(19) **Bundesrepublik Deutschland**
Deutsches Patent- und Markenamt

(10) **DE 699 31 783 T2** 2007.06.14

(12) # Übersetzung der europäischen Patentschrift

(97) **EP 1 094 446 B1**
(21) Deutsches Aktenzeichen: **699 31 783.5**
(96) Europäisches Aktenzeichen: **99 308 221.3**
(96) Europäischer Anmeldetag: **18.10.1999**
(97) Erstveröffentlichung durch das EPA: **25.04.2001**
(97) Veröffentlichungstag der Patenterteilung beim EPA: **07.06.2006**
(47) Veröffentlichungstag im Patentblatt: **14.06.2007**

(51) Int Cl.[8]: ***G10L 19/00*** (2006.01)

(73) Patentinhaber:
Lucent Technologies Inc., Murray Hill, N.J., US

(74) Vertreter:
Klunker, Schmitt-Nilson, Hirsch, 80797 München

(84) Benannte Vertragsstaaten:
DE, FR, GB

(72) Erfinder:
Rupinder, Judge, Bracknell, Berkshire RG42 4UE, GB

(54) Bezeichnung: **Verbesserung bei digitaler Kommunikationseinrichtung**

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] Diese Erfindung bezieht sich auf eine digitale Kommunikationsvorrichtung. Ein wünscheswertes Merkmal in einer digitalen Kommunikationsvorrichtung ist die Fähigkeit, Sprache aufzunehmen. In einem digitalen mobilen Telefon z.B. würde diese Fähigkeit dem Telefon ermöglichen, als ein Telefonanrufbeantworter zu fungieren, oder eine Sprachnotiz (Voice Memo) aufzunehmen, oder ein Gespräch aufzunehmen, das über das Telefon geführt wird. Eine wichtige Überlegung, insbesondere in einem digitalen mobilen Telefon, ist die Menge an Speicher, der für diesen Zweck zur Verfügung gestellt werden muss. Bisherige Ansätze zur Bereitstellung von Sprachaufzeichnungen in einem digitalen mobilen Telefon haben keinen effizienten Gebrauch von Speicherplatz gemacht, und verlangten deshalb entweder einen unakzeptablen großen Speicher oder benutzten einen kleinereren Speicher, welcher für manche Zwecke ausreichend war.

[0002] Das US-Patent 5 630 016 beschreibt ein digitales Kommunikationssystem für das Senden und Empfangen von Sprache. Der Sender sendet zwei Datenrahmen folgend auf eine Detektion von Stimmeninaktivität. Der Empfänger beinhaltet einen Generator für künstliches Geräusch, der die zwei Datenrahmen benutzt, um Geräusch an einen Lautsprecher während Stimmeninaktivität auszugeben.

[0003] Das IBM Technical Disclosure Bulletin, VOL.29, Nr. 4, 1. September 1986, mit dem Titel "Compression Method for Voice Preprocessing and Postprocessing", Seiten 1756 bis 1757, beschreibt ein Verfahren zur Komprimierung und Speicherung von Sprachdaten. Eine sogenannte Suprakomprimierung („super-compression") wird benutzt, um Gesprächspausenperioden in der Sprache zu komprimieren.

[0004] Gemäß einem Aspekt dieser Erfindung wird eine digitale Kommunikationsvorrichtung bereitgestellt, die einen Generator für künstliches Geräusch zur Bereitstellung von künstlichem Geräusch für eine Simulation von akustischem Hintergrundgeräusch und einen Sprachkodierer für eine Erzeugung von Sprachrahmen beinhaltet, dadurch gekennzeichnet dass die digitale Kommunikationsvorrichtung weiterhin eine Abschätzeinrichtung für künstliches Geräusch für eine Bereitstellung von Gesprächspausenrahmen aufweist, wobei die Gesprächspausenrahmen eine Information bezeichnend für akustisches Hintergrundgeräusch aufweisen; und eine Sprachaufnahme/-wiedergabe-Einrichtung, wobei die Sprachaufnahme/-wiedergabe-Einrichtung angepasst ist, bei Aufnahme Sprachrahmen während des Vorhandenseins von Sprache zu speichern, einen oder mehrere Gesprächspausenrahmen an dem Ende des Vorhandenseins von Sprache während dem Ausbleiben von Sprache zu speichern, und Daten zu speichern, die bezeichnend sind für die Dauer des Ausbleibens von Sprache; und wobei die Sprachaufnahme/-wiedergabe-Einrichtung angepasst ist, bei Wiedergabe als Ausgabe Sprachsignale bereitzustellen, die abgeleitet sind von den gespeicherten Sprachrahmen, und, in Abhängigkeit von dem gespeicherten einen oder mehreren Gesprächspausenrahmen, künstliches Geräusch von dem Generator für künstliches Geräusch für eine Dauer repräsentiert durch die gespeicherten Daten.

[0005] Die Vorrichtung kann ein digitales mobiles Telefon sein, das einen Sender beinhaltet, der angepasst ist, nur für das Senden von Rahmen eingeschaltet zu werden, die nützliche Information enthalten, und wobei der Generator für künstliches Geräusch angepasst ist, künstliches Geräusch in Abhängigkeit von Gesprächspausenrahmen bereitzustellen, wenn keine Sprachrahmen empfangen werden.

[0006] Die Vorrichtung kann einen Sprachkodierer für das Kodieren von Sprache in Sprachrahmen und für Bereitstellung einer Eingabe an die Abschätzeinrichtung für künstliches Geräusch für Bereitstellung der Gesprächspausenrahmen beinhalten, wodurch die Sprachrahmen, die durch die Sprachaufnahme/-wiedergabe-Einrichtung gespeichert werden, jene sind, die durch den Sprachkodierer kodiert werden, und einen Sprachdekodierer für eine Dekodierung von gespeicherten Sprachrahmen und für eine Bereitstellung von künstlichem Geräusch in der Ausgabe des Sprachdekodierers, das durch den Generator für künstliches Geräusch in Abhängigkeit von gespeicherten Gesprächspausenrahmen erzeugt wurde.

[0007] Das digitale mobile Telefon kann ein GSM-Telefon sein.

[0008] Die gespeicherten Daten können eine Anzahl von Sprachrahmen umfassen, die während des Ausbleibens von Sprache auftreten.

[0009] Die gespeicherten Daten können einen jeweiligen Leerrahmen einer minimalen Länge für jeden Sprachrahmen umfassen, der während des Ausbleibens von Sprache auftritt.

[0010] Gemäß einem anderen Aspekt dieser Erfindung wird ein Verfahren zur Aufnahme/Wiedergabe von Sprache in einer digitalen Kommunikationsvorrichtung bereitgestellt, wobei die Kommunikationsvorrichtung einen Sprachkodierer zur Erzeugung von Sprachrahmen und einen Generator für künstliches Geräusch zur Bereitstellung von künstlichem Geräusch für eine Simulation von akustischem Hintergrundgeräusch beinhaltet, dadurch gekennzeichnet, dass die digitale Kommunikationsvorrichtung weiter-

hin eine Abschätzeinrichtung für künstliches Geräusch und eine Sprachaufnahme/-wiedergabe-Einrichtung aufweist, wobei die Abschätzeinrichtung für künstliches Geräusch zur Bereitsstellung von Gesprächspausenrahmen dient, wobei die Gesprächspausenrahmen Information bezeichnend für akustisches Hintergrundgeräusch umfassen, wobei das Verfahren für eine Aufnahme beinhaltet das Speichern von Sprachrahmen während des Vorhandenseins von Sprache durch die Sprachaufnahme/-wiedergabe-Einrichtung, das Speichern eines oder mehrerer Gesprächspausenrahmen durch die Sprachaufnahme/-wiedergabe-Einrichtung, die Information bezeichnend für akustisches Hintergrundgeräusch beinhalten, an dem Ende des Vorhandenseins von Sprache während des Ausbleibens von Sprache, und das Speichern von Daten durch die Sprachaufnahme/-wiedergabe-Einrichtung, die bezeichnend sind für die Dauer des Ausbleibens von Sprache, und für die Wiedergabe das Bereitstellen von Sprachsignalen durch die Sprachaufnahme/-wiedergabe-Einrichtung als Ausgabe, die von den gespeicherten Sprachrahmen abgeleitet sind, und von künstlichem Geräusch in Abhängigkeit von dem gespeicherten Gesprächspausenrahmen oder den gespeicherten Gesprächspausenrahmen für das Simulieren von akustischem Hintergrundgeräusch für eine Dauer repräsentiert durch die gespeicherten Daten.

[0011] Die gespeicherten Daten können eine Anzahl von Sprachrahmen aufweisen, die während des Ausbleibens von Sprache auftreten.

[0012] Die gespeicherten Daten können einen jeweiligen Leerrahmen einer minimalen Länge für jeden Sprachrahmen umfassen, der während des Ausbleibens von Sprache auftritt.

[0013] Die Erfindung wird nun im Wege eines Beispiels mit Bezug auf die beiliegenden Zeichnungen beschrieben werden, in welchen:

[0014] Fig. 1 eine schematische Darstellung eines GSM-Telefons ist, das die Erfindung beinhaltet; und

[0015] Fig. 2 und Fig. 3 schematische Darstellungen sind, die detaillierter die Audio-Kodierfunktion („Audio Encode function") bzw. die Audio-Dekodierfunktion („Audio Decode function") darstellen, die in Fig. 1 gezeigt sind.

[0016] Während einer normalen Telefonkonversation wechseln sich die Teilnehmer ab, so dass im Durchschnitt jede Senderichtung mit ungefähr 50% der Zeit belegt ist. In dem GSM-System wird die sogenannte "diskontinuierliche Sendung" ("discontinuous transmission", DTX) als eine Betriebsart spezifiziert, wonach der Sender nur für jene Rahmen eingeschaltet wird, welche Sprache oder eine andere nützliche Information enthalten. Dies hat die Vorteile, dass das durchschnittliche Interferenzniveau "auf Sendung" reduziert wird, was zu besserer Spektrumeffizienz führt, und wobei in dem mobilen Telefon die Batterielebensdauer verlängert wird oder eine kleinere Batterie für eine gegebene Betriebsdauer benutzt werden kann. Die Ausführung der diskontinuierlichen Sendebetriebsart erfordert einen Stimmenaktivitätsdetektor auf der Sendeseite, um ein Ausbleiben von Sprache zu detektieren, eine Bewertung des akustischen Hintergrundgeräuschs auf der Sendeseite, um charakteristische Parameter des Hintergrundgeräuschs zu der Empfängerseite zu senden, und eine Erzeugung von einem ähnlichen Geräusch auf der Empfängerseite, d.h. sogenanntes künstliches Geräusch, während Perioden, in denen die Funkübertragung unterbrochen ist.

[0017] Das Aussenden von Charakteristiken bezüglich des künstlichen Geräuschs zu der Empfängerseite wird mittels eines speziellen Rahmens erreicht, ein sogenannter Gesprächspausen-Deskriptorrahmen („silence descriptor frame") oder SID-Rahmen. Dieser Rahmen wird ausgesendet an dem Ende einer jeden Sprach-Signalfolge (burst) und dient als ein "Ende von Sprache"-Marker für die Empfängerseite. Um die Charakteristiken bezüglich des künstlichen Geräuschs auf der Empfängerseite zu aktualisieren, werden SID-Rahmen ebenso an regulären Intervallen während Sprachpausen ausgesendet. Dies dient ebenfalls dem Zweck der Verbesserung der Messung der Funkverbindungsqualität durch das Funkgerätsubsystem.

[0018] In seiner Anwendung auf ein GSM-Telefon macht die vorliegende Erfindung von dem Stimmenaktivitätsdetektor, dem Generator für künstliches Geräusch und von verschiedenen anderen Komponenten Gebrauch, welche bereits in dem Telefon für den Zweck der diskontinuierlichen Sendebetriebsart, wie oben beschrieben, vorhanden sind.

[0019] Nun bezugnehmend auf Fig. 1 werden Sprachsignale von einem normalerweise vorgesehenen Mikrofon **101** in einem GSM-Handapparat **100** über einen Filter **102** zu einem Analog-Digital-Wandler **103** geleitet, wobei die digitale Ausgabe desselben an einen Audio-Kodierungs-Funktionsblock **104** bereitgestellt wird. Die Ausgabe von Block **104**, dessen Funktionen detallierter mit Bezug auf Fig. 2 beschrieben werden, wird als Eingabe an einen Sendeketten-Block **105** bereitgestellt, in dem die gebräuchliche Kanalkodierung, Verschränkung und Modulation bewirkt werden. Die Ausgabe von Block **105** wird als Eingabe an einen Funkvorrichtungs-Block **106** bereitgestellt, der die gewöhnlichen Funkfrequenz(HF)-Gruppen wie beispielsweise den Leistungsverstärker und die Antenne für eine Translation zu einem Funkfrequenz(HF)-Signal für das Aussenden aufweist. Empfangene Signale an dem Ausgang des Funkvorrichtungs-Blocks **106** werden als Einga-

be zu einem Empfangsketten-Block **107** gespeist, in dem die gewöhnlichen Entzerrungs- und Kanaldekodierungsoperationen durchgeführt werden, um Sprachdaten und verschiedene Hinweisbits (Flags) zu erzeugen, die hinweisend auf die Qualität des empfangenen Signals sind (d.h., ob es irgendwelche Fehler in den Sprach- und SID-Rahmen gibt und ob die Sprachdaten ein Sprachrahmen oder ein SID-Rahmen sind). Die Sprachdaten und Hinweisbits werden als Eingabe an einen Audio-Dekodierungs-Funktionsblock **108** geleitet. Die Ausgabe des Blocks **108**, dessen Funktionen noch mit Bezug auf Fig. 3 beschrieben werden, wird über einen Schalter **109** als Eingabe an einen Digital-Analog-Wandler **110** geleitet, dessen analoge Ausgabe über einen Filter **111** an einen elektroakustischen Signalgeber **112** gespeist wird, wie beispielsweise eine Hörkapsel oder ein Lautsprecher.

[0020] Die Eingabe an den Audio-Kodierungs-Funktionsblock **104** und die Ausgabe von dem Audio-Dekodierungs-Funktionsblock **108** werden außerdem als Eingaben zu einem Summierer **113** und zu einem Schalter **114** gespeist. Die Ausgabe von dem Summierer **113** wird außerdem als Eingabe an den Schalter **114** gespeist. Die Ausgabe von Schalter **114** wird als Eingabe zu einem Audio-Kodierungs-Funktionsblock **104'** gespeist, welcher eine zweite Instanz des bereits vorgesehenen Audio-Kodierungs-Funktionsblocks **104** sein kann und vorzugsweise ist, welcher, wie es für einen Fachmann offensichtlich ist, normalerweise in Software in dem digitalen Signalprozessor oder DSP (nicht gezeigt) in dem GSM-Handapparat eingebunden ist.

[0021] Nun bezugnehmend auf Fig. 2 wird die Eingabe an den Audio-Kodierungs-Funktionsblock **104'** zu einem Sprachkodierer **201** gespeist, welcher die Sprachsignale in Sprachrahmen kodiert. Signale von dem Sprachkodierer **201** werden zu einem Stimmenaktivitätsdetektor **202** gespeist, welcher ein Signal erzeugt, ein sogenanntes VAD-Flag, in Antwort auf das Ausbleiben von Sprache. Signale von dem Sprachkodierer **201** werden außerdem zu einer Abschätzeinheit für künstliches Geräusch **203** gespeist, welche SID-Rahmen erzeugt, welche Information hinweisend auf das akustische Hintergrundgeräusch enthalten.

[0022] Das VAD-Flag vom Stimmenaktivitätsdetektor **202**, Sprachrahmen von dem Sprachkodierer **201** und SID-Rahmen von der Abschätzeinheit für künstliches Geräusch **203** werden zu einem Sende-DTX- und Steuer-Block **204** gespeist, welcher an seinem Ausgang Sprachdaten und ein Sprach- oder SP-Flag bereitstellt, das darauf hinweist, ob die Sprachdatenausgabe ein Sprachrahmen oder ein SID-Rahmen ist. Bis der Stimmenaktivitätsdetektor **202** das Ausbleiben von Sprache detektiert, umfasst die Sprachdatenausgabe Sprachrahmen, wobei das SP-Flag darauf hinweist. An dem Ende einer Sprachsignalfolge braucht es typischerweise eine Zeit, die gleich ist zu N-Sprachrahmen (N>1), um einen neuen SID-Rahmen zu berechnen. Wenn daher der Stimmenaktivitätsdetektor **202** das Ausbleiben von Sprache detektiert, enthält die Ausgabe von Block **204** weiterhin Sprachrahmen, bis der neue SID-Rahmen verfügbar wird, woraufhin er an dem Ausgang von Block **204** auftritt und das SP-Flag wechselt, um diesen anzuzeigen. Es tritt eine Ausnahme davon auf, wenn am Ende einer Sprachsignalfolge weniger als M-Rahmen (M>1) verstrichen sind, seitdem der letzte SID-Rahmen berechnet wurde. In dieser Situation wird der letzte SID-Rahmen vorgesehen und das SP-Flag wechselt, um diesen anzuzeigen. Im Ergebnis wird, falls der letzte SID-Rahmen nicht zu alt ist, dieser behandelt, als ob er aktuell ist, und er wird deshalb benutzt, als dass auf einen neuen zu berechnenden gewartet wird. Dieses Merkmal vermeidet das Warten, welches andererseits auftreten würde für einen neuen zu berechnenden SID-Rahmen in dem Fall, in dem eine kurze Hintergrundgeräuschspitze für Sprache gehalten wird, so dass Sendeaktivität reduziert wird.

[0023] Zurückkommend wiederum auf Fig. 1 ist der Betrieb des Audio-Kodierungs-Funktionsblocks **104** ähnlich zu dem Betrieb, wie eben für Block **104'** beschrieben. Wie sich der Fachmann bewusst sein wird, kann der Block **104** mit halber Frequenz, voller Frequenz oder erhöhter voller Frequenz betrieben werden, abhängig von betrieblichen Umgebungsbedingungen oder Anforderungen, wohingegen Block **104'** normalerweise mit voller Frequenz betrieben wird. Block **104** dient dazu, für eine Übertragung zu einem entfernten Empfänger über Sendeketten-Block **105** und Funkvorrichtungsblock **106** Sprachrahmen und SID-Rahmen und ein SP-Flag bereitzustellen, das anzeigt, ob die Rahmen Sprache oder SID sind.

[0024] Für einen Aufnahmebetrieb wird die Sprachdaten- und SP-Flag-Ausgabe von Block **104'** an einen Aufnahme- und Wiedergabesteuerblock **115** bereitgestellt, welcher dazu dient, zu entscheiden, was in einem Speicher **116** über einen Mikrokontroller **117** zu speichern ist. Der Aufnahme-und-Wiedergabesteuerblock **115** veranlasst, dass in dem Speicher **116** alle Rahmen zu speichern sind, welche durch das SP-Flag als Sprachrahmen angezeigt werden. Wenn das SP-Flag einen SID-Rahmen anzeigt, wird der erste solche SID-Rahmen ebenfalls im Speicher **116** gespeichert. Danach besteht keine Notwendigkeit, weitere Rahmen zu speichern, bis das SP-Flag wiederum einen Sprachrahmen anzeigt. Wenn eine solche Anzeige eines Sprachrahmens empfangen wird, veranlasst Block **115**, dass Daten im Speicher **116** zu speichern sind, die bezeichnend sind für die Dauer des zwischenzeitlichen Ausbleibens von Sprache. Vorzugsweise sind solche Daten eine Anzahl

von Sprachrahmen, die nicht im Speicher **116** gespeichert wurden, d.h. einer Anzahl oder eines Zählerwerts von Sprachrahmen, in welchen Sprache nicht vorhanden war. Als eine Alternative zu der Speicherung einer solchen Anzahl können die gespeicherten Daten andere Formen annehmen, wie beispielsweise einen Leerrahmen (dummy frame) einer minimalen Länge für jeden ungespeicherten Rahmen. Man wird sich dabei bewusst sein, dass die Speicherung von Sprachrahmen nur dann, wenn Sprache vorhanden ist, zu einem sehr viel effizienteren Gebrauch von Speicherplatz führt.

[0025] Während langer Perioden des Ausbleibens von Sprache kann es wünschenswert sein, im Speicher **116** periodisch aktualisierte SID-Rahmen zu speichern, wobei in diesem Fall die Daten, die bezeichnend sind für die Dauer des Ausbleibens von Sprache, wie z.B. der oben erwähnte Rahmenzählwert, eine Anzahl oder ein Zählwert sein würden, von dem ersten gespeicherten SID-Rahmen zum zweiten SID-Rahmen, eine Anzahl oder ein Zählwert von dem zweiten gespeicherten SID-Rahmen zu dem dritten SID-Rahmen, usw., endend mit einer finalen gespeicherten Anzahl oder einem finalen gespeicherten Zählwert von dem letzten gespeicherten SID-Rahmen zu dem Beginn von Sprache.

[0026] Für den Wiedergabebetrieb veranlasst der Aufnahme-und-Wiedergabesteuerblock **115**, dass die Inhalte von Speicher **116** zum Audio-Dekodierungsblock **108'** zu senden sind. Deshalb wird, wenn ein Sprachrahmen von Speicher **116** abgerufen wird, dieser zum Block **108'** gesendet, wobei die Flags derart gesetzt sind, dass sie auf Sprache mit keinen Fehlern hinweisen. (An dem Ausgang des Empfangskettenblocks **107** ist Information betreffend das Vorhandensein oder Nichtvorhandensein von Fehlern in den empfangenen Signalen relevant für Dekodierungszwecke im Block **108**, jedoch ist für die Wiedergabefunktion solch eine Fehlerinformation irrelevant, so dass für Block **108'** nur das Sprach-/SID-Feld benutzt wird, wobei die Flags derart gesetzt sind, dass sie das Nichtvorhandensein von Fehlern anzeigen). Wenn ein SID-Rahmen vom Speicher **116** abgerufen wird, wird er zu dem Audio-Dekodierungs-Block **108'** gesendet, wobei Flags so gesetzt sind, dass sie einen SID-Rahmen mit keinen Fehlern anzeigen. Der gleiche SID-Rahmen wird wiederholt zum Block **108'** gesendet, für eine Periode, die festgelegt ist durch die Daten, die im Speicher **116** gespeichert sind, und die bezeichnend sind für die Dauer des Ausbleibens von Sprache, wobei diese gespeicherten Daten verändert werden, um die verbleibende Dauer des Ausbleibens von Sprache zu reflektieren. Aus diesem Grund wird, wo solche Daten ein Rahmenzählwert sind, der Zählwert jedesmal um 1 reduziert, wenn ein SID-Rahmen zum Block **108'** gesendet wird. Wo, wie oben beschrieben, während eines Sprachausbleibens mehr als ein SID-Rahmen im Speicher **116** gespeichert ist, wird der erste SID-Rahmen gesendet, bis der erste gespeicherte Zählwert auf 0 reduziert ist, dann wird der zweite gespeicherte SID-Rahmen gesendet, bis der zweite gespeicherte Zählwert auf 0 reduziert wird, usw.

[0027] Nunmehr bezugnehmend auf Fig. 3 kann der Audio-Dekodierungs-Funktionsblock **108'** eine zweite Instanz des bereits vorgesehenen Audio-Dekodierungs-Funktionsblocks **108** (Fig. 1) sein, und ist es vorzugsweise, welcher, wie es vom Fachmann verstanden wird, normalerweise in Software in dem DSP (nicht gezeigt) in dem GSM-Handapparat eingebettet ist. Demgemäß wird, obwohl die Beschreibung des Betriebs mit Bezug auf Block **108'** vorgenommen wird, vorausgesetzt, dass Block **108** (Fig. 1) in einer ähnlichen Weise fungiert.

[0028] Sprachdaten und Flags an dem Eingang von Block **108'** werden als Eingabe an einen Empfangs-DTX- und Steuer-Funktionsblock **301** bereitgestellt, welcher Sprachrahmen, die als solches durch die Flags angezeigt werden, direkt an einen Sprachdekodierer **302** leitet, dessen Ausgabe die Ausgabe von Block **108'** ist. Der oder jeder SID-Rahmen, der als solches durch die Flags angezeigt wird, wird zu einem Generator für künstliches Geräusch **303** gesendet, welcher künstliches Geräusch generiert, dessen Charakteristiken abhängen von Information, die in dem SID-Rahmen oder den SID-Rahmen enthalten ist. Das künstliche Geräusch wird an den Sprachdekodierer **302** bereitgestellt für die Dauer des Ausbleibens von Sprache, repräsentiert durch die gespeicherten Daten im Speicher **116** (Fig. 1), beispielsweise den Zählwert der Rahmen, während welcher Sprache nicht vorhanden war.

[0029] Deshalb umfasst die Ausgabe von Block **108'** die dekodierten Sprachrahmen, die in Speicher **116** (Fig. 1) gespeichert sind, wobei jedes Gesprächspausenintervall zwischen Sprachsignalfolgen künstliches Geräusch enthält, das durch den Generator für künstliches Geräusch **313** erzeugt wird, in Abhängigkeit von dem SID-Rahmen oder den SID-Rahmen, wie im Speicher **116** gespeichert, für eine Dauer repräsentiert durch die Daten, beispielsweise den Rahmenzählwert, wie im Speicher **116** gespeichert, der die Dauer des Gesprächspausenintervalls bezeichnet. Der Block **108'** weist außerdem einen Sprachrahmen-Substitutions-Block **304** auf, welcher nicht in der Wiedergabefunktion benutzt wird. Er ist lediglich eine Entsprechung eines entsprechenden Blocks in der ersten Instanz des Audio-Dekodierungs-Blocks **108** (Fig. 1), wo er dazu dient, einen oder mehrere Substitutionssprachrahmen einzufügen, oder eine "stumme" Ausgabe bereitzustellen, falls schlechte Daten von dem Empfangskettenblock **107** empfangen werden.

[0030] Wiederum bezugnehmend auf Fig. 1 wird

die Ausgabe von Block **108'** zum Digital-Analog-Wandler **110** über den Schalter **109** gespeist. Es versteht sich, dass für einen normalen Betrieb als ein Telefon der Schalter **109** den Ausgang des Audio-Dekodierungsblocks **108** mit dem Eingang des Digital-Analog-Wandlers **110** verbindet, wohingegen beim Abhören einer Aufnahme der Schalter **109** den Ausgang des Audio-Dekodierungs-Funktionsblocks **108'** mit dem Eingang des Wandlers **110** verbindet.

[0031] Es versteht sich, dass das Setzen des Schalters **114** steuert, was aufgenommen wird, deshalb kann, durch Setzen des Schalters **114**, um den Ausgang des Analog-Digital-Wandlers **103** mit dem Eingang des Audio-Kodierungsblocks **104'** zu verbinden, der Handapparat eine Stimmennotiz aufnehmen, oder abgehende Sprache während eines Telefongesprächs aufnehmen, oder eine Stimmenaufforderung für eine Aufforderung des Nutzers des Handgeräts aufnehmen, beispielsweise eine Aufforderung wie etwa "bitte sprechen Sie den Namen der Person, die Sie nun gerne anrufen möchten", oder "sprechen Sie 'Hilfe' für weitere Optionen". Wenn der Schalter derart gesetzt ist, eine Ausgabe von dem Audio-Dekodierungsblock **108** an dem Eingang des Audio-Kodierungsblocks **104'** bereitzustellen, kann der Handapparat einfließende Sprache während eines Telefongesprächs aufnehmen oder wenn er als ein Telefonanrufbeantworter fungiert. Wenn der Schalter **114** den Ausgang des Summierers **113** mit dem Eingang des Audio-Kodierungsblocks **104'** verbindet, ist eine Aufnahme einer Telefonkonversation möglich, d.h., sowohl einfließender als auch abgehender Sprache.

[0032] Wenn die Erfindung angewandt wird auf einen Funk-Sende-Empfänger, der eine diskontinuierliche Übertragung verwendet, können verschiedene Hardware- oder Softwareelemente "wiederverwendet" werden, die bereits in dem Sende-Empfänger existieren. Demnach existieren bereits in einem GSM-Handapparat alle Elemente, die in Fig. 1 dargestellt sind, mit der Ausnahme des Summierers **113**, der Schalter **109** und **114**, und des Aufnahme- und Wiedergabe-Steuer-Blocks **115**, für die normalen Sende- und Empfangsoperationen des Handapparats, so dass die Erfindung insbesondere hinsichtlich ihrer Wiederverwendung von solcher Hardware und/oder Software effizient ist.

[0033] Der Fachmann ist sich dabei bewusst, dass die verschiedenen Elemente, die in Fig. 1 gezeigt sind, in Hardware oder Software implementiert werden können, wie jeweils anwendbar. Normalerweise werden das Mikrofon **101**, der Filter **102**, der Analog-Digital-Wandler **103**, die Funkvorrichtung **106**, die Modulationsfunktion in der Sendekette **105**, der Digital-Analog-Wandler **110**, der Filter **111**, der Signalgeber **112** und der Speicher **116** in Hardware implementiert. Andere Funktionen in der Sendekette **105**, der Empfangskette **107**, Audio-Kodierungsfunktionen **104** und **104'** und Audio-Dekodierungsfunktionen **108** und **108'** werden normalerweise in Software in dem DSP in dem Handapparat implementiert. Der Aufnahme- und Wiedergabe-Steuer-Block **115** kann in Software in dem DSP implementiert werden, oder in dem Mikrokontroller **117**, oder in Hardware.

[0034] Obwohl die Erfindung im Wege ihrer Anwendung auf ein GSM-Telefon beschrieben wurde, versteht es sich, dass sie in anderen Systemen Anwendung finden wird, die diskontinuierliche Übertragung verwenden, wie z.B. PCS 1900 und IS-136, beispielsweise, und in Systemen, in denen Übertragung mit einer niedrigeren Rate während Gesprächspausenperioden auftritt, wie es z.B. in CDMA-Systemen auftreten kann. Die Erfindung kann ebenso Anwendung finden in dem sogenannten Voice over IP (Internet), wo ein GSM EFR (Enhanced Full Rate) Sprachkodierer benutzt werden kann, um Paketdaten über das Internet zu senden, oder ein GPRS-Anschluss, der in der Lage ist, als ein GSM-Anschluss sowie als ein Daten-Anschluss zu fungieren, kann benutzt werden, um Paketdaten, die Sprachrahmen beinhalten, zu transferieren, wobei ein GSM-Sprachkodierer benutzt wird, die Sprache zu kodieren.

Patentansprüche

1. Digitale Kommunikationsvorrichtung, beinhaltend einen Generator für künstliches Geräusch zur Bereitstellung eines künstlichen Geräuschs für Simulation eines akustischen Hintergrundgeräuschs und einen Sprachkodierer zur Erzeugung von Sprachrahmen, **dadurch gekennzeichnet**, dass
die digitale Kommunikationsvorrichtung weiterhin eine Abschätzeinrichtung für künstliches Geräusch umfasst zur Bereitstellung von Gesprächspausenrahmen, wobei die Gesprächspausenrahmen Information umfassen, die bezeichnend ist für akustisches Hintergrundgeräusch; und eine Sprachaufnahme/-wiedergabe-Einrichtung;
wobei die Sprachaufnahme/-wiedergabe-Einrichtung angepasst ist, bei Aufnahme Sprachrahmen während des Vorhandenseins von Sprache zu speichern, einen oder mehrere Gesprächspausenrahmen an dem Ende des Vorhandenseins von Sprache während dem Ausbleiben von Sprache zu speichern, und Daten bezeichnend für die Dauer des Ausbleibens von Sprache zu speichern; und
wobei die Sprachaufnahme/-wiedergabe-Einrichtung angepasst ist, bei Wiedergabe als Ausgabe Sprachsignale, abgeleitet aus den gespeicherten Sprachrahmen, bereitzustellen und, in Abhängigkeit von dem gespeicherten einen oder mehreren Gesprächspausenrahmen, künstliches Geräusch von dem Generator für künstliches Geräusch für eine Dauer repräsentiert durch die gespeicherten Daten.

2. Vorrichtung nach Anspruch 1, wobei die Vorrichtung ein digitales mobiles Telefon ist, beinhaltend

einen Sender, der angepasst ist, nur für Senden von Rahmen eingeschaltet zu werden, die nützliche Information enthalten, und wobei der Generator für künstliches Geräusch angepasst ist, künstliches Geräusch in Abhängigkeit von Gesprächspausenrahmen bereitzustellen, wenn keine Sprachrahmen empfangen werden.

3. Vorrichtung nach Anspruch 2, beinhaltend einen Sprachkodierer für eine Kodierung von Sprache in Sprachrahmen und für eine Bereitsstellung einer Eingabe an die Abschätzeinrichtung für künstliches Geräusch für die Bereitstellung der Gesprächspausenrahmen, wodurch die Sprachrahmen, die durch die Sprachaufnahme/-wiedergabe-Einrichtung gespeichert werden, jene sind, die durch den Sprachkodierer kodiert werden, und einen Sprachdekodierer für eine Dekodierung von gespeicherten Sprachrahmen und für eine Bereitstellung von künstlichem Geräusch in der Ausgabe des Sprachdekodierers, das durch den Generator für künstliches Geräusch in Abhängigkeit von gespeicherten Gesprächspausenrahmen erzeugt wird.

4. Vorrichtung nach Anspruch 2 oder 3, wobei das digitale mobile Telefon ein GSM-Telefon ist.

5. Vorrichtung nach Anspruch 1, 2, 3 oder 4, wobei die gespeicherten Daten eine Anzahl von Sprachrahmen umfassen, die während des Ausbleibens von Sprache auftreten.

6. Vorrichtung nach Anspruch 1, 2, 3 oder 4, wobei die gespeicherten Daten einen jeweiligen Leerrahmen einer minimalen Länge für jeden Sprachrahmen umfassen, der während des Ausbleibens von Sprache auftritt.

7. Verfahren zur Aufnahme/Wiedergabe von Sprache in einer digitalen Kommunikationsvorrichtung, wobei die Kommunikationsvorrichtung einen Sprachkodierer zur Erzeugung von Sprachrahmen und einen Generator für künstliches Geräusch zur Bereitstellung von künstlichem Geräusch für eine Simulation von akustischem Hintergrundgeräusch beinhaltet, dadurch gekennzeichnet, dass die digitale Kommunikationsvorrichtung weiterhin eine Abschätzeinrichtung für künstliches Geräusch und eine Sprachaufnahme/-wiedergabe-Einrichtung aufweist, wobei die Abschätzeinrichtung für künstliches Geräusch zur Bereitsstellung von Gesprächspausenrahmen dient, wobei die Gesprächspausenrahmen Information bezeichnend für akustisches Hintergrundgeräusch umfassen, wobei das Verfahren für eine Aufnahme beinhaltet:
Speichern von Sprachrahmen während des Vorhandenseins von Sprache durch die Sprachaufnahme/-wiedergabe-Einrichtung,
Speichern eines oder mehrerer Gesprächspausenrahmen durch die Sprachaufnahme/-wiedergabe-Einrichtung, die Information bezeichnend für akustisches Hintergrundgeräusch beinhalten, an dem Ende des Vorhandenseins von Sprache während des Ausbleibens von Sprache, und
Speichern von Daten durch die Sprachaufnahme/-wiedergabe-Einrichtung, die bezeichnend sind für die Dauer des Ausbleibens von Sprache, und für die Wiedergabe:
Bereitstellen von Sprachsignalen durch die Sprachaufnahme/-wiedergabe-Einrichtung als Ausgabe, die von den gespeicherten Sprachrahmen abgeleitet sind, und von künstlichem Geräusch in Abhängigkeit von dem gespeicherten Gesprächspausenrahmen oder den gespeicherten Gesprächspausenrahmen für das Simulieren von akustischem Hintergrundgeräusch für eine Dauer repräsentiert durch die gespeicherten Daten.

8. Verfahren nach Anspruch 7, wobei die gespeicherten Daten eine Anzahl von Sprachrahmen umfassen, die während des Ausbleibens von Sprache auftreten.

9. Verfahren nach Anspruch 7, wobei die gespeicherten Daten einen jeweiligen Leerrahmen einer minimalen Länge für jeden Sprachrahmen umfassen, der während des Ausbleibens von Sprache auftritt.

Es folgen 3 Blatt Zeichnungen

Anhängende Zeichnungen

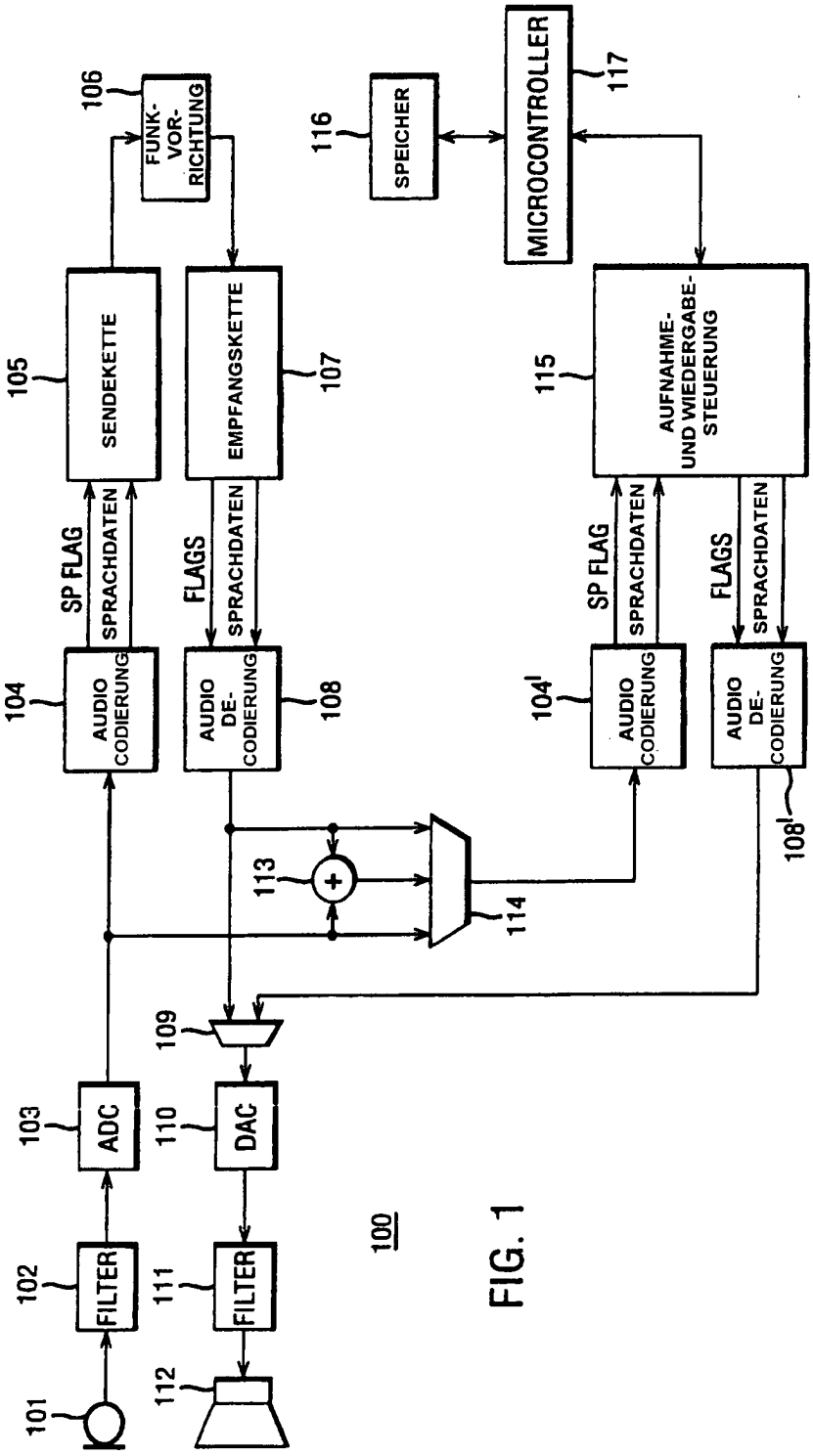


FIG. 1

FIG. 2

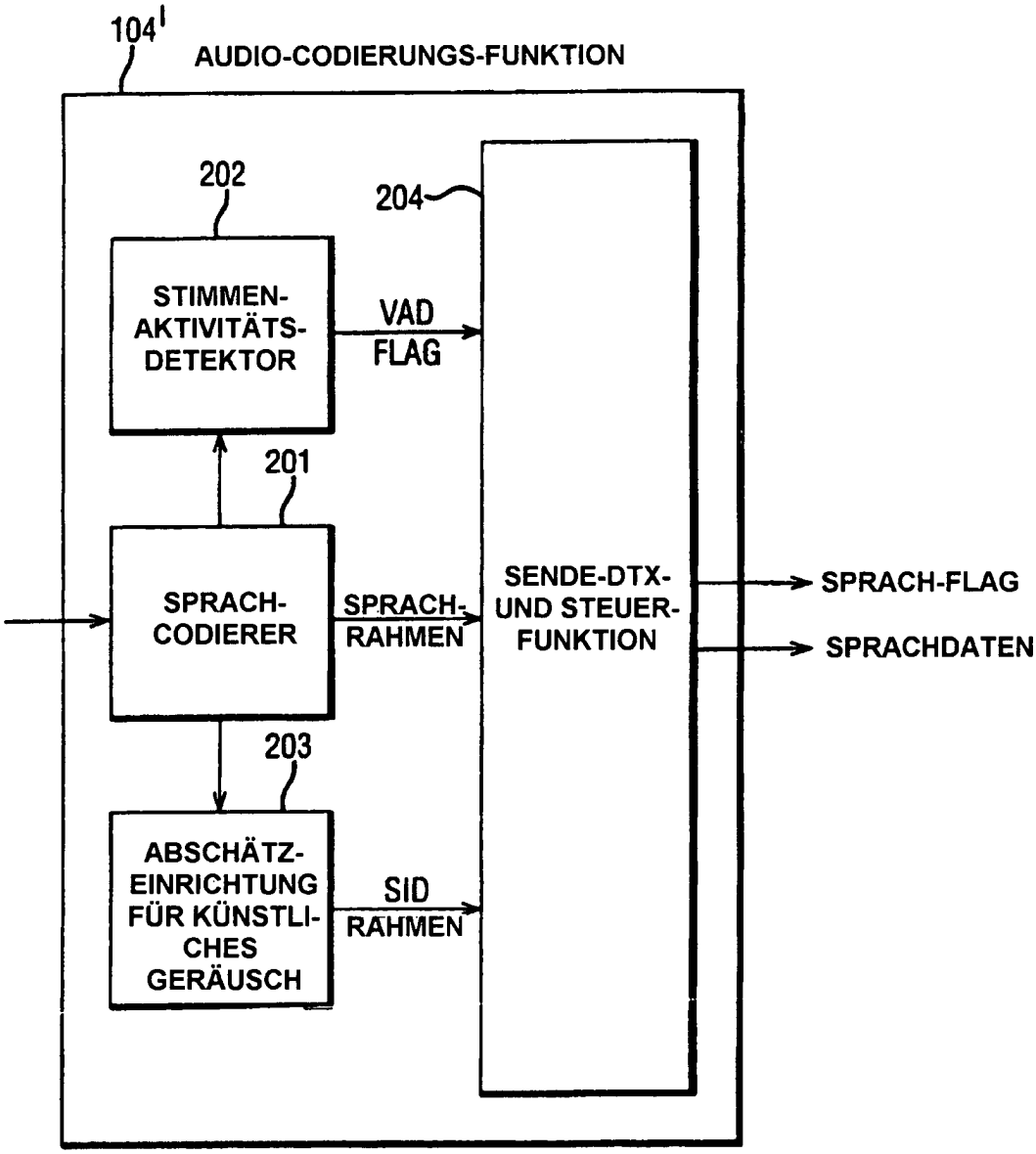
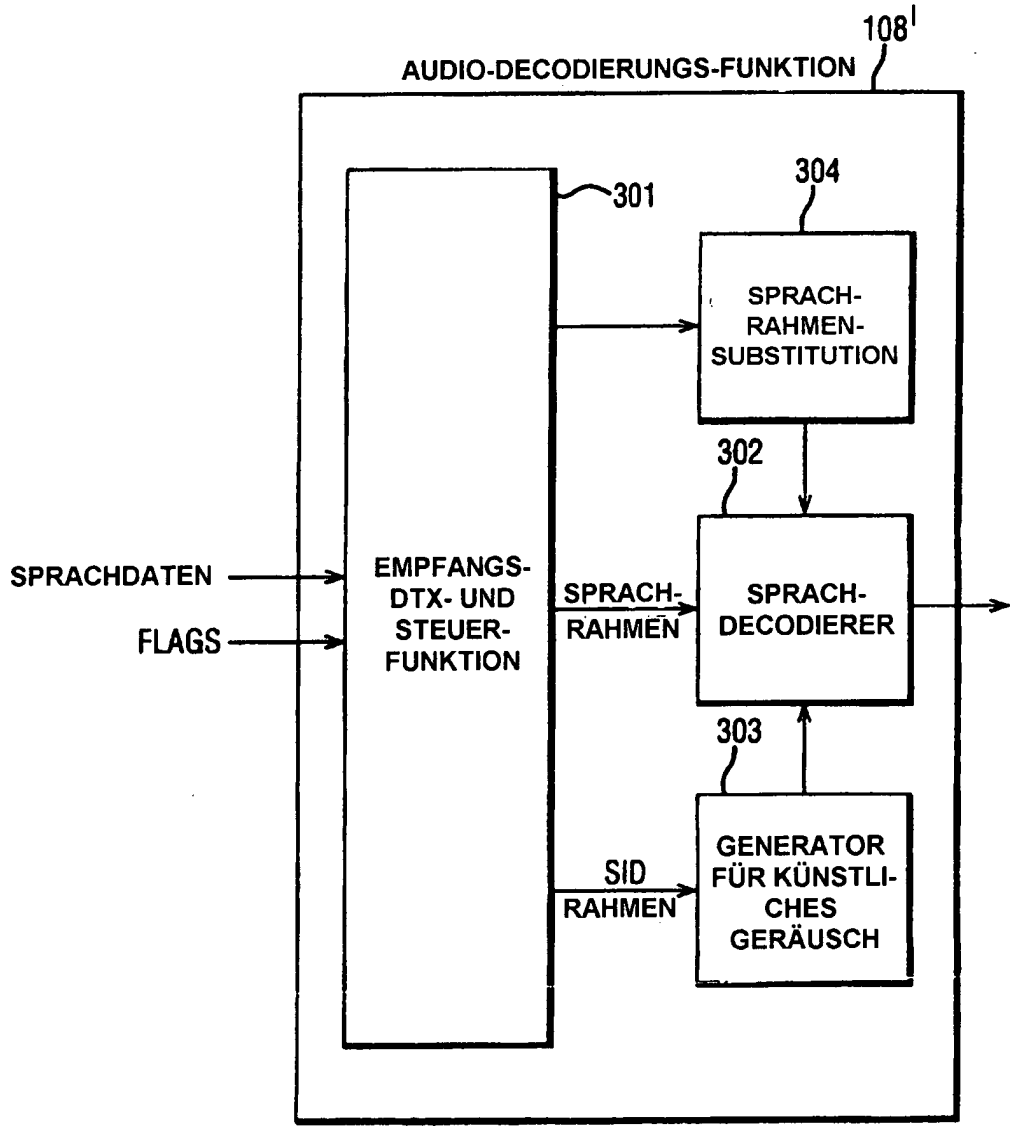

104'
AUDIO-CODIERUNGS-FUNKTION
202
204
STIMMEN-AKTIVITÄTS-DETEKTOR
VAD FLAG
201
SPRACH-CODIERER
SPRACH-RAHMEN
SENDE-DTX- UND STEUER-FUNKTION
SPRACH-FLAG
SPRACHDATEN
203
ABSCHÄTZ-EINRICHTUNG FÜR KÜNSTLI-CHES GERÄUSCH
SID RAHMEN

FIG. 3


108'
AUDIO-DECODIERUNGS-FUNKTION
301
304
SPRACH-RAHMEN-SUBSTITUTION
302
SPRACHDATEN
FLAGS
EMPFANGS-DTX- UND STEUER-FUNKTION
SPRACH-RAHMEN
SPRACH-DECODIERER
303
SID RAHMEN
GENERATOR FÜR KÜNSTLI-CHES GERÄUSCH