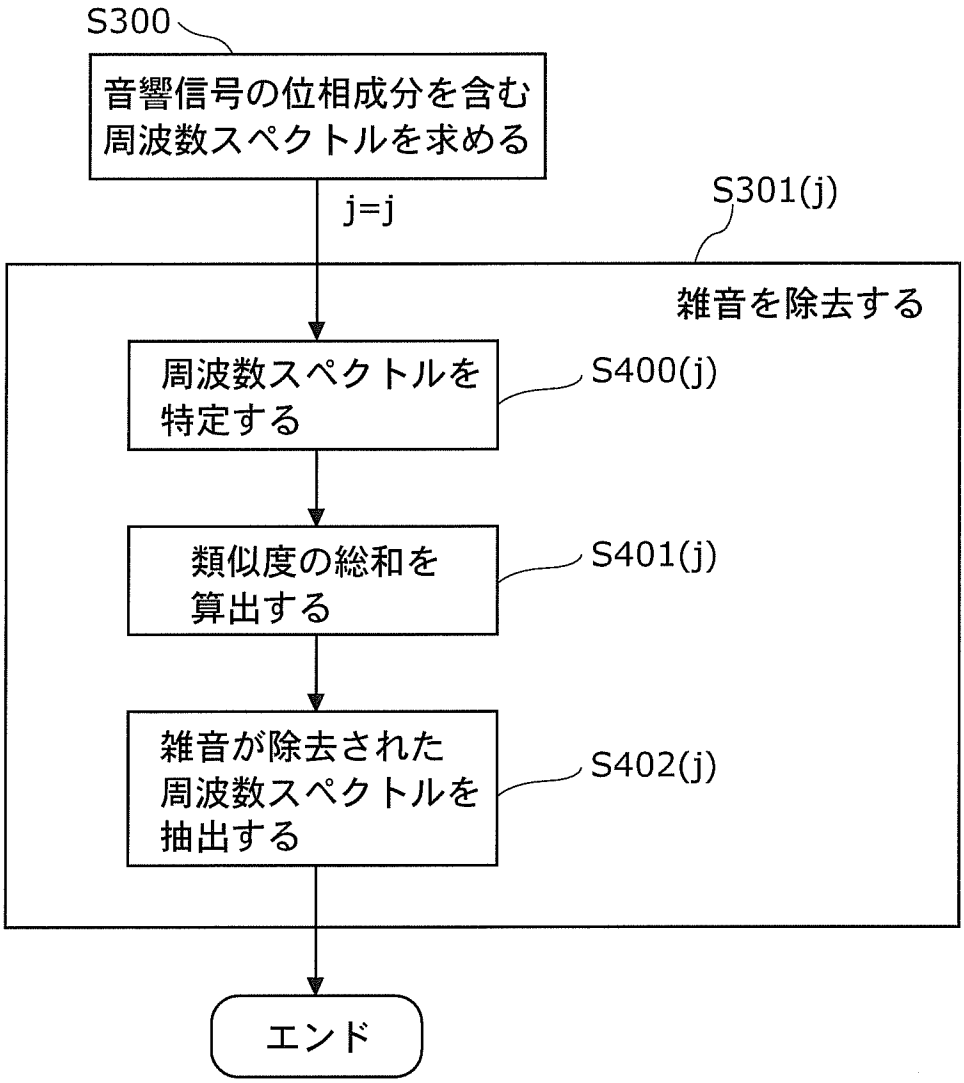
[図53]


(a)
振幅
x10^6
0.2
0
-0.2
0.930
0.935
0.940
0.945
0.950
0.955
0.960
0.965
0.970
0.975
時間（秒）
(b)
振幅
x10^6
0.2
0
0.930
0.935
0.940
0.945
0.950
0.955
0.960
0.965
0.970
0.975
時間（秒）

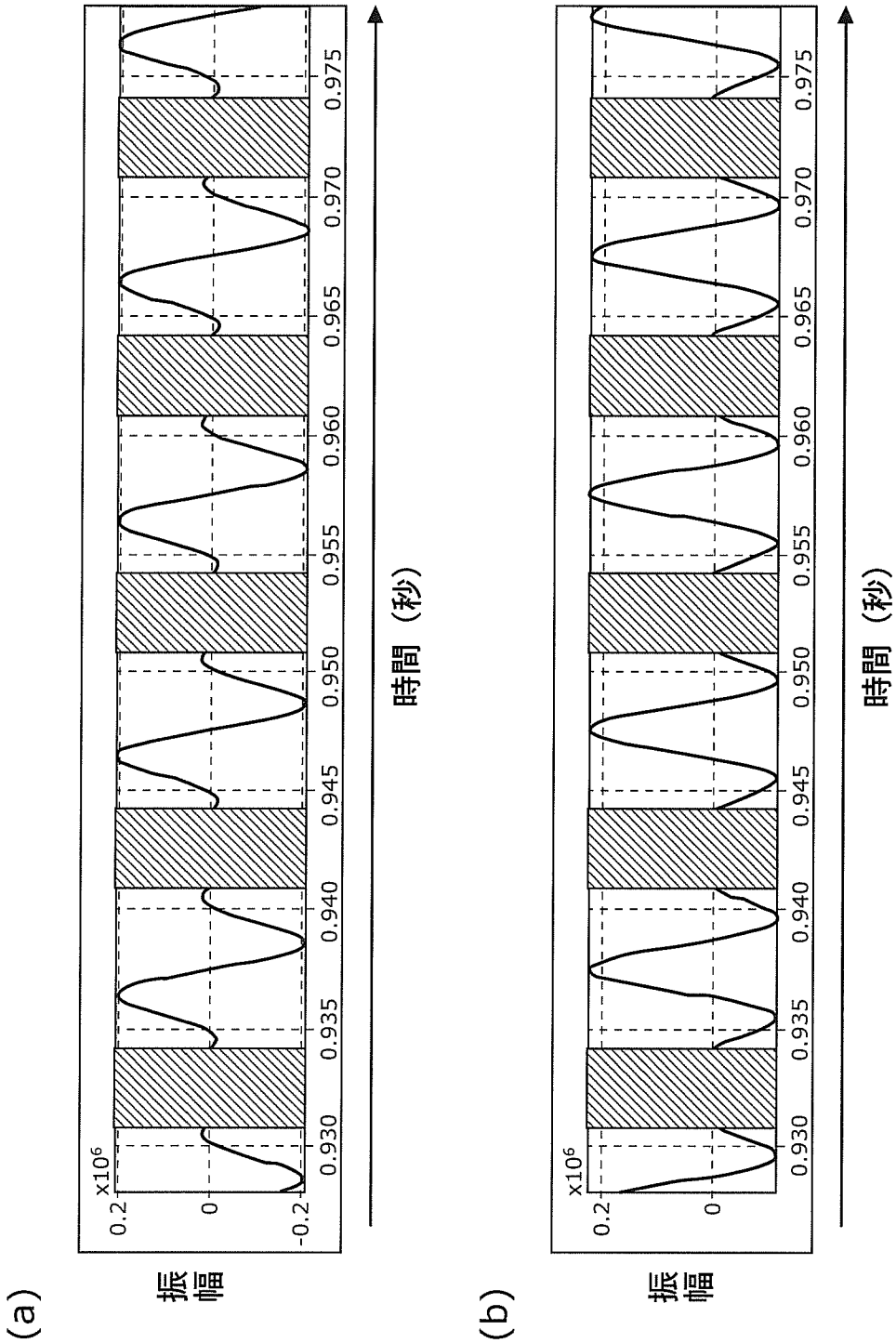
[図54]


(a)
振幅
x10^6
0.2
0
-0.2
0.930
0.935
0.940
0.945
0.950
0.955
0.960
0.965
0.970
0.975
時間（秒）
(b)
振幅
x10^6
0.2
0
-0.2
0.930
0.935
0.940
0.945
0.950
0.955
0.960
0.965
0.970
0.975
時間（秒）

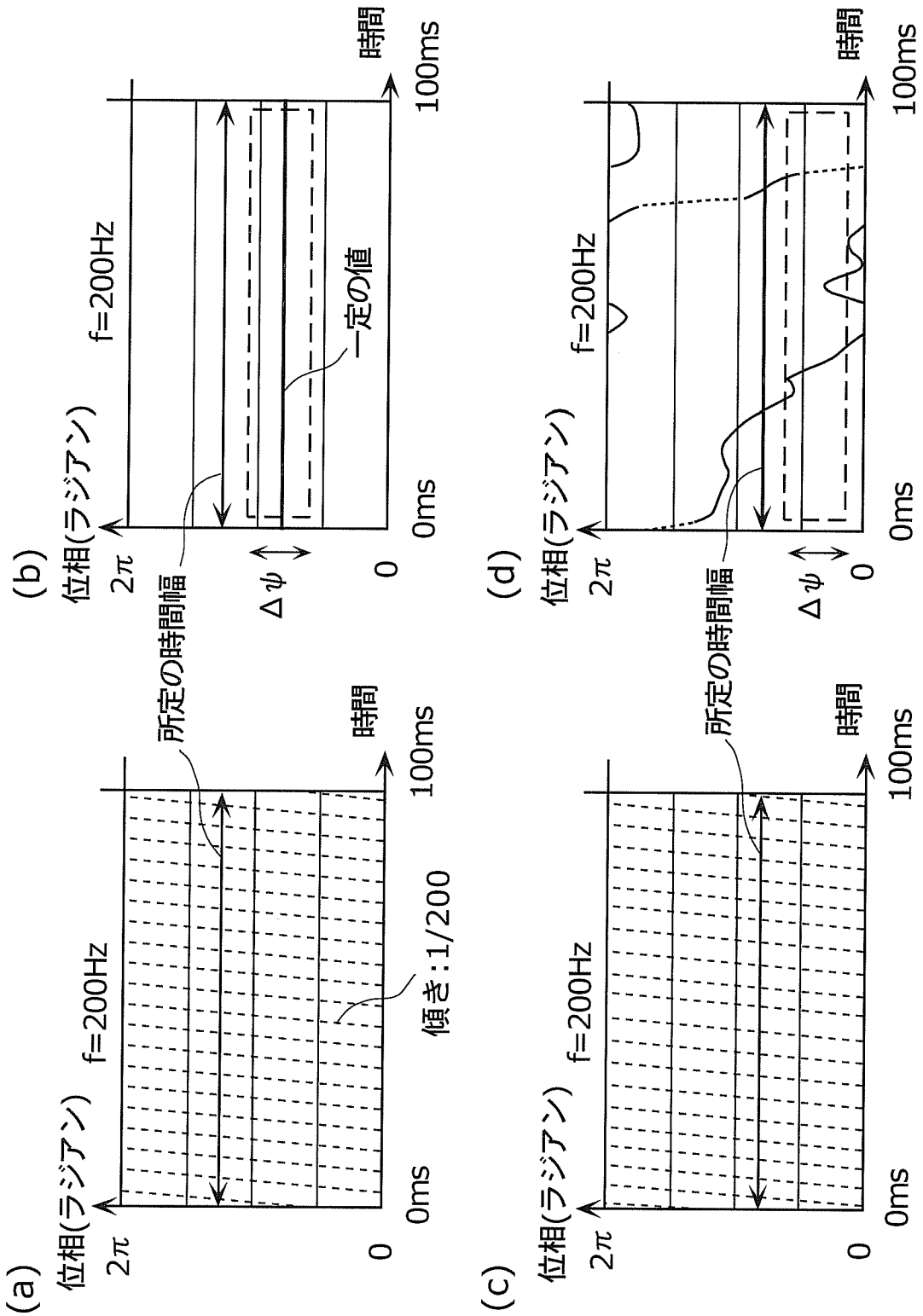
[図55]


S300
音響信号の位相成分を含む周波数スペクトルを求める
j=j
S301(j)
雑音を除去する
周波数スペクトルを特定する
S400(j)
類似度の総和を算出する
S401(j)
雑音が除去された周波数スペクトルを抽出する
S402(j)
エンド

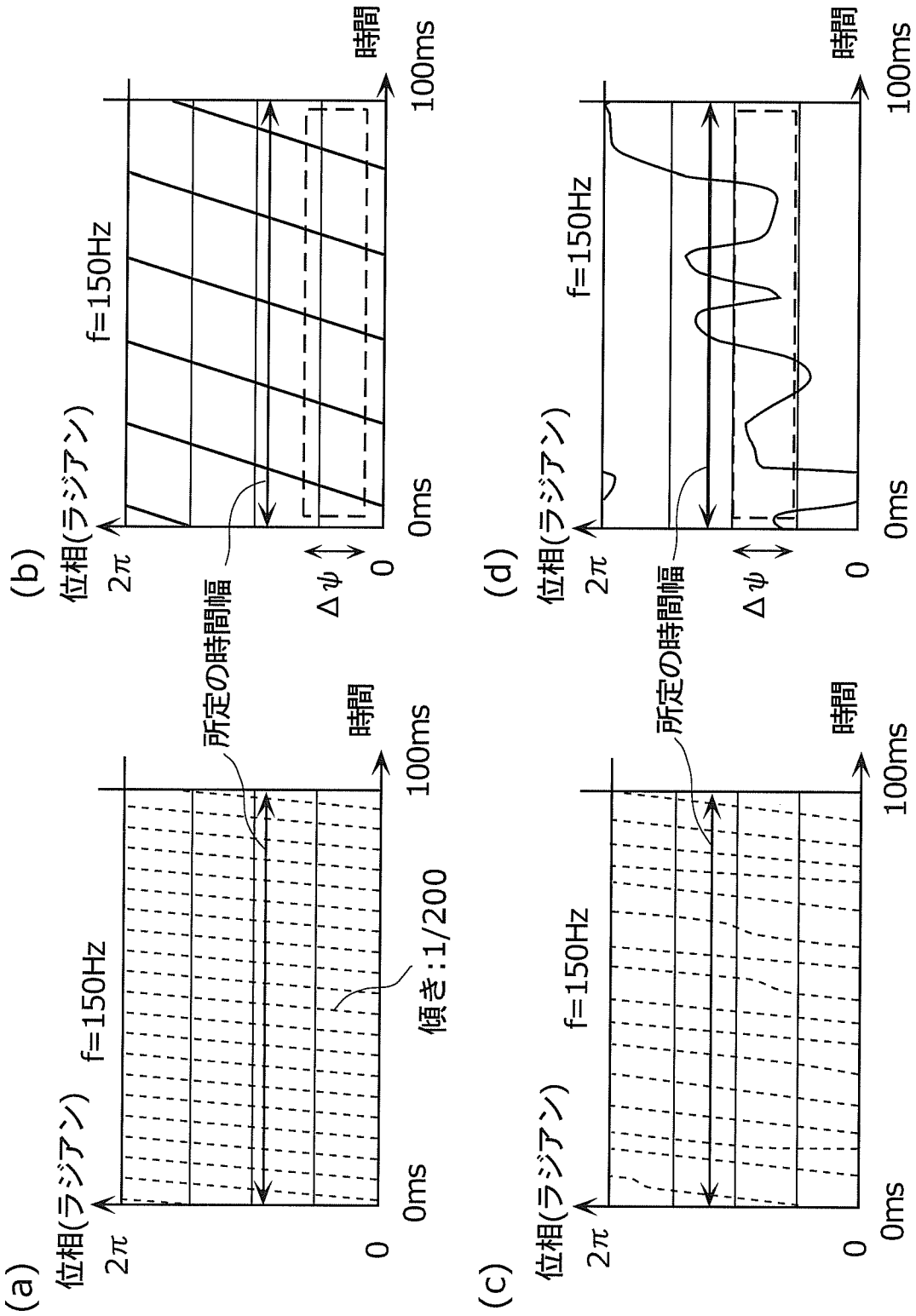
[図56]


(a)
振幅
x10^6
0.2
0
-0.2
0.930
0.935
0.940
0.945
0.950
0.955
0.960
0.965
0.970
0.975
時間（秒）
(b)
振幅
x10^6
0.2
0
0.930
0.935
0.940
0.945
0.950
0.955
0.960
0.965
0.970
0.975
時間（秒）

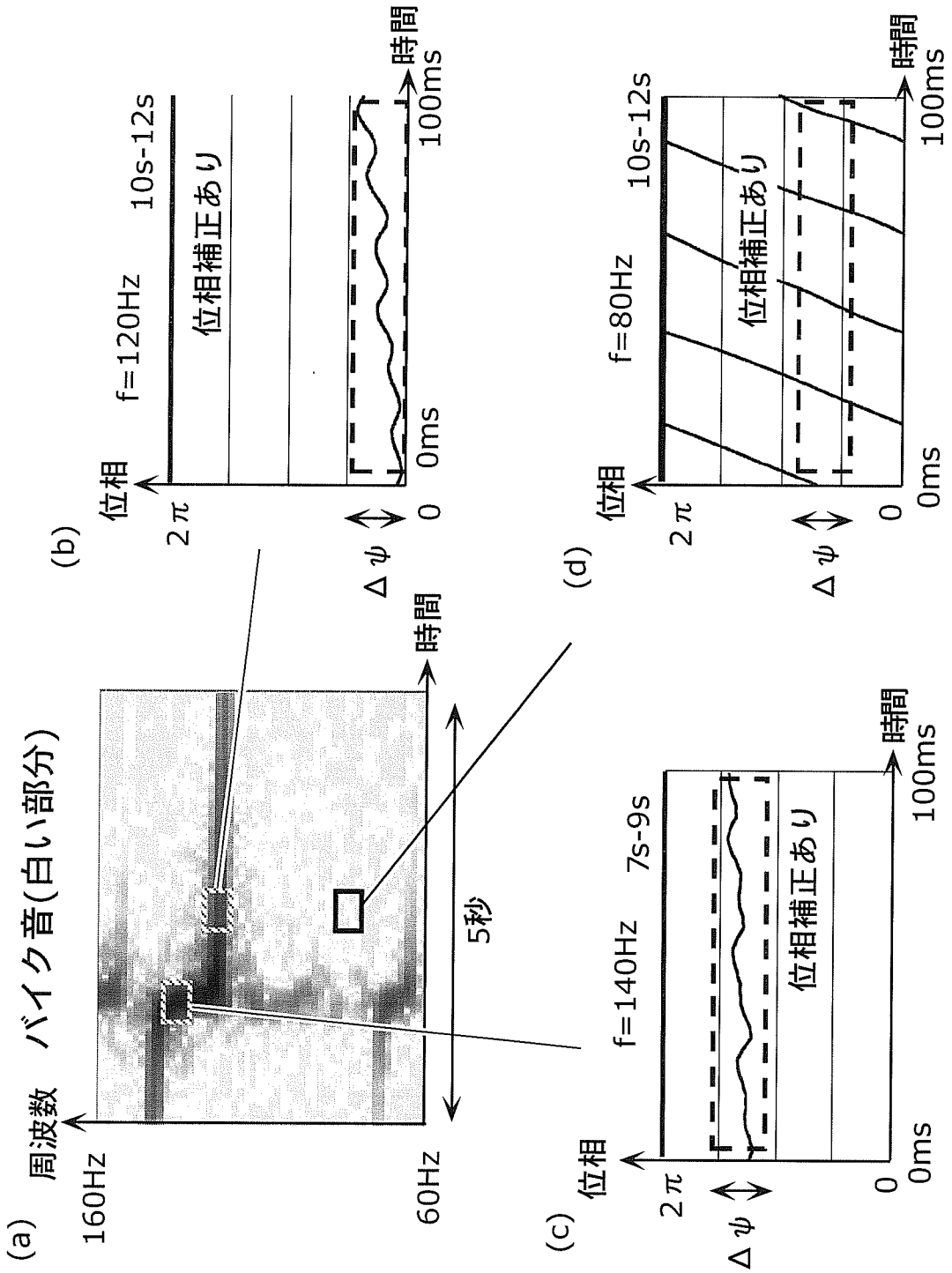
[図57]


(a)
位相(ラジアン)
2π
0
f=200Hz
所定の時間幅
傾き:1/200
0ms
100ms
時間
(b)
位相(ラジアン)
2π
0
f=200Hz
一定の値
Δψ
0ms
100ms
時間
(c)
位相(ラジアン)
2π
0
f=200Hz
所定の時間幅
0ms
100ms
時間
(d)
位相(ラジアン)
2π
0
f=200Hz
Δψ
0ms
100ms
時間

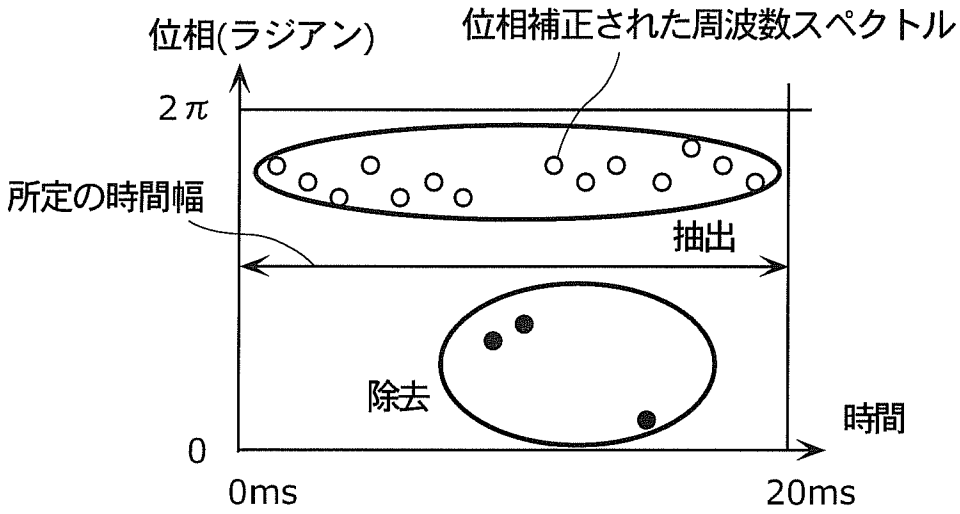
[図58]


(a)
位相(ラジアン)
2π
0
f=150Hz
0ms
100ms
時間
所定の時間幅
傾き:1/200
(b)
位相(ラジアン)
2π
0
f=150Hz
0ms
100ms
時間
Δψ
(c)
位相(ラジアン)
2π
0
f=150Hz
0ms
100ms
時間
所定の時間幅
(d)
位相(ラジアン)
2π
0
f=150Hz
0ms
100ms
時間
Δψ

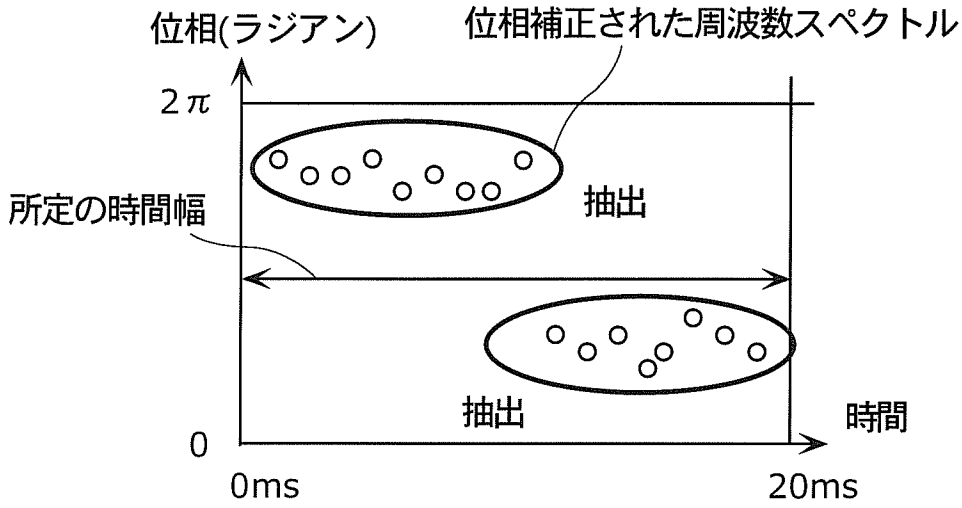
[図59]


(a)
周波数
バイク音(白い部分)
160Hz
60Hz
時間
5秒
(b)
位相
2π
f=120Hz
10s-12s
位相補正あり
Δψ
0
0ms
100ms
時間
(c)
位相
2π
f=140Hz
7s-9s
位相補正あり
Δψ
0
0ms
100ms
時間
(d)
位相
2π
f=80Hz
10s-12s
位相補正あり
Δψ
0
0ms
100ms
時間

[図60A]


位相(ラジアン)
位相補正された周波数スペクトル
2π
所定の時間幅
抽出
除去
時間
0
0ms
20ms

[図60B]


位相(ラジアン)
位相補正された周波数スペクトル
2π
所定の時間幅
抽出
抽出
時間
0
0ms
20ms

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2008/002497

A. CLASSIFICATION OF SUBJECT MATTER
G01S3/808(2006.01)i

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)
G01S1/72-1/82, G01S3/80-3/86, G01S5/18-5/30, G01S7/52-7/66, G01S15/00-15/96

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2008
Kokai Jitsuyo Shinan Koho	1971-2008	Toroku Jitsuyo Shinan Koho	1994-2008

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
IEEE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 09-159744 A (Director General of Technical Research and Development Institute of Defense Agency), 20 June, 1997 (20.06.97), Full text; all drawings (Family: none)	1-6
A	JP 2002-008189 A (Matsushita Electric Industrial Co., Ltd.), 11 January, 2002 (11.01.02), Full text; all drawings & US 2002/0011939 A1 & EP 1168274 A1	1-6

☒ 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 application or patent 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 07 October, 2008 (07.10.08)	Date of mailing of the international search report 21 October, 2008 (21.10.08)
Name and mailing address of the ISA/ Japanese Patent Office Facsimile No.	Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2008/002497

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 08-229039 A (Toshiba Medical Systems Engineering Co., Ltd.), 10 September, 1996 (10.09.96), Full text; all drawings (Family: none)	1-6
A	JP 2001-100800 A (Toshiba Corp.), 13 April, 2001 (13.04.01), Full text; all drawings (Family: none)	1-6
A	JP 05-273326 A (NEC Corp.), 22 October, 1993 (22.10.93), Full text; all drawings (Family: none)	1-6
A	JP 2002-062348 A (Sony Corp.), 28 February, 2002 (28.02.02), Full text; all drawings & US 2002/0048376 A1	1-6
A	Bhiksha Raj, et al., Missing-Feature Approaches in Speech Recognition, IEEE Signal Processing Magazine, 2005, pp.101-116	1-6

国際調査報告

国際出願番号　ＰＣＴ／ＪＰ２００８／００２４９７

Ａ．　発明の属する分野の分類（国際特許分類（ＩＰＣ））

Int.Cl.　G01S3/808(2006.01)i

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.Cl.　G01S 1/72-1/82, G01S3/80-3/86, G01S5/18-5/30, G01S 7/52-7/66, G01S 15/00-15/96

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	１９７１－２００８年
日本国実用新案登録公報	１９９６－２００８年
日本国登録実用新案公報	１９９４－２００８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）

IEEE

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
A	JP 09-159744 A（防衛庁技術研究本部長）1997.06.20, 全文, 全図（ファミリーなし）	1-6
A	JP 2002-008189 A（松下電器産業株式会社）2002.01.11, 全文, 全図 & US 2002/0011939 A1 & EP 1168274 A1	1-6

☑　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	国際調査報告の発送日
０７．１０．２００８	２１．１０．２００８

国際調査機関の名称及びあて先	特許庁審査官（権限のある職員）	２Ｓ	３７１０
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	川瀬　徹也 電話番号　０３－３５８１－１１０１　内線　３２５８		

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
A	JP 08-229039 A (東芝医用システムエンジニアリング株式会社) 1996.09.10, 全文, 全図 (ファミリーなし)	1-6
A	JP 2001-100800 A (株式会社東芝) 2001.04.13, 全文, 全図 (ファミリーなし)	1-6
A	JP 05-273326 A (日本電気株式会社) 1993.10.22, 全文, 全図 (ファミリーなし)	1-6
A	JP 2002-062348 A (ソニー株式会社) 2002.02.28, 全文, 全図 & US 2002/0048376 A1	1-6
A	Bhiksha Raj, et al., Missing-Feature Approaches in Speech Recognition, IEEE Signal Processing Magazine, 2005, pp.101-116	1-6