

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2010-507105  
(P2010-507105A)

(43) 公表日 平成22年3月4日(2010.3.4)

(51) Int. Cl.	F I	テーマコード (参考)
<b>G 1 0 L 21/02 (2006.01)</b>	G 1 0 L 21/02 1 0 3 Z	5 J 0 2 3
<b>H O 4 M 3/56 (2006.01)</b>	H O 4 M 3/56 B	5 K 0 4 6
<b>H O 4 M 1/60 (2006.01)</b>	H O 4 M 1/60 C	5 K 1 2 7
<b>H O 3 H 21/00 (2006.01)</b>	H O 4 M 1/60 A	5 K 2 0 1
<b>H O 4 B 3/23 (2006.01)</b>	G 1 0 L 21/02 1 0 1 B	
審査請求 有 予備審査請求 未請求 (全 24 頁) 最終頁に続く		

(21) 出願番号 特願2009-532431 (P2009-532431)  
 (86) (22) 出願日 平成19年10月12日 (2007.10.12)  
 (85) 翻訳文提出日 平成21年5月19日 (2009.5.19)  
 (86) 国際出願番号 PCT/US2007/021814  
 (87) 国際公開番号 W02008/045537  
 (87) 国際公開日 平成20年4月17日 (2008.4.17)  
 (31) 優先権主張番号 11/546,680  
 (32) 優先日 平成18年10月12日 (2006.10.12)  
 (33) 優先権主張国 米国 (US)

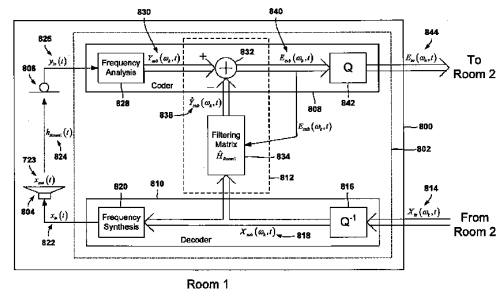
(71) 出願人 503003854  
 ヒューレット-パッカード デベロップメント カンパニー エル. ピー.  
 アメリカ合衆国 テキサス州 77070  
 ヒューストン コンパクト センタ ド ライブ ウェスト 11445  
 (74) 代理人 100075513  
 弁理士 後藤 政喜  
 (74) 代理人 100078053  
 弁理士 上野 英夫  
 (74) 代理人 100114236  
 弁理士 藤井 正弘  
 (74) 代理人 100120260  
 弁理士 飯田 雅昭

最終頁に続く

(54) 【発明の名称】 音声会議通信システムにおいて音響エコーをキャンセルするためのシステム及び方法

(57) 【要約】

本発明の種々の実施の形態は、音響エコーキャンセレーション機能を含む音声会議通信システムのための周波数領域符号器/復号器 802 に関する。本発明の一実施の形態では、音響エコーキャンセラ 812 は、周波数領域符号器/復号器 802 に組み込まれ、周波数領域に変換されており、且つ周波数領域符号器/復号器 802 によって部分帯域に分割されている音声信号からの音響エコーを緩和するか、又はその音響エコーを除去する。



## 【特許請求の範囲】

## 【請求項 1】

第 1 の場所 ( 8 0 0 ) にある音声会議通信システムの周波数領域符号器 / 復号器構成要素 ( 8 0 2 ) であって、前記周波数領域符号器 / 復号器構成要素 ( 8 0 2 ) は、

第 2 の場所 ( 1 0 4 ) から受信される量子化された周波数領域音声信号 ( 8 1 4 ) を 1 組の第 2 の場所の部分帯域信号 ( 8 1 8 ) に変換する復号器 ( 8 1 0 ) と、

前記第 1 の場所 ( 8 0 0 ) から受信される時間領域エコー音声信号 ( 8 2 6 ) を 1 組の第 1 の場所の周波数領域エコー部分帯域信号 ( 8 3 0 ) に変換する符号器 ( 8 0 8 ) と、

前記 1 組の第 2 の場所の部分帯域信号 ( 8 1 8 ) 及び前記 1 組の第 1 の場所の周波数領域エコー部分帯域信号 ( 8 3 0 ) に基づいて 1 組の周波数領域誤差音声部分帯域信号 ( 8 4 0 ) を生成し、前記生成された 1 組の周波数領域誤差音声部分帯域信号 ( 8 4 0 ) に基づいて、第 1 の場所のインパルス応答 ( 8 2 4 ) を追跡する音響エコーキャンセラ ( 8 1 2 ) と、

前記第 2 の場所 ( 1 0 4 ) に、量子化された周波数領域誤差音声部分帯域信号 ( 8 4 4 ) を出力する音声信号出力とを備えることを特徴とする周波数領域符号器 / 復号器構成要素。

10

## 【請求項 2】

周波数領域符号器 / 復号器構成要素 ( 8 0 2 ) であって、

前記復号器 ( 8 1 0 ) は、

前記第 2 の場所 ( 1 0 4 ) から受信される前記量子化された周波数領域音声信号 ( 8 1 4 ) を前記 1 組の第 2 の場所の部分帯域信号 ( 8 1 8 ) に変換するための逆量子化器 ( 8 1 6 ) と、

20

前記第 2 の場所の部分帯域信号 ( 8 1 8 ) を単一のサンプリングされた音声時間領域波形 ( 8 2 2 ) に変換するための周波数合成ステージ ( 8 2 0 ) とを備えることを特徴とする、請求項 1 に記載の周波数領域符号器 / 復号器構成要素。

## 【請求項 3】

周波数領域符号器 / 復号器構成要素 ( 8 0 2 ) であって、

前記符号器 ( 8 0 8 ) は、

前記第 1 の場所 ( 8 0 0 ) から受信される前記時間領域エコー音声信号 ( 8 2 6 ) を、前記音響エコーキャンセラ ( 8 1 2 ) に入力される、前記 1 組の第 1 の場所の周波数領域エコー部分帯域信号 ( 8 3 0 ) に変換するための周波数解析ステージ ( 8 2 8 ) と、

30

前記音響エコーキャンセラ ( 8 1 2 ) によって生成される前記 1 組の周波数領域誤差音声部分帯域信号 ( 8 4 0 ) を、前記第 2 の場所 ( 1 0 4 ) に出力される前記量子化された周波数領域誤差音声部分帯域信号 ( 8 4 4 ) に変換するための量子化器 ( 8 4 2 ) とを備えることを特徴とする請求項 1 に記載の周波数領域符号器 / 復号器構成要素。

## 【請求項 4】

周波数領域符号器 / 復号器構成要素 ( 8 0 2 ) であって、

前記 1 組の量子化された周波数領域誤差音声部分帯域信号 ( 8 4 0 ) が前記第 2 の場所 ( 1 0 4 ) に出力される前に、前記 1 組の周波数領域誤差音声部分帯域信号 ( 8 4 0 ) に関して、

40

知覚符号化、

雑音低減、及び

ウィーナタイプフィルタリングのうちの 1 つ又は複数が実施されることを特徴とする請求項 1 に記載の周波数領域符号器 / 復号器構成要素。

## 【請求項 5】

周波数領域符号器 / 復号器構成要素 ( 8 0 2 ) であって、

前記音響エコーキャンセラ ( 8 1 2 ) は、

前記生成された 1 組の周波数領域誤差音声部分帯域信号 ( 8 4 0 ) に基づいて前記第 1 の場所のインパルス応答 ( 8 2 4 ) を追跡し、1 組の第 1 の場所のエコー部分帯域信号推定値 ( 8 3 8 ) を出力する適応フィルタ ( 8 3 4 ) と、

50

前記受信された 1 組の第 1 の場所の周波数領域エコー部分帯域信号 ( 8 3 0 ) から前記受信された 1 組の第 1 の場所のエコー部分帯域信号推定値 ( 8 3 8 ) を減算し、前記 1 組の周波数領域誤差音声部分帯域信号 ( 8 4 0 ) を出力する加算接合部 ( 8 3 2 ) とをさらに備えることを特徴とする請求項 1 に記載の周波数領域符号器 / 復号器構成要素。

【請求項 6】

音声会議通信システムにおいて音響エコーをキャンセルする方法であって、

復号器 ( 8 1 0 ) と、符号器 ( 8 0 8 ) と、音響エコーキャンセラ ( 8 1 2 ) とを備える周波数領域符号器 / 復号器 ( 8 0 2 ) を第 1 の場所 ( 8 0 0 ) に設け、

第 2 の場所 ( 1 0 4 ) から前記復号器 ( 8 1 0 ) に、量子化された周波数領域音声信号 ( 8 1 4 ) を送信し、前記量子化された周波数領域音声信号 ( 8 1 4 ) を 1 組の第 2 の場所の部分帯域信号 ( 8 1 8 ) に変換し、

前記第 1 の場所 ( 8 0 0 ) から前記符号器 ( 8 0 8 ) に、時間領域エコー音声信号 ( 8 2 6 ) を送信し、前記時間領域エコー音声信号 ( 8 2 6 ) を 1 組の第 1 の場所の周波数領域エコー部分帯域信号 ( 8 3 0 ) に変換し、

前記音響エコーキャンセラ ( 8 1 2 ) によって、前記 1 組の第 2 の場所の部分帯域信号 ( 8 1 8 ) 及び前記 1 組の第 1 の場所の周波数領域エコー部分帯域信号 ( 8 3 0 ) に基づいて 1 組の周波数領域誤差音声部分帯域信号 ( 8 4 0 ) を生成し、前記生成された 1 組の周波数領域誤差部分帯域信号 ( 8 4 0 ) に基づいて、第 1 の場所のインパルス応答 ( 8 2 4 ) を追跡し、

前記第 2 の場所 ( 1 0 4 ) に、量子化された周波数領域誤差音声部分帯域信号 ( 8 4 4 ) を出力することを含むことを特徴とする方法。

【請求項 7】

前記復号器 ( 8 1 0 ) は、

前記第 2 の場所 ( 1 0 4 ) から受信される前記量子化された周波数領域音声信号 ( 8 1 4 ) を前記 1 組の第 2 の場所の部分帯域信号 ( 8 1 8 ) に変換するための逆量子化器 ( 8 1 6 ) と、

前記第 2 の場所の部分帯域信号 ( 8 1 8 ) を単一のサンプリングされた音声時間領域波形 ( 8 2 2 ) に変換するための周波数合成ステージ ( 8 2 0 ) とを備えることを特徴とする、請求項 6 に記載の方法。

【請求項 8】

前記符号器 ( 8 0 8 ) は、

前記第 1 の場所 ( 8 0 0 ) から受信される前記時間領域エコー音声信号 ( 8 2 6 ) を、前記音響エコーキャンセラ ( 8 1 2 ) に入力される、前記 1 組の第 1 の場所の周波数領域エコー部分帯域信号 ( 8 3 0 ) に変換するための周波数解析ステージ ( 8 2 8 ) と、

前記音響エコーキャンセラ ( 8 1 2 ) によって生成される前記 1 組の周波数領域誤差音声部分帯域信号 ( 8 4 0 ) を、前記第 2 の場所に出力される前記量子化された周波数領域誤差音声部分帯域信号 ( 8 4 4 ) に変換するための量子化器 ( 8 4 2 ) とを備えることを特徴とする請求項 6 に記載の方法。

【請求項 9】

前記 1 組の量子化された周波数領域誤差音声部分帯域信号 ( 8 4 0 ) が前記第 2 の場所 ( 1 0 4 ) に出力される前に、前記 1 組の周波数領域誤差音声部分帯域信号 ( 8 4 0 ) に関して、

知覚符号化、

雑音低減、及び

ウィーナタイプフィルタリングのうちの 1 つ又は複数が実施されることを特徴とする請求項 6 に記載の方法。

【請求項 10】

前記音響エコーキャンセラ ( 8 1 2 ) は、

前記生成された 1 組の周波数領域誤差音声部分帯域信号 ( 8 4 0 ) に基づいて前記第 1 の場所のインパルス応答 ( 8 2 4 ) を追跡し、 1 組の第 1 の場所のエコー部分帯域信号推

10

20

30

40

50

定値(838)を出力する適応フィルタ(834)と、

前記受信された1組の第1の場所の周波数領域エコー部分帯域信号(830)から前記受信された1組の第1の場所のエコー部分帯域信号推定値(838)を減算し、前記1組の周波数領域誤差音声部分帯域信号(840)を出力する加算接合部(832)とをさらに備えることを特徴とする請求項6に記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は音響エコーキャンセレーションに関し、詳細には、音声会議通信システムにおいて音響エコーをキャンセルするためのシステム及び方法に関する。

10

【背景技術】

【0002】

一般のインターネット、電子プレゼンテーション、ボイスメール及び音声会議通信システムのような通信媒体が、さらに良好な音声及び通信技術の需要を拡大している。現在、効率及び生産性を高めると同時に、コストを削減し、構成を簡単にするために、数多くの個人及び企業がこれらの通信媒体を使用する。音声会議通信システムによって、第1の場所にいる1人又は複数の人が、ヘッドセットを装着することなく、又は携帯通信デバイスを用いることなく、全二重通信線を通じて、他の場所にいる1人又は複数の人と同時に会話することができるようになる。典型的には、音声会議通信システムは、各場所に、複数のマイクロフォン及びラウドスピーカを備える。これらのマイクロフォン及びラウドスピーカは、他の場所との間で音声信号を送受信するために、多数の人によって使用される。音声信号を伝送するためにデジタル通信システムが用いられるとき、伝送前に音声信号を圧縮し、伝送後に音声信号を解凍するために、多くの場合に、音声会議通信システムの中に符号器/復号器が組み込まれる。

20

【0003】

最新の音声会議通信システムは、聞き取れるほどの歪み、背景雑音及び他の望ましくない音声アーティファクトがない、音声信号の明瞭な伝送の提供を試みている。1つの一般的なタイプの望ましくない音声アーティファクトは音響エコーである。マイクロフォンとスピーカとの結合に起因して、送信された音声信号が電話会議通信システムを通じて戻されるときに、音響エコーが生じることがある。たとえば、音声信号が、第1の場所にあるマイクロフォンから第2の場所にあるラウドスピーカまで伝送されるとき、その音声信号が、第2の場所において結合されるマイクロフォンに伝達されることがあり、その後、第1の場所にあるラウドスピーカに戻されることがある。そのような場合に、第1の場所においてマイクロフォンに向かって話をしている人は、もともと自分が送信した音声信号の遅延したエコーを聞くことがある。信号増幅又は利得、及び各場所におけるマイクロフォンのスピーカへの接近に応じて、第1の場所においてマイクロフォンに向かって話をしている人は、うるさいほどのハウリング音を聞く可能性さえある。

30

【0004】

音声会議通信システムの設計者は、種々の方法で音響エコーの補償を試みてきた。1つの補償技法は、エコーをキャンセルするために、「音響エコーキャンセラ」と呼ばれるフィルタリングシステムを使用する。音響エコーキャンセラは、元の音声信号の送り手に音響エコーが達する前に、音響エコーをキャンセルしようと試みる。典型的には、音響エコーキャンセラは適応フィルタを使用し、適応フィルタは、音響エコーの特性に影響を及ぼすことがある音声信号受信場所の条件変化に適応する。

40

【0005】

しかしながら、適応フィルタは一般的に、大量の計算を実行してフィルタ性能を調整するので、多くの場合に、条件変化に合わせるのに時間がかかる。それゆえ、電話会議通信システムの設計者、製造者及び使用者は、音声信号受信場所の条件変化にさらに迅速に適応し、電話会議通信システム内の望ましくないエコーを効率的に相殺することができる音響エコーキャンセラが必要であることを認識している。

50

## 【発明の概要】

## 【0006】

本発明の種々の実施の形態は、音響エコーキャンセレーション機能を含む音声会議通信システムのための周波数領域符号器／復号器に関する。本発明の一実施の形態では、音響エコーキャンセラは、周波数領域符号器／復号器に組み込まれ、音声信号からの音響エコーを緩和するか又はその音響エコーを除去し、その音声信号は、周波数領域に変換されており、且つ周波数領域符号器／復号器によって部分帯域に分割されている。

## 【図面の簡単な説明】

## 【0007】

【図1A】1つの例示的な2地点音声会議通信システムの概略図である。

10

【図1B】2地点のうち的一方において音響エコーキャンセラを使用する、1つの例示的な2地点音声会議通信システムの概略図である。

【図2】周波数領域音声符号器の全体的な構造を示すブロック図である。

【図3】図2に示される周波数領域音声符号器において音声信号の周波数解析を実行するのに適しているフィルタバンクシステムを示す図である。

【図4】図2に示される周波数領域音声符号器と共に用いるのに適している周波数領域音声復号器の全体的な構造を示すブロック図である。

【図5】図4に示される周波数領域音声復号器において音声信号の周波数合成を実行するのに適しているフィルタバンクシステムを示す図である。

【図6】音響エコーキャンセラ及び周波数領域符号器／復号器を使用する、図1A及び図1Bに示される例示的な2地点音声会議通信システムの概略図である。

20

【図7】図6に示される、周波数領域符号器／復号器を基にする、例示的な2地点音声会議通信システムの部屋1のさらに詳細な概略図である。

【図8】1つの例示的な2地点電話会議通信システムの部屋1内の周波数領域符号器／復号器に組み込まれ、本発明の一実施形態を表す音響エコーキャンセラの概略図である。

【図9A】線形フィルタリングと、これに続く周波数解析の概略図である。

【図9B】図9A及び図9Bの出力が均等であるようにする、周波数解析と、これに続く部分帯域信号の線形フィルタリングの概略図である。

## 【発明を実施するための形態】

## 【0008】

30

本発明の一実施形態は、周波数領域符号器／復号器内に組み込まれ、音声会議通信システム内に含まれる音響エコーキャンセラに関する。音響エコーキャンセラは、音声信号受信場所において、1つ又は複数のラウドスピーカが1つ又は複数のマイクロフォンに結合されるときに引き起こされる音響エコーをキャンセルする。音声信号受信場所の条件変化によって、音声信号受信場所において結合したラウドスピーカとマイクロフォンとの間のインパルス応答に変化が生じ、それにより、音響エコーの特性に変化が生じる。音響エコーキャンセラ内の適応フィルタは、音声信号受信場所のインパルス応答を追跡し、インパルス応答推定値を生成する。そのインパルス応答推定値を用いて、音響エコーキャンセラにおいてエコー信号推定値が生成される。その後、エコー信号推定値は、音声信号受信場所にあるマイクロフォンから伝搬する信号から減算され、結果として生成される誤差信号が出力され、音声信号送信場所に戻される。

40

## 【0009】

適応フィルタは、音声信号を圧縮するために、音声信号の符号化及び復号化を実施するために用いられるものと同じ周波数解析及び合成演算を用いることによって、周波数領域において実施される。適応フィルタは、周波数領域符号器／復号器内で一連の相対的に平坦なスペクトルの部分帯域に分割された周波数領域音声信号を入力及び出力する。部分帯域信号は、全帯域音声信号の場合に典型的に用いられるサンプリングレートよりもはるかに低いサンプリングレートにおいてサンプリングされる。さらに、本発明の代替的な実施形態では、音響エコーキャンセラは、音響エコーキャンセラ内の周波数領域符号器／復号器の既存の雑音低減構成要素及び知覚符号化構成要素を組み込み、それにより、エコーキ

50

キャンセリング性能を向上させることができる。

【0010】

本発明は、次の3つのサブセクション、すなわち(1)音響エコーキャンセレーションの概説、(2)音声信号圧縮の概説、及び(3)本発明の周波数領域音響エコーキャンセラ実施形態において以下に説明される。

【0011】

[音響エコーキャンセレーションの概説]

1つ又は複数の場所において、1つ又は複数のマイクロフォンと1つ又は複数のラウドスピーカとの間に結合が生じるために、音声会議通信システムにおいて音響エコーが生じる。図1Aは、1つの例示的な2地点音声会議通信システムの概略図を示す。音声会議通信システム100は、2つの場所、すなわち部屋1の102及び部屋2の104を含む。通信媒体106及び108によって、部屋1の102と部屋2の104との間で音声信号が伝送される。音声信号は、マイクロフォン110及び112によって通信媒体に入力され、音声信号は、ラウドスピーカ114及び116において通信媒体から出力される。

【0012】

図1Aでは、部屋2の104内の音声信号源118が音声信号 $s_{out}(t)$ 120を生成する。下付き文字「out」は、本出願全体を通して種々の図面において、その信号が通信媒体の外部で伝送されていることを表すために、いくつかの異なる信号を参照する際に用いられ、一方、下付き文字「in」は、通信媒体内部で伝送される信号を参照する際に用いられる。表記「(t)」は、本出願全体を通して種々の図面において、その信号が時間の関数であることを表すために、いくつかの異なる信号を参照する際に用いられる。部屋1の102及び部屋2の104内で生じる音響信号を検討するときに、「(t)」は連続した(アナログ)時間を表す。デジタル伝送及びデジタル信号処理の場合に用いられるように、サンプリングされた信号を検討するときには、「(t)」はサンプリング周期 $T_s = 1/f_s$ の間隔(又は倍数)だけ離れて位置する離散した時点を表す。

【0013】

音声信号 $s_{out}(t)$ 120は、部屋2の104内で多数の経路をとる。経路のうちの一つは、直接経路によって、又は部屋2の104内の物体からの反射によって、マイクロフォン110によって受信される。音声信号 $s_{out}(t)$ 120が音声信号源118からマイクロフォン110の出力までに辿る種々の経路は、まとめて、部屋2の104のインパルス応答と呼ばれる。図1Aにおいて、部屋2の104のインパルス応答、 $g_{Room2}(t)$ 122は、音声信号源118からマイクロフォン110までを指している点線によって表される。インパルス応答 $g_{Room2}(t)$ 122は、部屋2の104内部の条件が変化するのに応じて変化することがある。変化の例は、人の移動、ドアの開閉、及び部屋2の104内の家具の位置変更を含む。例示を簡単にするために、インパルス応答 $g_{Room2}(t)$ 122は、一本の線として示されるが、一般的には、数多くの異なる方向を有する数多くの異なる音波経路の複雑な重ね合わせである。

【0014】

標準的な条件下において、室内の音波伝送は、線形システムとして十分にモデル化することができる。線形システムは数学的に畳み込み演算によって記述されることがよく知られている。したがって、音声信号 $x_{in}(t)$ 124、すなわちマイクロフォン110の出力は、以下に記述される音声信号 $s_{out}(t)$ 120とインパルス応答 $g_{Room2}(t)$ 122との間の畳み込みの結果である。図1Aにおいて、音声信号 $x_{in}(t)$ 124は、以下のように表すことができる。

【0015】

【数1】

$$x_{in}(t) = s_{out}(t) * g_{Room2}(t) = \int_{-\infty}^{\infty} s_{out}(\tau) g_{Room2}(t-\tau) d\tau$$

【0016】

10

20

30

40

50

ただし、 $s_{out}(t)$  120は音声信号源118によって出力される音声信号であり、 $g_{Room2}(t)$  122は部屋2の104のインパルス応答であり、 $x_{in}(t)$  124は通信媒体106に入力される信号であり、「 $*$ 」は連続時間畳み込みを表す。上記の例では、 $g_{Room2}(t)$  122は、線形であると仮定されるマイクロフォン応答と、部屋2 104のマルチパル伝送とを含む。

【0017】

部屋2の104内の音声信号 $x_{in}(t)$  124は、マイクロフォン110から、通信媒体106を経由して、部屋1の102内のラウドスピーカ114に送られる。音声信号 $x_{in}(t)$  124は、ラウドスピーカ114を通り(図1Aにおいて、部屋1の102内の音声信号「 $x_{out}(t)$ 」として示される)、その後、部屋1の102を通過して、マイクロフォン112まで進む。音声信号 $x_{in}(t)$  124がラウドスピーカ114からマイクロフォン112の出力 $y_{in}(t)$  126までに辿る一連の経路はまとめて、部屋1 102のインパルス応答と呼ばれる。図1Aにおいて、部屋1の102のインパルス応答、すなわち $h_{Room1}(t)$  128は、ラウドスピーカ114からマイクロフォン112まで指している点線によって表される。例示を簡単にするために、インパルス応答 $h_{Room1}(t)$  128は、一本の線として示されるが、一般的には、数多くの異なる方向及び反射を有する数多くの異なる音波経路の複雑な重ね合わせである。ラウドスピーカ及びマイクロフォンはいずれも線形システムであり、その応答特性は、部屋2の102のマルチパスインパルス応答と線形結合することができるものと仮定されることに留意されたい。マイクロフォン112から出力される音声信号は、エコー信号 $y_{in}(t)$  126であり、音声信号 $x_{in}(t)$  124とインパルス応答 $h_{Room1}(t)$  128との間の畳み込みの結果である。誰かが部屋1の102において話をしているときなどのように、部屋1の102において音声信号が生じるとき、その音声信号もマイクロフォン112によって拾われることに留意されたい。マイクロフォン112が、部屋2の104からの音声信号及び部屋1の102からの音声信号の両方から伝送している音を拾っているとき、この条件は「ダブルトーク」として知られている。ダブルトーク状態は一般的に、音響エコーキャンセラによって検出され、エコーキャンセレーションが一時中止される。数多くのダブルトーク検出アルゴリズムが、音響エコーキャンセラの技術分野において知られており、本発明のための制御機構の一部として適用することができる。

10

20

30

【0018】

部屋1の102において、マイクロフォン112によって拾われている音声信号が生じていないものと仮定するとき、エコー信号 $y_{in}(t)$  126は、以下の式によって表すことができる。

【0019】

【数2】

$$y_{in}(t) = x_{in}(t) * h_{Room1}(t) = \int_{-\infty}^{\infty} x_{in}(\tau) h_{Room1}(t-\tau) d\tau$$

【0020】

ただし、 $x_{in}(t)$  124はラウドスピーカ114に入力される音声信号であり、 $h_{Room1}(t)$  128は部屋1の102のインパルス応答であり、 $y_{in}(t)$  126は通信媒体108に入力される信号であり、「 $*$ 」は連続時間畳み込みを表す。

40

【0021】

エコー信号 $y_{in}(t)$  126は、マイクロフォン112から、通信媒体108を経由して、部屋2の104内のラウドスピーカ116に送られる。ラウドスピーカ116が、エコー信号 $y_{out}(t)$  130を出力する。音声信号源118が話をしている人であるとき、その人は、自分が依然として話をしている間に、自分の声の時間遅れのエコーを聞くことがある。遅延時間は、部屋1の102と部屋2の104とを隔てている距離、場所間のデジタル伝送の前後に音声信号を処理するために電話会議通信システム100によって使用される周波数領域符号器/復号器(図1Aには示されない)のような、付加的な信号処

50

理によって必要とされる時間の長さなどの複数の要因によって異なることがある。マイクロフォンによる音声信号の増幅、及びラウドスピーカとマイクロフォンとの間の距離にもよるが、マイクロフォン 110 に向かって話をしている人は、自分の声の遅延したエコーを聞くことがあり、ループ利得が十分に高いときには、うるさいほどのハウリング音を聞くこともある。音声信号  $y_{out}(t)$  130 は、マイクロフォン 110 によって受信されることがあり、それにより、音響エコーを除去するために何もなされなければ、音声会議通信システム 100 の中で音響エコーが無限に繰り返されることがある。

【0022】

図 1 B は、2つの場所のうち的一方において音響エコーキャンセラを使用する1つの例示的な2地点音声会議通信システムの概略図を示す。図 1 B において破線の長方形によって表される音響エコーキャンセラ 134 が、通信媒体 106 と相互接続される通信媒体 136 を経由して、サンプリングされた音声信号  $x_{in}(t)$  124 を受信する。図 1 B において、音響エコーキャンセラは、アナログシステムとして現れる。しかしながら、電話会議通信システムのための適応フィルタは、通常、有限インパルス応答デジタルフィルタである。有限応答デジタルシステムの場合、音声信号は一般的にサンプリングされ、畳み込みは一般的に数値計算によって実行される。サンプリング及び数値計算は、たとえば、部屋 1 の 102 内のアナログ/デジタルコンバータを用いて、 $y_{in}(t)$  126 をサンプリングし、離散時間バージョンを生成することによって果たすことができる。同様に、部屋 2 の 104 内のアナログ/デジタルコンバータを用いて、離散時間バージョンの信号  $x_{in}(t)$  124 を生成することができる。図 1 B において、デジタル/アナログコンバータを用いて、 $x_{in}(t)$  124 をアナログ信号に変換し、ラウドスピーカ 114 に入力することができる。アナログ/デジタルコンバータ及びデジタル/アナログコンバータは図 1 B には示されないが、上記の説明において、図 1 B 内の信号は、適切なサンプリングレートにおいてサンプリングされること、部屋 1 の 102 と部屋 2 の 104 との間でデジタル伝送が用いられること、並びにエコーキャンセレーションを実施するために、デジタルフィルタリングが用いられることが仮定される。

10

20

【0023】

音響エコーキャンセラ 134 は、適応フィルタ 138 と、加算接合部 140 とを備える。適応フィルタ 138 は、2つの入力を介して、信号を受信する。第 1 の入力、通信媒体 136 を介して、音声信号  $x_{in}(t)$  124 を受信し、第 2 の入力、通信媒体 142 を介して、フィードバック信号、すなわち音響エコーキャンセラ 134 から出力される信号を受信する。適応フィルタ 138 は、2つの入力信号に含まれる情報を用いて、インパルス応答推定値

30

【0024】

【数 3】

$$\hat{h}_{Room1}(t)$$

【0025】

144 を生成し、部屋 1 の 102 内の条件変化に応じてインパルス応答  $h_{Room1}(t)$  128 が変化するとき、そのインパルス応答推定値は、インパルス応答  $h_{Room1}(t)$  128 を追跡するために調整される。音声信号  $x_{in}(t)$  124 が、音響エコーキャンセラ 134 によって、インパルス応答推定値

40

【0026】

【数 4】

$$\hat{h}_{Room1}(t)$$

【0027】

142 と畳み込まれ、離散畳み込み

50

【 0 0 2 8 】

【 数 5 】

$$\hat{y}_{in}(t) = x_{in}(t) * \hat{h}_{Room1}(t) = \sum_{\tau=0}^M \hat{h}_{Room1}(\tau) x_{in}(t-\tau).$$

【 0 0 2 9 】

によって、エコー信号推定値

【 0 0 3 0 】

【 数 6 】

10

$$\hat{y}_{in}(t)$$

【 0 0 3 1 】

1 4 6 が生成される。エコー信号推定値

【 0 0 3 2 】

【 数 7 】

$$\hat{y}_{in}(t)$$

20

【 0 0 3 3 】

1 4 6 は、通信媒体 1 4 8 を経由して、加算接合部 1 4 0 に送られ、加算接合部 1 4 0 には、通信線 1 5 0 を介して、マイクロフォン 1 1 2 から、エコー信号  $y_{in}(t)$  1 2 6 も入力される。加算接合部 1 4 0 は、エコー信号  $y_{in}(t)$  1 2 6 から、エコー信号推定値

【 0 0 3 4 】

【 数 8 】

$$\hat{y}_{in}(t)$$

【 0 0 3 5 】

1 4 6 を減算し、誤差音声信号  $e_{in}(t)$  1 5 2、すなわち部屋 2 の 1 0 4 に伝送されることになる信号

【 0 0 3 6 】

【 数 9 】

30

$$e_{in}(t) = y_{in}(t) - \hat{y}_{in}(t) = x_{in}(t) * h_{Room1}(t) - x_{in}(t) * \hat{h}_{Room1}(t)$$

【 0 0 3 7 】

を生成する。誤差音声信号  $e_{in}(t)$  1 5 2 は、通信線 1 5 4 を介して、ラウドスピーカ 1 1 6 に送られ、部屋 2 の 1 0 4 に誤差音声信号  $e_{out}(t)$  1 5 6 として出力される。

40

インパルス応答推定値

【 0 0 3 8 】

【 数 1 0 】

$$\hat{h}_{Room1}(t)$$

【 0 0 3 9 】

1 4 4 がインパルス応答  $h_{Room1}(t)$  1 2 8 に十分に近いとき、誤差音声信号  $e_{in}(t)$  1 5 2 の大きさは小さく、部屋 2 の 1 0 4 内に音響エコーはほとんど伝送されない。ダブルトークの状況中には、線形性によって、誤差信号は部屋 1 の 1 0 2 内の人の発話信号

50

(図1Bには示されない)も含み、これが適応フィルタ138の発散を引き起こすことがあるので、適応フィルタ138の適応を一時中止する必要があることに留意されたい。音響エコーキャンセラ134は、最新の導出された

【0040】

【数11】

$$\hat{h}_{Room1}(t)$$

【0041】

144を用いて、部屋2の104内の音声信号源118によって生成された音響エコーのキャンセルを試み続けることができるが、システムは全二重動作を利用するので、部屋1102内の人の発話(図1Bには示されない)はそれでも、部屋2の104に伝送される。

【0042】

フィルタ係数値

【0043】

【数12】

$$\hat{h}_{Room1}(t)$$

【0044】

144(ただし、 $t = 0, 1, 2, \dots, M$ )は、離散時間フィルタの特性を決定する。適応フィルタの場合、それらの係数は時間と共に調整される。フィルタ係数は、最小平均二乗アルゴリズム(「LSM」)又はアフィン投影のような、当該技術分野においてよく知られている技法を用いて導出される。そのようなアルゴリズムを用いて、適応フィルタ138のフィルタ係数を絶えず適応させて、インパルス応答推定値

【0045】

【数13】

$$\hat{h}_{Room1}(t)$$

【0046】

144を部屋1102のインパルス応答 $h_{Room1}(t)$ 128に近づけることができる。図1Bを参照しながら先に説明されたように、通信媒体142によって、適応フィルタ138にフィードバックが与えられ、通信媒体142は通信媒体154と接続し、誤差音声信号 $e_{in}(t)$ 152のための最新値を適応フィルタ138に戻す。

【0047】

図1Bを参照しながら説明された音響エコーキャンセラは、部屋2の104から生じる音声信号から導出される音響エコーをキャンセルようにだけ動作することに留意されたい。大部分の双方向の会話では、音声信号は、それぞれの場所において送信され、受信される。部屋1の102から生じる音響エコーをキャンセルするために、一般的には、部屋2の104において、第2の音響エコーキャンセラが使用される。

【0048】

[音声信号圧縮の概説]

音声会議通信システムを含む、デジタル電気通信技術の主な要素は、データを記憶すること及び場所間でデータを転送することである。データの記憶及び伝送は費用がかかり、時間を要することがあるので、記憶又は伝送前にデータを圧縮することによって、データ

10

20

30

40

50

をより効率的に格納し、伝送するために、種々の技法が生み出されてきた。圧縮されたデータの個々のユニットは一般的に、直にアクセスすることはできない。圧縮されたデータの伝送及び記憶は、より効率的であるが、データの個々のユニットにアクセスするには、圧縮されたデータが解凍される必要がある。

#### 【0049】

圧縮技法は一般的に、非可逆圧縮及び可逆圧縮に分けられる。非可逆圧縮は、可逆圧縮によって達成される圧縮比よりも高い圧縮比を達成するが、非可逆圧縮は、後に解凍する結果として、情報が失われる。音声信号の場合、圧縮/解凍される音声信号が聞き取れるほど劣化するのを避けるために、非可逆圧縮/解凍サイクルから生じるデータ損失は巧みに処理される必要がある。人の聴覚系の固有の限界を使用することによって、音質を犠牲にすることなく、音声信号を圧縮及び解凍することができる。知覚現象は多くの場合に、周波数領域において最もよく理解され、表現されるので、高品質音声符号化システムの大部分は、周波数解析を伴う。

10

#### 【0050】

図2は、周波数領域音声符号器の全体的な構造を示すブロック図を示す。ブロック図200は、単一のサンプリングされた時間波形 $x(t)$ 202を、時間及び周波数の両方の関数であるデジタルデータストリームに符号化するための過程を示す。そのような音声符号化システムのいくつかの例は、MPEG-2及びAACを含む。図2では、時間波形 $x(t)$ 202は、「周波数解析」を付されるブロック204に入力されるように示される。周波数解析ブロック204は、入力時間波形 $x(t)$ 202の時間と共に変化する周波数解析を得る。時間シフトブロック変換又はフィルタバンクを用いて、時間と共に変化する周波数解析を実行することができる。たとえば、フィルタバンクが利用されるとき、フィルタバンクは、各時刻 $t$ においてベクトル時間信号 $X_{sub}(k, t)$ 206(ただし、 $k = 0, 1, 2, \dots, N - 1$ )を形成する集合的な1組 $N$ 個の出力を出力する。下付き文字「sub」は、図2及び後続の図面においていくつかの異なる信号を参照する際に、その信号が部分帯域を集めたものであることを表すために用いられる。図2において、ベクトル信号 $X_{sub}(k, t)$ 206は、太い矢印として表される。図2及び後続の図面において、時間及び周波数の両方の関数である信号は、太い矢印として示される。

20

#### 【0051】

ベクトル信号 $X_{sub}(k, t)$ 206は、「Q」を付されたブロック208に入力され、そのブロックにおいて、ベクトル信号 $X_{sub}(k, t)$ 206は、量子化及び符号化されて、信号 $X_{in}(k, t)$ 210として出力される。特定の周波数の音が、近傍周波数の大きな音によって聞き取れなくなることがあること、すなわち「マスクされる」ことがあることが、信号処理の分野において十分に確立されている。図2において、時間波形 $x(t)$ 202が、「知覚モデル」を付されるブロック212に入力され、ブロック212は、補助的なきめの細かいスペクトル解析を用いて、マスク効果を計算し、周波数解析の量子化を導く。この音声知覚のモデルを用いて、知覚することができない周波数成分は、数ビット又は0ビットを与えられ、一方、最も知覚することができる周波数成分は、最大のビットを与えられる。

30

#### 【0052】

図3は、図2に示される周波数領域音声符号器において音声信号の周波数解析を実行するのに適しているフィルタバンクシステムを示す。図3において、時間波形 $x(t)$ 202が示されており、フィルタバンク300に入力され、ベクトル時間信号 $X_{sub}(k, t)$ 206(ただし、 $k = 0, 1, 2, \dots, N - 1$ )を形成する集合的な1組 $N$ 個の出力として出力される。フィルタバンク300は、 $N$ 個のバンドパスフィルタ $G_k$ 304を含み、その中心周波数は $f_k$ であり、その通過帯域は、表現されるべき音声周波数の所望の帯域を含む。図3は、 $N = 4$ の場合を示すが、典型的な値は一般的に $N = 32$ 以上である。バンドパスフィルタ304の出力 $x_k(t)$ 306は、サンプル/秒の合計数が一定のままであるように、 $N$ 分の1にダウンサンプリングされている(308)時間信号である。

40

50

## 【 0 0 5 3 】

一般的に、2つのタイプのマスキング、すなわち(1)空間的マスキング及び(2)時間的マスキングが考えられる。空間的マスキングでは、低い強度の音が、同時に生じている高い強度の音によってマスクされる。2つの音の周波数が近いほど、低い強度の音をマスクするのに要する音の強度の差が小さくなる。時間的マスキングは、低い強度の音が、高い強度の音の伝送直前又は直後に伝送されるときに、低い強度の音が高い強度の音によってマスクされる。2つ音の時間が近いほど、低い強度の音をマスクするのに要する音の強度の差が小さくなる。

## 【 0 0 5 4 】

通常、周波数領域符号化システムは、対応する周波数領域復号化システムを有する。図4は、図2において示される周波数領域音声符号器と共に用いるのに適している周波数領域音声復号器の全体的な構造を示すブロック図を示す。図4において、信号 $X_{in}(k, t)$  402が、「 $Q^{-1}$ 」を付されたブロック404に入力され、ブロック404は、符号化されたデジタル信号を取り込み、そのデータを変換して、周波数合成のための1組の適切な入力に戻す。図4において、周波数領域符号化信号 $X_{sub}(k, t)$  406(ただし、 $k = 0, 1, 2, \dots, N - 1$ )が、 $Q^{-1}$ ブロック404から出力され、「周波数合成」を付されたブロック406に入力され、そのブロック406において、信号 $X_{sub}(k, t)$  406(ただし、 $k = 0, 1, 2, \dots, N - 1$ )は、サンプリングされた音声時間波形 $x(t)$  410に再構成される。

10

## 【 0 0 5 5 】

図5は、図4に示される周波数領域音声復号器において音声信号の周波数合成を実行するのに適しているフィルタバンクシステムを示す。集合的な1組の信号 $X_{sub}(k, t)$  406(ただし、 $k = 0, 1, 2, \dots, N - 1$ )がアップサンプリングされ(502)、 $N$ 個のバンドパスフィルタ $G_k$  504を通じて送られ、その中心周波数は $k$ であり、その通過帯域は、表現されるべき音声周波数の所望の帯域を含む。その出力 $x_k(t)$  506は合算され(508)、サンプリングされた音声時間波形 $x(t)$  410が再構成される。バンドパスフィルタ504を適切に設計し、元の周波数解析データを細かく量子化することによって、サンプリングされた音声時間波形 $x(t)$  410は、ごくわずかな量の誤差しか含むことなく、再構成することができる。

20

## 【 0 0 5 6 】

[ 本発明の周波数領域音響エコーキャンセラ実施形態 ]

デジタル伝送を使用する音声会議通信システムでは、MPEG2及びAACに基づく周波数領域符号器/復号器のような、周波数領域符号器/復号器を用いることによって音声信号を圧縮することにより、高品質の音声伝送のために必要とされるビットレートを低減するのが一般的である。伝送前に、音声信号は最初に周波数領域符号器を通じて送られ、その後、受信時に、周波数領域復号器を通じて送られる。周波数領域符号器は、音声信号を送信する前に、送出される音声信号を圧縮されたデジタル音声信号に変換し、周波数領域復号器は、圧縮されている受信デジタル音声信号を解凍して、ラウドスピーカに送ることができるアナログ音声信号を復元する。

30

## 【 0 0 5 7 】

図6は、音響エコーキャンセラ及び周波数領域符号器/復号器を使用する、図1A及び図1Bに示される例示的な2地点電話会議通信システムの概略図である。部屋2の104内の周波数領域符号器602は、音声信号源118から生じる音声信号をデジタル化及び圧縮し、圧縮されたデジタル音声信号を、部屋1の102内の周波数領域復号器604に送信する。周波数領域復号器604は、圧縮されている受信デジタル音声信号を解凍することによって、アナログ音声信号を復元し、復元された音声信号は、離散時間形式において、適応フィルタ138に送られ、ラウドスピーカ114に送る前に、アナログ形式に変換される。エコー推定値信号

40

## 【 0 0 5 8 】

【数 1 4】

$$\hat{y}_{in}(t)$$

【0059】

146 がエコー信号  $y_{in}(t)$  126 から減算され、結果として生成された誤差音声信号  $e_{in}(t)$  152 が、部屋 1 102 内の周波数領域符号器 606 に送られる。誤差音声信号  $e_{in}(t)$  152 は、デジタル化及び圧縮され、部屋 2 の 104 内の周波数領域復号器 608 に送信され、そこで、誤差音声信号  $e_{in}(t)$  152 は、離散時間信号に復元され、アナログ形式に変換され、ラウドスピーカ 116 に送られる。

10

【0060】

図 7 は、図 6 に示される、周波数領域符号器 / 復号器に基づく例示的な 2 地点音声会議通信システムの部屋 1 のさらに詳細な概略図を示す。部屋 1 の 102 において点線の長方形として示される周波数領域符号器 / 復号器 700 は、周波数領域符号器 702 と、周波数領域復号器 704 とを備える。周波数領域符号器 702 は、音声信号が部屋 2 に送信される前に、音声信号をデジタル化及び圧縮し、周波数領域復号器 704 は、圧縮されている受信デジタル音声信号を解凍することによって、部屋 2 から受信される音声信号を復元する。

【0061】

図 2 において先に示されたように、図 7 に示される周波数領域符号器 702 は、周波数解析ステージ 706 及び量子化器 708 を備えており、量子化器は知覚モデル（図 7 には示されない）によって制御される。周波数解析ステージ 706 は、バンドパスフィルタの阵列、すなわち図 3 に示されるフィルタバンクに類似のフィルタバンクを使用することによって、入力音声信号を周波数領域に変換し、入力音声信号を、太い矢印としてまとめて示される、複数の類似の帯域制限された信号 710、すなわち部分帯域に分離する。各部分帯域は、入力音声信号の周波数範囲全体の周波数サブセットを含む。各部分帯域 710 内の分離された周波数成分は量子化器 708 に送られ、そこで、部分帯域は量子化及び符号化される。部分帯域は、量子化誤差が強い音声信号成分によってマスクされるように量子化される。図 2 に示されるように、音声信号内の情報ビットを捨てるために、知覚符号化が用いられており、知覚符号化は、信号が単一の音声波形に再構成されるときに、聞き取られる歪みを増大させることなく、音声信号のデータ速度を低減するように設計される。図 7 に示される概略図を簡単にするために、知覚モデル計算を省略した。しかしながら、量子化器を制御するために、通常知覚モデル計算が用いられる。可変ビット割当てを用いて信号が符号化され、一般的に、人の聴覚が最も敏感である中央の周波数範囲において、サンプル当たり、より多くのビットが用いられ、中央の周波数範囲において、より細かい分解能が与えられている。

20

30

【0062】

その後、圧縮されたデジタル音声信号は、部屋 2 内の周波数領域復号器に送信され、そこで、圧縮された音声信号は復元されることができ。部屋 1 の 102 では、復号器 704 が部屋 2 からの圧縮された入力音声信号に関して逆演算を実行する。復号器 704 は、逆量子化器 712 を備えており、逆量子化器において、量子化されている受信音声信号が逆量子化され、適切な共通振幅スケールにおいて、まとめて太い矢印として示される部分帯域 716 が生成される。部分帯域は周波数合成ステージ 714 に送られ、そこで、部分帯域は、たとえば、図 5 に示されるように、元の周波数帯域の場所にアップサンプリングすることによって周波数シフトされ、フィルタバンクを通じて送られ、単一の音声波形に合算され、変換されて、時間領域に戻される。解析及び合成フィルタバンク、並びに周波数領域符号器 / 復号器によって実行される圧縮及び解凍ルーチンは、電話会議通信システムの中に遅延をもたらすことに留意されたい。

40

【0063】

50

本発明の種々の実施形態は、音響エコーキャンセラ機能を含む音声会議通信システムのための周波数領域符号器／復号器に向けられる。音声会議通信システムに組み込まれる周波数領域符号器／復号器において一連の部分帯域に分割されるときに、音響エコーがキャンセルされる。畳み込みは線形演算であり、周波数解析及び周波数合成ステージも線形演算を利用するので、音響エコーキャンセレーションは、周波数領域において実行することができる。音響エコーキャンセレーションを周波数領域符号器／復号器に組み込むことによって、音響エコーキャンセレーションを周波数領域において実行することができ、その際に、音響エコーキャンセラのための冗長な音声信号変換装置を設ける必要はない。

【0064】

本発明では、音響エコーキャンセラは、部分帯域が音声会議通信システム内の周波数領域復号器内にある間に、一連の部分帯域に分割される音声信号を受信する。音響エコーキャンセラは、音声会議通信システム内の周波数領域符号器に一連の部分帯域を出力する。図8は、1つの例示的な2地点電話会議通信システムの部屋1内の周波数領域符号器／復号器に組み込まれ、本発明の一実施形態を表す音響エコーキャンセラの概略図を示す。部屋1の800は、点線の長方形として表される周波数領域符号器／復号器802と、ラウドスピーカ804と、マイクロフォン806とを備える。周波数領域符号器／復号器802は、周波数領域符号器808と、周波数領域復号器810と、破線の長方形によって表される音響エコーキャンセラ812とを備える。部屋2から入ってくる圧縮されたデジタル音声信号  $X_{in}(k, t)$  814が、周波数復号器810に入力される。デジタル音声信号  $X_{in}(k, t)$  814、すなわち周波数領域の音声信号は圧縮され、逆量子化器816によって受信され、部分帯域信号  $X_{sub}(k, t)$  818として図8において示される、一連の部分帯域信号に変換される。

【0065】

音声信号  $X_{sub}(k, t)$  818は2つの場所、すなわち周波数合成ステージ820及び音響エコーキャンセラ812に出力される。周波数合成ステージ820は、音声信号  $X_{sub}(k, t)$  818を音声信号  $x_{in}(t)$  822に変換する。音声信号  $X_{sub}(k, t)$  818は、再構成された1組のバンドパスフィルタ出力であり、音声信号  $x_{in}(t)$  822は、単一の離散時間領域信号であることに留意されたい。音声信号  $x_{in}(t)$  822は、周波数領域復号器810から出力され、デジタル／音声変換器(図8には示されない)を通じて送られ、その後、ラウドスピーカ804に送られ、部屋1の700の中に音響信号  $x_{out}(t)$  823として送出される。マイクロフォン806の出力はエコー信号  $y_{in}(t)$  826であり、これは、音声信号  $x_{in}(t)$  822とインパルス応答  $h_{Room1}(t)$  824との畳み込みである。エコー信号  $y_{in}(t)$  826は、周波数領域符号器808に入力され、周波数解析ステージ828によって変換され、一連の部分帯域、すなわちエコー信号  $Y_{sub}(k, t)$  830に分割され、そしてN個の部分帯域信号のベクトル減算を表す加算接合部832に送られる。

【0066】

音響エコーキャンセラ812は、音声信号  $X_{sub}(k, t)$  818を受信し、部分帯域信号に1組のフィルタを適用する。1組のフィルタは、図8において、フィルタリング行列

【0067】

【数15】

$$\hat{H}_{Room1}$$

【0068】

を付されたブロック834によって表される。フィルタリング行列

【0069】

10

20

30

40

【数 16】

$$\hat{H}_{Room1}$$

【0070】

834は、図1Bを参照しながら先に説明された

【0071】

【数17】

$$\hat{y}_{in}(t) = x_{in}(t) * \hat{h}_{Room1}(t)$$

10

【0072】

の演算に等価である。フィルタリング行列

【0073】

【数18】

$$\hat{H}_{Room1}$$

【0074】

834によって表されるフィルタは、音声信号  $X_{sub}(\omega_k, t)$  818に適用されて、エコー信号推定値

20

【0075】

【数19】

$$\hat{Y}_{sub}(\omega_k, t)$$

【0076】

838が生成され、これはフィルタリング行列

【0077】

【数20】

30

$$\hat{H}_{Room1}$$

【0078】

834から出力され、ベクトル加算接合部832によって受信される。エコー信号推定値

【0079】

【数21】

$$\hat{Y}_{sub}(\omega_k, t)$$

40

【0080】

838がエコー信号  $Y_{sub}(\omega_k, t)$  830から減算され、誤差音声信号  $E_{sub}(\omega_k, t)$  840が生成され、これが、フィードバックを与えるために適応フィルタ834に戻され、また量子化器842にも送られ、ここで、誤差音声信号  $E_{sub}(\omega_k, t)$  840は量子化され、その結果が、 $E_{in}(\omega_k, t)$  844として表される。誤差音声信号  $E_{in}(\omega_k, t)$  844は、周波数領域符号器808から出力され、部屋2に伝送される。

【0081】

誤差信号の量子化は知覚モデルによって導かれる。部屋2からの信号が存在しない場合に、信号  $y_{in}(t)$  826はまさに、部屋2に送られることになる所望の信号であるので、知覚モデルは一般的に、信号  $y_{in}(t)$  826から計算される高分解能スペクトルによ

50

って制御される。したがって、信号  $y_{in}(t)$  826 は、正確に量子化され、符号化される必要がある。部屋 1 において誰も話をしていない場合には、信号  $E_{sub}(k, t)$  840 はキャンセルされることが望ましいエコーを表すので、信号  $E_{sub}(k, t)$  840 を正確に量子化することは重要ではなくなる。この場合、誤差信号  $E_{sub}(k, t)$  840 は、信号  $y_{in}(t)$  826 を減衰させて、フィルタリングしたものである。それでも、信号  $y_{in}(t)$  826 に基づく知覚モデルを使用することは妥当である。図 8 に示される量子化演算は、音声会議信号の品質を高めるためのさらに別の機会をもたらす。部分帯域信号に関する音響エコーキャンセレーションの技術分野において、量子化過程の一部としてよく知られている非線形エコー抑圧技法を実施することによって、残留音響エコーのさらなるマスキングを組み込むことができる。

10

【0082】

線形フィルタリングの前後いずれかにおいて、周波数解析を実行することができる。図 9 A は、線形フィルタリングと、それに続く周波数解析の概略図を示す。図 9 A では、周波数解析は、畳み込み

【0083】

【数22】

$$\hat{y}_{in}(t) = x_{in}(t) * \hat{h}_{Room1}(t)$$

【0084】

の後に実行され、部分帯域信号

20

【0085】

【数23】

$$\hat{Y}_{sub}(\omega_k, t)$$

【0086】

が得られる。図 9 B は、図 9 A 及び図 9 B の出力が等価になるように、周波数解析の後に部分帯域信号の線形フィルタリングを実施する概略図を示す。C. A. Lanciani 及び R. W. Schafer 著「Psychoacoustically-based processing of MPEG-1 layer 1-2 signals」(IEE First Workshop on Multimedia Signal Processing, June 1997, pp 53-58)、並びに C. A. Lanciani 及び R. W. Schafer 著「Subband-domain filtering of MPEG audio signals」(Proc. IEEE ICASSP '99, vol. 2, March 1999, pp 917-920) において、Lanciani 及び Schafer は、周波数解析が線形フィルタリング前に実行されるときに、部分帯域信号に適用することができる 1 組のバンドパスフィルタを見つけることができることを示した。フィルタリング行列

30

【0087】

【数24】

$$\hat{H}_{Room1}$$

40

【0088】

によって表される、この 1 組の線形フィルタを求めることは、図 9 B において示される線形フィルタを実現するのに重要である。フィルタリング行列

【0089】

【数25】

$$\hat{H}_{Room1}$$

【0090】

50

に  $X_{sub}(\omega_k, t)$  が入力されるとき、図 9 B において得られる

【 0 0 9 1 】

【 数 2 6 】

$$\hat{Y}_{sub}(\omega_k, t)$$

【 0 0 9 2 】

が図 9 A において示される結果と等価になるように、フィルタリング行列

【 0 0 9 3 】

【 数 2 7 】

10

$$\hat{H}_{Room1}$$

【 0 0 9 4 】

を調整することができる。

【 0 0 9 5 】

一般的に、図 9 B の出力信号が図 9 A の出力信号に等価になる場合、

【 0 0 9 6 】

【 数 2 8 】

20

$$\hat{Y}_{sub}(\omega_k, t)$$

【 0 0 9 7 】

の各個別の部分帯域は、解析 / 合成フィルタバンクシステムのエイリアスキャンセレーション特性を保持するために、 $X_{sub}(\omega_k, t)$  の全ての部分帯域に依存する。しかしながら、C. A. Lanciani 及び R. W. Schafer 著「Subband-domain filtering of MPEG audio signals」(Proc. IEEE ICASSP '99, vol. 2, March 1999, pp 917-920) において、Lanciani 及び Schafer は、音声符号器において用いられるタイプのフィルタバンクの場合、隣接する部分帯域の影響しか含む必要がないことを示した。フィルタリング行列

【 0 0 9 8 】

【 数 2 9 】

30

$$\hat{H}_{Room1}$$

【 0 0 9 9 】

を含むインパルス応答は、音響エコーキャンセレーションの技術分野においてよく知られている技法を用いて適応させることができ、バンドパスフィルタが音声信号のサンプリングレートの  $1/N$  倍であるサンプリングレートにおいて動作するという利点、及び部分帯域信号が、その制限された周波数帯域にわたって比較的平坦なスペクトルを有するという利点がある。

【 0 1 0 0 】

電話会議通信システム内の周波数領域符号器 / 復号器によって実行される音声信号処理を用いて、音声信号が異なる場所に伝送される前に、音声信号内の可聴背景雑音の量を低減することもできる。1つの手法は、ウィーナタイプのフィルタリングを用いることである。ウィーナフィルタは、各信号の周波数スペクトルに基づいて信号を分離する。ウィーナフィルタは、主に音声信号を含む周波数を通し、主に雑音を含む周波数を遮断する。さらに、各周波数におけるウィーナフィルタの利得は、各周波数における音声信号及び雑音の相対的な量によって決定される。ウィーナフィルタは、音声信号と共に、信号対雑音比を最大にする。ウィーナタイプのフィルタリングを使用するために、信号は周波数領域内にある必要があり、且つ現在の場所内の雑音スペクトルがわかっている必要があり、それ

40

50

により、ウィーナフィルタの周波数応答を計算することができる。本発明の現在の実施形態では、音響エコーキャンセラの適応フィルタを使用して、周波数領域符号器／復号器が配置される場所における雑音スペクトルを推定することによって、音声信号においてウィーナタイプのフィルタリングを実行し、音声信号が別の場所に伝送される前に、雑音を低減することができる。

【0101】

本発明は、特定の実施形態に関して説明されてきたが、本発明がこの実施形態に限定されることを意図していない。本発明の精神の中にある変更が、当業者には明らかになるであろう。たとえば、電話会議通信システム内の場所の数は、2つよりも多くの数にすることができる。例示を明確にするために、上記の説明における例の多くにおいて、2つの場所が説明される。各場所において用いられるマイクロフォン及びラウドスピーカの数も変更することができる。例示を明確にするために、上記の説明における例の多くにおいて、1つのマイクロフォン及び1つのラウドスピーカが用いられる。各場所において、多数のマイクロフォン及び／又は多数のラウドスピーカを用いることができる。多数のマイクロフォン及び多数のラウドスピーカを有する場所の場合のインパルス応答はさらに複雑になることがあり、それに応じて、フィルタリング係数を調整して、音声信号受信場所のインパルス応答の変化に適応フィルタを適応させるために、さらに多くの計算が実行される必要があることに留意されたい。

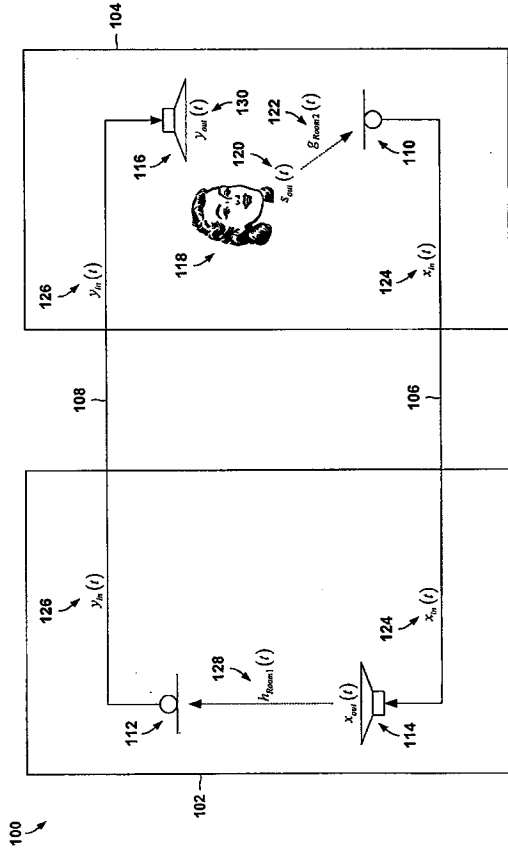
10

【0102】

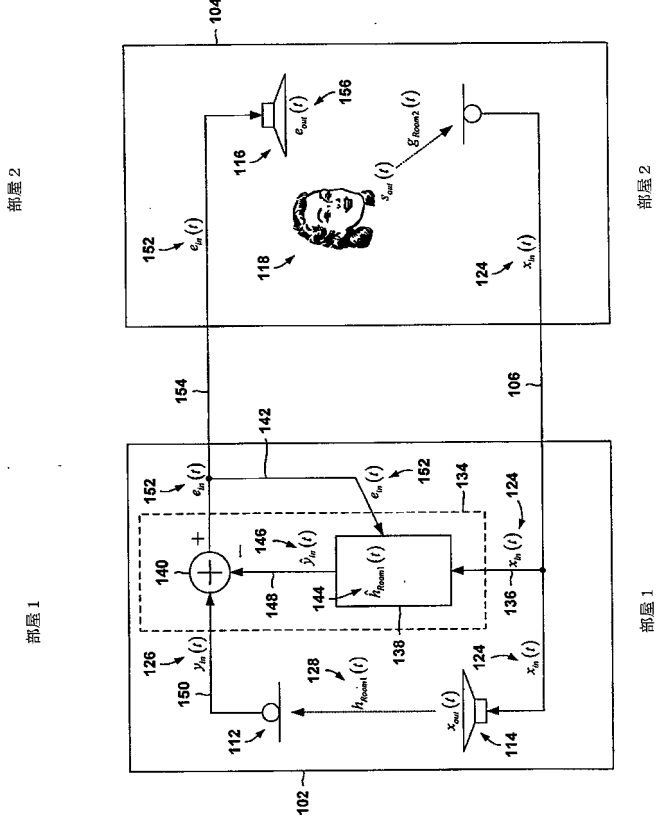
これまでの詳細な説明は、本発明を完全に理解してもらうために、例示するのを目的として、特有の用語を使用した。しかしながら、本発明を実施するのに、具体的な細部が不要であることは当業者には明らかであろう。したがって、本発明の具体的な実施形態のこれまでの説明は、例示し、説明するために提示される。それらの実施形態は、本発明を余す所なく述べることや、本発明を開示されるのと全く同じ形態に限定することは意図していない。上記の教示に鑑みて、数多くの変更及び変形が可能であることは明らかである。それらの実施形態は、本発明の原理及びその実用的な用途を最もわかりやすく説明し、それにより、当業者が、意図している特定の用途に相応しいように、本発明及び種々の実施形態に種々の変更を加えて最大限に利用することができるようにするために選択され、説明された。

20

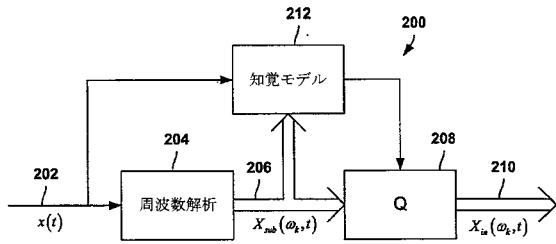
【図 1 A】



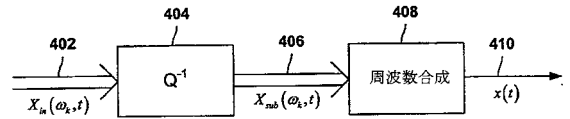
【図 1 B】



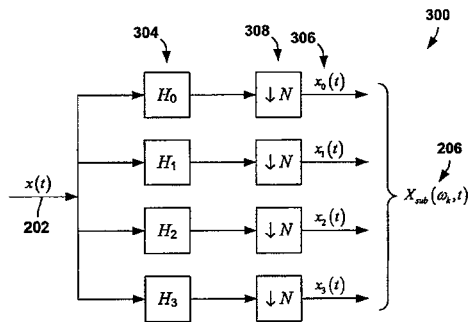
【図 2】



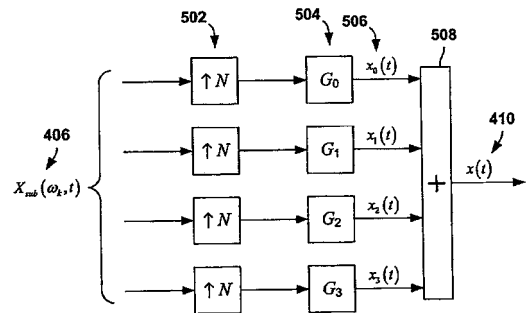
【図 4】



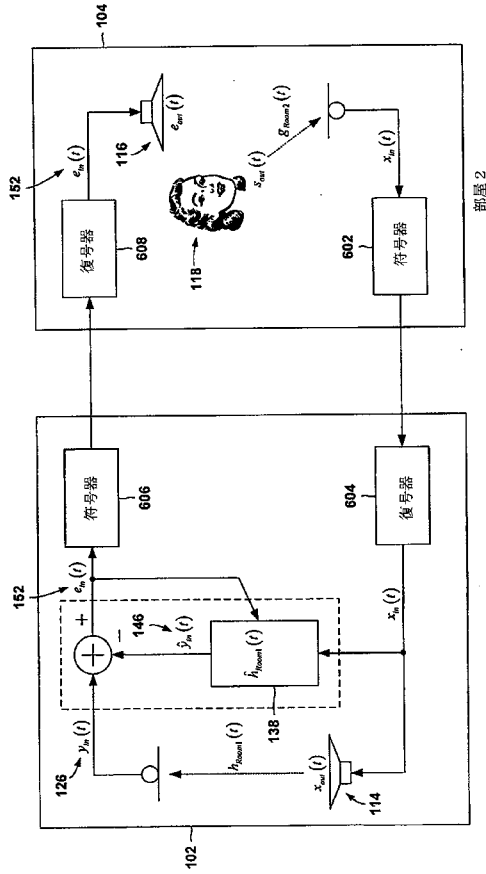
【図 3】



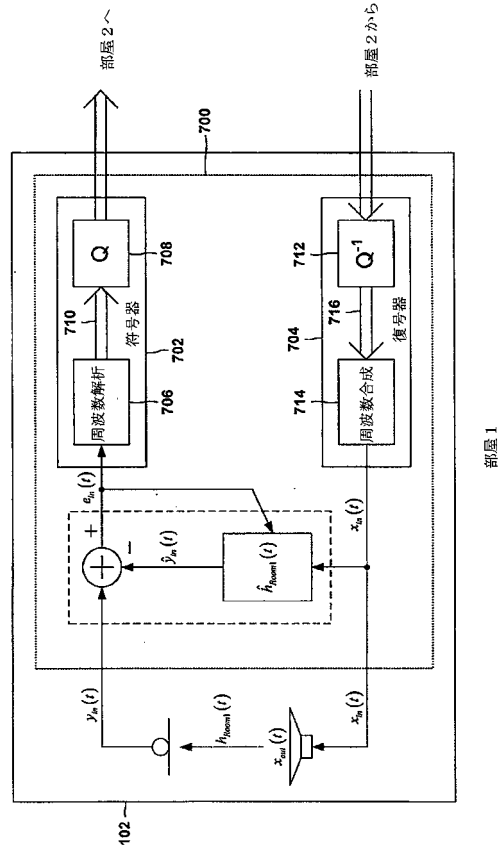
【図 5】



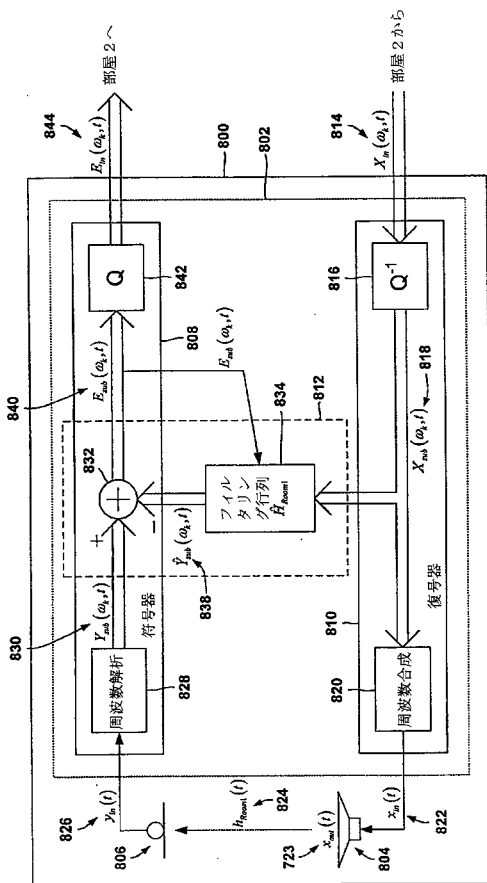
【図6】



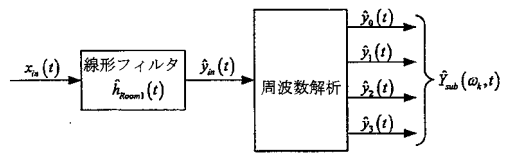
【図7】



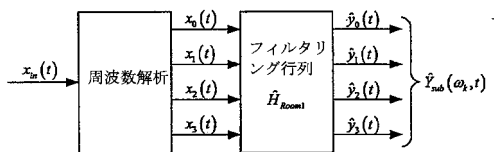
【図8】



【図9A】



【図9B】



【 国際調査報告 】

## INTERNATIONAL SEARCH REPORT

International application No  
PCT/US2007/021814

A. CLASSIFICATION OF SUBJECT MATTER INV. G10L21/02 G10L19/02		
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) G10L		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the International search (name of data base and, where practical, search terms used) EPO-Internal, INSPEC, WPI Data		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KELLERMANN W ED - VANDEWALLE J ET AL: "ON THE INTEGRATION OF SUBBAND ECHO CANCELLATION INTO SUBBAND CODINGSCHMES" SIGNAL PROCESSING THEORIES AND APPLICATIONS. BRUSSELS, AUG. 24 - 27, 1992; [PROCEEDINGS OF THE EUROPEAN SIGNAL PROCESSING CONFERENCE (EUSIPCO)], AMSTERDAM, ELSEVIER, NL, vol. 1, 24 August 1992 (1992-08-24), pages 123-126, XP000348629 ISBN: 978-0-444-89587-5 *Sections 1, 2* figure 1	1-10
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C.		<input checked="" type="checkbox"/> 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		*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
*E* earlier document but published on or after the international filing date		*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
*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)		*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
*O* document referring to an oral disclosure, use, exhibition or other means		*Z* document member of the same patent family
*P* document published prior to the international filing date but later than the priority date claimed		
Date of the actual completion of the international search	Date of mailing of the international search report	
14 May 2008	29/05/2008	
Name and mailing address of the ISA European Patent Office, P.B. 6818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Authorized officer  Bensa, Julien	

## INTERNATIONAL SEARCH REPORT

International application No

PCT/US2007/021814

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ENEROTH P: "Joint filterbanks for echo cancellation and audio coding" IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING, IEEE SERVICE CENTER, NEW YORK, NY, US, vol. 11, no. 4, 1 July 2003 (2003-07-01), pages 342-354, XP011099061 ISSN: 1063-6676 *Sections II, III.C, V* figure 1	1-10
X	US 4 644 108 A (CROUSE WILLIAM G [US] ET AL) 17 February 1987 (1987-02-17) abstract figure 7 column 6, line 48 - column 7, line 24	1-10

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International application No  
PCT/US2007/021814

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 4644108	A	17-02-1987	CA 1201541 A1 04-03-1986
			DE 3366616 D1 06-11-1986
			EP 0107122 A1 02-05-1984
			JR 59080029 A 09-05-1984

## フロントページの続き

(51)Int.Cl. F I テーマコード(参考)  
 G 1 0 L 21/02 1 0 2 B  
 H 0 3 H 21/00  
 H 0 4 B 3/23

(81)指定国 AP(BW,GH,GM,KE,LS,MW,MZ,NA,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,HU,IE,IS,IT,LT,LU,LV,MC,MT,NL,PL,PT,RO,SE,SI,SK,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,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,SG,SK,SL,SM,SV,SY,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,ZA,ZM,ZW

(72)発明者 シェーファー, ロナルド  
 アメリカ合衆国 カリフォルニア 9 4 3 0 4, パロ アルト, ページ ミル ロード 1 5  
 0 1

Fターム(参考) 5J023 DB01 DB05 DC07  
 5K046 CC01 HH11 HH79  
 5K127 AA03 GB47 MA02 MA03 MA32 MA34  
 5K201 AA01 BB09 CA01 EE20