



(10) **DE 10 2012 220 796 B4** 2015.08.27

Patentschrift

(51) Int Cl.: **H04M 3/56** (2006.01)
H04M 1/60 (2006.01)
H04M 1/26 (2006.01)
G10L 15/00 (2006.01)

Innerhalb von neun Monaten nach Veröffentlichung der Patenterteilung kann nach § 59 Patentgesetz gegen das Patent Einspruch erhoben werden. Der Einspruch ist schriftlich zu erklären und zu begründen. Innerhalb der Einspruchsfrist ist eine Einspruchsgebühr in Höhe von 200 Euro zu entrichten (§ 6 Patentkostengesetz in Verbindung mit der Anlage zu § 2 Abs. 1 Patentkostengesetz).

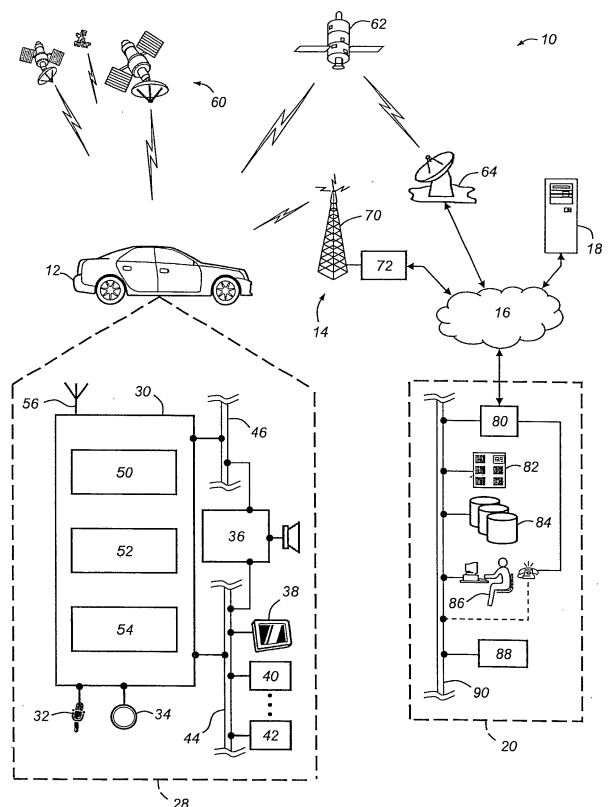
(72) Erfinder:
Peirce, Kenneth L., Grosse Pointe Farms, Mich.,
US; Arun, Uma, Novi, Mich., US; Madhavan, Sethu
K., Erie, Pa., US; Yi, Ki Hak, Windsor, Ontario, CA

(56) Ermittelter Stand der Technik:

DE	102 25 597	A1
US	2010 / 0 223 389	A1

(54) Bezeichnung: **Verfahren zum Initiieren eines Freisprechkonferenzgesprächs**

(e) der gespeicherte Konferenzgesprächzugangscode auf der Grundlage der Ermittlung der Spracherkennungsfunktion an den ATA gesendet wird.



Beschreibung**TECHNISCHES GEBIET**

[0001] Die vorliegende Erfindung bezieht sich allgemein auf Fahrzeuge und insbesondere auf Verfahren zum Betreiben einer Fahrzeugtelematikeinheit.

HINTERGRUND DER ERFINDUNG

[0002] Fahrzeughersteller statten ihre Fahrzeuge mit einer steigenden Anzahl an Fähigkeiten für drahtlose Kommunikationen aus. In modernen Fahrzeugen eingebaute Telematikeinheiten können drahtlos sowohl Sprach- als auch Datenübermittlungen zwischen dem Fahrzeug und einer Vielzahl von Empfängern, wie beispielsweise einer zentralen Entität oder einem Anrufer von außerhalb, übermitteln. Beispielsweise können Fahrzeugtelematikeinheiten Telefonanrufe an dem Fahrzeug sowohl tätigen als auch empfangen. Diese Anrufe können unter Verwendung von verbalen Befehlen, die durch einen Fahrzeuginsassen bereitgestellt werden, oder über physisch bediente Eingaben, die in dem Fahrzeug mitgeführt und durch den Fahrzeuginsassen betätigt werden, initiiert werden. Einige Anrufe, wie beispielsweise Konferenzgespräche, können zusätzlich zu dem Anrufer, der sich im Fahrzeug befindet, mehrere andere Anrufer umfassen. Die Teilnahme an Konferenzgesprächen kann das Bereitstellen einer Information für den Host dieser Anrufe voraussetzen, wie beispielsweise die Identität des Anrufs unter Verwendung eines Konferenzgesprächzugangscodes oder eines bestimmten Sicherheitscodes. Beispielsweise ist aus der Druckschrift DE 102 25 597 A1 ein Verfahren zum Betreiben einer Fahrzeugtelematikeinheit bekannt, das die Schritte umfasst, dass ein Konferenzgespräch von einem Fahrzeug auf der Grundlage von verbalen Befehlen initiiert wird, die von einem Fahrzeuginsassen empfangen werden, eine Anforderung hinsichtlich des Konferenzgesprächzugangscodes von einem automatisierten Telefonteilnehmer (ATA) empfangen wird, während das Konferenzgespräch initiiert wird, und der Konferenzgesprächzugangscodes an den ATA gesendet wird.

[0003] Ferner ist aus der US 2010/0223 389 A1 ein Verfahren bekannt, das die Schritte umfasst, dass ein Konferenzgespräch auf der Grundlage der Befehle initiiert wird, die von einem Nutzer empfangen werden, und eine Anforderung hinsichtlich des Konferenzgesprächzugangscodes von einem automatisierten Telefonteilnehmer (ATA) empfangen wird, während das Konferenzgespräch initiiert wird und der Konferenzgesprächzugangscodes an den ATA gesendet wird.

[0004] Für den Fahrzeuginsassen kann das Bereitstellen dieser Information eine Herausforderung sein.

ZUSAMMENFASSUNG DER ERFINDUNG

[0005] Gemäß einem Aspekt der Erfindung wird ein Verfahren zum Betreiben einer Fahrzeugtelematikeinheit vorgesehen. Die Schritte umfassen, dass ein Konferenzgesprächzugangscodes gespeichert wird; ein Konferenzgespräch von einem Fahrzeug auf der Grundlage von verbalen Befehlen, die von einem Fahrzeuginsassen empfangen werden, initiiert wird; eine Anforderung hinsichtlich des Konferenzgesprächzugangscodes von einem automatisierten Telefonteilnehmer (ATA von automated telephone attendant) empfangen wird, während das Konferenzgespräch initiiert wird; ermittelt wird, dass der ATA den Konferenzgesprächzugangscodes unter Verwendung einer Spracherkennungsfunktion an dem Fahrzeug angefordert hat; und der gespeicherte Konferenzgesprächzugangscodes auf der Grundlage der Ermittlung der Spracherkennungsfunktion an den ATA gesendet wird.

[0006] Gemäß einem weiteren Aspekt der Erfindung wird ein Verfahren zum Betreiben einer Fahrzeugtelematikeinheit bereitgestellt. Die Schritte umfassen, dass ein Konferenzgesprächzugangscodes von einem Telematikdienstbenutzer an dem Fahrzeug erhalten wird; der erhaltene Konferenzgesprächzugangscodes an dem Fahrzeug gespeichert wird; unter Verwendung der Spracherkennung an dem Fahrzeug automatisch ermittelt wird, dass ein automatisierter Telefonteilnehmer (ATA) einen Konferenzgesprächzugangscodes anfordert; DTMF-Töne oder DTMF-basierte Befehle auf der Grundlage des gespeicherten Konferenzgesprächzugangscodes erzeugt werden; und die DTMF-Töne oder DTMF-basierten Befehle drahtlos an den ATA übertragen werden, um einen Zugang zu dem Konferenzgespräch herzustellen.

[0007] Gemäß noch einem weiteren Aspekt der Erfindung wird ein Verfahren zum Betreiben einer Fahrzeugtelematikeinheit bereitgestellt. Die Schritte umfassen, dass eine Spracherkennungsfunktion an dem Fahrzeug optimiert wird, um eine Gruppe von Stimmen zu erkennen, die bei einem oder mehreren automatisierten Telefonteilnehmern (ATA) verwendet wird; ein Konferenzgesprächzugangscodes, ein Sicherheitscode oder beides von einem Telematikdienstbenutzer erhalten werden, um ein Konferenzgespräch an dem Fahrzeug herzustellen; der Konferenzgesprächzugangscodes, der Sicherheitscode oder beide an dem Fahrzeug gespeichert werden; unter Verwendung der optimierten Spracherkennungsfunktion automatisch ermittelt wird, dass ein ATA einen Konferenzgesprächzugangscodes anfordert; DTMF-Töne oder DTMF-Befehle auf der Grundlage des gespeicherten Konferenzgesprächzugangscodes, des Sicherheitscodes oder beider in Anspruch auf die Ermittlung erzeugt werden; und die DTMF-Töne oder DTMF-Befehle drahtlos an den ATA

übertragen werden, um einen Zugang zu dem Konferenzgespräch herzustellen.

KURZBESCHREIBUNG DER ZEICHNUNGEN

[0008] Hierin nachfolgend werden eine oder mehrere bevorzugte beispielhafte Ausführungsformen der Erfindung in Verbindung mit den beigefügten Zeichnungen beschrieben, wobei gleiche Bezugszeichen gleiche Elemente bezeichnen und wobei:

[0009] Fig. 1 ein Blockdiagramm ist, das eine beispielhafte Ausführungsform eines Kommunikationssystems zeigt, das das hierin offenbarte Verfahren verwenden kann;

[0010] Fig. 2 ein Blockdiagramm einer beispielhaften Architektur für ein ASR-System ist; und

[0011] Fig. 3 ein Flussdiagramm eines Verfahrens zum Betreiben einer Fahrzeugtelematikeinheit ist.

DETAILLIERTE BESCHREIBUNG DER GEZEIGTEN AUSFÜHRUNGSFORMEN

[0012] Das nachstehend beschriebene Verfahren interpretiert Sprachaufforderungen, die an einem Fahrzeug von einem automatisierten Telefonteilnehmer (ATA) empfangen werden und verwendet werden, um ein Konferenzgespräch aufzubauen, und antwortet automatisch auf jene Sprachaufforderungen, indem dem ATA ein oder mehrere Konferenzgesprächzugangscodes geliefert werden, die zuvor durch einen Telematikdienstbenutzer bereitgestellt wurden. Konferenzgespräche können mehrere Anrufer mit einem einzelnen Telefongespräch verbinden. Möglicherweise möchte eine anrufende Seite, dass mehr als eine angerufene Seite mithört und/oder an dem Anruf teilnimmt. Das Initiieren von Konferenzgesprächen kann umfassen, dass eine Telefonnummer angerufen wird, die speziell für Konferenzgespräche vorgesehen ist, und eine durch den ATA erzeugte Begrüßung gehört wird. Der ATA kann dann den Anrufer bitten, eine Information, wie beispielsweise einen Konferenzgesprächzugangscode, einen Sicherheitscode oder beides, zu liefern, die das bestimmte Konferenzgespräch identifiziert, an dem der Benutzer versucht teilzunehmen.

[0013] Gespräche wie diese können hergestellt werden, indem Zahlen empfangen werden, die von dem Telematikdienstbenutzer an dem Fahrzeug ausgesprochen werden. Wenn der Benutzer Zahlen ausspricht um eine Telefonnummer zu wählen, kann das Spracherkennungssystem des Fahrzeugs jene Zahlen erkennen und die Telefonnummer wählen. Bei einem Beispiel kann der Benutzer einen Knopf drücken, der sich in dem Fahrzeug befindet und das System informiert, dass der Benutzer einen Anruf tätigen möchte. Dann bereitet sich das System darauf vor, ei-

ne Telefonnummer zu empfangen, und kennt für gewöhnlich vorzeitig die Anzahl von Ziffern, die der Benutzer ausspricht. Beispielsweise können Fahrzeuge, die in einem Bereich betrieben werden, der durch den North American Numbering Plan (NANP) betreut wird, erwarten, dass sie eine zehn- oder elfstellige Nummer empfangen. Nach dem Empfangen jener Ziffern kann das System dann ermitteln, dass das Gespräch hergestellt wurde, und es hört dann damit auf, auf weitere gesprochene Zahlen zu hören. Der Benutzer möchte jedoch möglicherweise auf eine Anforderung von dem ATA unter Verwendung von zusätzlichen Ziffern, wie beispielsweise dem Konferenzgesprächzugangscode, antworten, der verwendet werden kann, um ein Konferenzgespräch aufzubauen. In diesem Fall müsste der Benutzer das System auf irgendeine Weise warnen, dass es bereit sein sollte, um weitere Ziffern von dem Benutzer zu empfangen, auch wenn dies für gewöhnlich nicht vorkommt. Oder der Benutzer muss möglicherweise dem System signalisieren, dass er das Tastenfeld verwenden möchte, um weitere Ziffern einzugeben. Gleichzeitig wartet der ATA möglicherweise auf eine Antwort von dem Benutzer. Und in Anbetracht der Zeitdauer, die der Benutzer möglicherweise benötigt, um das System vorzubereiten, um weitere gesprochene Ziffern zu empfangen, kann der ATA möglicherweise ermitteln, dass der Benutzer nicht antwortet, und das Gespräch beenden, während der Benutzer versucht zu antworten.

[0014] Stattdessen ist es möglich, den/die Konferenzgesprächzugangscode(s) und/oder Sicherheitscode(s) von dem Telematikdienstbenutzer zu erhalten, bevor ein Konferenzgespräch getätigt wird. Diese(r) Zugangscode(s)/Sicherheitscode(s) kann/können an dem Fahrzeug gespeichert sein, wie beispielsweise in einem speziellen Konferenzgesprächprofil. Wenn der Benutzer einen Anruf tätigen möchte, kann sich der Benutzer auf die Spracherkennungsfunktion des Fahrzeugs stützen, um nicht nur den Anruf zu tätigen, sondern auch um Informationsanforderungen zum Herstellen des Anrufs von dem ATA zu interpretieren, indem auf den/die Konferenzgesprächzugangscode(s)/Sicherheitscode(s) etc. zugegriffen wird und zumindest ein Teil dieser Information automatisch an den ATA gesendet wird. Hierbei kann die Spracherkennungsfunktion des Fahrzeugs angewiesen werden, auf sowohl den Benutzer in dem Fahrzeug als auch die eingehenden Anforderungen von dem ATA zu hören.

[0015] Bezug nehmend auf Fig. 1 ist eine beispielhafte Betriebsumgebung gezeigt, die ein Mobilfahrzeugkommunikationssystem **10** umfasst und die verwendet werden kann, um das hierin offenbarte Verfahren zu realisieren. Das Kommunikationssystem **10** umfasst allgemein ein Fahrzeug **12**, ein oder mehrere drahtlose Trägersysteme **14**, ein Bodenkommunikationsnetz **16**, einen Computer **18** und ein Call

Center **20**. Es ist zu verstehen, dass das offenbarte Verfahren mit jeder Anzahl von verschiedenen Systemen verwendet werden kann und nicht speziell auf die hier gezeigte Betriebsumgebung beschränkt ist. Auch sind die Architektur, die Konstruktion, der Aufbau und der Betrieb des Systems **10** sowie seine einzelnen Komponenten in der Technik allgemein bekannt. Somit liefern die folgenden Absätze lediglich einen kurzen Überblick über solch ein beispielhaftes System **10**; andere Systeme, die hier nicht gezeigt sind, könnten jedoch auch das offenbarte Verfahren einsetzen.

[0016] Das Fahrzeug **12** ist bei der dargestellten Ausführungsform als ein Personenkraftwagen gezeigt, es sei jedoch angemerkt, dass auch jedes andere Fahrzeug verwendet werden kann, das Motorräder, Lastwagen, Geländewagen (SUVs von sports utility vehicles), Wohnmobile (RVs von recreational vehicles), Schiffe, Luftfahrzeuge etc. umfasst. Ein Teil der Fahrzeugelektronik **28** ist in **Fig. 1** allgemein gezeigt und umfasst eine Telematikeinheit **30**, ein Mikrofon **32**, einen oder mehrere Druckknöpfe oder andere Steuereingabeeinrichtungen **34**, ein Audiosystem **36**, eine visuelle Anzeige **38** und ein GPS-Modul **40** sowie eine Anzahl von Fahrzeugsystemmodulen (VSMs von vehicle system modules) **42**. Einige dieser Einrichtungen können direkt mit der Telematikeinheit verbunden sein, wie beispielsweise das Mikrofon **32** und der Druckknopf/die Druckknöpfe **34**, wohingegen andere indirekt unter Verwendung einer oder mehrerer Netzverbindungen, wie beispielsweise eines Kommunikationsbusses **44** oder eines Unterhaltungsbusses **46**, verbunden sind. Beispiele geeigneter Netzverbindungen umfassen ein Controller Area Network (CAN), einen Media Oriented System Transfer (MOST), ein Local Interconnection Network (LIN), ein Local Area Network (LAN) und andere geeignete Verbindungen, wie beispielsweise Ethernet oder andere, die den bekannten ISO-, SAE- und IEEE-Standards und -Spezifikationen entsprechen, nur um einige zu nennen.

[0017] Die Telematikeinheit **30** kann eine OEM-Einrichtung (eingebaut) oder eine nachgerüstete Einrichtung sein, die eine drahtlose Sprach- und/oder Datenkommunikation über das drahtlose Trägersystem **14** und über einen drahtlosen Netzbetrieb ermöglicht, sodass das Fahrzeug mit dem Call Center **20**, anderen telematikfähigen Fahrzeugen oder einer anderen Entität oder Einrichtung kommunizieren kann. Die Telematikeinheit verwendet vorzugsweise Funkübertragungen, um einen Kommunikationskanal (einen Sprachkanal und/oder einen Datenkanal) mit dem drahtlosen Trägersystem **14** herzustellen, sodass Sprach- und/oder Datenübertragungen über den Kanal gesendet und empfangen werden können. Durch Bereitstellen von sowohl einer Sprach- als auch einer Datenkommunikation ermöglicht die Telematikeinheit **30** dem Fahrzeug, eine An-

zahl von verschiedenen Diensten anzubieten, die jene umfassen, die mit Navigation, Telefonie, Notfallunterstützung, Diagnose, Infotainment etc. in Beziehung stehen. Die Daten können entweder über eine Datenverbindung, wie beispielsweise über eine Paketdatenübertragung über einen Datenkanal, oder über einen Sprachkanal unter Verwendung von in der Technik bekannten Techniken gesendet werden. Für kombinierte Dienste, die sowohl eine Sprachkommunikation (z. B. mit einem menschlichen Berater oder einer Sprachausgabeeinheit an dem Call Center **20**) als auch eine Datenkommunikation (z. B. um GPS-Ortsdaten oder Fahrzeugdiagnosedaten für das Call Center **20** bereitzustellen) umfassen, kann das System einen einzelnen Anruf über einen Sprachkanal verwenden und nach Bedarf zwischen einer Sprach- und einer Datenübertragung über den Sprachkanal umschalten, und dies kann unter Verwendung von Fachleuten bekannten Techniken erfolgen.

[0018] Gemäß einer Ausführungsform verwendet die Telematikeinheit **30** eine zellulare Kommunikation gemäß entweder GSM- oder CDMA-Standards und umfasst somit einen standardisierten zellularen Chipsatz **50** für Sprachübermittlungen wie Freisprechanrufe, ein Drahtlosmodem für eine Datenübertragung, eine elektronische Verarbeitungseinrichtung **52**, eine oder mehrere digitale Speichereinrichtungen **54** und eine Dualantenne **56**. Es sei angemerkt, dass das Modem entweder durch eine Software realisiert sein kann, die in der Telematikeinheit gespeichert ist und durch den Prozessor **52** ausgeführt wird, oder dass es eine separate Hardwarekomponente sein kann, die sich in der Telematikeinheit **30** oder außerhalb dieser befindet. Das Modem kann unter Verwendung jeder Anzahl von verschiedenen Standards oder Protokollen arbeiten, wie beispielsweise EVDO, CDMA, GPRS und EDGE. Ein drahtloser Netzbetrieb zwischen dem Fahrzeug und anderen vernetzten Einrichtungen kann auch unter Verwendung der Telematikeinheit **30** ausgeführt werden. Zu diesem Zweck kann die Telematikeinheit **30** ausgestaltet sein, um gemäß einem oder mehreren drahtlosen Protokollen, wie beispielsweise einem beliebigen der IEEE 802.11-Protokolle, WiMAX oder Bluetooth, drahtlos zu kommunizieren. Bei einer Verwendung für eine paketvermittelte Datenübertragung, wie beispielsweise TCP/IP, kann die Telematikeinheit mit einer statischen IP-Adresse konfiguriert sein oder kann sie aufgebaut sein, um automatisch eine zugeordnete IP-Adresse von einer anderen Einrichtung an dem Netzwerk, wie beispielsweise einem Router, oder von einem Netzadressenserver zu empfangen.

[0019] Der Prozessor **52** kann jeder Typ von Einrichtung sein, der elektronische Anweisungen verarbeiten kann, und kann Mikroprozessoren, Mikrocontroller, Host-Prozessoren, Controller, Fahrzeugkommunikationsprozessoren und anwendungsspezifische integrierte Schaltkreise (ASICs von applicati-

on specific integrated circuits) umfassen. Er kann ein dedizierter Prozessor sein, der nur für die Telematikeinheit **30** verwendet wird, oder er kann von anderen Fahrzeugsystemen gemeinsam genutzt werden. Der Prozessor **52** führt verschiedene Typen von digital gespeicherten Anweisungen aus, wie beispielsweise Software- oder Firmware-Programme, die in dem Speicher **54** gespeichert sind und der Telematikeinheit ermöglichen, eine große Vielzahl von Diensten bereitzustellen. Beispielsweise kann der Prozessor **52** Programme ausführen oder Daten verarbeiten, um mindestens einen Teil des hierin erläuterten Verfahrens auszuführen.

[0020] Die Telematikeinheit **30** kann verwendet werden, um einen vielseitigen Bereich von Fahrzeugdiensten bereitzustellen, die eine drahtlose Übermittlung zu und/oder von dem Fahrzeug umfassen. Solche Dienste umfassen: Turn-by-Turn-Anweisungen und andere navigationsbezogene Dienste, die in Verbindung mit dem GPS-basierten Fahrzeugnavigationsmodul **40** bereitgestellt werden; eine Airbag-Einsatzbenachrichtigung und andere Notfall- oder Pannehilfedienste, die in Verbindung mit einem oder mehreren Kollisionssensorschnittstellenmodulen bereitgestellt werden, wie beispielsweise einem Karosseriesteuermodul (nicht gezeigt); eine Diagnoseberichterstattung unter Verwendung eines oder mehrerer Diagnosemodule; und Infotainment-bezogene Dienste, bei denen Musik, Webseiten, Filme, Fernsehprogramme, Videospiele und/oder andere Informationen durch ein Infotainment-Modul (nicht gezeigt) heruntergeladen werden und für eine sofortige oder spätere Wiedergabe gespeichert werden. Die oben aufgelisteten Dienste sind keineswegs eine vollständige Liste aller Fähigkeiten der Telematikeinheit **30**, sondern sind lediglich eine Aufzählung einiger der Dienste, die die Telematikeinheit anbieten kann. Ferner sei angemerkt, dass mindestens einige der zuvor genannten Module in Form von Softwareanweisungen realisiert sein könnten, die innerhalb oder außerhalb der Telematikeinheit **30** gespeichert sind, dass sie Hardwarekomponenten sein könnten, die sich innerhalb oder außerhalb der Telematikeinheit **30** befinden, oder dass sie miteinander oder mit anderen Systemen, die sich in dem Fahrzeug befinden, integriert sein könnten und/oder von diesen gemeinsam genutzt werden könnten, nur um einige Möglichkeiten zu nennen. In dem Fall, dass die Module als VSMs **42** realisiert sind, die außerhalb der Telematikeinheit **30** angeordnet sind, könnten sie den Fahrzeugbus **44** verwenden, um Daten und Befehle mit der Telematikeinheit auszutauschen.

[0021] Das GPS-Modul **40** empfängt Funksignale von einer Konstellation **60** von GPS-Satelliten. Aus diesen Signalen kann das Modul **40** die Fahrzeugposition ermitteln, die verwendet wird, um dem Fahrer des Fahrzeugs Navigations- und andere positionsbezogene Dienste bereitzustellen. Eine Navigationsin-

formation kann an der Anzeige **38** (oder an einer anderen Anzeige innerhalb des Fahrzeugs) dargestellt werden oder kann verbal dargestellt werden, wie es bei einem Bereitstellen einer Turn-by-Turn-Navigation der Fall ist. Die Navigationsdienste können unter Verwendung eines dedizierten fahrzeuginternen Navigationsmoduls (das Teil des GPS-Moduls **40** sein kann) bereitgestellt werden, oder es können einige oder alle Navigationsdienste über die Telematikeinheit **30** ausgeführt werden, wobei die Positionsinformation zu Zwecken des Bereitstellens von Navigationskarten, Kartenanmerkungen (Punkte von Interesse, Restaurants etc.), Routenberechnungen und dergleichen für das Fahrzeug an einen entfernten Ort gesendet wird. Die Positionsinformation kann dem Call Center **20** oder einem anderen entfernten Computersystem, wie beispielsweise einem Computer **18**, zu anderen Zwecken, wie beispielsweise einer Flottenverwaltung, bereitgestellt werden. Es können auch neue oder aktualisierte Kartendaten von dem Call Center **20** über die Telematikeinheit **30** auf das GPS-Modul **40** heruntergeladen werden.

[0022] Abgesehen von dem Audiosystem **36** und dem GPS-Modul **40** kann das Fahrzeug **12** andere Fahrzeugsystemmodule (VSMs) **42** in Form von elektronischen Hardwarekomponenten umfassen, die an dem Fahrzeug angeordnet sind und typischerweise einen Eingang von einem oder mehreren Sensoren empfangen und den erfassten Eingang verwenden, um Diagnose-, Überwachungs-, Steuerungs-, Berichterstattungs- und/oder andere Funktionen durchzuführen. Jedes der VSMs **42** ist vorzugsweise durch den Kommunikationsbus **44** mit den anderen VSMs sowie mit der Telematikeinheit **30** verbunden und kann programmiert sein, um Fahrzeugsystem- und -teilsystemdiagnosetests auszuführen. Beispielsweise kann ein VSM **42** ein Motorsteuermodul (ECM von engine control module) sein, das verschiedene Aspekte des Motorbetriebs steuert, wie beispielsweise Kraftstoffzündung und Zündzeitpunkt, kann ein anderes VSM **42** ein Antriebsstrangsteuermodul sein, das den Betrieb einer oder mehrerer Komponenten des Fahrzeugantriebsstrangs reguliert, und kann ein anderes VSM **42** ein Karosseriesteuermodul sein, das verschiedene elektrische Komponenten überwacht, die sich an dem Fahrzeug befinden, wie beispielsweise die Zentralverriegelung und die Scheinwerfer des Fahrzeugs. Gemäß einer Ausführungsform ist das Motorsteuermodul mit fahrzeugeigenen Diagnosemerkmalen (OBD-Merkmalen von on-board diagnostic features) ausgestattet, die eine Vielzahl von Echtzeitdaten bereitstellen, wie beispielsweise die, die von verschiedenen Sensoren einschließlich Fahrzeugemissionssensoren empfangen werden und eine standardisierte Reihe von Diagnosefehlercodes (DTCs von diagnostic trouble codes) bereitstellen, die einem Ingenieur ermöglichen, Fehlfunktionen in dem Fahrzeug schnell zu identifizieren und zu beheben. Fachleute werden erkennen, dass die oben erwähn-

ten VSMs nur Beispiele einiger der Module sind, die in dem Fahrzeug **12** verwendet werden können, da auch zahlreiche andere möglich sind.

[0023] Die Fahrzeugelektronik **28** umfasst auch eine Anzahl von Fahrzeugbenutzerschnittstellen, die Fahrzeuginsassen ein Mittel zum Bereitstellen und/oder Empfangen einer Information bereitstellen und das Mikrofon **32**, den Druckknopf/die Druckknöpfe **34**, das Audiosystem **36** und die visuelle Anzeige **38** umfassen. Wie hierin verwendet umfasst der Begriff "Fahrzeugbenutzerschnittstelle" breit jede geeignete Form von elektronischer Einrichtung, die sowohl Hardware- als auch Softwarekomponenten umfasst und sich an dem Fahrzeug befindet und einem Fahrzeugbenutzer ermöglicht, mit einer oder über eine Komponente des Fahrzeugs zu kommunizieren. Das Mikrofon **32** stellt einen Audioeingang für die Telematikeinheit bereit, um dem Fahrer oder einem anderen Insassen zu ermöglichen, Sprachbefehle bereitzustellen und über das drahtlose Trägersystem **14** Freisprechanrufe auszuführen. Zu diesem Zweck kann es mit einer fahrzeugeigenen automatisierten Sprachverarbeitungseinheit verbunden sein, die eine in der Technik bekannte Mensch-Maschine-Schnittstellentechnologie (HMI-Technologie von human-machine interface technology) verwendet. Der Druckknopf/die Druckknöpfe **34** ermöglicht/ermöglichen eine manuelle Benutzereingabe in die Telematikeinheit **30**, um drahtlose Telefonanrufe zu initiieren und andere Daten, eine Antwort oder einen Steuerungseingang bereitzustellen. Es können separate Druckknöpfe verwendet werden, um im Gegensatz zu regulären Dienstunterstützungsanrufen an das Call Center **20** Notrufe zu initiieren. Das Audiosystem **36** stellt einen Audioausgang für einen Fahrzeuginsassen bereit und kann ein dediziertes, unabhängiges System oder ein Teil des primären Fahrzeugaudiosystems sein. Gemäß der bestimmten hier gezeigten Ausführungsform ist das Audiosystem **36** funktional mit sowohl dem Fahrzeugbus **44** als auch dem Unterhaltungsbus **46** gekoppelt und kann es eine AM-, FM- und Satellitenradio-, CD-, DVD- und eine andere Multimedialfunktionalität bereitstellen. Diese Funktionalität kann in Verbindung mit oder unabhängig von dem oben beschriebenen Infotainment-Modul bereitgestellt werden. Die visuelle Anzeige **38** ist vorzugsweise eine Graphikanzeige, wie beispielsweise ein berührungsempfindlicher Bildschirm an dem Armaturenbrett, oder eine Head-Up-Anzeige, die an der Windschutzscheibe reflektiert wird, und kann verwendet werden, um eine Vielzahl von Eingabe- und Ausgabefunktionen bereitzustellen. Es können auch verschiedene andere Fahrzeugbenutzerschnittstellen verwendet werden, da die Schnittstellen von **Fig. 1** nur ein Beispiel einer bestimmten Realisierung sind.

[0024] Das drahtlose Trägersystem **14** ist vorzugsweise ein Mobiltelefonsystem, das mehrere Mobil-

funktürme **70** (nur einer gezeigt), eine oder mehrere Mobilfunkvermittlungsstellen (MSCs von mobile switching centers) **72** sowie beliebige andere Netzkomponenten umfasst, die erforderlich sind, um das drahtlose Trägersystem **14** mit dem Bodennetz **16** zu verbinden. Jeder Mobilfunkurm **70** umfasst sendende und empfangende Antennen und eine Basisstation, wobei die Basisstationen von unterschiedlichen Mobilfunktürmen entweder direkt oder über ein Zwischengerät, wie beispielsweise einen Basisstations-Controller, mit der MSC **72** verbunden sind. Das Zellularsystem **14** kann jede geeignete Kommunikationstechnologie realisieren, die beispielsweise analoge Technologien, wie beispielsweise AMPS, oder die neueren digitalen Technologien, wie beispielsweise CDMA (z. B. CDMA2000) oder GSM/GPRS, umfasst. Fachleute werden erkennen, dass verschiedene Mobilfunkurm/Basisstation/MSC-Anordnungen möglich sind und mit dem drahtlosen System **14** verwendet werden könnten. Beispielsweise könnten die Basisstation und der Mobilfunkurm zusammen an dem gleichen Ort angeordnet sein, oder sie könnten entfernt voneinander angeordnet sein, könnte jede Basisstation für einen einzelnen Mobilfunkurm verantwortlich sein oder könnte eine einzelne Basisstation verschiedene Mobilfunktürme bedienen und könnten verschiedene Basisstationen mit einer einzelnen MSC gekoppelt sein, nur um einige der möglichen Anordnungen zu nennen.

[0025] Abgesehen von einem Verwenden des drahtlosen Trägersystems **14** kann ein anderes drahtloses Trägersystem in Form einer Satellitenkommunikation verwendet werden, um eine unidirektionale oder bidirektionale Kommunikation mit dem Fahrzeug bereitzustellen. Dies kann unter Verwendung eines oder mehrerer Kommunikationssatelliten **62** und einer Uplink-Übertragungsstation **64** erfolgen. Eine unidirektionale Kommunikation kann beispielsweise Satellitenfunkdienste umfassen, bei denen Programminhalt (Nachrichten, Musik, etc.) durch die Übertragungsstation **64** empfangen wird, für ein Hochladen verpackt wird und dann an den Satelliten **62** gesendet wird, der die Programme an Teilnehmer ausstrahlt. Eine bidirektionale Kommunikation kann beispielsweise Satellitentelefonien umfassen, die den Satelliten **62** verwenden, um Telefonverkehr zwischen dem Fahrzeug **12** und der Station **64** weiterzuleiten. Bei einer Verwendung kann diese Satellitentelefonie entweder zusätzlich zu dem drahtlosen Trägersystem **14** oder anstatt diesem verwendet werden.

[0026] Das Bodennetz **16** kann ein herkömmliches bodenbasiertes Telekommunikationsnetz sein, das mit einem oder mehreren Festnetztelefonen verbunden ist und das drahtlose Trägersystem **14** mit dem Call Center **20** verbindet. Beispielsweise kann das Bodennetz **16** ein Fernsprechnetz (PSTN von public switched telephone network) umfassen, wie

beispielsweise jenes, das verwendet wird, um eine Festnetztelefonie, paketvermittelte Datenkommunikationen und die Internet-Infrastruktur bereitzustellen. Ein oder mehrere Segmente des Bodennetzes **16** könnten durch die Verwendung eines standardisierten drahtgebundenen Netzes, eines Faser- oder anderen optischen Netzes, eines Kabelnetzes, von Hochspannungsleitungen, anderen drahtlosen Netzen, wie beispielsweise Wireless Local Area Networks (WLANs), oder Netzen, die einen drahtlosen Breitbandzugriff (BWA von broadband wireless access) bereitstellen, oder jeder Kombination hiervon realisiert sein. Ferner muss das Call Center **20** nicht über das Bodennetz **16** verbunden sein, sondern könnte es ein Drahtlostelefoniegerät umfassen, sodass es direkt mit einem drahtlosen Netz, wie beispielsweise dem drahtlosen Trägersystem **14**, kommunizieren kann.

[0027] Der Computer **18** kann einer einer Anzahl von Computern sein, auf die über ein privates oder öffentliches Netz, wie beispielsweise das Internet, zugegriffen werden kann. Jeder solche Computer **18** kann für einen oder mehrere Zwecke, wie beispielsweise einen Web-Server, verwendet werden, auf den durch das Fahrzeug über die Telematikeinheit **30** und den drahtlosen Träger **14** zugegriffen werden kann. Andere solche Computer **18**, auf die zugegriffen werden kann, können beispielsweise umfassen: einen Computer einer Dienstzentrale, an dem Diagnoseinformationen und andere Fahrzeugdaten von dem Fahrzeug über die Telematikeinheit **30** hochgeladen werden können; einen Client-Computer, der durch den Fahrzeughalter oder einen anderen Teilnehmer zu Zwecken wie beispielsweise Zugreifen auf oder Empfangen von Fahrzeugdaten oder Einstellen oder Konfigurieren von Teilnehmervorlieben oder Steuern von Fahrzeugfunktionen verwendet wird; oder einen dritten Speicher, für den oder von dem Fahrzeugdaten oder andere Informationen geliefert werden, entweder durch Kommunizieren mit dem Fahrzeug **12** oder dem Call Center **20** oder beiden. Ein Computer **18** kann auch zum Bereitstellen einer Internet-Konnektivität, wie beispielsweise von DNS-Diensten, oder als ein Netzadressenserver, der DHCP oder ein anderes geeignetes Protokoll verwendet, um dem Fahrzeug **12** eine IP-Adresse zuzuordnen, verwendet werden.

[0028] Das Call Center **20** ist entworfen, um der Fahrzeugelektronik **28** eine Anzahl von verschiedenen System-Backend-Funktionen bereitzustellen und umfasst gemäß der hier gezeigten, beispielhaften Ausführungsform allgemein eine(n) oder mehrere Schalter **80**, Server **82**, Datenbanken **84**, menschliche Berater **86** sowie ein automatisiertes Sprachausgabesystem (VRS von voice response system) **88**, die alle in der Technik bekannt sind. Diese verschiedenen Call Center-Komponenten sind vorzugsweise über ein drahtgebundenes oder drahtloses lokales Netz **90** miteinander gekoppelt. Der Schalter **80**,

der ein Telekommunikationsanlagenschalter (PBX-Schalter von private branch exchange switch) sein kann, leitet eingehende Signale derart weiter, dass Sprachübertragungen für gewöhnlich entweder durch ein normales Telefon an den menschlichen Berater **86** oder unter Verwendung von VoIP an das automatisierte Sprachausgabesystem **88** gesendet werden. Das Telefon des menschlichen Beraters kann auch VoIP verwenden, wie es durch die gestrichelte Linie in **Fig. 1** gezeigt ist. VoIP und andere Datenkommunikationen über den Schalter **80** werden über ein Modem (nicht gezeigt) realisiert, das zwischen dem Schalter **80** und dem Netz **90** verbunden ist. Die Datenübertragungen werden über das Modem an den Server **82** und/oder die Datenbank **84** weitergeleitet. Die Datenbank **84** kann eine Kontoinformation, wie beispielsweise eine Teilnehmerauthentifizierungsinformation, Fahrzeugidentifikatoren, Profilaufzeichnungen, Verhaltensmuster und andere entsprechende Teilnehmerinformationen, speichern. Datenübertragungen können auch durch drahtlose Systeme, wie beispielsweise 802.11x, GPRS und dergleichen, ausgeführt werden. Obwohl die gezeigte Ausführungsform als in Verbindung mit einem mit Personal besetzten Call Center **20** unter Verwendung des menschlichen Beraters **86** verwendet beschrieben wurde, sei angemerkt, dass das Call Center stattdessen das VRS **88** als einen automatisierten Berater verwenden kann oder eine Kombination aus dem VRS **88** und dem menschlichen Berater **86** verwendet werden kann.

[0029] Nun auf **Fig. 2** Bezug nehmend ist eine beispielhafte Architektur für ein ASR-System **210** gezeigt, das verwendet werden kann, um das vorliegend offenbarte Verfahren zu ermöglichen. Im Allgemeinen tritt ein Fahrzeuginsasse zu einem oder mehreren der folgenden grundlegenden Zwecke sprachlich mit einem automatischen Spracherkennungssystem (ASR) in Interaktion: Trainieren des Systems, um die bestimmte Stimme eines Fahrzeuginsassen zu verstehen; Speichern einer diskreten Sprache, wie beispielsweise eines gesprochenen Nametags oder eines gesprochenen Steuerworts, wie einer Zahl oder eines Schlüsselworts; oder Erkennen der Sprache des Fahrzeuginsassen zu einem beliebigen geeigneten Zweck, wie beispielsweise Sprachwahl, Menünavigation, Transkription, Dienstanforderungen, Fahrzeugeinrichtungs- oder -einrichtungsfunktionssteuerung oder dergleichen. Im Allgemeinen extrahiert die ASR akustische Daten aus menschlicher Sprache, vergleicht die akustischen Daten mit gespeicherten Subwortdaten und stellt diese gegenüber, wählt ein geeignetes Subwort aus, das mit anderen ausgewählten Subworten verknüpft werden kann, und gibt die verknüpften Subworte oder Worte zur Nachverarbeitung, wie beispielsweise Diktat oder Transkription, zur Adressbuchwahl, zum Speichern im Speicher, zum Trainieren von ASR-Modellen oder Abstimmungsparametern oder dergleichen, aus.

[0030] ASR-Systeme sind Fachleuten im Allgemeinen bekannt, und **Fig. 2** zeigt lediglich ein spezifisches beispielhaftes ASR-System **210**. Das System **210** umfasst eine Einrichtung zum Empfangen von Sprache, wie beispielsweise das Telematikmikrophon **32**, und eine Akustikschnittstelle **33**, wie beispielsweise eine Soundkarte der Telematikeinheit **30** mit einem Analog-Digital-Wandler, um die Sprache in akustische Daten zu digitalisieren. Das System **210** umfasst auch einen Speicher, wie beispielsweise den Telematikspeicher **54**, zum Speichern der akustischen Daten und Speichern von Spracherkennungssoftware und Datenbanken, und einen Prozessor, wie beispielsweise die elektronische Verarbeitungseinrichtung **52**, zum Verarbeiten der akustischen Daten. Der Prozessor arbeitet mit dem Speicher und in Verbindung mit den folgenden Modulen: einem oder mehreren Front-End-Prozessor(en) oder Präprozessor-Softwaremodul(en) **212** zum Parsen von Strömen der akustischen Daten der Sprache in parametrische Darstellungen, wie beispielsweise Akustikmerkmale; ein oder mehrere Decodierersoftwaremodul(e) **214** zum Decodieren der Akustikmerkmale, um ein digitales Subwort oder Wortausgangsdaten zu erhalten, die den eingegebenen Sprachäußerungen entsprechen; und ein oder mehrere Postprozessor-Softwaremodul(e) **216** zum Verwenden der Ausgangsdaten von dem Decodierersoftwaremodul/den Decodierersoftwaremodulen **214** für einen beliebigen geeigneten Zweck.

[0031] Das System **210** kann auch Sprache von einer beliebigen anderen geeigneten Audioquelle/beliebigen anderen geeigneten Audioquellen **31** empfangen, die direkt mit dem Präprozessorsoftwaremodul/den Präprozessorsoftwaremodulen **212** übermittelt werden kann, wie es als durchgehende Linie gezeigt ist, oder indirekt damit über die Akustikschnittstelle **33** übermittelt werden kann. Die Audioquelle(n) **31** kann/können beispielsweise eine telefonische Audioquelle, wie beispielsweise ein Sprachmitteilungssystem, oder andere telefonische Dienste jeder Art umfassen.

[0032] Es können ein oder mehrere Module oder Modelle als Eingang in das Decodierersoftwaremodul/die Decodierersoftwaremodule **214** verwendet werden. Erstens kann ein Grammatik- und/oder Lexikonmodell/können Grammatik- und/oder Lexikonmodelle **218** Regeln bereitstellen, die überwachen, welche Wörter anderen Wörtern logisch folgen können, um gültige Sätze zu bilden. In einem breiten Sinne kann eine Grammatik eine Gesamtheit von Vokabular definieren, das das System **210** zu einem beliebigen gegebenen Zeitpunkt in einem beliebigen gegebenen ASR-Modus erwartet. Wenn sich das System **210** beispielsweise in einem Trainingsmodus zum Trainieren von Befehlen befindet, kann das Grammatikmodell/können die Grammatikmodelle **218** alle Befehle umfassen, die dem System **210** bekannt sind und durch dieses verwendet werden. Bei ei-

nem anderen Beispiel kann das aktive Grammatikmodell/können die aktiven Grammatikmodelle **218**, wenn sich das System **210** in einem Hauptmenümodus befindet, alle Hauptmenübefehle umfassen, die von dem System **210** erwartet werden, wie beispielsweise Anruf, Wahl, Verlassen, Löschen, Verzeichnis oder dergleichen. Zweitens unterstützt ein Akustikmodell/unterstützen Akustikmodelle **220** die Wahl der wahrscheinlichsten Subworte oder Worte, die einer Eingabe von dem Präprozessormodul/den Präprozessormodulen **212** entsprechen. Drittens stellt/stellen ein Wortmodell/Wortmodelle **222** und (ein) Satz-/Sprachmodell(e) **224** Regeln, Syntax und/oder Semantik beim Anordnen der ausgewählten Subworte oder Worte in einem Wort- oder Satzkontext bereit. Das/die Satz-/Sprachmodell(e) **224** kann/können eine Gesamtheit von Sätzen definieren, die das System **210** zu einem beliebigen gegebenen Zeitpunkt in einem beliebigen gegebenen ASR-Modus erwartet, und/oder kann/können Regeln bereitstellen etc., die überwachen, welche Sätze anderen Sätzen logisch folgen können, um eine gültige erweiterte Sprache zu bilden. Bei jedem dieser Beispiele kann das relevante Modell trainiert werden, um die Stimmen zu lernen, die von einem oder mehreren ATAs erzeugt werden. Beispielsweise variiert weder die Identität der Stimmen, die durch die ATAs verwendet werden, noch die Anzahl von Befehlen stark. Folglich können die Grammatik- oder Lexikonmodelle **218** schnell und leicht vorbereitet werden, um eine einer begrenzten Anzahl von bekannten und identifizierten Stimmen, die durch ATAs verwendet werden, und auch die wahrscheinlichen Worte und/oder Befehle, die ATAs häufig verwenden, zu identifizieren. Beispielsweise könnte ein ATA einen Anrufer bitten: "Geben Sie bitte Ihren Konferenzgesprächzugangscode ein". Das ASR-System **210** kann trainiert werden, um nach jedem/jeder dieser Worte/Phrasen und/oder der Stimme, die sie wahrscheinlich spricht, zu suchen.

[0033] Gemäß einer alternativen beispielhaften Ausführungsform kann sich ein Teil des oder das gesamte ASR-System **210** in einem Rechenggerät an einem Ort, der von dem Fahrzeug **12** entfernt ist, wie beispielsweise dem Call Center **20**, befinden und unter dessen Verwendung verarbeitet werden. Beispielsweise können Grammatikmodelle, Akustikmodelle und dergleichen in dem Speicher eines der Server **82** und/oder einer der Datenbanken **84** in dem Call Center **20** gespeichert und zur fahrzeuginternen Sprachverarbeitung an die Fahrzeugtelematikeinheit **30** übermittelt werden. Ähnlich kann eine Spracherkennungssoftware unter Verwendung von Prozessoren eines der Server **82** in dem Call Center **20** verarbeitet werden. Mit anderen Worten kann sich das ASR-System **210** in der Telematikeinheit **30** befinden oder auf eine beliebige gewünschte Weise in dem Call Center **20** und dem Fahrzeug **12** verteilt sein.

[0034] Zuerst werden akustische Daten von menschlicher Sprache extrahiert, wobei ein Fahrzeuginsasse in das Mikrophon **32** spricht, das die Äußerungen in elektrische Signale umwandelt und solche Signale an die Akustikschnittstelle **33** übermittelt. Ein auf Töne ansprechendes Element in dem Mikrophon **32** erfasst die Sprachäußerungen eines Insassen als Schwankungen des Luftdrucks und wandelt die Äußerungen in entsprechende Schwankungen analoger elektrischer Signale, wie beispielsweise Gleichstrom oder -spannung, um. Die Akustikschnittstelle **33** empfängt die analogen elektrischen Signale, die zuerst derart abgetastet werden, dass Werte des analogen Signals zu diskreten Zeitpunkten erfasst werden, und dann quantisiert werden, so dass die Amplituden der analogen Signale zu jedem Abtastzeitpunkt in einen kontinuierlichen Strom von digitalen Sprachdaten umgewandelt werden. Mit anderen Worten wandelt die Akustikschnittstelle **33** die analogen elektrischen Signale in digitale elektronische Signale um. Die digitalen Daten sind binäre Bits, die in dem Telematikspeicher **54** gepuffert werden und dann durch den Telematikprozessor **52** verarbeitet werden oder in Echtzeit verarbeitet werden können, wenn sie anfänglich durch den Prozessor **52** empfangen werden.

[0035] Zweitens transformiert/transformieren das Präprozessormodul/die Präprozessormodule **212** den kontinuierlichen Strom von digitalen Sprachdaten in diskrete Sequenzen von Akustikparametern. Im Spezielleren lässt der Prozessor **52** das Präprozessormodul/die Präprozessormodule **212** arbeiten, um die digitalen Sprachdaten in sich überschneidende phonetische oder akustische Frames mit einer Dauer von beispielsweise 10–30 ms zu segmentieren. Die Frames entsprechen akustischen Subworten, wie beispielsweise Silben, Halbsilben, Phonemen, Diphonemen oder dergleichen. Das Präprozessormodul/die Präprozessormodule **212** führt/führen auch eine phonetische Analyse durch, um aus jedem Frame Akustikparameter von der Sprache des Insassen, wie beispielsweise zeitvariable Merkmalsvektoren, zu extrahieren. Äußerungen innerhalb der Sprache des Insassen können als Sequenzen dieser Merkmalsvektoren dargestellt werden. Beispielsweise und wie es Fachleuten bekannt ist, können Merkmalsvektoren extrahiert werden und können diese beispielsweise eine vokale Tonhöhe, Energieprofile, Spektralattribute und/oder Cepstral-Koeffizienten umfassen, die durch Durchführen von Fourier-Transformationen der Frames und Dekorrelieren von akustischen Spektren unter Verwendung von Kosinus-Transformationen erhalten werden können. Akustische Frames und entsprechende Parameter, die eine bestimmte Sprachdauer abdecken, werden zu unbekannten zu decodierenden Sprachtestmustern verknüpft.

[0036] Drittens lässt der Prozessor das Decodierermodule/die Decodierermodule **214** arbeiten, um

die eingehenden Merkmalsvektoren jedes Testmusters zu verarbeiten. Das Decodierermodule/die Decodierermodule **214** ist/sind auch als Erkennungsmaschine oder Klassifizierer bekannt und verwendet/verwenden gespeicherte bekannte Sprachreferenzmuster. Wie die Testmuster sind die Referenzmuster als Verknüpfung von in Beziehung stehenden akustischen Frames mit entsprechenden Parametern definiert. Das Decodierermodule/die Decodierermodule **214** vergleicht/vergleichen die Akustikmerkmalsvektoren eines zu erkennenden Subworttestmusters mit gespeicherten Subwortreferenzmustern und stellt/stellen diese gegenüber, bewertet/bewerten den Umfang der Unterschiede oder Ähnlichkeiten zwischen ihnen und verwendet/verwenden schließlich eine Entscheidungslogik, um ein am besten übereinstimmendes Subwort als das erkannte Subwort auszuwählen. Im Allgemeinen ist das am besten übereinstimmende Subwort jenes, welches dem gespeicherten bekannten Referenzmuster entspricht, das eine minimale Unähnlichkeit zu dem Testmuster aufweist oder die größte Wahrscheinlichkeit hat, dieses zu sein, wie es durch eine beliebige von verschiedenen Fachleuten bekannten Techniken zum Analysieren und Erkennen von Subworten ermittelt wird. Solche Techniken können Dynamic Time Warping-Klassifizierer, Techniken künstlicher Intelligenz, neuronale Netze, Einrichtungen für eine freie Phonemerkennung und/oder Einrichtungen für eine wahrscheinlichkeitstheoretische Inübereinstimmungbringung von Mustern, wie beispielsweise Hidden Markov Modell-Maschinen (HMM-Maschinen), umfassen.

[0037] HMM-Maschinen sind Fachleuten bezüglich eines Erzeugens von Mehrfachspracherkennungsmodeilhypothesen einer akustischen Eingabe bekannt. Die Hypothesen werden beim schließlichen Identifizieren und Auswählen jenes Erkennungsausgangs, der die Decodierung der Akustikeingabe darstellt, die am wahrscheinlichsten korrekt ist, über eine Merkmalsanalyse der Sprache betrachtet. Spezieller erzeugt eine HMM-Maschine statistische Modelle in Form einer Liste der "N besten" von Subwortmodellhypothesen, die entsprechend durch ein HMM berechneten Konfidenzwerten oder Wahrscheinlichkeiten einer beobachteten Sequenz von akustischen Daten, wobei ein oder ein anderes Subwort gegeben ist, wie beispielsweise durch die Anwendung des Bayestheorems, eingeordnet werden.

[0038] Ein Bayes'scher HMM-Prozess identifiziert eine beste Hypothese entsprechend der wahrscheinlichsten Äußerung oder Subwortsequenz für eine gegebene Beobachtungssequenz von Akustikmerkmalsvektoren, und ihre Konfidenzwerte können von einer Vielzahl von Faktoren abhängen, die akustische Rauschabstände umfassen, die mit eingehenden akustischen Daten in Verbindung stehen. Das HMM kann auch eine statistische Verteilung umfas-

sen, welche eine Mischung von diagonalen Gauss-Verteilungen (Mixture of diagonal Gaussians) genannt wird, die einen Wahrscheinlichkeitspunktwert für jeden beobachteten Merkmalsvektor jedes Subworts liefert, wobei die Punktwerte verwendet werden können, um die Liste der N besten von Hypothesen neu zu ordnen. Die HMM-Maschine kann auch ein Subwort identifizieren und auswählen, dessen Modellwahrscheinlichkeitspunktwert der höchste ist.

[0039] Auf ähnliche Weise können einzelne HMMs für eine Sequenz von Subworten verknüpft werden, um Einzel- oder Mehrwort-HMMs herzustellen. Danach kann eine Liste der N besten von Einzel- oder Mehrwortreferenzmustern und zugehörigen Parameterwerten erzeugt und weiter bewertet werden.

[0040] Bei einem Beispiel verarbeitet der Spracherkennungsdecodierer **214** die Merkmalsvektoren unter Verwendung der geeigneten Akustikmodelle, Grammatiken und Algorithmen zum Erzeugen einer Liste der N besten von Referenzmustern. Wie hierin verwendet, kann der Begriff Referenzmuster durch Modelle, Wellenformen, Vorlagen, Rich-Signal-Modelle, Beispiele, Hypothesen oder andere Typen von Referenzen ausgetauscht werden. Ein Referenzmuster kann eine Reihe von Merkmalsvektoren umfassen, die ein oder mehrere Worte oder Subworte darstellen, und kann auf bestimmten Sprechern, Sprecharten und hörbaren Umgebungsbedingungen basieren. Fachleute werden erkennen, dass Referenzmuster durch ein geeignetes Referenzmustertraining des ASR-Systems erzeugt und in einem Speicher gespeichert werden können. Fachleute werden auch erkennen, dass gespeicherte Referenzmuster verändert werden können, wobei Parameterwerte der Referenzmuster auf der Grundlage von Spracheingabesignalunterschieden zwischen einem Referenzmustertraining und einer tatsächlichen Verwendung des ASR-Systems abgestimmt werden. Beispielsweise kann ein Satz von Referenzmustern, die für einen Fahrzeuginsassen oder bestimmte akustische Bedingungen trainiert werden, abgestimmt und als ein anderer Satz von Referenzmustern für einen anderen Fahrzeuginsassen oder andere akustische Bedingungen gespeichert werden, und zwar auf der Grundlage einer begrenzten Menge von Trainingsdaten von dem anderen Fahrzeuginsassen oder den anderen akustischen Bedingungen. Mit anderen Worten sind die Referenzmuster nicht notwendigerweise fest und können während der Spracherkennung angepasst werden.

[0041] Unter Verwendung der vokabelinternen Grammatik und eines beliebigen geeigneten Decodiereralgorithmus/beliebiger geeigneter Decodiereralgorithmen und eines beliebigen geeigneten Akustikmodells/beliebiger geeigneter Akustikmodelle ruft der Prozessor von dem Speicher verschiedene Referenzmuster ab, die das Testmuster in-

terpretieren. Beispielsweise kann der Prozessor eine Liste von N besten Vokabelergebnissen oder Referenzmustern zusammen mit entsprechenden Parameterwerten erzeugen und im Speicher speichern. Beispielsweise können Konfidenzpunktwerte jedes Referenzmusters in der Vokabelliste der N besten und zugehörige Segmentdauern, Wahrscheinlichkeitspunktwerte, Rauschabstandswerte (SNR-Werte) und/oder dergleichen umfassen. Die Vokabelliste der N besten kann mit absteigender Größe des Parameterwerts/der Parameterwerte geordnet werden. Beispielsweise ist das Vokabelreferenzmuster mit dem höchsten Konfidenzpunktwert das erstbeste Referenzmuster usw. Sobald eine Kette von erkannten Subworten hergestellt ist, können sie verwendet werden, um Worte mit einer Eingabe von den Wortmodellen **222** zu konstruieren und Sätze mit der Eingabe von den Sprachmodellen **224** zu konstruieren.

[0042] Schließlich empfängt/empfangen das Postprozessor-Softwaremodul/die Postprozessor-Softwaremodule **216** die Ausgangsdaten von dem Decodierermodule/den Decodierern **214** für einen beliebigen geeigneten Zweck. Bei einem Beispiel kann/können das Postprozessor-Softwaremodul/die Postprozessor-Softwaremodule **216** eines der Referenzmuster aus der Liste der N besten von Einzel- oder Mehrwortreferenzmustern als erkannte Sprache identifizieren oder auswählen. Bei einem anderen Beispiel kann/können das Postprozessormodul/die Postprozessormodule **216** verwendet werden, um akustische Daten in Text oder Ziffern zur Verwendung bei anderen Aspekten des ASR-Systems oder anderen Fahrzeugsystemen umzuwandeln. Bei einem anderen Beispiel kann/können das Postprozessormodul/die Postprozessormodule **216** verwendet werden, um dem Decodierer **214** oder dem Präprozessor **212** eine Trainingsrückmeldung zu liefern. Genauer gesagt kann der Postprozessor **216** verwendet werden, um Akustikmodelle für das Decodierermodule/die Decodierern **214** zu trainieren, oder um Abstimmungsparameter für das Präprozessormodul/die Präprozessormodule **212** zu trainieren.

[0043] Nun auf **Fig. 3** Bezug nehmend ist ein Verfahren **300** zum Betreiben einer Fahrzeugtelematikeinheit **30** gezeigt. Das Verfahren **300** beginnt in Schritt **310** durch Optimieren einer Spracherkennungsfunktion an dem Fahrzeug **12**, um eine Gruppe von Stimmen zu erkennen, die bei einem oder mehreren automatisierten Telefonteilnehmern (ATAs) verwendet wird. ATAs, die auch als Sprachdialogsysteme (IVR-Systeme von interactive voice response systems) bezeichnet werden können, leiten im Allgemeinen Anrufer an ihre gewünschte Nummer weiter, ohne menschliche Berater zu verwenden. Das heißt, Anrufer können über eine Sprachausgabe und/oder DTMF-Töne oder Befehle mit Computer, die durch die ATAs verwendet werden, in Interaktion treten. Zumindest

ein Teil hiervon kann durch Verwenden des oben beschriebenen ASR-Systems **210** erreicht werden. Das ASR-System **210** kann nicht nur Sprache von einem Fahrzeuginsassen empfangen, sondern kann auch angewiesen werden, durch einen ATA erzeugte Sprache zu empfangen. Auf diese Weise kann das ASR-System **210** wechselweise angewiesen werden, in Abhängigkeit davon, welche Quelle spricht, entweder auf den Telematikdienstbenutzer/Fahrzeuginsassen oder den ATA zu hören. Der Begriff "Anrufer", wie er hierin verwendet wird, kann auch als "Fahrzeuginsasse" oder "Telematikdienstbenutzer" umfassend verstanden werden.

[0044] Ferner kann bei denn verwendeten ATAs nur eine begrenzte Anzahl von verschiedenen Stimmen verwendet werden, um eine Ausgabe von dem ATA zu erzeugen. Oder mit anderen Worten, es gibt möglicherweise nur eine Handvoll eindeutiger Stimmen, die zur Verwendung bei einem ATA aufgezeichnet werden. Und die Anzahl an Anforderungen, die von dem ATA ausgegeben werden kann, kann auch begrenzt werden. Beispielsweise kann erwartet werden, dass der ATA Befehle wie beispielsweise "Bitte geben Sie den Konferenzgesprächzugangscode an" oder "Bitte geben Sie den Sicherheitscode an" äußert. Diese Befehle können auf der Grundlage des Kontexts der Konversation, die verwendet wird, um ein Konferenzgespräch aufzubauen, in vernünftiger Weise vorhergesagt werden. Mit der begrenzten Anzahl an verschiedenen Stimmen und/oder den begrenzten Phrasen oder Befehlen, die der ATA erzeugen kann, kann das ASR-System **210** programmiert (z. B. speziell trainiert) werden, um diese Stimmen und/oder Befehle wie oben erläutert vorherzusehen. Das Verfahren **300** fährt mit Schritt **320** fort.

[0045] In Schritt **320** werden ein Konferenzgesprächzugangscode, ein Sicherheitscode oder beides zum Herstellen eines Konferenzgesprächs von einem Telematikdienstbenutzer erhalten und gespeichert. Der Konferenzgesprächzugangscode kann ein Code mit mehreren Ziffern sein, den ein Anrufer dem ATA liefert, wenn er an einem Konferenzgespräch teilnimmt. Ähnlich könnte der Konferenzgesprächzugangscode zusammen mit dem Sicherheitscode verwendet werden, um unter jenen, die an einem bestimmten Konferenzgespräch teilnehmen, einen Anrufer, wie beispielsweise einen Fahrzeuginsassen oder einen Telematikdienstteilnehmer, weiter zu identifizieren. Beispielsweise könnte der Anrufer den Konferenzgesprächzugangscode liefern, um ein bestimmtes Konferenzgespräch zu identifizieren, und der Sicherheitscode könnte auch geliefert werden, um die Identität des Anrufers anzugeben; die Anruferidentität kann mit zusätzlichen Privilegien in Verbindung stehen, wie beispielsweise als Host des Konferenzgesprächs vorgesehen zu sein. Der Konferenzgesprächzugangscode und/oder der Sicherheitscode können auf eine Vielzahl von Arten und an einer Viel-

zahl von Orten gespeichert sein. Beispielsweise können der Konferenzgesprächzugangscode und/oder der Sicherheitscode durch den Fahrzeuginsassen an dem Fahrzeug **12** geäußert werden, durch das ASR-System **210** über das Mikrofon **32** empfangen werden, durch das ASR-System **210** interpretiert werden und in einem Speicher an dem Fahrzeug **12** gespeichert werden. Bei einem anderen Beispiel könnten der Konferenzgesprächzugangscode und/oder der Sicherheitscode unter Verwendung eines Web-Portals gespeichert werden, auf das ein Benutzer über einen Personalcomputer zugreifen kann. Der Konferenzgesprächzugangscode und/oder der Sicherheitscode können in einem Konferenzgesprächprofil gespeichert werden, das dann für eine Verwendung auf Anforderung drahtlos an das Fahrzeug **12** gesendet werden kann, wenn der Fahrzeuginsasse ein Konferenzgespräch herstellen möchte. Das Konferenzgesprächprofil kann eine(n) oder mehrere Konferenzgesprächzugangscodes, Sicherheitscodes oder Telefonnummern, die verwendet werden können, um das Konferenzgespräch zu starten, und/oder Regeln/Protokolle umfassen, die bei jedem Konferenzgesprächdienstanbieter verwendet werden können. Der Konferenzgesprächzugangscode und/oder der Sicherheitscode können mit einem gesprochenen Nametag verknüpft sein, wie beispielsweise einem verbalen Befehl, der ein Wort oder Worte umfasst, welche durch den Fahrzeuginsassen geäußert werden können, wenn ein Konferenzgespräch gewünscht wird. Das gesprochene Nametag kann eine oder mehrere Telefonnummern umfassen, die verwendet werden können, um mit dem Konferenzgespräch zu beginnen, und kann an dem Fahrzeug **12** gespeichert sein. Während das Verfahren **300** hierin hinsichtlich des Konferenzgesprächzugangscode beschrieben wird, ist zu verstehen, dass auch der Sicherheitscode oder der Konferenzgesprächzugangscode und der Sicherheitscode oder eine andere Kombination einer eindeutigen Information gemeint sein kann, die verwendet wird, um den Anrufer und/oder das Konferenzgespräch, an dem der Anrufer versucht teilzunehmen, zu identifizieren. Das Verfahren **300** fährt mit Schritt **330** fort.

[0046] In Schritt **330** wird ein Konferenzgespräch von einem Fahrzeug auf der Grundlage von verbalen Befehlen initiiert, die von einem Fahrzeuginsassen empfangen werden. Der Fahrzeuginsasse oder Telematikdienstbenutzer kann das Konferenzgespräch durch Drücken des Knopfs **32** und äußern der Telefonnummer, die das Konferenzgespräch bedient, initiieren. Bei einem anderen Beispiel kann der Fahrzeuginsasse das gesprochene Nametag, wie beispielsweise ein Schlüsselwort, das mit dem Konferenzgesprächzugangscode in Verbindung steht, äußern, um den Prozess des Initiierens des Konferenzgesprächs zu starten. Der Initiierungsprozess kann sehr ähnlich zu einem Initiieren eines normalen Freisprechgesprächs von dem Fahrzeug **12** sein – wenn

er nicht exakt der gleiche ist. Im Gegensatz zu einem Tätigen normaler Telefongespräche kann das Verfahren **300** jedoch ein Halten des ASR-Systems **210** in einem bereiten Zustand umfassen, der automatisch eine weitere Information erzeugen kann, um das Konferenzgespräch herzustellen. Die Initiierung des Konferenzgesprächs kann auch das Anweisen des ASR-Systems **210**, mit einem Hören auf Sprache von dem ATA zu beginnen, umfassen. Beispielsweise kann der gespeicherte Konferenzgesprächzugangscod in einem Profil gespeichert sein, das eine von einem Computer lesbare Anweisung zum Initiieren des ASR-Systems **210**, wenn die zum Herstellen des Konferenzgesprächs verwendete Telefonnummer angerufen wird, umfasst. Und wenn das ASR-System **210** initiiert ist, kann es angewiesen werden, das Hören auf eine Eingabe von dem Fahrzeuginsassen über das Mikrofon **32** zu beenden, sobald das Konferenzgespräch initiiert (z. B. gewählt) wurde, und mit dem Hören auf die Eingabe, die von einem sich außerhalb befindlichen Anrufer empfangen wird (z. B. einem ATA), zu beginnen. In Abhängigkeit von der Quelle des akustischen Geräuschs kann das ASR-System **210** angewiesen werden, wechselweise auf den ATA oder den Anrufer zu hören. Das Verfahren **300** fährt mit Schritt **340** fort.

[0047] In Schritt **340** wird automatisch ermittelt, dass ein ATA einen Konferenzgesprächzugangscod anfordert, wobei die optimierte Spracherkennungsfunktion verwendet wird. Dies kann während des Konferenzgesprächinitiierungsprozesses oder kurz danach stattfinden. Beispielsweise kann der ATA, nachdem der Konferenzgesprächanruf getätigt wurde, automatisch den Anruf beantworten und eine Aufforderung für den Anrufer erzeugen, die eine Information anfordert, bevor ein Konferenzgespräch hergestellt wird. Diese Aufforderung kann den ATA identifizieren und den Anrufer bitten, den Konferenzgesprächzugangscod und/oder den Sicherheitscod einzugeben. Das ASR-System **210** kann auf die Aufforderung von dem ATA hören und basierend auf dieser Aufforderung auf eine Information zugreifen, die an dem Fahrzeug gespeichert ist, wie beispielsweise den Konferenzgesprächzugangscod und/oder den Sicherheitscod. Das Verfahren **300** fährt mit Schritt **350** fort.

[0048] In Schritt **350** wird der gespeicherte Konferenzgesprächzugangscod auf der Grundlage der Ermittlung der Spracherkennungsfunktion an den ATA gesendet. Sobald die ATA-Anforderung interpretiert und auf den Konferenzgesprächzugangscod/Sicherheitscod zugegriffen wurde, kann die Telematikeinheit **30** den Konferenzgesprächzugangscod/Sicherheitscod an den ATA liefern. Dies kann umfassen, dass DTMF-Töne oder DTMF-Befehle auf der Grundlage des gespeicherten Konferenzgesprächzugangscodes in Ansprechen auf die Ermittlung, dass der ATA einen Konferenzgesprächzugangscod anfordert, erzeugt werden. Die DTMF-

Töne oder DTMF-Befehle können drahtlos an den ATA übertragen werden, um einen Zugang zum Konferenzgespräch herzustellen. Bei einem anderen Beispiel kann das ASR-System **210** verbale Antworten auf ATA-Anforderungen erzeugen. Wenn beispielsweise der ATA fordert, dass der Anrufer den Konferenzgesprächzugangscod nennt, kann das ASR-System **210** den Konferenzgesprächzugangscod/Sicherheitscod empfangen und diese Information für den ATA hörbar erzeugen. Der ATA kann den Anrufer nach mehr als einer Antwort fragen. In diesem Fall kann das ASR-System **210** zumindest einige der obigen Schritte für weitere ATA-Abfragen wiederholen. Dann endet das Verfahren **300**.

[0049] Es ist zu verstehen, dass das Vorangehende eine Beschreibung einer oder mehrerer bevorzugter beispielhafter Ausführungsformen der Erfindung ist. Die Erfindung ist nicht auf die hierin offenbarte(n) bestimmte(n) Ausführungsform(en) beschränkt, sondern ist lediglich durch die nachstehenden Ansprüche definiert. Ferner betreffen die in der vorangehenden Beschreibung enthaltenen Aussagen bestimmte Ausführungsformen und sollen nicht als Einschränkungen des Schutzzumfangs der Erfindung oder der Definition von Begriffen, die in den Ansprüchen verwendet werden, betrachtet werden, außer, wenn oben ein Ausdruck oder eine Phrase ausdrücklich definiert ist. Für den Fachmann werden verschiedene andere Ausführungsformen und verschiedene Änderungen und Abwandlungen der offenbarten Ausführungsform(en) ersichtlich werden. Alle anderen Ausführungsformen, Änderungen und Abwandlungen sollen innerhalb des Schutzzumfangs der beigefügten Ansprüche liegen.

[0050] Wie in dieser Beschreibung und den Ansprüchen verwendet, sollen die Begriffe "zum Beispiel", "beispielsweise", "wie beispielsweise" und "wie" und die Verben "umfassen", "aufweisen", "einschließen" und ihre anderen Verbformen, wenn sie in Verbindung mit einer Auflistung einer oder mehrerer Komponenten oder eines oder mehrerer anderer Elemente verwendet werden, jeweils als ein offenes Ende aufweisend betrachtet werden, was bedeutet, dass die Auflistung nicht als andere, zusätzliche Komponenten oder Elemente ausschließend betrachtet werden soll. Andere Begriffe sollen als ihre breiteste vernünftige Bedeutung umfassend betrachtet werden, wenn sie nicht in einem Kontext verwendet werden, der eine andere Interpretation erfordert.

Patentansprüche

1. Verfahren zum Betreiben einer Fahrzeugtelematikeinheit, das die Schritte umfasst, dass:
 - (a) ein Konferenzgesprächzugangscod gespeichert wird, der von einem Telematikdienstbenutzer erhalten wird;

(b) ein Konferenzgespräch von einem Fahrzeug auf der Grundlage von verbalen Befehlen initiiert wird, die von einem Fahrzeuginsassen empfangen werden;
 (c) eine Anforderung hinsichtlich des Konferenzgesprächzugangscodes von einem automatisierten Telefonteilnehmer (ATA) empfangen wird, während das Konferenzgespräch initiiert wird;
 (d) unter Verwendung einer Spracherkennungsfunktion an dem Fahrzeug ermittelt wird, dass der ATA den Konferenzgesprächzugangscodes angefordert hat; und
 (e) der gespeicherte Konferenzgesprächzugangscodes auf der Grundlage der Ermittlung der Spracherkennungsfunktion an den ATA gesendet wird.

2. Verfahren nach Anspruch 1, wobei die Spracherkennungsfunktion programmiert ist, um eine oder mehrere Stimmen zu erkennen, die durch den ATA verwendet werden.

3. Verfahren nach Anspruch 1, wobei die Spracherkennungsfunktion programmiert ist, um einen oder mehrere Befehle, die durch den ATA erzeugt werden, zu erkennen.

4. Verfahren nach Anspruch 1, das ferner den Schritt umfasst, dass die Spracherkennungsfunktion angewiesen wird, in Abhängigkeit davon, wer spricht, wechselweise auf den Fahrzeuginsassen oder den ATA zu hören.

5. Verfahren nach Anspruch 1, wobei Schritt (a) ferner umfasst, dass:
 darauf gehört wird, dass der Fahrzeuginsasse den Konferenzgesprächzugangscodes in dem Fahrzeug ausspricht;
 der gesprochene Konferenzgesprächzugangscodes unter Verwendung der Spracherkennungsfunktion an dem Fahrzeug interpretiert wird; und
 der interpretierte Konferenzgesprächzugangscodes an dem Fahrzeug gespeichert wird.

6. Verfahren nach Anspruch 1, wobei der Konferenzgesprächzugangscodes als Teil eines Konferenzgesprächprofils gespeichert ist, das ein gesprochenes Nametag oder zumindest eine Regel umfasst, um mit dem ATA in Interaktion zu treten.

7. Verfahren nach Anspruch 1, wobei die Initiierung des Konferenzgesprächs auf der Grundlage der Identität einer Telefonnummer, die von dem Fahrzeuginsassen gewählt wird, ermittelt wird.

8. Verfahren nach Anspruch 1, wobei der Konferenzgesprächzugangscodes drahtlos von dem Fahrzeug an den ATA als Doppelton-Mehrfrequenz-Töne (DTMF-Töne) oder DTMF-Befehle gesendet wird.

9. Verfahren nach Anspruch 1, wobei der Konferenzgesprächzugangscodes durch die Spracherken-

nungsfunktion verbal für den ATA ausgesprochen wird.

10. Verfahren zum Betreiben einer Fahrzeugtelematikeinheit, das die Schritte umfasst, dass:

(a) eine Spracherkennungsfunktion an dem Fahrzeug optimiert wird, um eine Gruppe von Stimmen zu erkennen, die bei einem oder mehreren automatisierten Telefonteilnehmern (ATA) verwendet werden;
 (b) ein Konferenzgesprächzugangscodes, ein Sicherheitscodes oder beides von einem Telematikdienstbenutzer erhalten werden, um ein Konferenzgespräch an dem Fahrzeug herzustellen;
 (c) der Konferenzgesprächzugangscodes, der Sicherheitscodes oder beides an dem Fahrzeug gespeichert werden;
 (d) unter Verwendung der optimierten Spracherkennungsfunktion automatisch ermittelt wird, dass ein ATA einen Konferenzgesprächzugangscodes anfordert;
 (e) DTMF-Töne oder DTMF-Befehle auf der Grundlage des gespeicherten Konferenzgesprächzugangscodes, Sicherheitscodes oder beider in Ansprechen auf die Ermittlung in Schritt (d) erzeugt werden; und
 (f) die DTMF-Töne oder DTMF-Befehle drahtlos an den ATA übertragen werden, um einen Zugang zu dem Konferenzgespräch herzustellen.

Es folgen 3 Seiten Zeichnungen

Anhängende Zeichnungen

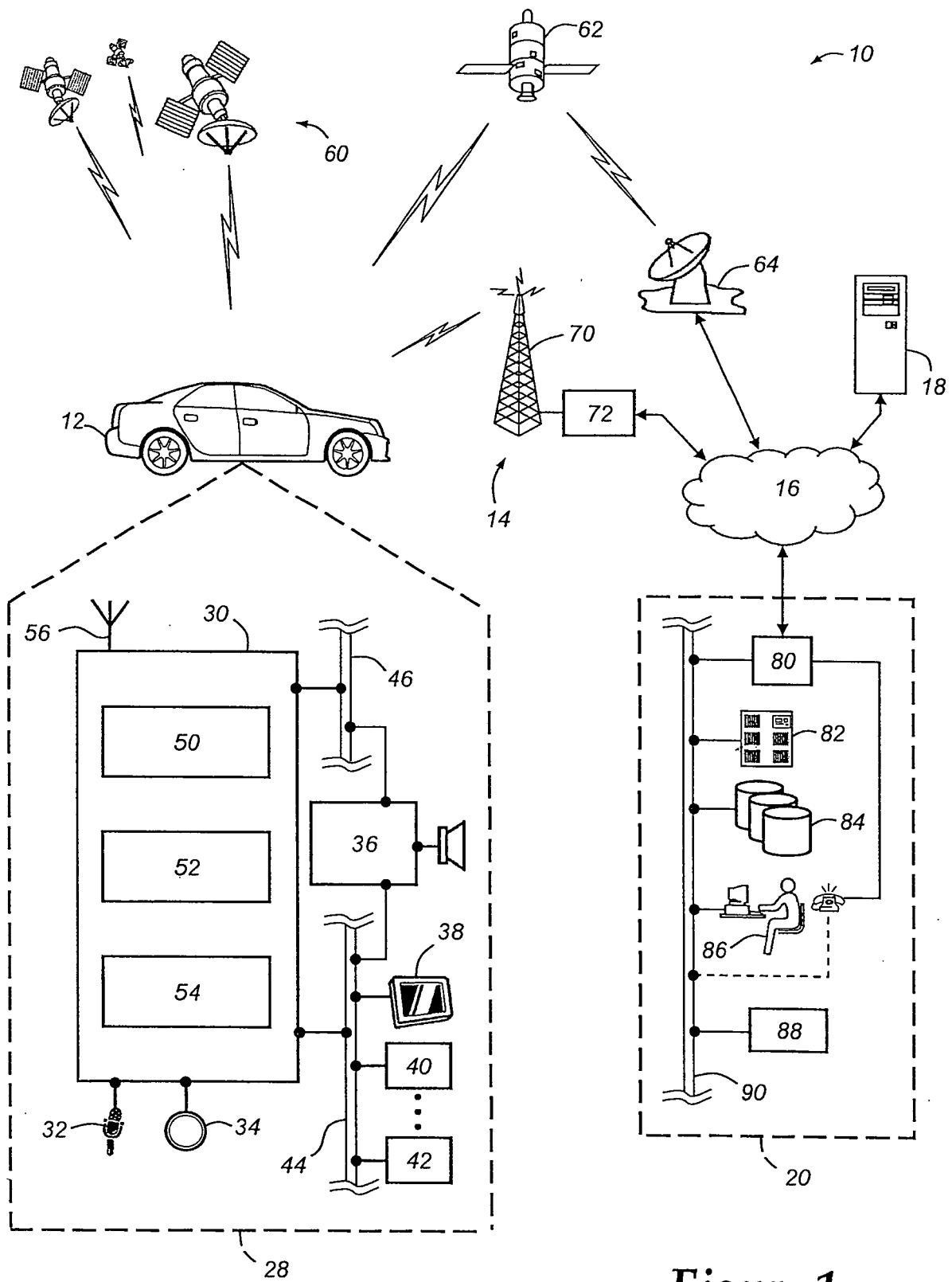
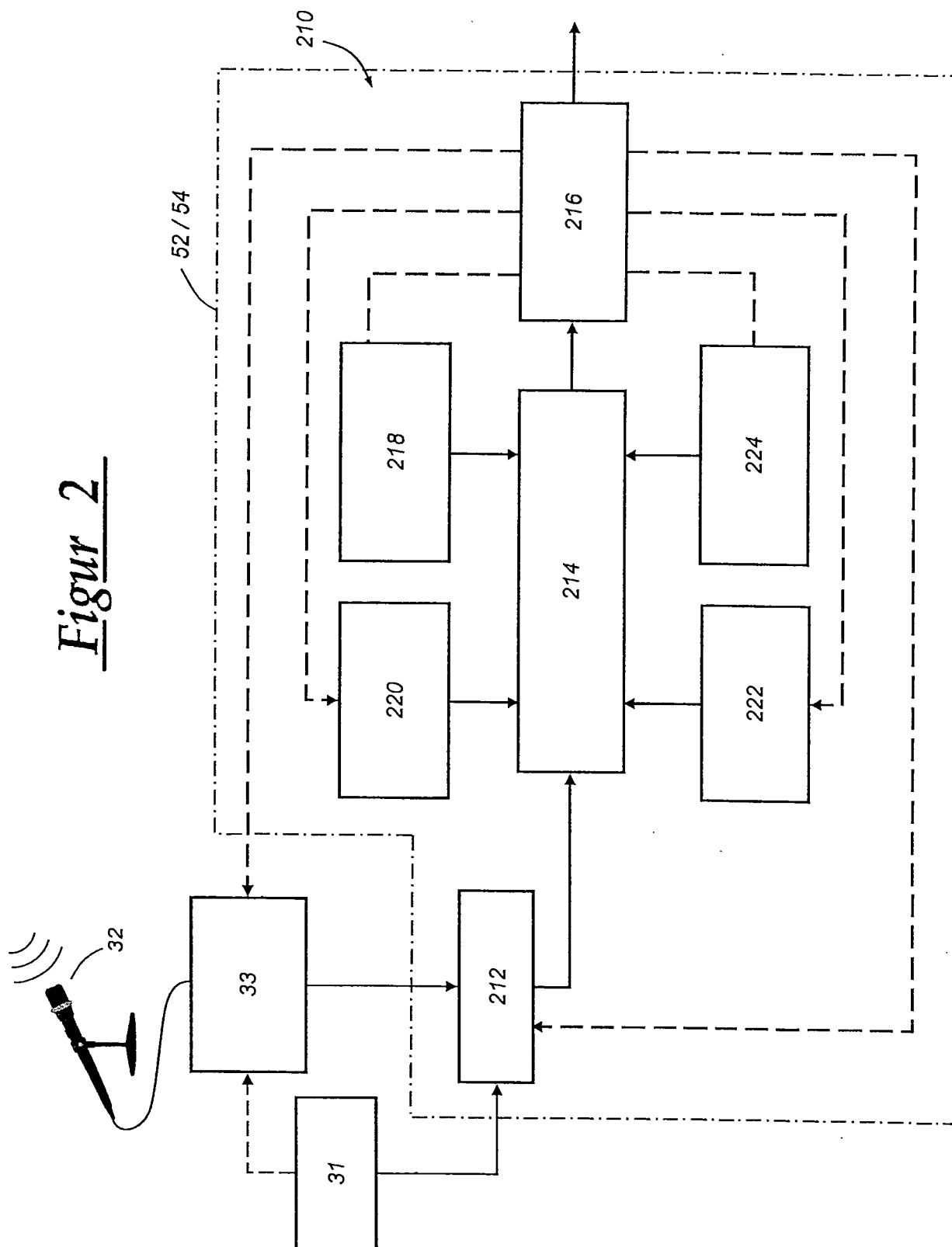
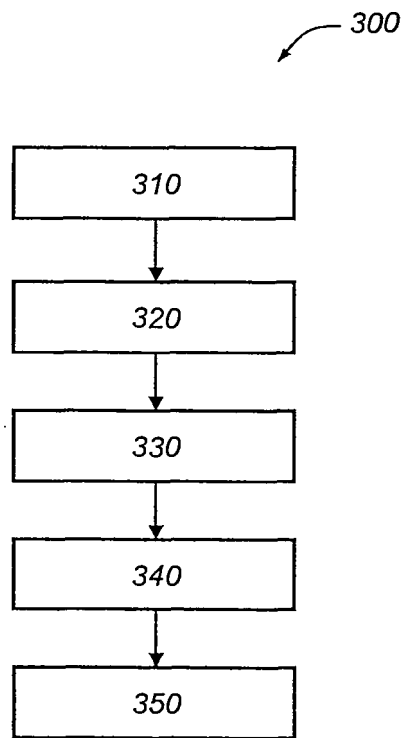
Figur 1

Figure 2





Figur 3