

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges

Eigentum

Internationales Büro

(43) Internationales

Veröffentlichungsdatum

2. April 2015 (02.04.2015)



(10) Internationale Veröffentlichungsnummer

WO 2015/044097 A1

(51) Internationale Patentklassifikation:

G10L 15/30 (2013.01) B60R 16/037 (2006.01)

G10L 15/18 (2013.01)

(21) Internationales Aktenzeichen: PCT/EP2014/070164

(22) Internationales Anmeldedatum:

23. September 2014 (23.09.2014)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:

10 2013 219 649.0

27. September 2013 (27.09.2013)

DE

(71) Anmelder: CONTINENTAL AUTOMOTIVE GMBH

[DE/DE]; Vahrenwalder Straße 9, 30165 Hannover (DE).

(72) Erfinder: WALTHER, Jens; Waldstraße 11, 35112

Fronhausen (DE).

DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Bestimmungsstaaten (soweit nicht anders angegeben, für

jede verfügbare regionale Schutzrechtsart): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), eurasisches (AM, AZ, BY, KG, KZ, RU, TJ, TM), europäisches (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

— mit internationalem Recherchenbericht (Artikel 21 Absatz 3)

(81) Bestimmungsstaaten (soweit nicht anders angegeben, für

jede verfügbare nationale Schutzrechtsart): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK,

(54) Title: METHOD AND SYSTEM FOR CREATING OR AUGMENTING A USER-SPECIFIC SPEECH MODEL IN A LOCAL DATA MEMORY THAT CAN BE CONNECTED TO A TERMINAL

(54) Bezeichnung : VERFAHREN UND SYSTEM ZUM ERSTELLEN ODER ERGÄNZEN EINES BENUTZERSPEZIFISCHEN SPRACHMODELLS IN EINEM MIT EINEM ENDGERÄT VERBINDBAREN LOKALEN DATENSPEICHER

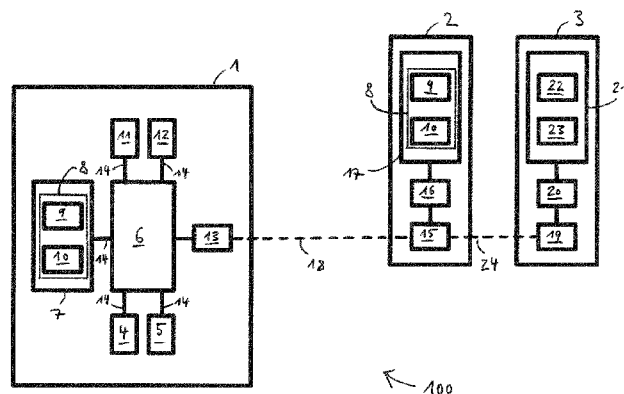


Fig. 1

(57) Abstract: The invention relates to a method for creating or augmenting a user-specific speech model (8) in a local data memory (7) that can be connected to a terminal (4, 5), wherein the speech model (8) is set up to assign control commands for controlling the terminal (4, 5) to natural-speech utterances from a user, comprising the steps of: – a natural-speech utterance being spoken by the user; – the natural-speech utterance being transmitted to an external data processing unit (2) and the natural-speech utterance being transcribed; – the speech model (8) being created or augmented on the external data processing unit (2) using the transcribed natural-speech utterance; – the speech model (8) being transmitted fully or at least partially from the external data processing unit (2) to the local data memory (7) and the speech model (8) being stored and/or augmented in the local data memory (7). The invention also relates to a system (100) that can be used to carry out the method.

(57) Zusammenfassung:

[Fortsetzung auf der nächsten Seite]

WO 2015/044097 A1



Die Erfindung betrifft ein Verfahren zum Erstellen oder Ergänzen eines benutzerspezifischen Sprachmodells (8) in einem mit einem Endgerät (4, 5) verbindbaren lokalen Datenspeicher (7), wobei das Sprachmodell (8) eingerichtet ist, natürlichsprachlichen Äußerungen eines Benutzers Steuerbefehle zum Steuern des Endgerätes (4, 5) zuzuordnen, umfassend die Schritte: - Einsprechen einer natürlichsprachlichen Äußerung durch den Benutzer; - Übermitteln der natürlichsprachlichen Äußerung an eine externe Datenverarbeitungseinheit (2) und Transkribieren der natürlichsprachlichen Äußerung; - Erstellen oder Ergänzen des Sprachmodells (8) auf der externen Datenverarbeitungseinheit (2) anhand der transkribierten natürlichsprachlichen Äußerung; - vollständiges oder wenigstens teilweises Übertragen des Sprachmodells (8) von der externen Datenverarbeitungseinheit (2) an den lokalen Datenspeicher (7) sowie Speichern und/oder Ergänzen des Sprachmodells (8) in dem lokalen Datenspeicher (7). Die Erfindung betrifft ferner ein System (100), mit dem das Verfahren durchführbar ist.

Beschreibung

Verfahren und System zum Erstellen oder Ergänzen eines benutzerspezifischen Sprachmodells in einem mit einem Endgerät
5 verbindbaren lokalen Datenspeicher

Die Erfindung betrifft ein Verfahren zum Erstellen oder Ergänzen eines benutzerspezifischen Sprachmodells in einem mit einem Endgerät verbindbaren lokalen Datenspeicher, wobei das Sprach-
10 modell eingerichtet ist, natürlichsprachlichen Äußerungen eines Benutzers Steuerbefehle zum Steuern des Endgerätes zuzuordnen. Die Erfindung betrifft ferner ein System mit dem das Verfahren durchführbar ist.

15 Aus dem Stand der Technik ist es bekannt, Endgeräte, beispielsweise Infotainmentsysteme oder Navigationssysteme für PKW oder LKW, durch Eingabe natürlichsprachlicher Äußerungen zu steuern. Zu diesem Zweck werden Spracherkenner gewöhnlich mit statistischen Sprachmodellen (SLM) und mit semantischen Modellen
20 (SEM) ergänzt. Bei einer natürlichsprachlichen Äußerung handelt es sich normalerweise um einen von dem Benutzer gesprochenen Text oder auch von dem Benutzer geäußerte Laute. Die genannten Modelle sind darauf ausgerichtet, eine von dem Benutzer z. B. in eine entsprechende akustische Eingabeeinheit eingegebene natür-
25 lichsprachliche Äußerung zu erkennen und der Äußerung einen Steuerbefehl zum Steuern des Endgerätes zuzuordnen. Gewöhnlich wird durch diesen Steuerbefehl dann eine definierte Aktion oder Funktion des Endgerätes ausgelöst.

30 Die genannten Modelle, mit denen bekannte Spracherkenner gewöhnlich ergänzt werden, haben jedoch den Nachteil, dass sie sehr umfangreich sind und viel Speicherplatz benötigen, damit sie in der Lage sind, den Wortschatz einer Vielzahl von Benutzern abzudecken. Die Folge sind lange Latenzzeiten bei der Benutzung.
35 Ferner bedingt die Verwendung dieser Modelle den Einsatz leistungsstarker und entsprechend teurer Prozessoren im Endgerät.

Der vorliegenden Erfindung liegt somit die Aufgabe zugrunde, ein Verfahren zur Erstellung eines möglichst leistungsstarken und gleichzeitig möglichst preisgünstig implementierbaren Sprachmodells zu entwickeln, wobei das Modell eingerichtet sein soll, natürlichsprachlichen Äußerungen eines Benutzers Steuerbefehle zum Steuern eines Endgerätes zuzuordnen. Ferner soll ein System angegeben werden, mit dem dieses Verfahren durchführbar ist.

Diese Aufgabe wird gelöst durch ein Verfahren und ein System gemäß den unabhängigen Ansprüchen. Spezielle Ausgestaltungen des erfindungsgemäßen Verfahrens und des erfindungsgemäßen Systems sind in den Unteransprüchen beschrieben.

Vorgeschlagen wird also ein Verfahren zum Erstellen oder Ergänzen eines benutzerspezifischen Sprachmodells in einem mit wenigstens einem Endgerät verbindbaren lokalen Datenspeicher, wobei das Sprachmodell eingerichtet ist, natürlichsprachlichen Äußerungen eines Benutzers Steuerbefehle zum Steuern des Endgerätes zuzuordnen, umfassend die Schritte:

- Einsprechen einer natürlichsprachlichen Äußerung durch den Benutzer;
- Übermitteln der natürlichsprachlichen Äußerung an eine externe Datenverarbeitungseinheit und Transkribieren der natürlichsprachlichen Äußerung, vorzugsweise Transkribieren in eine Textform;
- Erstellen oder Ergänzen des benutzerspezifischen Sprachmodells auf der externen Datenverarbeitungseinheit anhand der transkribierten natürlichsprachlichen Äußerung;
- vollständiges oder wenigstens teilweises Übertragen des Sprachmodells von der externen Datenverarbeitungseinheit an den mit dem Endgerät verbindbaren lokalen Datenspeicher sowie Speichern und/oder Ergänzen des Sprachmodells in dem lokalen Datenspeicher.

Dadurch, dass das Erstellen und/oder Ergänzen des benutzerspezifischen Sprachmodells auf der externen Datenverarbei-

tungseinheit durchgeführt wird und dass das Sprachmodell anschließend vollständig oder wenigstens teilweise von der externen Datenverarbeitungseinheit an den lokalen Datenspeicher übertragen sowie in dem lokalen Datenspeicher gespeichert und/oder ergänzt wird, wird ein zur Steuerung des Endgerätes nutzbares Sprachmodell zur Verfügung gestellt, das auf den jeweiligen Benutzer zugeschnitten ist. Bei vergleichsweise geringem Speicherbedarf kann das derart im lokalen Datenspeicher zur Verfügung gestellte Sprachmodell den von diesem Benutzer zur Steuerung des Endgerätes verwendeten Wortschatz abdecken. Der lokale Datenspeicher kann also kleiner ausgelegt sein als dies bei ähnlich leistungsstarken Spracherkennern der Fall ist, die aus dem Stand der Technik bekannt sind. Gleichzeitig können gute Ergebnisse und kurze Latenzzeiten auch mit vergleichsweise leistungsschwachen und damit preisgünstigen Prozessoren erzielt werden.

Der Vorgang des Transkribierens kann insbesondere das Erstellen einer orthografischen Darstellung der natürlichsprachlichen Äußerung umfassen, z. B. in Form einer Folge von Worten und/oder Zeichen. Dabei können beliebige Schriftarten und Alphabete verwendet werden. Vorzugsweise liegt die Darstellung nach dem Transkribieren in digitaler Form vor, z. B. in Gestalt einer Textdatei.

Vorgeschlagen wird ferner ein System, mit dem das hier vorgeschlagene Verfahren durchführbar ist. Es handelt es sich also um ein System zum Erstellen oder Ergänzen eines benutzerspezifischen Sprachmodells in einem mit einem Endgerät verbindbaren lokalen Datenspeicher, wobei das Sprachmodell eingerichtet ist, natürlichsprachlichen Äußerungen eines Benutzers Steuerbefehle zum Steuern des Endgerätes zuzuordnen. Das System umfasst:

- wenigstens ein Endgerät;
- einen mit dem Endgerät verbindbaren lokalen Datenspeicher zum Speichern des Sprachmodells;
- eine akustische Eingabeeinheit, über die natürlichsprachliche Äußerungen zur Steuerung des Endgerätes eingebar sind;

- eine externe Datenverarbeitungseinheit, an die die natürlichsprachlichen Äußerungen übermittelbar sind, wobei die externe Datenverarbeitungseinheit eingerichtet ist, anhand der natürlichsprachlichen Äußerungen ein benutzerspezifisches Sprachmodell zu erstellen oder zu ergänzen und das Sprachmodell vollständig oder wenigstens teilweise zum Speichern in dem lokalen Datenspeicher an den lokalen Datenspeicher zu übertragen und
- eine mit der akustischen Eingabeeinheit und mit dem lokalen Datenspeicher verbindbare lokale Sende- und Empfangseinheit zum Übermitteln der natürlichsprachlichen Äußerung an die externe Datenverarbeitungseinheit und zum Empfangen des Sprachmodells von der externen Datenverarbeitungseinheit.

Vorzugsweise sind das Endgerät, der lokale Datenspeicher, die akustische Eingabeeinheit und die lokale Sende- und Empfangseinheit in einem Kraftfahrzeug anordenbar, angeordnet oder eingebaut, z. B. in einem PKW oder LKW. Der lokale Datenspeicher zum Speichern des Sprachmodells kann in das Endgerät integriert sein. Das Endgerät und der lokale Datenspeicher können jedoch auch als separate Einheiten ausgeführt sein. Auch die akustische Eingabeeinheit und die lokale Sende- und Empfangseinheit können in das Endgerät integriert oder als separate Einheiten ausgeführt sein. Das Endgerät und der lokale Datenspeicher können über eine drahtgebundene und/oder über eine drahtlose Datenverbindung verbindbar oder verbunden sein. Die akustische Eingabeeinheit umfasst typischerweise wenigstens ein Mikrofon.

Gewöhnlich umfasst das System zudem wenigstens einen lokalen Prozessor, der mit dem Endgerät, dem lokalen Datenspeicher, der akustischen Eingabeeinheit und der lokalen Sende- und Empfangseinheit drahtlos oder drahtgebunden verbindbar oder verbunden ist. Auch der lokale Prozessor ist vorzugsweise in dem PKW oder LKW angeordnet. Insbesondere kann der lokale Prozessor eingerichtet sein, durch Ausführen einer entsprechenden Spracherkennungssoftware das Transkribieren der eingegebenen natürlichsprachlichen Äußerung durchzuführen. Anhand des im

lokalen Datenspeicher gespeicherten benutzerspezifischen Sprachmodells kann der lokale Prozessor der Äußerung dann wenigstens einen Steuerbefehl aus einer Menge von Steuerbefehlen zum Steuern des Endgerätes zuordnen und diesen Steuerbefehl an das Endgerät senden. Das Endgerät kann dann eine durch den Steuerbefehl veranlasste Aktion oder Funktion ausführen.

Das Endgerät kann z. B. wenigstens

- ein Audiogerät zum Abspielen von Audiodateien und/oder
- 10 - ein Videogerät zum Abspielen von Bild- und/oder Videodateien und/oder
- ein Telekommunikationsgerät zum Empfangen und/oder Versenden von Text-, Sprach-, Bild- oder Videonachrichten und/oder
- ein Navigationsgerät und/oder
- 15 - einen Stellantrieb zum Verstellen der Position einer Fens- terscheibe, eines Spiegels, eines Sitzes oder einer Luft- zuführvorrichtung und/oder
- eine Anlage zur Regelung einer Temperatur und/oder einer Luftfeuchtigkeit und/oder
- 20 - eine Steuerung von aus dem Internet heruntergeladenen Ap- plikationssoftwareeinheiten

umfassen. Bei dem Endgerät kann es sich auch um einen Bordcomputer oder um ein tragbares Infotainmentgerät, wie beispielsweise ein Smartphone, handeln.

Die externe Datenverarbeitungseinheit kann z. B. ein zentraler Server eines Anbieters von Spracherkennern sein. Insbesondere kann die externe Datenverarbeitungseinheit eingerichtet sein, benutzerspezifische Sprachmodelle für eine Vielzahl ver-
30 schiedener Benutzer zu erstellen, zu ergänzen, zu speichern und an verschiedene lokale Datenspeicher zum Speichern auf diesen lokalen Datenspeichern zu senden. Natürlich können auch ver- schiedene benutzerspezifische Sprachmodelle verschiedener Benutzer auf demselben lokalen Datenspeicher gespeichert werden.
35 Dies kann sinnvoll sein, wenn mehrere Personen abwechselnd dasselbe Endgerät benutzen, z. B. wenn sie abwechselnd dasselbe Fahrzeug benutzen, in dem das Endgerät angeordnet oder eingebaut

ist. Nur der Einfachheit halber wird hier und im Folgenden nur das Erstellen und Ergänzen eines benutzerspezifischen Sprachmodells für nur einen gegebenen Benutzer beschrieben.

- 5 Die externe Datenverarbeitungseinheit umfasst normalerweise wenigstens eine externe Sende- und Empfangseinheit, wenigstens einen externen Prozessor zum Erstellen und/oder Ergänzen des benutzerspezifischen Sprachmodells und wenigstens einen externen Datenspeicher, in dem das benutzerspezifische Sprachmodell dauerhaft oder wenigstens zeitweise speicherbar ist. Vorzugsweise sind die lokale Sende- und Empfangseinheit und die externe Sende- und Empfangseinheit über eine drahtlose Datenverbindung, z. B. über eine drahtlose Internetverbindung oder über eine sonstige drahtlose Telekommunikationsverbindung
10 miteinander zum Austauschen von Daten verbindbar.
15

- Bei einer speziellen Ausführungsform des vorgeschlagenen Verfahrens umfasst das Erstellen oder das Ergänzen des benutzerspezifischen Sprachmodells auf der externen Datenverarbeitungseinheit ein Erstellen oder ein Ergänzen eines benutzerspezifischen statistischen Sprachmodells (US-SLM), wobei das US-SLM aus der transkribierten natürlichsprachlichen Äußerung des Benutzers gebildet wird oder um diese ergänzt wird und wobei die transkribierte natürlichsprachliche Äußerung transkribierte Worte und/oder transkribierte Laute und/oder transkribierte Wortfolgen und/oder transkribierte Lautfolgen umfasst. Alternativ oder zusätzlich kann das Erstellen oder das Ergänzen des benutzerspezifischen Sprachmodells auf der externen Datenverarbeitungseinheit ein Erstellen oder Ergänzen eines semantischen Sprachmodells (US-SEM) umfassen, wobei in der transkribierten natürlichsprachlichen Äußerung des Benutzers Schlüsselworte und/oder Wortstämme von Schlüsselworten und/oder Wortfolgen identifiziert werden und das US-SEM aus den Schlüsselworten und/oder den Wortstämmen und/oder den Wortfolgen gebildet wird oder um diese ergänzt wird. Schlüsselworte des semantischen Modells lassen sich trainieren, indem der Benutzer eine Aktion ausführt und dieser ein Schlüsselwort zuordnet.
20
25
30
35

In beiden Fällen basiert das erstellte benutzerspezifische Sprachmodell also vor allem auf dem Vokabular, der Ausdrucksweise und den Sprachgewohnheiten des jeweiligen Benutzers. Das be-
5 nutzerspezifische Sprachmodell kommt also ohne Vokabeln und Ausdrucksweisen aus, die der Benutzer zum Eingeben der Steuerbefehle gar nicht verwendet. Dies trägt in vorteilhafter Weise dazu bei, den Speicherbedarf des benutzerspezifischen Sprachmodells zu verringern, so dass der lokale Datenspeicher und der
10 lokale Prozessor kleiner und kostengünstiger ausgelegt sein können.

Eine weitere spezielle Ausführungsform des Verfahrens zeichnet sich dadurch aus, dass das Erstellen oder das Ergänzen des US-SEM
15 umfasst, dass den Schlüsselworten und/oder den Wortstämmen und/oder den Wortfolgen statistische Gewichte und/oder Attribute zugeordnet werden, wobei mittels der statistischen Gewichte und/oder mittels der Attribute eine Zuordnung des jeweiligen Schlüsselworts oder Wortstamms oder der jeweiligen Wortfolge zu
20 einer Teilmenge aus einer Menge von Steuerbefehlen herstellbar ist.

Insbesondere kann der eingesprochenen natürlichsprachlichen Äußerung wenigstens ein Steuerbefehl aus einer Menge von
25 Steuerbefehlen zugeordnet werden, wobei die Zuordnung des Steuerbefehls zur eingesprochenen natürlichsprachlichen Äußerung mittels des im Datenspeicher gespeicherten benutzerspezifischen Sprachmodells vorgenommen wird und/oder wobei die Zuordnung des Steuerbefehls zur eingesprochenen natürlich-
30 sprachlichen Äußerung mittels der externen Datenverarbeitungseinheit vorgenommen wird und der Steuerbefehl an das Endgerät übermittelt wird.

Die Zuordnung des Steuerbefehls zur eingesprochenen natürlichsprachlichen Äußerung kann dann insbesondere in Abhängigkeit
35 von den in der natürlichsprachlichen Äußerung identifizierten Schlüsselworten und/oder Wortstämmen und/oder Wortfolgen

und/oder in Abhängigkeit von den diesen Schlüsselworten und/oder Wortstämmen und/oder Wortfolgen im US-SEM jeweils zugeordneten statistischen Gewichten und/oder Attributen vorgenommen werden.

5 So kann einem Schlüsselwort oder einem entsprechenden Wortstamm, das oder der eines von mehreren Endgeräten bezeichnet, mittels eines entsprechenden Attributes die Menge derjenigen Steuerbefehle zugeordnet werden, die zur Steuerung dieses Endgerät dienen. Sind in oder an einem Fahrzeug beispielsweise ein
10 Audiogerät und ein Außenspiegel angeordnet, so kann mittels des semantischen Sprachmodells festgestellt werden, dass es sich bei einer Äußerung, die das Schlüsselwort „Außenspiegel“ enthält, um eine Äußerung zum Steuern des Außenspiegels und nicht zum Steuern des Audiogerätes handelt. Die Zuordnung eines Steuerbefehls zu
15 einer gegebenen Äußerung kann anhand der Attribute also auf eine oder mehrere (echte) Untermengen der Menge aller möglichen Steuerbefehle eingeschränkt werden.

Es ist denkbar, dass ein und dasselbe Schlüsselwort sich auf
20 unterschiedliche Steuerbefehle beziehen kann. Beispielsweise kann das Schlüsselwort „verstellen“ sich im Zusammenhang mit einem Außenspiegel eines Fahrzeugs auf das Verstellen bezüglich unterschiedlicher Achsen beziehen. Einem gegebenen Schlüsselwort können also mehrere Attribute zugeordnet sein, wobei
25 durch jedes der Attribute die Zuordnung zu unterschiedlichen Teilmengen aller möglichen Steuerbefehle zum Ausdruck kommt. Den Schlüsselworten können auch statistische Gewichte zugeordnet werden. Diese können die unterschiedlichen Wahrscheinlichkeiten kennzeichnen, mit denen ein gegebenes Schlüsselwort ver-
30 schiedenen Steuerbefehlen zugeordnet wird.

Beispielsweise kann das Schlüsselwort „Suchen“ sich auf die Suche in einem Adressbuch oder auf die Suche in einer Liste von
Audiotiteln beziehen. In dem Fall, dass ein gegebener Benutzer
35 mit dem Schlüsselwort „Suchen“ in der großen Mehrzahl der Fälle die Suche in einer Liste von Audiotiteln initiieren möchte und nicht die Suche in einem Adressbuch, kann dem Schlüsselwort

„Suchen“ in dem semantischen Modell für diesen Benutzer beispielsweise ein erstes Attribut „Suche in Adressbuch“ und ein zweites Attribut „Suche in Titelliste“ zugeordnet sein, wobei das zweite Attribut für diesen Benutzer z. B. ein größeres statistisches Gewicht hat als das erste Attribut.

Enthält eine Äußerung mehrere Schlüsselworte, von denen jedes Schlüsselwort einem anderen Steuerbefehl zugeordnet ist, können anhand der den Schlüsselworten zugeordneten statistischen Gewichte verschiedene Wahrscheinlichkeiten für die Zuordnung der Äußerung zu den verschiedenen in Frage kommenden Steuerbefehlen ermittelt werden. Zum Beispiel kann einer gegebenen Äußerung derjenige Steuerbefehl zugeordnet werden, für den die ermittelte Wahrscheinlichkeit am größten ist. Es kann auch vorgesehen sein, dass eine endgültige Zuordnung nur dann vorgenommen wird, wenn die so ermittelte Wahrscheinlichkeit einen vorgegebenen Schwellwert überschreitet.

Kann eine endgültige Zuordnung der Äußerung zu einem bestimmten Steuerbefehl nicht vorgenommen werden, so kann der Benutzer aufgefordert werden, einen Steuerbefehl aus einer Auswahl von mehreren in Frage kommenden Steuerbefehlen auszuwählen. Dazu kann das System eine zusätzliche Ein- und Ausgabeeinheit umfassen, z. B. in Gestalt eines Sensorbildschirms (Touchscreen). Das benutzerspezifische Sprachmodell kann dann anhand der Interaktion mit dem Benutzer und unter Berücksichtigung von dessen Eingabe entsprechend angepasst oder ergänzt werden. Das Erstellen oder Ergänzen des benutzerspezifischen Sprachmodells kann also insbesondere eine Interaktion mit dem Benutzer umfassen. Verwendet der Benutzer beispielsweise erstmalig einen Begriff, der noch nicht Teil des Sprachmodells ist, so kann der Benutzer aufgefordert werden, eine Zuordnung dieses Begriffes zu einem bestimmten Steuerbefehl oder zu einer bestimmten Teilmenge von Steuerbefehlen vorzunehmen. Der neue Begriff kann dann dem US-SLM und/oder dem US-SEM hinzugefügt werden. Gegebenenfalls können dem Begriff im US-SEM dann Attribute und/oder statistische Gewichte zugeordnet werden. So kann das benutzerspezifische

Sprachmodell schrittweise erweitert und auf den Benutzer zugeschnitten werden.

- Bei einer speziellen Ausführungsform des vorgeschlagenen Verfahrens kann das Transkribieren der natürlichsprachlichen Äußerung auf der externen Datenverarbeitungseinheit und/oder auf einer mit der externen Datenverarbeitungseinheit über eine Datenverbindung verbindbaren weiteren externen Datenverarbeitungseinheit vorgenommen werden. Auf der externen Datenverarbeitungseinheit und/oder auf der weiteren externen Datenverarbeitungseinheit kann also eine Spracherkennungssoftware ausgeführt werden oder ausführbar sein. Bei der weiteren externen Datenverarbeitungseinheit kann es sich z. B. um einen internetbasierten Diktierservice handeln. In diesem Fall müssen in der (ersten) externen Datenverarbeitungseinheit also weniger Diktierressourcen zum Transkribieren vorgehalten werden, so dass die erste externe Datenverarbeitungseinheit kostengünstiger ausgeführt sein kann.
- Insbesondere kann es vorgesehen sein, dass das Transkribieren der natürlichsprachlichen Äußerung und/oder das Erstellen oder das Ergänzen des Sprachmodells mittels eines auf der externen Datenverarbeitungseinheit oder mittels eines auf der weiteren Datenverarbeitungseinheit gespeicherten weiteren statistischen Sprachmodells und/oder weiteren semantischen Sprachmodells und/oder eines neuronalen Netzes vorgenommen wird. Bei dem weiteren statistischen Sprachmodell und dem weiteren semantischen Sprachmodell handelt es sich dabei normalerweise um nicht-benutzerspezifische Modelle, also um sprecherunabhängige Modelle. Da sie auf der externen Datenverarbeitungseinheit bzw. auf der weiteren externen Datenverarbeitungseinheit implementiert sind, können diese Modelle umfangreich sein, ohne dass der lokale Datenspeicher und/oder der lokale Prozessor deswegen größer oder leistungstärker ausgebildet sein müssen. Damit kann die Wahrscheinlichkeit dafür erhöht werden, dass von dem Benutzer verwendete Begriffe oder Ausdrucksweisen auch dann transkribiert werden können, wenn diese Begriffe noch nicht Teil des weniger

umfangreichen benutzerspezifischen Sprachmodells sind, das in der externen Datenverarbeitungseinheit gespeichert ist. Eine zusätzliche Interaktion mit dem Benutzer, bei der der Benutzer wie oben beschrieben zur Eingabe über die weitere Eingabeeinheit
5 aufgefördert wird, kann damit weitgehend vermieden werden.

Eine weitere spezielle Ausführungsform des Verfahrens zeichnet sich dadurch aus, dass mit der natürlichsprachlichen Äußerung eine Benutzerkennung zur Identifikation des Benutzers an die
10 Datenverarbeitungseinheit übermittelt wird, wobei die Benutzerkennung

- durch den Benutzer eingegeben wird und/oder
- einen akustischen Sprachabdruck umfasst, der anhand der durch den Benutzer eingesprochenen natürlichsprachlichen Äußerung
15 bestimmt wird, und/oder
- beim Einstecken eines Schlüssels in eine Einstecköffnung anhand einer Form des Schlüssels bestimmt wird und/oder
- mittels Gesichtserkennung bestimmt wird und/oder
- von einem mobilen Kommunikationsgerät des Benutzers, ins-
20 besondere von einem Mobiltelefon, übermittelt wird.

Dies ermöglicht in vorteilhafter Weise eine Identifikation des Benutzers. So kann die von dem Benutzer eingesprochene natürlichsprachliche Äußerung in der externen Datenverarbeitungseinheit zum Erstellen oder Ergänzen desjenigen benutzerspezifischen Sprachmodells verwendet werden, das genau diesem Benutzer in der externen Datenverarbeitungseinheit zugeordnet ist. Dies ist besonders dann vorteilhaft, wenn in der externen Datenverarbeitungseinheit eine Vielzahl von benutzerspezifischen Sprachmodellen jeweils unterschiedlicher Benutzer gespeichert sind, erstellt oder ergänzt werden. Dies ist auch vorteilhaft für den oben geschilderten Fall, in dem verschiedene Benutzer abwechselnd dasselbe Endgerät verwenden, z. B. wenn sie abwechselnd dasselbe Fahrzeug benutzen, in dem das Endgerät
35 angeordnet oder eingebaut ist.

Normalerweise ist es vorgesehen, dass die natürlichsprachliche Äußerung und/oder die Benutzerkennung über eine drahtlose Datenverbindung an die externe Datenverarbeitungseinheit übermittelt wird und/oder dass das benutzerspezifische Sprachmodell über eine drahtlose Datenverbindung von der Datenverarbeitungseinheit an das Endgerät übertragen wird. Vorzugsweise handelt es sich dabei um die bereits beschriebene drahtlose Internetverbindung oder sonstige drahtlose Telekommunikationsverbindung.

Ein Ausführungsbeispiel der Erfindung ist in den Zeichnungen dargestellt und wird anhand der folgenden Beschreibung näher erläutert. Es zeigen:

Fig. 1 schematisch ein erfindungsgemäßes System zum Erstellen oder Ergänzen eines benutzerspezifischen Sprachmodells in einem lokalen Datenspeicher, wobei das Sprachmodell eingerichtet ist, natürlichsprachlichen Äußerungen eines Benutzers Steuerbefehle zum Steuern eines ersten Endgerätes und/oder eines zweiten Endgerätes zuzuordnen,

Fig. 2 schematisch Schritte eines Verfahrens zum Erstellen oder Ergänzen des benutzerspezifischen Sprachmodells in dem lokalen Datenspeicher aus Fig. 1 sowie

Fig. 3 schematisch weitere Schritte des in Fig. 2 dargestellten Verfahrens, wobei einer von dem Benutzer eingegebenen natürlichsprachlichen Äußerung anhand des in dem lokalen Datenspeicher gespeicherten benutzerspezifischen Sprachmodells ein Steuerbefehl zum Steuern eines der Endgeräte zugeordnet wird, der Steuerbefehl an dieses Endgerät übermittelt wird und dieses Endgerät eine durch den Steuerbefehl verursachte Aktion durchführt.

Fig. 1 zeigt ein erfindungsgemäßes System 100. Bei dem gezeigten Ausführungsbeispiel umfasst das System 100 ein Kraftfahrzeug 1, eine erste externe Datenverarbeitungseinheit 2 und eine zweite externe Datenverarbeitungseinheit 3. In einem Innenraum des Kfz 1 sind ein erstes Endgerät 4 und ein zweites Endgerät 5 angeordnet. Bei dem ersten Endgerät 4 handelt es sich um ein Audiogerät mit einer Vielzahl von darin gespeicherten Audio-dateien, die mittels des Audiogerätes abspielbar sind. Jeder Audiodatei ist z. B. ein Titel, ein Interpret, ein Aufnahmejahr und eine Dateigröße zugeordnet. Bei dem zweiten Endgerät 5 handelt es sich um einen Rückspiegel im Innenraum des Kfz 1, der z. B. mittels eines elektrischen Stellantriebes um eine erste und eine zweite Achse schwenkbar ist. Eine Stellung des zweiten Endgerätes 5 ist beispielsweise durch einen ersten Drehwinkel bezüglich der ersten Achse und durch einen zweiten Drehwinkel bezüglich der zweiten Achse gegeben.

Die Endgeräte 4 und 5 sind über Steuerbefehle steuerbar. Steuerbefehle, die das erste Endgerät 4 betreffen, sind beispielsweise ein Befehl zum Ein- und Ausschalten des ersten Endgerätes 4, zum Anzeigen aller Titel eines bestimmten Interpreten, zum Abspielen eines bestimmten Titels, zum Springen innerhalb eines gegebenen Titels, zum Verstellen einer Lautstärke oder zum Ordnen der Audiodateien anhand vorgegebener Kriterien wie Titel, Interpret, Aufnahmejahr, Dateigröße etc. Steuerbefehle, die das zweite Endgerät 5 betreffen, sind Befehle zum Verstellen des Spiegels, wobei diese Befehle z. B. eine der beiden Achsen als Verstellachse, einen Verstellwinkel und eine Verstellrichtung umfassen können.

Das System 100 umfasst ferner einen lokalen Prozessor 6, einen lokalen Datenspeicher 7 zum Speichern wenigstens eines benutzerspezifischen Sprachmodells 8, das hier ein benutzerspezifisches statistisches Sprachmodell (US-SLM) 9 und ein benutzerspezifisches semantisches Sprachmodell (US-SEM) 10 umfasst, eine erste Eingabeeinheit 11, eine zweite Eingabeeinheit 12 sowie eine lokale Sende- und Empfangseinheit 13. Der lokale Prozessor

6, der lokale Datenspeicher 7, die Eingabeeinheiten 8 und 9 sowie die lokale Sende- und Empfangseinheit 10 sind ebenfalls im Innenraum des Kfz 1 angeordnet. Über Datenleitungen 14 sind die Eingabeeinheiten 11 und 12, die Endgeräte 4 und 5, der lokale
5 Datenspeicher 7 und die lokale Sende- und Empfangseinheit 13 jeweils mit dem lokalen Prozessor 6 verbunden.

Die erste Eingabeeinheit 11 umfasst wenigstens eine akustische Eingabeeinheit, z. B. in Form eines Mikrophons, über das ein
10 Benutzer natürlichsprachliche Äußerungen eingeben kann. Diese können dann in der ersten Eingabeeinheit 11 selbst oder durch den lokalen Prozessor 6 in eine digitale Audiodatei umgewandelt werden. Die zweite Eingabeeinheit 12 ist hier als Ein- und
Ausgabeeinheit ausgebildet und umfasst wenigstens einen Sen-
15 sorbildschirm (Touchscreen), über den der Benutzer mit dem System 100 interagieren kann.

Die erste externe Datenverarbeitungseinheit 2 umfasst wenigstens eine erste externe Sende- und Empfangseinheit 15, wenigstens
20 einen ersten externen Prozessor 16 und wenigstens einen ersten externen Datenspeicher 17 zum dauerhaften oder wenigstens zeitweisen Speichern des benutzerspezifischen Sprachmodells 8 mit dem US-SLM 9 und dem US-SEM 10. Die erste externe Sende- und
Empfangseinheit 15 ist über eine erste drahtlose Internetver-
25 bindung 18 mit der lokalen Sende- und Empfangseinheit 13 verbunden. Bei der ersten externen Datenverarbeitungseinheit 2 handelt es sich z. B. um einen Server eines Anbieters von
Spracherkennungssoftware.

30 Die zweite externe Datenverarbeitungseinheit 3 umfasst wenigstens eine zweite externe Sende- und Empfangseinheit 19, wenigstens einen zweiten externen Prozessor 20 und wenigstens einen zweiten externen Datenspeicher 21. In dem Datenspeicher 21 sind z. B. ein nicht-benutzerspezifisches, d. h. sprecheru-
35 nabhängiges weiteres statistisches Sprachmodell 22 und ein ebenfalls nicht-benutzerspezifisches weiteres semantisches Sprachmodell 23 gespeichert. Die zweite externe Sende- und

Empfangseinheit 19 ist über eine zweite drahtlose Internet-
verbindung 24 mit der ersten externen Sende- und Empfangseinheit
15 verbunden. Bei der zweiten externen Datenverarbeitungseinheit
3 handelt es sich z. B. um einen Server eines internetbasierten
5 Diktierdienstes, der Audiodateien in Textdateien transkribieren
kann.

Fig. 2 zeigt Schritte eines erfindungsgemäßen Verfahrens zum
Erstellen oder Ergänzen des benutzerspezifischen Sprachmodells
10 8 in dem lokalen Datenspeicher 7. Hier und im Folgenden sind
wiederkehrenden Merkmale jeweils mit identischen Bezugszeichen
versehen. Das benutzerspezifische Sprachmodell 8 umfasst di-
gitale Daten und/oder ein auf dem lokalen Prozessor 6 und auf dem
ersten externen Prozessor 16 ausführbares Computerprogramm. Das
15 Sprachmodell 8 ist eingerichtet, den durch den Benutzer über die
erste Eingabeeinheit 11 eingesprochenen natürlichsprachlichen
Äußerungen einen oder mehrere einer Vielzahl von Steuerbefehlen
zum Steuern des ersten Endgerätes 4 und/oder des zweiten End-
gerätes 5 zuzuordnen.

20 In einem ersten Schritt 25 gibt der Benutzer über den Sensor-
bildschirm der zweiten Eingabeeinheit 12 eine in digitaler Form
speicherbare Benutzerkennung ein. Dies kann z. B. ein Benut-
zernamen in Verbindung mit einem Passwort sein, durch welche der
25 Benutzer eindeutig identifizierbar ist. Damit wird dem System 100
mitgeteilt, dass im Weiteren das benutzerspezifische Sprach-
modell 8 für genau diesen Benutzer erstellt oder ergänzt werden
soll. Bei abgewandelten Ausführungsformen kann die Benutzer-
kennung auch durch einen akustischen Sprachabdruck gegeben sein,
30 der z. B. anhand eines Stimmenspektrums der durch den Benutzer
eingesprochenen natürlichsprachlichen Äußerung ermittelt wird.
Denkbar ist es ebenso, dass im Innenraum des Fahrzeugs eine Kamera
angeordnet ist, die ein Bild des Gesichts des Benutzers aufnimmt,
wobei der Benutzer mittels einer Gesichtserkennungssoftware
35 identifiziert wird. Die Benutzerkennung kann auch von einem
Mobiltelefon des Benutzers an das System 100 übermittelt werden.
Der Benutzer kann auch anhand eines individuellen mechanischen

Schlüssels identifiziert werden, den er zum Zwecke der Identifizierung in eine dafür vorgesehene Öffnung einsteckt.

- 5 Im nächsten Schritt 26 spricht der Benutzer eine natürlichsprachliche Äußerung in das Mikrophon der ersten Eingabeeinheit 11 ein. Mit der Eingabe der natürlichsprachlichen Äußerung beabsichtigt der Benutzer, eine bestimmte Aktion oder Funktion eines der Endgeräte 4 oder 5 auszulösen. Im vorliegenden Beispiel spricht der Benutzer den Satz „Ich möchte jetzt X von Y hören.“
- 10 Damit soll das erste Endgerät 4 veranlasst werden, die im Endgerät 4 gespeicherte Audiodatei abzuspielen, der der Titel X und der Interpret Y zugeordnet sind. Hier sind X und Y z. B. Einwortfolgen, bestehen also jeweils nur aus einem Wort.
- 15 Im Schritt 27 wird die eingesprochene Äußerung des Benutzers in eine digitale Audiodatei umgewandelt. Diese wird zusammen mit der Benutzerkennung von der lokalen Sende- und Empfangseinheit 13 über die drahtlose Internetverbindung 18 an die erste externe Sende- und Empfangseinheit 15 der ersten externen Datenverarbeitungseinheit 2 übermittelt. Dort wird die Audiodatei mit der natürlichsprachlichen Äußerung anhand der Benutzerkennung dem in dem ersten externen Datenspeicher 17 gespeicherten benutzerspezifischen Sprachmodell 8 zugeordnet.
- 20
- 25 Im Schritt 28 wird die Audiodatei mit der natürlichsprachlichen Äußerung „Ich möchte jetzt X von Y hören“ über die zweite drahtlose Internetverbindung 24 an die zweite externe Datenverarbeitungseinheit 3 übermittelt.
- 30 Im Schritt 29 wird sie dort mittels des zweiten externen Prozessors 20 unter Verwendung der im Datenspeicher 21 gespeicherten nicht-benutzerspezifischen Sprachmodelle SLM 22 und SEM 23 transkribiert. Dies bedeutet hier, dass die zweite externe Datenverarbeitungseinheit 3 z. B. eine digitale Textdatei erstellt, in der die genannte natürlichsprachliche Äußerung des Benutzers in Form einer Folge von Worten gespeichert ist, die durch eine Folge von Zeichen gegeben ist. Zum Beispiel umfasst
- 35

die transkribierte natürlichsprachliche Äußerung eine Folge von Buchstaben des lateinischen Alphabets, Satzzeichen und Leerzeichen.

- 5 Im Schritt 30 wird die so transkribierte natürlichsprachliche Äußerung dann wiederum in digitaler Form von der zweiten externen Datenverarbeitungseinheit 3 an die erste externe Datenverarbeitungseinheit 2 übermittelt.
- 10 Bei alternativen Ausführungsformen ist es auch denkbar, dass das im Schritt 29 beschriebene Transkribieren unmittelbar mittels der ersten externen Datenverarbeitungseinheit 2 vorgenommen wird. Das SLM 22 und das SEM 23 können also auch in der ersten externen Datenverarbeitungseinheit 2 gespeichert sein. In diesem
- 15 Fall müssen die Schritte 28 und 30 nicht durchgeführt werden.

Im Schritt 31 wird das benutzerspezifische Sprachmodell 8 unter Verwendung der transkribierten natürlichsprachlichen Äußerung zunächst im ersten externen Datenspeicher 17 der ersten externen

20 Datenverarbeitungseinheit 2 ergänzt. Sofern das Sprachmodell 8 noch nicht besteht, wird es im Schritt 31 erstmalig erstellt. Der Schritt 31 umfasst Schritte 31a und 31b. Im Schritt 31a wird dabei das US-SLM 9 im Datenspeicher 17 ergänzt oder erstellt. Im Schritt 31b wird das US-SEM im Datenspeicher 17 ergänzt oder erstellt.

25 Im vorliegenden Fall umfasst das US-SLM 9 des Sprachmodells 8 im Datenspeicher 17 schon eine Vielzahl von transkribierten Wortfolgen, z. B. Einwortfolgen, Zweiwortfolgen, Dreiwortfolgen oder längere Wortfolgen, die derselbe Benutzer in vorherigen

30 Sitzungen eingegeben bzw. ausgesprochen hat. Im Schritt 31a wird das US-SLM 9 nur insoweit durch die transkribierte natürlichsprachliche Äußerung „Ich möchte jetzt X von Y hören“ ergänzt, als das US-SLM 9 diese noch nicht umfasst. Beispielsweise umfasst das US-SLM 9 die gesamte Wortfolge „Ich möchte jetzt X von Y hören“

35 noch nicht, so dass diese dem US-SLM 9 im Datenspeicher 17 hinzugefügt wird. Ergänzt werden aber beispielsweise nur solche Einwortfolgen, die noch nicht von dem US-SLM 9 im Datenspeicher

17 umfasst sind. Hier sind dies beispielsweise die Einwortfolgen „X“ und „Y“, die dem US-SLM 9 im Datenspeicher 17 im Schritt 31a hinzugefügt werden. Statistische Gewichte von solchen Worten oder Wortfolgen der Äußerung „Ich möchte jetzt X von Y hören“, die das US-SLM 9 im Datenspeicher 17 bereits umfasst, können anhand der neu eingesprochenen Äußerung entsprechend angepasst, also z. B. erhöht werden.

Ebenso umfasst das US-SEM 10 im Datenspeicher 17 schon eine Vielzahl von Schlüsselworten, denen jeweils Attribute und statistische Gewichte zugeordnet sind. Hier umfasst das US-SEM 10 im Datenspeicher 17 beispielsweise bereits das Schlüsselwort „hören.“ Diesem sind im US-SEM 10 schon die Attribute „Audiogerät“ und „Abspielen“ mit einem bestimmten statistischen Gewicht zugeordnet. Dies bedeutet z. B., dass natürlichsprachlichen Äußerungen, die das Schlüsselwort „hören“ beinhalten, mit einer von Null verschiedenen Wahrscheinlichkeit diejenige Untermenge aller möglichen Steuerbefehle zugeordnet wird, die das erste Endgerät 4, also das Audiogerät, betreffen.

Im vorliegenden Beispiel sind der Titel „X“ und der Interpret „Y“ bisher weder Teil des US-SLM 9 noch des US-SEM 10 im Datenspeicher 17. Wegen des gleichzeitigen Auftretens des dem US-SLM 9 und dem US-SEM 10 bereits bekannten Schlüsselwortes „hören“ werden das Wort „X“ und das Wort „Y“ im Schritt 31b jedoch dem US-SEM 10 im Datenspeicher 17 hinzugefügt und erhalten dort ebenfalls die dem bekannten Schlüsselwort „hören“ zugeordneten Attribute „Audiogerät“ und „Abspielen.“ Dabei wird dem Attribut „Audiogerät“ jedoch gleichzeitig ein größeres statistisches Gewicht zugeordnet als dem Attribut „Abspielen.“ So ist mit großer Wahrscheinlichkeit jeder Äußerung, die den Titel „X“ enthält, ein Befehl zuzuordnen, der das Audiogerät betrifft. Jedoch wird nicht jede Äußerung, die den Titel „X“ enthält, auf das Abspielen des Titels „X“ gerichtet sein. Denkbar ist z. B., dass der Benutzer lediglich das Anzeigen des dem Titel „X“ zugeordneten Aufnahmehandbuches wünscht.

Da es sich bei den Worten „X“ und „Y“ um neu in das US-SEM 10 im
Datenspeicher 17 aufgenommene Schlüsselworte handelt, wird der
Benutzer über den Sensorbildschirm der zweiten Eingabeeinheit 12
aufgefordert, die Aufnahme der Worte „X“ und „Y“ als Schlüs-
5 selworte in das US-SEM 10 im Datenspeicher 17 und die Zuordnung
der Attribute „Audiogerät“ und „Abspielen“ zu den neuen
Schlüsselworten „X“ und „Y“ zu bestätigen. Zusätzlich kann dem
Benutzer die Möglichkeit gegeben werden, den neuen Schlüssel-
worten „X“ und „Y“ weitere Attribute hinzuzufügen. Hier ordnet
10 er dem neuen Schlüsselwort „X“ beispielsweise manuell das weitere
Attribut „Titel“ und dem neuen Schlüsselwort „Y“ das weitere
Attribut „Interpret“ zu.

Auf die hier beschriebene Weise können den Schlüsselworten
15 und/oder Wortfolgen von Schlüsselworten des US-SEM 10 im Da-
tenspeicher 17 also statistische Gewichte und Attribute zu-
geordnet werden, anhand derer mit einer gegebenen Wahr-
scheinlichkeit eine Zuordnung des jeweiligen Schlüsselwortes
oder der jeweiligen Wortfolge von Schlüsselworten zu einem
20 bestimmten Steuerbefehl oder wenigstens zu einer Teilmenge aus
der Menge aller möglichen Steuerbefehle zum Steuern der Endgeräte
4 und 5 herstellbar ist. Obwohl hier lediglich das Ergänzen der
im Datenspeicher 17 bereits vorhandenen Modelle US-SLM 9 und
US-SLM 10 beschrieben ist, wird es sich dem Fachmann unmittelbar
25 erschließen, wie das Verfahren umgesetzt werden kann, wenn das
Sprachmodell 8 anhand einer eingesprochenen natürlichsprach-
lichen Äußerung ganz von Neuem erstellt wird, z. B. bei einer
allerersten Sitzung des Benutzers.

30 Im Schritt 32 wird das benutzerspezifische Sprachmodell 8 mit dem
US-SLM 9 und dem US-SLM 10 vollständig oder wenigstens teilweise
von der ersten externen Datenverarbeitungseinheit 2 über die
drahtlose Internetverbindung 18 an den lokalen Datenspeicher 7
übertragen und in diesem gespeichert. Vorzugsweise handelt es
35 sich dabei um eine Aktualisierung einer bereits im lokalen
Datenspeicher 7 gespeicherten vorherigen Version des Sprach-
modells 8. Es werden also vorzugsweise nur Änderungen und/oder

Ergänzungen der vorherigen Version übertragen. Eine solche Aktualisierung kann in regelmäßigen Abständen automatisch oder z. B. auf eine entsprechende Eingabe des Benutzers hin vorgenommen werden. Diese kann der Benutzer z. B. über die zweite
5 Eingabeeinheit 12 vornehmen. Es ist auch denkbar, dass die Aktualisierung des Sprachmodells 8 auch dem lokalen Datenspeicher 7 im Anschluss an jede Sitzung des Benutzers durchgeführt wird, also jedes Mal, nachdem der Benutzer eine natürlichsprachliche Äußerung ausgesprochen hat.

10

In Fig. 3 sind weitere Verfahrensschritte gezeigt, die zeitgleich mit den in Fig. 2 dargestellten Verfahrensschritten durchführbar sind. Diese weiteren Verfahrensschritte beschreiben, wie der
eingesprochenen natürlichsprachlichen Äußerung anhand des bereits im lokalen Datenspeichers 7 gespeicherten Modells 8 ein
15 Steuerbefehl zugeordnet wird und gemäß diesem Steuerbefehl eine Aktion in einem der Endgeräte 4 und 5 ausgelöst wird. Das in Fig. 2 beschriebene Ergänzen des Modells 8 im lokalen Datenspeicher 7 und das Steuern der Endgeräte 4 und 5 anhand der gesprochenen
20 Äußerung und der bereits im Datenspeicher 7 gespeicherten Version des Modells 8 können also parallel durchgeführt werden.

Nach der Eingabe der Benutzerkennung und der natürlich sprachlichen Äußerung in den Schritten 25 und 26 wird in einem
25 Schritt 33 das Transkribieren der natürlichsprachlichen Äußerung durch den lokalen Prozessor 6 unter Verwendung des im lokalen Datenspeicher 7 gespeicherten US-SLM 9 durchgeführt. Im Schritt 34 wird der transkribierten Äußerung dann unter Verwendung des im lokalen Datenspeicher 7 gespeicherten US-SEM 10 ein Steuerbefehl zugeordnet. Im Schritt 35 wird dieser Steuerbefehl dann
30 an das entsprechende Endgerät gesendet. Im Schritt 36 führt dieses Endgerät dann eine durch diesen Steuerbefehl veranlasste Aktion aus. Diese Aktion kann z. B. das Abspielen einer Audiodatei durch das erste Endgerät 4 oder das Verstellen des Endgerätes 5
35 umfassen.

In dem Fall, dass das Transkribieren im Schritt 33 unter Ver-

wendung des im lokalen Datenspeicher 7 gespeicherten US-SLM 9 oder die Zuordnung im Schritt 34 unter Verwendung des im lokalen Datenspeicher 7 gespeicherten US-SEM 10 nicht durchführbar sind, weil die im lokalen Datenspeicher 7 gespeicherten Versionen des
5 US-SLM 9 oder des US-SEM 10 noch nicht mächtig genug sind, um die aktuell ausgesprochene Äußerung des Benutzers zu verarbeiten, werden zunächst die in Fig. 2 gezeigten Verfahrensschritte durchgeführt, was in Fig. 3 in Gestalt der Verfahrensschritte 37 und 38 gezeigt ist. In diesem Fall erfolgt also zunächst eine
10 Aktualisierung des im lokalen Datenspeicher 7 gespeicherten Sprachmodells 8.

In dem Fall, dass die drahtlose Internetverbindung 18 zwischen der lokalen Sende- und Empfangseinheit 13 und der ersten externen
15 Sende- und Empfangseinheit 15 nicht herstellbar ist, weil das Fahrzeug 1 z. B. durch einen Tunnel fährt oder sich in einer Tiefgarage befindet, werden nur die Verfahrensschritte 25, 26, 33, 34, 35 und 36 durchgeführt.

Patentansprüche

1. Verfahren zum Erstellen oder Ergänzen eines benutzer-
spezifischen Sprachmodells (8) in einem mit einem Endgerät
5 (4, 5) verbindbaren lokalen Datenspeicher (7), wobei das
Sprachmodell (8) eingerichtet ist, natürlichsprachlichen
Äußerungen eines Benutzers Steuerbefehle zum Steuern des
Endgerätes (4, 5) zuzuordnen, umfassend die Schritte:
- Einsprechen einer natürlichsprachlichen Äußerung durch
10 den Benutzer;
 - Übermitteln der natürlichsprachlichen Äußerung an eine
externe Datenverarbeitungseinheit (2) und Transkribieren
der natürlichsprachlichen Äußerung;
 - Erstellen oder Ergänzen des Sprachmodells (8) auf der
15 externen Datenverarbeitungseinheit (2) anhand der
transkribierten natürlichsprachlichen Äußerung;
 - vollständiges oder wenigstens teilweises Übertragen des
Sprachmodells (8) von der externen Datenverarbeitungs-
einheit (2) an den lokalen Datenspeicher (7) sowie
20 Speichern und/oder Ergänzen des Sprachmodells (8) in dem
lokalen Datenspeicher (7).
2. Verfahren nach Anspruch 1, d a d u r c h g e k e n n -
z e i c h n e t , dass das Erstellen oder das Ergänzen des
25 Sprachmodells (8) auf der externen Datenverarbeitungs-
einheit (2) ein Erstellen oder ein Ergänzen eines be-
nutzerspezifischen
- statistischen Sprachmodells (US-SLM) (9) umfasst, wobei
das US-SLM (9) aus der transkribierten natürlichsprach-
30 lichen Äußerung gebildet wird oder um diese ergänzt wird
und wobei die transkribierte natürlichsprachliche Äu-
ßerung transkribierte Worte und/oder transkribierte
Laute und/oder transkribierte Wortfolgen und/oder
transkribierte Lautfolgen umfasst, und/oder
 - 35 - semantischen Sprachmodells (US-SEM) (10) umfasst, wobei
in der transkribierten natürlichsprachlichen Äußerung
Schlüsselworte und/oder Wortstämme von Schlüsselworten

und/oder Wortfolgen identifiziert werden und das US-SEM aus den Schlüsselworten und/oder den Wortstämmen und/oder den Wortfolgen gebildet wird oder um diese ergänzt wird.

- 5 3. Verfahren nach Anspruch 1 oder 2, d a d u r c h g e -
k e n n z e i c h n e t , dass das Erstellen oder das Er-
gänzen des US-SEM (10) umfasst, dass den Schlüsselworten
und/oder den Wortstämmen und/oder den Wortfolgen sta-
tistische Gewichte und/oder Attribute zugeordnet werden,
10 wobei mittels der statistischen Gewichte und/oder mittels
der Attribute eine Zuordnung des jeweiligen Schlüsselworts
oder Wortstamms oder der jeweiligen Wortfolge zu einer
Teilmenge aus einer Menge von Steuerbefehlen herstellbar
ist.
- 15 4. Verfahren nach einem der vorhergehenden Ansprüche,
d a d u r c h g e k e n n z e i c h n e t , dass der ein-
gesprochenen natürlichsprachlichen Äußerung wenigstens ein
Steuerbefehl aus einer Menge von Steuerbefehlen zugeordnet
20 wird, wobei die Zuordnung des Steuerbefehls zur einge-
sprochenen natürlichsprachlichen Äußerung mittels des im
lokalen Datenspeicher (7) gespeicherten benutzerspezi-
fischen Sprachmodells (8) vorgenommen wird und/oder wobei
die Zuordnung des Steuerbefehls zur eingesprochenen na-
25 türlichsprachlichen Äußerung mittels der externen Daten-
verarbeitungseinheit (2) vorgenommen wird und der Steu-
erbefehl an das Endgerät (4, 5) übermittelt wird.
- 30 5. Verfahren nach den Ansprüchen 3 und 4, d a d u r c h
g e k e n n z e i c h n e t , dass die Zuordnung des
Steuerbefehls zur eingesprochenen natürlichsprachlichen
Äußerung in Abhängigkeit von den in der natürlich-
sprachlichen Äußerung identifizierten Schlüsselworten
und/oder Wortstämmen und/oder Wortfolgen und/oder in
35 Abhängigkeit von den diesen Schlüsselworten und/oder
Wortstämmen und/oder Wortfolgen im US-SEM (10) jeweils

zugeordneten statistischen Gewichten und/oder Attributen vorgenommen wird.

6. Verfahren nach einem der vorhergehenden Ansprüche,
5 d a d u r c h g e k e n n z e i c h n e t , dass das
Transkribieren der natürlichsprachlichen Äußerung auf der
externen Datenverarbeitungseinheit (2) und/oder auf einer
mit der externen Datenverarbeitungseinheit (2) über eine
Datenverbindung (24) verbindbaren weiteren externen Da-
10 tenverarbeitungseinheit (3) vorgenommen wird.
7. Verfahren nach Anspruch 6, d a d u r c h g e k e n n -
z e i c h n e t , dass das Transkribieren der natürlich-
sprachlichen Äußerung und/oder das Erstellen oder das
15 Ergänzen des Sprachmodells (8) mittels eines auf der
externen Datenverarbeitungseinheit (2) oder mittels eines
auf einer weiteren externen Datenverarbeitungseinheit (3)
gespeicherten
- weiteren statistischen Sprachmodells (22) und/oder
20 - weiteren semantischen Sprachmodells (23) und/oder
- neuronalen Netzes
vorgenommen wird.
8. Verfahren nach einem der vorhergehenden Ansprüche,
25 d a d u r c h g e k e n n z e i c h n e t , dass mit der
natürlichsprachlichen Äußerung eine Benutzerkennung zur
Identifikation des Benutzers an die externe Datenverar-
beitungseinheit (2) übermittelt wird, wobei die Benut-
zerkennung
30 - durch den Benutzer eingegeben wird und/oder
- einen akustischen Sprachabdruck umfasst, der anhand der
durch den Benutzer eingesprochenen natürlichsprachlichen
Äußerung bestimmt wird, und/oder
- beim Einstecken eines Schlüssels in eine Einstecköffnung
35 anhand einer Form des Schlüssels bestimmt wird und/oder
- mittels Gesichtserkennung bestimmt wird und/oder
- von einem mobilen Kommunikationsgerät des Benutzers,

insbesondere von einem Mobiltelefon, übermittelt wird.

9. Verfahren nach einem der vorhergehenden Ansprüche,
dadurch gekennzeichnet, dass die natürlichsprachliche Äußerung und/oder die Benutzerkennung
5 über eine drahtlose Internetverbindung (18) an die externe Datenverarbeitungseinheit (2) übermittelt wird und/oder dass das benutzerspezifische Sprachmodell (8) über eine drahtlose Internetverbindung (18) von der externen Datenverarbeitungseinheit (2) an das Endgerät (4, 5) übertragen wird.
10. System (100) zum Erstellen oder Ergänzen eines benutzerspezifischen Sprachmodells (8) in einem mit einem
15 Endgerät (4, 5) verbindbaren lokalen Datenspeicher (7), wobei das Sprachmodell (8) eingerichtet ist, natürlichsprachlichen Äußerungen eines Benutzers Steuerbefehle zum Steuern des Endgerätes (4, 5) zuzuordnen, umfassend:
- wenigstens ein Endgerät (4, 5);
 - 20 - einen mit dem Endgerät (4, 5) verbindbaren lokalen Datenspeicher (7) zum Speichern des Sprachmodells (8);
 - eine akustische Eingabeeinheit (11), über die natürlichsprachliche Äußerungen zur Steuerung des Endgerätes (4, 5) eingebbar sind;
 - 25 - eine externe Datenverarbeitungseinheit (2), an die die natürlichsprachlichen Äußerungen übermittelbar sind, wobei die externe Datenverarbeitungseinheit (2) eingerichtet ist, anhand der natürlichsprachlichen Äußerungen das benutzerspezifische Sprachmodell (8) zu
 - 30 erstellen oder zu ergänzen und das Sprachmodell (8) vollständig oder wenigstens teilweise zum Speichern in dem lokalen Datenspeicher (7) an den lokalen Datenspeicher (7) zu übertragen; und
 - eine mit der akustischen Eingabeeinheit (11) und mit dem lokalen Datenspeicher (7) verbindbare lokale Sende- und
 - 35 Empfangseinheit (13) zum Übermitteln der natürlich-

sprachlichen Äußerung an die externe Datenverarbeitungseinheit (2) und zum Empfangen des Sprachmodells (8) von der externen Datenverarbeitungseinheit (2).

- 5 11. System nach Anspruch 10, d a d u r c h g e k e n n -
z e i c h n e t , dass das Endgerät (4, 5)
- ein Audiogerät zum Abspielen von Audiodateien und/oder
 - ein Videogerät zum Abspielen von Bild- und/oder Video-
dateien und/oder
 - 10 - ein Telekommunikationsgerät zum Empfangen und/oder
Versenden von Text-, Sprach-, Bild- oder Videonachrichten
und/oder
 - ein Navigationsgerät und/oder
 - einen Stellantrieb zum Verstellen der Position einer
15 Fensterscheibe, eines Spiegels, eines Sitzes oder einer
Luftzuführvorrichtung und/oder
 - eine Anlage zur Regelung einer Temperatur und/oder einer
Luftfeuchtigkeit
- 20 umfasst, wobei das Endgerät vorzugsweise in einem Fahrzeug
(1) angeordnet oder eingebaut ist, insbesondere in einem PKW
oder LKW.

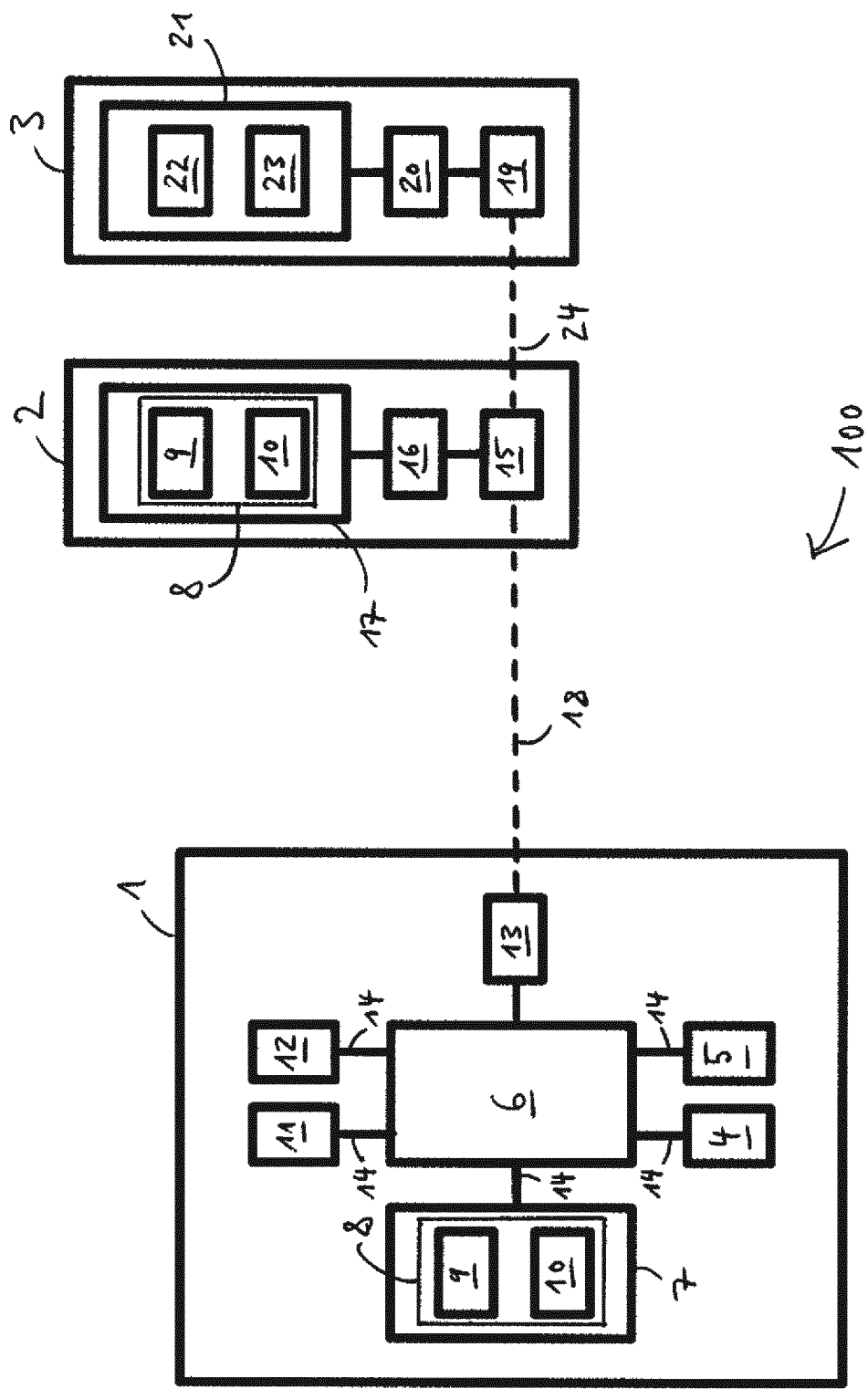


Fig. 1

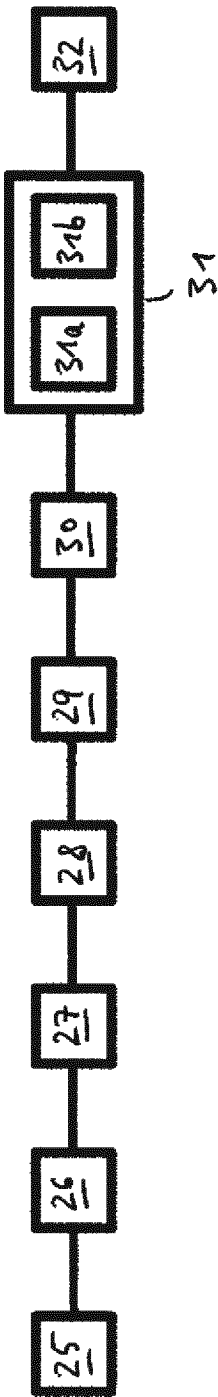


Fig. 2

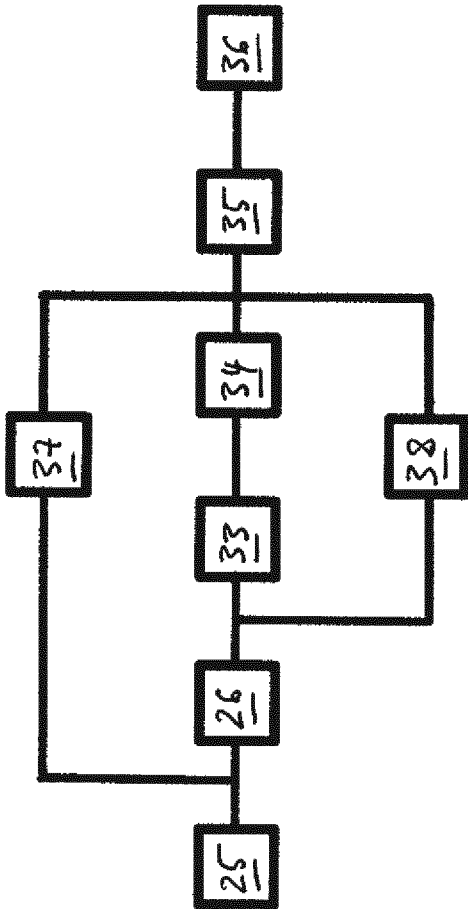


Fig. 3

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2014/070164

A. CLASSIFICATION OF SUBJECT MATTER
INV. G10L15/30 G10L15/18
ADD. B60R16/037

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G10L G01C B60R

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, WPI Data, COMPENDEX, INSPEC, IBM-TDB

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2005/102142 A1 (SOUFFLET FREDERIC [FR] ET AL) 12 May 2005 (2005-05-12)	1,2,4-11
Y	paragraphs [0005], [0009] - [0014], [0018], [0035], [0040] - [0044], [0069] - [0097], [0108], [0131], [0167], [0179]	3
X	----- US 2007/233487 A1 (COHEN MICHAEL H [US] ET AL) 4 October 2007 (2007-10-04) paragraphs [0013], [0014], [0025] - [0027], [0032], [0039], [0040], [0060], [0062] - [0065] figures 1,2,4	1,10
X	----- US 2012/130709 A1 (BOCCHIERI ENRICO [US] ET AL) 24 May 2012 (2012-05-24) paragraphs [0011], [0030], [0031] ----- -/--	1,10

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

4 November 2014

Date of mailing of the international search report

11/11/2014

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Geißler, Christian

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2014/070164

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2007/276651 A1 (BLISS HARRY M [US] ET AL) 29 November 2007 (2007-11-29) paragraph [0034]	1, 10
Y	----- US 2007/156682 A1 (HOWELL DAVID A [US] ET AL) 5 July 2007 (2007-07-05) paragraphs [0012], [0014], [0018] -----	3

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2014/070164

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2005102142 A1	12-05-2005	CN 1491412 A	21-04-2004
		DE 60222093 T2	05-06-2008
		EP 1362343 A1	19-11-2003
		ES 2291440 T3	01-03-2008
		FR 2820872 A1	16-08-2002
		JP 4751569 B2	17-08-2011
		JP 2004530149 A	30-09-2004
		KR 20030076661 A	26-09-2003
		MX PA03007178 A	04-12-2003
		US 2005102142 A1	12-05-2005
		WO 02065454 A1	22-08-2002

US 2007233487 A1	04-10-2007	AT 524777 T	15-09-2011
		EP 2008189 A2	31-12-2008
		EP 2437181 A1	04-04-2012
		EP 2453436 A2	16-05-2012
		US 2007233487 A1	04-10-2007
		US 2011213613 A1	01-09-2011
		US 2013006640 A1	03-01-2013
		US 2013246065 A1	19-09-2013
		WO 2007118100 A2	18-10-2007

US 2012130709 A1	24-05-2012	NONE	

US 2007276651 A1	29-11-2007	CN 101454775 A	10-06-2009
		US 2007276651 A1	29-11-2007
		WO 2007140047 A2	06-12-2007

US 2007156682 A1	05-07-2007	NONE	

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen

PCT/EP2014/070164

A. KLASSIFIZIERUNG DES ANMELDUNGSGEGENSTANDES
INV. G10L15/30 G10L15/18
ADD. B60R16/037

Nach der Internationalen Patentklassifikation (IPC) oder nach der nationalen Klassifikation und der IPC

B. RECHERCHIERTE GEBIETE

Recherhierter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)
G10L G01C B60R

Recherhierte, aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherhierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)

EPO-Internal, WPI Data, COMPENDEX, INSPEC, IBM-TDB

C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X	US 2005/102142 A1 (SOUFFLET FREDERIC [FR] ET AL) 12. Mai 2005 (2005-05-12)	1,2,4-11
Y	Absätze [0005], [0009] - [0014], [0018], [0035], [0040] - [0044], [0069] - [0097], [0108], [0131], [0167], [0179] -----	3
X	US 2007/233487 A1 (COHEN MICHAEL H [US] ET AL) 4. Oktober 2007 (2007-10-04) Absätze [0013], [0014], [0025] - [0027], [0032], [0039], [0040], [0060], [0062] - [0065] Abbildungen 1,2,4 -----	1,10
X	US 2012/130709 A1 (BOCCHIERI ENRICO [US] ET AL) 24. Mai 2012 (2012-05-24) Absätze [0011], [0030], [0031] ----- -/--	1,10



Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen



Siehe Anhang Patentfamilie

* Besondere Kategorien von angegebenen Veröffentlichungen :

"A" Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist

"E" frühere Anmeldung oder Patent, die bzw. das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist

"L" Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)

"O" Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht

"P" Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist

"T" Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist

"X" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden

"Y" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist

"&" Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der internationalen Recherche

4. November 2014

Absenddatum des internationalen Recherchenberichts

11/11/2014

Name und Postanschrift der Internationalen Recherchenbehörde
Europäisches Patentamt, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Bevollmächtigter Bediensteter

Geißler, Christian

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen

PCT/EP2014/070164

C. (Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X	US 2007/276651 A1 (BLISS HARRY M [US] ET AL) 29. November 2007 (2007-11-29) Absatz [0034] -----	1, 10
Y	US 2007/156682 A1 (HOWELL DAVID A [US] ET AL) 5. Juli 2007 (2007-07-05) Absätze [0012], [0014], [0018] -----	3

INTERNATIONALER RECHERCHENBERICHT

Angaben zu Veröffentlichungen, die zur selben Patentfamilie gehören

Internationales Aktenzeichen

PCT/EP2014/070164

Im Recherchenbericht angeführtes Patentedokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
US 2005102142 A1	12-05-2005	CN 1491412 A	21-04-2004
		DE 60222093 T2	05-06-2008
		EP 1362343 A1	19-11-2003
		ES 2291440 T3	01-03-2008
		FR 2820872 A1	16-08-2002
		JP 4751569 B2	17-08-2011
		JP 2004530149 A	30-09-2004
		KR 20030076661 A	26-09-2003
		MX PA03007178 A	04-12-2003
		US 2005102142 A1	12-05-2005
		WO 02065454 A1	22-08-2002

US 2007233487 A1	04-10-2007	AT 524777 T	15-09-2011
		EP 2008189 A2	31-12-2008
		EP 2437181 A1	04-04-2012
		EP 2453436 A2	16-05-2012
		US 2007233487 A1	04-10-2007
		US 2011213613 A1	01-09-2011
		US 2013006640 A1	03-01-2013
		US 2013246065 A1	19-09-2013
		WO 2007118100 A2	18-10-2007

US 2012130709 A1	24-05-2012	KEINE	

US 2007276651 A1	29-11-2007	CN 101454775 A	10-06-2009
		US 2007276651 A1	29-11-2007
		WO 2007140047 A2	06-12-2007

US 2007156682 A1	05-07-2007	KEINE	
