

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第6995254号
(P6995254)

(45)発行日 令和4年1月14日(2022.1.14)

(24)登録日 令和3年12月16日(2021.12.16)

(51)国際特許分類		F I			
H 0 4 R	3/00 (2006.01)	H 0 4 R	3/00	3 1 0	
B 6 0 R	11/02 (2006.01)	B 6 0 R	11/02		
G 1 0 K	11/178 (2006.01)	G 1 0 K	11/178		

請求項の数 9 (全17頁)

(21)出願番号	特願2021-541855(P2021-541855)	(73)特許権者	000006013 三菱電機株式会社 東京都千代田区丸の内二丁目7番3号
(86)(22)出願日	令和1年8月28日(2019.8.28)	(74)代理人	100088672 弁理士 吉竹 英俊
(86)国際出願番号	PCT/JP2019/033619	(74)代理人	100088845 弁理士 有田 貴弘
(87)国際公開番号	WO2021/038736	(72)発明者	高橋 立明 東京都千代田区丸の内二丁目7番3号 三菱電機株式会社内
(87)国際公開日	令和3年3月4日(2021.3.4)	(72)発明者	下谷 光生 東京都千代田区丸の内二丁目7番3号 三菱電機株式会社内
審査請求日	令和3年10月13日(2021.10.13)	審査官	堀 洋介
早期審査対象出願			

最終頁に続く

(54)【発明の名称】 音場制御装置及び音場制御方法

(57)【特許請求の範囲】

【請求項1】

車両内のスピーカによって生成される、前記車両内の特定の位置での音場を制御する音場制御装置であって、

前記車両内の複数の乗員の撮影画像から、前記複数の乗員の顔情報を取得する取得部と、前記取得部で取得された前記顔情報と、前記車両内のマイクで受け付けた音声とに基づいて、前記複数の乗員の中から話者と聴者とを判定する判定制御部と、前記判定制御部による判定の結果に基づいて、前記複数の乗員の位置における音場を制御する音場制御部と

を備え、

前記音場制御部は、

前記判定制御部による判定の結果に基づいて、前記話者からの音声を打ち消す音声を、前記スピーカから前記聴者に出力させる駆動を行う消音音声駆動部と、

前記判定制御部による判定の結果に基づいて、前記話者からの音声を、前記スピーカから前記聴者に出力させる駆動を行う音声駆動部と、

前記消音音声駆動部の駆動及び前記音声駆動部の駆動を制御する駆動制御を行う駆動制御部と

を含む、音場制御装置。

【請求項2】

請求項1に記載の音場制御装置であって、

前記消音音声駆動部は、
前記車両の外部からの音声を打ち消さずに前記話者からの音声を打ち消す音声を、前記スピーカから前記聴者に出力させる駆動を行う、音場制御装置。

【請求項 3】

請求項 1 に記載の音場制御装置であって、
前記音声駆動部は、
前記聴者の特性に基づいて、前記スピーカから当該聴者に出力される音声の周波数特性を変更する、音場制御装置。

【請求項 4】

請求項 1 に記載の音場制御装置であって、
前記駆動制御部は、
前記取得部で取得された前記顔情報に基づいて、前記話者の顔の表情及び前記聴者の顔の表情、前記話者の顔の向き及び前記聴者の顔の向き、または、前記話者の顔の動き及び前記聴者の顔の動きを判定し、当該判定の結果に基づいて前記駆動制御を行う、音場制御装置。

10

【請求項 5】

請求項 1 に記載の音場制御装置であって、
前記駆動制御部は、
前記話者がハンドフリー通話をしているか否かに基づいて前記駆動制御を行う、音場制御装置。

20

【請求項 6】

請求項 1 に記載の音場制御装置であって、
前記駆動制御部は、
前記話者及び前記聴者からの操作に基づいて前記駆動制御を行う、音場制御装置。

【請求項 7】

請求項 1 に記載の音場制御装置であって、
前記駆動制御部は、
前記話者及び前記聴者からの音声の内容に基づいて前記駆動制御を行う、音場制御装置。

【請求項 8】

請求項 1 に記載の音場制御装置であって、
前記駆動制御部は、前記駆動制御を前記聴者ごとに行う、音場制御装置。

30

【請求項 9】

車両内のスピーカによって生成される、前記車両内の特定の位置での音場を制御する音場制御方法であって、
前記車両内の複数の乗員の撮影画像から、前記複数の乗員の顔情報を取得し、
取得された前記顔情報と、前記車両内のマイクで受け付けた音声とに基づいて、前記複数の乗員の中から話者と聴者とを判定し、
前記話者及び前記聴者の判定の結果に基づいて、前記複数の乗員の位置における音場を制御し、

前記複数の乗員の位置における前記音場の前記制御は、
前記判定の結果に基づいて、前記話者からの音声を打ち消す音声を前記スピーカから前記聴者に出力させる駆動と、前記判定の結果に基づいて、前記話者からの音声を前記スピーカから前記聴者に出力させる駆動と、を制御する駆動制御を含む、音場制御方法。

40

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、車両内のスピーカによって生成される、車両内の特定の位置での音場を制御する音場制御装置及び音場制御方法に関する。

【背景技術】

【0002】

50

従来、車両内において快適な会話環境を提供するシステムが提案されている。例えば、マイク（マイクロホン）とスピーカとを備えたハンドフリー電話システムでは、車両内のノイズを除去するノイズキャンセリング技術や、車両内のエコーを除去するエコーキャンセル技術によって快適な通話環境を実現している。また、車両内での乗員の会話をスムーズに行うために、それぞれの乗員近傍にマイク及びスピーカを配置し、一方の乗員の音声をマイクロホンから取り込み、他方の乗員近傍のスピーカから当該音声を出力する技術が提案されている。

【 0 0 0 3 】

なお、マイク及びスピーカを用いて車両内の音場を制御する技術として、例えば特許文献 1 及び特許文献 2 の技術が提案されている。特許文献 1 には、頭の回転に対応して音場を制御する技術が提案されている。特許文献 2 には、座席の位置に応じて音場を制御する技術が提案されている。

10

【先行技術文献】

【特許文献】

【 0 0 0 4 】

【文献】特開 2 0 0 9 - 2 5 3 5 2 6 号公報

特開平 0 7 - 1 8 4 2 9 8 号公報

【発明の概要】

【発明が解決しようとする課題】

【 0 0 0 5 】

以上のような従来技術では、運転者という特定の乗員からの音声のみに関して音場を適切に制御することを想定している。このため、車両内の任意の乗員からの音声に関して車両内の音場を適切に制御することができないという問題があった。

20

【 0 0 0 6 】

そこで、本発明は、上記のような問題点を鑑みてなされたものであり、車両内の任意の乗員からの音声に関して車両内の音場を適切に制御することが可能な技術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】

【 0 0 0 7 】

本発明に係る音場制御装置は、車両内の複数の乗員の撮影画像から、複数の乗員の顔情報を取得する取得部と、取得部で取得された顔情報と、車両内のマイクで受け付けた音声とに基づいて、複数の乗員の中から話者と聴者とを判定する判定制御部と、判定制御部による判定の結果に基づいて、複数の乗員の位置における音場を制御する音場制御部とを備え、音場制御部は、判定制御部による判定の結果に基づいて、話者からの音声を打ち消す音声を、スピーカから聴者に出力させる駆動を行う消音音声駆動部と、判定制御部による判定の結果に基づいて、話者からの音声を、スピーカから聴者に出力させる駆動を行う音声駆動部と、消音音声駆動部の駆動及び音声駆動部の駆動を制御する駆動制御を行う駆動制御部とを含む。

30

【発明の効果】

【 0 0 0 8 】

本発明によれば、取得された顔情報と、マイクで受け付けた音声とに基づいて、複数の乗員の中から話者と聴者とを判定し、話者及び聴者の判定の結果に基づいて、複数の乗員の位置における音場を制御する。このような構成によれば、車両内の任意の乗員からの音声に関して車両内の音場を適切に制御することができる。

40

【 0 0 0 9 】

本発明の目的、特徴、態様及び利点は、以下の詳細な説明と添付図面とによって、より明白となる。

【図面の簡単な説明】

【 0 0 1 0 】

【図 1】実施の形態 1 に係る音場制御装置の構成を示すブロック図である。

50

- 【図 2】実施の形態 2 に係る音場制御装置の構成を示すブロック図である。
 【図 3】実施の形態 2 に係る音場制御装置の動作を示すフローチャートである。
 【図 4】実施の形態 2 に係る音場制御装置の動作を説明するための図である。
 【図 5】実施の形態 3 に係る音場制御装置の動作を説明するための図である。
 【図 6】その他の変形例に係る音場制御装置のハードウェア構成を示すブロック図である。
 【図 7】その他の変形例に係る音場制御装置のハードウェア構成を示すブロック図である。
 【図 8】その他の変形例に係るサーバの構成を示すブロック図である。
 【図 9】その他の変形例に係る通信端末の構成を示すブロック図である。
 【発明を実施するための形態】

【 0 0 1 1 】

<実施の形態 1 >

図 1 は、本発明の実施の形態 1 に係る音場制御装置 1 の構成を示すブロック図である。音場制御装置 1 は、例えば D M S (Driver Monitoring System) などに適用することができる。図 1 の音場制御装置 1 は、車両内のマイク (マイクホン) 5 1 及びスピーカ 5 2 と無線または有線を介して接続されている。

【 0 0 1 2 】

マイク 5 1 は、車両内の複数の乗員の口から発せられた音声を受け付ける。マイク 5 1 には、例えば、各乗員からの音声のみを受け付け可能な超指向性マイクなどが適用される。複数の乗員には、例えば、運転者、助手席の同乗者、及び、後部席の同乗者などが含まれる。

【 0 0 1 3 】

スピーカ 5 2 は、音場制御装置 1 の制御によって、車両内に音場を生成する。このスピーカ 5 2 には、例えば、各乗員の位置において異なる音場を生成することが可能なパラメトリックアレイスピーカなどの超指向性スピーカなどが適用される。

【 0 0 1 4 】

次に、音場制御装置 1 について説明する。以下で説明するように、音場制御装置 1 は、スピーカ 5 2 によって生成される、車両内の特定の位置での音場を制御することが可能となっている。

【 0 0 1 5 】

図 1 の音場制御装置 1 は、取得部 1 1 と、判定制御部 1 2 と、音場制御部 1 3 とを備える。

【 0 0 1 6 】

取得部 1 1 は、車両内の複数の乗員の撮影画像から複数の乗員の顔情報を取得する。顔情報は、乗員の顔の特徴点及び状態を含む情報であり、例えば、乗員の口の位置、乗員の耳 (例えば両耳) の位置、及び、口の動きに関する情報である。取得部 1 1 には、カメラなどの撮影装置で撮影された乗員の画像に認識処理が可能な装置、及び、そのインターフェースの少なくともいずれかが用いられる。

【 0 0 1 7 】

判定制御部 1 2 は、取得部 1 1 で取得された顔情報と、マイク 5 1 で受け付けた音声とに基づいて、複数の乗員の中から話者と聴者とを判定する。例えば、判定制御部 1 2 は、時系列の顔情報から乗員の口が開閉していることを認識し、かつ、マイク 5 1 で受け付けた音声の波形が母音波形または子音波形であることを認識した場合に、当該乗員を話者と判定し、それ以外の乗員を聴者と判定する。また例えば、判定制御部 1 2 は、時系列の顔情報にリップリーディング機能を用いて乗員が発話していると推定し、かつ、マイク 5 1 で一定以上の音圧を検出した場合に、当該乗員を話者と判定し、それ以外の乗員を聴者と判定する。

【 0 0 1 8 】

音場制御部 1 3 は、判定制御部 1 2 による判定の結果に基づいてスピーカ 5 2 を制御することにより、車両内の特定の位置での音場、ひいては複数の乗員の位置における音場を制御する。例えば、音場制御部 1 3 は、判定制御部 1 2 で判定された話者の口の位置と、判定制御部 1 2 で判定された聴者の耳の位置とに基づいて、スピーカ 5 2 から聴者の耳に対

10

20

30

40

50

して、話者の音声、及び、話者の音声を打ち消す音声を選択的に出力する制御を行う。

【0019】

<実施の形態1のまとめ>

以上のような本実施の形態1に係る音場制御装置1によれば、取得された顔情報と、マイク51で受け付けた音声とに基づいて、複数の乗員の中から話者と聴者とを判定し、話者及び聴者の判定の結果に基づいて、複数の乗員の位置における音場を制御する。このような構成によれば、例えば、車両内で音声を発した任意の乗員を話者として判定し、それ以外の乗員を聴者として判定することができるので、任意の乗員からの音声に関して、複数の乗員の位置における音場を適切に制御することができる。このことは、近年提案されているライドシェアにおいて特に有効である。

10

【0020】

<実施の形態2>

図2は、本発明の実施の形態2に係る音場制御装置1の構成を示すブロック図である。以下、本実施の形態2に係る構成要素のうち、上述の構成要素と同じまたは類似する構成要素については同じまたは類似する参照符号を付し、異なる構成要素について主に説明する。

【0021】

図2の音場制御装置1は、指向性マイク51a、スピーカアレイ52a、及び、画像認識装置53と通信可能に接続されている。指向性マイク51a及びスピーカアレイ52aは、図1のマイク51及びスピーカ52の概念に含まれる。具体的には、指向性マイク51aは、実施の形態1で説明したマイク51と同様に、車両内の複数の乗員の口から発せられた音声を受け付ける。スピーカアレイ52aは、実施の形態1で説明したスピーカ52と同様に、音場制御装置1の制御によって、車両内に音場を生成する。画像認識装置53は、車両内の複数の乗員の撮影画像に認識処理を行うことにより、複数の乗員の顔情報を生成する。

20

【0022】

次に、音場制御装置1について説明する。図2の音場制御装置1は、顔情報取得部11aと、発話音声判定部12aと、話者聴者認識部12bと、消音音声駆動部13aと、音声駆動部13bと、駆動制御部13cと、マイク制御部14とを備える。なお、顔情報取得部11aは、図1の取得部11の概念に含まれ、発話音声判定部12a及び話者聴者認識部12bは、図1の判定制御部12の概念に含まれる。消音音声駆動部13a、音声駆動部13b及び駆動制御部13cは、図1の音場制御部13の概念に含まれる。

30

【0023】

顔情報取得部11aは、画像認識装置53で生成された顔情報を取得する。顔情報は、実施の形態1と同様に乗員の顔の特徴点及び状態を含む情報である。

【0024】

マイク制御部14は、指向性マイク51aが乗員の口の位置の音声を精度よく受け付けることができるように、顔情報取得部11aで取得された顔情報に基づいて、指向性マイク51aの指向性を制御する。なお、指向性マイク51aの指向性の調整ロジックは限定されるものではなく、既存の任意の調整ロジックを用いることができる。

【0025】

発話音声判定部12aは、顔情報取得部11aで取得された顔情報と、指向性マイク51aで受け付けた音声とに基づいて、指向性マイク51aで受け付けた音声が、乗員が発した音声であるか否かを判定する。このような判定には、例えば、実施の形態1の判定制御部12で説明した判定が用いられる。

40

【0026】

話者聴者認識部12bは、指向性マイク51aで受け付けた音声が乗員からの音声であると発話音声判定部12aにて判定された場合に、顔情報取得部11aで取得された顔情報と、発話音声判定部12aの判定の結果とに基づいて話者と聴者とを特定する。例えば、話者聴者認識部12bは、顔情報取得部11aで顔情報が取得された複数の乗員のうち、発話音声判定部12aで音声を発した乗員と判定された話者と判定し、それ以外の乗員を

50

聴者と判定する。この際、話者聴者認識部 1 2 b は、話者の口の位置と聴者の耳の位置も判定する。話者聴者認識部 1 2 b による判定の結果は、駆動制御部 1 3 c を介して消音音声駆動部 1 3 a 及び音声駆動部 1 3 b に出力される。

【 0 0 2 7 】

消音音声駆動部 1 3 a は、話者聴者認識部 1 2 b による判定の結果に基づいて、話者からの音声を打ち消す音声を、スピーカアレイ 5 2 a から聴者に出力させる駆動を行う。具体的には、消音音声駆動部 1 3 a は、話者聴者認識部 1 2 b で判定された話者の口の位置及び聴者の耳の位置と、指向性マイク 5 1 a で受け付けた話者の音声とに基づいて、聴者の耳の位置に話者の音声を打ち消す音声が出力されるように、スピーカアレイ 5 2 a を制御する。話者の音声の打ち消しは、聴者が話者の音声を聞き取れない程度の打ち消しであればよく、完全な打ち消しでなくてもよい。

10

【 0 0 2 8 】

音声駆動部 1 3 b は、話者聴者認識部 1 2 b による判定の結果に基づいて、話者からの音声を、スピーカアレイ 5 2 a から聴者に出力させる駆動を行う。具体的には、音声駆動部 1 3 b は、話者聴者認識部 1 2 b で判定された話者の口の位置及び聴者の耳の位置と、指向性マイク 5 1 a で受け付けた話者の音声とに基づいて、聴者の耳の位置に話者の音声が出力されるように、スピーカアレイ 5 2 a を制御する。

【 0 0 2 9 】

駆動制御部 1 3 c は、消音音声駆動部 1 3 a の駆動及び音声駆動部 1 3 b の駆動を制御する駆動制御を行う。本実施の形態 2 では、駆動制御部 1 3 c は、話者がハンドフリー通話をしているか否かに基づいて駆動制御を行う。ここでいう話者は、運転者であってもよいし、運転者以外の乗員であってもよい。

20

【 0 0 3 0 】

例えば、駆動制御部 1 3 c は、話者がハンドフリー通話をしていると判定した場合に、全ての聴者に対して音声駆動部 1 3 b を駆動せずに消音音声駆動部 1 3 a を駆動する制御を行う。また例えば、駆動制御部 1 3 c は、話者がハンドフリー通話をしていないと判定した場合に、全ての聴者に対して消音音声駆動部 1 3 a を駆動せずに音声駆動部 1 3 b を駆動する制御を行う。

【 0 0 3 1 】

なお、話者がハンドフリー通話をしているか否かは、ハンドフリー通話の制御装置（図示せず）から入力されたハンドフリー通話のステータス情報に基づいて判定されてもよいし、顔情報取得部 1 1 a で取得された顔情報に基づいて判定されてもよい。

30

【 0 0 3 2 】

< 動作 >

図 3 は、本実施の形態 2 に係る音場制御装置 1 の動作を示すフローチャートである。

【 0 0 3 3 】

まずステップ S T 1 にて、顔情報取得部 1 1 a は、画像認識装置 5 3 から顔情報を取得する。

【 0 0 3 4 】

ステップ S T 2 にて、マイク制御部 1 4 は、顔情報取得部 1 1 a で取得された顔情報に基づいて指向性マイク 5 1 a の指向性を制御する。

40

【 0 0 3 5 】

ステップ S T 3 にて、発話音声判定部 1 2 a は、顔情報取得部 1 1 a で取得された顔情報と、指向性マイク 5 1 a で受け付けた音声とに基づいて、指向性マイク 5 1 a で受け付けた音声が、乗員が発した音声であるか否かを判定する。つまり、発話音声判定部 1 2 a は、乗員が発話しているか否かを判定する。乗員が発話していると判定された場合には処理がステップ S T 4 に進み、乗員が発話していないと判定された場合には処理がステップ S T 1 に戻る。

【 0 0 3 6 】

ステップ S T 4 にて、話者聴者認識部 1 2 b は、顔情報取得部 1 1 a で取得された顔情報

50

と、発話音声判定部 1 2 a の判定の結果とに基づいて話者と聴者とを判定する。

【 0 0 3 7 】

ステップ S T 5 にて、駆動制御部 1 3 c は、話者がハンドフリー通話をしているか否かを判定する。話者がハンドフリー通話をしていると判定された場合には処理がステップ S T 6 に進み、話者がハンドフリー通話をしていないと判定された場合には処理がステップ S T 7 に進む。

【 0 0 3 8 】

ステップ S T 6 にて、駆動制御部 1 3 c は消音音声駆動部 1 3 a を駆動する制御を行う。これにより、消音音声駆動部 1 3 a は、話者聴者認識部 1 2 b による判定の結果に基づいて、話者からの音声を打ち消す音声を、スピーカアレイ 5 2 a から全ての聴者に出力させる駆動を行う。その後、処理がステップ S T 1 に戻る。

10

【 0 0 3 9 】

ステップ S T 7 にて、駆動制御部 1 3 c は音声駆動部 1 3 b を駆動する制御を行う。これにより、音声駆動部 1 3 b は、話者聴者認識部 1 2 b による判定の結果に基づいて、話者からの音声を、スピーカアレイ 5 2 a から全ての聴者に出力させる駆動を行う。その後、処理がステップ S T 1 に戻る。

【 0 0 4 0 】

< 実施の形態 2 のまとめ >

以上のような本実施の形態 2 に係る音場制御装置 1 によれば、話者が運転者であるか運転者以外の乗員であるかに関わらず、話者がハンドフリー通話をしているか否かに基づいて、消音音声駆動部 1 3 a の駆動及び音声駆動部 1 3 b の駆動を制御する駆動制御を行う。このような構成によれば、任意の乗員がハンドフリー通話を行っている場合に、他の乗員に対してその会話の秘匿を行うことができる。

20

【 0 0 4 1 】

< 変形例 1 >

駆動制御部 1 3 c は、ハンドフリー通話以外の動作に基づいて、消音音声駆動部 1 3 a の駆動及び音声駆動部 1 3 b の駆動を制御する駆動制御を行ってもよい。以下、その例について説明する。

【 0 0 4 2 】

< 変形例 1 A >

駆動制御部 1 3 c は、話者及び聴者からの操作に基づいて駆動制御を行ってもよい。例えば、駆動制御部 1 3 c は、話者及び聴者の両者が図示しないアイコンやキーに対して会話許諾の意思を示す操作を行った場合には、消音音声駆動部 1 3 a を駆動せずに音声駆動部 1 3 b を駆動する制御を行ってもよい。また例えば、駆動制御部 1 3 c は、話者及び聴者の両者が図示しないアイコンやキーに対して会話拒否の意思を示す操作を行った場合には、音声駆動部 1 3 b を駆動せずに消音音声駆動部 1 3 a を駆動する制御を行ってもよい。

30

【 0 0 4 3 】

なお、話者の意思と聴者の意思とが異なる場合には、会話拒否優先規則、話者優先規則などの予め定められた規則にしたがって駆動制御を行ってもよい。ここで、会話拒否優先規則とは、話者及び聴者の少なくとも一方が会話拒否の意思を示す操作を行った場合に、音声駆動部 1 3 b を駆動せずに消音音声駆動部 1 3 a を優先的に駆動する規則である。話者優先規則とは、聴者の操作よりも話者の操作を優先して消音音声駆動部 1 3 a 及び音声駆動部 1 3 b の駆動制御を行う規則である。

40

【 0 0 4 4 】

なお、駆動制御部 1 3 c は、話者及び聴者からの操作による設定を記憶することにより、それ以降、当該設定に基づいて駆動制御を行ってもよい。また、話者及び聴者からの操作は、図示しないアイコンやキーに対する操作に限ったものではなく、例えば、カメラなどを用いて話者及び聴者から検出されたジェスチャー操作などであってもよい。ジェスチャー操作は、例えば、会話許諾を示すジェスチャー操作として、指で O K マークを作成するジェスチャー操作を含んでもよいし、会話拒否を示すジェスチャー操作として、手で払い

50

のけるようなジェスチャー操作を含んでもよい。なお、ジェスチャー操作は、音場制御装置 1 が位置する範囲（例えば地域、国）によって適宜変更されてもよい。

【 0 0 4 5 】

< 変形例 1 B >

駆動制御部 1 3 c は、顔情報取得部 1 1 a で取得された顔情報に基づいて、話者の顔の表情及び聴者の顔の表情を判定し、当該判定の結果に基づいて駆動制御を行ってもよい。例えば、駆動制御部 1 3 c は、話者及び聴者の一方の顔がウインクをしていると判定した場合には、会話許諾の意思が示されていると判定して、消音音声駆動部 1 3 a を駆動せずに音声駆動部 1 3 b を駆動する制御を行ってもよい。また例えば、駆動制御部 1 3 c は、話者及び聴者の一方の顔がゆがんだ表情をしていると判定した場合には、会話拒否の意思が示されていると判定して、音声駆動部 1 3 b を駆動せずに消音音声駆動部 1 3 a を駆動する制御を行ってもよい。

10

【 0 0 4 6 】

または、駆動制御部 1 3 c は、顔情報取得部 1 1 a で取得された顔情報に基づいて、話者の顔の向き及び聴者の顔の向きを判定し、当該判定の結果に基づいて駆動制御を行ってもよい。例えば、駆動制御部 1 3 c は、話者及び聴者の一方の顔が他方に向いていると判定した場合には、会話許諾の意思が示されていると判定して、消音音声駆動部 1 3 a を駆動せずに音声駆動部 1 3 b を駆動する制御を行ってもよい。また例えば、駆動制御部 1 3 c は、話者及び聴者の一方の顔が他方から背けていると判定した場合には、会話拒否の意思が示されていると判定して、音声駆動部 1 3 b を駆動せずに消音音声駆動部 1 3 a を駆動する制御を行ってもよい。

20

【 0 0 4 7 】

または、駆動制御部 1 3 c は、顔情報取得部 1 1 a で取得された顔情報に基づいて、話者の顔の動き及び聴者の顔の動きを判定し、当該判定の結果に基づいて駆動制御を行ってもよい。例えば、駆動制御部 1 3 c は、話者及び聴者の一方の顔が他方に近づいた、または、縦方向に揺れたと判定した場合には、会話許諾の意思が示されていると判定して、消音音声駆動部 1 3 a を駆動せずに音声駆動部 1 3 b を駆動する制御を行ってもよい。また例えば、駆動制御部 1 3 c は、話者及び聴者の一方の顔が他方から離れた、または、横方向に揺れたと判定した場合には、会話拒否の意思が示されていると判定して、音声駆動部 1 3 b を駆動せずに消音音声駆動部 1 3 a を駆動する制御を行ってもよい。

30

【 0 0 4 8 】

なお本変形例 1 B において、話者の意思と聴者の意思とが異なる場合には、変形例 1 A と同様に会話拒否優先規則、話者優先規則などの予め定められた規則にしたがって駆動制御を行ってもよい。

【 0 0 4 9 】

< 変形例 1 C >

駆動制御部 1 3 c は、話者及び聴者からの音声の内容に基づいて駆動制御を行ってもよい。例えば、駆動制御部 1 3 c は、話者が聴者の名前を呼んだ、または、聴者が名前の呼びかけに返事したと判定した場合には、会話許諾の意思が示されていると判定して、消音音声駆動部 1 3 a を駆動せずに音声駆動部 1 3 b を駆動する制御を行ってもよい。また例えば、駆動制御部 1 3 c は、話者が聴者の名前を呼ばなかった、または、聴者が名前の呼びかけに返事しなかったと判定した場合には、会話拒否の意思が示されていると判定して、音声駆動部 1 3 b を駆動せずに消音音声駆動部 1 3 a を駆動する制御を行ってもよい。

40

【 0 0 5 0 】

なお本変形例 1 C において、話者の意思と聴者の意思とが異なる場合には、変形例 1 A と同様に会話拒否優先規則、話者優先規則などの予め定められた規則にしたがって駆動制御を行ってもよい。

【 0 0 5 1 】

< 変形例 2 >

実施の形態 2 では、駆動制御部 1 3 c は、話者がハンドフリー通話をしていると判定した

50

場合に、全ての聴者に対して一律に駆動制御を行ったが、これに限ったものではない。例えば、駆動制御部 1 3 c は、駆動制御を聴者ごとに行ってもよい。つまり、駆動制御部 1 3 c は、ある聴者に対しては音声駆動部 1 3 b を駆動する制御を行い、それと同時に、別の聴者に対しては消音音声駆動部 1 3 a を駆動する制御を行ってもよい。このような構成によれば、消音音声駆動部 1 3 a の駆動及び音声駆動部 1 3 b の駆動という、相反する駆動を同時に行うことができるので、車両内の快適な通話空間を実現することができる。

【 0 0 5 2 】

< 変形例 3 >

例えば、車両内において音楽が出力されている間に、音声駆動部 1 3 b を駆動して話者からの音声をスピーカアレイ 5 2 a から聴者に出力させる場合、音場制御装置 1 は、音楽のボリュームを下げる制御を行ってもよい。また、駆動制御部 1 3 c は、音声駆動部 1 3 b を駆動して話者からの音声を出力する場合に、消音音声駆動部 1 3 a を駆動してノイズキャンセル（例えば車両のエンジン音や走行に伴う騒音のキャンセル）を行ってもよい。

10

【 0 0 5 3 】

< 変形例 4 >

音声駆動部 1 3 b は、聴者の聴力及び嗜好などの特性に基づいて、スピーカアレイ 5 2 a から当該聴者に出力される音声の周波数特性を変更してもよい。例えば、聴者が低周波数の音を聞き取りにくい人である場合には、音声駆動部 1 3 b は、スピーカアレイ 5 2 a から当該聴者に出力される音声の低周波数成分を大きくしてもよい。

【 0 0 5 4 】

なお、聴者の特性は、当該聴者のスマートフォン及び補聴器から取得してもよいし、ユーザからの操作によって音場制御装置 1 に予め記憶されていてもよい。

20

【 0 0 5 5 】

< 実施の形態 3 >

図 4 は、上述した実施の形態 2 に係る音場制御装置 1 の動作を説明するための図であり、図 5 は、本発明の実施の形態 3 に係る音場制御装置 1 の動作を説明するための図である。なお、本実施の形態 3 に係る音場制御装置 1 のブロック構成は、実施の形態 2 に係る音場制御装置 1 のブロック構成（図 2）と同じである。このため、以下、本実施の形態 3 に係る構成要素のうち、上述の構成要素と同じまたは類似する構成要素については同じまたは類似する参照符号を付し、異なる構成要素について主に説明する。

30

【 0 0 5 6 】

まず図 4 を用いて実施の形態 2 に係る音場制御装置 1 の動作について説明する。指向性マイク 5 1 a で受け付ける音声 V_s は、話者 2 1 から指向性マイク 5 1 a に達した音声 S_1 と、車両の外部から指向性マイク 5 1 a に達した車外音 N_1 とが重畳した波形を有する。

このため、音声 V_s は、

$$V_s = S_1 + N_1 \quad \dots (1)$$

と表すことができる。

【 0 0 5 7 】

消音音声駆動部 1 3 a が駆動しない場合に、聴者 2 2 が聞こえる音声 h_1 は、話者 2 1 から聴者 2 2 の耳の位置に達した音声 S_2 と、車両の外部から聴者 2 2 の耳の位置に達した車外音 N_2 とが重畳した波形を有する。このため、音声 h_1 は、

40

$$h_1 = S_2 + N_2 \quad \dots (2)$$

と表すことができる。

【 0 0 5 8 】

音場制御装置 1 は、上式 (1) の音声 V_s に基づいてスピーカアレイ 5 2 a から聴者 2 2 の耳の位置に出力される音声 V_a を生成する。この音声 V_a は、

$$V_a = f(V_s) = f(S_1 + N_1) = f_1(S_1) + f_2(N_1) \quad \dots (3)$$

と表すことができる。

【 0 0 5 9 】

この音声 V_a は、音声 S_2 及び車外音 N_1 を打ち消すので、

50

$$f_1(S_1) - S_2 \cdots (4)$$

$$f_2(N_1) - N_1 \cdots (5)$$

という関係が成り立つ。

【0060】

このため、消音音声駆動部 13a が駆動している場合に、聴者 22 が聞こえる音声 h_2 は、

$$h_2 = S_2 + N_2 + V_a$$

$$= S_2 + N_2 + f_1(S_1) + f_2(N_1)$$

$$= \{S_2 + f_1(S_1)\} + \{N_2 + f_2(N_1)\}$$

$$0 + (N_2 - N_1) \cdots (6)$$

となる。なお、式(6)の最後の行では、式(4)及び式(5)を適用した。

10

【0061】

車外音は乗員の位置によらずほぼ同一であると仮定すると、 $N_2 = N_1$ となることから、

式(6)の音声 h_2 は、

$$h_2 = 0 \cdots (7)$$

となる。

【0062】

このため、聴者 22 には、話者 21 からの音声 S_2 だけでなく車外音 N_2 も消去されることになる。しかしながら、例えば聴者 22 が運転者である場合には、他車両のクラクション、衝突音、急ブレーキ音や、緊急車両のサイレンなどの運転に係る車外音が、運転者(聴者 22)に聞こえることが望ましい。

20

【0063】

そこで、本実施の形態 3 では、消音音声駆動部 13a は、車両の外部からの音声を打ち消さずに話者からの音声を打ち消す音声を、スピーカアレイ 52a から聴者に出力させる駆動を行う。これにより、条件付きまたは条件なしで、聴者 22 は車外音を聞くことが可能となっている。

【0064】

以下、図 5 を用いて本実施の形態 3 に係る音場制御装置 1 の動作について説明する。本実施の形態 3 では、車外音を積極的に受け付けるためのマイク 56 が車両に設けられている。マイク 56 は、例えば、無指向性のマイクであり、指向性マイク 51a よりも車外音を受け付けやすい位置に設けられる。

30

【0065】

このマイク 56 で受け付ける音声 V_r は、話者 21 からマイク 56 に達した音声 S_3 と、車両の外部からマイク 56 に達した車外音 N_3 とが重畳した波形を有する。このため、音声 V_r は、

$$V_r = S_3 + N_3 \cdots (8)$$

と表すことができる。

【0066】

本実施の形態 3 に係る音場制御装置 1 は、音声 V_r と音声 V_s とを入力音声として扱い、その差分の音声 V_{mix} を示す信号を生成する。この音声 V_{mix} は、

$$V_{mix} = V_s - V_r$$

$$= S_1 + N_1 - (S_3 + N_3) \cdots (9)$$

と表すことができる。

40

【0067】

通常であれば、音場制御装置 1 は、音声 S_3 を含めた音響フィルタを考慮して駆動制御ひいては音場制御を行うが、説明を簡単にするため $S_3 = 0$ とする。また、車外音は乗員の位置によらずほぼ同一であると仮定すると、 $N_1 = N_3$ となる。これらから式(9)の音声 V_{mix} は、

$$V_{mix} = S_1 \cdots (10)$$

と表すことができる。

【0068】

50

音場制御装置 1 が、この音声 V_{mix} の信号を用いて消音音声駆動部 13a を駆動すると、スピーカアレイ 52a から聴者 22 の耳の位置に出力される音声 V_a は、

$$V_a = f(V_{mix}) \quad f_1(S_1) \quad - S_2 \quad \dots (11)$$
と表すことができる。

【0069】

したがって、消音音声駆動部 13a が駆動している場合に、聴者 22 が聞こえる音声 h_2 は、

$$h_2 = S_2 + N_2 + V_a \quad N_2 \quad \dots (12)$$

となる。

【0070】

このため、聴者 22 には、話者 21 からの音声 S_2 は聞こえないが、車外音 N_2 を聞くことができる。

【0071】

<実施の形態 3 のまとめ>

以上のような本実施の形態 3 に係る音場制御装置 1 によれば、車外音を打ち消さずに話者 21 からの音声を打ち消す音声を、スピーカアレイ 52a から聴者 22 に出力させる駆動を行う。このような構成によれば、聴者 22 は、他車両のクラクション、衝突音、急ブレーキ音や、緊急車両のサイレンなど、運転者の運転に関係する車外音を聞くことができる。

【0072】

<その他の変形例>

上述した図 1 の取得部 11、判定制御部 12、及び、音場制御部 13 を、以下「取得部 11 等」と記す。取得部 11 等は、図 6 に示す処理回路 81 により実現される。すなわち、処理回路 81 は、車両内の複数の乗員の撮影画像から、複数の乗員の顔情報を取得する取得部 11 と、取得部 11 で取得された顔情報と、車両内のマイクで受け付けた音声とに基づいて、複数の乗員の中から話者と聴者とを判定する判定制御部 12 と、判定制御部 12 による判定の結果に基づいて、複数の乗員の位置における音場を制御する音場制御部 13 と、を備える。処理回路 81 には、専用のハードウェアが適用されてもよいし、メモリに格納されるプログラムを実行するプロセッサが適用されてもよい。プロセッサには、例えば、中央処理装置、処理装置、演算装置、マイクロプロセッサ、マイクロコンピュータ、DSP (Digital Signal Processor) などが該当する。

【0073】

処理回路 81 が専用のハードウェアである場合、処理回路 81 は、例えば、単回路、複合回路、プログラム化したプロセッサ、並列プログラム化したプロセッサ、ASIC (Application Specific Integrated Circuit)、FPGA (Field Programmable Gate Array)、またはこれらを組み合わせたものが該当する。取得部 11 等の各部の機能それぞれは、処理回路を分散させた回路で実現されてもよいし、各部の機能をまとめて一つの処理回路で実現されてもよい。

【0074】

処理回路 81 がプロセッサである場合、取得部 11 等の機能は、ソフトウェア等との組み合わせにより実現される。なお、ソフトウェア等には、例えば、ソフトウェア、ファームウェア、または、ソフトウェア及びファームウェアが該当する。ソフトウェア等はプログラムとして記述され、メモリに格納される。図 7 に示すように、処理回路 81 に適用されるプロセッサ 82 は、メモリ 83 に記憶されたプログラムを読み出して実行することにより、各部の機能を実現する。すなわち、音場制御装置 1 は、処理回路 81 により実行されるときに、車両内の複数の乗員の撮影画像から、複数の乗員の顔情報を取得するステップと、取得された顔情報と、車両内のマイクで受け付けた音声とに基づいて、複数の乗員の中から話者と聴者とを判定するステップと、話者及び聴者の判定の結果に基づいて、複数の乗員の位置における音場を制御するステップと、が結果的に実行されることになるプログラムを格納するためのメモリ 83 を備える。換言すれば、このプログラムは、取得部 11 等の手順や方法をコンピュータに実行させるものであるともいえる。ここで、メモリ 8

10

20

30

40

50

3は、例えば、RAM (Random Access Memory)、ROM (Read Only Memory)、フラッシュメモリ、EPROM (Erasable Programmable Read Only Memory)、EEPROM (Electrically Erasable Programmable Read Only Memory)などの、不揮発性または揮発性の半導体メモリ、HDD (Hard Disk Drive)、磁気ディスク、フレキシブルディスク、光ディスク、コンパクトディスク、ミニディスク、DVD (Digital Versatile Disc)、そのドライブ装置等、または、今後使用されるあらゆる記憶媒体であってもよい。

【0075】

以上、取得部11等の各機能が、ハードウェア及びソフトウェア等のいずれか一方で実現される構成について説明した。しかしこれに限ったものではなく、取得部11等の一部を専用のハードウェアで実現し、別の一部をソフトウェア等で実現する構成であってもよい。例えば、取得部11については専用のハードウェアとしての処理回路81、インターフェース及びレシーバなどでその機能を実現し、それ以外についてはプロセッサ82としての処理回路81がメモリ83に格納されたプログラムを読み出して実行することによってその機能を実現することが可能である。

10

【0076】

以上のように、処理回路81は、ハードウェア、ソフトウェア等、またはこれらの組み合わせによって、上述の各機能を実現することができる。

【0077】

また、以上で説明した音場制御装置1は、DMS (Driver Monitoring System) 装置などの車両装置と、携帯電話、スマートフォン及びタブレットなどの携帯端末を含む通信端末と、車両装置及び通信端末の少なくとも1つにインストールされるアプリケーションの機能と、サーバとを適宜に組み合わせてシステムとして構築される音場制御システムにも適用することができる。この場合、以上で説明した音場制御装置1の各機能あるいは各構成要素は、前記システムを構築する各機器に分散して配置されてもよいし、いずれかの機器に集中して配置されてもよい。

20

【0078】

図8は、本変形例に係るサーバ91の構成を示すブロック図である。図8のサーバ91は、通信部91aと制御部91bとを備えており、車両92のDMS装置などの車両装置93と無線通信を行うことが可能となっている。

30

【0079】

取得部である通信部91aは、車両装置93と無線通信を行うことにより、車両装置93で取得された車両内の複数の乗員の顔情報と車両内の音声とを受信する。

【0080】

制御部91bは、サーバ91の図示しないプロセッサなどが、サーバ91の図示しないメモリに記憶されたプログラムを実行することにより、図1の判定制御部12及び音場制御部13と同様の機能を有している。つまり、制御部91bは、通信部91aで受信された顔情報及び音声に基づいて車両内の複数の乗員の中から話者と聴者とを判定し、その判定の結果に基づいて複数の乗員の位置における音場を制御するための制御信号を生成する。そして、通信部91aは、制御部91bで生成された制御信号を車両装置93に送信する。このように構成されたサーバ91によれば、実施の形態1で説明した音場制御装置1と同様の効果を得ることができる。

40

【0081】

図9は、本変形例に係る通信端末96の構成を示すブロック図である。図9の通信端末96は、通信部91aと同様の通信部96aと、制御部91bと同様の制御部96bとを備えており、車両97の車両装置98と無線通信を行うことが可能となっている。なお、通信端末96には、例えば車両97の運転者が携帯する携帯電話、スマートフォン、及びタブレットなどの携帯端末が適用される。このように構成された通信端末96によれば、実施の形態1で説明した音場制御装置1と同様の効果を得ることができる。

【0082】

50

なお、本発明は、その発明の範囲内において、各実施の形態及び各変形例を自由に組み合わせたり、各実施の形態及び各変形例を適宜、変形、省略したりすることが可能である。

【0083】

本発明は詳細に説明されたが、上記した説明は、すべての態様において、例示であって、本発明がそれに限定されるものではない。例示されていない無数の変形例が、本発明の範囲から外れることなく想定され得るものと解される。

【符号の説明】

【0084】

1 音場制御装置、11 取得部、12 判定制御部、13 音場制御部、13a 消音音声駆動部、13b 音声駆動部、13c 駆動制御部、51 マイク、52 スピーカ。

10

20

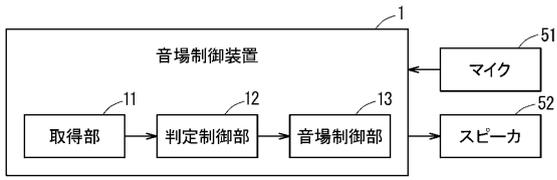
30

40

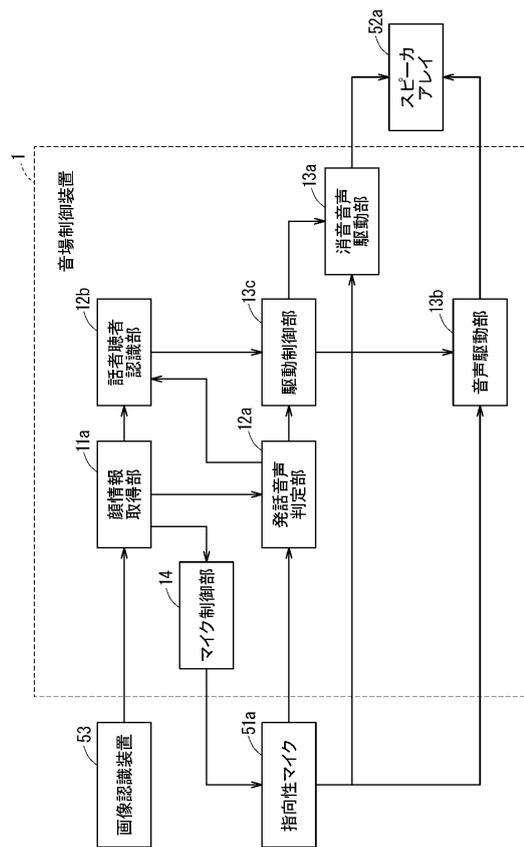
50

【図面】

【図1】



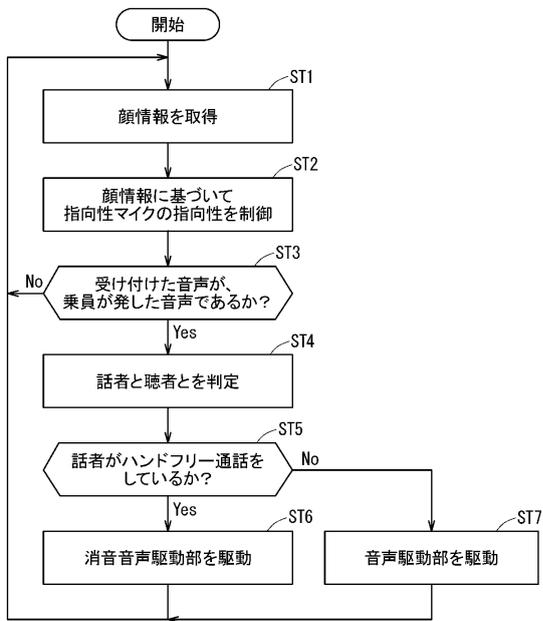
【図2】



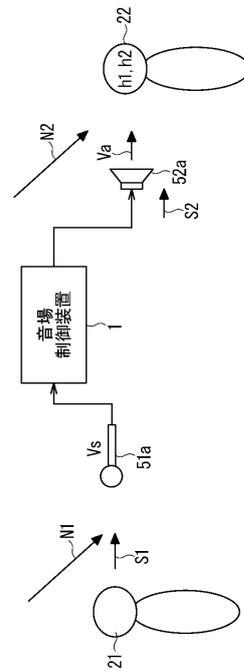
10

20

【図3】



【図4】

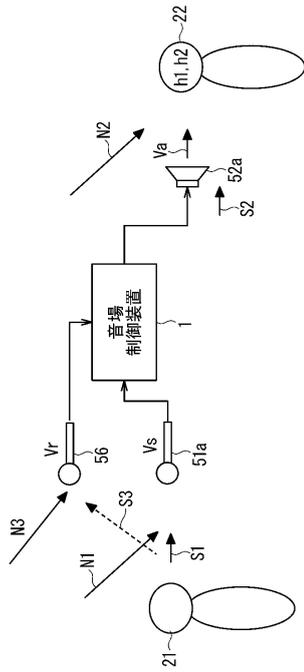


30

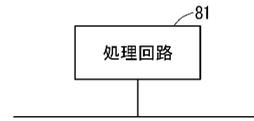
40

50

【図5】



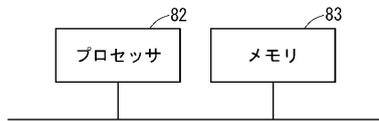
【図6】



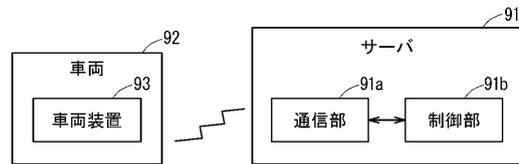
10

20

【図7】



【図8】

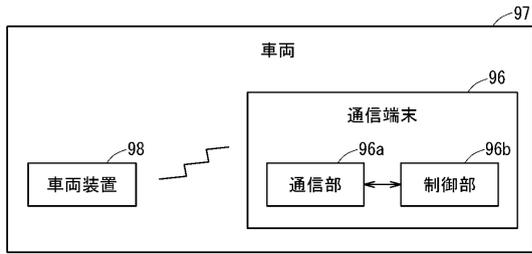


30

40

50

【図 9】



10

20

30

40

50

フロントページの続き

- (56)参考文献 特開2017-193207(JP,A)
特開2010-023639(JP,A)
特開平11-165594(JP,A)
特開2016-180971(JP,A)
特開2015-071320(JP,A)

- (58)調査した分野 (Int.Cl., DB名)
H04R 3/00
B60R 11/02
G10K 11/178