

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2010-28653

(P2010-28653A)

(43) 公開日 平成22年2月4日(2010.2.4)

(51) Int.Cl.	F I	テーマコード (参考)
H04R 3/02 (2006.01)	H04R 3/02	5D018
H04B 3/23 (2006.01)	H04B 3/23	5D020
H04R 3/00 (2006.01)	H04R 3/00 320	5K027
H04R 1/40 (2006.01)	H04R 1/40 320A	5K046
H04M 1/60 (2006.01)	H04M 1/60	

審査請求 未請求 請求項の数 8 O L (全 16 頁)

(21) 出願番号	特願2008-189910 (P2008-189910)	(71) 出願人	000004226
(22) 出願日	平成20年7月23日 (2008.7.23)		日本電信電話株式会社
			東京都千代田区大手町二丁目3番1号
		(74) 代理人	100121706
			弁理士 中尾 直樹
		(74) 代理人	100066153
			弁理士 草野 卓
		(74) 代理人	100128705
			弁理士 中村 幸雄
		(72) 発明者	福井 勝宏
			東京都千代田区大手町二丁目3番1号 日
			本電信電話株式会社内
		(72) 発明者	中川 朗
			東京都千代田区大手町二丁目3番1号 日
			本電信電話株式会社内

最終頁に続く

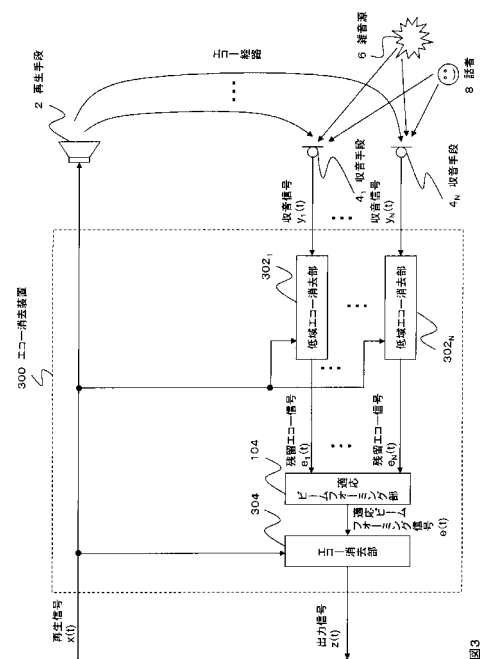
(54) 【発明の名称】 エコー消去装置、エコー消去方法、そのプログラム、記録媒体

(57) 【要約】

【課題】低計算コストで、適応ビームフォーミング処理により音響指向性の変動が生じたとしても、所望のエコー消去量に達する。

【解決手段】N個の收音手段で收音された收音信号に含まれるエコー信号を消去し、当該收音信号に含まれる雑音信号を抑圧して、出力信号として出力するものであり、收音手段毎に收音信号の、予め定められた基準値より低い周波数帯域に含まれるエコー信号の成分を消去し、残留エコー信号を出力し、N個の残留エコー信号から雑音信号を抑圧して、適応ビームフォーミング信号を出力し、適応ビームフォーミング信号から、エコー信号を消去することで出力信号を出力する。

【選択図】 図3



【特許請求の範囲】**【請求項 1】**

N 個の收音手段で收音された收音信号に含まれるエコー信号を消去し、当該收音信号に含まれる雑音信号を抑圧して、出力信号として出力するエコー消去装置において、

それぞれが前記收音信号の、予め定められた基準値より低い周波数帯域に含まれるエコー信号の成分を消去し、残留エコー信号を出力する N 個の低域エコー消去部と、

N 個の前記残留エコー信号から前記雑音信号を抑圧して、適応ビームフォーミング信号を出力する適応ビームフォーミング部と、

前記適応ビームフォーミング信号から、エコー信号を消去することで前記出力信号を出力するエコー消去部と、

を有するエコー消去装置。

10

【請求項 2】

請求項 1 記載のエコー消去装置であって、

前記低域エコー消去部は、

前記再生信号の低域周波数帯の信号（以下、「低域再生信号」という。）を出力するローパスフィルタ手段と、

前記收音信号を帯域分割することで、低域收音信号と高域收音信号に分けて出力する帯域分割手段と、

前記低域再生信号に適応フィルタを畳み込むことで、低域擬似エコー信号を出力する適応フィルタ手段と、

20

前記低域收音信号と前記低域擬似エコー信号との差をとることで、低域残留エコー信号を出力する減算手段と、

前記低域残留エコー信号と前記高域收音信号について帯域合成を行うことで、前記残留エコー信号を出力する帯域合成手段と、を有し、

前記適応フィルタは、過去の低域残留エコー信号と前記低域再生信号とを用いて更新されることを特徴とするエコー消去装置。

【請求項 3】

請求項 1 または 2 記載のエコー消去装置であって、

前記エコー消去部は、

前記基準値より高い周波数帯域の適応ビームフォーミング信号から、エコー信号を消去するものであることを特徴とするエコー消去装置。

30

【請求項 4】

N 個の收音手段で收音された收音信号に含まれるエコー信号を消去し、当該收音信号に含まれる雑音信号を抑圧して、出力信号として出力するエコー消去方法において、

收音手段毎に、前記收音信号の予め定められた基準値より低い周波数帯域に含まれるエコー信号の成分を消去し、残留エコー信号を出力する低域エコー消去過程と、

N 個の前記残留エコー信号から前記雑音信号を抑圧して、適応ビームフォーミング信号を出力する適応ビームフォーミング過程と、

前記適応ビームフォーミング信号から、エコー信号を消去することで前記出力信号を出力するエコー消去過程と、

40

を有するエコー消去方法。

【請求項 5】

請求項 4 記載のエコー消去方法であって、

前記低域エコー消去過程は、

前記再生信号の低域周波数帯の信号（以下、「低域再生信号」という。）を出力するローパスフィルタステップと、

前記收音信号を帯域分割することで、低域收音信号と高域收音信号に分けて出力する帯域分割ステップと、

前記低域再生信号に適応フィルタを畳み込むことで、低域擬似エコー信号を出力する適応フィルタステップと、

50

前記低域收音信号と前記低域擬似エコー信号との差をとることで、低域残留エコー信号を出力する減算ステップと、

前記低域残留エコー信号と前記高域收音信号について帯域合成を行うことで、前記残留エコー信号を出力する帯域合成ステップと、を有し、

前記適応フィルタは、過去の低域残留エコー信号と前記低域再生信号とを用いて更新されることを特徴とするエコー消去方法。

【請求項 6】

請求項 4 または 5 記載のエコー消去方法であって、

前記エコー消去過程は、

前記基準値より高い周波数帯域の適応ビームフォーミング信号から、エコー信号を消去することを特徴とするエコー消去方法。

【請求項 7】

請求項 1 ～ 3 に記載のエコー消去装置としてコンピュータを動作させるプログラム。

【請求項 8】

請求項 7 記載のプログラムを記録したコンピュータ読み取り可能な記録媒体。

10

20

【発明の詳細な説明】

【技術分野】

【0001】

この発明は、複数の收音手段で收音された收音信号に含まれるエコー信号を消去し、收音信号に含まれる雑音信号を抑圧して出力信号として出力するエコー消去装置、エコー消去方法、そのプログラム、記録媒体に関する。

【背景技術】

【0002】

従来の電話会議やテレビ会議に加えて、近年 PC（パーソナルコンピュータ）を用いたテレビ電話デスクトップ会議などハンズフリー拡声通話が増加している。ハンズフリー拡声通話においては、受信した相手の声は再生手段（例えば、スピーカ）から放出される一方、受話者の音声は收音手段（例えば、マイクロホン）で收音されることから、通話しながら議事録の筆記や PC の操作ができるという大きな利点がある。

【0003】

しかし、スピーカで放出された相手の声は、受話者に届くだけでなく、マイクロホンにも回り込んで收音され、音響エコーとして通話品質を劣化させるとともに、ハウリングの原因にもなる。そこで、このようなエコー信号の成分を消去し、ハウリングの発生を未然に防ぐために、ハンズフリー拡声通話においてエコー消去技術が導入されている。

【0004】

また、種々の環境雑音の中から利用者の音声を選択的に受音、追従し、雑音抑圧を行ったあと、高精度に話者の音声を收音する適応ビームフォーミング技術がある。この技術は特許文献 1 の收音方法及び收音装置「マイクロホンアレイによるエコー抑圧、方向別 AGC マイクロホンアレイ」に記載されている。この技術は、複数のマイクロホンを用いて音響的指向性を制御することにより、雑音源からの雑音信号を抑圧し、高精度に話者の音声を收音するものである。近年、この適応ビームフォーミング技術はエコー消去装置にも備えられている（非特許文献 1 参照）。

【0005】

図 1 に、適応ビームフォーミングを備えた従来のエコー消去装置 100 の機能構成例を示す。エコー消去装置 100 は、N（N は 1 以上の整数）個の收音手段（例えば、マイクロホン）に接続されている。エコー消去装置 100 は、N 個のエコー消去部 102_n（n

30

40

50

$= 1, \dots, N$) と適応ビームフォーミング部 104 とを有する。また、以下の説明では、再生手段 2 から放出された再生信号を $x(t)$ とし、再生信号 $x(t)$ が收音手段 4_n に回り込んだ信号をエコー信号 $a(t)$ とし、雑音源よりの雑音信号を $b(t)$ とし、話者 8 からの話者信号を $c(t)$ とする。エコー信号 $a(t)$ 、雑音信号 $b(t)$ 、話者信号 $c(t)$ からなる信号を收音信号とし、收音手段 4_n で收音された信号を收音信号を $y_n(t)$ とし、エコー信号、雑音信号、話者信号をそれぞれ $a_n(t)$ $b_n(t)$ $c_n(t)$ と示す。 t は離散時刻を示し、連続時間信号を一定のサンプル間隔 T でサンプリングしたことにより示されるものである。例えば、サンプリング周波数 $f_s = 48 \text{ kHz}$ (1 秒間に 48000 回) とする。サンプリング周波数 $f_s = 1/T$ である。また、エコー信号、雑音信号、話者信号が收音手段まで到達するまでは離散時刻ではなく、実時刻で表すべきであるが、説明簡略化のために、全て離散時刻 t で示す。エコー消去装置 100、以下で説明するエコー消去装置 200、300 においても離散時刻 t で表された信号であるデジタル信号を扱うので、入力されるアナログ信号を DA 変換器によりデジタル信号に変換し、出力するデジタル信号を AD 変換器によりアナログ信号に変換する必要があるのであるが、このことは当然のことであるので、図 1 ~ 図 3 において、AD 変換器、DA 変換器については図示しない。また、雑音源 6、話者 8 が存在しない場合は、收音信号には雑音信号、話者信号は含まれない。

10

20

30

40

50

【0006】

説明を図 1 に戻す。收音信号 $y_n(t)$ は、エコー消去部 102_n に入力される。エコー消去部 102_n は收音信号 $y_n(t)$ からエコー信号 $a_n(t)$ を抑圧する。そして、エコー信号が消去された信号である残留エコー信号 $e_n(t)$ が適応ビームフォーミング部 104 に入力される。

【0007】

適応ビームフォーミング部 104 は N 個の残留エコー信号 $e_n(t)$ ($= 1, \dots, N$) から雑音信号を抑圧し、出力信号 $z(t)$ として出力する。エコー消去部 102_n、適応ビームフォーミング部 104 の詳細な処理については、以下の「発明を実施するための最良の形態」で説明する。

【0008】

エコー消去装置 100 の構成は、 N 個のエコー消去部 102_n の後段に適応ビームフォーミング部 104 を設置している。従って、適応ビームフォーミング部 104 の適応ビームフォーミングによる音響的指向性の変動をエコー消去部 102_n のエコー経路モデルに含まないようにできる。

【0009】

次に、図 2 に、適応ビームフォーミングを備えた従来のエコー消去装置 200 の機能構成例を示す。エコー消去装置 200 の構成は、適応ビームフォーミング部 202 が前段に、エコー消去部 204 が後段に設けられる。このように構成することで、適応ビームフォーミング部 202 よりの出力信号が 1 つである場合には、エコー消去部 204 は、その出力信号に対して 1 回のエコー消去処理 (適応フィルタ処理) を行うだけでよく、低演算・低メモリ量の構成となっている。

【特許文献 1】特許第 4104626 号

【非特許文献 1】Kellermann, W., "Strategies for combining acoustic echo cancellation and adaptive beamforming microphone arrays," IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 1, 21-24, pp. 219-222, Apr. 1997.

【発明の開示】

【発明が解決しようとする課題】

【0010】

上記エコー消去装置 100 の構成であると、計算コストの高いエコー消去部 102_n を收音手段の数である N 個分、備える必要があるため、膨大な計算コストが必要であった。また、上記エコー消去装置 200 の構成であると、エコー消去部のエコー経路モデルにおいて、時変ビームフォーミングを考慮しなければならないため、適応ビームフォーミング

部 2 0 2 によって音響的指向性の変動が生じれば、エコー消去部 2 0 4 がエコー経路モデル（適応フィルタ）の再推定を行わなければならない、所望のエコー消去量に達しない場合がある。

【 0 0 1 1 】

この発明の目的は、（ 1 ）計算コストを少なくし、（ 2 ）エコー消去部の所望エコー消去量に達するまでの時間を短くすることで、適応ビームフォーミング処理による音響的指向性の変動に関わらず、所望のエコー消去量に達するエコー消去装置を提供することである。

【課題を解決するための手段】

【 0 0 1 2 】

この発明のエコー消去装置は、N 個の收音手段で收音された收音信号に含まれるエコー信号を消去し、当該收音信号に含まれる雑音信号を抑圧して、出力信号として出力するものである。そして当該エコー消去装置は、N 個の低域エコー消去部と、適応ビームフォーミング部と、エコー消去部と、を有する。低域エコー消去部は、收音信号の、予め定められた基準値より低い周波数帯域に含まれるエコー信号の成分を消去し、残留エコー信号を出力する。適応ビームフォーミング部は、N 個の残留エコー信号から雑音信号を抑圧して、適応ビームフォーミング信号を出力する。エコー消去部は、適応ビームフォーミング信号から、エコー信号を消去することで出力信号を出力する。

【発明の効果】

【 0 0 1 3 】

この発明では、パワーが低周波数帯域（以下、単に「低域」という。）に集中する話者音声の性質から低周波数帯域の所望エコー消去量が高周波数帯域（以下、単に「高域」という。）の所望エコー消去量より大きいことに着目する。そして、N 個の低域エコー消去部は、入力された收音信号の予め定められた基準値より低い低域についてのみエコーを消去するので、エコー消去装置 1 0 0 中の N 個のエコー消去部と比較して計算コストを削減できる。従って、エコー消去装置 1 0 0 の問題点を解決できる。

【 0 0 1 4 】

また、この発明のエコー消去装置の構成は、前段から N 個の低域エコー消去部、低域ビームフォーミング部、エコー消去部、の順番で構成されている。従って、低域エコー消去部のエコー経路モデルでは、適応ビームフォーミング部 1 0 4 の音響的指向性の変動による再推定を回避できる。また、エコー消去部については、低域エコー消去部によって、低域のエコーを消去したもの、つまり、大半のエコー信号の成分が消去された信号（適応ビームフォーミング信号）についてエコー信号の成分を消去する。従って、エコー消去部内で用いる全域残留エコー信号のパワーを最初から小さくでき、結果として、エコー消去部の所望エコー消去量に達するまでの時間を短くできる。よって、適応ビームフォーミング部による音響指向性変動が生じたとしても、エコー消去部は安定的に所望のエコー消去量を達成できる。また、上記非特許文献 1 には、エコー消去技術、適応ビームフォーミング技術のそれぞれの処理手順については記載されていない。

【発明を実施するための最良の形態】

【 0 0 1 5 】

以下に、発明を実施するための最良の形態を示す。なお、同じ機能を持つ構成部や同じ処理を行う過程には同じ番号を付し、重複説明を省略する。

【実施例 1】

【 0 0 1 6 】

図 3 にエコー消去装置 3 0 0 の機能構成例を示し、図 4 に処理フローを示す。エコー消去装置 3 0 0 は N 個の低域エコー消去部 3 0 2_n、適応ビームフォーミング部 1 0 4、エコー消去部 3 0 4 を有する。図 5 に低域エコー消去部 3 0 2_n の機能構成例を示し、図 6 に適応ビームフォーミング部 1 0 4 の機能構成例を示し、図 7 にエコー消去部 3 0 4 の機能構成例を示す。なお、全ての低域エコー消去部 3 0 2_n（ $n = 1, \dots, N$ ）の構成手段は全て同じであるので、図 5 記載の各構成手段の参照番号については添え字「n」を

10

20

30

40

50

省略している。また、エコー消去部 304 は、図 2 記載のエコー消去部 204 と同一である。

【0017】

図 5 記載のように、この実施例の低域エコー消去部 302 は、ローパスフィルタ手段 3014 と、ダウンサンプリング手段 3016 と、適応フィルタ手段 3018 と、帯域分割手段 3012 と、減算手段 3020 と、帯域合成手段 3022 と、により構成される。帯域分割手段 3012 は、ローパスフィルタ手段 30122、ハイパスフィルタ手段 30124、2つのダウンサンプリング手段 30126 により構成される。また、図 6 記載のように、この実施例の適応ビームフォーミング部 104 は、周波数領域変換手段 1042 と、適応ビームフォーミング手段 1044 と、時間領域変換手段 1046 と、インパルス応答記憶部 1050（またはステアリングベクトル記憶部 1048）と、により構成される。また、図 7 記載のように、この実施例のエコー消去部 304 は、適応フィルタ手段 3048、減算手段 3051、により構成される。

【0018】

[低域エコー消去部 302_n の処理（ステップ S102）]

まず、收音信号 $y_n(t)$ が、低域エコー消去部 302_n に入力される。エコー消去部 302_n は、收音信号 $y_n(t)$ の予め定められた基準値 α より低い周波数帯域に含まれるエコー信号 $a_n(t)$ の成分を消去し、残留エコー信号 $e_n(t)$ を出力する。まず、基準値 α について説明する。この発明では、パワーが低域に集中する話者音声の性質から、低域の所望エコー消去量が高域の所望エコー消去量より大きいことに着目する。例えば、全周波数帯域を 24 kHz とすると、再生信号の周波数帯域 12 kHz 未満に話者音声は集中する場合が多く、結果として、收音信号の周波数帯域 12 kHz 未満に含まれるエコー信号の成分を消去すれば、大半のエコー信号の成分を消去できることになる。基準値 α とは、エコー信号の元となる信号（例えば、話者音声の信号）が集中している帯域の上限の周波数値を示すものである。基準値 α は予め定められるものであり、上記の例では $\alpha = 12 \text{ kHz}$ となる。このように、低域エコー消去部 302_n は、計算コスト削減のために、低域の收音信号 $y_n(t)$ についてのみ、エコー消去処理を行う。この周波数帯域の基準値 α は事前に設定しておけばよい。以下、低域エコー消去部 302 の処理を図 5 を用いて具体的に説明する。また、收音信号 $y_n(t)$ について、低域エコー消去部 302 に入力される收音信号を説明簡略化のために「 $y(t)$ 」と示す。

【0019】

收音手段により收音され、入力された收音信号 $y(t)$ は、帯域分割手段 3012 中のローパスフィルタ手段 30122 およびハイパスフィルタ手段 30124 に入力される。ローパスフィルタ手段 30122 は收音信号 $y(t)$ に対して、ローパスフィルタ（上記の例では、12 kHz 未満の信号のみ通過させるもの）の係数を畳み込み、低域の收音信号 $y_L(t)$ を出力する。

【0020】

一方、ハイパスフィルタ手段 30124 は收音信号 $y(t)$ に対して、ハイパスフィルタ（上記の例では、12 kHz 以上の信号のみ通過させるもの）の係数を畳み込み、高域の收音信号 $y_H(t)$ を出力する。ローパスフィルタ手段 30122、ローパスフィルタ手段 3014 において、通過させる周波数帯域の基準値 α を同一にすることが好ましい。

【0021】

低域收音信号 $y_L(t)$ 、高域收音信号 $y_H(t)$ はそれぞれ、ダウンサンプリング手段 30126、30128 に入力される。そして、ダウンサンプリング手段 30126 は、低域收音信号 $y_L(t)$ について、例えば $M - 1$ 個、間引くダウンサンプリング処理を行う。間引き数 M は 1 以上の正の実数であり、例えば「2」とする。ダウンサンプリング処理後の信号を $y_L(Mt)$ と表す。同様に、ダウンサンプリング手段 30128 は、高域收音信号 $y_H(t)$ について、例えば $M - 1$ 個、間引くダウンサンプリング処理を行う。ダウンサンプリング処理後の信号を $y_H(Mt)$ と表す。低域收音信号 $y_L(Mt)$ は減算手段 3020 に入力され、高域收音信号 $y_H(Mt)$ は帯域合成手段 3022 に入力

10

20

30

40

50

される。

【 0 0 2 2 】

一方、再生信号 $x(t)$ は、再生手段 2 およびローパスフィルタ手段 3 0 1 4 に入力される。ローパスフィルタ手段 3 0 1 4 は、ローパスフィルタ（上記の例では、1 2 k H z 未満の信号のみ通過させるもの）の係数を再生信号 $x(t)$ に対して畳み込むことで、再生信号 $x(t)$ の低域の信号（以下、「低域再生信号 $x_L(t)$ 」という。）を出力する。低域再生信号 $x_L(t)$ はダウンサンプリング手段 3 0 1 6 に入力される。ダウンサンプリング手段 3 0 1 6 は、低域再生信号 $x_L(t)$ について、 $M - 1$ 個、間引くダウンサンプリング処理を行う。ダウンサンプリング処理後の信号 $x_L(Mt)$ は適応フィルタ手段 3 0 1 8 に入力される。

10

【 0 0 2 3 】

ここでダウンサンプリング手段 3 0 1 6、3 0 1 2 6、3 0 1 2 8 によるダウンサンプリング処理については、元の標本化周波数より低くなればよい。有理数のダウンサンプリングに関しては、例えば、特願 2 0 0 7 - 2 6 2 9 8 6 号のアップダウンサンプリング装置とアップダウンサンプリング装置およびそのプログラムに詳細に記載されている。また、ダウンサンプリング手段 3 0 1 6、3 0 1 2 6、3 0 1 2 8 は無くても良い。

【 0 0 2 4 】

適応フィルタ手段 3 0 1 8 は、低域再生信号 $x_L(Mt)$ に適応フィルタの係数 $h_L(Mt)$ を畳み込むことで、低域擬似エコー信号 $d_L(Mt)$ を出力する。当該畳み込みの式は以下になる。

20

【 数 1 】

$$d_L(Mt) = \sum_{k=0}^{S-1} h_{L,k}(Mt) x_L(M(t-k))$$

ここで、 S はタップ数を示す。

【 0 0 2 5 】

減算手段 3 0 2 0 は、低域收音信号 $y_L(Mt)$ と、低域擬似エコー信号 $d_L(Mt)$ との差をとることで、低域残留エコー信号 $e_L(Mt)$ を出力する。図 5 の例では、低域收音信号 $y_L(Mt)$ から低域擬似エコー信号 $d_L(Mt)$ が差し引かれる。低域残留エコー信号 $e_L(Mt)$ は帯域合成手段 3 0 2 2 および適応フィルタ手段 3 0 1 8 に入力される。適応フィルタ手段 3 0 1 8 の適応フィルタの係数は、低域再生信号 $x_L(Mt)$ および過去の低域残留エコー信号 $e_L(Mt)$ により更新される。この更新については、様々なアルゴリズムが知られており、例えば、学習同定アルゴリズム（N L M S : Normalized Least-Mean-Squares）や指数重みつけアルゴリズムがあるが、念のため学習同定アルゴリズムについて簡単に説明する。学習同定アルゴリズムにより低域再生信号 $x_L(Mt)$ のベクトル $X_L(Mt)$ を用いて、以下の式で適応フィルタ係数ベクトル $H_L(Mt)$ は更新される。

30

【 数 2 】

$$H_L(M(t+1)) = H_L(Mt) + \beta \frac{e_L(Mt)}{X_L^T(Mt)X_L(Mt) + \delta} X_L(Mt)$$

40

ただし、

$$X_L(Mt) = [x_L(Mt), x_L(M(t-1)), \dots, x_L(M(t-S+1))]^T$$

$H_L(Mt) = [h_{L,0}(Mt), h_{L,1}(Mt), \dots, h_{L,S-1}(Mt)]^T$ であり、 β はスカラ量のステップサイズであって、取りうる範囲は $0 < \beta < 2$ である。また、 δ は分母が 0 になることを防止するための微小な定数である。＜学習同定アルゴリズムの説明は以上＞

【 0 0 2 6 】

帯域合成手段 3 0 2 2 は、低域残留エコー信号 $e_L(Mt)$ と高域收音信号 $y_H(Mt)$

50

）について、帯域合成を行うことで、残留エコー信号 $e(t)$ を出力する。処理の詳細の具体例として、低域残留エコー信号 $e_L(Mt)$ と高域収音信号 $y_H(Mt)$ に対して、サンプル間に $M-1$ 個の零値を挿入し、アップサンプルで生じるイメージング（アップサンプリングする前のサンプリング周波数の折り返しスペクトル）を取り除くフィルタを積み込んだそれぞれの出力信号を加算することで、残留エコー信号 $e(t)$ を出力する。そして、上記の処理が全ての低域エコー消去部 302_n ($n=1, \dots, N$) で行われ、それぞれの低域エコー消去部 302_n から残留エコー信号 $e_n(t)$ が出力される。

【0027】

[適応ビームフォーミング部 104 の処理 (ステップ S104)]

全ての N 個の残留エコー信号 $e_n(t)$ ($n=1, \dots, N$) は、適応ビームフォーミング部 104 に入力される。適応ビームフォーミング部 104 は、 N 個の残留エコー信号 $e_n(t)$ から雑音信号を抑圧して、適応ビームフォーミング信号 $s(t)$ を出力する。以下、図 6 を用いて具体例を説明する。なお、以下の適応ビームフォーミング部 104 の説明の詳細については、特開 2008-60635 号の段落 [0004] ~ [0010] 等に記載されている。

【0028】

入力された N 個の残留エコー信号 $e_n(t)$ は周波数領域変換手段 1042 に入力される。周波数領域変換手段 1042 は、時間領域の信号である残留エコー信号 $e_n(t)$ を周波数領域信号 $E_n(f, \tau)$ に変換する。ここで、 f は周波数を示し、 τ は任意の時間を示す。周波数領域の変換については、公知の技術であるフーリエ変換などを用いればよい。 $E_n(f, \tau)$ ($n=1, \dots, N$) は全て適応ビームフォーミング手段 1044 に入力される。

【0029】

適応ビームフォーミング手段 1044 は、目的信号である話者信号 $c(t)$ を強調し、不要信号である雑音信号 $b(t)$ をできるだけ抑圧するフィルタ ($W_1(f), \dots, W_N(f)$)^T を推定することで実現される。適応ビームフォーミング手段 1044 を設計する際には、「話者から各収音手段へのインパルス応答ベクトル $g(f)$ もしくは、 $g(f)$ の近似であるステアリングベクトル $a(f) = [\exp(-i2\pi f\tau_1), \dots, \exp(-i2\pi f\tau_N)]^T$ が既知である」ということを仮定する。ここで τ_n は話者 8 よりの話者信号 (目的信号) が収音手段 n に達する時刻と原点 0 に達する時間差である。そして、図 8 に示すように収音手段 4_n を直線状に配置させることが多い。話者 8 の方向を θ とし、収音手段 4_1 を基準 (原点) とした場合の収音手段 4_n の座標 d_n とすると、上述の τ_n は、 $\tau_n = d_n \cos \theta / c$ で与えられる。ここで c は信号の速度である。

【0030】

図 6 に説明を戻すと、不要信号を抑圧する適応ビームフォーミング手段 1044 として、以下の式で表される出力パワー $A(W(f))$ を最小にする適応フィルタ群 (ベクトル) 1048_n ($n=1, \dots, N$) である $W(f) = [W_1(f), \dots, W_N(f)]$ を推定する。

$$\begin{aligned} A(W(f)) &= V\{|S|^2(f, \tau)\} \\ &= V\{S(f, \tau)S^*(f, \tau)\} \\ &= V\{W^T(f)E(f, \tau)E^T(f, \tau)W(f)\} \\ &= W^T(f)R_E(f)W(f) \quad (1) \end{aligned}$$

【0031】

ここで、 $V\{\}$ は時間 τ に関する平均操作、 A^* は A の複素共役、 $R_E(f) = V\{E(f, \tau)E^T(f, \tau)\}$ は入力信号の相関行列、 $S(f, \tau)$ は適応ビームフォーミング手段 1044 の出力であり、以下の式で表すことが出来る。

$$S(f, \tau) = W^T(f)E(f, \tau) \quad (2)$$

ここで、意味のない解 ($W(f) = 0 = [0, \dots, 0]^T$) を回避するために、目的信号が無歪みで得られるという以下の式に示す拘束条件を付与する。

10

20

30

40

50

$$W^T(f)g(f) = 1 \quad (3)$$

【 0 0 3 2 】

これにより、式 (3) を満たし、かつ上記式 (1) の $A(W(f))$ の値が最小となる $W(f)$ の値を求める問題は $Lagrange$ の未定乗数 p を用いて、以下の式 (4) で表すことができる。

$$A'(W(f)) = A(W(f)) + p(W^T(f)g(f) - 1) \quad (4)$$

式 (4) を解くことにより、適応フィルタ群 (ベクトル) $W(f)$ は、以下の式 (5) で得られる。

【 数 3 】

$$W(f) = \frac{R_E^{-1}(f)g(f)}{g^T(f)R_E^{-1}(f)g(f)} \quad (5)$$

10

【 0 0 3 3 】

適応ビームフォーミング手段 1 0 4 4 から出力された周波数領域の適応ビームフォーミング信号 $S(f, \tau)$ は時間領域変換手段 1 0 4 6 に入力される。時間領域変換手段 1 0 4 6 は、例えば公知の技術である逆フーリエ変換などを用いて、周波数領域の信号である適応ビームフォーミング信号 $S(f, \tau)$ を時間領域の適応ビームフォーミング信号 $s(t)$ に変換して出力する。

【 0 0 3 4 】

[エコー消去部 3 0 4 の処理 (ステップ S 1 0 6)]

20

適応ビームフォーミング信号 $s(t)$ は、エコー消去部 3 0 4 に入力される。エコー消去部 3 0 4 は、適応ビームフォーミング信号 $s(t)$ から、エコー信号を消去することで、出力信号 $z(t)$ として出力する。エコー消去部 3 0 4 の処理内容は公知であるが、図 7 を用いて念のため説明する。

【 0 0 3 5 】

再生信号 $x(t)$ は、適応フィルタ手段 3 0 4 8 に入力される。適応フィルタ手段 3 0 4 8 は、以下の式のように、再生信号 $x(t)$ に適応フィルタの係数 $h(t)$ を畳み込むことで擬似エコー信号のベクトル $d(t)$ を出力する。

【 数 4 】

$$d(t) = \sum_{k=0}^{S-1} h_k(t)x(t-k)$$

30

【 0 0 3 6 】

そして、減算手段 3 0 5 1 は、適応ビームフォーミング信号 $s(t)$ から擬似エコー信号 $d(t)$ 差し引くことで、残留エコー信号 $e(t)$ を出力する。そして、適応フィルタ手段 3 0 4 8 の適応フィルタのベクトルは、過去の残留エコー信号と再生信号により、上述した学習同定法等を用いて更新される。

【 0 0 3 7 】

ここで、 N 個の低域エコー消去部 3 0 2_n により、収音信号の低周波数帯域に含まれるエコー信号の成分は消去されており、つまり、適応ビームフォーミング信号に含まれるエコー信号の成分は大半が消去されているので、残留エコー信号 $e(t)$ のパワーはかなり小さくなる。従って、所望エコー消去量に達するまでの時間を短くでき、適応ビームフォーミング部 1 0 4 の音響的指向性の変動が生じたとしても、所望のエコー消去量に達することが出来る。

40

【 0 0 3 8 】

そして、残留エコー信号 $e(t)$ は出力信号 $z(t)$ として出力される。また、話者信号 $c(t)$ が存在する場合には、残留エコー信号 $e(t)$ と話者信号 $c(t)$ とが出力信号 $z(t)$ として出力される。

【 実施例 2 】

【 0 0 3 9 】

50

実施例 1 のエコー消去部 304 は入力信号の全周波数帯域についてエコー信号の成分を消去した。しかし、N 個の低域エコー消去部 302_n で、低域のエコー信号の成分を消去されているので、エコー消去部では、基準値 α (上記の例では $\alpha = 12 \text{ kHz}$) より高い周波数帯域の適応ビームフォーミング信号から、エコー信号を消去するようにしてもよい。このようにすることで、エコー消去部の計算コストを大幅に削減できる。この場合のエコー消去部の参照符号を 306 とし、エコー消去部 306 の機能構成例を図 9 に示す。エコー消去部 306 は低域エコー消去部 302 (図 5 参照) のハイパスフィルタ手段とローパスフィルタ手段とを入れ替えた構成になっている。エコー消去部 306 の処理内容を簡単に説明する。

【0040】

まず、適応ビームフォーミング信号 $s(t)$ が、帯域分割手段 3062 中のハイパスフィルタ手段 30622 およびローパスフィルタ手段 30624 に入力される。ハイパスフィルタ手段 30622 は適応ビームフォーミング $s(t)$ に対して、ハイパスフィルタ (上記の例では、 12 kHz (= 基準値 α) 以上の信号のみ通過させるもの) の係数を畳み込み、高域の適応ビームフォーミング信号 $s_H(t)$ を出力する。ローパスフィルタ手段 30624 は適応ビームフォーミング信号 $s(t)$ に対して、ローパスフィルタ (上記の例では、 12 kHz 未満の信号のみ通過させるもの) の係数を畳み込み、低域の適応ビームフォーミング信号 $s_L(t)$ を出力する。

【0041】

高域適応ビームフォーミング信号 $s_H(t)$ 、低域適応ビームフォーミング信号 $s_L(t)$ はそれぞれ、ダウンサンプリング手段 30626、30628 に入力される。そして、ダウンサンプリング手段 30626、30628 はそれぞれ、高域適応ビームフォーミング信号 $s_H(t)$ 、低域適応ビームフォーミング信号 $s_L(t)$ について、ダウンサンプリング処理を行う。ダウンサンプリング処理後の高域適応ビームフォーミング信号 $s_H(Mt)$ は、減算手段 3070 に入力され、ダウンサンプリング処理後の低域適応ビームフォーミング信号 $s_L(Mt)$ は、帯域合成手段 3072 に入力される。

【0042】

一方、ハイパスフィルタ手段 3064 は、再生信号 $x(t)$ に対して、ハイパスフィルタ (上記の例では、 12 kHz (= 基準値 α) 以上の信号のみ通過させるもの) の係数を再生信号 $x(t)$ に対して畳み込むことで、高域再生信号 $x_H(t)$ を出力する。高域再生信号 $x_H(t)$ はダウンサンプリング手段 3066 に入力される。ダウンサンプリング手段 3066 は、高域再生信号 $x_H(t)$ について、ダウンサンプリング処理を行う。ダウンサンプリング処理後の信号 $x_H(Mt)$ は適応フィルタ手段 3068 に入力される。また、ダウンサンプリング手段 3066、30626、30628 は無くても良い。

【0043】

適応フィルタ手段 3068 は、低域再生信号に適応フィルタを畳み込むことで、高域擬似エコー信号 $d_H(Mt)$ を出力する。減算手段 3070 は、高域適応ビームフォーミング信号 $s_H(Mt)$ から高域擬似エコー信号 $d_H(Mt)$ を差し引くことで、高域残留エコー信号 $e_H(Mt)$ を出力する。高域残留エコー信号 $e_H(Mt)$ は帯域合成手段 3072 および適応フィルタ手段 3068 に入力される。適応フィルタ手段 3068 の適応フィルタ係数は、高域再生信号 $x_H(Mt)$ および過去の高域残留エコー信号 $e_H(Mt)$ により、上記学習同定アルゴリズム等を用いて、更新される。

【0044】

帯域合成手段 3072 は、高域残留エコー信号 $e_H(Mt)$ と低域適応ビームフォーミング信号 $s_L(Mt)$ について、帯域合成およびアップサンプリング処理を行うことで、残留エコー信号 $e(t)$ を出力する。帯域合成手段 3072 の処理の詳細は、帯域合成手段 3022 の処理と同様なので、省略する。そして、帯域合成手段 3072 から残留エコー信号 $e(t)$ が、出力信号 $z(t)$ として出力される。＜エコー消去部 306 の処理の説明以上＞

【0045】

10

20

30

40

50

上記のエコー消去装置 300 は、再生手段 2 が 1 つの場合について説明したが、再生手段が 2 以上の場合であっても、適用できる。

また、低域エコー消去部 302_n、エコー消去部 304 の処理は時間領域で行い、適応ビームフォーミング部 104 の処理は周波数領域で行った例を説明した。しかし、低域エコー消去部 302_n、エコー消去部 304 の処理を周波数領域で行い、適応ビームフォーミング部 104 の処理は時間領域で行ってもよい。

【0046】

この発明では、パワーが低域に集中する話者音声の性質から、低域の所望エコー消去量が高域の所望エコー消去量より大きいことに着目する。そして、本実施例のエコー消去装置は、「低域エコー消去部 302_nによる低域のエコー消去処理」→「適応ビームフォーミング部 104 による適応ビームフォーミング処理」→「エコー消去部 304 によるエコー消去」、の順番で処理を行う。まず、低域エコー消去部 302_nは、收音信号の低周波数帯域のみに含まれるエコー信号の成分を消去する。従って、エコー消去装置 100 よりも計算コストを削減できる。また、適応ビームフォーミング処理の前に、低域のエコー消去処理を行う。従って、低域エコー消去部のエコー経路モデルでは、適応ビームフォーミング部 104 の音響的指向性の変動による再推定を回避できる。また、低域のエコー消去処理により收音信号に含まれる大半のエコー信号の成分を消去している。従って、エコー消去部 304 中の減算手段 3051 よりの残留エコー信号 $e(t)$ のパワーは小さいものとなり、結果として、エコー消去部の所望エコー消去量に達するまでの時間を短くできる。よって、適応ビームフォーミング部 104 による音響的指向特性の変動が生じたとしても、所望のエコー消去量に達することが出来る。

【0047】

また、この実施例 1 のエコー消去装置 300 が従来のエコー消去装置 100 よりも優れていることを示す実験結果について説明する。マイクロホン数 $N = 4$ とし、全周波数帯域を 24 kHz とし、周波数帯域の基準値 $\alpha = 12 \text{ kHz}$ 、つまり、低域を 12 kHz 未満、高域を $12 \sim 24 \text{ kHz}$ とした場合に、本実施例のエコー消去装置 300 の計算コストは、エコー消去装置 100 のエコー消去処理にかかる計算コストより約 40% 削減できる。

【0048】

< ハードウェア構成 >

本発明は上述の実施の形態に限定されるものではない。また、上述の各種の処理は、記載に従って時系列に実行されるのみならず、処理を実行する装置の処理能力あるいは必要に応じて並列的にあるいは個別に実行されてもよい。その他、本発明の趣旨を逸脱しない範囲で適宜変更が可能であることはいうまでもない。

また、上述の構成をコンピュータによって実現する場合、エコー消去装置 300 が有すべき機能の処理内容はプログラムによって記述される。そして、このプログラムをコンピュータで実行することにより、処理機能がコンピュータ上で実現される。

【0049】

この処理内容を記述したプログラムは、コンピュータで読み取り可能な記録媒体に記録しておくことができる。コンピュータで読み取り可能な記録媒体としては、例えば、磁気記録装置、光ディスク、光磁気記録媒体、半導体メモリ等のようなものでもよいが、具体的には、例えば、磁気記録装置として、ハードディスク装置、フレキシブルディスク、磁気テープ等を、光ディスクとして、DVD (Digital Versatile Disc)、DVD-RAM (Random Access Memory)、CD-ROM (Compact Disc Read Only Memory)、CD-R (Recordable) / RW (ReWritable) 等を、光磁気記録媒体として、MO (Magnetooptical disc) 等を、半導体メモリとしてEEP-ROM (Electrically Erasable and Programmable-Read Only Memory) 等を用いることができる。

【0050】

また、このプログラムの流通は、例えば、そのプログラムを記録したDVD、CD-ROM等の可搬型記録媒体を販売、譲渡、貸与等することによって行う。さらに、このプロ

グラムをサーバコンピュータの記憶装置に格納しておき、ネットワークを介して、サーバコンピュータから他のコンピュータにそのプログラムを転送することにより、このプログラムを流通させる構成としてもよい。

【0051】

このようなプログラムを実行するコンピュータは、例えば、まず、可搬型記録媒体に記録されたプログラムもしくはサーバコンピュータから転送されたプログラムを、一旦、自己の記憶装置に格納する。そして、処理の実行時、このコンピュータは、自己の記録媒体に格納されたプログラムを読み取り、読み取ったプログラムに従った処理を実行する。また、このプログラムの別の実行形態として、コンピュータが可搬型記録媒体から直接プログラムを読み取り、そのプログラムに従った処理を実行することとしてもよく、さらに、このコンピュータにサーバコンピュータからプログラムが転送されるたびに、逐次、受け取ったプログラムに従った処理を実行することとしてもよい。また、サーバコンピュータから、このコンピュータへのプログラムの転送は行わず、その実行指示と結果取得のみによって処理機能を実現する、いわゆる A S P (Application Service Provider) 型のサービスによって、上述の処理を実行する構成としてもよい。なお、本形態におけるプログラムには、電子計算機による処理の用に供する情報であってプログラムに準ずるもの(コンピュータに対する直接の指令ではないがコンピュータの処理を規定する性質を有するデータ等)を含むものとする。

【0052】

また、この形態では、コンピュータ上で所定のプログラムを実行させることにより、本装置を構成することとしたが、これらの処理内容の少なくとも一部をハードウェア的に実現することとしてもよい。

また、本実施例で説明したエコー消去装置 300 は、C P U (Central Processing Unit)、入力部、出力部、補助記憶装置、R A M (Random Access Memory)、R O M (Read Only Memory) 及びバスを有している(何れも図示せず)。

C P U は、読み込まれた各種プログラムに従って様々な演算処理を実行する。補助記憶装置は、例えば、ハードディスク、M O (Magneto-Optical disc)、半導体メモリ等であり、R A M は、S R A M (Static Random Access Memory)、D R A M (Dynamic Random Access Memory) 等である。また、バスは、C P U、入力部、出力部、補助記憶装置、R A M 及び R O M を通信可能に接続している。

【0053】

< ハードウェアとソフトウェアとの協働 >

本実施例の単語追加装置は、上述のようなハードウェアに所定のプログラムが読み込まれ、C P U がそれを実行することによって構築される。以下、このように構築される各装置の機能構成を説明する。

エコー消去装置 300 の低域エコー消去部 302_n、適応ビームフォーミング部 104、エコー消去部 304 は、所定のプログラムが C P U に読み込まれ、実行されることによって構築される演算部である。エコー消去装置 300 の記憶部(図示せず)は上記補助記憶装置として機能する。

【図面の簡単な説明】

【0054】

【図1】従来のエコー消去装置の機能構成例を示した図。

【図2】従来のエコー消去装置の機能構成例を示した図。

【図3】実施例1のエコー消去装置の機能構成例を示した図。

【図4】実施例1のエコー消去装置の処理フローを示した図。

【図5】実施例1の低域エコー消去部の機能構成例を示した図。

【図6】実施例1の適応ビームフォーミング部の機能構成例を示した図。

【図7】実施例1のエコー消去部の機能構成例を示した図。

【図8】適応ビームフォーミング処理で用いる τ を説明するための図。

【図9】実施例2のエコー消去部の機能構成例を示した図。

【図 1】

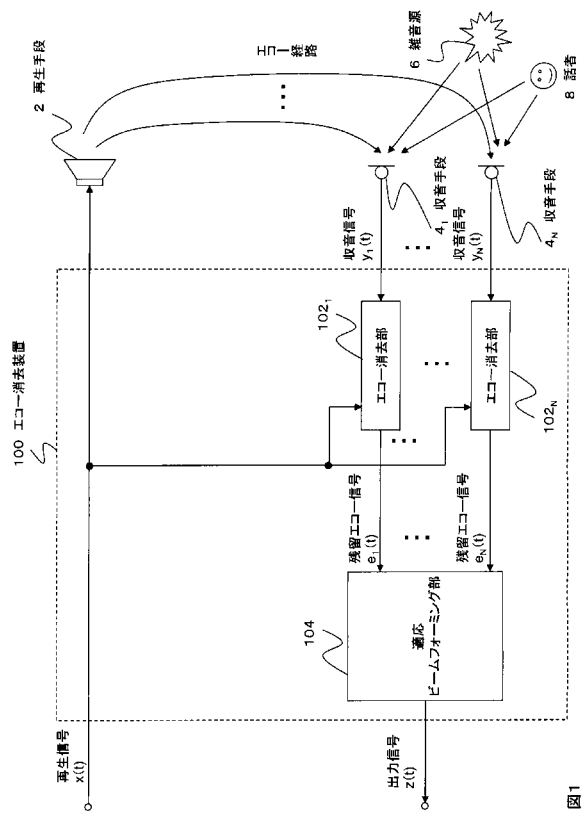


図1

【図 2】

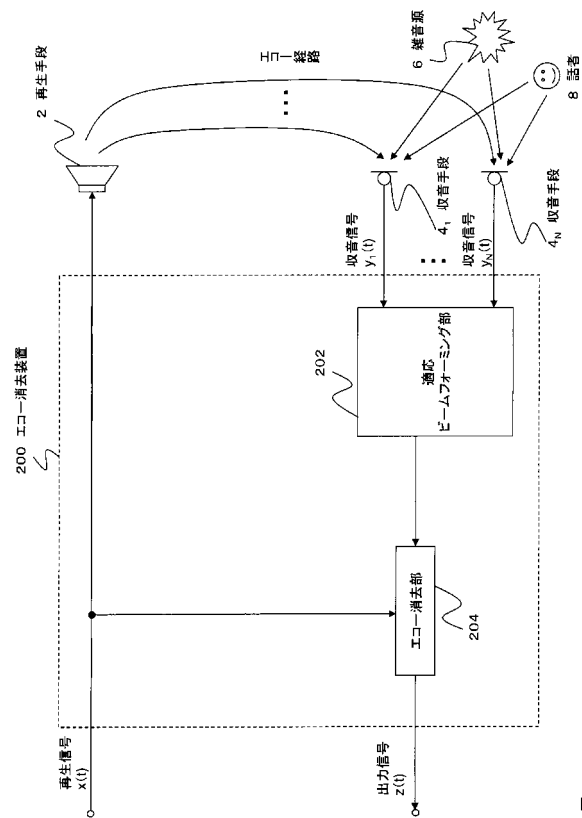


図2

【図 3】

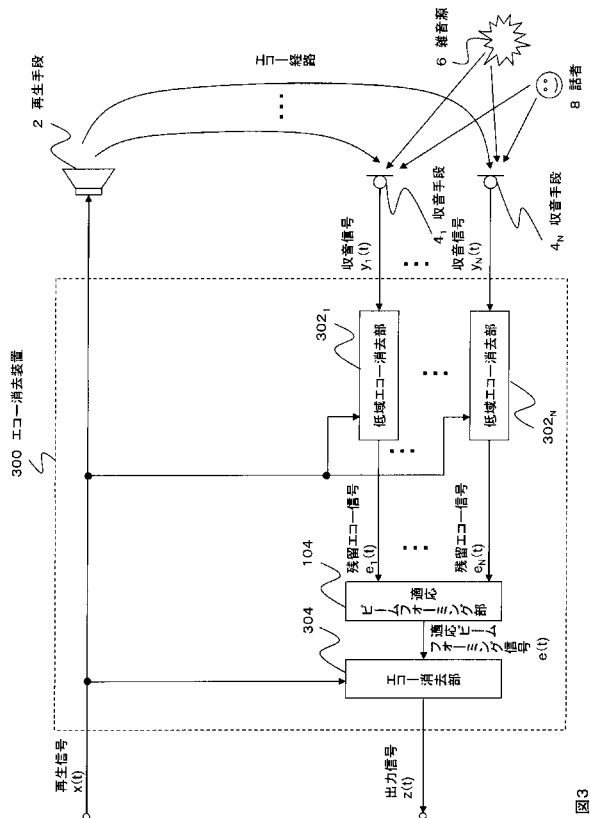


図3

【図 4】

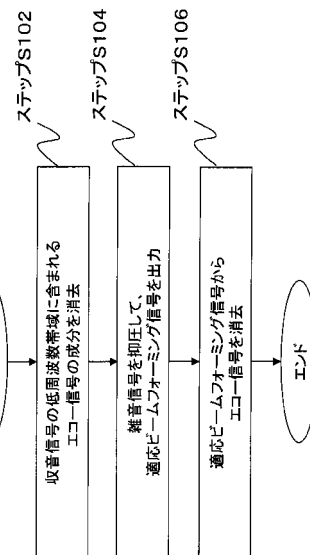


図4

フロントページの続き

(72)発明者 羽田 陽一

東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内

Fターム(参考) 5D018 BB22

5D020 BB04 BB07 CC04

5K027 BB14 DD10

5K046 HH11 HH24 HH41 HH57 HH58 HH71 HH79