



(12) **Veröffentlichung**

der internationalen Anmeldung mit der
(87) Veröffentlichungs-Nr.: **WO 2022/059245**
in der deutschen Übersetzung (Art. III § 8 Abs. 2
IntPatÜbkG)
(21) Deutsches Aktenzeichen: **11 2021 004 878.3**
(86) PCT-Aktenzeichen: **PCT/JP2021/016088**
(86) PCT-Anmeldetag: **20.04.2021**
(87) PCT-Veröffentlichungstag: **24.03.2022**
(43) Veröffentlichungstag der PCT Anmeldung
in deutscher Übersetzung: **10.08.2023**

(51) Int Cl.: **G10L 15/28 (2013.01)**
B60R 16/02 (2006.01)

(30) Unionspriorität:
2020-157743 18.09.2020 JP

(71) Anmelder:
**Panasonic Intellectual Property Management Co.,
Ltd., Osaka, JP**

(74) Vertreter:
**Eisenführ Speiser Patentanwälte Rechtsanwälte
PartGmbH, 28217 Bremen, DE**

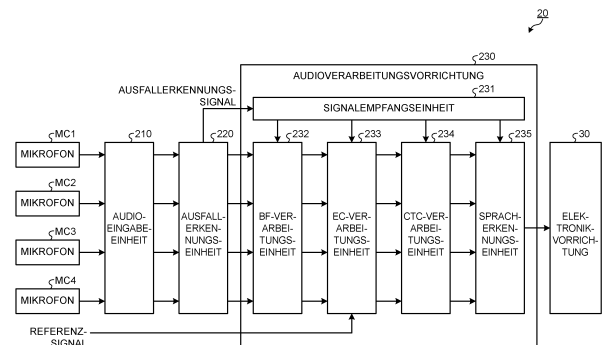
(72) Erfinder:
**Yamanashi, Tomofumi, Kadoma-shi, Osaka, JP;
Mochiki, Naoya, Kadoma-shi, Osaka, JP**

Prüfungsantrag gemäß § 44 PatG ist gestellt.

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen.

(54) Bezeichnung: **Audioverarbeitungssystem, Audioverarbeitungsvorrichtung und Audioverarbeitungsverfahren**

(57) Zusammenfassung: Ein Audioverarbeitungssystem gemäß der vorliegenden Offenbarung enthält eine Eingabeeinheit, eine Bestimmungseinheit und eine Spracherkennungseinheit. Die Eingabeeinheit empfängt eine erste Sprachäußerung, die eine durch einen ersten Sprecher abgegebene Sprachäußerung ist. Die Bestimmungseinheit bestimmt, ob eine Position des ersten Sprechers ermittelt werden kann oder nicht. Die Spracherkennungseinheit gibt an eine Objektivorrichtung einen Sprachbefehl aus, der durch eine Sprachäußerung angegeben ist und ein Signal zum Steuern der Objektivorrichtung ist, und schränkt die Ausgabe eines auf eine Position eines Sprechers bezogenen Sprecherpositionsbefehls unter den Sprachbefehlen in einem Fall ein, wo die Bestimmungseinheit bestimmt, dass die Position des ersten Sprechers nicht ermittelt werden kann.



Beschreibung

Technisches Gebiet

[0001] Die vorliegende Offenbarung betrifft ein Audioverarbeitungssystem, eine Audioverarbeitungsvorrichtung und ein Audioverarbeitungsverfahren.

Technischer Hintergrund

[0002] Ein Audioverarbeitungssystem ist bekannt, das einen Spracherkennungsbefehl auf Grundlage einer durch einen Sprecher abgegebenen Sprachäußerung verarbeitet.

[0003] Die Patentschrift 1 offenbart ein Audioverarbeitungssystem, das einen Spracherkennungsbefehl verarbeitet, der auf einer Position basiert, wo sich ein Sprecher geäußert hat.

Auflistung von Entgegenhaltungen

Patentliteratur

[0004] Patentschrift 1: JP 2007-90611 A

Kurzdarstellung der Erfindung

[0005] Jedoch offenbart die Patentschrift 1 nicht die Steuerung in einem Fall, wo die Position des Sprechers nicht ermittelt werden kann. Wenn nicht angenommen wird, dass die Position des Sprechers nicht ermittelt werden kann, könnte das Audioverarbeitungssystem eine unbeabsichtigte Verarbeitung ausführen.

[0006] Eine Aufgabe der vorliegenden Offenbarung ist es, eine angemessene Verarbeitung sogar in einem Fall auszuführen, wo in einem Audioverarbeitungssystem eine Position eines Sprechers nicht ermittelt werden kann.

[0007] Ein Audioverarbeitungssystem gemäß der vorliegenden Offenbarung enthält eine Eingabeeinheit, eine Bestimmungseinheit und eine Spracherkennungseinheit. Die Eingabeeinheit empfängt eine erste Sprachäußerung, die eine durch einen ersten Sprecher abgegebene Sprachäußerung ist. Die Bestimmungseinheit bestimmt, ob eine Position des ersten Sprechers ermittelt werden kann oder nicht. Die Spracherkennungseinheit gibt an eine Objektivvorrichtung einen Sprachbefehl aus, der durch eine Sprachäußerung angegeben ist und ein Signal zum Steuern der Objektivvorrichtung ist, wobei die Spracherkennungseinheit die Ausgabe eines auf die Position eines Sprechers bezogenen Sprecherpositionsbefehls unter dem Sprachbefehl in einem Fall einschränkt, wo die Bestimmungseinheit bestimmt,

dass die Position des ersten Sprechers nicht ermittelt werden kann.

[0008] Gemäß der vorliegenden Offenbarung ist es möglich, eine angemessene Verarbeitung sogar in einem Fall auszuführen, wo in einem Audioverarbeitungssystem eine Position eines Sprechers nicht ermittelt werden kann.

Figurenliste

Fig. 1 ist ein Diagramm, das ein Beispiel eines schematischen Aufbaus eines Fahrzeug-Audioverarbeitungssystems gemäß einer ersten Ausführungsform darstellt;

Fig. 2 ist ein Diagramm, das ein Beispiel eines Hardwareaufbaus eines Audioverarbeitungssystems gemäß der ersten Ausführungsform darstellt;

Fig. 3 ist ein Blockschaltbild, das ein Beispiel eines Aufbaus des Audioverarbeitungssystems gemäß der ersten Ausführungsform darstellt;

Fig. 4 ist ein Flussdiagramm, das ein Beispiel eines Betriebs des Audioverarbeitungssystems gemäß der ersten Ausführungsform darstellt;

Fig. 5 ist ein Blockschaltbild, das ein Beispiel eines Aufbaus eines Audioverarbeitungssystems gemäß einer zweiten Ausführungsform darstellt; und

Fig. 6 ist ein Flussdiagramm, das ein Beispiel eines Betriebs des Audioverarbeitungssystems gemäß der zweiten Ausführungsform darstellt.

Beschreibung von Ausführungsformen

[0009] Nachstehend sind Ausführungsformen der vorliegenden Offenbarung unter Bezugnahme auf die Zeichnung, soweit jeweils anwendbar, genau beschrieben. Jedoch kann eine unnötig genaue Beschreibung weggelassen sein. Es ist anzumerken, dass die begleitende Zeichnung und die folgende Beschreibung für Fachleute auf dem Gebiet vorgesehen sind, damit sie die vorliegende Offenbarung gründlich verstehen, und den in den Ansprüchen beschriebenen Gegenstand nicht einschränken sollen.

(Erste Ausführungsform)

[0010] **Fig. 1** ist ein Diagramm, das ein Beispiel eines schematischen Aufbaus eines Audiosystems 5 gemäß einer ersten Ausführungsform darstellt. Das Audiosystem 5 ist beispielsweise in einem Fahrzeug 10 montiert. Nachfolgend ist ein Beispiel beschrieben, in dem das Audiosystem 5 in dem Fahrzeug 10 montiert ist.

[0011] Eine Vielzahl von Sitzen ist in einem Inneren des Fahrzeugs 10 vorgesehen. Die Vielzahl von Sitzen enthält beispielsweise vier Sitze, umfassend einen Fahrersitz, einen Beifahrersitz sowie einen linken und einen rechten Rücksitz. Anzumerken ist, dass die Anzahl von Sitzen nicht hierauf beschränkt ist. Nachstehend ist eine Person, die auf dem Fahrersitz sitzt, als ein Insasse hm1 bezeichnet, ist eine Person, die auf dem Beifahrersitz sitzt, als ein Insasse hm2 bezeichnet, ist eine Person, die auf der rechten Seite des Rücksitzes sitzt, als ein Insasse hm3 bezeichnet, und ist eine Person, die auf der linken Seite des Rücksitzes sitzt, als ein Insasse hm4 bezeichnet.

[0012] Das Audiosystem 5 enthält ein Mikrofon MC1, ein Mikrofon MC2, ein Mikrofon MC3, ein Mikrofon MC4, ein Audioverarbeitungssystem 20 und eine Elektronikvorrichtung 30. Das in **Fig. 1** dargestellte Audiosystem 5 enthält die gleiche Anzahl von Mikrofonen wie die Anzahl von Sitzen, das heißt, vier Mikrofone, aber die Anzahl von Mikrofonen braucht nicht gleich der Anzahl von Sitzen zu sein.

[0013] Das Mikrofon MC1, das Mikrofon MC2, das Mikrofon MC3 und das Mikrofon MC4 geben Audiosignale an das Audioverarbeitungssystem 20 aus. Dann gibt das Audioverarbeitungssystem 20 das Spracherkennungsergebnis an die Elektronikvorrichtung 30 aus. Die Elektronikvorrichtung 30 führt eine durch das Spracherkennungsergebnis angegebene Verarbeitung aus, die auf dem eingegebenen Spracherkennungsergebnis basiert.

[0014] Das Mikrofon MC1 ist ein Mikrofon, das eine durch den Insassen hm1 abgegebene Sprachäußerung aufnimmt. Mit anderen Worten, das Mikrofon MC1 erlangt ein Audiosignal, das eine durch den Insassen hm1 geäußerte Audiokomponente enthält. Das Mikrofon MC1 ist beispielsweise auf der rechten Seite der obenliegenden Konsole angeordnet. Das Mikrofon MC2 nimmt eine durch den Insassen hm2 abgegebene Sprachäußerung auf. Mit anderen Worten, das Mikrofon MC2 ist ein Mikrofon, das ein Audiosignal erlangt, das eine durch den Insassen hm2 geäußerte Audiokomponente enthält. Das Mikrofon MC2 ist beispielsweise auf der linken Seite der obenliegenden Konsole angeordnet. Das heißt, das Mikrofon MC1 und das Mikrofon MC2 sind an Positionen nah beieinander angeordnet.

[0015] Das Mikrofon MC3 ist ein Mikrofon, das eine durch den Insassen hm3 abgegebene Sprachäußerung aufnimmt. Mit anderen Worten, das Mikrofon MC3 erlangt ein Audiosignal, das eine durch den Insassen hm3 geäußerte Audiokomponente enthält. Das Mikrofon MC3 ist beispielsweise rechts in der Mitte nahe dem Rücksitz am Himmel angeordnet. Das Mikrofon MC4 ist ein Mikrofon, das eine durch den Insassen hm4 abgegebene Sprachäußerung

aufnimmt. Mit anderen Worten, das Mikrofon MC4 erlangt ein Audiosignal, das eine durch den Insassen hm4 geäußerte Audiokomponente enthält. Das Mikrofon MC4 ist beispielsweise links in der Mitte nahe dem Rücksitz am Himmel angeordnet. Das heißt, das Mikrofon MC3 und das Mikrofon MC4 sind an Positionen nah beieinander angeordnet.

[0016] Ferner sind die in **Fig. 1** dargestellten Anordnungspositionen des Mikrofons MC1, des Mikrofons MC2, des Mikrofons MC3 und des Mikrofons MC4 ein Beispiel, und sie können an anderen Positionen angeordnet sein.

[0017] Jedes Mikrofon kann ein Richtmikrofon oder ein omnidirektionales Mikrofon sein. Jedes Mikrofon kann ein MEMS-Mikrofon (Small Micro Electro Mechanical System, kleines mikroelektromechanisches System) oder ein Elektretkondensatormikrofon (Electret Condenser Microphone, ECM) sein. Jedes Mikrofon kann ein Richtmikrofon (Beamforming-Mikrofon) sein. Zum Beispiel kann jedes Mikrofon eine Mikrofonanordnung mit einer Richtwirkung in der Richtung eines entsprechenden Sitzes sein und in der Lage sein, die Sprachäußerung aus der Richtwirkungsrichtung aufzunehmen.

[0018] Das in **Fig. 1** dargestellte Audiosystem 5 enthält eine Vielzahl von Audioverarbeitungssystemen 20, die den jeweiligen Mikrofonen entsprechen. Genauer enthält das Audiosystem 5 ein Audioverarbeitungssystem 21, ein Audioverarbeitungssystem 22, ein Audioverarbeitungssystem 23 und ein Audioverarbeitungssystem 24. Das Audioverarbeitungssystem 21 entspricht dem Mikrofon MC1. Das Audioverarbeitungssystem 22 entspricht dem Mikrofon MC2. Das Audioverarbeitungssystem 23 entspricht dem Mikrofon MC3. Das Audioverarbeitungssystem 24 entspricht dem Mikrofon MC4. Nachstehend können das Audioverarbeitungssystem 21, das Audioverarbeitungssystem 22, das Audioverarbeitungssystem 23 und das Audioverarbeitungssystem 24 kollektiv als ein Audioverarbeitungssystem 20 bezeichnet sein.

[0019] Ein von dem Audioverarbeitungssystem 20 ausgegebenes Signal wird in die Elektronikvorrichtung 30 eingegeben. Die Elektronikvorrichtung 30 führt eine Verarbeitung entsprechend dem von dem Audioverarbeitungssystem 20 ausgegebenen Signal aus. Hier ist das von dem Audioverarbeitungssystem 20 ausgegebene Signal beispielsweise ein Sprachbefehl, der ein durch Sprachäußerung eingegebener Befehl ist. Der Sprachbefehl ist ein Signal, das durch Sprachäußerung angegeben ist und eine Objektivrichtung steuert. Mit anderen Worten, die Elektronikvorrichtung 30 führt eine Verarbeitung entsprechend dem von dem Audioverarbeitungssystem 20 ausgegebenen Sprachbefehl aus. Zum Beispiel führt die Elektronikvorrichtung 30 eine Verarbeitung zum Öff-

nen und Schließen eines Fensters, eine Verarbeitung bezogen auf das Fahren des Fahrzeugs 10, eine Verarbeitung zum Ändern der Temperatur der Klimaanlage und eine Verarbeitung zum Ändern der Lautstärke der Audio-Vorrichtung auf Grundlage des Sprachbefehls aus. Die Elektronikvorrichtung 30 ist ein Beispiel der Objektvorrichtung.

[0020] Obwohl **Fig. 1** einen Fall darstellt, wo vier Personen in das Fahrzeug 10 einsteigen, ist die Anzahl der Personen, die in das Fahrzeug einsteigen, nicht darauf beschränkt. Die Anzahl von Insassen kann gleich oder kleiner als eine maximale Insassenkapazität des Fahrzeugs 10 sein. Wenn beispielsweise die maximale Insassenkapazität des Fahrzeugs 10 sechs beträgt, kann die Anzahl von Insassen sechs oder fünf oder weniger betragen.

[0021] **Fig. 2** ist ein Diagramm, das ein Beispiel eines Hardwareaufbaus des Audioverarbeitungssystems 20 gemäß der ersten Ausführungsform darstellt. In dem in **Fig. 2** dargestellten Beispiel enthält das Audioverarbeitungssystem 20 einen digitalen Signalprozessor (DSP) 2001, einen Direktzugriffsspeicher (RAM) 2002, einen Nur-Lese-Speicher (ROM) 2003 und eine Ein-/Ausgabeschnittstelle (E/A-Schnittstelle) 2004.

[0022] Der DSP 2001 ist ein Prozessor, der imstande ist, ein Computerprogramm auszuführen. Anzumerken ist, dass die Art des in dem Audioverarbeitungssystem 20 enthaltenen Prozessors nicht auf den DSP 2001 beschränkt ist. Zum Beispiel kann das Audioverarbeitungssystem 20 eine Zentraleinheit (CPU) oder eine andere Hardware sein. Ferner kann das Audioverarbeitungssystem 20 eine Vielzahl von Prozessoren enthalten.

[0023] Das RAM 2002 ist ein flüchtiger Speicher, benutzt als Cache, als Puffer oder dergleichen. Anzumerken ist, dass die Art des in dem Audioverarbeitungssystem 20 enthaltenen flüchtigen Speichers nicht auf das RAM 2002 beschränkt ist. Das Audioverarbeitungssystem 20 kann ein Register anstelle des RAMs 2002 enthalten. Das Audioverarbeitungssystem 20 kann eine Vielzahl von flüchtigen Speichern enthalten.

[0024] Das ROM 2003 ist ein nichtflüchtiger Speicher, der verschiedene Arten von Informationen speichert, einschließlich eines Computerprogramms. Der DSP 2001 führt die Funktionen des Audioverarbeitungssystems 20 aus, indem er ein spezielles Computerprogramm aus dem ROM 2003 ausliest und das Programm ausführt. Die Funktion des Audioverarbeitungssystems 20 ist weiter unten beschrieben. Anzumerken ist, dass die Art des in dem Audioverarbeitungssystem 20 enthaltenen nichtflüchtigen Speichers nicht auf das ROM 2003 beschränkt ist. Zum Beispiel kann das Audioverarbeitungssystem

20 einen Flash-Speicher anstelle des ROMs 2003 enthalten. Das Audioverarbeitungssystem 20 kann eine Vielzahl von nichtflüchtigen Speichern enthalten.

[0025] Die E/A-Schnittstelle 2004 ist eine Schnittstellenvorrichtung, an die eine externe Vorrichtung angeschlossen wird. Hier ist die externe Vorrichtung beispielsweise eine Vorrichtung, wie etwa das Mikrofon MC1, das Mikrofon MC2, das Mikrofon MC3, das Mikrofon MC4 oder die Elektronikvorrichtung 30. Das Audioverarbeitungssystem 20 kann eine Vielzahl der E/A-Schnittstellen 2004 enthalten.

[0026] Wie oben beschrieben, enthält das Audioverarbeitungssystem 20 den Speicher, in dem das Computerprogramm gespeichert ist, und den Prozessor, der imstande ist, das Computerprogramm auszuführen. Das heißt, das Audioverarbeitungssystem 20 kann als ein Computer betrachtet werden. Anzumerken ist, dass die Anzahl von zum Verwirklichen der Funktion als das Audioverarbeitungssystem 20 erforderlichen Computern nicht auf einen einzigen beschränkt ist. Die Funktion als das Audioverarbeitungssystem 20 kann durch ein Zusammenwirken von zwei oder mehr Computern verwirklicht werden.

[0027] **Fig. 3** ist ein Blockschaltbild, das ein Beispiel eines Aufbaus des Audioverarbeitungssystems 20 gemäß der ersten Ausführungsform darstellt. Audiosignale werden in das Audioverarbeitungssystem 20 von dem Mikrofon MC1, dem Mikrofon MC2, dem Mikrofon MC3 und dem Mikrofon MC4 eingegeben. Dann gibt das Audioverarbeitungssystem 20 das Spracherkennungsergebnis an die Elektronikvorrichtung 30 aus. Das Audioverarbeitungssystem 20 enthält eine Audioeingabeeinheit 210, eine Ausfallerkennungseinheit 220 und eine Audioverarbeitungsvorrichtung 230.

[0028] Das Mikrofon MC1 erzeugt ein Audiosignal durch ein Umwandeln einer aufgenommenen Sprachäußerung in ein elektrisches Signal. Dann gibt das Mikrofon MC1 das Audiosignal an die Audioeingabeeinheit 210 aus. Das Audiosignal ist ein Signal, das eine Sprachäußerung des Insassen hm1, eine Sprachäußerung einer anderen Person als des Insassen hm1, von einer Audio-Vorrichtung ausgegebene Musik oder Geräusche, wie etwa ein Fahrgeräusch, enthält.

[0029] Das Mikrofon MC2 erzeugt ein Audiosignal durch ein Umwandeln einer aufgenommenen Sprachäußerung in ein elektrisches Signal. Dann gibt das Mikrofon MC2 das Audiosignal an die Audioeingabeeinheit 210 aus. Das Audiosignal ist ein Signal, das eine Sprachäußerung des Insassen hm2, eine Sprachäußerung einer anderen Person als des Insassen hm2, von einer Audio-Vorrichtung

ausgegebene Musik oder Geräusche, wie etwa ein Fahrgeräusch, enthält.

[0030] Das Mikrofon MC3 erzeugt ein Audiosignal durch ein Umwandeln aufgenommener Sprachäußerung in ein elektrisches Signal. Dann gibt das Mikrofon MC3 das Audiosignal an die Audioeingabeeinheit 210 aus. Das Audiosignal ist ein Signal, das eine Sprachäußerung des Insassen hm3, eine Sprachäußerung einer anderen Person als des Insassen hm3, von einer Audio-Vorrichtung ausgegebene Musik oder Geräusche, wie etwa ein Fahrgeräusch, enthält.

[0031] Das Mikrofon MC4 erzeugt ein Audiosignal durch ein Umwandeln aufgenommener Sprachäußerung in ein elektrisches Signal. Dann gibt das Mikrofon MC4 das Audiosignal an die Audioeingabeeinheit 210 aus. Das Audiosignal ist ein Signal, das eine Sprachäußerung des Insassen hm4, eine Sprachäußerung einer anderen Person als des Insassen hm4, von einer Audio-Vorrichtung ausgegebene Musik oder Geräusche, wie etwa ein Fahrgeräusch, enthält.

[0032] Die Audiosignale werden in die Audioeingabeeinheit 210 von jedem aus dem Mikrofon MC1, dem Mikrofon MC2, dem Mikrofon MC3 und dem Mikrofon MC4 eingegeben. Das heißt, die Audioeingabeeinheit 210 empfängt eine erste Sprachäußerung, die eine durch einen ersten Sprecher abgegebene Sprachäußerung ist. Mit anderen Worten, die Audioeingabeeinheit 210 empfängt eine durch einen ersten Sprecher, einen beliebigen aus der Vielzahl von Sprechern, abgegebene Sprachäußerung. Die Audioeingabeeinheit 210 ist ein Beispiel für die Eingabeeinheit. Dann gibt die Audioeingabeeinheit 210 das Audiosignal an die Ausfallerkennungseinheit 220 aus.

[0033] Die Ausfallerkennungseinheit 220 erkennt einen Ausfall jedes aus dem Mikrofon MC1, dem Mikrofon MC2, dem Mikrofon MC3 und dem Mikrofon MC4. Ferner bestimmt die Ausfallerkennungseinheit 220, ob eine Position des ersten Sprechers ermittelt werden kann oder nicht. Die Ausfallerkennungseinheit 220 ist ein Beispiel einer Bestimmungseinheit. Hier vergleicht das Audioverarbeitungssystem 20 die von jedem aus dem Mikrofon MC1, dem Mikrofon MC2, dem Mikrofon MC3 und dem Mikrofon MC4 ausgegebenen Audiosignale, um die Position des Sprechers zu ermitteln, der die in jedem Audiosignal enthaltene Sprachäußerung abgegeben hat. In einigen Fällen kann das Audioverarbeitungssystem 20 die Position des Sprechers nicht ermitteln, wenn ein beliebiges aus dem Mikrofon MC1, dem Mikrofon MC2, dem Mikrofon MC3 und dem Mikrofon MC4 ausgefallen ist. Daher erkennt die Ausfallerkennungseinheit 220 das Vorliegen oder Nichtvorliegen eines Ausfalls in der Vielzahl von Mikrofonen und bestimmt auf Grundlage des Erkennungsergebnis-

ses, ob die Position des ersten Sprechers ermittelt werden kann oder nicht.

[0034] Das Bestimmen des Vorliegens oder Nichtvorliegens eines Mikrofonausfalls ist nachstehend genauer beschrieben. Das Mikrofon MC1 und das Mikrofon MC2 sind an Positionen nah beieinander angeordnet. Daher sind der durch das Mikrofon MC1 empfangene Schalldruck und der durch das Mikrofon MC2 empfangene Schalldruck im Wesentlichen derselbe. Daher sind die Pegel der von dem Mikrofon MC1 und dem Mikrofon MC2 ausgegebenen Audiosignale im Wesentlichen dieselben. Jedoch kann in einem Fall, wo eins aus dem Mikrofon MC1 und dem Mikrofon MC2 ausfällt, eins aus dem Mikrofon MC1 und dem Mikrofon MC2 Schall nicht auf normale Weise aufnehmen. Daher tritt eine Differenz bei den Pegeln der von dem Mikrofon MC1 und dem Mikrofon MC2 ausgegebenen Audiosignale auf. In einem Fall, wo eine Differenz zwischen dem Pegel des von dem Mikrofon MC1 ausgegebenen Audiosignals und dem Pegel des von dem Mikrofon MC2 ausgegebenen Audiosignals gleich oder größer als ein Schwellenwert ist, bestimmt die Ausfallerkennungseinheit 220, dass ein Ausfall in einem aus dem Mikrofon MC1 und dem Mikrofon MC2 aufgetreten ist. Zum Beispiel bestimmt die Ausfallerkennungseinheit 220, dass das Mikrofon, das das Audiosignal mit niedrigerem Pegel aus den beiden Audiosignalen ausgegeben hat, ausgefallen ist.

[0035] Aus dem ähnlichen Grund bestimmt in einem Fall, wo eine Differenz zwischen dem Pegel des von dem Mikrofon MC3 ausgegebenen Audiosignals und dem Pegel des von dem Mikrofon MC4 ausgegebenen Audiosignals gleich oder größer als ein Schwellenwert ist, die Ausfallerkennungseinheit 220, dass ein Ausfall in einem aus dem Mikrofon MC3 und dem Mikrofon MC4 aufgetreten ist.

[0036] In einem Fall, wo ein Ausfall in mindestens einem aus dem Mikrofon MC1, dem Mikrofon MC2, dem Mikrofon MC3 und dem Mikrofon MC4 erkannt ist, gibt die Ausfallerkennungseinheit 220 ein Ausfallerkennungssignal aus, das angibt, dass der Ausfall erkannt wurde. Das heißt, die Ausfallerkennungseinheit 220 gibt ein Ausfallerkennungssignal aus, das angibt, ob die Position des Sprechers, der die durch die Audioeingabeeinheit 210 empfangene Sprachäußerung abgegeben hat, ermittelt werden kann oder nicht. Das Ausfallerkennungssignal ist ein Beispiel des ersten Signals. Ferner gibt die Ausfallerkennungseinheit 220 Audiosignale, die von dem Mikrofon MC1, dem Mikrofon MC2, dem Mikrofon MC3 und dem Mikrofon MC4 ausgegeben sind, an die Audioverarbeitungsvorrichtung 230 aus.

[0037] Die Audioverarbeitungsvorrichtung 230 enthält eine Signalempfangseinheit 231, eine Strahlformungs-Verarbeitungseinheit (BF-Verarbeitungsein-

heit) 232, eine Echokompensations-Verarbeitungseinheit (EC-Verarbeitungseinheit) 233, eine Übersprechkompensations-Verarbeitungseinheit (CTC-Verarbeitungseinheit) 234 und eine Spracherkennungseinheit 235.

[0038] Die Signalempfangseinheit 231 empfängt ein Ausfallerkennungssignal, das angibt, ob die Position des Sprechers, der die durch die Audioeingabeeinheit 210 empfangene Sprachäußerung abgegeben hat, ermittelt werden kann oder nicht. Die Signalempfangseinheit 231 ist ein Beispiel einer Empfangseinheit. Die Signalempfangseinheit 231 empfängt das Ausfallerkennungssignal von der Ausfallerkennungseinheit 220. Die Signalempfangseinheit 231 sendet das Ausfallerkennungssignal an die BF-Verarbeitungseinheit 232, die EC-Verarbeitungseinheit 233, die CTC-Verarbeitungseinheit 234 und die Spracherkennungseinheit 235.

[0039] Die BF-Verarbeitungseinheit 232 hebt die Sprachäußerung in der Richtung vom Objektsitz durch die Richtwirkungssteuerung an. Der Betrieb der BF-Verarbeitungseinheit 232 ist im Folgenden am Beispiel eines Falls beschrieben, wo eine Sprachäußerung in der Richtung des Fahrersitzes unter den von dem Mikrofon MC1 ausgegebenen Audiosignalen angehoben wird. Das Mikrofon MC1 und das Mikrofon MC2 sind an Positionen nah beieinander angeordnet. Daher ist angenommen, dass das von dem Mikrofon MC1 ausgegebene Audiosignal Sprachäußerungen des Insassen hm1 auf dem Fahrersitz und des Insassen hm2 auf dem Beifahrersitz enthält. Ähnlich ist angenommen, dass das von dem Mikrofon MC2 ausgegebene Audiosignal Sprachäußerungen des Insassen hm1 auf dem Fahrersitz und des Insassen hm2 auf dem Beifahrersitz enthält.

[0040] Jedoch befindet sich das Mikrofon MC1 weiter weg vom Beifahrersitz als das Mikrofon MC2. Wenn sich daher der Insasse hm2 auf dem Beifahrersitz äußert, wird die in dem von dem Mikrofon MC1 ausgegebenen Audiosignal enthaltene Sprachäußerung des Insassen hm2 auf dem Beifahrersitz verzögert gegenüber der in dem von dem Mikrofon MC2 ausgegebenen Audiosignal enthaltenen Sprachäußerung des Insassen hm2 auf dem Beifahrersitz. Daher hebt die BF-Verarbeitungseinheit 232 die Sprachäußerung in der Richtung des Objektsitzes beispielsweise durch ein Anwenden der Zeitverzögerungsverarbeitung an dem Audiosignal an. Dann gibt die BF-Verarbeitungseinheit 232 das Audiosignal, das die Sprachäußerung in der Richtung des Objektsitzes anhebt, an die EC-Verarbeitungseinheit 233 aus. Jedoch ist das Verfahren, nach dem die BF-Verarbeitungseinheit 232 die Sprachäußerung in der Richtung des Objektsitzes anhebt, nicht auf das Obige beschränkt.

[0041] Die EC-Verarbeitungseinheit 233 löscht eine Audio-Komponente außer der Sprachäußerung, die durch den Sprecher abgegeben ist, in dem von der BF-Verarbeitungseinheit 232 ausgegebenen Audiosignal. Hier ist die Audio-Komponente außer der durch den Sprecher abgegebenen Sprachäußerung beispielsweise durch eine Audio-Vorrichtung des Fahrzeugs 10 ausgegebene Musik, Fahrgeräusch oder dergleichen. Mit anderen Worten, die EC-Verarbeitungseinheit 233 führt eine Echokompensationsverarbeitung aus.

[0042] Genauer löscht die EC-Verarbeitungseinheit 233 die durch das Referenzsignal angegebene Audio-Komponente aus dem Audiosignal, das von der BF-Verarbeitungseinheit 232 ausgegeben ist. Als Ergebnis löscht die EC-Verarbeitungseinheit 233 eine Audio-Komponente außer der durch den Sprecher abgegebenen Sprachäußerung. Hier ist das Referenzsignal ein Signal, das die Audio-Komponente außer der durch den Sprecher abgegebenen Sprachäußerung angibt. Zum Beispiel ist das Referenzsignal ein Signal, das eine Audio-Komponente von durch die Audio-Vorrichtung ausgegebener Musik angibt. Als Ergebnis kann die EC-Verarbeitungseinheit 233 die Audio-Komponente außer der durch den Sprecher abgegebenen Sprachäußerung löschen durch ein Löschen der durch das Referenzsignal angegebenen Audio-Komponente.

[0043] Die CTC-Verarbeitungseinheit 234 löscht die Sprachäußerung, die aus einer beliebigen anderen Richtung als vom Objektsitz ausgegeben ist. Mit anderen Worten, die CTC-Verarbeitungseinheit 234 führt eine Übersprechkompensations-Verarbeitung aus. Die Audiosignale von allen Mikrofonen werden in die CTC-Verarbeitungseinheit 234 eingegeben, nachdem sie der Echokompensationsverarbeitung durch die EC-Verarbeitungseinheit 233 unterzogen wurden. Die CTC-Verarbeitungseinheit 234 löscht eine aus einer beliebigen Richtung außer vom Objektsitz aufgenommene Audio-Komponente, indem sie ein von einem Mikrofon außer dem Objektsitz ausgegebenes Audiosignal unter den eingegebenen Audiosignalen als Referenzsignal verwendet. Das heißt, die CTC-Verarbeitungseinheit 234 löscht die Audio-Komponente, die durch das Referenzsignal angegeben ist, aus dem auf das Mikrofon am Objektsitz bezogenen Audiosignal. Dann gibt die CTC-Verarbeitungseinheit 234 das Audiosignal nach der Übersprechkompensations-Verarbeitung an die Spracherkennungseinheit 235 aus.

[0044] Die Spracherkennungseinheit 235 gibt einen Sprachbefehl an die Elektronikvorrichtung 30 auf Grundlage des Audiosignals und des Ausfallerkennungssignals aus. Genauer erkennt die Spracherkennungseinheit 235 einen in dem Audiosignal enthaltenen Sprachbefehl durch ein Ausführen einer Spracherkennungsverarbeitung an dem von der

CTC-Verarbeitungseinheit 234 ausgegebenen Audiosignal. Ferner enthält der Sprachbefehl einen Sprecherpositionsbefehl, der ein auf die Position des Sprechers bezogener Befehl ist. Die Elektronikvorrichtung 30 führt eine Verarbeitung entsprechend dem Sprecherpositionsbefehl aus. Zum Beispiel führt die Elektronikvorrichtung 30 eine Verarbeitung zum Ändern der Temperatur der Klimaanlage, eine Verarbeitung zum Ändern der Lautstärke des Lautsprechers und eine Verarbeitung zum Öffnen und Schließen des Fensters auf Grundlage des Sprecherpositionsbefehls aus.

[0045] Der Sprecherpositionsbefehl ist ein Befehl, in dem die auszuführende Verarbeitung gemäß der Position des Sprechers bestimmt wird. Wenn beispielsweise der Insasse hm2 auf dem Beifahrersitz „Fenster öffnen“ äußert, bestimmt die Spracherkennungseinheit 235, dass eine Sprachäußerung in der Äußerung ein Sprecherpositionsbefehl ist, der eine Verarbeitung zum Öffnen des Fensters auf der linken Seite des Beifahrersitzes angibt. Wenn ferner der Insasse hm3 auf dem Sitz auf der rechten Seite des Rücksitzes „Fenster öffnen“ äußert, bestimmt die Spracherkennungseinheit 235, dass eine Sprachäußerung in der Äußerung ein Sprecherpositionsbefehl ist, der eine Verarbeitung zum Öffnen des Fensters auf der rechten Seite des Rücksitzes angibt.

[0046] Ferner enthält der Sprecherpositionsbefehl einen Fahrtbefehl, der ein auf das Fahren bezogener Befehl ist. Der Fahrtbefehl ist ein auf das Fahren des Fahrzeugs 10 bezogener Befehl. Wenn beispielsweise eine auf das Fahren des Fahrzeugs 10 bezogene Vorrichtungssteuerung durch eine Äußerung des Insassen hm3 auf einem Rücksitz usw. durchgeführt wird, von dem ursprünglich nicht angenommen ist, dass er fährt, besteht eine Möglichkeit, dass eine Steuerung in anderer Weise durchgeführt werden kann als der Absicht des Insassen hm1 auf dem Fahrersitz, was sich als gefährlich erweist. Daher ist die Spracherkennungseinheit 235 ausgelegt, imstande zu sein, zwischen dem Fahrtbefehl und anderen Sprecherpositionsbefehlen zu unterscheiden. Zum Beispiel ist der Fahrtbefehl ein Befehl zum Steuern eines Autonavigationssystems, ein Befehl zum Steuern einer Fahrzeuggeschwindigkeit durch eine Gassteuerung oder ein Befehl zum Steuern einer Fahrzeuggeschwindigkeit durch eine Bremsensteuerung.

[0047] Ferner bestimmt die Spracherkennungseinheit 235 auf Grundlage der Mikrofonposition, in die das Audiosignal, das den Sprecherpositionsbefehl enthält, eingegeben wurde, von welcher Position aus das Audiosignal geäußert wurde. Die Spracherkennungseinheit 235 bestimmt, dass das auf dem Mikrofon MC1 beruhende Audiosignal eine aus der Richtung des Fahrersitzes abgegebene Sprachäußerung ist. Die Spracherkennungseinheit 235

bestimmt, dass das auf dem Mikrofon MC2 beruhende Audiosignal eine aus der Richtung des Beifahrersitzes abgegebene Sprachäußerung ist. Die Spracherkennungseinheit 235 bestimmt, dass das auf dem Mikrofon MC3 beruhende Audiosignal eine aus der Richtung der rechten Seite des Rücksitzes abgegebene Sprachäußerung ist. Die Spracherkennungseinheit 235 bestimmt, dass das auf dem Mikrofon MC4 beruhende Audiosignal eine aus der Richtung der linken Seite des Rücksitzes abgegebene Sprachäußerung ist.

[0048] Ferner bestimmt die Spracherkennungseinheit 235 auf Grundlage des von der Ausfallerkennungseinheit 220 ausgegebenen Ausfallerkennungssignals, ob die Position des Sprechers ermittelt werden kann oder nicht. Hier können in einem Fall, wo ein beliebiges aus dem Mikrofon MC1, dem Mikrofon MC2, dem Mikrofon MC3 und dem Mikrofon MC4 ausgefallen ist, die BF-Verarbeitungseinheit 232 und die CTC-Verarbeitungseinheit 234 in einigen Fällen eine Verarbeitung nicht in einer normalen Weise ausführen. Zum Beispiel nimmt das Mikrofon MC1 eine durch den Insassen hm1 auf dem Fahrersitz abgegebene Sprachäußerung und eine durch den Insassen hm2 auf dem Beifahrersitz abgegebene Sprachäußerung auf. Wenn das Mikrofon MC2 in diesem Fall ausfällt, können die BF-Verarbeitungseinheit 232 und die CTC-Verarbeitungseinheit 234 eine Verarbeitung nicht in einer normalen Weise ausführen. Das heißt, die CTC-Verarbeitungseinheit 234 kann die Audio-Komponente der Sprachäußerung, die durch das Mikrofon MC2 aufgenommen sein sollte, nicht aus dem von dem Mikrofon MC1 ausgegebenen Audiosignal löschen. Daher wird das von dem Mikrofon MC1 ausgegebene Audiosignal in die Spracherkennungseinheit 35 eingegeben, während sowohl die durch den Insassen hm1 auf dem Fahrersitz abgegebene Sprachäußerung als auch die durch den Insassen hm2 auf dem Beifahrersitz abgegebene Sprachäußerung enthalten sind. In diesem Fall behandelt die Spracherkennungseinheit 235 auch eine durch den Insassen hm2 auf dem Beifahrersitz abgegebene Sprachäußerung, die in dem von dem Mikrofon MC1 ausgegebenen Audiosignal enthalten ist, als eine durch den Insassen hm1 auf dem Fahrersitz abgegebene Sprachäußerung. Daher bestimmt die Spracherkennungseinheit 235, ob die Position des Sprechers auf Grundlage des Ausfallerkennungssignals ermittelt werden kann oder nicht.

[0049] Hier kann in einem Fall, wo bestimmt ist, dass der in dem Audiosignal enthaltene Sprachbefehl der Sprecherpositionsbefehl ist, wenn die Position des Sprechers, der den Sprecherpositionsbefehl geäußert hat, nicht ermittelt werden kann, die Spracherkennungseinheit 235 den auszugebenden Sprecherpositionsbefehl nicht bestimmen. Zum Beispiel kann in einem Fall, wo bestimmt ist, dass das

Audiosignal den Sprecherpositionsbefehl „Fenster öffnen“ enthält, und die Position des Sprechers nicht ermittelt werden kann, die Spracherkennungseinheit 235 als den auszugebenden Sprecherpositionsbefehl nicht angeben, welches Fenster zu öffnen ist.

[0050] Daher gibt die Spracherkennungseinheit 235 an die Elektronikvorrichtung 30 einen Sprachbefehl aus, der durch Sprachäußerung angegeben ist und ein Signal zum Steuern der Elektronikvorrichtung 30 ist, und schränkt die Ausgabe eines auf eine Position eines Sprechers bezogenen Sprecherpositionsbefehls unter den Sprachbefehlen in einem Fall ein, wo das Ausfallerkennungssignal angibt, dass die Position des ersten Sprechers nicht ermittelt werden kann. Die Spracherkennungseinheit 235 ist ein Beispiel einer Spracherkennungseinheit. Mit anderen Worten, die Spracherkennungseinheit 235 gibt an die Elektronikvorrichtung 30 einen Sprachbefehl aus, der durch Sprachäußerung angegeben ist und ein Signal zum Steuern der Elektronikvorrichtung 30 ist, und schränkt die Ausgabe eines auf eine Position eines Sprechers bezogenen Sprecherpositionsbefehls unter den Sprachbefehlen in einem Fall ein, wo die Ausfallerkennungseinheit 220 bestimmt, dass die Position des ersten Sprechers nicht ermittelt werden kann.

[0051] Als Nächstes ist ein Verfahren zum Einschränken der Ausgabe des Sprecherpositionsbefehls beschrieben.

[0052] Zum Beispiel gibt in einem Fall, wo die Ausfallerkennungseinheit 220 bestimmt, dass die Position des Sprechers nicht ermittelt werden kann, die Spracherkennungseinheit 235 den Sprecherpositionsbefehl nicht aus. As Ergebnis führt die Elektronikvorrichtung 30 keine Verarbeitung entsprechend dem Sprecherpositionsbefehl aus. Daher kann die Spracherkennungseinheit 235 die Ausführung einer unbeabsichtigten Verarbeitung durch die Elektronikvorrichtung 30 sperren.

[0053] Alternativ schränkt in einem Fall, wo die Ausfallerkennungseinheit 220 den Ausfall des Mikrofons erkennt und dadurch bestimmt, dass die Position des Sprechers nicht ermittelt werden kann, die Spracherkennungseinheit 235 die Ausgabe des Sprecherpositionsbefehls ein, der durch die Sprachäußerung angegeben ist, die von dem Mikrofon eingegeben ist, das dem ausgefallenen Mikrofon beigeordnet ist. Mit anderen Worten, in einem Fall, wo ein beliebiges Mikrofon ausfällt, das zu einer Gruppe gehört, die eine Vielzahl von benachbarten Mikrofonen enthält, gibt die Spracherkennungseinheit 235 an die Elektronikvorrichtung 30 einen Sprecherpositionsbefehl nicht aus, der durch eine von den anderen zu der Gruppe gehörenden Mikrofonen eingegebenen Sprachäußerung angegeben ist. Andererseits

schränkt die Spracherkennungseinheit 235 die Ausgabe des Sprecherpositionsbefehls nicht ein, der durch die Sprachäußerung angegeben ist, die von dem Mikrofon eingegeben ist, das zu einer anderen Gruppe gehört. Das heißt, die Spracherkennungseinheit 235 gibt den Sprecherpositionsbefehl aus, der durch die Sprachäußerung angegeben ist, die von dem Mikrofon eingegeben ist, das zu einer anderen Gruppe gehört.

[0054] Zum Beispiel bilden das Mikrofon MC1 und das Mikrofon MC2 eine Gruppe. Die Audioeingabeinheit 210 empfängt die Sprachäußerung, die die erste Sprachäußerung enthält, die von der Vielzahl von Mikrofonen ausgegeben ist, die das erste Mikrofon und das dem ersten Mikrofon beigeordnete zweite Mikrofon enthält. Das erste Mikrofon ist beispielsweise das Mikrofon MC2. Das zweite Mikrofon ist beispielsweise das Mikrofon MC1. Die erste Sprachäußerung ist beispielsweise eine durch den Insassen hm2 auf dem Beifahrersitz abgegebene Sprachäußerung. Zum Beispiel nimmt das Mikrofon MC1 eine durch den Insassen hm1 auf dem Fahrersitz abgegebene Sprachäußerung und eine durch den Insassen hm2 auf dem Beifahrersitz abgegebene Sprachäußerung auf. Wenn in diesem Fall das Mikrofon MC2 ausfällt, können die BF-Verarbeitungseinheit 232 und die CTC-Verarbeitungseinheit 234 eine Verarbeitung nicht in einer normalen Weise ausführen. Daher wird das auf dem Mikrofon MC1 basierende Audiosignal in die Spracherkennungseinheit 235 eingegeben, während die durch den Insassen hm1 auf dem Fahrersitz abgegebene Sprachäußerung und die durch den Insassen hm2 auf dem Beifahrersitz abgegebene Sprachäußerung enthalten sind. Daher kann die Spracherkennungseinheit 235 fälschlicherweise die durch den Insassen hm2 auf dem Beifahrersitz abgegebene Sprachäußerung als die durch den Insassen hm1 auf dem Fahrersitz abgegebene Sprachäußerung bestimmen. Da andererseits das Mikrofon MC3 und das Mikrofon MC4 von dem Insassen hm1 auf dem Fahrersitz und dem Insassen hm2 auf dem Beifahrersitz getrennt sind, besteht eine geringe Möglichkeit, dass die durch den Insassen hm1 auf dem Fahrersitz abgegebene Sprachäußerung und die durch den Insassen hm2 auf dem Beifahrersitz abgegebene Sprachäußerung aufgenommen werden. Daher gibt in einem Fall, wo die Ausfallerkennungseinheit 220 den Ausfall des ersten Mikrofons erkennt und bestimmt, dass die Position des ersten Sprechers nicht ermittelt werden kann, die Spracherkennungseinheit 235 den Sprecherpositionsbefehl unter den Sprecherpositionsbefehlen nicht aus, der durch die Sprachäußerung angegeben ist, die von dem zweiten Mikrofon eingegeben ist. Der erste Sprecher ist beispielsweise der Insasse hm2 auf dem Beifahrersitz.

[0055] Alternativ ändert in einem Fall, wo die Ausfallerkennungseinheit 220 bestimmt, dass die Position des ersten Sprechers nicht ermittelt werden kann, die Spracherkennungseinheit 235 die Priorität der Ausgabe des auf das Fahren bezogenen Fahrtbefehls unter den Sprecherpositionsbefehlen. Wenn sie beispielsweise eine Vielzahl von Sprecherpositionsbefehlen empfängt, weist die Spracherkennungseinheit 235 den Sprecherpositionsbefehl einer der Prioritätsstufen zu, die in eine Vielzahl von Stufen aufgeteilt sind. Dann gibt die Spracherkennungseinheit 235 einen Sprecherpositionsbefehl mit einer höheren Priorität als ein Schwellenwert an die Elektronikvorrichtung 30 aus. Das heißt, die Spracherkennungseinheit 235 veranlasst die Elektronikvorrichtung 30, den Sprecherpositionsbefehl vorzugsweise auszuführen. Andererseits gibt die Spracherkennungseinheit 235 einen Sprecherpositionsbefehl mit einer niedrigeren Priorität als der Schwellenwert nicht an die Elektronikvorrichtung 30 aus. Wie oben beschrieben, ändert in einem Fall, wo die Ausfallerkennungseinheit 220 bestimmt, dass die Position des Sprechers nicht angegeben werden kann, die Spracherkennungseinheit 235 die Priorität der Ausgabe des Fahrtbefehls.

[0056] Zum Beispiel erhöht in einem Fall, wo die Ausfallerkennungseinheit 220 bestimmt, dass die Position des ersten Sprechers nicht angegeben werden kann, die Spracherkennungseinheit 235 die Priorität der Ausgabe des Fahrtbefehls. Als Ergebnis verhindert, wenn ein beliebiges Mikrofon ausfällt, die Spracherkennungseinheit 235, dass der auf das Fahren bezogene Vorgang durch Sprachäußerung deaktiviert wird.

[0057] Alternativ verringert in einem Fall, wo die Ausfallerkennungseinheit 220 bestimmt, dass die Position des Sprechers nicht ermittelt werden kann, die Spracherkennungseinheit 235 die Priorität der Ausgabe des Fahrtbefehls. Als Ergebnis verhindert, wenn ein beliebiges Mikrofon ausfällt, die Spracherkennungseinheit 235, dass ein auf das Fahren bezogener Vorgang durch eine Sprachäußerung einer Person durchgeführt wird, die ursprünglich nicht am Fahren beteiligt ist, wie etwa des Insassen hm4 auf dem Rücksitz.

[0058] Als Nächstes ist eine Arbeitsweise des Audioverarbeitungssystems 20 gemäß der ersten Ausführungsform beschrieben. **Fig. 4** ist ein Flussdiagramm, das ein Beispiel eines Betriebs des Audioverarbeitungssystems 20 gemäß der ersten Ausführungsform darstellt.

[0059] Die Audioeingabeeinheit 210 empfängt Audiosignale von dem Mikrofon MC1, dem Mikrofon MC2, dem Mikrofon MC3 und dem Mikrofon MC4 (Schritt S11).

[0060] Die Ausfallerkennungseinheit 220 bestimmt, ob ein beliebiges aus dem Mikrofon MC1, dem Mikrofon MC2, dem Mikrofon MC3 und dem Mikrofon MC4 ausgefallen ist oder nicht, auf Grundlage des von der Audioeingabeeinheit 210 ausgegebenen Audiosignals (Schritt S12).

[0061] Die Ausfallerkennungseinheit 220 gibt ein Ausfallerkennungssignal aus, das angibt, ob ein beliebiges aus dem Mikrofon MC1, dem Mikrofon MC2, dem Mikrofon MC3 und dem Mikrofon MC4 ausgefallen ist oder nicht, an die Signalempfangseinheit 231 der Audioverarbeitungsvorrichtung 230 aus (Schritt S13).

[0062] Die Signalempfangseinheit 231 sendet das Ausfallerkennungssignal, das angibt, ob ein beliebiges aus dem Mikrofon MC1, dem Mikrofon MC2, dem Mikrofon MC3 und dem Mikrofon MC4 ausgefallen ist oder nicht, an die BF-Verarbeitungseinheit 232, die EC-Verarbeitungseinheit 233, die CTC-Verarbeitungseinheit 234 und die Spracherkennungseinheit 235 (Schritt S14).

[0063] Die Spracherkennungseinheit 235 bestimmt auf Grundlage des von der Signalempfangseinheit 231 ausgegebenen Ausfallerkennungssignals, ob die Position des Sprechers der Sprachäußerung, die in dem über die BF-Verarbeitungseinheit 232, die EC-Verarbeitungseinheit 233 und die CTC-Verarbeitungseinheit 234 eingegebenen Audiosignal enthalten ist, ermittelt werden kann oder nicht (Schritt S15).

[0064] In einem Fall, wo die Position des Sprechers angegeben werden kann (Ja in Schritt S15), gibt die Spracherkennungseinheit 235 den in dem Audiosignal enthaltenen Sprachbefehl an die Elektronikvorrichtung 30 aus (Schritt S16). Als Ergebnis veranlasst die Spracherkennungseinheit 235 die Elektronikvorrichtung 30, die durch den Sprachbefehl angegebene Verarbeitung auszuführen.

[0065] In einem Fall, wo die Position des Sprechers nicht angegeben werden kann (Nein in Schritt S15), bestimmt die Spracherkennungseinheit 235, ob der in dem Audiosignal enthaltene Sprachbefehl ein anderer Befehl als der Sprecherpositionsbefehl ist oder nicht (Schritt S17). In einem Fall, wo der Befehl ein anderer Befehl ist als der Sprecherpositionsbefehl (Ja in Schritt S17) fährt die Spracherkennungseinheit 235 mit Schritt S16 fort.

[0066] In einem Fall, wo der in dem Audiosignal enthaltene Sprachbefehl der Sprecherpositionsbefehl ist (Nein in Schritt S17), schränkt die Spracherkennungseinheit 235 die Ausgabe des Sprecherpositionsbefehls ein (Schritt S18). Das heißt, wie in Schritt S16 gezeigt, die Spracherkennungseinheit 235 gibt an die Elektronikvorrichtung 30 einen

Sprachbefehl aus, der durch Sprachäußerung angegeben ist und ein Signal zum Steuern der Objektvorrichtung ist. Jedoch schränkt in einem Fall, wo in Schritt S15 bestimmt ist, dass die Position des ersten Sprechers nicht angegeben werden kann, die Spracherkennungseinheit 235 die Ausgabe des auf die Position des Sprechers bezogenen Sprecherpositionsbefehls unter den Sprachbefehlen ein. Als Ergebnis schränkt die Spracherkennungseinheit 235a die Ausführung der durch den Sprachbefehl angegebenen Verarbeitung ein.

[0067] Wie oben beschrieben, beendet das Audioverarbeitungssystem 20 die Verarbeitung.

[0068] Wie oben beschrieben, empfängt gemäß der ersten Ausführungsform die Audioeingabeeinheit 210 die durch den ersten Sprecher, der ein beliebiger aus der Vielzahl von Sprechern ist, abgegebene erste Sprachäußerung. Die Ausfallerkennungseinheit 220 bestimmt, ob es möglich ist oder nicht, die Position des ersten Sprechers, der die durch die Audioeingabeeinheit 210 empfangene erste Sprachäußerung abgegeben hat, durch ein Erkennen eines Ausfalls des Mikrofons MC1, des Mikrofons MC2, des Mikrofons MC3 und des Mikrofons MC4 zu ermitteln. Dann gibt die Spracherkennungseinheit 235, die durch Sprachäußerung angegeben ist, an die Elektronikvorrichtung 30 einen Sprachbefehl aus, der ein Signal zum Steuern der Objektvorrichtung ist, schränkt die Ausgabe eines auf eine Position eines Sprechers bezogenen Sprecherpositionsbefehls unter den durch Sprachäußerung angegebenen Sprachbefehlen ein, wenn sie bestimmt, dass die Position des ersten Sprechers nicht ermittelt werden kann. Dies bewirkt, dass die Ausführung einer unbeabsichtigten Verarbeitung eingeschränkt wird, und das Audioverarbeitungssystem 20 kann eine angemessene Verarbeitung sogar in einem Fall ausführen, wo die Position des Sprechers nicht ermittelt werden kann.

(Zweite Ausführungsform)

[0069] Als Nächstes ist ein Audioverarbeitungssystem 20a gemäß einer zweiten Ausführungsform beschrieben. In der zweiten Ausführungsform sind Aspekte beschrieben, die sich von denjenigen der ersten Ausführungsform unterscheiden, und die gleichen Aspekte wie die der ersten Ausführungsform sind kurz beschrieben, oder ihre Beschreibung ist weggelassen.

[0070] **Fig. 5** ist ein Blockschaltbild, das ein Beispiel eines Aufbaus des Audioverarbeitungssystems 20a gemäß der zweiten Ausführungsform darstellt. Eine Audioverarbeitungsvorrichtung 230a des Audioverarbeitungssystems 20a in der zweiten Ausführungsform unterscheidet sich von dem Audioverarbeitungssystem 20 in der ersten Ausführungsform

darin, dass eine Sprechererkennungseinheit 236 vorgesehen ist.

[0071] Die Sprechererkennungseinheit 236 bestimmt, ob die erste Stimme, die eine durch den ersten Sprecher abgegebene Stimme ist, der ein beliebiger aus der Vielzahl von Sprechern ist, eine Stimme eines vorab registrierten Registranten ist oder nicht. Die Sprechererkennungseinheit 236 ist ein Beispiel einer Sprechererkennungseinheit. Genauer vergleicht die Sprechererkennungseinheit 236 das Audiosignal des vorab registrierten Registranten mit dem von der CTC-Verarbeitungseinheit 234 ausgegebenen Audiosignal, um zu bestimmen, ob die in dem von der CTC-Verarbeitungseinheit 234 ausgegebenen Audiosignal enthaltene Stimme eine auf der Äußerung des vorab registrierten Registranten basierende Stimme ist. Zum Beispiel bestimmt die Sprechererkennungseinheit 236, ob die in dem Audiosignal enthaltene Stimme die Stimme des Besitzers des Fahrzeugs 10 ist. Dann gibt die Sprechererkennungseinheit 236 ein Erkennungsergebnissignal, das angibt, ob der Sprecher, der die in dem Audiosignal enthaltene Sprachäußerung abgegeben hat, als der Registrant bestimmt wurde oder nicht, an die Spracherkennungseinheit 235a aus.

[0072] Die Spracherkennungseinheit 235a gibt den Sprecherpositionsbefehl unter der Bedingung aus, dass die Sprechererkennungseinheit 236 bestimmt hat, dass die erste Sprachäußerung eine Äußerung durch den Registranten ist. Genauer gibt in einem Fall, wo die Ausfallerkennungseinheit 220 bestimmt, dass die Position des Sprechers ermittelt werden kann, die Spracherkennungseinheit 235a den Sprecherpositionsbefehl ungeachtet dessen aus, ob die Äußerung durch den vorab registrierten Registranten erfolgt ist oder nicht. Ferner gibt in einem Fall, wo die Ausfallerkennungseinheit 220 bestimmt, dass die Position des Sprechers nicht ermittelt werden kann, die Spracherkennungseinheit 235a den Sprecherpositionsbefehl unter der Bedingung aus, dass die Sprechererkennungseinheit 236 erkannt hat, dass die Äußerung durch den Registranten abgegeben ist. Zum Beispiel führt die Spracherkennungseinheit 235a die Verarbeitung des Sprecherpositionsbefehls unter der Bedingung aus, dass die Stimme die Stimme des vorab registrierten Besitzers des Fahrzeugs 10 ist. Andererseits schränkt in einem Fall, wo die Ausfallerkennungseinheit 220 bestimmt, dass die Position des Sprechers nicht ermittelt werden kann, die Spracherkennungseinheit 235a die Ausgabe des Sprecherpositionsbefehls unter der Bedingung ein, dass die Sprechererkennungseinheit 236 erkannt hat, dass die Äußerung durch den Registranten gemacht ist.

[0073] Als Nächstes ist eine Arbeitsweise der Audioverarbeitungssystems 20a gemäß der zweiten Ausführungsform beschrieben. **Fig. 6** ist ein Flussdia-

gramm, das ein Beispiel eines Betriebs des Audioverarbeitungssystems 20a gemäß der zweiten Ausführungsform darstellt.

[0074] Die Audioeingabeeinheit 210 empfängt Audiosignale von dem Mikrofon MC1, dem Mikrofon MC2, dem Mikrofon MC3 und dem Mikrofon MC4 (Schritt S21).

[0075] Die Ausfallerkennungseinheit 220 bestimmt, ob ein beliebiges aus dem Mikrofon MC1, dem Mikrofon MC2, dem Mikrofon MC3 und dem Mikrofon MC4 ausgefallen ist oder nicht, auf Grundlage des von der Audioeingabeeinheit 210 ausgegebenen Audiosignals (Schritt S22).

[0076] Die Ausfallerkennungseinheit 220 gibt das Ausfallerkennungssignal aus, das angibt, ob ein beliebiges aus dem Mikrofon MC1, dem Mikrofon MC2, dem Mikrofon MC3 und dem Mikrofon MC4 ausgefallen ist oder nicht, an die Signalempfangseinheit 231 der Audioverarbeitungsvorrichtung 230a aus (Schritt S23).

[0077] Die Signalempfangseinheit 231 sendet das Ausfallerkennungssignal, das angibt, ob das Mikrofon MC1, das Mikrofon MC2, das Mikrofon MC3 oder das Mikrofon MC4 ausgefallen ist oder nicht, an die BF-Verarbeitungseinheit 232, die EC-Verarbeitungseinheit 233, die CTC-Verarbeitungseinheit 234 und die Spracherkennungseinheit 235a (Schritt S24).

[0078] Die Spracherkennungseinheit 235a bestimmt auf Grundlage des von der Signalempfangseinheit 231 ausgegebenen Ausfallerkennungssignals, ob die Position des Sprechers der Sprachäußerung, die in dem über die BF-Verarbeitungseinheit 232, die EC-Verarbeitungseinheit 233 und die CTC-Verarbeitungseinheit 234 eingegebenen Audiosignal enthalten ist, ermittelt werden kann oder nicht (Schritt S25).

[0079] In einem Fall, wo die Position des Sprechers ermittelt werden kann (Ja in Schritt S25), gibt die Spracherkennungseinheit 235a den in dem Audiosignal enthaltenen Sprachbefehl an die Elektronikvorrichtung 30 aus (Schritt S26). Als Ergebnis veranlasst die Spracherkennungseinheit 235a die Elektronikvorrichtung 30, die durch den Sprachbefehl angegebene Verarbeitung auszuführen.

[0080] In einem Fall, wo die Position des Sprechers nicht angegeben werden kann (Nein in Schritt S25), bestimmt die Spracherkennungseinheit 235a auf Grundlage des Erkennungsergebnissignals, ob die in dem Audiosignal enthaltene Sprachäußerung die durch den Registranten gemachte Äußerung ist oder nicht (Schritt S27).

[0081] In einem Fall, wo die in dem Audiosignal enthaltene Sprachäußerung auf der Äußerung des Registranten beruht (Ja in Schritt S27), fährt die Spracherkennungseinheit 235a mit Schritt S26 fort.

[0082] In einem Fall, wo die in dem Audiosignal enthaltene Sprachäußerung nicht die durch den Registranten gemachte Äußerung ist (Nein in Schritt S27), bestimmt die Spracherkennungseinheit 235a, ob der in dem Audiosignal enthaltene Sprachbefehl ein anderer Befehl als der Sprecherpositionsbefehl ist oder nicht (Schritt S28). In einem Fall, wo der in dem Audiosignal enthaltene Sprachbefehl ein anderer Befehl ist als der Sprecherpositionsbefehl (Ja in Schritt S28), fährt die Spracherkennungseinheit 235a mit Schritt S26 fort.

[0083] In einem Fall, wo der in dem Audiosignal enthaltene Sprachbefehl der Sprecherpositionsbefehl ist (Nein in Schritt S28), schränkt die Spracherkennungseinheit 235a die Ausgabe des Sprecherpositionsbefehls ein (Schritt S29). Als Ergebnis schränkt die Spracherkennungseinheit 235a die Ausführung der durch den Sprachbefehl angegebenen Verarbeitung ein.

[0084] Wie oben beschrieben, beendet das Audioverarbeitungssystem 20a die Verarbeitung.

[0085] Wie oben beschrieben, bestimmt gemäß der zweiten Ausführungsform die Sprechererkennungseinheit 236, ob die durch den ersten Sprecher abgegebene erste Sprachäußerung, der ein beliebiger aus der Vielzahl von Sprechern ist, eine Sprachäußerung eines vorab registrierten Registranten ist oder nicht. Dann gibt die Spracherkennungseinheit 235a an die Elektronikvorrichtung 30 den Sprecherpositionsbefehl unter der Bedingung aus, dass die Sprechererkennungseinheit 236 bestimmt hat, dass die erste Sprachäußerung eine durch den Registranten abgegebene Sprachäußerung ist. Als Ergebnis führt die Spracherkennungseinheit 30 die Verarbeitung des Sprecherpositionsbefehls unter der Bedingung aus, dass die Sprachäußerung durch einen bestimmten Registranten abgegeben ist, wie etwa den Besitzer des Fahrzeugs 10. Andererseits schränkt in einem Fall einer durch eine andere Person als den Registranten abgegebenen Sprachäußerung die Spracherkennungseinheit 235a die Ausgabe des Sprecherpositionsbefehls ein. Dies bewirkt, dass die Ausführung einer unbeabsichtigten Verarbeitung eingeschränkt wird, und das Audioverarbeitungssystem 20a kann eine angemessene Verarbeitung sogar in einem Fall ausführen, wo die Position des Sprechers nicht ermittelt werden kann.

(Erste Modifikation)

[0086] Nachfolgend ist eine erste Modifikation der ersten Ausführungsform oder der zweiten Ausführungsform beschrieben.

[0087] Die Audioverarbeitungsvorrichtung 230 in der ersten Ausführungsform und die Audioverarbeitungsvorrichtung 230a in der zweiten Ausführungsform enthalten die CTC-Verarbeitungseinheit 234. Jedoch brauchen die Audioverarbeitungsvorrichtung 230 und die Audioverarbeitungsvorrichtung 230a die CTC-Verarbeitungseinheit 234 nicht zu enthalten. Ferner enthalten die in **Fig. 3** dargestellte Audioverarbeitungsvorrichtung 230 und die in **Fig. 5** dargestellte Audioverarbeitungsvorrichtung 230a die EC-Verarbeitungseinheit 233 an einer nachfolgenden Stufe der BF-Verarbeitungseinheit 232. Jedoch können die Audioverarbeitungsvorrichtung 230 und die Audioverarbeitungsvorrichtung 230a die BF-Verarbeitungseinheit 232 an einer nachfolgenden Stufe der EC-Verarbeitungseinheit 233 enthalten.

(Zweite Modifikation)

[0088] Nachfolgend ist eine zweite Modifikation der ersten Ausführungsform oder der zweiten Ausführungsform beschrieben.

[0089] In einem Fall, wo das Mikrofon MC3 oder das Mikrofon MC4 ausfällt, das nahe dem in **Fig. 1** dargestellten Rücksitz eingebaut ist, können die Audioverarbeitungsvorrichtung 230 gemäß der ersten Ausführungsform und die Audioverarbeitungsvorrichtung 230a gemäß der zweiten Ausführungsform eine teilweise Mehrzonen-Schallaufnahme mit einem nicht ausgefallenen Mikrofon ausführen. Genauer nehmen, wenn das Mikrofon MC3 ausfällt, die Audioverarbeitungsvorrichtung 230 und die Audioverarbeitungsvorrichtung 230a die Sprachäußerung auf dem Rücksitz unter Verwendung des Mikrofons MC4 auf. Alternativ nehmen, wenn das Mikrofon MC4 ausfällt, die Audioverarbeitungsvorrichtung 230 und die Audioverarbeitungsvorrichtung 230a die Sprachäußerung auf dem Rücksitz unter Verwendung des Mikrofons MC3 auf.

[0090] In der ersten Ausführungsform, der zweiten Ausführungsform und der ersten und der zweiten Modifikation davon ist beschrieben, dass die Funktionen des Audioverarbeitungssystems 20 und des Audioverarbeitungssystems 20a durch den DSP 2001 umgesetzt werden, der ein bestimmtes Computerprogramm ausführt. Ein Computerprogramm, um einen Computer zu veranlassen, die Funktionen des Audioverarbeitungssystems 20 und des Audioverarbeitungssystems 20a zu verwirklichen, kann vorgesehen sein, indem es vorab im ROM 2003 gespeichert ist. Das Computerprogramm zum Veranlassen eines Computers, die Funktionen des Audioverarbei-

tungssystems 20 und des Audioverarbeitungssystems 20a zu verwirklichen, kann vorgesehen sein, indem es als eine Datei in einem installierbaren Format oder einem ausführbaren Format auf einem computerlesbaren Aufzeichnungsmedium aufgezeichnet ist, wie etwa einem Compact-Disk-Nur-Lesespeicher (CD-ROM), einer flexiblen Platte (FD), einer beschreibbaren Compact-Disk (CD-R), einer Digital Versatile Disc (DVD), einem USB-Speicher (Universal Serial Bus) oder einer SD-Karte (Secure Digital Card).

[0091] Ferner kann ein Computerprogramm, um einen Computer zu veranlassen, die Funktionen des Audioverarbeitungssystems 20 und des Audioverarbeitungssystems 20a zu verwirklichen, auf einem Computer gespeichert werden, der mit einem Netzwerk, wie etwa dem Internet, verbunden ist, und vorgesehen werden, indem es über das Netzwerk heruntergeladen wird. Ferner kann ein Computerprogramm, um einen Computer zu veranlassen, die Funktionen des Audioverarbeitungssystems 20 und des Audioverarbeitungssystems 20a zu verwirklichen, über ein Netzwerk, wie etwa das Internet, vorgesehen oder verteilt werden.

[0092] Ferner können einige oder alle der Funktionen des Audioverarbeitungssystems 20 und des Audioverarbeitungssystems 20a durch eine Logikschaltung verwirklicht sein. Einige oder alle der Funktionen des Audioverarbeitungssystems 20 und des Audioverarbeitungssystems 20a können durch eine analoge Schaltung verwirklicht sein. Einige oder alle der Funktionen des Audioverarbeitungssystems 20 und des Audioverarbeitungssystems 20a können durch ein Field Programmable Gate Array (FPGA), einen anwendungsspezifischen integrierten Schaltkreis (ASIC) oder dergleichen verwirklicht sein.

[0093] Obwohl vorstehend mehrere Ausführungsformen der vorliegenden Offenbarung beschrieben sind, sind diese Ausführungsformen als Beispiele dargelegt und sollen den Geltungsbereich der Erfindung nicht einschränken. Diese Ausführungsformen können in verschiedenen anderen Formen umgesetzt werden, und verschiedene Auslassungen, Ersetzungen und Änderungen können vorgenommen werden, ohne vom Kern der Erfindung abzuweichen. Diese Ausführungsformen und Modifikationen davon gehören zum Geltungsbereich und Kern der Erfindung und sind in der in den Ansprüchen beschriebenen Erfindung und dem äquivalenten Geltungsbereich davon enthalten.

Bezugszeichenliste

| | |
|----|-------------|
| 5 | Audiosystem |
| 10 | Fahrzeug |

| | |
|-------------------------|---|
| 20, 20a, 21, 22, 23, 24 | Audioverarbeitungssystem |
| 30 | Elektronikvorrichtung |
| 210 | Audioeingabeeinheit |
| 220 | Ausfallerkennungseinheit |
| 230, 230a | Audioverarbeitungsvorrichtung |
| 231 | Signalempfangseinheit |
| 232 | BF-Verarbeitungseinheit (Strahlformungs-Verarbeitungseinheit) |
| 233 | EC-Verarbeitungseinheit (Echokompensations-Verarbeitungseinheit) |
| 234 | CTC-Verarbeitungseinheit (Übersprechkompensations-Verarbeitungseinheit) |
| 235, 235a | Spracherkennungseinheit |
| 236 | Sprechererkennungseinheit |
| hm1, hm2, hm3, hm4 | Insasse |
| MC1, MC2, MC3, MC4 | Mikrofon |
| 2001 DSP | (digitaler Signalprozessor) |
| 2002 RAM | (Direktzugriffsspeicher) |
| 2003 ROM | (Nur-Lese-Speicher) |
| 2004 | E/A-Schnittstelle (Ein-/Ausgabeschnittstelle) |

ZITATE ENTHALTEN IN DER BESCHREIBUNG

Diese Liste der vom Anmelder aufgeführten Dokumente wurde automatisiert erzeugt und ist ausschließlich zur besseren Information des Lesers aufgenommen. Die Liste ist nicht Bestandteil der deutschen Patent- bzw. Gebrauchsmusteranmeldung. Das DPMA übernimmt keinerlei Haftung für etwaige Fehler oder Auslassungen.

Zitierte Patentliteratur

- JP 200790611 A [0004]

Patentansprüche

1. Audioverarbeitungssystem, umfassend:
 eine Eingabeeinheit, die eine erste Sprachäußerung empfängt, die eine durch einen ersten Sprecher abgegebene Sprachäußerung ist;
 eine Bestimmungseinheit, die bestimmt, ob eine Position des ersten Sprechers ermittelt werden kann oder nicht; und
 eine Spracherkennungseinheit, die an eine Objektvorrichtung einen Sprachbefehl ausgibt, der durch eine Sprachäußerung angegeben ist und ein Signal zum Steuern der Objektvorrichtung ist, wobei die Spracherkennungseinheit die Ausgabe eines auf eine Position eines Sprechers bezogenen Sprecherpositionsbefehls unter dem Sprachbefehl in einem Fall einschränkt, wo die Bestimmungseinheit bestimmt, dass die Position des ersten Sprechers nicht ermittelt werden kann.

2. Audioverarbeitungssystem nach Anspruch 1, wobei die Spracherkennungseinheit den Sprecherpositionsbefehl in einem Fall nicht ausgibt, wo die Bestimmungseinheit bestimmt, dass die Position des ersten Sprechers nicht ermittelt werden kann.

3. Audioverarbeitungssystem nach Anspruch 1, wobei
 die Eingabeeinheit eine Sprachäußerung empfängt, die die erste Sprachäußerung enthält, die von einer Vielzahl von Mikrofonen ausgegeben ist, die ein erstes Mikrofon und ein dem ersten Mikrofon beigeordnetes zweites Mikrofon enthält,
 die Bestimmungseinheit ein Erkennen des Vorliegens oder Nichtvorliegens eines Ausfalls in der Vielzahl von Mikrofonen durchführt und auf Grundlage eines Erkennungsergebnisses bestimmt, ob die Position des ersten Sprechers ermittelt werden kann oder nicht, und
 die Spracherkennungseinheit einen Sprecherpositionsbefehl, der durch eine Sprachäußerung angegeben ist, die von dem zweiten Mikrofon unter dem Sprecherpositionsbefehl eingegeben ist, in einem Fall nicht ausgibt, wo die Bestimmungseinheit einen Ausfall des ersten Mikrofons erkennt und bestimmt, dass die Position des ersten Sprechers nicht ermittelt werden kann.

4. Audioverarbeitungssystem nach einem beliebigen der Ansprüche 1 bis 3, wobei die Spracherkennungseinheit die Priorität der Ausgabe eines auf einen Fahrersitz bezogenen Fahrersitzbefehls unter dem Sprecherpositionsbefehl in einem Fall ändert, wo die Bestimmungseinheit bestimmt, dass die Position des ersten Sprechers nicht ermittelt werden kann.

5. Audioverarbeitungssystem nach Anspruch 4, wobei die Spracherkennungseinheit die Priorität

der Ausgabe des Fahrersitzbefehls in einem Fall erhöht, wo die Bestimmungseinheit bestimmt, dass die Position des ersten Sprechers nicht ermittelt werden kann.

6. Audioverarbeitungssystem nach einem beliebigen der Ansprüche 1 bis 5, weiter umfassend eine Sprecherbestimmungseinheit, die bestimmt, ob die erste Stimme eine Stimme eines vorab registrierten Registranten ist, wobei die Spracherkennungseinheit den Sprecherpositionsbefehl unter der Bedingung ausgibt, dass die Sprecherbestimmungseinheit bestimmt, dass die erste Stimme eine Stimme des Registranten ist.

7. Audioverarbeitungsvorrichtung, umfassend:
 eine Empfangseinheit, die ein erstes Signal empfängt, das angibt, ob eine Position eines ersten Sprechers, der eine erste Sprachäußerung abgegeben hat, ermittelt werden kann oder nicht; und
 eine Spracherkennungseinheit, die an eine Objektvorrichtung einen Sprachbefehl ausgibt, der durch eine Sprachäußerung angegeben ist und ein Signal zum Steuern der Objektvorrichtung ist, wobei die Spracherkennungseinheit die Ausgabe eines auf eine Position eines Sprechers bezogenen Sprecherpositionsbefehls unter dem Sprachbefehl in einem Fall einschränkt, wo das erste Signal angibt, dass die Position des ersten Sprechers nicht ermittelt werden kann.

8. Audioverarbeitungsverfahren, umfassend:
 einen Eingabeschritt, der eine durch einen ersten Sprecher abgegebene erste Sprachäußerung empfängt;
 einen Bestimmungsschritt, der bestimmt, ob eine Position des ersten Sprechers ermittelt werden kann oder nicht; und
 einen Ausgabeschritt, der an eine Objektvorrichtung einen Sprachbefehl ausgibt, der durch eine Sprachäußerung angegeben ist und ein Signal zum Steuern der Objektvorrichtung ist, wobei der Ausgabeschritt die Ausgabe eines auf eine Position eines Sprechers bezogenen Sprecherpositionsbefehls unter dem Sprachbefehl in einem Fall einschränkt, wo im Bestimmungsschritt bestimmt ist, dass die Position des ersten Sprechers nicht ermittelt werden kann.

Es folgen 6 Seiten Zeichnungen

Anhängende Zeichnungen

FIG.1

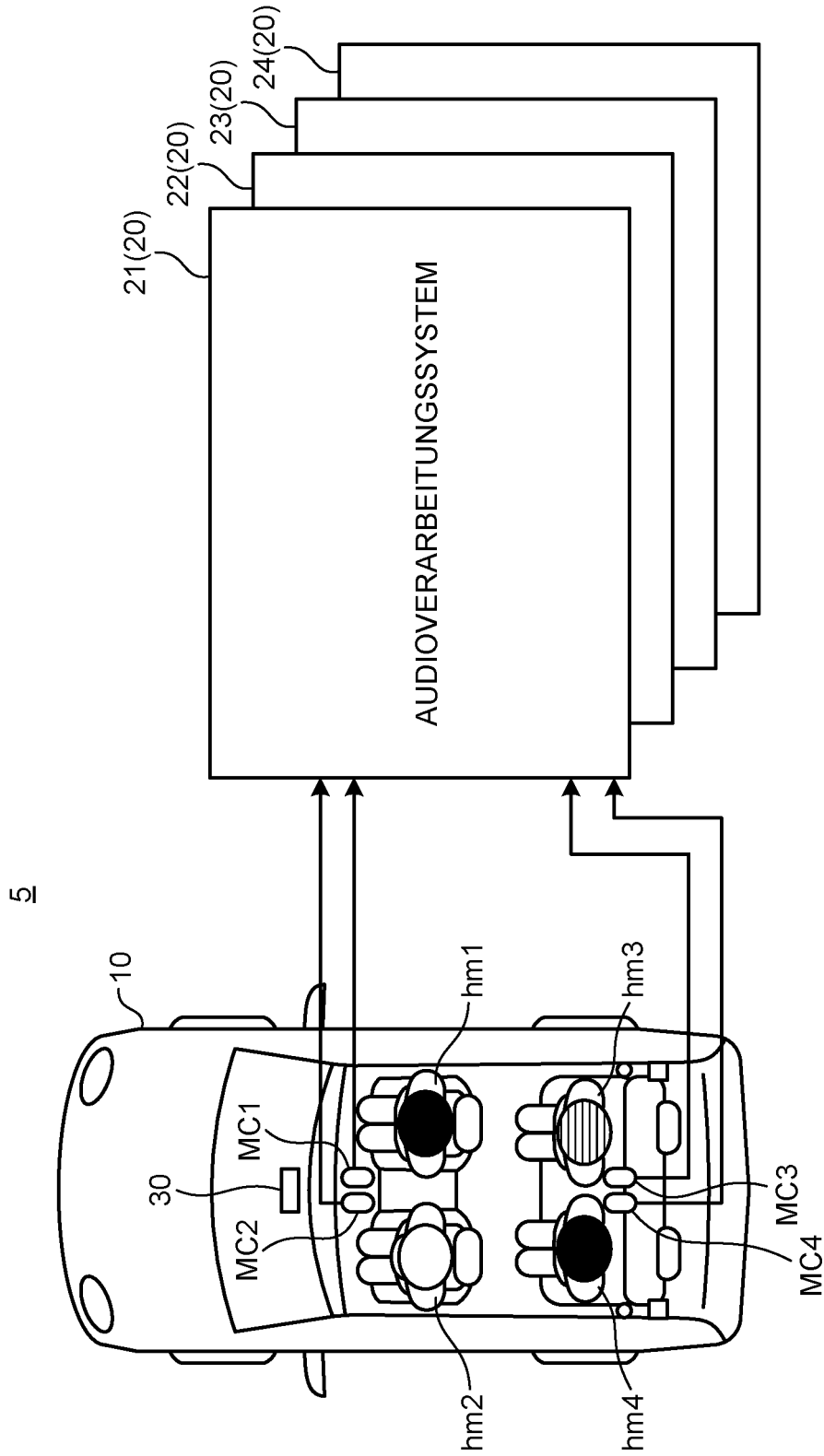


FIG.2

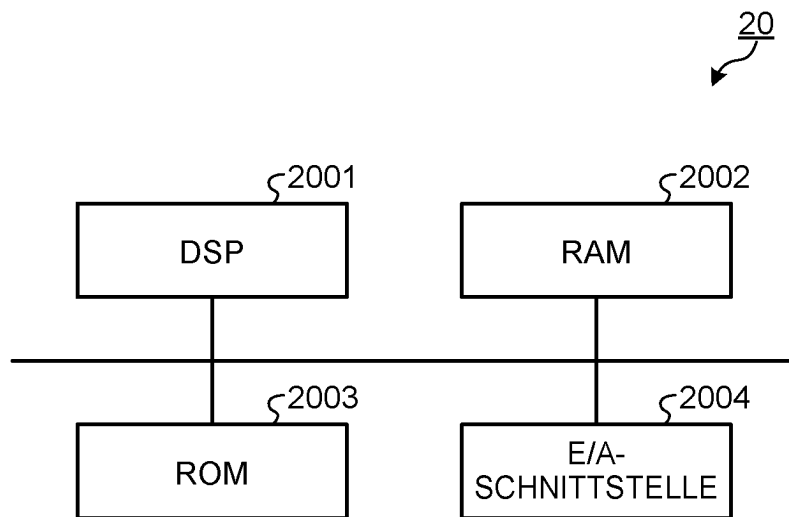


FIG.3

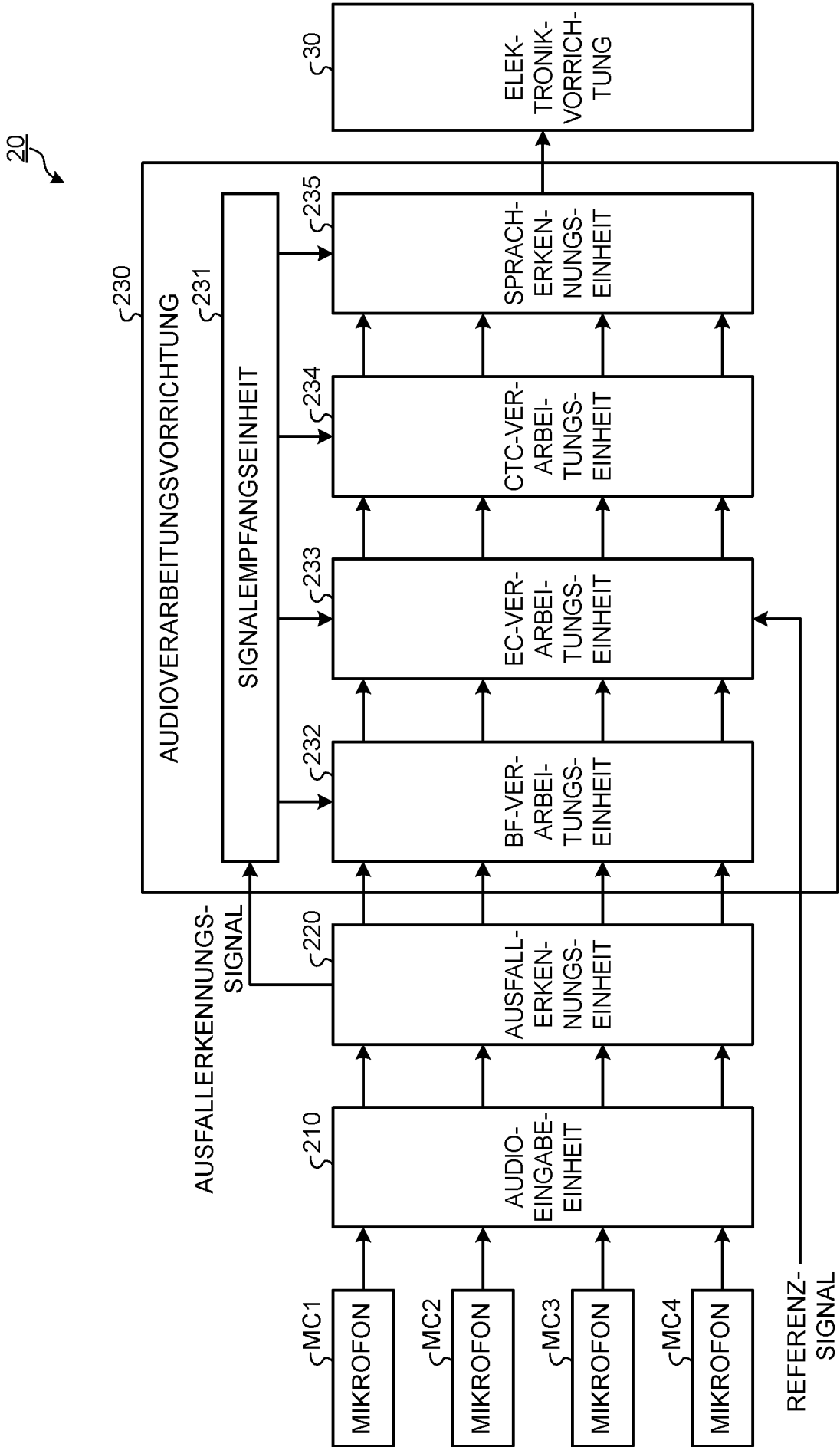


FIG.4

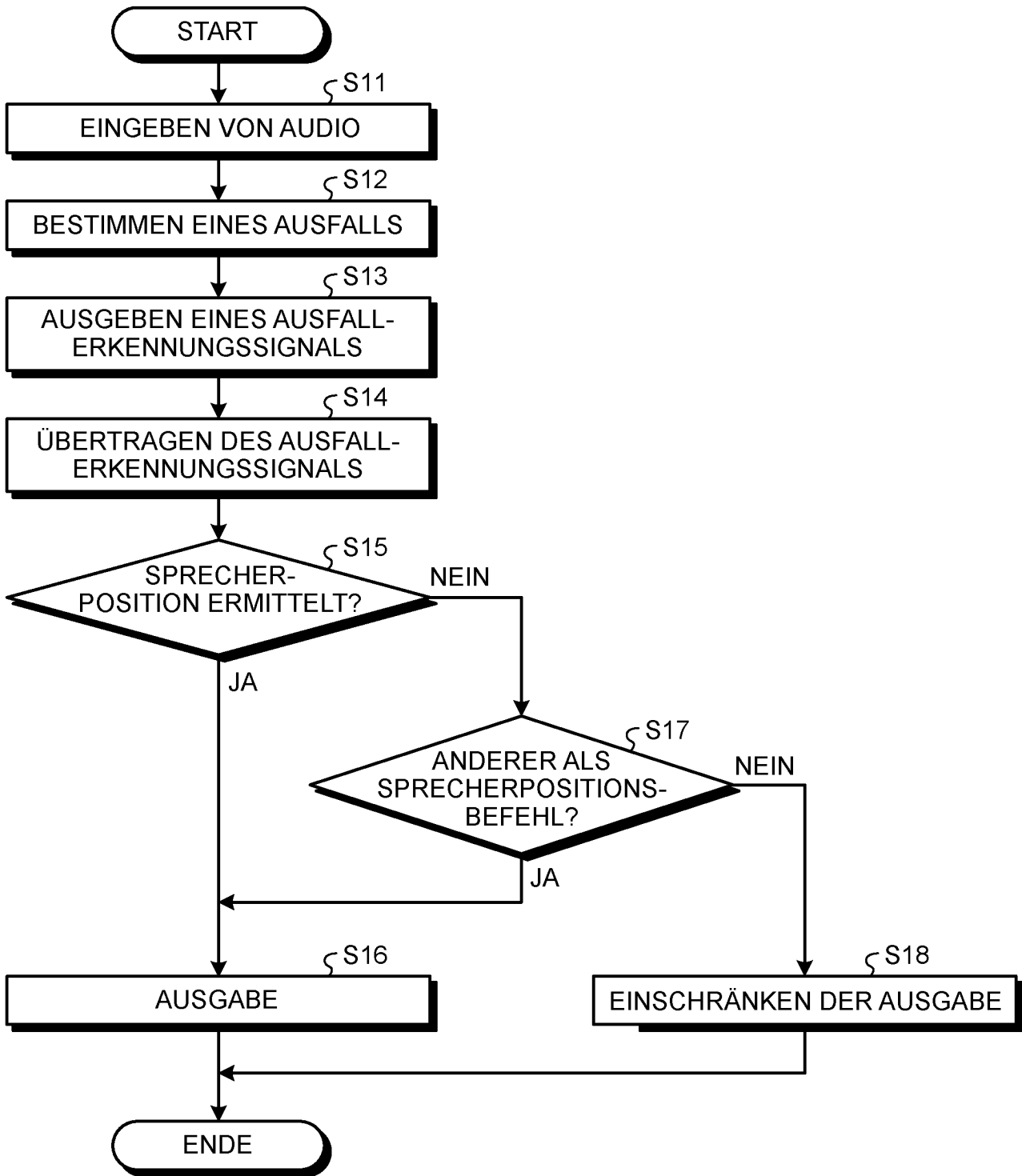


FIG.5

20a

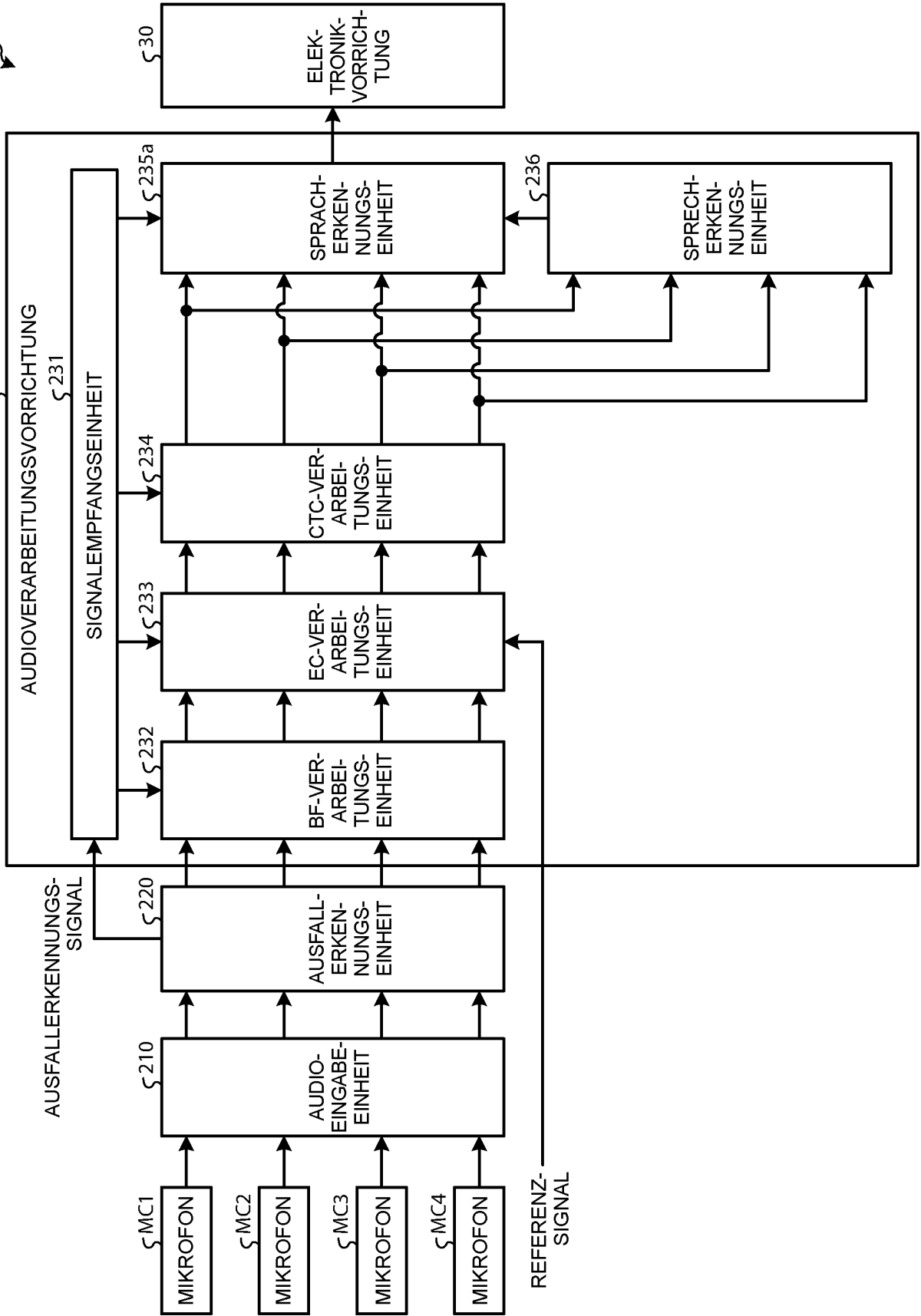


FIG.6

