



(19)  
Bundesrepublik Deutschland  
Deutsches Patent- und Markenamt

(10) **DE 698 39 068 T2** 2009.01.22

(12) **Übersetzung der europäischen Patentschrift**

(97) **EP 0 935 378 B1**

(51) Int Cl.<sup>8</sup>: **H04M 3/436** (2006.01)

(21) Deutsches Aktenzeichen: **698 39 068.7**

(96) Europäisches Aktenzeichen: **98 310 516.4**

(96) Europäischer Anmeldetag: **21.12.1998**

(97) Erstveröffentlichung durch das EPA: **11.08.1999**

(97) Veröffentlichungstag

der Patenterteilung beim EPA: **30.01.2008**

(47) Veröffentlichungstag im Patentblatt: **22.01.2009**

(30) Unionspriorität:

**8124                      16.01.1998      US**

(84) Benannte Vertragsstaaten:

**DE, FR, GB**

(73) Patentinhaber:

**International Business Machines Corp., Armonk,  
N.Y., US**

(72) Erfinder:

**Epstein, Mark Edward, Katonah, NY 10536, US;  
Kanevsky, Dimitri, Ossining, NY 10562, US; Maes,  
Stephan Herman, Danbury, CT06811, US**

(74) Vertreter:

**Duscher, R., Dipl.-Phys. Dr.rer.nat., Pat.-Anw.,  
70176 Stuttgart**

(54) Bezeichnung: **System und Verfahren zur automatischen Verarbeitung von Anruf und Datenübertragung**

Anmerkung: Innerhalb von neun Monaten nach der Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents kann jedermann beim Europäischen Patentamt gegen das erteilte europäische Patent Einspruch einlegen. Der Einspruch ist schriftlich einzureichen und zu begründen. Er gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist (Art. 99 (1) Europäisches Patentübereinkommen).

Die Übersetzung ist gemäß Artikel II § 3 Abs. 1 IntPatÜG 1991 vom Patentinhaber eingereicht worden. Sie wurde vom Deutschen Patent- und Markenamt inhaltlich nicht geprüft.

## Beschreibung

**[0001]** Die vorliegende Erfindung betrifft ein System und Verfahren zum automatischen Verarbeiten von Anrufen und Datenübertragungen und insbesondere ein System und Verfahren zum automatischen Verarbeiten von Anrufen und Datenübertragungen nach einer vorprogrammierten Prozedur auf der Grundlage der Identität eines Anrufers oder eines Verfassers, des Themas und Inhalts eines Anrufs oder einer Nachricht und/oder der Uhrzeit eines solchen Anrufs oder einer solchen Nachricht.

**[0002]** Früher wurden Anrufe im Allgemeinen manuell entweder von einem Geschäftsinhaber, einer Sekretärin oder einer örtlichen Telefonzentrale verarbeitet. Es gibt einige herkömmliche Einrichtungen zur teilweisen Ausführung bestimmter Telefonverarbeitungsfunktionen. Zum Beispiel zeichnen herkömmliche Anrufbeantworter und Voicemail-Dienste (Sprachpost) ankommende Telefonnachrichten auf, die vom Benutzer solcher Einrichtungen oder Dienste wiedergegeben werden. Außerdem bieten die Software von Telefonapparaten oder örtliche Nebenstellenanlagen (Private Branch Exchange, PBX) Vermittlungsfunktionen für Telefonnetze. Diese herkömmlichen Anrufbeantworter, Voicemail-Dienste und Vermittlungssysteme sind jedoch nicht in der Lage, automatisch bestimmte Verarbeitungsprozeduren auszuführen, die auf die Identität des Anrufers reagieren oder den Inhalt oder das Thema des Anrufs oder der Nachricht bewerten und einen solchen Anruf oder eine solche Nachricht dann entsprechend zu verarbeiten. Stattdessen muss der Benutzer zuerst seine Telefonanrufe manuell beantworten oder solche Anrufe vom Anrufbeantworter oder vom Voicemail-Dienst abrufen und dann für jeden einzelnen Anruf entscheiden, wie mit diesem zu verfahren ist. Die vorliegende Erfindung erübrigt oder erleichtert eine solche aufwändige manuelle Verarbeitung.

**[0003]** Obwohl Anrufbeantworter und Voicemail-Dienste durch die Tasteneingabe im Mehrfrequenzwahlverfahren (Dual Tone Multi-Frequency, DTMF) geschützt sind, können sie die Identität des Anrufers nicht prüfen, wenn der Anrufer von außen anruft oder sie von ihm durch eine gültige persönliche Kennnummer (Personal Identification Number, PIN) umprogrammiert werden, die durch Tasten nach dem Mehrfrequenzwahlverfahren eingegeben wird. Ferner erfordern Telefonkonferenzzentralen für den Zugriff DTMF-PINs, jedoch können sie die Identität des Sprechers während einer Telefonkonferenz nicht prüfen. Solche Anrufbeantworter, Voicemail-Dienste und Telefonkonferenzzentralen können deshalb von unberechtigten Personen abgehört werden, die Zugriff auf eine ansonsten gültige PIN erlangen können.

**[0004]** In der US-Patentschrift 5 327 486 werden ein Verfahren und ein System zur Verwaltung beispiels-

weise von Telefongesprächen beschrieben und kurz die Anwendung der Stimm- und Spracherkennung erwähnt. Diese Veröffentlichung beschreibt jedoch nicht die Vorteile und neue Lösungen zur Verwaltung eines Telefongesprächs unter Verwendung der Stimm- und Spracherkennung.

**[0005]** Deshalb besteht eine Aufgabe der vorliegenden Erfindung darin, ein System und Verfahren zur automatischen Verarbeitung von Anrufen und Datenübertragungen nach einer vorher programmierten Prozedur auf der Grundlage der Identität des Anrufers oder Verfassers, des Themas des Anrufs oder der Nachricht und/oder der Uhrzeit bereitzustellen.

**[0006]** Eine weitere Aufgabe der vorliegenden Erfindung besteht darin, ein Anrufverarbeitungssystem bereitzustellen, das zuerst durch ein Telefon empfangene Nachrichten, Fax-Mitteilungen und eMail-Nachrichten sowie andere durch das System elektronisch empfangene Daten umsetzen, dann die Identität des Anrufers (oder Teilnehmers einer Telefonkonferenz) oder des Urhebers solcher eMail-Nachrichten oder Fax-Mitteilungen kennzeichnen und dann solche Anrufe, Gespräche und Nachrichten nach ihrem Ursprung oder ihrem Thema indexieren kann, sodass ein berechtigter Benutzer dann entweder an Ort und Stelle oder von einem fernen Standort auf das System zugreifen kann, um solche Telefongespräche oder -nachrichten wiederzugeben oder solche eMail- oder Fax-Nachrichten in Form synthetisierter Sprache abzurufen.

**[0007]** Noch eine weitere Aufgabe der vorliegenden Erfindung besteht darin, ein System bereitzustellen, das auf sprachaktivierte Befehle von einem berechtigten Benutzer reagiert (d. h. den Zugriff erlaubt oder programmiert werden kann), wobei das System den Benutzer erkennen und prüfen kann, bevor dieser auf Anrufe oder Nachricht zugreifen oder das System programmieren kann.

**[0008]** Gemäß einem Aspekt der vorliegenden Erfindung wird ein automatisches System zur Verarbeitung von Anrufen und Datenübertragungen bereitgestellt, das Folgendes umfasst: ein Server-Mittel zum Empfangen eines ankommenden Anrufs; ein Spracherkennungsmodul, das funktionell mit dem Server-Mittel verbunden ist, zum Erkennen eines Anrufers dieses Anrufs; ein Spracherkennungsmittel, das funktionell mit dem Server-Mittel verbunden ist, zum Ermitteln des Themas des Anrufs; ein Vermittlungsmittel, das als Reaktion auf das Spracherkennungsmodul und das Spracherkennungsmittel den Anruf entsprechend der Erkennung des Anrufers und dem ermittelten Thema verarbeitet; ein Programmiermittel, das mit dem Server-Mittel, dem Spracherkennungsmodul, dem Spracherkennungsmittel und dem Vermittlungsmittel funktionell verbunden ist, mit dem ein Benutzer das System programmieren kann;

wobei das System durch Folgendes gekennzeichnet ist: ein Identitätsmarkierungsmittel, das als Reaktion auf das Sprechererkennungsmodul die Identität des Anrufers markiert; ein Umsetzungsmittel, das als Reaktion auf das Spracherkennungsmittel ein Telefongespräch oder eine Telefonnachricht des Anrufers umsetzt; und ein Audioindexierungsmittel, das funktionell mit dem Identitätsmarkierungsmittel und dem Umsetzungsmittel verbunden ist, zum Indexieren der Nachrichten und der Gespräche des Anrufers entsprechend dem Thema des Gesprächs und der Nachricht sowie der Identität des Anrufers.

**[0009]** Gemäß einem anderen Aspekt der vorliegenden Erfindung wird ein Verfahren zur Verarbeitung von Anrufen bereitgestellt, das folgende Schritte umfasst: Empfangen eines ankommenden Anrufs; Erkennen eines Anrufers des Anrufs und Ermitteln des Themas des Anrufs, wenn ein ankommender Anruf empfangen wird; Verarbeiten des Anrufs und der Nachricht entsprechend der Identität des Anrufers und des Verfassers und dem Thema des Anrufs und der Nachricht; wobei das System durch die folgenden Schritte gekennzeichnet ist: Markieren der ermittelten Identität eines Anrufers; Umsetzen des ermittelten Themas des Anrufs; Indexieren der sich aus der Markierung und der Umsetzung ergebenden Informationen gemäß einem ermittelten Thema und der ermittelten Identität.

**[0010]** Gemäß noch einem weiteren Aspekt der vorliegenden Erfindung ist das Verfahren der vorliegenden Erfindung zur Verarbeitung von Anrufen in der Lage, die Identität eines Anrufers oder der Teilnehmer eines Telefonkonferenz zu markieren, während die Nachricht oder die Gespräche solcher Anrufer und Teilnehmer umgesetzt werden. Folglich kann das System automatisch Telefonnachrichten und -gespräche sowie Voicemail-, eMail- und Fax-Nachrichten verarbeiten, indem es solche Anrufe und Nachrichten nach ihrem Thema oder der Identität des Anrufers oder des Verfassers oder beider speichert. Genauer gesagt, die vorliegende Erfindung kann in Verbindung mit einer solchen Erkennung und Umsetzung automatisch die empfangenen Telefonanrufe und eMail- und Fax-Nachrichten nach ihrem Ursprung und/oder ihrem Thema indexieren oder ihnen Prioritäten zuweisen, sodass ein berechtigter Benutzer bestimmte Nachrichten, z. B. von einer bestimmten Quelle stammende Nachrichten oder Nachrichten mit einem ähnlichen oder einem bestimmten Thema, abrufen kann.

**[0011]** Gemäß einem weiteren Aspekt der vorliegenden Erfindung beinhaltet das System Text-Sprache-Umsetzungsfähigkeiten, damit das System den Benutzer oder Anrufer in Form synthetisierter Sprache auffordern (d. h. abfragen), Antworten auf Fragen oder Anforderungen des Benutzers in Form synthetisierter Sprache geben und eMail- und Fax-Nachricht-

ten in Form synthetisierter Sprache wiedergeben kann. Das System beinhaltet auch Wiedergabefähigkeiten zum Wiedergeben aufgezeichneter Telefonnachrichten und anderer aufgezeichneter Audiodaten.

**[0012]** Diese sowie weitere Aufgaben, Merkmale und Vorteile der vorliegenden Erfindung werden aus der folgenden detaillierten Beschreibung deren anschaulicher Ausführungsarten in Verbindung mit den beiliegenden Zeichnungen klar.

**[0013]** [Fig. 1](#) ist ein Blockschaltbild, das die allgemeinen Funktionen eines Systems gemäß der vorliegenden Erfindung zur automatischen Verarbeitung von Anrufen und Datenübertragungen veranschaulicht;

**[0014]** [Fig. 2](#) ist ein Blockschaltbild sowie ein Ablaufplan, der die funktionelle Verknüpfung zwischen den Modulen für ein System gemäß einer Ausführungsart der vorliegenden Erfindung zur Verarbeitung von Anrufen und Datenübertragungen veranschaulicht; und

**[0015]** [Fig. 3a](#) und [Fig. 3b](#) sind Ablaufpläne, die ein Verfahren gemäß der vorliegenden Erfindung zur Verarbeitung von Anrufen und Datenübertragungen veranschaulichen.

**[0016]** [Fig. 1](#) ist ein Blockschaltbild, das die allgemeinen Funktionen eines Systems gemäß der vorliegenden Erfindung zur automatischen Verarbeitung von Anrufen und Datenübertragungen zeigt. Bei der vorliegenden Erfindung handelt es sich um eine Maschine zur automatischen Verarbeitung von Anrufen und Datenübertragungen, die von einem berechtigten Benutzer (Kasten **12**) so programmiert werden kann, dass ankommende Telefonanrufe in einer von einem solchen Benutzer vorgegebenen Weise verarbeitet werden. Obwohl die vorliegende Erfindung zur Verarbeitung beliebiger Sprachdaten, die über digitale oder analoge Kanäle empfangen werden können, sowie von Daten genutzt werden kann, die elektronisch empfangen und (gemäß den folgenden Erläuterungen) anderweitig in lesbaren Text umgesetzt werden können, betrifft eine Ausführungsart der vorliegenden Erfindung die Verarbeitung von Telefongesprächen. Insbesondere beantwortet das System **10** automatisch einen ankommenden Telefonanruf von einem Anrufer (Kasten **14**) und kann in Abhängigkeit davon, wie der Benutzer (Kasten **12**) das System **10** programmiert hat, den Telefonanruf verarbeiten, indem es den Anruf zum Beispiel an ein anderes Telefonsystem oder einen Anrufbeantworter (Kasten **18**) weiterleitet oder indem es den Anruf direkt verarbeitet, z. B. den Anrufer durchstellt, trennt oder in die Warteschleife versetzt (Kasten **16**). Außerdem kann das System **10** so programmiert werden, dass es einen ankommenden Telefonanruf in einer bestimmten

Reihenfolge an verschiedene Telefonsysteme (z. B. durch Weiterleiten des Anrufs an verschiedene vorgegebene Rufnummern, bis der Anruf beantwortet wird) oder gleichzeitig an alle solche Systeme weiterleitet. Es ist klar, dass die in Kasten **18** aufgeführten Telefonsysteme sowie die in Kasten **16** von [Fig. 1](#) gezeigten Optionen nur zur Veranschaulichung dienen und keine Beschränkung der Verarbeitungsprozeduren darstellen, für deren Ausführung das System **10** programmiert werden kann.

**[0017]** Gemäß einer weiteren Ausführungsart der vorliegenden Erfindung kann das System **10** so programmiert werden, dass es ankommende Fax- und eMail-Nachrichten verarbeiten oder Nachrichten von eMail- oder Voicemail-Systemen automatisch abrufen kann. Somit sollte klar sein, dass die bidirektionalen Leitungen von [Fig. 1](#) zum Verbinden des Systems **10** mit den Telefonsystemen in Kasten **18** (z. B. eMail, Voicemail, Fax/Modem und Anrufbeantworter) anzeigen, dass das System **10** zum Senden von Daten (z. B. Anrufen oder Nachrichten) an solche Systeme sowie zum Abrufen und Verarbeiten der in solchen Systemen gespeicherten oder aufgezeichneten Daten eingerichtet ist. Zum Beispiel kann das System **10** so programmiert werden, dass es einen bestimmten Anruf dadurch verarbeitet, dass es den Anruf zum Aufzeichnen an einen Anrufbeantworter (Kasten **18**) weiterleitet. Das System **10** kann anschließend die aufgezeichnete Nachricht vom Anrufbeantworter abrufen, codieren und in einer bestimmten Weise verarbeiten. Ferner kann das System **10** so programmiert werden, dass es einen ankommenden Telefonanruf oder Nachrichten in eine Textseite umsetzt, die dann an den Personrufempfänger, ein Mobiltelefon oder an ein eMail-System des Benutzers übertragen werden kann.

**[0018]** Die Funktionsmodule des Systems **10** und deren Zusammenwirken gemäß einer Ausführungsart der vorliegenden Erfindung wird im Folgenden unter Bezug auf [Fig. 2](#) erläutert. Es sollte klar sein, dass gleiche oder ähnliche in den Figuren dargestellte Komponenten mit derselben Bezugsnummer bezeichnet werden. Ferner sollte klar sein, dass die hier beschriebenen Funktionsmodule gemäß der vorliegenden Erfindung in Form von Hardware, Software oder einer Kombination beider realisiert werden können. Zum Beispiel werden die Hauptmodule zur Sprach- und Sprechererkennung, zur Sprachenerkennung und zum Indexieren vorzugsweise in Form von Software auf einem oder mehreren in geeigneter Weise programmierten digitalen Universalcomputern realisiert, die jeweils einen Prozessor, einen zugehörigen Speicher und Eingabe-/Ausgabe-Schnittstellen zum Ausführen der Elemente der vorliegenden Erfindung aufweisen. Es sollte klar sein, dass die Funktionselemente von [Fig. 2](#), obwohl die Erfindung vorzugsweise auf einem in geeigneter Weise programmierten Universalcomputer realisiert wird, eine ge-

eignete und bevorzugte Prozessorarchitektur zum Ausführen der Erfindung enthalten können und Beispiele für Funktionselemente darstellen, die durch Programmierung in einem solchen Computer realisiert werden können. Ferner können die Funktionselemente von [Fig. 2](#) durch Programmieren eines oder mehrerer Universal-Mikroprozessoren realisiert werden. Zum Realisieren der Erfindung können natürlich auch Spezial-Mikroprozessoren verwendet werden. Ausgehend von den hier dargelegten Lehren der Erfindung kann sich ein Fachmann diese und ähnliche Realisierungsformen der Elemente der Erfindung vorstellen.

**[0019]** Gemäß [Fig. 2](#) beinhaltet das System **10** einen Server **20**, der vorzugsweise mit verschiedenen Telefonsystemen verbunden ist, darunter, aber nicht darauf beschränkt, eine oder mehrere Telefonleitungen (Kasten **14**) und eine oder mehrere Faxleitungen und eine Modemleitung ([Fig. 1](#) und [Fig. 2](#), Kasten **18**) zum Empfangen und Senden von Telefonanrufen bzw. Nachrichtendaten. Der Server **20** ist so programmiert, dass er automatisch ankommende Telefonanrufe beantwortet und ankommende Fax-Sendungen empfängt. Das System **10** kann auch eine Internet-/Intranet-Standleitung zum Zugreifen auf einen örtlichen Mail-Server beinhalten, wobei der Server **20** so programmiert sein kann, dass er (über das TCP/IP-Protokoll) periodisch eine Verbindung zu diesem örtlichen Mail-Server herstellt, um sowohl ankommende eMails zu empfangen und zu verarbeiten als auch eMail-Nachrichten zu senden. Wenn das System **10** nicht ständig mit einem lokalen Netz-Server verbunden ist, kann der System-Server **20** alternativ so programmiert werden, dass er periodisch eine Zugangsnummer zu einem Internetanbieter wählt, um eMail-Nachrichten abzurufen oder zu senden. Solche Prozeduren können (im Gegensatz zur automatischen Überwachung solcher eMail-Konten) auch auf Wunsch des Benutzers ausgeführt werden, wenn der Benutzer auf das System **10** zugreift.

**[0020]** Gemäß den [Fig. 1](#) und [Fig. 2](#) (Kasten **18**) kann der Server **20** auch direkt mit Voicemail-Systemen und Anrufbeantwortern verbunden werden, damit der Benutzer Nachrichten, die auf solchen Voicemail-Systemen und Anrufbeantwortern aufgezeichnet wurden, abrufen und verarbeiten kann. Wenn das System **10** mit einem lokalen Netz verbunden ist, kann der Server **20** so programmiert werden, dass er periodisch Nachrichten von anderen Voicemail-Systemen oder Anrufbeantwortern abrufen, die nicht direkt mit dem Server **20** verbunden sind, auf die aber ansonsten über das lokale Netz zugegriffen werden kann, sodass das System **10** automatisch Nachrichten von solchen Voicemail-Systemen oder Anrufbeantwortern abrufen und überwachen kann.

**[0021]** Der Server **20** beinhaltet eine Aufzeichnungseinheit **40** zum Aufzeichnen und Speichern von

Audiodaten vorzugsweise in digitaler Form (z. B. von ankommenden Telefonanrufen oder Nachrichten, die von Voicemail-System oder Anrufbeantwortern abgerufen wurden). Außerdem beinhaltet der Server **20** vorzugsweise ein Komprimierungs-/Dekomprimierungsmodul **42** zum Komprimieren der digitalisierten Audiodaten sowie der als eMail und Fax-Nachricht empfangenen Nachrichtendaten, um die Datenspeicherkapazität eines (nicht gezeigten) Speichers des Systems **10** zu erhöhen, sowie zum Dekomprimieren solcher Daten vor der Wiederherstellung, wenn solche Daten vom Speicher abgerufen werden.

**[0022]** Ein Sprechererkennungsmodul **22** und ein Modul **24** zur automatischen Spracherkennung/Sprachenerkennung (Automatic Speech Recognizer/Natural Language Understanding, ASR/NLU) sind funktionell mit dem Server **20** verbunden. Das Sprechererkennungsmodul **22** ermittelt die Identität des Anrufers **14** und der Teilnehmer einer Konferenzschaltung aus den durch den Server **20** empfangenen Sprachdaten sowie die Identität des Urhebers einer empfangenen Fax- oder eMail-Nachricht. Das ASR/NLU-Modul **24** setzt die Sprachdaten oder andere vom Server **20** empfangene Nachrichtendaten in lesbaren Text um, um den Inhalt und das Thema solcher Anrufe, Gespräche oder Nachrichten zu ermitteln. Ferner wird im Folgenden gezeigt, dass das ASR/NLU-Modul **24** Sprachbefehle von einem berechtigten Benutzer verarbeitet, damit dieser das System **10** programmieren sowie Nachrichten erzeugen oder abrufen kann. Das ASR/NLU-Modul **24** verarbeitet auch Sprachdaten von Anrufern und berechtigten Benutzern, um interaktive Sprachantwortfunktionen (Interactive Voice Response, IVR) auszuführen. Ein Spracherkennungs-/Übersetzungsmodul **26**, das funktionell mit dem ASR/NLU-Modul **24** verbunden ist, wird bereitgestellt, damit das System **10** fremdsprachige Nachrichten verstehen und ordnungsgemäß beantworten kann, wenn das System zum Beispiel in einem mehrsprachigen Land wie beispielsweise Kanada genutzt wird.

**[0023]** Ein Vermittlungsmodul **28**, das funktionell mit dem Sprechererkennungsmodul **22** und dem ASR/NLU-Modul **24** verbunden ist, verarbeitet die durch das Sprechererkennungsmodul **22** und/oder durch das ASR/NLU-Modul **24** empfangenen Daten. Das Vermittlungsmodul führt mit den ankommenden Telefonanrufen oder den Fax- oder eMail-Nachrichten gemäß einer vorprogrammierten Prozedur eine Verarbeitungsprozedur durch (z. B., indem es einen Anruf an ein Voicemail-System oder einen Anrufbeantworter weiterleitet).

**[0024]** Ein Identitätsmarkierungsmodul (ID-Markierungsmodul) **30**, das funktionell mit dem Sprechererkennungsmodul **22** verbunden ist, wird zum elektronischen Markieren der Identität des Anrufers in dessen Nachricht oder Gespräch oder zum Markieren

der Identität des Urhebers einer eMail- oder Fax-Nachricht bereitgestellt. Wenn das System im Hintergrund einer Telefonkonferenz aktiv ist, markiert das ID-Markierungsmodul **30** außerdem die Identität der gerade Sprechenden Person. Ein Umsetzungsmodul **32**, das funktionell mit dem ASR/NLU-Modul **24** verbunden ist, wird zum Umsetzen von Telefonnachrichten oder -gesprächen, Telefonkonferenz- und/oder Fax-Nachrichten bereitgestellt. Außerdem kann das Umsetzungsmodul **32** eine vom Benutzer diktierte Sprachnachricht umsetzen, die dann vom System **10** per Telefon, Fax oder eMail an eine andere Person gesendet werden kann.

**[0025]** Ein Audioindexierungs-/Prioritätsmodul **34** ist funktionell mit dem ID-Markierungsmodul **30** und dem Umsetzungsmodul **32** verbunden. Das ID-Markierungsmodul **34** speichert in einer vorprogrammierten Weise die Umsetzungsdaten und die Anruferkennungsdaten, die vom Umsetzungsmodul **32** beziehungsweise vom ID-Markierungsmodul **30** verarbeitet werden, sowie die Uhrzeit des Anrufs und die Rufnummer des Absenders (sofern verfügbar über die automatische Nummernerkennung (Automatic Number Identification, ANI), damit der Benutzer bestimmte Anrufe oder Nachrichten von einem bestimmten Absender oder Nachrichten mit einem bestimmten Thema abrufen kann. Ferner kann das Audioindexierungs-/Prioritätsmodul so programmiert werden, dass es bestimmten Anrufen oder Nachrichten eine Priorität zuweist und den Benutzer über solche Anrufe oder Nachrichten informiert.

**[0026]** Ein Sprachsynthesemodul **36**, das funktionell mit dem Audioindexierungs-/Prioritätsmodul **34** verbunden ist, ermöglicht dem Benutzer das Abrufen von Nachrichten (eMails oder Fax-Nachrichten) in Form von Audiosignalen (d. h. in Form synthetisierter Sprache). Das Sprachsynthesemodul ist auch mit dem ASR/NLU-Modul funktionell verbunden, um (im Gegensatz zur Anzeige zum Beispiel auf einem Computermonitor) Systemeingabeaufforderungen (d. h. Abfragen) in Form synthetisierter Sprache bereitzustellen.

**[0027]** Eine Programmieroberfläche **38**, die funktionell mit dem Server **20**, dem Sprechererkennungsmodul **22**, dem Spracherkennungs-/Übersetzungsmodul **26**, dem ASR/NLU-Modul **24**, dem Audioindexierungs-/Prioritätsmodul **34** und dem Vermittlungsmodul verbunden ist, wird zum Programmieren des Systems **10** bereitgestellt, um Anrufe und Nachrichten nach einer vorgegebenen Prozedur zu verarbeiten. Im Folgenden wird ausführlich erläutert, dass ein Benutzer das System unter Verwendung der Programmierschnittstelle **38** entweder über Sprachbefehle oder über eine grafische Benutzeroberfläche (Graphical User Interface, GUI) oder über beide programmieren kann. Bei einer bevorzugten Ausführungsart wird das System **10** durch Sprachbefehle

vom Benutzer (d. h. im Sprachbefehlsmodus) programmiert. Genauer gesagt, der Benutzer kann das System **10** entweder von fern durch Anrufen des Systems **10** oder an Ort und Stelle mit einem Mikrofon programmieren. Die Programmierschnittstelle **38** ist mit dem Server **20** verbunden, der in Verbindung mit dem Sprechererkennungsmodul **22** und dem ASR/NLU-Modul **24** die Identität des Benutzers prüft, bevor dieser die vom Benutzer gesprochenen Programmierbefehle verarbeitet. Das System **10** kann Informationen über die gesprochenen Programmierbefehle (d. h., ob das System **10** einen solchen Befehl akzeptiert) sowie die aktuelle Programmstruktur des Systems **10** entweder (über die GUI) anzeigen oder (über das Sprachsynthesemodul **36**) wiedergeben.

**[0028]** Gemäß einer weiteren Ausführungsart kann das System **10** an Ort und Stelle über einen Personal Computer und einen GUI-Bildschirm oder von fern programmiert werden, indem von einem fernen Standort über ein Computernetz auf das System **10** zugegriffen wird. Ähnlich wie bei der herkömmlichen Windows-Oberfläche kann der Benutzer das System **10** programmieren, indem er bestimmte auf der GUI angezeigte Felder auswählt. Es sollte klar sein, dass das System **10** durch eine Kombination von Sprachbefehlen und einer GUI programmiert werden kann. In einem solchen Fall kann die GUI zum Beispiel den Benutzer unterstützen, indem sie ihm die zum Programmieren des Systems **10** zur Verfügung stehenden Sprachbefehle anzeigt. Darüber hinaus kann das System **10** durch Bearbeiten einer entsprechenden Programmstrukturdatei programmiert werden, welche die Funktionsmodule von [Fig. 2](#) steuert.

**[0029]** Im Folgenden wird unter Bezug auf [Fig. 2](#) und die [Fig. 3a](#) und [Fig. 3b](#) die Funktionsweise der vorliegenden Erfindung beschrieben. Es sollte klar sein, dass die Darstellung der vorliegenden Erfindung in [Fig. 2](#) als Ablaufplan zur Veranschaulichung der Arbeitsschritte der vorliegenden Erfindung sowie als Blockschaltbild zur Darstellung einer Ausführungsart der vorliegenden Erfindung zu verstehen ist. Der Server **20** wird so programmiert, dass er einen ankommenden Telefonanruf, eine eMail, ein Fax/Modem oder andere elektronische Sprach- oder Nachrichtendaten automatisch beantwortet (Schritt **100**). Der Server **20** unterscheidet anhand spezieller Codes, d. h. Protokolle am Anfang jeder Nachricht, welche die Quelle anzeigen, zwischen ankommenden Telefonanrufen, eMail-Nachrichten, Fax-Nachrichten usw. Insbesondere geht der Server **20** zuerst davon aus, dass es sich bei dem ankommenden Anruf um ein Telefongespräch handelt und verfährt entsprechend (Schritt **110**), sofern der Server **20** nicht zum Beispiel ein Quittungssignal von einem Modem empfängt, worauf das System **10** den Anruf als Computer-Verbindungsprotokoll verarbeitet. Es sollte klar sein, dass das System **10** so programmiert werden

kann, dass es andere Voicemail- oder eMail-Konten durch periodisches Anrufen solcher Konten und durch Abrufen von Voicemail- und eMail-Nachrichten von solchen Konten überwachen kann.

**[0030]** Wenn festgestellt wird, dass es sich bei dem vom Server **20** empfangenen ankommenden Anruf um einen Telefonanruf handelt, werden die Audiodaten (z. B. ankommende Anrufe sowie von Voicemail-Systemen oder Anrufbeantwortern abgerufene Anrufe) von der Aufzeichnungseinheit **40** aufgezeichnet (Schritt **112**). Bei der Aufzeichnungseinheit **40** kann es sich um eine herkömmliche Einheit wie beispielsweise eine analoge Aufzeichnungseinheit oder ein digitales Audioband (Digital Audio Tape, DAT) handeln. Vorzugsweise wird als Aufzeichnungseinheit **40** eine digitale Aufzeichnungseinheit verwendet, d. h. ein Analog-Digital-Umsetzer zum Umsetzen der Audiodaten in digitale Daten. Die digitalisierten Audiodaten können dann durch das Komprimierungs-/Dekomprimierungsmodul **42** (Schritt **114**) komprimiert werden, bevor sie in dem (in [Fig. 2](#) nicht gezeigten Speicher) gespeichert werden (Schritt **116**). Es sollte klar sein, dass vom Komprimierungs-/Dekomprimierungsmodul **42** zur Verarbeitung der Nachrichtendaten ein beliebiger herkömmlicher Algorithmus verwendet werden kann, zum Beispiel der in „Digital Signal Processing, Synthesis and Recognition“, von S. Furui, Dekker, 1989, beschriebene Algorithmus.

**[0031]** Als Nächstes wird gleichzeitig mit dem Aufzeichnen und Speichern der Audiodaten die Identität des Anrufers ermittelt, indem die Audionachrichten und/oder Audioantworten des Anrufers auf Abfragen seitens des Systems **10** verarbeitet werden. Genauer gesagt, die verbalen Äußerungen und Antworten werden durch den Server **20** empfangen und an das Sprechererkennungsmodul **22** gesendet, wo solche verbalen Äußerungen und Antworten verarbeitet und mit früher gespeicherten Sprechermodellen verglichen werden (Schritt **120**). Wenn der Sprecher durch Übereinstimmung der empfangenen Sprachdaten mit einem früher gespeicherten Sprachmodell eines solchen Sprechers erkannt wird (Schritt **130**) und das System **10** so programmiert ist, dass es Anrufe entsprechend der Identität eines Anrufers verarbeitet, verarbeitet das System **10** den Telefonanruf gemäß einer solchen vorher programmierten Prozedur (Schritt **152**).

**[0032]** Wenn der Sprecher (z. B. ein Sprecher, der zum ersten Mal anruft) andererseits nicht anhand der zuvor gespeicherten Sprachmodelle erkannt werden kann, kann eine Sprechererkennung sowohl durch das Sprechererkennungsmodul **22** als auch durch das ASR/NLU-Modul **26** erfolgen, wobei der Inhalt der Telefonnachricht durch das ASR/NLU-Modul **26** verarbeitet werden kann, um den Namen des Anrufers zu ermitteln, der dann mit zuvor gespeicherten Namen verglichen wird, um die Identität eines sol-

chen Anrufers zu ermitteln (Schritt 140). Wenn dann die Identität des Anrufers ermittelt wird, verarbeitet das System 10 den Telefonanruf entsprechend der Identität des Anrufer (Schritt 152).

**[0033]** Falls das System 10 nicht in der Lage ist, die Identität des Anrufers entweder anhand der gespeicherten Sprachmodelle oder des Inhalts der Telefonnachricht zu ermitteln, sendet das Sprechererkennungsmodul 22 ein Signal an den Server 20, der wiederum den Anrufer mit einer Abfrage, z. B. „Wer sind Sie“, auffordert, seinen Namen anzugeben (Schritt 150), worauf der obige Erkennungsprozess wiederholt wird (Schritt 120). Der Server 20 erhält die Abfrage vom Sprachsynthesemodul 36 in Form synthetisierter Sprache. Gemäß der obigen Beschreibung sollte klar sein, dass das System 10 so programmiert werden kann, dass es zu Anfang den Anrufer nach seinem Namen oder nach Einzelheiten über den Grund seines Anrufs befragt.

**[0034]** Sobald das Sprechererkennungsmodul 22 die Identität des Anrufers oder Urhebers ermittelt hat, wird vom Sprechererkennungsmodul 22 ein Signal an das Vermittlungsmittel 28 gesendet, wodurch das Vermittlungsmittel 28 den Anruf oder die Nachricht anhand der Identität des Anrufers oder Urhebers gemäß einer vorher programmierten Prozedur verarbeitet (Schritt 152). Wenn die Identität des Anrufers jedoch überhaupt nicht ermittelt werden kann, kann das System 10 so programmiert sein, dass es den Anruf als Anruf von einem unbekanntem Anrufer verarbeitet (Schritt 154), indem es den Anruf z. B. an ein Voicemail-System weiterleitet. Eine solche Programmierung, die näher erläutert werden muss, wird vom Benutzer 12 über das Programmieroberflächenmodul 38 ausgeführt. Gemäß der obigen Beschreibung können die Verarbeitungsschritte des Systems 10 so programmiert werden, dass es unter anderem, aber nicht ausschließlich, das Vermitteln des Anrufs an ein anderes System, das Weiterleiten des Anrufs an ein anderes Telefonendgerät (Fig. 1 und Fig. 2, Kasten 18) oder die direkte Verarbeitung des Anrufs entweder durch das Durchstellen des Anrufs an einen bestimmten Teilnehmer, das Trennen des Anrufs oder das Versetzen des Anrufs in eine Warteschleife (Fig. 1 und Fig. 2, Kasten 16) ausführt.

**[0035]** Es sollte klar sein, dass immer, wenn ein neuer Anrufer zum ersten Mal mit dem System 10 in Verbindung tritt, Sprechermodelle erstellt und im Sprechererkennungsmodul 22 gespeichert werden, sofern sie nicht auf Wunsch des Benutzers gelöscht werden. Solche Modelle werden dann vom Sprechererkennungsmodul 22 zur Prüfung und Erkennung verwendet, wenn der Anrufer zu einem späteren Zeitpunkt mit dem System 10 in Verbindung tritt.

**[0036]** Es sollte klar sein, dass das System 10 die Sprechererkennung unter Verwendung anderer Ver-

fahren durchführen kann, die nicht auf akustischen Merkmalen beruhen, wenn die erforderlichen Sprachmodelle nicht vorliegen. Beispielsweise kann das System 10 bei Telefonanrufen weitere Informationen (z. B. die Anrufer-ID) nutzen, um die Treffergenauigkeit des Systems 10 zu erhöhen und/oder erstmalig anrufende Anrufer zu erkennen.

**[0037]** Gemäß der folgenden Erläuterung kann das System 10 so programmiert werden, dass der Name und die Rufnummer jedes Anrufers (oder bestimmter Anrufer) gespeichert werden. Durch eine solche Funktion kann der Benutzer automatisch Antwortnachrichten an Anrufer versenden sowie dynamisch ein Adressbuch erstellen (das im System 10 gespeichert wird), auf das der Benutzer später zugreifen kann, um eine Nachricht an eine bestimmte Person zu senden.

**[0038]** Es sollte klar sein, dass es je nach Anwendung nicht in jedem Fall erforderlich ist, dass das System 10 die Sprechererkennung und Spracherkennung in Echtzeit (d. h. gleichzeitig mit dem Aufzeichnen und während des Telefongesprächs selbst) durchführt. Zum Beispiel kann das System 10 so programmiert werden, dass es beim Eingang des Anrufs den Anrufer (über die IVR-Programmierung) zur Gewinnung wichtiger Informationen (d. h. des Namens und des Grundes für den Anruf) abfragt und diese Informationen speichert. Nach dem Anruf kann dann der Erkennungsprozess durch das Sprechererkennungsmodul 22 oder das ASR/NLU-Modul 24 durchgeführt werden, indem die gespeicherten Audiodaten vom Speicher abgerufen werden (Schritt 118) (siehe gestrichelte Linie in Fig. 3a).

**[0039]** Es sollte klar sein, dass als Sprechererkennungsmodul 22 zur Erkennung des Anrufers ein beliebiges Sprechererkennungssystem verwenden kann. Vorzugsweise nutzt das gemäß der vorliegenden Erfindung angewendete Sprechererkennungssystem ein System, das eine textunabhängige Sprechererkennung durchführt und Zufallsfragen stellt, d. h. eine Kombination von Spracherkennung, textunabhängiger Sprechererkennung und Spracherkennung gemäß der Beschreibung in der US-Patentanmeldung mit der Serien-Nr. 08/871 784, eingereicht am 11. Juni 1997, mit dem Titel „Apparatus and Methods for Speaker Verification/Identification/Classification Employing Non-Acoustic and/or Acoustic Models and Databases“. Insbesondere beruht die textunabhängige Sprechererkennung vorzugsweise auf einer Merkmalsklassifikation einzelner Ausschnitte (frames), die ausführlich in der US-Patentanmeldung mit der Serien-Nr. 08/788 471, eingereicht am 28. Januar 1977, mit dem Titel „Text Independent Speaker Recognition for Transparent Command Ambiguity and Continuous Access Control“ beschrieben wird, die hier durch Bezugnahme einbezogen ist.

**[0040]** Gemäß der Erläuterung in der oben erwähnten einbezogenen US-Patentanmeldung mit der Serien-Nr. 08/871 784 wird der textunabhängigen Sprechererkennung der Vorzug vor der textabhängigen oder der textgebundenen Sprechererkennung gegeben, da die Sprechererkennungsfunktion bei Textunabhängigkeit parallel mit anderen Spracherkennungsfunktionen in einer für den Anrufer transparenten Weise ausgeführt werden kann, ohne das Gespräch durch neue Befehle oder die Erkennung eines neuen Anrufers zu unterbrechen, wenn ein neuer Anrufer in Erscheinung tritt.

**[0041]** Gemäß [Fig. 3b](#) (und unter der Annahme, dass das System **10** zur Verarbeitung von Anrufen auf der Grundlage der Identität eines Anrufers oder Urhebers programmiert ist) wird als Nächstes ermittelt, ob es sich bei dem ankommenden Anruf um eine Fax- oder eine eMail-Nachricht handelt, worauf die Nachrichtendaten (z. B. ankommende eMails oder Nachrichten, die von eMail-Konten abgerufen wurden) durch das ASR/NLU-Modul **24** (Schritt **190**) verarbeitet, komprimiert (Schritt **192**) und in einem (nicht gezeigten) Speicher gespeichert werden (Schritt **194**). Bei eMail-Nachrichten werden die Daten direkt verarbeitet (da solche Daten bereits im Textformat vorliegen). Bei Fax-Nachrichten nutzt das ASR/NLU-Modul **24** bekannte Techniken zur optischen Zeichenerkennung (Optical Character Recognition, OCR), um die empfangene Fax-Nachricht in lesbaren Text umzusetzen (d. h., die Fax-Nachricht in eine ASCII-Datei umzusetzen.)

**[0042]** Als Nächstes kann gleichzeitig mit dem Umsetzen und Speichern der Daten der ankommenden Nachricht die Identität des Urhebers einer solchen Nachricht durch das ASR/NLU-Modul **24** ermittelt werden, wobei der Inhalt der ankommenden Nachricht analysiert wird (Schritt **200**), um den Namen des Urhebers oder die Quelle der Nachricht zu entnehmen, der dann zur Ermittlung der Identität eines solchen Urhebers mit den zuvor gespeicherten Namen verglichen wird (Schritt **210**). Wenn der Urheber erkannt worden ist (Schritt **210**), kann die Nachricht nach einer vorher programmierten Prozedur auf der Grundlage der Identität des Urhebers verarbeitet werden (Schritt **222**). Wenn die Identität des Urhebers jedoch nicht ermittelt werden kann, kann die Nachricht nach der vorher programmierten Prozedur für einen nicht erkannten Urheber verarbeitet werden (Schritt **224**).

**[0043]** Nach den obigen Erläuterungen sollte klar sein, dass das System **10** die ankommende oder abgerufene Nachricht nicht in jedem Fall in Echtzeit (d. h. gleichzeitig mit dem Umsetzen der ursprünglichen Nachricht) verarbeiten muss. Die Verarbeitung kann vom ASR/NLU-Modul **24** nach dem Empfangen der eMail- oder Fax-Nachricht erfolgen, indem die Daten der umgesetzten Nachricht vom Speicher abgerufen

werden (Schritt **196**)(siehe die gestrichelte Linie in [Fig. 3b](#)).

**[0044]** Zusätzlich zur Erkennung der Identität des Anrufers oder Urhebers kann das System **10** ferner vom Benutzer **12** so programmiert werden, dass ein ankommender Telefonanruf oder eine Fax- oder eMail-Nachricht auf der Grundlage des Inhalts oder des Themas des Anrufs oder der Nachricht und/oder der Uhrzeit verarbeitet wird, zu der ein solcher Anruf oder eine solche Nachricht empfangen wurde. Gemäß den [Fig. 2](#), [Fig. 3a](#) und [Fig. 3b](#) sendet der Server **20** nach dem Empfangen eines ankommenden Telefonanrufs oder einer eMail- oder Fax-Nachricht oder nach dem Abrufen einer aufgezeichneten Nachricht vom einem Anrufbeantworter oder einem Voicemail-System die Daten des Anrufs oder der Nachricht an das ASR/NLU-Modul **24**. Im Fall von Sprachdaten (z. B. Telefonanrufen oder Nachrichten, die von einem Voicemail-System oder einem Anrufbeantworter abgerufen wurden) setzt das ASR/NLU-Modul **24** solche Daten in eine Symbolsprache oder einen lesbaren Text um. Gemäß der obigen Erläuterung werden eMail-Nachrichten direkt verarbeitet (da sie bereits in einem lesbaren Textformat vorliegen) und Fax-Nachrichten durch das ASR/NLU-Modul **24** unter Verwendung bekannter optischer Zeichenerkennungsverfahren (OCR) in lesbaren Text (d. h. ASCII-Dateien) umgesetzt. Das ASR/NLU-Modul **24** analysiert dann die Daten des Anrufs oder der Nachricht unter Verwendung einer Kombination zur Spracherkennung, um bestimmte Schlüsselwörter oder Themen herauszufiltern und die gesprochene Sprache zu erkennen und so das Thema und den Inhalt des Anrufs (Schritt **160** in [Fig. 3a](#) für Telefonanrufe) oder der Nachricht (Schritt **200** in [Fig. 3b](#) für eMails oder Fax-Nachrichten) zu ermitteln.

**[0045]** Sobald das ASR/NLU-Modul das Thema des Anrufs (Schritt **170** in [Fig. 3a](#)) oder der Nachricht (Schritt **220** in [Fig. 3b](#)) ermittelt hat, wird vom ASR/NLU-Modul **24** ein Signal an das Vermittlungsmittel **28** gesendet, in dem der Anruf oder die Nachricht nach einer vorgegebenen Weise auf der Grundlage des Themas oder des Inhalts des Anrufs (Schritt **158** in [Fig. 3a](#)) oder des Inhalts der Nachricht (Schritt **228** in [Fig. 3b](#)) verarbeitet wird. Wenn eine Nachricht oder ein Anruf beispielsweise einen Notfall oder einen Unfall betrifft, kann das Vermittlungsmittel **28** so programmiert sein, dass der Anruf sofort an eine bestimmte Person übertragen wird.

**[0046]** Wenn das ASR/NLU-Modul **24** nicht in der Lage ist, das Thema oder den Inhalt eines Telefonanrufs zu ermitteln, sendet es ein Signal an das Sprachsynthesemodul **36**, das wiederum eine Nachricht an den Server **20** sendet, damit dieser den Anrufer auffordert, den Grund für den Anruf in wenigen Worten zu beschreiben (Schritt **180**), z. B. „Warum rufen Sie an?“ Auch hier sollte klar sein, dass das System **10**



so programmiert werden kann, dass es gleich zu Anfang den Anrufer auffordert, den Grund für den Anruf anzugeben. Wenn das System **10** dann immer noch nicht in der Lage ist, das Thema eines solchen Anrufs zu ermitteln, kann der Anruf nach einer vorher programmierten Prozedur auf der Grundlage eines unbekanntes Themas verarbeitet werden (Schritt **156**). Desgleichen kann die Nachricht nach einer vorher programmierten Prozedur auf der Grundlage eines unbekanntes Themas verarbeitet werden (Schritt **226**), wenn das Thema einer eMail- oder Fax-Nachricht nicht ermittelt werden kann (Schritt **220**).

**[0047]** Wenn der ankommende Anruf oder die eMail-Nachricht ferner in einer Fremdsprache in Bezug auf das System **10** (d. h., einer Fremdsprache für den Benutzer) abgefasst ist, teilt das ASR/NLU-Modul **24** dies dem Sprachenerkennungs-/Übersetzungsmodul **26** mit, um die betreffende Sprache des Anrufs oder der Nachricht zu ermitteln und die geforderte Übersetzung an das ASR/NLU-Modul **24** zu liefern, damit das System **10** den Anruf verstehen und dem Anrufer die Antwort in der richtigen Sprache zukommen lassen kann. Es sollte klar sein, dass das System **10** vorher auch so programmiert werden kann, dass Anrufe oder Nachrichten mit einer unbekanntes Sprache in einer bestimmten Weise verarbeitet werden.

**[0048]** Es sollte klar sein, dass in der vorliegenden Erfindung jedes beliebige herkömmliche Verfahren zur Sprachenerkennung und Übersetzung verwendet werden kann, zum Beispiel das bestens bekannte maschinelle Sprachenerkennungsverfahren, das von J. Hieronymus und S. Kadambe in dem Artikel „Robust Spoken Language Identification using Large Vocabulary Speech Recognition“, Proceedings of ICASSP 97, Bd. 2, S. 1111 ff., beschrieben wird, sowie das Sprachenübersetzungsverfahren, das von Hutchins und Somers in „An Introduction to Machine Translation“, Academic Press, London, 1992 (enzyklopädische Übersicht) beschrieben wird.

**[0049]** Abgesehen von den obigen Literaturangaben kann die Sprachenerkennung auch unter Verwendung verschiedener statistischer Verfahren durchgeführt werden. Erstens kann das System **10** den eingegebenen Text in jeder der verschiedenen Sprachen (unter Verwendung verschiedener ASR-Systeme) decodieren, wenn das System **10** zur Verarbeitung einer kleinen Anzahl verschiedener Sprachen konfiguriert ist (z. B. in Kanada, wo im Wesentlichen nur Englisch oder Französisch gesprochen wird). Dann werden die verschiedenen decodierten Scripts analysiert, um statistische Muster zu ermitteln (d. h. die statistische Verteilung decodierter Wörter in jedem Script wird analysiert). Wenn die Decodierung in der falschen Sprache erfolgt war, wäre die Verwirrung des decodierten Scripts sehr hoch, und diese Sprache würde nicht mehr berücksichtigt.

**[0050]** Zweitens kann die Sprachenerkennung auf phonetischer Ebene erfolgen, indem das System einen Satz von Phonemen erkennt (entweder unter Verwendung eines universellen phonetischen Systems oder mehrerer Systeme für verschiedene Sprachen). Dann untersucht das System die Häufigkeiten der decodierten Phonemsequenzen für jede Sprache. Wenn sich eine bestimmte decodierte Sequenz als ungebräuchlich herausstellt, würde das System eine solche Sprache ausschließen. Es kann aber auch Sequenzen geben, die für eine bestimmte Sprache typisch sind. Dadurch wird das System die mit hoher Wahrscheinlichkeit infrage kommende Sprache erkennen.

**[0051]** Es sollte klar sein, dass die vorliegende Erfindung die Identität des Anrufers zur Sprachenerkennung ausnutzen kann. Genauer gesagt, wenn das Sprecherprofil eines bestimmten Anrufers (das im System **10** gespeichert ist) besagt, dass der Anrufer in einer bestimmten Sprache spricht, kann diese Information zur Erkennung der Sprache herangezogen werden. Umgekehrt kann das System **10** unter Verwendung der oben erwähnten Verfahren eine bestimmte Sprache erkennen und dann die Identität eines Anrufers ermitteln, indem es die Sprecherprofile durchsucht und ermittelt, welche Sprecher eine solche erkannte Sprache verwenden.

**[0052]** Es sollte klar sein, dass sowohl die Sprachenerkennung als auch die Sprachenerkennung vom ASR/NLU-Modul **24** zur Verarbeitung der vom Server **20** empfangenen Daten verwendet werden kann. Die vorliegende Erfindung nutzt vorzugsweise die Verfahren zur Sprachenerkennung, die in der US-Patentanmeldung mit der Serien-Nr. 08/859 586 mit dem Titel „A Statistical Translation System with Features Based on Phrases or Groups of Words“, eingereicht am 20. Mai 1997, und in der US-Patentanmeldung mit der Serien-Nr. 08/593 032 mit dem Titel „Statistical Natural Language Understanding Using Hidden Clumpings“, eingereicht am 29. Januar 1996, beschrieben werden. Die oben einbezogenen Erfindungen betreffen Verfahren zur Sprachenerkennung, um Texteingaben (unter Verwendung bestimmter Algorithmen) zu parametrisieren, d. h. in eine Sprache umzusetzen, die vom System **10** verstanden und verarbeitet werden kann. Im Zusammenhang mit der vorliegenden Erfindung liefert die ASR-Komponente des ASR/NLU-Moduls **24** beispielsweise die NLU-Komponente eines solchen Moduls für eine unbeschränkte Texteingabe wie „Erste Nachricht von Bob wiedergeben“. Ein solcher Text kann von der NLU-Komponente des ASR/NLU-Moduls **24** in die Form „Nachricht-abrufen(Absender=Bob, Nachrichtennummer=1)“ umgesetzt werden. Ein solcher parametrisierter Vorgang kann vom System **10** verstanden und ausgeführt werden.

**[0053]** Die in dem Artikel von Zeppenfeld et al. mit

dem Titel „Recognition of Conversational Telephone Speech Using the Janus Speech Engine“, Proceedings of ICASSP 97, Bd. 3, S. 1815 ff., 1997, beschriebenen bekannten automatischen Spracherkennungsfunktionen und die in dem Artikel von K. Shirai und S. Furui mit dem Titel „Special Issue on Spoken Dialog“, 15, (3 bis 4) Speech Communication, 1994, beschriebenen bekannten Sprachenerkennungsfunktionen können ebenfalls in der vorliegenden Erfindung verwendet werden. Zur Vereinfachung der Programmierung des ASR/NLU-Moduls **24** und um sicherzustellen, dass gewisse kritische Nachrichten in ausreichender Weise bearbeitet werden, kann ferner die Suche nach Schlüsselwörtern vorzugsweise auf der Grundlage der Erkennungsverfahren erfolgen, die von Richard C. Cross in „Word Spotting from Continuous Speech Utterances“, Automatic Speech and Speaker Recognition, Advanced Topics, S. 303 bis 327, Herausgeber Chin-Hui Lee, Frank K. Soong, Kuldip K. Paliwal (Huwer Academic Publishers), 1996, beschrieben werden.

**[0054]** Es sollte klar sein, dass das System **10** durch die Nutzung der Spracherkennung gemäß der obigen Darstellung in der Lage ist, interaktive Sprachantwortfunktionen (Interactive Voice Response, IVR) auszuführen, um einen Dialog mit dem Benutzer oder Anrufer zu eröffnen, um den Dialog zu führen und die Anforderungen zu verstehen. Dadurch kann das System **10** zur Auftragsannahme und zum Ausfüllen von Formularen im Verlauf des Dialogs genutzt werden. Ferner versetzen solche Funktionen den Anrufer in die Lage zu entscheiden, wie der Anruf verarbeitet werden soll (unter der Annahme, dass das System **10** dementsprechend programmiert ist), d. h. durch Hinterlegen einer eMail- oder Voice-mail-Nachricht, durch Senden einer Textseite oder Übermitteln des Anrufs an eine andere Rufnummer. Gemäß der obigen Erläuterung kann das System **10** dadurch außerdem vom Benutzer von einem fernen Standort durch Sprachbefehle programmiert werden.

**[0055]** Ferner sollte klar sein, dass das System **10** zum Schutz vor unberechtigten Zugriffen auf das System **10** beiträgt. Insbesondere muss ein Benutzer, der auf das System **10** zugreifen und sich an diesem beteiligen will, den Anmeldeprozess des Systems durchlaufen. Dieser Prozess kann auf verschiedene Weise ablaufen. Zum Beispiel kann die Anmeldung von einem fernen Standort erfolgen, indem ein neuer Benutzer anrufen und eine zuvor erteilte persönliche Kennnummer (PIN) eingeben muss, wobei der Server **20** so programmiert sein kann, dass er auf die PIN antwortet, die der neue Benutzer über DTMF-Tasten seines Telefons in das System **10** eingibt. Das System **10** kann dann Sprachmodelle des neuen Benutzers erstellen, um den neuen Benutzer zu prüfen und zu erkennen, wenn er zu einem späteren Zeitpunkt versucht, auf das System **10** zuzugreifen oder es zu programmieren. Alternativ kann zum

Erstellen der für die zukünftige Erkennung und Prüfung nötigen Sprechermodelle entweder ein aufgezeichnetes oder ein laufendes Telefongespräch des neuen Benutzers genutzt werden.

**[0056]** Es sollte klar sein, dass der Server **20** der vorliegenden Erfindung gemäß den Lehren der Patentanmeldung (IBM-Aktenzeichen Y0977-313) mit dem Titel „Apparatus and Methods for Providing Repetitive Enrollment in a Plurality of Biometric Recognition Systems Bases on an Initial Enrollment“ aufgebaut sein, um die Sprechermodelle (d. h. biometrische Daten) berechtigter Benutzer (die im Server **20** gespeichert sind) anderen biometrischen Erkennungssystemen zur Verfügung zu stellen, damit der Benutzer automatisch angemeldet wird, ohne dass zur Anmeldung in solchen Systemen systematisch neue biometrische Modelle erstellt werden müssen.

**[0057]** Der Prozess der Programmierung des Systems **10** kann vom Benutzer entweder an Ort und Stelle über eine grafische Benutzeroberfläche oder über Sprachbefehle oder von einem fernen Standort über eine Telefonleitung (durch Sprachbefehle) oder über ein mit dem System verbundenes Netz ausgeführt werden. In allen Fällen wird dies über eine Programmieroberfläche **38** abgewickelt. Wie oben gezeigt wurde, erfolgt das Programmieren des Systems **10** z. B. durch Auswählen der Namen von Personen, die an eine bestimmte Rufnummer, ein Voice-mail-System oder einen Anrufbeantworter vermittelt werden sollen, durch Eingeben bestimmter Schlüsselwörter oder Themen, aus denen das System **10** die Notwendigkeit bestimmter Verarbeitungsprozeduren erkennt und/oder durch Programmieren des Systems **10** in der Weise, dass Notfallanrufe oder geschäftliche Anrufe zwischen 8:00 Uhr und 12:00 Uhr sofort durchgestellt werden. [Fig. 2](#) zeigt, dass die Programmieroberfläche **38** solche Informationen an den Server **20**, an das Spracherkennungsmodul **22**, das ASR/NLU-Modul **24**, das Sprachenerkennungs-/Übersetzungsmodul **26**, das Audioindexierungs-/Prioritätsmodul **34** und das Vermittlungsmittel **28** sendet, welches das System **10** veranlasst, Anrufe gemäß den vom Benutzer programmierten Anweisungen zu verarbeiten.

**[0058]** Die Programmieroberfläche reagiert entweder auf DTMF-Tasteneingaben oder Sprachbefehle eines berechtigten Benutzers. Das bevorzugte Verfahren zur Programmierung des Systems **10** wird nicht über DTMF-Tasteneingaben oder über eine GUI-Oberfläche, sondern über sprachaktivierte Befehle durch einen Prozess zur Spracherkennung und zur Sprachenerkennung abgewickelt. Durch diesen Prozess kann das System **10** den Benutzer prüfen und erkennen, bevor dieser auf das System **10** zugreifen kann. Diese Maßnahme trägt zum Schutz vor unberechtigten Benutzern bei, die möglicherweise Kenntnis von einer ansonsten gültigen PIN erlangt

haben könnten. Genauer gesagt, bevor der Benutzer das System **10** durch Sprachbefehle programmieren kann, wird seine Stimme zuerst durch den Server **20** empfangen und dann durch das Sprechererkennungsmodul **22** geprüft und erkannt. Nach der Prüfung der Identität des Benutzers gibt der Server **20** die Programmieroberfläche **38** zur Programmierung des Systems **10** durch den Benutzer frei.

**[0059]** Die Sprachbefehle zur Programmierung des Systems **10** werden im ASR/NLU-Modul **24** verarbeitet. Insbesondere befindet sich das ASR/NLU-Modul **24** während einer solchen Programmierung in einem Befehls- und Steuermodus, wobei jede durch die Programmieroberfläche **38** empfangene Sprachanweisung bzw. jeder Sprachbefehl an das ASR/NLU-Modul **24** gesendet, in eine Symbolsprache umgesetzt und als Befehl gedeutet wird. Wenn der Benutzer beispielsweise möchte, dass das System **10** alle Anrufe von seiner Ehefrau auf seine Leitung legt, kann er z. B. anordnen „Alle Anrufe von meiner Frau Jane sofort durchstellen“, und das System **10** erkennt einen solchen Programmierbefehl und verarbeitet ihn entsprechend.

**[0060]** Darüber hinaus kann der Benutzer über das ASR/NLU-Modul **24** und das Sprachsynthesemodul **36** einen Dialog mit dem System **10** führen. Der Benutzer kann das aktuelle Programm überprüfen, indem er an die Programmieroberfläche **38** z. B. die Frage stellt: „Welche Anrufe werden an meinen Anrufbeantworter geleitet“. Diese Anfrage wird dann vom Server **20** (wenn der Benutzer das System **10** von außen anruft) oder von der Programmieroberfläche **28** über den Server **20** (wenn sich der Benutzer im Büro aufhält) an das ASR/NLU-Modul **24** gesendet, wo die Anfrage verarbeitet wird. Das ASR/NLU-Modul **24** erzeugt dann eine Antwort auf die Anfrage und sendet diese an das Sprachsynthesemodul **36**, um eine synthetisierte Nachricht zu erzeugen, die dann für den Benutzer wiedergegeben wird.

**[0061]** Desgleichen kann das ASR/NLU-Modul **24**, wenn das System **10** eine gesprochene Programmieranfrage von einem berechtigten Benutzer nicht verstehen kann, eine Aufforderung an den Benutzer richten, z. B. „Bitte formulieren Sie ihre Anfrage neu“, die dann vom Sprachsynthesemodul **36** verarbeitet wird. Genauer gesagt, der Server **20** sendet während einer solchen Programmierung eine Programmieranfrage an die Programmieroberfläche **38**. Wenn das System **10** die Anfrage nicht entschlüsseln kann, sendet die Programmieroberfläche **38** eine Fehlermeldung an den Server **20** zurück, der diese Meldung an das ASR/NLU-Modul **24** weiterleitet. Das ASR/NLU-Modul **24** kann dann die Abfrage nach einer anderen möglichen Bedeutung untersuchen oder den Benutzer (über das Sprachsynthesemodul **36**) auffordern, eine neue Programmieranfrage zu stel-

len.

**[0062]** Es sollte klar sein, dass das System **10** so programmiert werden kann, dass es durch das Audioindexierungs-/Prioritätsmodul **34** verschiedene Nachrichten und Anrufe verwalten kann, die über mit dem System **10** verbundene Voicemail-Systeme, Telefonleitungen, Fax-/Modemeinheiten, eMail-Programme und andere Datenübertragungseinheiten empfangen wurden. Insbesondere kann das Audioindexierungs-/Prioritätsmodul **34** so programmiert werden, dass es solche Nachrichten und Telefongespräche automatisch nach ihrem Thema und Inhalt oder ihrer Herkunft oder nach beidem sortiert und indiziert. Vorzugsweise kann das System **10** ferner so programmiert werden, dass es bestimmten Anrufen und Nachrichten von einer bestimmten Person eine Priorität zuweist.

**[0063]** Das Audioindexierungs-/Prioritätssystem gemäß [Fig. 2](#) funktioniert wie folgt. Nachdem der Anrufer durch das Sprechererkennungsmodul **22** geprüft und erkannt wurde, sendet dieses Modul Signale an das ID-Markierungsmodul **30**, das automatisch die Identität des Anrufers oder die Identität des aktuellen Sprechers einer Gruppe von Teilnehmern einer Telefonkonferenz markiert. Zeitgleich mit dem ID-Markierungsprozess setzt das Umsetzungsmodul **32** das Telefongespräch oder die Nachricht um. Zum Markierungsprozess gehört, dass die umgesetzte Nachricht der Identität des Sprechers oder Anrufers zugeordnet wird. Zum Beispiel wird während einer Telefonkonferenz jedes der dem aktuellen Sprecher entsprechende Segment des umgesetzten Gesprächs mit der Identität dieses Sprechers zusammen mit dem Anfangs- und Endzeitpunkt dieses Segments markiert.

**[0064]** Die im ID-Markierungsmodul **30** und im Umsetzungsmodul **32** verarbeiteten Informationen werden an das Audioindexierungs-/Prioritätsmodul **34** gesendet, in dem die empfangenen Informationen nach einer zuvor programmierten Prozedur verarbeitet und gespeichert werden. Das Audioindexierungs-/Prioritätsmodul **34** kann so programmiert werden, dass es die Nachrichten und Gespräche in einer vom Benutzer gewünschten Weise indiziert. Zum Beispiel kann der Benutzer in die Lage versetzt werden, entweder Nachrichten von einem bestimmten Anrufer oder alle dringenden Anrufe oder Nachrichten zu einem bestimmten Thema abzurufen. Ferner kann das Audioindexierungs-/Prioritätsmodul **34** so programmiert werden, dass es Anrufen von einem Anrufer eine Priorität zuweist, der entweder viele oder dringende Nachrichten hinterlegt hat.

**[0065]** Der Benutzer kann entweder an Ort und Stelle oder von einem fernen Standort auf die im Audioindexierungs-/Prioritätsmodul **34** gespeicherten Informationen zugreifen, um sie abzurufen. Wenn der Benutzer auf solche Informationen zugreift, sendet

das Audioindexierungs-/Prioritätsmodul **34** die angeforderten Informationen an das Sprachsynthesemodul **38**, in welchem eine Text-Sprache-Umsetzung erfolgt, damit der Benutzer die Nachricht in Form synthetisierter Sprach abhören kann. Es sollte klar sein, dass in der vorliegenden Erfindung jedes beliebige herkömmliche Sprachsyntheseverfahren verwendet werden kann, zum Beispiel die „Eloquent engine“ (Sprachmaschine), die mit der handelsüblichen Software VIAVOICEGOLD von IBM geliefert wird.

**[0066]** Es sollte klar sein, dass die Informationen unter Verwendung verschiedener Verfahren wie beispielsweise über eine GUI-Oberfläche, PINs und DTMF-Tasteneingaben vom Audioindexierungs-/Prioritätsmodul **34** abgerufen werden können. Das bevorzugte Verfahren in der vorliegenden Erfindung zum Abrufen solcher Informationen bedient sich jedoch sprachaktivierter Befehle. Unter Verwendung eines solchen Verfahrens kann das System **10** den Benutzer prüfen und erkennen, bevor dieser auf die im Audioindexierungs-/Prioritätsmodul **34** gespeicherten und indexierten Nachrichten oder Gespräche zugreifen kann. Das Audioindexierungs-/Prioritätsmodul **34** kann so programmiert werden, dass es bestimmte Sprachbefehle des Benutzers erkennt und beantwortet, die durch das ASR/NLU-Modul **24** und das Audioindexierungs-/Prioritätsmodul **34** verarbeitet werden, um bestimmte Nachrichten und Gespräche abzurufen. Zum Beispiel kann der Benutzer durch einen Sprachbefehl der Form „Alle Nachrichten von Herrn Smith wiedergeben“ alle Nachrichten von Herrn Smith abrufen, die im Audioindexierungs-/Prioritätsmodul **34** gespeichert sind. Dieser Befehl wird durch den Server **20** empfangen und zur Verarbeitung an das ASR/NLU-Modul **24** gesendet. Wenn das ASR/NLU-Modul **24** die Abfrage versteht, sendet es eine Antwort an den Server **20** zurück, um die Abfrage zu verarbeiten. Dann weist der Server **20** das Audioindexierungs-/Prioritätsmodul **34** an, die angeforderten Nachrichten an das Sprachsynthesemodul zu senden, um synthetisierte eMail- oder Fax-Nachrichten zu erzeugen, oder aufgezeichnete Telefon- oder Voicemail-Nachrichten direkt an den Server **20** zu senden, die einfach wiedergegeben werden.

**[0067]** Es sollte klar sein, dass ein Fachmann in der vorliegenden Erfindung verschiedene alternative Programmierstrategien zur Verarbeitung von Anrufen anwenden kann. Zum Beispiel kann das System **10** so programmiert werden, dass es den Benutzer bei einem wichtigen oder dringenden ankommenden Telefonanruf alarmiert. Genauer gesagt, das System **10** kann so programmiert werden, dass es den Benutzer auf einem Bildschirm benachrichtigt, damit der Benutzer selbst über die Verarbeitung des Anrufs entscheiden oder den Anruf einfach nach einer vorher programmierten Prozedur in der oben gezeigten Weise verarbeiten kann. Darüber hinaus kann das System **10** so programmiert werden, dass es einen drin-

genden oder wichtigen Anruf an das Signalgerät des Benutzers weiterleitet, wenn dieser nicht zu Hause ist oder sich nicht im Büro aufhält. Der Benutzer kann das System **10** auch so programmieren, dass es (nach dem Entgegennehmen eines ankommenden Telefonanrufs) eine Reihe von Rufnummern anwählt, an denen sich der Benutzer im Tagesverlauf aufhalten kann. Außerdem kann die Folge (d. h. die Liste) der vorher programmierten Rufnummern entsprechend dem letzten ermittelten Aufenthaltsort des Benutzers automatisch durch das System **10** aktualisiert werden. Wenn der Benutzer dies wünscht, können Personen, die das System **10** anrufen, auch auf eine solche Liste zugreifen, sodass solche Anrufer versuchen können, den Benutzer an einem der verschiedenen Standorte ihrer Wahl zu erreichen.

**[0068]** Außerdem sollte klar sein, dass das System **10** so programmiert werden kann, dass die Namen aller das System **10** anrufenden Personen (unter Verwendung der automatischen Rufnummererkennung ANI) zusammen mit ihren Rufnummern sowie die eMail-Adressen von Personen, die eine elektronische Nachricht gesendet haben, gespeichert werden. Dadurch kann der Benutzer des Systems **10** automatisch auf aktuelle Anrufe oder Nachrichten antworten, ohne vorher die Rufnummer oder die eMail-Adresse der Person zu ermitteln, der er antworten möchte. Ferner sorgt eine solche Programmierung für die dynamische Erstellung eines ständig aktuellen Adressbuchs, auf das ein berechtigter Benutzer zugreifen kann, um Nachrichten zu senden oder Anrufe zu tätigen. Genauer gesagt, der Benutzer kann auf das System **10** zugreifen, den Namen einer bestimmten anzurufenden Person auswählen und dann das System **10** anweisen, eine bestimmte Nachricht (z. B. eine eMail- oder eine Fax-Nachricht) an diese Person zu senden.

**[0069]** Außerdem kann das System **10** so programmiert werden, dass der Benutzer auf bestimmte Funktionen des Systems **10** zugreifen und diese nutzen kann. Zum Beispiel kann das System **10** dem Benutzer die Möglichkeit anbieten, einen beabsichtigten Termin mit dem Benutzer zu planen, der im System **10** gespeichert und dann vom Benutzer akzeptiert oder abgelehnt werden kann. Dem Benutzer kann auch die Gelegenheit gegeben werden, das Verfahren (z. B. per Telefonanruf, Fax- oder eMail-Nachricht) auszuwählen, mit dem er einen solchen Termin bestätigen, ablehnen oder verschieben kann. Außerdem kann das System **10** so programmiert werden, dass ein bestimmter berechtigter Anrufer auf den Terminkalender des Benutzers zugreifen darf, damit solche Termine unkompliziert geplant werden können.

**[0070]** Ferner sollte klar sein, dass die vorliegende Erfindung in kleinem Maßstab bei einer privat genutzten Anwendung oder in großem Maßstab bei einer Verwaltungs- oder Unternehmensanwendung einge-

setzt werden kann. Dem Fachmann sollte ferner klar sein, dass das System **10** auch bei anderen Anwendungen eingesetzt werden kann. Zum Beispiel kann das System **10** unter Verwendung des NLU-Merkmals mit Geräten wie Tonband-, Hörfunk- und Fernsehgeräten verbunden werden, um den Benutzer zu alarmieren, wenn auf einem Kanal über ein bestimmtes Thema berichtet oder wenn eine bestimmte Person interviewt wird. Es sollte klar sein, dass das System **10** nicht auf Telefonübertragungen beschränkt ist. Das System **10** kann für Web-Telefone, Internetgespräche, Telefonkonferenzen und diverse andere Datenübertragungen genutzt werden, bei denen Sprache über einen digitalen oder analogen Kanal übertragen wird. Auf dieselbe Weise können auch weitere elektronische Daten wie beispielsweise ASCII-Zeichen, Fax-Nachrichten und der Inhalt von Internetseiten und Datenbanksuchen verarbeitet werden. Durch die Verknüpfung der optischen Zeichenerkennung (OCR) mit Fax-Empfangsfunktionen kann das System **10** den Inhalt von Nachrichten umsetzen, die per Fax oder eMail empfangen und im Audioindexierungs-/Prioritätsmodul **34** gespeichert wurden. Wie oben gezeigt wurde, kann der Benutzer diese Nachrichten dann durch das Sprachsynthesemodul **36** abrufen, um den Inhalt solcher Nachrichten abzuhören.

**[0071]** Zusammenfassend ist zu sagen, dass die vorliegende Erfindung ein programmierbares System zur Verarbeitung von Anrufen und Nachrichten bereitstellt, das von einem Benutzer zur Verarbeitung ankommender Telefonanrufe, eMail-Nachrichten, Fax-Nachrichten und der Daten anderer elektronischer Daten auf eine vorgegebene Weise programmiert werden kann, ohne dass der Benutzer zuerst manuell einen Telefonanruf entgegennehmen oder eine eMail- oder Fax-Nachricht abrufen, den Anrufer oder den Urheber der Nachricht ermitteln und dann entscheiden muss, wie er einen solchen Anruf weiterleiten oder auf eine solche Nachricht antworten muss. Die vorliegende Erfindung kann so programmiert werden, dass Telefongespräche oder Telefonkonferenzen umgesetzt, die Identität des Anrufers oder der Teilnehmer der Telefonkonferenz markiert und solche Nachrichten und Gespräche entsprechend der Identität des Anrufers oder Urhebers und/oder dem Thema und dem Inhalt des Gesprächs oder der Nachricht gespeichert werden. Der Benutzer kann dann jede beliebige gespeicherte Nachricht oder jedes Gespräch auf der Grundlage der Identität des Anrufers oder eine Gruppe ähnlicher Nachrichten auf der Grundlage ihres Themas abrufen.

**[0072]** Die Erfindung kann ferner folgende Merkmale aufweisen:

Das Servermittel empfängt ferner und reagiert auf eine ankommende Fax-Nachricht und/oder eMail-Nachricht und/oder Sprachdaten und/oder in Text umsetzbare Daten.

**[0073]** Das Sprechererkennungsmodul beruht auf der textunabhängigen Sprechererkennung.

**[0074]** Das Sprechererkennungsmodul nutzt zur Ermittlung des Themas und des Inhalts des Anrufs die Spracherkennung und die Sprachenerkennung.

**[0075]** Das System beinhaltet ein Sprachenerkennungsmittel, das funktionell mit dem Sprechererkennungsmodul verbunden ist, um die Sprachen des ankommenden Anrufs zu erkennen und zu verstehen.

**[0076]** Das Erkennungsmittel führt die Sprachenübersetzung durch.

**[0077]** Die Identität des Anrufers wird aus der erkannten Sprache des Anrufs ermittelt.

**[0078]** Das Sprachenerkennungsmittel nutzt die Identität des Anrufers zur Erkennung der Sprache des Anrufs.

**[0079]** Das System enthält ein Anmeldeinstrument zum Anmelden eines neuen Benutzer, der auf das System zugreifen will.

**[0080]** Der neue Benutzer kann automatisch angemeldet werden.

**[0081]** Es wird ein Mittel zum Ermitteln eines Zeitpunkts des Anrufs bereitgestellt, wobei das System ferner so programmiert werden kann, dass der Anruf entsprechend dem Zeitpunkt des Anrufs verarbeitet wird.

**[0082]** Das Programmiermittel beinhaltet eine GUI-Oberfläche und/oder eine Sprachschnittstelle und/oder eine Programm-Konfigurationsdatei.

**[0083]** Das Programmieren kann an Ort und Stelle oder von einem fernen Standort oder von beiden erfolgen.

**[0084]** Es wird ein Mittel bereitgestellt, das als Reaktion auf den ankommenden Anruf dynamisch ein Adressbuch erstellt.

**[0085]** Es wird ein Mittel zum Zugreifen auf das Adressbuch bereitgestellt, um eine Nachricht an eine ausgewählte Person zu senden.

**[0086]** Die Verarbeitung des Anrufs beinhaltet das sequenzielle oder gleichzeitige Weiterleiten eines ankommenden Telefonanrufs an eine Vielzahl verschiedener Rufnummern.

**[0087]** Es wird ein Mittel bereitgestellt, um den Anrufer zur Angabe seiner Identität und des Themas des Anrufs aufzufordern. Diese Aufforderung kann erfolgen, wenn das System entweder die Identität

oder das Thema des Anrufs nicht ermitteln kann.

**[0088]** Alternativ erfolgt die Aufforderung, wenn der Anruf empfangen wird, um die Identität des Anrufers und das Thema des Anrufs zu ermitteln.

**[0089]** Ferner kann das System ein funktionell mit dem Umsetzungsmittel verbundenes Mittel umfassen, damit ein Benutzer des Systems Nachrichten diktieren und an eine ausgewählte Person senden kann. Die Nachricht kann per Fax und/oder eMail und/oder Telefonanruf an die ausgewählte Person gesendet werden.

**[0090]** Ferner kann das System ein Mittel umfassen, um in den Kopfdaten der Fax- und eMail-Nachricht Markierungen zur Stimmungslage oder zur Dringlichkeit/Vertraulichkeit anzubringen.

**[0091]** Der Schritt zur Ermittlung der Identität des Anrufers kann durch eine textabhängige Sprechererkennung ausgeführt werden.

**[0092]** Der Schritt zur Ermittlung des Themas des Anrufs kann durch eine Spracherkennung und eine Sprachenerkennung ausgeführt werden.

**[0093]** Das Verfahren kann den zur Übersetzung des Anrufs in eine von der Sprache des Anrufs verschiedene Sprache beinhalten.

**[0094]** Der ankommende Anruf kann aufgezeichnet werden.

**[0095]** Die Aufzeichnung erfolgt gleichzeitig mit dem Schritt zur Ermittlung der Identität des Anrufers und kann vor dem Schritt zur Ermittlung der Identität des Anrufers durchgeführt werden.

**[0096]** Das Verfahren kann ferner die folgenden Schritte umfassen: Ermitteln eines Zeitpunkts des Anrufs; und Verarbeiten des Anrufs auf der Grundlage des ermittelten Zeitpunkts des Anrufs.

**[0097]** Der Schritt zum Abrufen der indexierten Informationen wird durch Sprachbefehle ausgeführt.

**[0098]** Das Verfahren kann die Ermittlung des Zeitpunkts des Anrufs oder der Nachricht und die Verarbeitung des Anrufs oder der Nachricht entsprechend dem ermittelten Zeitpunkt beinhalten.

### Patentansprüche

1. System zur automatischen Verarbeitung von Anrufen und Datenübertragungen, wobei das System Folgendes umfasst:

ein Servermittel (**20**) zum Empfangen eines ankommenden Anrufs;

ein funktionell mit dem Servermittel verbundenes

Sprechererkennungsmittel (**22**) zum Erkennen eines Anrufers des Anrufs;

ein funktionell mit dem Servermittel verbundenes Spracherkennungsmittel (**24**) zum Ermitteln des Themas des Anrufs;

ein Vermittlungsmittel (**28**) zum Verarbeiten des Anrufs gemäß der Erkennung des Anrufers und dem ermittelten Thema als Reaktion auf das Sprechererkennungsmittel und das Spracherkennungsmittel;

ein funktionell mit dem Servermittel, dem Sprechererkennungsmittel und dem Spracherkennungsmittel verbundenes Programmiermittel (**38**), mit welchem ein Benutzer das System programmieren kann; und gekennzeichnet durch ein Identitätsmarkierungsmittel (**30**), das die Identität des Anrufers als Reaktion auf das Sprechererkennungsmittel markiert;

ein Umsetzungsmittel (transcribing means) (**32**), das ein Telefongespräch oder eine Nachricht des Anrufers als Reaktion auf das Spracherkennungsmittel umsetzt; und ein funktionell mit dem Identitätsmarkierungsmittel und dem Umsetzungsmittel verbundenes Audioindexierungsmittel (**34**) zum Indexieren der Nachrichten und der Gespräche des Anrufers gemäß dem Thema des Gesprächs und der Nachricht und der Identität des Anrufers.

2. System nach Anspruch 1, dadurch gekennzeichnet, dass das Servermittel ein Mittel zum Aufzeichnen (**40**) des ankommenden Anrufs beinhaltet.

3. Verfahren nach Anspruch 2, dadurch gekennzeichnet, dass das Servermittel ferner ein Mittel (**42**) zum Komprimieren und Speichern der aufgezeichneten Daten und ein Mittel zum Dekomprimieren der komprimierten Daten beinhaltet.

4. System nach Anspruch 1, 2 oder 3, das ferner durch ein Mittel (**118**) zum Abrufen der indexierten Nachrichten vom Audioindexierungsmittel gekennzeichnet ist.

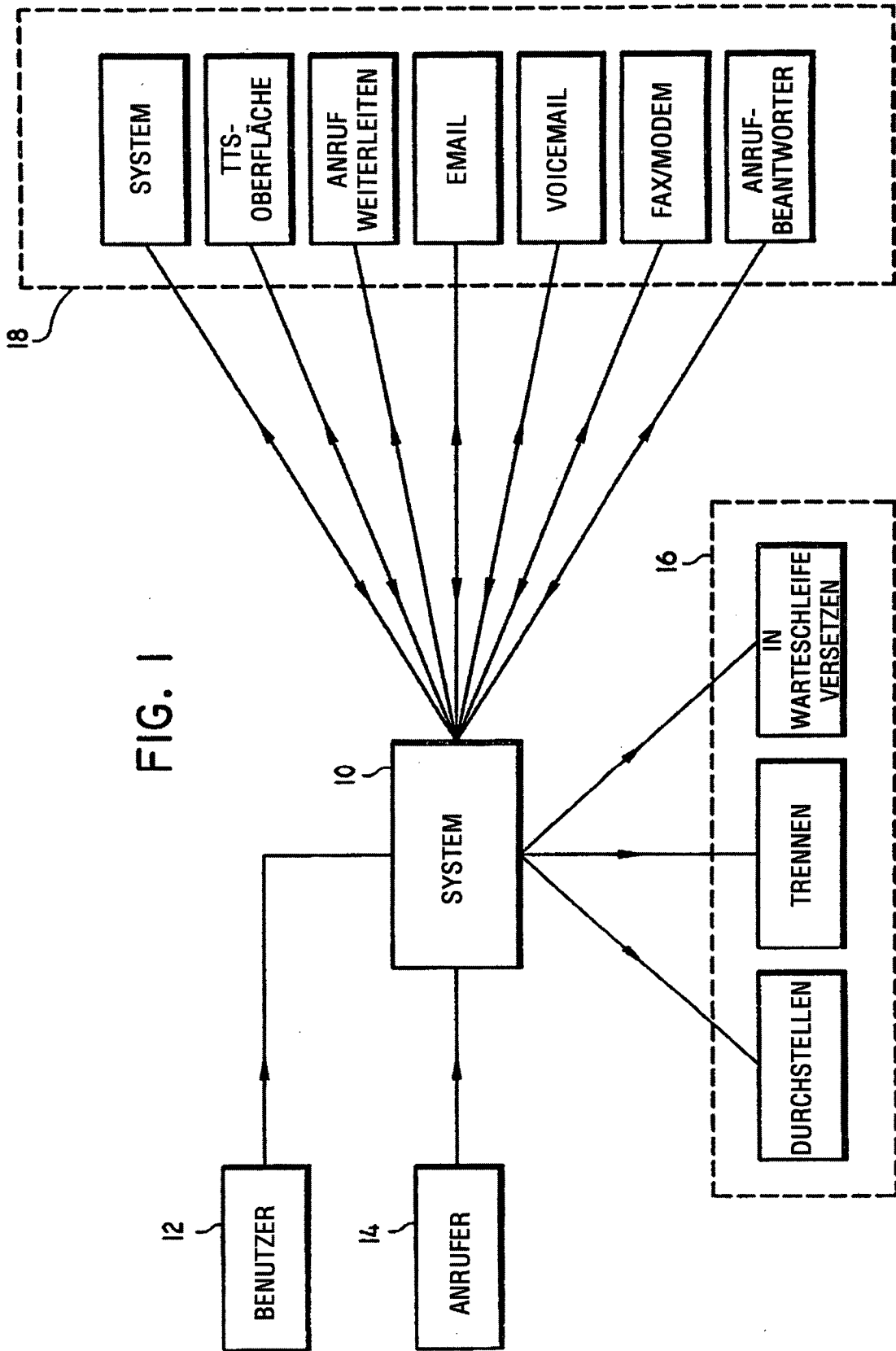
5. System nach Anspruch 1, 2, 3 oder 4, das ferner durch ein funktionell mit dem Servermittel, dem Spracherkennungsmittel und dem Audioindexierungsmittel verbundenes Sprachsynthesemittel (**36**) zum Umwandeln der im Audioindexierungsmittel gespeicherten Daten in synthetisierte Sprache gekennzeichnet ist.

6. Verfahren zum Bereitstellen einer Anrufverarbeitung, das die folgenden Schritte umfasst: Empfangen eines ankommenden Anrufs (**100**); Erkennen eines Anrufers des Anrufs, wenn ein ankommender Anruf empfangen wird (**130**), und Ermitteln (**160**) des Themas des Anrufs; Verarbeiten (**152, 154, 156, 158**) des Anrufs und der Nachricht gemäß der Identität des Anrufers und Urhebers bzw. dem Thema des Anrufs und der Nachricht; und gekennzeichnet durch die folgenden Schritte: Markieren der ermittelten Identität eines solchen Anrufers; Umsetzen des ermittelten

Themas eines solchen Anrufs; Indexieren der sich aus dem Markieren und dem Umsetzen ergebenden Daten gemäß dem ermittelten Thema und der ermittelten Identität.

7. Verfahren nach Anspruch 6, gekennzeichnet durch das Abrufen der indexierten Daten und das Umwandeln der indexierten Daten in synthetisierte Sprache.

Es folgen 4 Blatt Zeichnungen





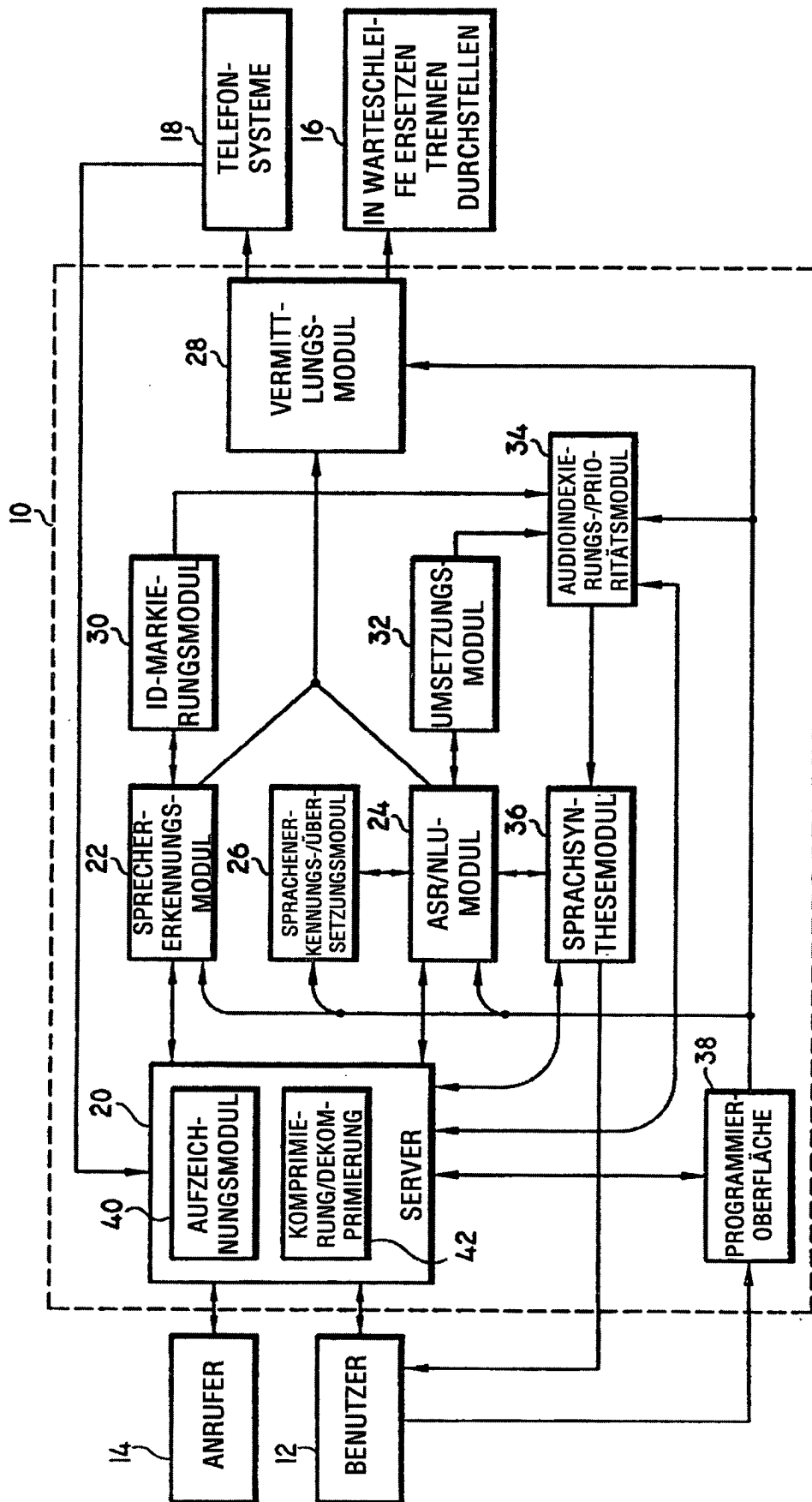


FIG. 2

FIG. 3A

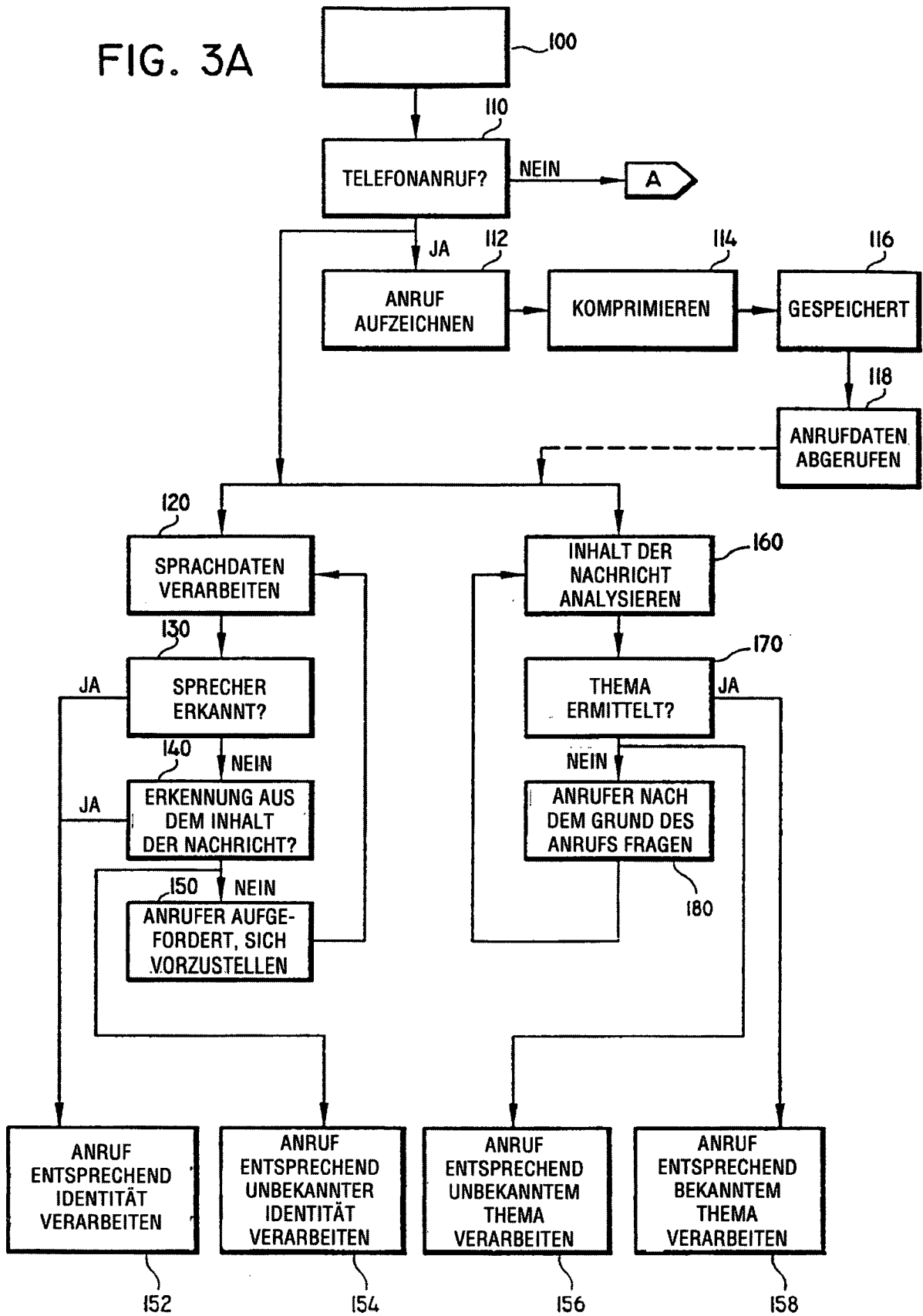


FIG. 3B

