

(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
2012 年 11 月 22 日(22.11.2012)



(10) 国際公開番号
WO 2012/157788 A1

- (51) 国際特許分類:
H04R 3/02 (2006.01) *H04B 3/23* (2006.01)
G10L 21/02 (2006.01)
- (21) 国際出願番号: PCT/JP2012/063408
- (22) 国際出願日: 2012 年 5 月 18 日(18.05.2012)
- (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 2011-112078 2011 年 5 月 19 日(19.05.2011) JP
- (71) 出願人(米国を除く全ての指定国について): 日本電気株式会社(NEC CORPORATION) [JP/JP]; 〒1088001 東京都港区芝五丁目 7 番 1 号 Tokyo (JP).
- (72) 発明者; および
- (75) 発明者/出願人(米国についてのみ): 宝珠山 治 (HOUSHUYAMA, Osamu) [JP/JP]; 〒1088001 東京都港区芝五丁目 7 番 1 号日本電気株式会社内 Tokyo (JP).
- (74) 代理人: 下坂 直樹(SHIMOSAKA, Naoki); 〒1088001 東京都港区芝五丁目 7 番 1 号日本電気株式会社内 Tokyo (JP).

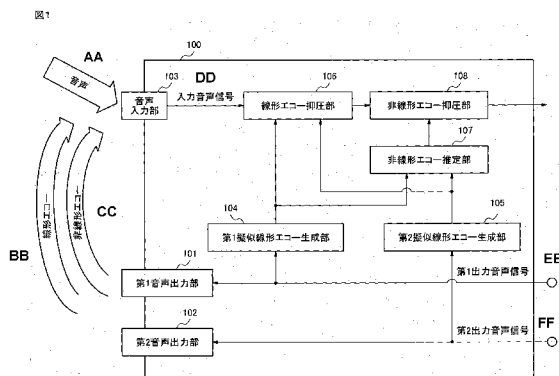
- (81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, 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, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, 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(条約第 21 条(3))

(54) Title: AUDIO PROCESSING DEVICE, AUDIO PROCESSING METHOD, AND RECORDING MEDIUM ON WHICH AUDIO PROCESSING PROGRAM IS RECORDED

(54) 発明の名称: 音声処理装置、音声処理方法および音声処理プログラムを記録した記録媒体



- 101 First audio output unit
102 Second audio output unit
103 Audio input unit
104 First artificial-linear-echo generation unit
105 Second artificial-linear-echo generation unit
106 Linear-echo suppression unit
107 Nonlinear-echo estimation unit
108 Nonlinear-echo suppression unit
AA Audio
BB Linear echo
CC Nonlinear echo
DD Input audio signal
EE First output audio signal
FF Second output audio signal

(57) Abstract: The present invention provides an audio processing device that appropriately suppresses echoes in stereo audio output. Said audio processing device is provided with the following: a means that generates first and second artificial linear echo signals corresponding to estimated echoes caused by feedback in first and second audio signals received by an audio input means; a means that uses said first and second artificial linear echo signals to suppress a linear echo signal in an input audio signal; a means that uses the first and second artificial linear echo signals to estimate a nonlinear echo signal; and a means that suppresses said nonlinear echo signal.

(57) 要約: 本発明は、ステレオ音声出力で発生したエコーを適正に抑圧する音声処理装置を提供する。その音声処理装置は、音声入力手段に対する第1音声および第2音声の回り込みにより発生したと推定される、第1擬似線形エコー信号および第2擬似線形エコー信号を生成する手段と、その第1擬似線形エコー信号およびその第2擬似線形エコー信号に基づいて、入力音声信号に混在する線形エコー信号を抑圧する手段と、その第1擬似線形エコー信号およびその第2擬似線形エコー信号に基づいて、非線形エコー信号を推定する手段と、その非線形エコー信号を抑圧する手段と、を備える。

WO 2012/157788 A1

明細書

発明の名称

音声処理装置、音声処理方法および音声処理プログラムを記録した記録媒体

5 技術分野

本発明は、音声中のエコーを抑圧する技術に関する。

背景技術

上記技術分野において、特許文献1に示されているように、エコーを抑圧する技術が知られている。この技術は、適応フィルタを用いて出力音声信号(遠端信号)から擬似線形エコー信号を生成し、入力音声信号中の線形エコー成分を抑圧した上で、さらに非線形エコー成分を抑圧する技術である。特に、入力音声信号に混在する非線形エコー信号を、擬似線形エコー信号を用いて推定することにより、比較的クリアに入力音声信号から近端音声信号を抽出している。

[特許文献]

15 [特許文献1] 再公表WO09-051197号公報

発明の概要

[発明が解決しようとする課題]

しかしながら、特許文献1に記載の技術では、ステレオ音声出力で発生したエコーを適正に抑圧することはできなかった。

その理由は、特許文献1記載のエコー抑圧装置が、入力音声信号に対する出力音声信号(特許文献1では遠端信号)が、2以上ある場合を想定していないからである。

本発明の目的は、上述の課題を解決する技術を提供することにある。

[課題を解決するための手段]

25 本発明の一態様における装置は、

第1出力音声信号に基づいて第1音声を出力する第1音声出力手段と、

第2出力音声信号に基づいて第2音声を出力する第2音声出力手段と、

音声を入力して入力音声信号を出力する音声入力手段と、

前記音声入力手段に対する前記第1音声の回り込みにより発生したと推定される第1擬似線形エコー信号を、前記第1出力音声信号から生成し、出力する第1擬似線形

エコー生成手段と、

前記音声入力手段に対する前記第2音声の回り込みにより発生したと推定される第2擬似線形エコー信号を、前記第2出力音声信号から生成し、出力する第2擬似線形エコー生成手段と、

- 5 前記第1擬似線形エコー生成手段および前記第2擬似線形エコー生成手段の出力に基づいて、前記入力音声信号に混在する線形エコー信号を抑圧した信号を生成し、出力する線形エコー抑圧手段と、

前記第1擬似線形エコー信号および前記第2擬似線形エコー信号に基づいて、非線形エコー信号を推定する非線形エコー推定手段と、

- 10 前記非線形エコー推定手段によって推定された非線形エコー信号に基づいて、前記線形エコー抑圧手段が出力した信号を抑圧する非線形エコー抑圧手段と、

を備えたことを特徴とする。

本発明の一態様における方法は、

- 第1出力音声信号および第2出力音声信号に基づいて2つの音声出力手段から出力された第1音声および第2音声を音声入力手段で入力して、入力音声信号を出力する音声入力ステップと、
- 15

前記音声入力手段に対する前記第1音声の回り込みにより発生したと推定される第1擬似線形エコー信号を、前記第1出力音声信号から生成し、出力する第1擬似線形エコー生成ステップと、

- 20 前記音声入力手段に対する前記第2音声の回り込みにより発生したと推定される第2擬似線形エコー信号を、前記第2出力音声信号から生成し、出力する第2擬似線形エコー生成ステップと、

- 前記第1擬似線形エコー信号および前記第2擬似線形エコー信号の出力に基づいて、前記入力音声信号に混在する線形エコー信号を抑圧した信号を生成し、出力する線形エコー抑圧ステップと、
- 25

前記第1擬似線形エコー信号および前記第2擬似線形エコー信号に基づいて、非線形エコー信号を推定する非線形エコー推定ステップと、

- 前記非線形エコー推定ステップにおいて推定された非線形エコー信号に基づいて、前記線形エコー抑圧ステップにおいて出力された信号を抑圧する非線形エコー抑圧ステップと、
- 30

を含むことを特徴とする。

本発明の一態様における不揮発性媒体に記録されたプログラムは、

第1出力音声信号および第2出力音声信号に基づいて2つの音声出力手段から出力された第1音声および第2音声を音声入力手段で入力して、入力音声信号を出力する音声入力ステップと、

前記音声入力手段に対する前記第1音声の回り込みにより発生したと推定される第1擬似線形エコー信号を、前記第1出力音声信号から生成し、出力する第1擬似線形エコー生成ステップと、

前記音声入力手段に対する前記第2音声の回り込みにより発生したと推定される第2擬似線形エコー信号を、前記第2出力音声信号から生成し、出力する第2擬似線形エコー生成ステップと、

前記第1擬似線形エコー信号および前記第2擬似線形エコー信号に基づいて、前記入力音声信号に混在する線形エコー信号を抑圧した信号を生成し、出力する線形エコー抑圧ステップと、

前記第1擬似線形エコー信号および前記第2擬似線形エコー信号に基づいて、非線形エコー信号を推定する非線形エコー推定ステップと、

前記非線形エコー推定ステップにおいて推定された非線形エコー信号に基づいて、前記線形エコー抑圧ステップにおいて出力された信号を抑圧する非線形エコー抑圧ステップと、

をコンピュータに実行させることを特徴とする。

〔発明の効果〕

本発明によれば、ステレオ音声出力で発生したエコーを適正に抑圧することができる。

25 図面の簡単な説明

〔図1〕 図1は、本発明の第1実施形態に係る音声処理装置の構成を示すブロック図である。

〔図2〕 図2は、本発明の第2実施形態に係る音声処理装置の機能構成を示すブロック図である。

30 〔図3〕 図3は、本発明の第2実施形態に係る音声処理装置の回路構成を示すブ

ロック図である。

〔図4〕 図4は、本発明の第3実施形態に係る音声処理装置の機能構成を示すブロック図である。

〔図5〕 図5は、本発明の第3実施形態に係る音声処理装置の回路構成を示すブロック図である。

〔図6〕 図6は、本発明の他の実施形態に係る情報処理装置の構成を示すブロック図である。

〔図7〕 図7は、本発明のプログラムを記録した記録媒体を示す図である。

発明を実施するための形態

以下に、図面を参照して、本発明の実施の形態について例示的に詳しく説明する。ただし、以下の実施の形態に記載されている構成要素はあくまで例示であり、本発明の技術範囲をそれらのみに限定する趣旨のものではない。

(第1実施形態)

本発明の第1実施形態としての音声処理装置100について、図1を用いて説明する。音声処理装置100は、2つの音声出力部から出力される音声に起因して、発生する非線形エコー信号を抑圧する装置である。

図1に示すように、音声処理装置100は、第1音声出力部101と第2音声出力部102と音声入力部103とを含む。さらに、音声処理装置100は、第1擬似線形エコー生成部104と第2擬似線形エコー生成部105と線形エコー抑圧部106と非線形エコー推定部107と非線形エコー抑圧部108とを含む。

これらのうち、第1音声出力部101および第2音声出力部102は、それぞれ第1出力音声信号および第2出力音声信号に応じた音声を出力する。

音声入力部103は、音声を入力する。

第1擬似線形エコー生成部104は、第1音声出力部101への第1出力音声信号に基づいて、第1擬似線形エコー信号を生成し、出力する。

第2擬似線形エコー生成部105は、第2音声出力部102への第2出力音声信号に基づいて、第2擬似線形エコー信号を生成し、出力する。

線形エコー抑圧部106は、第1擬似線形エコー信号および第2擬似線形エコー信号に基づいて、入力音声信号に混在する線形エコー信号を抑圧し、出力する。

非線形エコー推定部107は、第1擬似線形エコー信号および第2擬似線形エコー信

号に基づいて、非線形エコー信号を推定し、出力する。

非線形エコー抑圧部108は、非線形エコー信号を推定した結果に基づいて、線形エコー信号が抑圧された入力音声信号に混在する、非線形エコー信号を抑圧し、出力する。

- 5 以上の構成により、2つの音声入力手段を有する装置、即ちステレオ音声出力、によって発生したエコーを、適正に抑圧できる。

その理由は、以下のような構成を含むからである。即ち、第一に、第1擬似線形エコー生成部104および第2擬似線形エコー生成部105それぞれが、第1出力音声信号および第2出力音声信号それぞれに基づいて、第1擬似線形エコー信号および第2擬似線形エコー信号を生成し、出力する。第二に、線形エコー抑圧部106が、第1
10 擬似線形エコー信号および第2擬似線形エコー信号に基づいて、入力音声信号に混在する線形エコー信号を抑圧する。第三に、非線形エコー推定部107が、第1擬似線形エコー信号および第2擬似線形エコー信号に基づいて、非線形エコー信号を推定するし、非線形エコー抑圧部108が、非線形エコー信号を抑圧し、出力する。

- 15 (第2実施形態)

次に本発明の第2実施形態に係る音声処理装置200について、図2を用いて説明する。図2は、本実施形態に係る音声処理装置200の構成を説明するための図である。

図2に示すように、音声処理装置200は、音声入力部としてのマイクロフォン203と、
20 第1および第2音声出力部としてのスピーカ201および202と、を含む。スピーカ201および202は、それぞれ第1出力信号 $x_R(k)$ および第2出力信号 $x_L(k)$ に応じた音声を出力する。例えば、第1出力信号 $x_R(k)$ および第2出力信号 $x_L(k)$ は、ステレオ音声の信号である。この場合、スピーカ201および202は、ステレオ音声を出力する。

また、音声処理装置200は、適応フィルタ214、適応フィルタ224および加算部205
25 を含む。適応フィルタ214および224は、それぞれ第1出力信号 $x_R(k)$ および第2出力信号 $x_L(k)$ を入力して、擬似線形エコー信号を生成し、出力する。加算部205は、適応フィルタ214および適応フィルタ224それぞれから出力された擬似線形エコー信号を加算し、合成擬似線形エコー信号として出力する。

音声処理装置200は、さらに、線形エコーキャンセラ206、非線形エコー推定部207
30 、フロアリング部208、および非線形エコーサプレッサ209を備えている。加算部205

で生成された合成擬似線形エコー信号は、線形エコーキャンセラ206と非線形エコー推定部207の両方に供給される。

これらのうち、線形エコーキャンセラ206は、加算部205で合成された擬似線形エコー信号を、混在信号 $P(k)$ から減算し、出力する。一方、非線形エコー推定部207は、
5 加算部205で合成された擬似線形エコー信号に基づいて、非線形エコー信号を推定する。そして、フロアリング部208は、非線形エコー推定部207が推定した非線形エコー信号を、フロアリングし、フロアリング結果を出力する。非線形エコーサプレッサ209は、フロアリング結果に基づいて、線形エコーキャンセラ206の出力信号から、利得制御によって非線形エコー信号を抑圧し、出力する。

10 以上の構成は、2つのスピーカによるエコーの影響を、1つのスピーカによる線形エコーからの影響と見なして抑圧を行うという新しい発想に基づくものであり、非常にシンプルな構成で、2つのスピーカによるエコーを抑圧することができる。

次に、図3を用いて音声処理装置200の回路構成について説明する。図3は、音声処理装置200のより具体的な回路構成を示す図である。

15 図2でも説明したように、適応フィルタ214および適応フィルタ224それぞれは、第1出力信号 $x_R(k)$ および第2出力信号 $x_L(k)$ を入力して、擬似線形エコー信号を生成する。適応フィルタについての詳しい説明は米国公開公報2010-0260352A1号公報に開示されているので、ここでは省略する。

加算部205は、生成された擬似線形エコー信号を加算して合成擬似線形エコー信号を生成する。
20

線形エコーキャンセラ206として減算器は、マイクロフォン203によって出力された入力音声信号から合成擬似線形エコー信号を減算して、残差信号 $d(k)$ を生成し、出力する。

残差信号 $d(k)$ は、高速フーリエ変換部(Fast Fourier Transform:FFT)301に
25 入力され、合成擬似線形エコー信号 $y(k)$ は、高速フーリエ変換部302に入力される。

音声処理装置200は、高速フーリエ変換部301、高速フーリエ変換部302、非線形エコー推定部207と、フロアリング部208と、非線形エコーサプレッサ209と、逆高速フーリエ変換部(Inverse Fast Fourier Transform:IFFT)306と、を更に備える。

高速フーリエ変換部301および302それぞれは、残差信号 $d(k)$ および擬似線形エ
30 コー信号 $y(k)$ それぞれを周波数スペクトルに変換する。

非線形エコー推定部207と、フロアリング部208と、非線形エコーサプレッサ209とは周波数成分ごとに用意されている。

逆高速フーリエ変換部306は、周波数成分ごとに導き出された振幅スペクトルを対応する位相と統合して逆高速フーリエ変換し、時間領域の出力信号 $z_i(k)$ に再合成する。尚、時間領域の出力信号 $z_i(k)$ は、つまり、通話相手に送る音声波形の信号である。

線形エコー信号と非線形エコー信号とはまったく違う波形であるが、周波数ごとにスペクトル振幅を見ると、擬似線形エコー信号が大きい時は非線形エコー信号も大きくなる傾向、すなわち振幅の相関がある。つまり、擬似線形エコー信号に基づいて、非線形エコー信号の量を推定することができる。

そこで、非線形エコー推定部207は、推定した非線形エコー信号の量に基づいて、所望の音声信号のスペクトル振幅を推定する。推定された音声信号のスペクトル振幅には誤差があるが、その推定誤差が主観的に不快にならないようフロアリング部208でフロアリング処理を加えている。

例えば、音声信号の推定スペクトル振幅が過剰に小さく、背景雑音のスペクトル振幅を下回る場合、エコーの有無で信号レベルが変動し、違和感を生じさせる。その対策としてフロアリング部208では、背景雑音レベルを推定して、推定スペクトル振幅の下限とすることにより、レベル変動を低減する。

一方、推定誤差により推定スペクトル振幅にエコーが大きく残留してしまった場合、残留したエコーは断続的かつ急激に変化して、ミュージカルノイズと呼ばれる、人工的な付加音となる。その対策として非線形エコーサプレッサ209は、エコーを消去するために、推定した非線形エコー信号を減算するのではなく、減算された程度の振幅になるように利得を乗じるスペクトル利得計算部として機能する。利得の急激な変化を防止する平滑化を行うことにより、残留エコーの断続的変化を抑えることができる。

以下、非線形エコー推定部207、フロアリング部208、非線形エコーサプレッサ209の内部構成について数式を用いて説明する。

高速フーリエ変換部301に入力される残差信号 $d(k)$ は、近端信号 $s(k)$ と、残留非線形エコー信号 $q(k)$ の和である。

$$d(k) = s(k) + q(k) \quad \dots(1)$$

適応フィルタ214、適応フィルタ224および減算器(線形エコーキャンセラ206)によ

って線形エコーがほぼ完全に除去されていると仮定して、非線形成分のみを周波数領域で考える。高速フーリエ変換部301および302によって、式(1)は周波数領域に変換され、以下の式となる。

$$D(m) = S(m) + Q(m) \quad \dots(2)$$

- 5 ここでmはフレーム番号、ベクトルD(m)、S(m)およびQ(m)それぞれは、d(k)、s(k)および q(k)それぞれを、周波数領域に変換した表現である。各周波数を独立に考えて式(2)を変形すると、i番目の周波数では、以下の式となる。

$$S_i(m) = D_i(m) - Q_i(m) \quad \dots(3)$$

- 10 適応フィルタ214、適応フィルタ224および減算器(線形エコーキャンセラ206)は相関除去を行うため、 $D_i(m)$ と $Y_i(m)$ の間にはほとんど相関はない。したがって、減算器276は、以下のように $\overline{|S_i(m)|^2}$ を算出する。

$$\overline{|S_i(m)|^2} = \overline{|D_i(m)|^2} - \overline{|Q_i(m)|^2} \quad \dots(4)$$

- 15 $\overline{|D_i(m)|^2}$ は、絶対値化回路271と平均化回路273とを用いて $D_i(m)$ から導かれる。一方、非線形エコー信号 $|Q_i(m)|$ は、回帰係数 a_i と平均エコーレプリカ $\overline{|Y_i(m)|}$ の積として以下の様にモデル化できる。

$$|Q_i(m)| \cong a_i \cdot \overline{|Y_i(m)|} \quad \dots(5)$$

- 20 そこで、絶対値化回路272と平均化回路274とが、 $Y_i(m)$ から平均エコーレプリカ $\overline{|Y_i(m)|}$ を導き、更に積算部275が、回帰係数 a_i を乗算する。ここで回帰係数 a_i は、 $|Q_i(m)|$ と $|Y_i(m)|$ との相関を示す回帰係数である。このモデルは、 $|Q_i(m)|$ と $|Y_i(m)|$ との間に有意な相関があるという実験結果に基づいている。

式(3)は、ノイズ抑圧において広く用いられている加法型のモデルである。図3のスペクトル整形では、ノイズ抑圧において、不快なミュージカルノイズを生じにくい、スペクトル乗算型の構成をとる。スペクトル乗算を用いて、出力信号の振幅 $|Z_i(m)|$ を、スペクトルゲイン $G_i(m)$ と残差信号 $|D_i(m)|$ の積として得る。

$$25 \quad |Z_i(m)| = G_i(m) \cdot |D_i(m)| \quad \dots(6)$$

式(6)の平方根をとり、式(3)の二乗平均をとって式(4)の $\overline{|Q_i(m)|^2}$ に $a_i^2 \cdot \overline{|Y_i(m)|^2}$ を代入することにより、以下のように $|S_i(m)|$ の推定値 $\overline{|S_i(m)|}$ を得るようにしてもよい。そのようにすることでより一層効果的に非線形エコー信号を抑圧することができる。

$$30 \quad \overline{|S_i(m)|} \cong \sqrt{\overline{|D_i(m)|^2} - a_i^2 \cdot \overline{|Y_i(m)|^2}} \quad \dots(7)$$

モデルが精緻ではないため、推定した振幅 $|\hat{S}_i(m)|$ には無視できない誤差がある。誤差が大きく、オーバーサブラクションがおきると、近端信号において、高域成分の減少、あるいは変調感を生じることになる。特に、近端信号が空調音のように定常である場合、変調感は不快である。この変調感を主観的に低減するために、フロアリング部208でスペクトル上のフロアリングを用いる。

フロアリング部208では、まず、平均化回路281が、近端信号 $D_i(m)$ の定常成分 $|N_i(m)|$ を推定する。次に、最大値選択回路282が、定常成分 $|N_i(m)|$ を下限とするフロアリングを行う。その結果、よりよい近端信号の振幅推定値 $|\hat{S}_i(m)|$ が得られる。その後、除算器291が、 $|\hat{S}_i(m)|$ と $|D_i(m)|$ の比を求める。さらに、平均化回路292が、その比を平均化して、スペクトルゲイン $G_i(m)$ を得る。

最後に数式(5)に示したように、積算器293が、スペクトルゲイン $G_i(m)$ と残差信号 $|D_i(m)|$ の積を求める。こうすることで、振幅 $|Z_i(m)|$ を出力信号として得ることができる。逆高速フーリエ変換部306は、振幅 $|Z_i(m)|$ を逆フーリエ変換し、非線形のエコーが効果的に抑圧された音声信号 $z_i(k)$ を出力する。

回帰係数 a_i は、それぞれ、スピーカから音声出力させた場合のマイクロフォン203の入力から推定することができる。再公表2009/051197に開示されているように、状況に応じて回帰係数を更新させてもよい。

以上の構成によれば、2つのスピーカ201および202による線形エコー信号および非線形エコー信号を効果的に抑圧することが可能となる。

その理由は、適応フィルタ214および適応フィルタ224の出力を合成した、合成擬似線形エコー信号に基づいて、線形エコーキャンセラ206、高速フーリエ変換部301、高速フーリエ変換部302、非線形エコー推定部207、フロアリング部208、非線形エコーサプレッサ209および逆高速フーリエ変換部306が、エコーの抑圧を行うようにしたからである。

また、以上の構成によれば、より効率的な回路設計とすることができる。

その理由は、2つのスピーカへの第1出力信号 $x_R(k)$ および第2出力信号 $x_L(k)$ について、線形エコーキャンセラ206、高速フーリエ変換部301、高速フーリエ変換部302、非線形エコー推定部207、フロアリング部208、非線形エコーサプレッサ209および逆高速フーリエ変換部306を共有化しているため、

(第3実施形態)

次に本発明の第3実施形態に係る音声処理装置400について、図4および図5を用いて説明する。図4は、本実施形態に係る音声処理装置400の機能構成を説明するための図である。本実施形態に係る音声処理装置400は、上記第2実施形態の音声処理装置200と比べると、非線形エコー推定部207に替えて、非線形エコー推定部417と非線形エコー推定部427とを含む点で異なる。非線形エコー推定部417は、第1擬似線形エコー信号から第1非線形エコー信号を推定する第1非線形エコー推定手段としての機能し、非線形エコー推定部427は第2擬似線形エコー信号から第2非線形エコー信号を推定する第2非線形エコー推定手段として機能する。その他の構成および動作は、第2実施形態と同様であるため、同じ構成および動作については同じ符号を付してその詳しい説明を省略する。

図5は、音声処理装置400の回路構成を示す図である。

音声処理装置400は、高速フーリエ変換部301、高速フーリエ変換部502および高速フーリエ変換部503を含む。また、音声処理装置400は、非線形エコー推定部507および非線形エコー推定部508と、フロアリング部208と、非線形エコーサプレッサ209と、逆高速フーリエ変換部306とを含む。

高速フーリエ変換部301は、残差信号 $d(k)$ を周波数スペクトル $D_i(m)$ に変換する。高速フーリエ変換部502および高速フーリエ変換部503は、それぞれ2つの擬似線形エコー信号 $y_1(k)$ 、 $y_2(k)$ を周波数スペクトル $Y_{i1}(m)$ 、 $Y_{i2}(m)$ に変換する。

非線形エコー推定部507と非線形エコー推定部508とフロアリング部208と非線形エコーサプレッサ209とは、周波数成分ごとに用意されている。

逆高速フーリエ変換部306は、周波数成分ごとに導き出された振幅スペクトルを対応する位相と統合して逆高速フーリエ変換し、時間領域の出力信号 $z_i(k)$ に再合成する。尚、時間領域の出力信号 $z_i(k)$ は、つまり、通話相手に送る音声波形の信号である。

非線形エコー推定部507および508は、それぞれ、推定した非線形エコー信号の量に基づいて、所望の音声信号のスペクトル振幅を推定する。

適応フィルタ214、適応フィルタ224および減算器（線形エコーキャンセラ206）は相関除去を行うため、 $D_i(m)$ と $Y_i(m)$ の間にはほとんど相関はない。したがって、減算器276で以下のように $|S_i(m)|^2$ が得られる。

$$|S_i(m)|^2 = |D_i(m)|^2 - |Q_{i1}(m)|^2 - |Q_{i2}(m)|^2$$

非線形エコー信号 $|Q_{i1}(m)|$ 、 $|Q_{i2}(m)|$ は、回帰係数 a_{i1} および a_{i2} それぞれと、平均エコーレプリカ $\overline{|Y_{i1}(m)|}$ および $\overline{|Y_{i2}(m)|}$ それぞれとの積として以下のようモデル化できる。

$$|Q_{i1}(m)| \doteq a_{i1} \cdot \overline{|Y_{i1}(m)|}$$

$$5 \quad |Q_{i2}(m)| \doteq a_{i2} \cdot \overline{|Y_{i2}(m)|}$$

そこで、絶対値化回路572と平均化回路574とが、 $Y_{i1}(m)$ から平均エコーレプリカ $\overline{|Y_{i1}(m)|}$ を導き、さらに積算部575が、回帰係数 a_{i1} を乗算する。また、絶対値化回路582と平均化回路584が、 $Y_{i2}(m)$ から平均エコーレプリカ $\overline{|Y_{i2}(m)|}$ を導き、さらに積算部585が、回帰係数 a_{i2} を乗算する。

- 10 一方、以下のように $|S_i(m)|$ の推定値 $\overline{|S_i(m)|}$ を得てもよい。そのようにすることでより一層効果的に非線形エコー信号を抑圧することができる。

$$\overline{|S_i(m)|} \doteq \sqrt{|\overline{D_i(m)}|^2 - a_{i1}^2 \cdot \overline{|Y_{i1}(m)|}^2 - a_{i2}^2 \cdot \overline{|Y_{i2}(m)|}^2}$$

- 変調感を主観的に低減するために、フロアリング部208でスペクトル上のフロアリングを行う。積算器293が、スペクトルゲイン $G_i(m)$ と残差信号 $|D_i(m)|$ の積を求めること
15 とで、振幅 $|Z_i(m)|$ を出力信号として出力する。逆高速フーリエ変換部306は、振幅 $|Z_i(m)|$ を逆フーリエ変換し、非線形のエコーが効果的に抑圧された音声信号 $z_i(k)$ を出力する。

- 回帰係数 a_{i1} および a_{i2} は、それぞれ、スピーカ201および202のいずれか一方からのみ音声を出力させた場合のマイクロフォン203の入力から別々に推定することができる。再公表2009/051197に開示されているように、状況に応じてこれらの回帰係数を更新させてもよい。
20

以上の構成によれば、第二実施形態と同様の効果を得ることができる。

その理由は、非線形エコー推定部207に替えて、非線形エコー推定部417と非線形エコー推定部427とを含むようにしたからである。

- 25 (他の実施形態)

以上、本発明の実施形態について詳述したが、それぞれの実施形態に含まれる別々の特徴を如何様に組み合わせたシステムまたは装置も、本発明の範疇に含まれる。

- また、本発明は、複数の機器から構成されるシステムに適用されてもよいし、単体の装置に適用されてもよい。さらに、本発明は、実施形態の機能を実現する情報処理プログラムが、システムあるいは装置に直接あるいは遠隔から供給される場合にも適用
30

可能である。

したがって、本発明の機能をコンピュータで実現するために、コンピュータにインストールされるプログラム、あるいはそのプログラムを格納した媒体、そのプログラムをダウンロードさせるWWW(World Wide Web)サーバも、本発明の範疇に含まれる。

- 5 以下、一例として、第2実施形態で説明した音声処理をソフトウェアで実現する場合に、コンピュータ600に設けられたCPU(Central Processing Unit)602で実行する処理の流れを、図6を用いて説明する。

まず、CPU602は、第1出力音声信号および第2出力音声信号に基づいて2つのスピーカ201および202それぞれから出力された第1音声および第2音声を、マイクロフ
10 オン203から入力して、入力音声信号を出力する(S601)。

CPU602は、マイクロフォン203に対する、スピーカ201からの音声の回り込みにより発生したと推定される第1擬似線形エコー信号を、第1出力音声信号から生成する(S603)。

CPU602は、マイクロフォン203に対する、スピーカ202からの音声の回り込みにより発生したと推定される第2擬似線形エコー信号を、第2出力音声信号から生成する(S605)。
15

CPU602は、第1擬似線形エコー信号および第2擬似線形エコー信号に基づいて、入力音声信号に混在する線形エコー信号を抑圧する(S607)。

CPU602は、第1擬似線形エコー信号および第2擬似線形エコー信号に基づいて、
20 非線形エコー信号を推定する(S609)。そして、推定された非線形エコー信号を抑圧する(S611)。

以上の処理により、第2実施形態と同様の効果を得ることができる。

尚、入力部601は、音声入力部103およびマイクロフォン203を含んでよい。出力部603は、第1音声出力部101、第2音声出力部102、スピーカ201およびスピーカ202を含んでよい。メモリ604は、情報を記憶する。CPU602は、各ステップの動作を実行する場合に、メモリ604に必要な情報を書き込み、またメモリ604から必要な情報を読み出す。
25

図7は、プログラムを記録(記憶)する、記録媒体(記憶媒体)707の例を示す図である。記録媒体707は、情報を非一時的に記憶する不揮発性記録媒体である。尚、記録媒体707は、情報を一時的に記憶する記録媒体であってもよい。記録媒体707は、
30

図6に示す動作をコンピュータ600(CPU602)に実行させるプログラム(ソフトウェア)を記録する。尚、記録媒体707は、さらに、任意のプログラムやデータを記録してよい。

上述のプログラム(ソフトウェア)のコードを記録した記録媒体707が、コンピュータ600に供給され、CPU602は、記録媒体707に格納されたプログラムのコードを読み出して実行するようにしてもよい。あるいは、CPU602は、記録媒体707に格納されたプログラムのコードを、メモリ604に格納するようにしてもよい。すなわち、本実施形態は、コンピュータ600(CPU602)が実行するプログラムを、一時的にまたは非一時的に、記憶する記録媒体707の実施形態を含む。

以上、実施形態を参照して本願発明を説明したが、本願発明は上記実施形態に限定されるものではない。本願発明の構成や詳細には、本願発明の範囲内で当業者が理解し得る様々な変更をすることができる。

この出願は、2011年5月19日に提出された日本出願特願2011-112078を基礎とする優先権を主張し、その開示の全てをここに取り込む。

15 符号の説明

- 100 音声処理装置
- 101 第1音声出力部
- 102 第2音声出力部
- 103 音声入力部
- 20 104 第1擬似線形エコー生成部
- 105 第2擬似線形エコー生成部
- 106 線形エコー抑圧部
- 107 非線形エコー推定部
- 108 非線形エコー抑圧部
- 25 200 音声処理装置
- 201 スピーカ
- 202 スピーカ
- 203 マイクロフォン
- 205 加算部
- 30 206 線形エコーキャンセラ

	207	非線形エコー推定部
	208	フロアリング部
	209	非線形エコーサプレッサ
	214	適応フィルタ
5	224	適応フィルタ
	271	絶対値化回路
	272	絶対値化回路
	273	平均化回路
	274	平均化回路
10	275	積算部
	276	減算器
	281	平均化回路
	282	最大値選択回路
	291	除算器
15	292	平均化回路
	293	積算器
	301	高速フーリエ変換部
	302	高速フーリエ変換部
	306	逆高速フーリエ変換部
20	400	音声処理装置
	417	非線形エコー推定部
	427	非線形エコー推定部
	502	高速フーリエ変換部
	503	高速フーリエ変換部
25	507	非線形エコー推定部
	508	非線形エコー推定部
	572	絶対値化回路
	574	平均化回路
	575	積算部
30	582	絶対値化回路

	584	平均化回路
	585	積算部
	600	コンピュータ
	602	CPU
5	707	記録媒体

請求の範囲

[請求項1]

第1出力音声信号に基づいて第1音声を出力する第1音声出力手段と、

第2出力音声信号に基づいて第2音声を出力する第2音声出力手段と、

5 音声を入力して入力音声信号を出力する音声入力手段と、

前記音声入力手段に対する前記第1音声の回り込みにより発生したと推定される第1擬似線形エコー信号を、前記第1出力音声信号から生成し、出力する第1擬似線形エコー生成手段と、

10 前記音声入力手段に対する前記第2音声の回り込みにより発生したと推定される第2擬似線形エコー信号を、前記第2出力音声信号から生成し、出力する第2擬似線形エコー生成手段と、

前記第1擬似線形エコー生成手段および前記第2擬似線形エコー生成手段の出力に基づいて、前記入力音声信号に混在する線形エコー信号を抑圧した信号を生成し、出力する線形エコー抑圧手段と、

15 前記第1擬似線形エコー信号および前記第2擬似線形エコー信号に基づいて、非線形エコー信号を推定する非線形エコー推定手段と、

前記非線形エコー推定手段によって推定された非線形エコー信号に基づいて、前記線形エコー抑圧手段が出力した信号を抑圧する非線形エコー抑圧手段と、

を備えたことを特徴とする音声処理装置。

20 [請求項2]

前記第1擬似線形エコー信号と前記第2擬似線形エコー信号とを加算する加算手段をさらに有することを特徴とする請求項1に記載の音声処理装置。

[請求項3]

25 前記加算手段での加算結果を、前記線形エコー抑圧手段と前記非線形エコー推定手段とに入力することを特徴とする請求項2に記載の音声処理装置。

[請求項4]

前記非線形エコー推定手段での推定結果にフロアリング処理を施すフロアリング手段をさらに有することを特徴とする請求項1乃至3のいずれか1項に記載の音声処理装置。

30 [請求項5]

前記非線形エコー抑圧手段は、

前記フロアリング手段でのフロアリング結果に基づいて前記非線形エコー信号を抑圧することを特徴とする請求項1乃至4のいずれか1項に記載の音声処理装置。

[請求項6]

5 前記非線形エコー推定手段は、

前記第1擬似線形エコー信号から第1非線形エコー信号を推定する第1非線形エコー推定手段と、

前記第2擬似線形エコー信号から第2非線形エコー信号を推定する第2非線形エコー推定手段と、

10 を含むことを特徴とする請求項1乃至5のいずれか1項に記載の音声処理装置。

[請求項7]

第1出力音声信号および第2出力音声信号に基づいて2つの音声出力手段から出力された第1音声および第2音声を音声入力手段で入力して、入力音声信号を出力する音声入力ステップと、

15 前記音声入力手段に対する前記第1音声の回り込みにより発生したと推定される第1擬似線形エコー信号を、前記第1出力音声信号から生成し、出力する第1擬似線形エコー生成ステップと、

前記音声入力手段に対する前記第2音声の回り込みにより発生したと推定される第2擬似線形エコー信号を、前記第2出力音声信号から生成し、出力する第2擬似線形

20 エコー生成ステップと、

前記第1擬似線形エコー信号および前記第2擬似線形エコー信号に基づいて、前記入力音声信号に混在する線形エコー信号を抑圧した信号を生成し、出力する線形エコー抑圧ステップと、

25 前記第1擬似線形エコー信号および前記第2擬似線形エコー信号に基づいて、非線形エコー信号を推定する非線形エコー推定ステップと、

前記非線形エコー推定ステップにおいて推定された非線形エコー信号に基づいて、前記線形エコー抑圧ステップにおいて出力された信号を抑圧する非線形エコー抑圧ステップと、

を含むことを特徴とする音声処理方法。

30 [請求項8]

第1出力音声信号および第2出力音声信号に基づいて2つの音声出力手段から出力された第1音声および第2音声を音声入力手段で入力して、入力音声信号を出力する音声入力ステップと、

5 前記音声入力手段に対する前記第1音声の回り込みにより発生したと推定される第1擬似線形エコー信号を、前記第1出力音声信号から生成し、出力する第1擬似線形エコー生成ステップと、

前記音声入力手段に対する前記第2音声の回り込みにより発生したと推定される第2擬似線形エコー信号を、前記第2出力音声信号から生成し、出力する第2擬似線形エコー生成ステップと、

10 前記第1擬似線形エコー信号および前記第2擬似線形エコー信号に基づいて、前記入力音声信号に混在する線形エコー信号を抑圧した信号を生成し、出力する線形エコー抑圧ステップと、

前記第1擬似線形エコー信号および前記第2擬似線形エコー信号に基づいて、非線形エコー信号を推定する非線形エコー推定ステップと、

15 前記非線形エコー推定ステップにおいて推定された非線形エコー信号に基づいて、前記線形エコー抑圧ステップにおいて出力された信号を抑圧する非線形エコー抑圧ステップと、

をコンピュータに実行させることを特徴とする音声処理プログラムを記録した不揮発性媒体。

図1

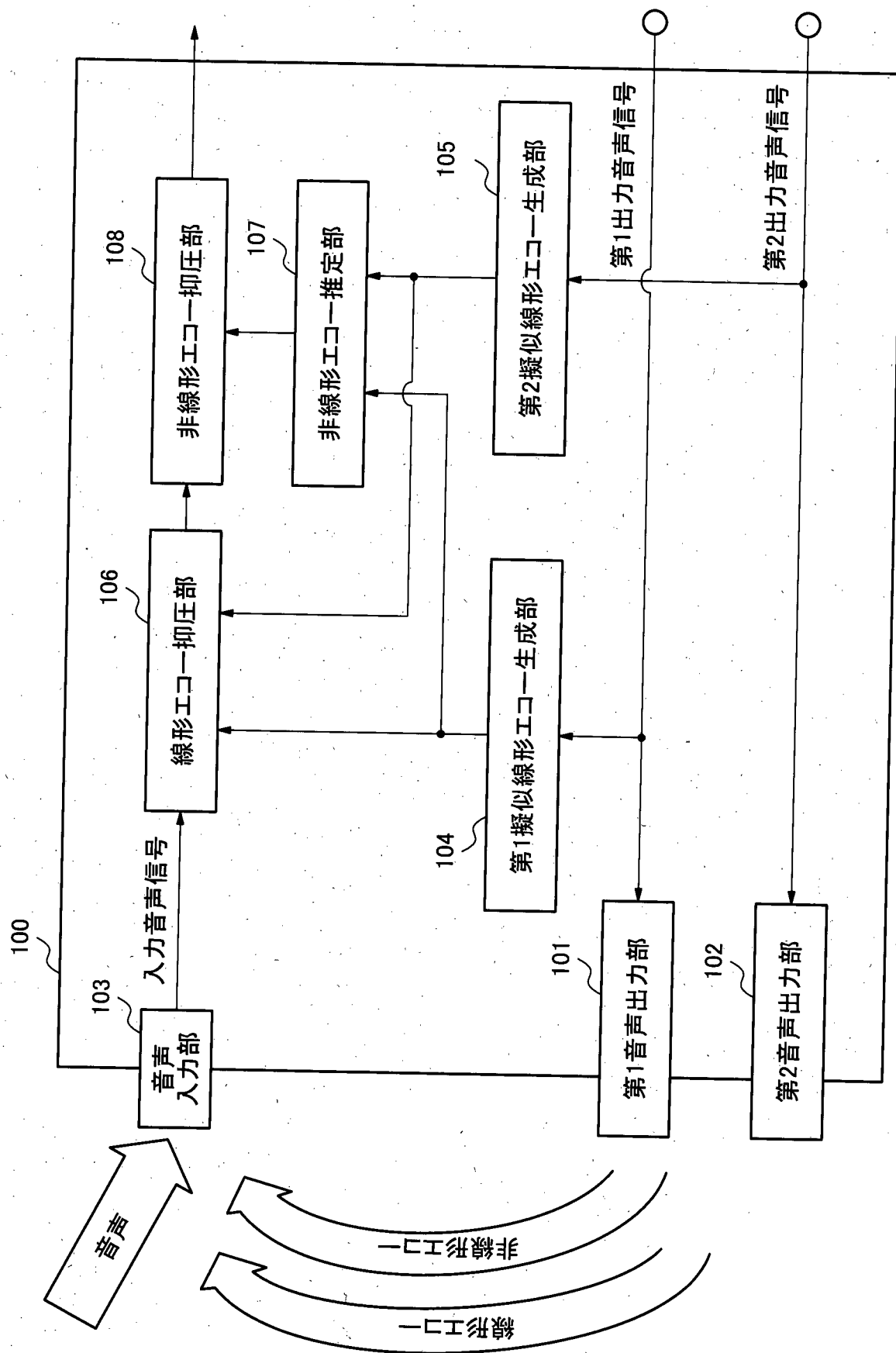


図2

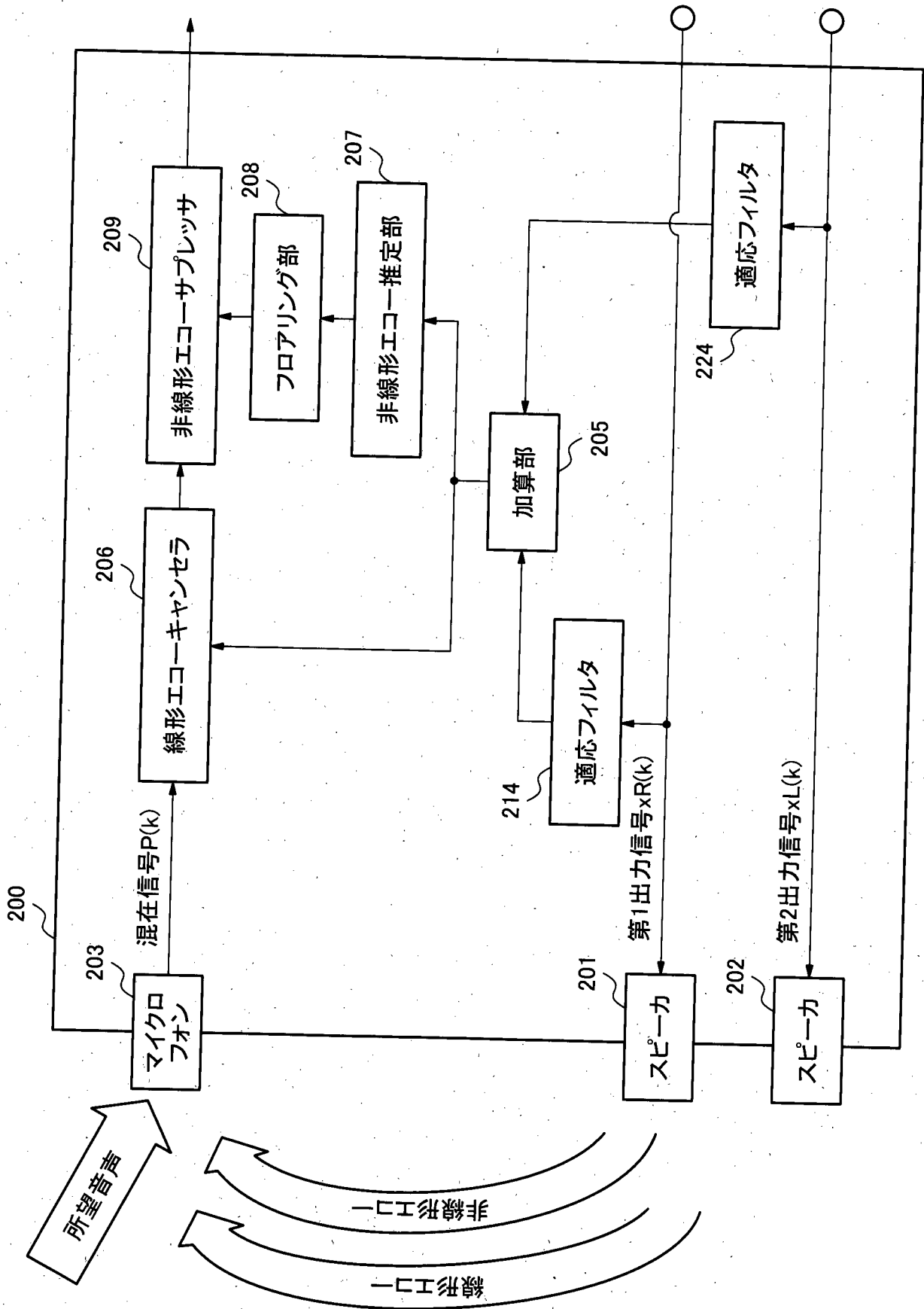


図3

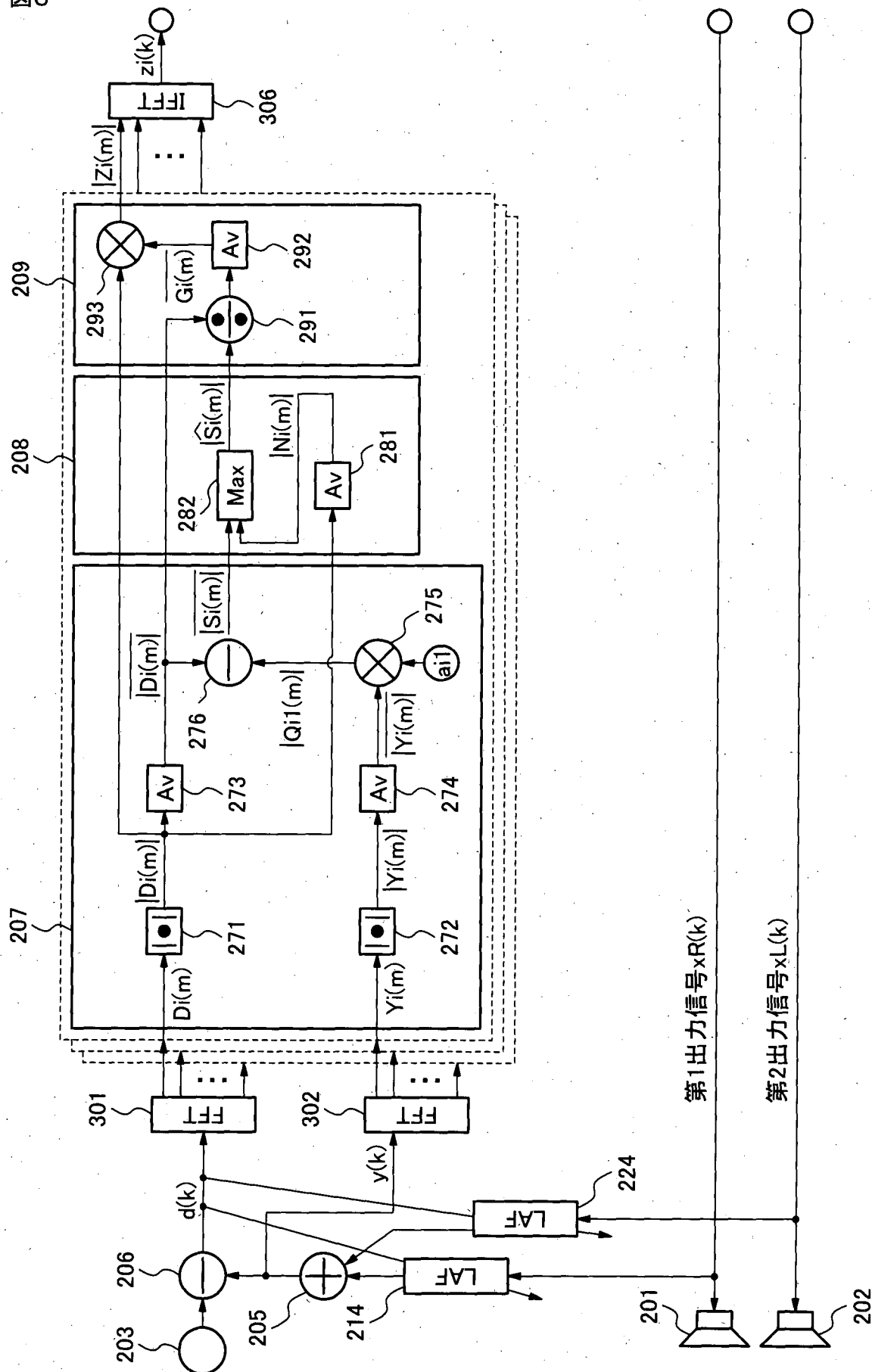


図4

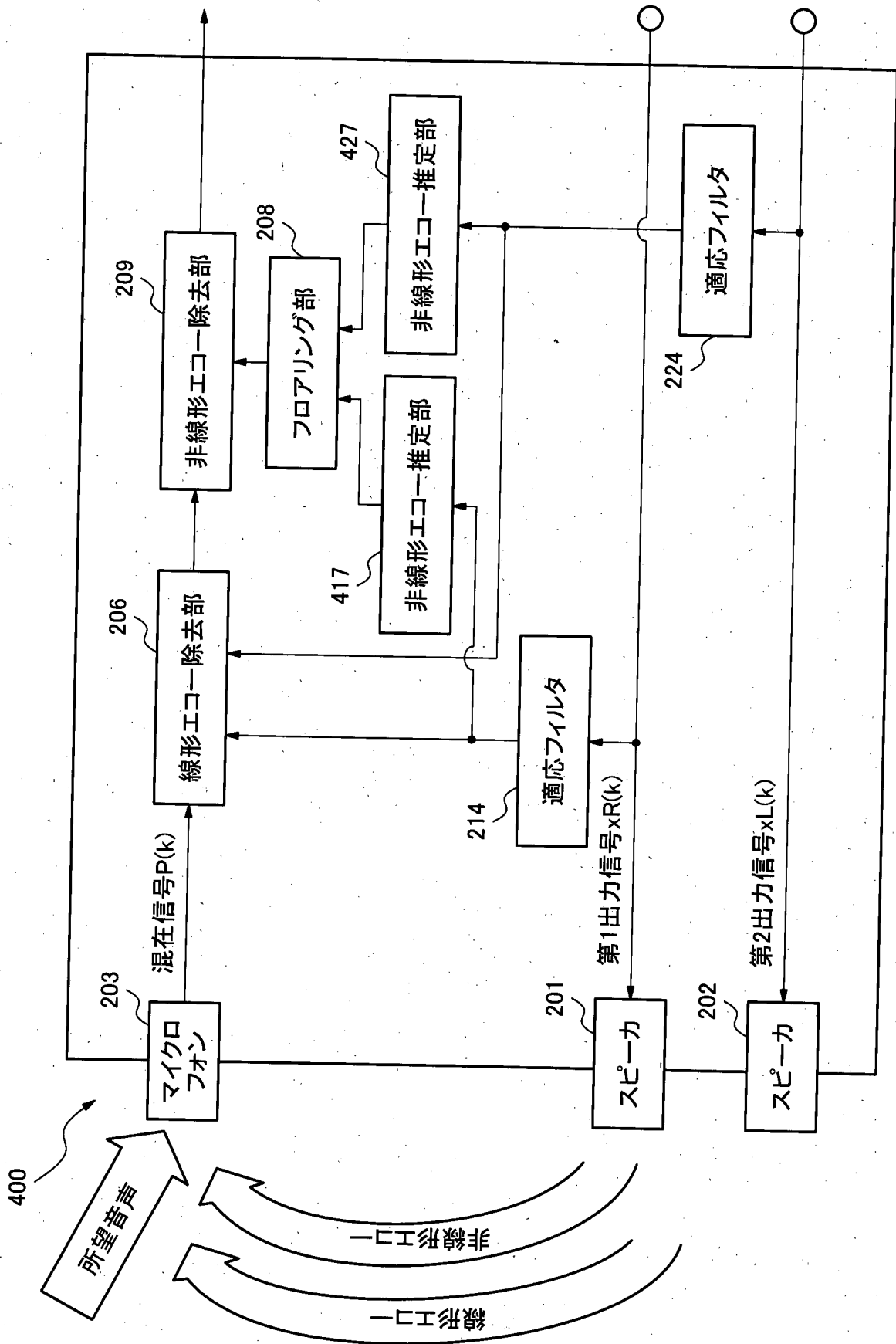


図5

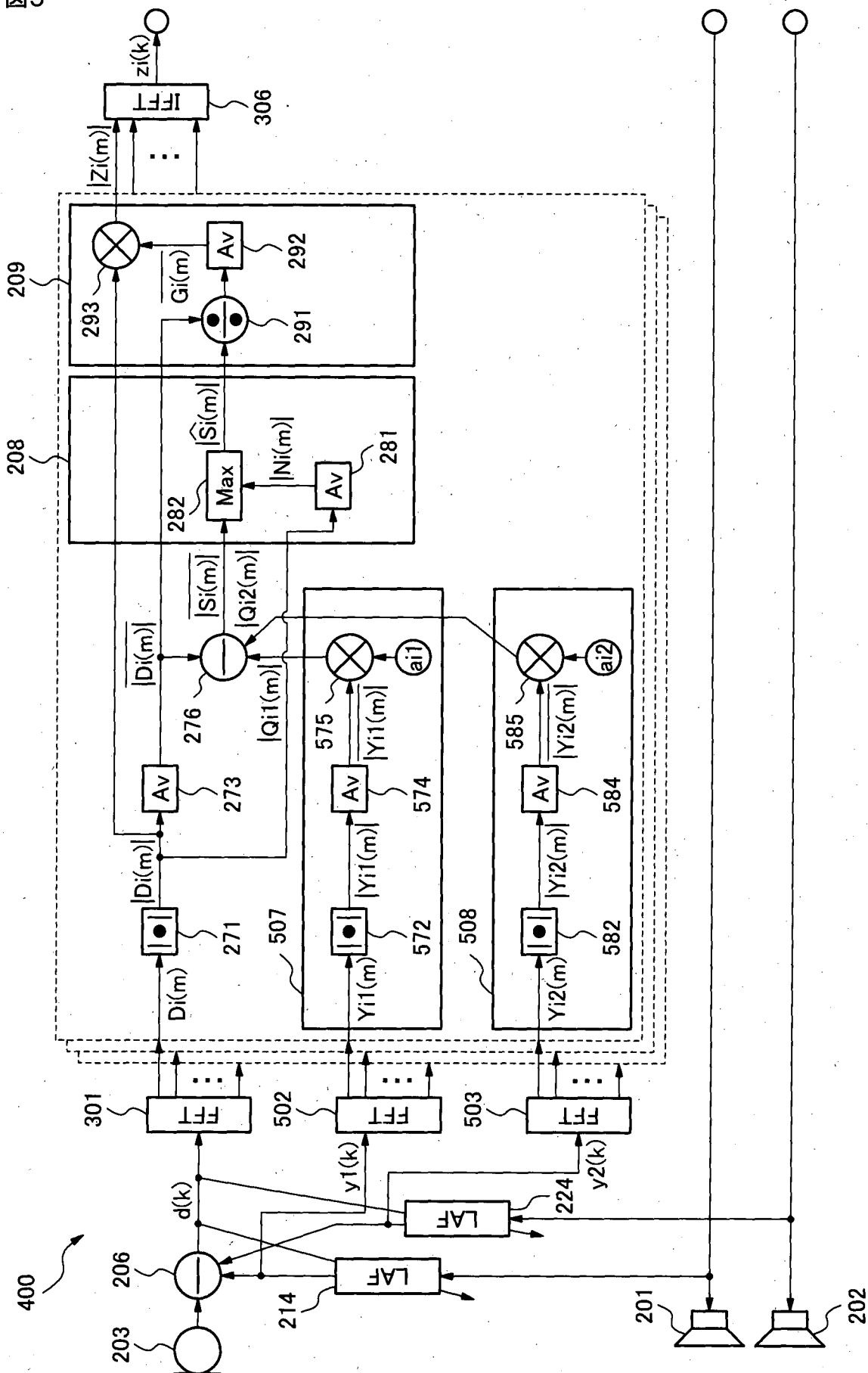
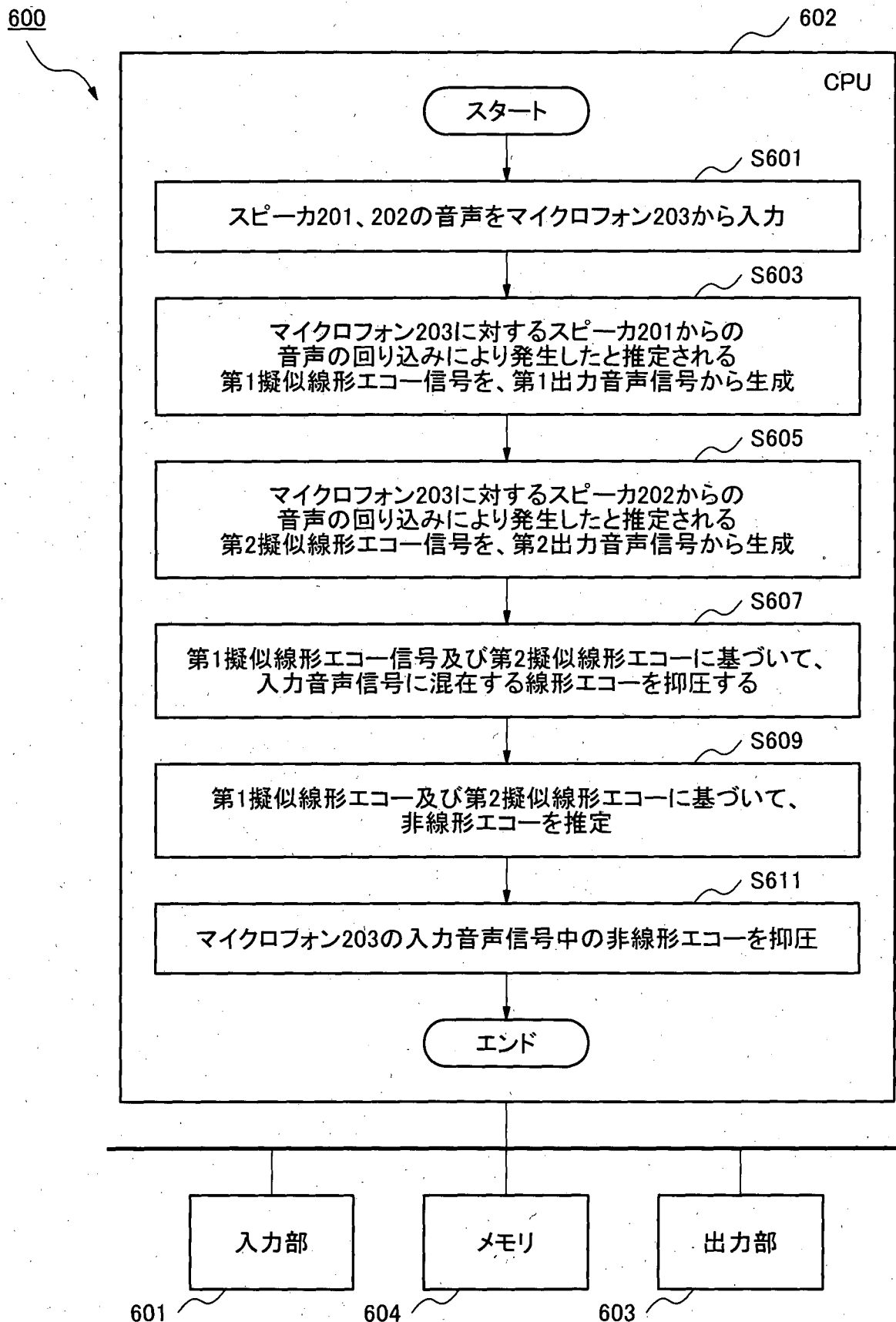
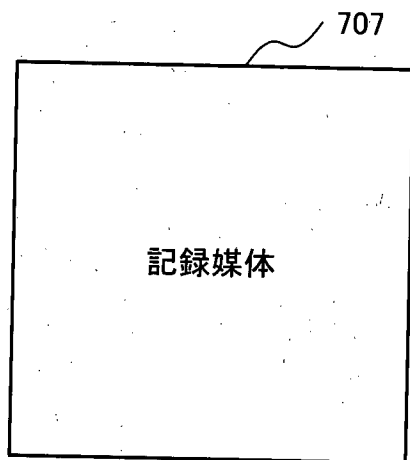


図6



7/7

図7



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2012/063408

A. CLASSIFICATION OF SUBJECT MATTER

H04R3/02(2006.01) i, G10L21/02(2006.01) i, H04B3/23(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)

H04R3/02, G10L21/02, H04B3/23

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	1996-2012
Kokai Jitsuyo Shinan Koho	1971-2012	Toroku Jitsuyo Shinan Koho	1994-2012

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y A	WO 2009/051197 A1 (NEC Corp.), 23 April 2009 (23.04.2009), paragraphs [0016] to [0019], [0023] to [0031]; fig. 1 to 4 & US 2010/0208908 A1	1-3, 7, 8 4-6
Y A	JP 2010-68213 A (Mitsubishi Electric Corp.), 25 March 2010 (25.03.2010), paragraphs [0002], [0003]; fig. 3 (Family: none)	1-3, 7, 8 4-6
A	JP 2010-220087 A (Yamaha Corp.), 30 September 2010 (30.09.2010), entire text; all drawings (Family: none)	1-8



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
29 June, 2012 (29.06.12)

Date of mailing of the international search report
10 July, 2012 (10.07.12)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

A. 発明の属する分野の分類（国際特許分類（I P C））

Int.Cl. H04R3/02(2006.01)i, G10L21/02(2006.01)i, H04B3/23(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（I P C））

Int.Cl. H04R3/02, G10L21/02, H04B3/23

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	1 9 2 2 - 1 9 9 6 年
日本国公開実用新案公報	1 9 7 1 - 2 0 1 2 年
日本国実用新案登録公報	1 9 9 6 - 2 0 1 2 年
日本国登録実用新案公報	1 9 9 4 - 2 0 1 2 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）

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 請求項の番号
Y A	WO 2009/051197 A1（日本電気株式会社）2009.04.23, [0016] - [0019], [0023] - [0031], 第1-4図 & US 2010/0208908 A1	1-3, 7, 8 4-6
Y A	JP 2010-68213 A（三菱電機株式会社）2010.03.25, 【0002】, 【0003】, 第3図（ファミリーなし）	1-3, 7, 8 4-6
A	JP 2010-220087 A（ヤマハ株式会社）2010.09.30, 全文, 全図（ファミリーなし）	1-8

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

2 9 . 0 6 . 2 0 1 2

国際調査報告の発送日

1 0 . 0 7 . 2 0 1 2

国際調査機関の名称及びあて先

日本国特許庁（I S A / J P）

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官（権限のある職員）

渡邊 正宏

電話番号 03-3581-1101 内線 3591

5 Z

4 5 4 6