



(19)  
Bundesrepublik Deutschland  
Deutsches Patent- und Markenamt

(10) **DE 699 08 723 T2** 2004.05.13

(12) **Übersetzung der europäischen Patentschrift**

(97) **EP 0 984 426 B1**

(51) Int Cl.<sup>7</sup>: **G10L 13/06**

(21) Deutsches Aktenzeichen: **699 08 723.6**

(96) Europäisches Aktenzeichen: **99 306 925.1**

(96) Europäischer Anmeldetag: **31.08.1999**

(97) Erstveröffentlichung durch das EPA: **08.03.2000**

(97) Veröffentlichungstag

der Patenterteilung beim EPA: **11.06.2003**

(47) Veröffentlichungstag im Patentblatt: **13.05.2004**

(30) Unionspriorität:

**24595198      31.08.1998      JP**

(84) Benannte Vertragsstaaten:

**DE, FR, GB**

(73) Patentinhaber:

**Canon K.K., Tokio/Tokyo, JP**

(72) Erfinder:

**Okutani, Yasuo, Ohta-ku, Tokyo, JP; Yamada,  
Masayuki, Ohta-ku, Tokyo, JP**

(74) Vertreter:

**Tiedtke, Bühling, Kinne & Partner GbR, 80336  
München**

(54) Bezeichnung: **Verfahren und Vorrichtung zur Sprachsynthese, sowie Speichermedium**

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****HINTERGRUND DER ERFINDUNG**

[0001] Die Erfindung bezieht sich auf eine Sprachsynthesevorrichtung mit einer Datenbank zum Verwalten von Phonemdaten, bei der die Vorrichtung unter Verwendung der durch die Datenbank verwalteten Phonemdaten eine Sprachsynthese durchführt. Die Erfindung bezieht sich ferner auf ein Verfahren zur Synthese von Sprache unter Verwendung dieser Vorrichtung und auf ein Programm zum Realisieren dieses Verfahrens speicherndes Speichermedium.

[0002] Ein Verfahren zur Sprachsynthese (auf das nachstehend als das "verkettende Syntheseverfahren" Bezug genommen ist), das einen Signalverlauf, das einen Signalverkauf verkettet, ist bei dem Stand der Technik als ein Verfahren zur Synthese von Sprache verfügbar. Das verkettende Syntheseverfahren ändert den Satzrhythmus mit einem Tönhöhensynchronüberlappungsadditionsverfahren (P-SOLA), das den Satzrhythmus durch ein Plazieren von aus der ursprünglichen Signalverlaufseinheit extrahierten Tönhöhensignalverlaufseinheiten gemäß einem gewünschten Tönhöhenzeitverlauf ändert. Ein Vorteil des verkettenden Syntheseverfahrens besteht darin, daß die erhaltene synthetisierte Sprache natürlicher ist als die durch ein Syntheseverfahren auf der Grundlage von Parametern bereitgestellte. Ein Nachteil besteht darin, daß der zulässige Bereich für die Änderung des Satzrhythmus eng ist.

[0003] Entsprechend wird die Tonqualität verbessert, indem Sprachdaten einer breiten Vielfalt von Variationen vorbereitet werden, diese richtig ausgewählt werden und sie verwendet werden. Informationen wie beispielsweise die Phonemumgebung (das Phonem, das das Ziel der Synthese ist, oder mehrere Phoneme einschließlich beider Seiten davon) und die Grundfrequenz  $F_0$  werden als die Kriterien zum Auswählen der Syntheseeinheit verwendet.

[0004] Das vorstehend beschriebene bekannte Verfahren zur Synthese von Sprache bringt jedoch eine Anzahl von Problemen mit sich.

[0005] Falls beispielsweise eine Datenbank eine Vielzahl von Einheiten von Phonemdaten enthält, die einer bestimmten Phonemumgebung und der Grundfrequenz  $F_0$  entsprechen, ist die bei der Synthese verwendete Phonemeinheit eine aus diesen Einheiten von Phonemdaten zufällig ausgewählte Phonemeinheit (zum Beispiel die Phonemeinheit, die in der Datenbank als erstes auftaucht). Da die Datenbank eine Sammlung von durch menschliche Wesen hervorgebrachter Sprache ist, sind nicht notwendigerweise alle Phonemdaten stabil (das heißt nicht notwendigerweise von guter Qualität). Die Datenbank kann Phonemdaten enthalten, die das Ergebnis eines Murmelns, einer stockenden Stimme, einer Langsamkeit der Sprache oder einer Heiserkeit sind. Falls eine Einheit von Phonemdaten aus einer derartigen Sammlung von Daten zufällig ausgewählt wird,

besteht natürlich die Möglichkeit, daß die Tonqualität bei der Erzeugung von synthetisierter Sprache abnimmt.

[0006] GB 2313530 beschreibt eine Sprachsyntheseeinrichtung, die eine Gewichtungskoeffiziententrainingssteuereinrichtung verwendet, die akustische Abstände zwischen einem Zielphonem und Phonemkandidaten auf der Grundlage von akustischen Merkmalsparametern und prosodischen Merkmalsparametern berechnet und die Grade des Beitrags zu den zweiten akustischen Merkmalsparametern für jeweilige Phonemkandidaten definierende Gewichtungskoeffizientenvektoren für jeweilige Zielphoneme bestimmt, indem eine vorbestimmte statistische Analyse ausgeführt wird. Eine Auswahleinrichtung sucht nach einer Kombination von Phonemkandidaten, die einer Phonemfolge einer Eingabefolge entspricht und die einen Näherungsaufwände zwischen einem Zielphonem und den Phonemkandidaten darstellenden Zielaufwand und einen Näherungsaufwände zwischen zwei angrenzend zu verkettenden Phonemkandidaten darstellenden Verkettungsaufwand minimiert, und gibt Indexinformationen über die gesuchte Ausgabekombination von Phonemkandidaten aus. Eine Syntheseeinrichtung synthetisiert daraufhin ein der Eingabephonemfolge entsprechendes Sprachsignal, indem den Indexinformationen entsprechende Sprachsegmente von Sprachsignalverlaufssignalen sequentiell ausgelesen und die gelesenen Sprachsegmente der Sprachsignalverlaufssignale verkettet werden.

[0007] Gemäß einer Ausgestaltung stellt die vorliegende Erfindung eine Sprachsynthesevorrichtung bereit, mit:

einer Speichereinrichtung zum Speichern mehrerer Einheiten von Phonemdaten;

einer Wiedergewinnungseinrichtung zum Wiedergewinnen von Phonemdaten aus den in der Speichereinrichtung gespeicherten mehreren Einheiten von Phonemdaten gemäß gegebenen Wiedergewinnungsbedingungen;

einer ersten Strafuweisungseinrichtung zum Sortieren von durch die Wiedergewinnungseinrichtung wiedergewonnenen Phonemdaten auf der Grundlage eines vorgeschriebenen Eigenschaftswerts und zum Zuweisen einer auf einem Eigenschaftswert basierenden Strafe zu jeder Einheit der Phonemdaten auf der Grundlage einer durch das Sortieren erhaltenen Reihenfolge; und

einer Auswahleinrichtung zum Auswählen von bei einer Synthese eines Sprachsignalverlaufs zu verwendenden Phonemdaten aus den durch die Wiedergewinnungseinrichtung wiedergewonnenen Phonemdaten und auf der Grundlage der durch die erste Strafuweisungseinrichtung zugewiesenen Strafe.

[0008] Gemäß einer weiteren Ausgestaltung stellt die vorliegende Erfindung ein Sprachsyntheseverfahren bereit, mit:

einem Speicherschritt des Speicherns mehrerer Einheiten von Phonemdaten;

einem Wiedergewinnungsschritt des Wiedergewinnens von Phonemdaten aus den in dem Speicherschritt gespeicherten mehreren Einheiten von Phonemdaten gemäß gegebenen Suchwiedergewinnungsbedingungen;

einem ersten Strafzuweisungsschritt, der in dem Wiedergewinnungsschritt wiedergewonnene Phonemdaten auf der Grundlage eines vorgeschriebenen Eigenschaftswerts sortiert und der auf der Grundlage einer durch das Sortieren erhaltenen Reihenfolge jeder Einheit der Phonemdaten eine auf einem Eigenschaftswert basierende Strafe zuweist; und

einem Auswahlsschritt des Auswählens von bei einer Synthese eines Sprachsignalverlaufs verwendeten Phonemdaten aus den in dem Wiedergewinnungsschritt wiedergewonnenen Phonemdaten und auf der Grundlage der in dem Strafzuweisungsschritt zugewiesenen Strafe.

[0009] Die vorliegende Erfindung stellt ferner ein Speichermedium bereit, das ein Steuerprogramm zur Veranlassung eines Computers zum Ausführen des vorstehend beschriebenen Verfahrens zur Synthese von Sprache speichert.

[0010] Andere Merkmale und Vorteile der vorliegenden Erfindung sind aus der in Verbindung mit den beigefügten Zeichnungen, in denen gleiche Bezugszeichen überall in den Figuren davon die gleichen oder ähnliche Teile bezeichnen, genommenen folgenden Beschreibung ersichtlich.

#### KURZBESCHREIBUNG DER ZEICHNUNGEN

[0011] Die beigefügten Zeichnungen, die in die Patentbeschreibung integriert sind und einen Teil von ihr bilden, veranschaulichen Ausführungsbeispiele der Erfindung und dienen zusammen mit der Beschreibung zur Erläuterung der Prinzipien der Erfindung.

[0012] **Fig. 1** zeigt ein Blockschaltbild, das den Aufbau einer Sprachsynthesevorrichtung gemäß einem ersten Ausführungsbeispiel der vorliegenden Erfindung darstellt;

[0013] **Fig. 2** zeigt ein Blockschaltbild, das Funktionen bezüglich einer Phonemdatenauswahlverarbeitung gemäß dem ersten Ausführungsbeispiel veranschaulicht;

[0014] **Fig. 3** zeigt ein Flußdiagramm, das eine Prozedur bezüglich einer Phonemdatenauswahlverarbeitung gemäß dem ersten Ausführungsbeispiel veranschaulicht;

[0015] **Fig. 4** zeigt ein Blockschaltbild, das Funktionen bezüglich einer Phonemdatenauswahlverarbeitung gemäß dem zweiten Ausführungsbeispiel veranschaulicht;

[0016] **Fig. 5** zeigt ein Flußdiagramm, das eine Prozedur bezüglich einer Phonemdatenauswahlverarbeitung gemäß dem zweiten Ausführungsbeispiel veranschaulicht; und

[0017] **Fig. 6** zeigt ein Flußdiagramm, das bei der Beschreibung eines Überblicks der Sprachsynthese-

verarbeitung verwendbar ist.

#### BESCHREIBUNG DER BEVORZUGTEN AUSFÜHRUNGSBEISPIELE

[0018] Nachstehend sind bevorzugte Ausführungsbeispiele der Erfindung gemäß den beigefügten Zeichnungen im einzelnen beschrieben.

##### [Erstes Ausführungsbeispiel]

[0019] **Fig. 1** zeigt ein Blockschaltbild, das den Aufbau einer Sprachsynthesevorrichtung gemäß einem ersten Ausführungsbeispiel der Erfindung veranschaulicht.

[0020] Wie es in **Fig. 1** gezeigt ist, umfaßt die Vorrichtung einen Speicherspeicher (ROM) **101**, der ein Steuerprogramm zum Veranlassen eines Computers zur Realisierung einer Steuerung gemäß einer in **Fig. 3** gezeigten Steuerprozedur speichert, eine Zentraleinheit **102** zum Ausführen einer Verarbeitung wie beispielsweise Entscheidungen und Berechnungen gemäß der in dem Speicherspeicher **101** gehaltenen Steuerprozedur, und einen Speicher (RAM) **103**, der einen Arbeitsbereich dafür bereitstellt, wenn die Zentraleinheit **102** verschiedene Steueroperationen ausführt. Dem Speicher **103** sind ein Bereich **202** zum Halten der Ergebnisse einer Phonemwiedergewinnung, ein Bereich **204** zum Halten der Ergebnisse einer Strafzuweisung, ein Bereich **207** zum Halten der Ergebnisse einer Sortierung und ein Bereich **209** zum Halten repräsentativer Phonemdaten zugeordnet. Diese Bereiche sind später unter Bezugnahme auf **Fig. 2** beschrieben. Die Vorrichtung umfaßt ferner eine Plattenvorrichtung **104**, die bei diesem Ausführungsbeispiel eine Festplatte ist. Die Plattenvorrichtung **104** speichert eine Datenbank **200**, die später unter Bezugnahme auf **Fig. 2** beschrieben ist. Die Daten der Datenbank **200** werden in dem Speicher **103** gespeichert, wenn die Daten verwendet werden. Ein Bus **105** verbindet die vorstehend angeführten Komponenten.

[0021] Die Sprachsynthesevorrichtung gemäß diesem Ausführungsbeispiel verwendet Informationen wie beispielsweise die Phonemumgebung und die Grundfrequenz zur Auswahl der passenden Phonemdaten aus Sprachdaten, die in der Datenbank **200** (**Fig. 2**) aufgezeichnet worden sind, und führt unter Verwendung der ausgewählten Daten eine Signalverlaufsbearbeitungssynthese durch.

[0022] **Fig. 6** zeigt ein Flußdiagramm, das einen Überblick der Sprachsyntheseverarbeitung gemäß diesem Ausführungsbeispiel veranschaulicht. Die Phonemumgebung und die Grundfrequenz eines zu verwendenden Phonems werden in einem Schritt S11 in **Fig. 6** bestimmt. Dies kann ausgeführt werden, indem die Phonemumgebung und die Grundfrequenz als eine Parameterdatei in der Plattenvorrichtung **104** gespeichert werden oder indem sie über eine Tastatur eingegeben werden. Als nächstes wer-

den in einem Schritt S12 zu verwendende Phonemdaten aus der Datenbank **200** ausgewählt. Dem folgt ein Schritt S13, in dem es bestimmt wird, ob weitere zu verarbeitende Phonemdaten vorhanden sind. Die Steuerung kehrt zu dem Schritt S11 zurück, falls derartige Daten vorhanden sind. Falls es bestimmt wird, daß alle notwendigen Phonemdaten ausgewählt worden sind, geht die Steuerung demgegenüber von dem Schritt S13 zu einem Schritt S14 über, und es wird eine Sprachsynthese durch eine Signalverlaufsbearbeitung unter Verwendung der ausgewählten Phonemdaten ausgeführt.

[0023] Die Einzelheiten der Verarbeitung zum Auswählen der Phonemdaten in dem Schritt S12 sind nachstehend beschrieben. In dem nachstehend beschriebenen Fall wird die Auswahl von Phonemdaten unter Verwendung der Phonemumgebung (drei Phoneme, die aus dem Phonem von Interesse und einem Phonem auf jeder Seite davon bestehen, wobei auf diese als ein so genanntes "Triphon" Bezug genommen wird) und der Durchschnittsgrundfrequenz des Phonems als Kriterien zum Auswählen von Phonemdaten ausgeführt.

[0024] **Fig. 2** zeigt ein Blockschaltbild, das Funktionen bezüglich einer Phonemdatenauswahlverarbeitung zum Auswählen der optimalen Phonemdaten aus einem Satz von Phonemdaten, bei dem die Phonemumgebungen und Grundfrequenzen identisch sind, veranschaulicht. Die Funktionen sind die einer Sprachsynthesevorrichtung gemäß dem ersten Ausführungsbeispiel.

[0025] Die Datenbank **200** in **Fig. 2** speichert Sprachdaten, bei denen jeder Einheit von Phonemdaten eine Phonemumgebung, eine Phonemgrenze und eine Grundfrequenz, eine Energie und eine Phonemdauer zugewiesen worden sind. Eine Phonemwiedergewinnungseinheit **201** gewinnt Phonemdaten, die einer spezifischen Phonemumgebung und Grundfrequenz entsprechen, aus der Datenbank **200** wieder. Der Bereich **202** speichert einen Satz von Phonemdaten, das heißt die Ergebnisse der durch die Phonemwiedergewinnungseinheit **201** durchgeführten Wiedergewinnung. Eine Energiestrafzuweisungseinheit **203** weist jeder Einheit von Phonemdaten des in dem Bereich **202** gespeicherten Satzes von Phonemdaten eine auf die Energie bezogene Strafe zu. Der Bereich **204** hält die Ergebnisse der Zuweisung von Strafen zu den Phonemdaten. Eine Dauerstrafzuweisungseinheit **205** weist jeder Einheit von Phonemdaten eine Strafe bezüglich der Phonemdauer zu.

[0026] Eine Sortierungsverarbeitungseinheit **206** unterzieht den Satz von Phonemdaten einer Sortierungsverarbeitung hinsichtlich spezifischer Informationen (Energie oder Phonemdauer usw.), wenn eine Strafe zugewiesen wird. Der Bereich **207** hält die Ergebnisse der Sortierung. Hinsichtlich der durch die Zuweisung von Strafen erhaltenen Ergebnisse wählt eine Datenbestimmungsverarbeitungseinheit **208** Phonemdaten mit der kleinsten Strafe als repräsen-

tative Phonemdaten aus. Der Bereich **209** hält die repräsentativen Phonemdaten, die bestimmt worden sind.

[0027] Von der vorstehend dargelegten Sprachsyntheseverarbeitung ist als nächstes die durch den vorstehend beschriebenen Funktionsaufbau realisierte Verarbeitung zum Auswählen von Phonemdaten erörtert. **Fig. 3** zeigt ein Flußdiagramm, das eine Prozedur bezüglich einer Phonemdatenauswahlverarbeitung zum Auswählen der optimalen Phonemdaten aus dem Satz von Phonemdaten mit identischen Phonemumgebungen und Grundfrequenzen veranschaulicht.

[0028] Zuerst werden in einem Schritt S301 alle Phonemdaten aus der Datenbank **200** extrahiert und in dem Bereich **202** gespeichert, die der Phonemumgebung (Triphon) und der Grundfrequenz  $F_0$  entsprechen, die in dem Schritt S11 bestimmt wurden. Als nächstes weist in einem Schritt S302 die Energiestrafzuweisungseinheit **203** dem Satz von Phonemdaten, der in dem Bereich **202** gespeichert worden ist, energiebezogene Strafen zu.

[0029] Die energiebezogene Strafen umfassende Richtlinie besteht darin, Phonemdaten mit Energiewerten, die von einem Durchschnittswert der Energie abweichen, große Strafen zuzuweisen, da das Ziel darin besteht, Phonemdaten mit einem Durchschnittswert der Energie in dem Satz von Phonemdaten auszuwählen. Die Energiestrafzuweisungseinheit **203** weist die Sortierungsverarbeitungseinheit **206** zur Sortierung des Phonemdatensatzes, der aus dem die Ergebnisse der Wiedergewinnung haltenden Bereich **202** extrahiert worden ist, auf der Grundlage von Werten der Energie an. Die Energie, auf die in diesem Fall Bezug genommen ist, kann die Energie der Phonemdaten oder die Durchschnittsenergie pro Zeiteinheit sein.

[0030] Die Sortierungsverarbeitungseinheit **206** reagiert, indem sie den Phonemdatensatz auf der Grundlage der Energie sortiert und die Ergebnisse in dem Bereich **207** speichert, der zum Halten der Ergebnisse der Sortierung dient. Die Energiestrafzuweisungseinheit **203** wartet darauf, daß die Sortierung endet, und weist daraufhin den sortierten Phonemdaten, die in dem Bereich **207** gespeichert worden sind, eine Strafe zu. Eine Strafe wird gemäß der vorstehend angeführten Richtlinie zugewiesen. Beispielsweise wird unter Einheiten von Phonemdaten, die in der Reihenfolge abnehmender Energie sortiert worden sind, Phonemdaten, deren Energiewerte in das kleinere Drittel von Werten fallen, und Phonemdaten, deren Energiewerte in das größere Drittel von Werten fallen, eine Strafe (zum Beispiel 2,0 Punkte) hinzugefügt. Mit anderen Worten wird anderen Phonemdaten als dem mittleren Drittel von Phonemdaten eine Strafe zugewiesen.

[0031] Als nächstes weist in einem Schritt S303 die Dauerstrafzuweisungseinheit **205** eine Strafe bezüglich der Phonemdauer durch eine Prozedur zu, die der der Energiestrafzuweisungseinheit

tungseinheit **203** ähnelt. Im einzelnen weist die Dauerstrafzuweisungsverarbeitungseinheit **205** die Sortierungsverarbeitungseinheit **206** zur Durchführung einer Sortierung auf der Grundlage der Phonemdauer an und speichert die Ergebnisse in dem Bereich **207**. Auf der Grundlage der sortierten Ergebnisse fügt die Dauerstrafzuweisungsverarbeitungseinheit **205** Phonemdaten, deren Phonemdauern in das kleinere Drittel von Dauern fallen, und Phonemdaten, deren Phonemdauern in das größere Drittel von Dauern fallen, eine Strafe (zum Beispiel 2,0 Punkte) hinzu. Die durch die Zuweisung der Strafe erhaltenen Ergebnisse werden in dem Bereich **204** gehalten. Die Steuerung geht daraufhin zu einem Schritt S304 über.

[0032] Der Schritt S304 verlangt es von der Datenbestimmungsverarbeitungseinheit **208**, eine hinsichtlich der Phonemumgebung und der Grundfrequenz, die derzeit von Interesse sind, repräsentative Phonemeinheit zu bestimmen. In diesem Fall wird der in dem Bereich **204** gespeicherte Satz von Phonemdaten, dem eine Strafe auf der Grundlage der Energie und der Phonemdauer zugewiesen ist, der Sortierungsverarbeitungseinheit **206** zugeführt, und die Sortierungsverarbeitungseinheit **206** wird zur Sortierung der Ergebnisse gemäß dem Strafwert angewiesen. Die Sortierungsverarbeitungseinheit **206** führt eine Sortierung auf der Grundlage der zwei Typen von Strafen bezüglich der Energie und der Phonemdauer durch (zum Beispiel unter Verwendung der Summe der zwei Strafwerte) und speichert die sortierten Ergebnisse in dem Bereich **207**. Wenn die Sortierungsverarbeitung endet, wählt die Datenbestimmungsverarbeitungseinheit **208** Phonemdaten mit der kleinsten Strafe aus und speichert sie in dem Bereich **209**, um diese Daten als repräsentative Phonemdaten zu verwenden. Falls eine Vielzahl von Phonemeinheiten mit dem minimalen Strafwert erscheint, wählt die Datenbestimmungsverarbeitungseinheit **208** die Phonemeinheit aus, die sich an der Spitze der sortierten Ergebnisse befindet. Dies ist äquivalent dazu, eine Phonemeinheit aus denen mit der kleinsten Strafe zufällig auszuwählen.

[0033] Somit werden gemäß dem ersten Ausführungsbeispiel die optimalen Phonemdaten auf der Grundlage einer Strafe bezüglich der Energie und einer Strafe bezüglich der Phonemdauer aus einem Phonemdatensatz ausgewählt, bei dem die Phonemumgebungen und die Grundfrequenzen identisch sind.

#### [Zweites Ausführungsbeispiel]

[0034] Das erste Ausführungsbeispiel ist hinsichtlich eines Falls beschrieben, in dem die Phonemumgebung (das "Triphon", das heißt das Phonem von Interesse und ein Phonem auf jeder Zeit davon) und die Durchschnittsgrundfrequenz  $F_0$  des Phonems als Kriterien zum Auswählen von Phonemdaten verwendet werden. In Fällen, in denen das Triphon einer nicht in

der Datenbank enthaltenen Kombination erforderlich ist, entsteht jedoch der Bedarf an der Verwendung eines alternativen "Links-Phons" (einer Phonemumgebung mit dem Phonem von Interesse und dem Phonem zu seiner Linken), "Rechts-Phons" (einer Phonemumgebung mit dem Phonem von Interesse und dem Phonem zu seiner Rechten) oder "Phons" (des Phonems von Interesse alleine). Bei dem zweiten Ausführungsbeispiel ist daher ein Fall beschrieben, in dem die Auswahl von anderen Phonemdaten als einem bestimmten Triphon (auf derartige ausgewählte Phonemdaten ist als ein "Triphon-Ersatz" Bezug genommen) berücksichtigt wird.

[0035] **Fig. 4** zeigt ein Blockschaltbild, das Funktionen bezüglich einer Phonemdatenauswahlverarbeitung zum Auswählen der optimalen Phonemdaten aus einem Satz von Phonemdaten, bei dem die Phonemumgebungen und die Grundfrequenzen identisch sind, veranschaulicht. Die Funktionen sind die einer Sprachsynthesevorrichtung gemäß dem zweiten Ausführungsbeispiel. Dieses Ausführungsbeispiel unterscheidet sich von dem ersten Ausführungsbeispiel in **Fig. 2** dahingehend, daß die Vorrichtung zudem eine Verarbeitungseinheit zum Zuweisen einer Elementanzahlstrafe umfaßt. Andere Bereiche oder Einheiten **400** bis **409** entsprechen jeweils den Bereichen oder Einheiten **200** bis **209** gemäß **Fig. 2**. Die Verarbeitungseinheit **410** weist in Abhängigkeit von der Anzahl von Elementen in einem Satz von Phonemdaten eine Strafe zu.

[0036] Die Sprachsyntheseverarbeitung umfaßt eine Prozedur bezüglich einer Phonemdatenauswahlverarbeitung zum Auswählen optimaler Phonemdaten aus einem Satz von Phonemdaten mit identischen Phonemumgebungen und Grundfrequenzen, die durch die vorstehend beschriebenen Funktionsblöcke realisiert wird. Diese Prozedur ist nachstehend beschrieben. **Fig. 5** zeigt ein Flußdiagramm, das eine Prozedur gemäß dem zweiten Ausführungsbeispiel bezüglich einer Phonemdatenauswahlverarbeitung zum Auswählen der optimalen Phonemdaten aus dem Satz von Phonemdaten mit identischen Phonemumgebungen und Grundfrequenzen veranschaulicht.

[0037] Schritte S501 bis S503 ähneln den Schritten S301 bis S303 (**Fig. 3**) bei dem ersten Ausführungsbeispiel. Es ist zu beachten, daß die Triphon-Wiedergewinnung in dem Schritt S501 die Wiedergewinnung der alternativen Kandidaten Links-Phon, Rechts-Phon oder Phon (des vorstehend erwähnten "Triphon-Ersatzes") umfaßt, falls ein bestimmtes Triphon nicht in der Datenbank vorhanden ist. In diesem Fall wird zum Beispiel zuerst die Wiedergewinnung des Links-Phons ausgeführt. Wenn das Links-Phon nicht in der Datenbank vorhanden ist, dann wird die Wiedergewinnung des Rechts-Phons ausgeführt. Wenn das Rechts-Phon nicht vorhanden ist, dann wird die Wiedergewinnung des Phons ausgeführt. Alternativ kann die Folge der Wiedergewinnung zwischen einem Vokal und einem Konsonanten ver-

schieden sein. Beispielsweise wird hinsichtlich eines Vokals die Wiedergewinnung in der Folge von Links-Phon, Rechts-Phon und Phon ausgeführt. Hinsichtlich eines Konsonanten wird die Wiedergewinnung in der Folge von Rechts-Phon, Links-Phon und Phon ausgeführt.

[0038] Bei dem zweiten Ausführungsbeispiel bedeutet es die Verwendung eines Triphon-Ersatzes, daß ein bestimmtes Triphon nicht vorhanden ist. Solange ein bestimmtes Triphon in der Datenbank enthalten ist, wird jedoch dieses Triphon angewendet. In einem Schritt S504 wird es daher bestimmt, ob als das Ergebnis der Wiedergewinnung ein Triphon-Ersatz erhalten worden ist. Falls kein Triphon-Ersatz erhalten worden ist, das heißt falls das bestimmte Triphon erhalten worden ist, überspringt die Steuerung einen Schritt S505 und geht zu einem Schritt S506 über. Wenn das bestimmte Triphon wiedergewonnen wird, wird daher eine Verarbeitung ausgeführt, die der des ersten Ausführungsbeispiels ähnelt. Falls es in dem Schritt S504 bestimmt wird, daß ein Triphon-Ersatz wiedergewonnen worden ist, geht die Steuerung demgegenüber zu dem Schritt S505 über. In diesem Fall weist die Verarbeitungseinheit **410** in Abhängigkeit von der Anzahl von Elementen in dem Satz von Phonemdaten eine Strafe zu. In einem Fall, in dem das bestimmte Triphon fehlt, zählt die Verarbeitungseinheit **505** die Anzahl von in dem Phonemdatensatz enthaltenen Elementen, wobei die Zählung pro jeder Triphon-Phonemumgebungsgruppe (einer durch die Umgebung mit dem betroffenen Phonem und einem Phonem auf jeder Site davon klassifizierten Gruppe) des alternativen Kandidaten Links-Phon (oder Rechts-Phon oder Phon) durchgeführt wird. Wenn die Anzahl von Einheiten von Phonemdaten einer anwendbaren Triphon-Phonemumgebung klein (zwei oder weniger) ist, dann fügt bei diesem Ausführungsbeispiel die Verarbeitungseinheit **505** allen betroffenen Phonemdaten eine Strafe (0,5 Punkte) hinzu. Mit anderen Worten entscheidet es die Verarbeitungseinheit **505**, daß Daten, die in einer ausreichend großen Datenbank nur eine niedrige Erscheinungshäufigkeit aufweisen, nicht zuverlässig sind.

[0039] Beispielsweise wird ein Fall betrachtet, in dem ein Triphon t.A.k nicht in der Datenbank vorhanden ist und durch ein Links-Phon t.A.\* zu ersetzen ist. Falls zwei Triphone t.A.p und 20 Triphone t.A.t in der Datenbank vorhanden sind, stellt ein Zuordnen eines Triphon-Ersatzes, der das Triphon t.A.k ersetzen soll, aus Triphonen t.A.t, von denen 20 vorhanden sind, eine höhere Wahrscheinlichkeit des Erhaltens von Phonemdaten von guter Qualität bereit.

[0040] Falls so eine Strafe auf der Grundlage der Anzahl von Elementen zugewiesen wird, wird das Ergebnis in dem Bereich **404** gespeichert, der zum Halten der Ergebnisse der Strafzuweisung dient, und daraufhin geht die Steuerung zu dem Schritt S506 über. Der Schritt S506 umfaßt eine Verarbeitung, die zu der des Schritts S304 bei dem ersten Ausführungsbeispiel äquivalent ist. Bei dem zweiten Ausführungsbeispiel

wird zusätzlich zu der Strafe auf der Grundlage der Energie und der Strafe auf der Grundlage der Phonemdauer eine Strafe auf der Grundlage der Anzahl von Elementen zugewiesen. Folglich werden Phonemdaten unter Berücksichtigung aller dieser drei Strafen ausgewählt. In einem Fall, in dem ein spezifisches Triphon wiedergewonnen wird und die Verarbeitung direkt von dem Schritt S504 zu dem Schritt S506 übergeht, wird die Strafe auf der Grundlage der Anzahl von Elementen nicht berücksichtigt. [0041] Somit ist es gemäß dem zweiten Ausführungsbeispiel möglich, die richtigen Phonemdaten einschließlich Triphonen, die Alternativen sein können, auszuwählen.

[0042] Bei den vorstehend dargelegten Ausführungsbeispielen ist ein Fall beschrieben, in dem die Strafzuweisungsverarbeitung in der Reihenfolge der Energiestrafe und der Phonemdauerstrafe (und daraufhin der Elementanzahlstrafe bei dem zweiten Ausführungsbeispiel) ausgeführt wird. Dies legt der Erfindung jedoch keine Beschränkung auf, da die Verarbeitung in jeder Reihenfolge ausgeführt werden kann. Ferner kann ein Aufbau angewendet werden, bei dem diese Strafzuweisungsverarbeitungsoperationen gleichzeitig ausgeführt werden.

[0043] Ferner werden bei jedem der vorstehenden Ausführungsbeispiele 2,0 Punkte als der Strafwert für die Energie- und Phonemdauerstrafen angewendet. Dies erlegt jedoch der Erfindung keine Beschränkung auf, da es offensichtlich ist, daß ein geeigneter Wert eingestellt werden kann. Darüber hinaus müssen keine gleichen Strafen als die Strafen bezüglich beider Eigenschaften angewendet werden.

[0044] Bei dem zweiten Ausführungsbeispiel ist ein Fall beschrieben, in dem 0,5 als der Wert der Elementanzahlstrafe eingestellt ist. Dies erlegt jedoch der Erfindung keine Beschränkung auf, da ein geeigneter Wert eingestellt werden kann.

[0045] Überdies ist bei jedem der vorstehenden Ausführungsbeispiele ein Fall beschrieben, in dem dem hinsichtlich der sortierten Ergebnisse bei kleineren Werten beginnenden einen Drittel von Phonemdaten (oder dem bei größeren Werten beginnenden einen Drittel von Phonemdaten) eine Strafe zugewiesen wird. Dies erlegt jedoch der Erfindung keine Beschränkung auf. Beispielsweise ist es möglich, das Verfahren zur Strafzuweisung abhängig von der Anzahl von Einheiten von Phonemdaten oder den Eigenschaften der in der Datenbank enthaltenen Phonemdaten zu ändern. In einem derartigen Fall kann Daten eine Strafe zugewiesen werden, für die die Differenz relativ zu einem Durchschnittswert größer als ein Schwellenwert ist.

[0046] Ferner ist bei den vorstehenden Ausführungsbeispielen ein Verfahren zur Auswahl repräsentativer Phonemdaten beschrieben, bei dem das Ziel ein Phonemdatensatz ist, der einer spezifischen Phonemumgebung und Grundfrequenz entspricht. Dies erlegt jedoch der Erfindung keine Beschränkung auf. Beispielsweise ist es möglich, einen Phonemdaten-

satz zu verwenden, für den der Gegenstand des Interesses allein die Phonemumgebung ist, und die Grundfrequenz als einen Faktor zum Zuweisen einer Strafe anzuwenden.

[0047] Ferner ist bei jedem der vorstehenden Ausführungsbeispiele ein Verfahren zur Auswahl einer repräsentativen Phonemeinheit auf Verlangen beschrieben, wobei das Ziel ein Phonemdatensatz ist, der einer spezifischen Phonemumgebung und Grundfrequenz entspricht. Es kann jedoch ein Aufbau angewendet werden, bei dem ein durch ein Anwenden der Verarbeitung gemäß dem ersten Ausführungsbeispiel im voraus erhaltenes Phonemlexikon auf der Grundlage aller vorstellbaren Phonemumgebungen und Grundfrequenzen erzeugt wird.

[0048] Ferner ist bei jedem der vorstehenden Ausführungsbeispiele ein Fall beschrieben, in dem die Sortierungsverarbeitungseinheit und der Bereich zum Halten der sortierten Ergebnisse zur universellen Verwendung entworfen sind. Dies erlegt jedoch der Erfindung keine Beschränkung auf. Beispielsweise kann ein Aufbau angewendet werden, bei dem eine Sortierungsverarbeitungseinrichtung ausschließlich für die Verarbeitungseinheit bereitgestellt ist, die die Energiestrafen zuweist, und eine Sortierungsverarbeitungseinrichtung ausschließlich für die Verarbeitungseinheit bereitgestellt ist, die die Phonemdauerstrafen zuweist.

[0049] Bei jedem der vorstehenden Ausführungsbeispiele ist ein Fall beschrieben, in dem die Bereiche zum Speichern von Daten durch Speicher (RAM) realisiert sind. Dies erlegt jedoch der Erfindung keine Beschränkung auf, da jedes Speichermedium verwendet werden kann.

[0050] Ferner ist bei jedem der vorstehenden Ausführungsbeispiele ein Fall beschrieben, in dem die Komponenten durch den gleichen Computer gebildet werden. Dies erlegt jedoch der Erfindung keine Beschränkung auf, da diese Komponenten durch Computer oder Verarbeitungseinrichtungen realisiert werden können, die über ein Netz verteilt sind.

[0051] Ferner ist bei jedem der vorstehenden Ausführungsbeispiele ein Fall beschrieben, in dem ein Programm in einem Speicherspeicher (ROM) gespeichert ist. Dies erlegt jedoch der Erfindung keine Beschränkung auf, da das Programm in jedem Speichermedium gespeichert werden kann. Die durch das Programm durchgeführten gleichen Operationen können durch Schaltungen ausgeführt werden.

[0052] Die Erfindung kann auf ein durch eine Vielzahl von Vorrichtungen gebildetes System oder auf ein Gerät mit einer einzelnen Vorrichtung (zum Beispiel einen Kopierer oder ein Faxgerät usw.) angewendet werden.

[0053] Ferner ist es selbstverständlich, daß die Erfindung auch auf einen Fall anwendbar ist, in dem das Ziel der Erfindung erreicht wird, indem ein die Programmcodes der Software zum Durchführen der Funktionen des vorstehenden Ausführungsbeispiels bei einem System oder einem Gerät speicherndes

Speichermedium oder ein diese tragendes Trägersignal zugeführt wird, die Programmcodes mit einem Computer (zum Beispiel einer CPU oder MPU) des Systems oder Geräts von dem Speichermedium gelesen werden und daraufhin die Programmcodes ausgeführt werden.

[0054] In diesem Fall realisieren die von dem Speichermedium gelesenen Programmcodes die neuen Funktionen der Erfindung, und das die Programmcodes speichernde Speichermedium bildet die Erfindung.

[0055] Ferner kann das Speichermedium wie beispielsweise eine Diskette, eine Festplatte, eine optische Platte, eine magneto-optische Platte, eine CD-ROM, eine CD-R, ein Magnetband, eine Speicherkarte des nichtflüchtigen Typs oder ein ROM zur Bereitstellung der Programmcodes verwendet werden.

[0056] Ferner ist es neben dem Fall, in dem die vorstehenden Funktionen gemäß dem Ausführungsbeispiel durch ein Ausführen der durch einen Computer gelesenen Programmcodes realisiert werden, selbstverständlich, daß die Erfindung einen Fall abdeckt, in dem ein Betriebssystem oder dergleichen, das auf dem Computer läuft, einen Teil des Prozesses oder den ganzen Prozeß gemäß der Bezeichnung von Programmcodes durchführt und die Funktionen gemäß den Ausführungsbeispielen realisiert.

[0057] Es ist selbstverständlich, daß die Erfindung zudem einen Fall abdeckt, in dem nach dem Schreiben der von dem Speichermedium gelesenen Programmcodes in eine in den Computer eingesetzte Funktionserweiterungsplatine oder in einen in einer mit dem Computer verbundenen Funktionserweiterungseinheit bereitgestellten Speicher eine in der Funktionserweiterungsplatine oder Funktionserweiterungseinheit enthaltene CPU oder dergleichen einen Teil des Prozesses oder den ganzen Prozeß gemäß der Bezeichnung von Programmcodes durchführt und die Funktion des vorstehenden Ausführungsbeispiels realisiert.

[0058] Somit ist es gemäß der Erfindung wie vorstehend beschrieben möglich, eine Sprachsynthesevorrichtung bereitzustellen, die zur Auswahl besserer Phonemeinheiten in der Lage ist, wobei als ein Ergebnis davon synthetisierte Sprache von hervorragender Qualität erzeugt werden kann. Die Erfindung stellt auch ein Verfahren zur Steuerung dieser Vorrichtung und eine ein Programm zum Realisieren dieses Steuerverfahrens speichernde Speichereinheit bereit.

[0059] Da viele offensichtlich sehr verschiedene Ausführungsbeispiele der Erfindung ausgebildet werden können, ohne von ihrem Bereich abzuweichen, ist es selbstverständlich, daß die Erfindung nicht auf die vorstehend beschriebenen spezifischen Ausführungsbeispiele beschränkt ist.

## Patentansprüche

1. Sprachsynthesevorrichtung, mit:  
 einer Speichereinrichtung (**200**, **400**) zum Speichern mehrerer Einheiten von Phonemdaten;  
 einer Wiedergewinnungseinrichtung (S11, S12, **201**, **401**, S301, S501) zum Wiedergewinnen von Phonemdaten aus den in der Speichereinrichtung gespeicherten mehreren Einheiten von Phonemdaten gemäß gegebenen Wiedergewinnungsbedingungen;  
 einer ersten Strafuweisungseinrichtung (**203–207**, S302, S303, **403–407**, S502, S503) zum Sortieren von durch die Wiedergewinnungseinrichtung wiedergewonnenen Phonemdaten auf der Grundlage eines vorgeschriebenen Eigenschaftswerts und zum Zuweisen einer auf einem Eigenschaftswert basierenden Strafe zu jeder Einheit der Phonemdaten auf der Grundlage einer durch das Sortieren erhaltenen Reihenfolge; und  
 einer Auswahlrichtung (**208**, S304, **408**, S506) zum Auswählen von bei einer Synthese eines Sprachsignalverlaufs zu verwendenden Phonemdaten aus den durch die Wiedergewinnungseinrichtung wiedergewonnenen Phonemdaten und auf der Grundlage der durch die erste Strafuweisungseinrichtung zugewiesenen Strafe.

2. Vorrichtung nach Anspruch 1, wobei die Speichereinrichtung (**200**, **400**) jeweilige Einheiten von Eigenschaftsinformationen zusammen mit den mehreren Einheiten von Phonemdaten speichert; und die erste Strafuweisungseinrichtung (**203–207**, S302, S303, **403–407**, S502, S503) aus den in der Speichereinrichtung gespeicherten Eigenschaftsinformationen einen Eigenschaftswert erhält.

3. Vorrichtung nach Anspruch 2, wobei die Eigenschaftsinformationen eine Phonemumgebung, eine Phonemgrenze, eine Grundfrequenz, eine Energie und eine Phonemdauer umfassen.

4. Vorrichtung nach einem der vorstehenden Ansprüche, wobei die Wiedergewinnungseinrichtung (S11, S12, **201**, **401**, S301, S501) Phonemdaten wiedergewinnt, die eine bestimmte Phonemumgebung erfüllen.

5. Vorrichtung nach einem der vorstehenden Ansprüche, wobei die Wiedergewinnungseinrichtung (S11, S12, **201**, **401**, S301, S501) Phonemdaten wiedergewinnt, die eine bestimmte Phonemumgebung und Grundfrequenz erfüllen.

6. Vorrichtung nach einem der vorstehenden Ansprüche, wobei die erste Strafuweisungseinrichtung (**203–207**, S302, S303, **403–407**, S502) unter Verwendung einer Energie und einer Phonemdauer jeder Einheit von Phonemdaten als die Eigenschaftswerte eine Strafe zuweist.

7. Vorrichtung nach einem der vorstehenden Ansprüche, wobei die erste Strafuweisungseinrichtung (**203–207**, S302, S303, **403–407**, S502):

die Einheiten von Phonemdaten in einer Reihenfolge abnehmender Energie sortiert und eine energiebezogene Strafe auf der Grundlage der durch das Sortieren erhaltenen Reihenfolge derart zuweist, daß Phonemdaten, deren Energie dicht bei einem Durchschnittswert ist, eine kleine Strafe zugewiesen wird; und

die Einheiten von Phonemdaten in einer Reihenfolge abnehmender Phonemdauer sortiert und eine phonemdauerbezogene Strafe auf der Grundlage der durch das Sortieren erhaltenen Reihenfolge derart zuweist, daß Phonemdaten, deren Phonemdauer dicht bei einem Durchschnittswert ist, eine kleine Strafe zugewiesen wird.

8. Vorrichtung nach einem der vorstehenden Ansprüche, ferner mit:

einer Alternativwiedergewinnungseinrichtung (**401**, S501) zum Wiedergewinnen von Phonemdaten, die einige der Wiedergewinnungsbedingungen erfüllen, in einem Fall, in dem keine Phonemdaten vorhanden sind, die den Wiedergewinnungsbedingungen in der Wiedergewinnungseinrichtung entsprechen;  
 einer Zähleinrichtung (S504, S505) zum Gruppieren von Phonemdaten, die durch die Alternativwiedergewinnungseinrichtung wiedergewonnen worden sind, auf der Grundlage einer Phonemumgebung und zum Zählen der Einheiten von Phonemdaten auf einer Pro-Gruppe-Grundlage; und  
 einer zweiten Strafuweisungseinrichtung (**410**, S505) zum Zuweisen einer Strafe zu den durch die Alternativwiedergewinnungseinrichtung wiedergewonnenen Phonemdaten auf der Grundlage eines durch die Zähleinrichtung erhaltenen Zählwerts, wobei diese Strafe zusätzlich zu der durch die erste Strafuweisungseinrichtung zugewiesenen Strafe zugewiesen wird.

9. Vorrichtung nach Anspruch 8, wobei die Wiedergewinnungsbedingungen eine Phonemumgebung umfassen; und

die Alternativwiedergewinnungseinrichtung (**401**, S501) Phonemdaten wiedergewinnt, die mit einem Teil einer in den Wiedergewinnungsbedingungen bestimmten Phonemumgebung übereinstimmen.

10. Vorrichtung nach Anspruch 9, wobei die in den Wiedergewinnungsbedingungen bestimmte Phonemumgebung ein aus einem anwendbaren Phonem und Phonemen auf beiden Seiten davon bestehendes Triphon ist; und

die Alternativwiedergewinnungseinrichtung (**401**, S501) Phonemdaten, für die das anwendbare Phonem und das Phonem zu seiner Linken mit den Wiedergewinnungsbedingungen übereinstimmen, oder Phonemdaten, für die das anwendbare Phonem und das Phonem zu seiner Rechten mit den Wiederge-



winnungsbedingungen übereinstimmen, wiedergewinnt.

11. Sprachsyntheseverfahren, mit:  
 einem Speicherschnitt des Speicherns mehrerer Einheiten von Phonemdaten;  
 einem Wiedergewinnungsschnitt (S11, S12, S301, S501) des Wiedergewinnens von Phonemdaten aus den in dem Speicherschnitt gespeicherten mehreren Einheiten von Phonemdaten gemäß gegebenen Suchwiedergewinnungsbedingungen;  
 einem ersten Strafuweisungsschnitt (S302, S303, S502, S503), der in dem Wiedergewinnungsschnitt wiedergewonnene Phonemdaten auf der Grundlage eines vorgeschriebenen Eigenschaftswerts sortiert und der auf der Grundlage einer durch das Sortieren erhaltenen Reihenfolge jeder Einheit der Phonemdaten eine auf einem Eigenschaftswert basierende Strafe zuweist; und einem Auswahlsschnitt (**S304, S506**) des Auswählens von bei einer Synthese eines Sprachsignalverlaufs verwendeten Phonemdaten aus den in dem Wiedergewinnungsschnitt wiedergewonnenen Phonemdaten und auf der Grundlage der in dem Strafuweisungsschnitt zugewiesenen Strafe.

12. Verfahren nach Anspruch 11, wobei der Speicherschnitt jeweilige Einheiten von Eigenschaftsinformationen zusammen mit den mehreren Einheiten von Phonemdaten speichert; und der erste Strafuweisungsschnitt (**S302, S303, S502, S503**) aus den in dem Speicherschnitt gespeicherten Eigenschaftsinformationen einen Eigenschaftswert erhält.

13. Verfahren nach Anspruch 12, wobei die Eigenschaftsinformationen ein Phonemetikett, eine Phonemgrenze, eine Grundfrequenz, eine Energie und eine Phonemdauer umfassen.

14. Verfahren nach einem der Ansprüche 11 bis 13, wobei der Wiedergewinnungsschnitt (**S12, S301, S501**) Phonemdaten wiedergewinnt, die eine bestimmte Phonemumgebung erfüllen.

15. Verfahren nach einem der Ansprüche 11 bis 14, wobei der Wiedergewinnungsschnitt (**S12, S301, S501**) Phonemdaten wiedergewinnt, die eine bestimmte Phonemumgebung und Grundfrequenz erfüllen.

16. Verfahren nach einem der Ansprüche 11 bis 15, wobei der erste Strafuweisungsschnitt (**S302, S303, S502, S503**) unter Verwendung einer Energie und einer Phonemdauer jeder Einheit von Phonemdaten als die Eigenschaftswerte eine Strafe zuweist.

17. Verfahren nach Anspruch 16, wobei der erste Strafuweisungsschnitt (**S302, S303, S502, S503**): die Einheiten von Phonemdaten in einer Reihenfolge abnehmender Energie sortiert und eine energiebezogene Strafe auf der Grundlage der durch das Sortie-

ren erhaltenen Reihenfolge derart zuweist, daß Phonemdaten, deren Energie dicht bei einem Durchschnittswert ist, eine kleine Strafe zugewiesen wird; und die Einheiten von Phonemdaten in einer Reihenfolge abnehmender Phonemdauer sortiert und eine phonemdauerbezogene Strafe auf der Grundlage der durch das Sortieren erhaltenen Reihenfolge derart zuweist, daß Phonemdaten, deren Phonemdauer dicht bei einem Durchschnittswert ist, eine kleine Strafe zugewiesen wird.

18. Verfahren nach einem der Ansprüche 11 bis 17, ferner mit:  
 einem Alternativwiedergewinnungsschnitt (**S501**) des Wiedergewinnens von Phonemdaten, die einige der Wiedergewinnungsbedingungen erfüllen, in einem Fall, in dem keine Phonemdaten vorhanden sind, die den Wiedergewinnungsbedingungen in dem Wiedergewinnungsschnitt entsprechen;  
 einem Zählsschnitt (**S504, S505**) des Gruppierens von Phonemdaten, die in dem Alternativwiedergewinnungsschnitt wiedergewonnen worden sind, auf der Grundlage einer Phonemumgebung, und des Zählens der Einheiten von Phonemdaten auf einer Pro-Gruppe-Grundlage; und  
 einem zweiten Strafuweisungsschnitt (**S505**) des Zuweisens einer Strafe zu den in dem Alternativwiedergewinnungsschnitt wiedergewonnenen Phonemdaten auf der Grundlage eines in dem Zählsschnitt erhaltenen Zählwerts, wobei diese Strafe zusätzlich zu der in dem ersten Strafuweisungsschnitt zugewiesenen Strafe zugewiesen wird.

19. Verfahren nach Anspruch 18, wobei die Wiedergewinnungsbedingungen eine Phonemumgebung umfassen; und der Alternativwiedergewinnungsschnitt Phonemdaten wiedergewinnt, die mit einem Teil einer in den Wiedergewinnungsbedingungen bestimmten Phonemumgebung übereinstimmen.

20. Verfahren nach Anspruch 19, wobei die in den Wiedergewinnungsbedingungen bestimmte Phonemumgebung ein aus einem anwendbaren Phonem und Phonemen auf beiden Seiten davon bestehendes Triphon ist; und der Alternativwiedergewinnungsschnitt (**S501**) Phonemdaten, für die das anwendbare Phonem und das Phonem zu seiner Linken mit den Wiedergewinnungsbedingungen übereinstimmen, oder Phonemdaten, für die das anwendbare Phonem und das Phonem zu seiner Rechten mit den Wiedergewinnungsbedingungen übereinstimmen, wiedergewinnt.

21. Speichermedium, das ein Steuerprogramm zur Veranlassung eines Computers zum Ausführen einer Sprachsynthese unter Verwendung von Phonemdaten speichert, wobei das Steuerprogramm aufweist:

Code eines Speicherschritts des Speicherns mehrerer Einheiten von Phonemdaten;

Code eines Wiedergewinnungsschritts des Wiedergewinnens von Phonemdaten aus den in dem Speicherschritt gespeicherten mehreren Einheiten von Phonemdaten gemäß gegebenen Suchwiedergewinnungsbedingungen;

Code eines ersten Strafzuweisungsschritts, der in dem Wiedergewinnungsschritt wiedergewonnene Phonemdaten auf der Grundlage eines vorgeschriebenen Eigenschaftswerts sortiert und der auf der Grundlage einer durch das Sortieren erhaltenen Reihenfolge jeder Einheit der Phonemdaten eine auf einem Eigenschaftswert basierende Strafe zuweist; und

Code eines Auswahlsschritts des Auswählens von bei einer Synthese eines Sprachsignalverlaufs verwendeten Phonemdaten aus den in dem Wiedergewinnungsschritt wiedergewonnenen Phonemdaten und auf der Grundlage der in dem ersten Strafzuweisungsschritt zugewiesenen Strafe.

22. Speichermedium nach Anspruch 21, wobei das Steuerprogramm ferner aufweist:

Code eines Alternativwiedergewinnungsschritts des Wiedergewinnens von Phonemdaten, die einige der Wiedergewinnungsbedingungen erfüllen, in einem Fall, in dem keine Phonemdaten vorhanden sind, die den Wiedergewinnungsbedingungen in dem Wiedergewinnungsschritt entsprechen;

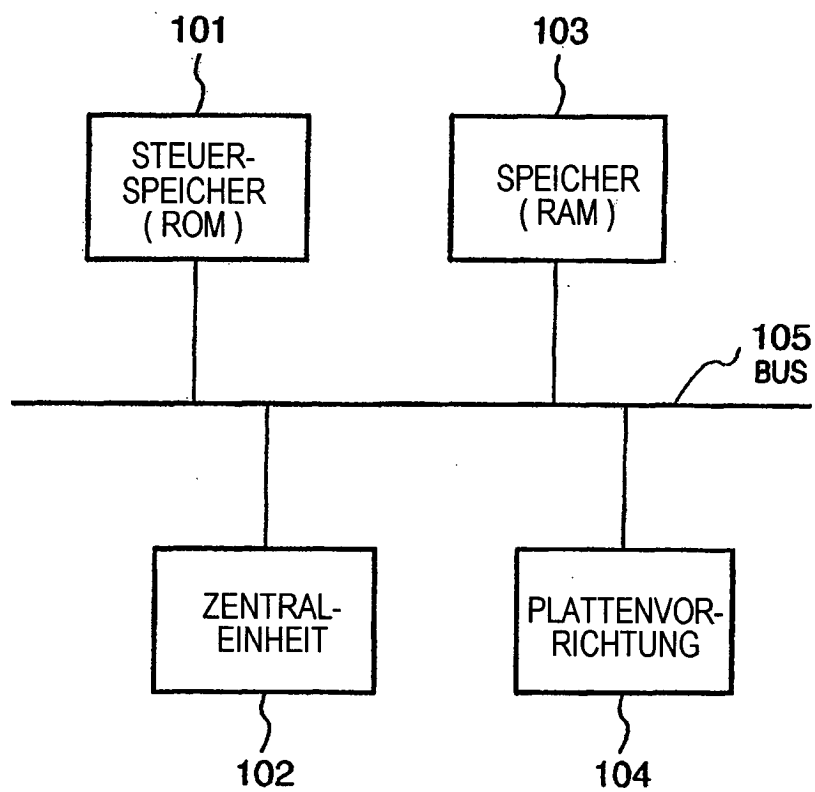
Code eines Zählsschritts des Gruppierens von Phonemdaten, die in dem Alternativwiedergewinnungsschritt wiedergewonnen worden sind, auf der Grundlage einer Phonenumgebung, und des Zählens der Einheiten von Phonemdaten auf einer Pro-Gruppe-Grundlage; und

Code eines zweiten Strafzuweisungsschritts des Zuweisens einer Strafe zu den in dem Alternativwiedergewinnungsschritt wiedergewonnenen Phonemdaten auf der Grundlage eines in dem Zählschritt erhaltenen Zählwerts, wobei diese Strafe zusätzlich zu der in dem ersten Strafzuweisungsschritt zugewiesenen Strafe zugewiesen wird.

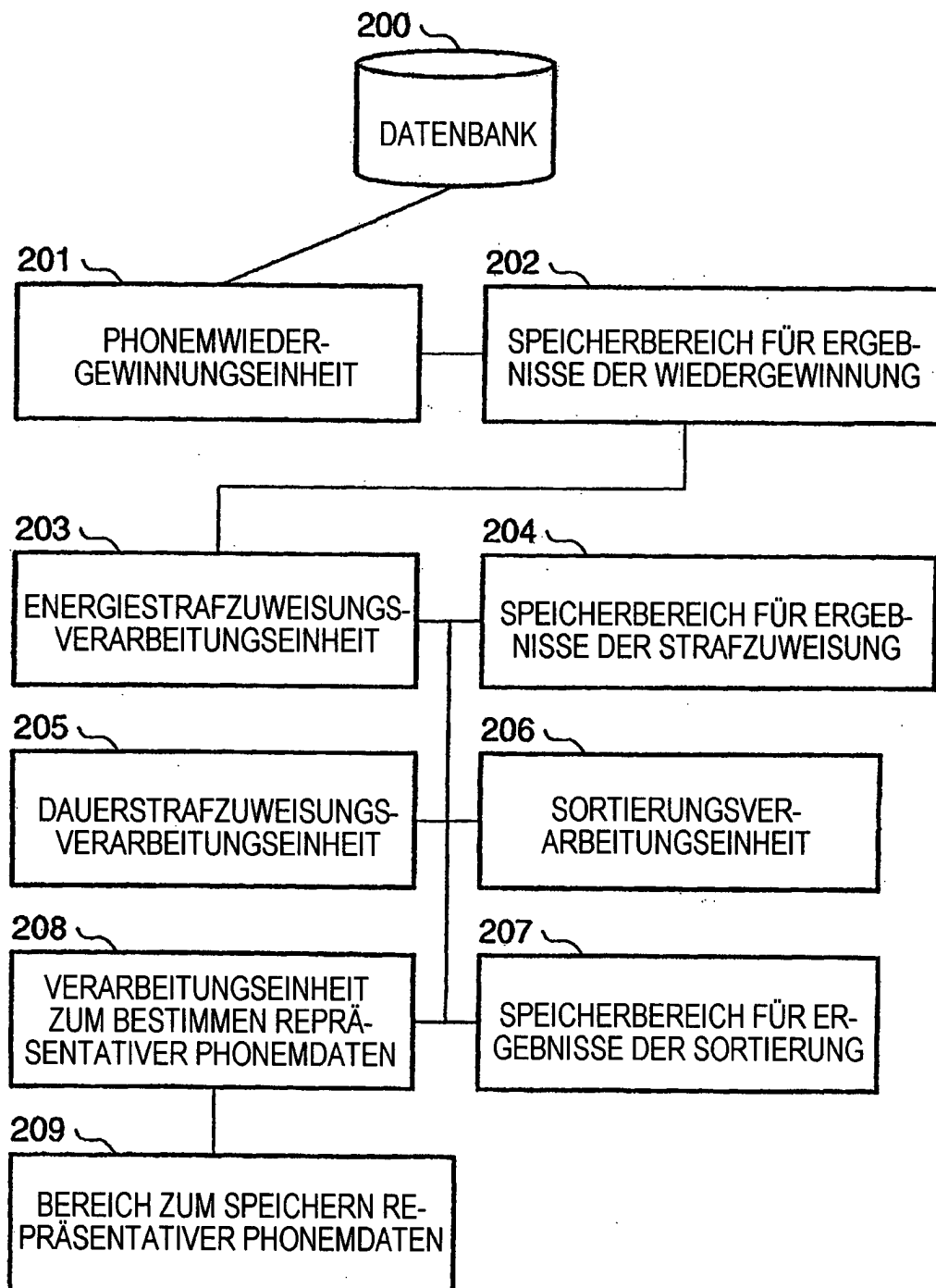
23. Von einer Verarbeitungseinrichtung ausführbare Anweisungen zur Steuerung einer Verarbeitungseinrichtung zum Ausführen aller Schritte des Verfahrens nach einem der Ansprüche 11 bis 20.

Es folgen 6 Blatt Zeichnungen

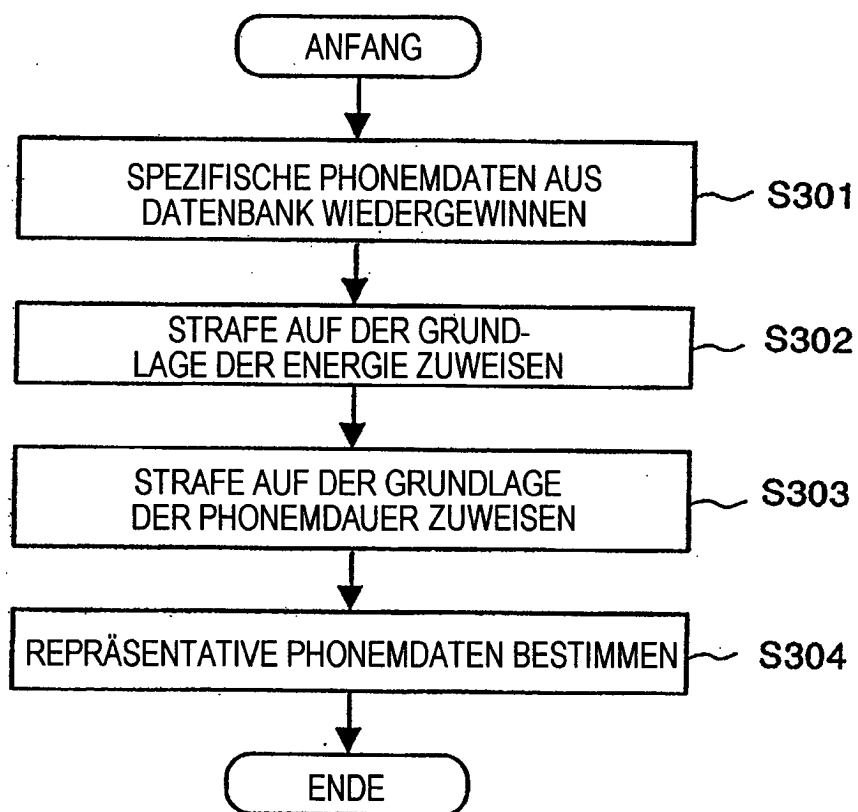
**FIG. 1**

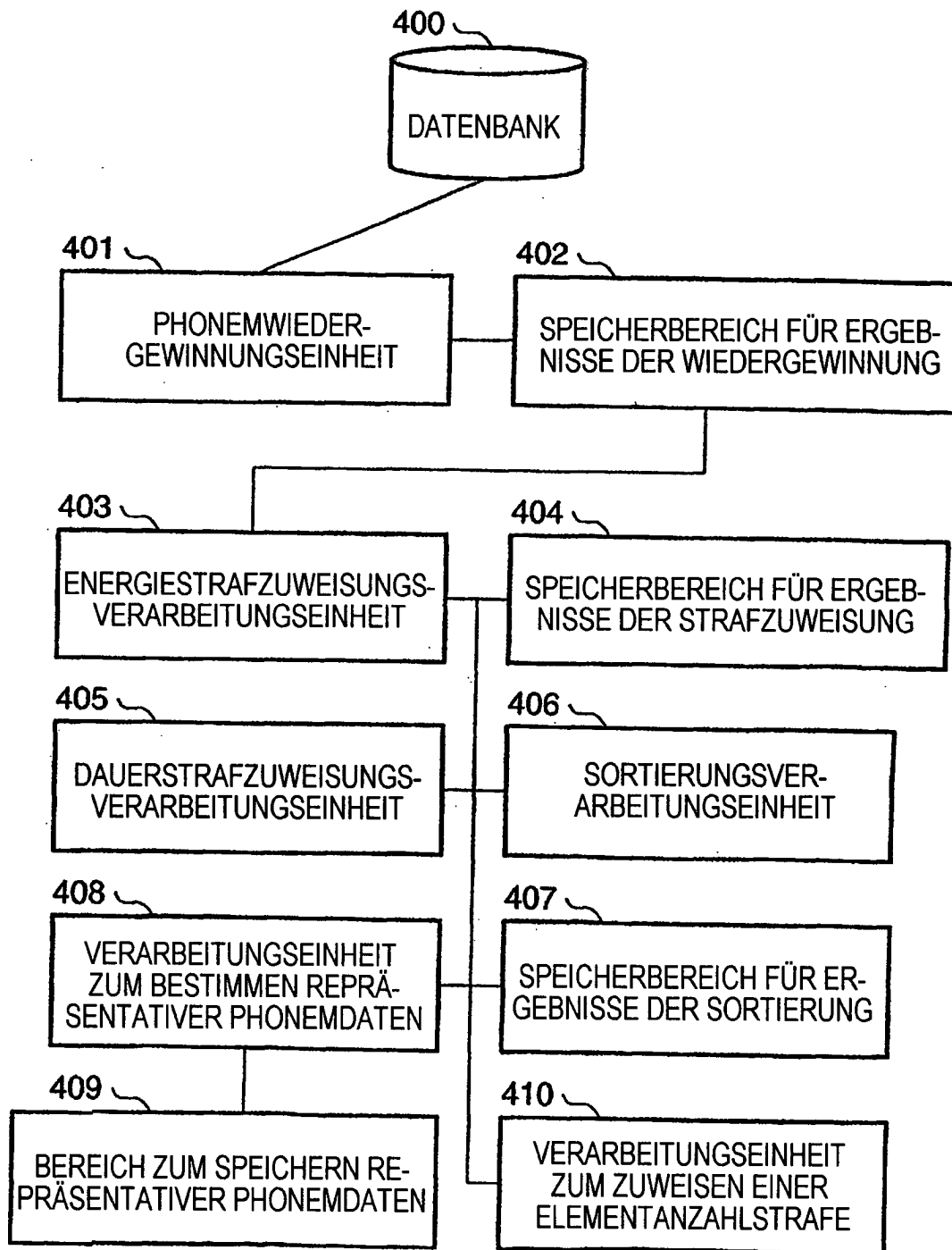


**FIG. 2**

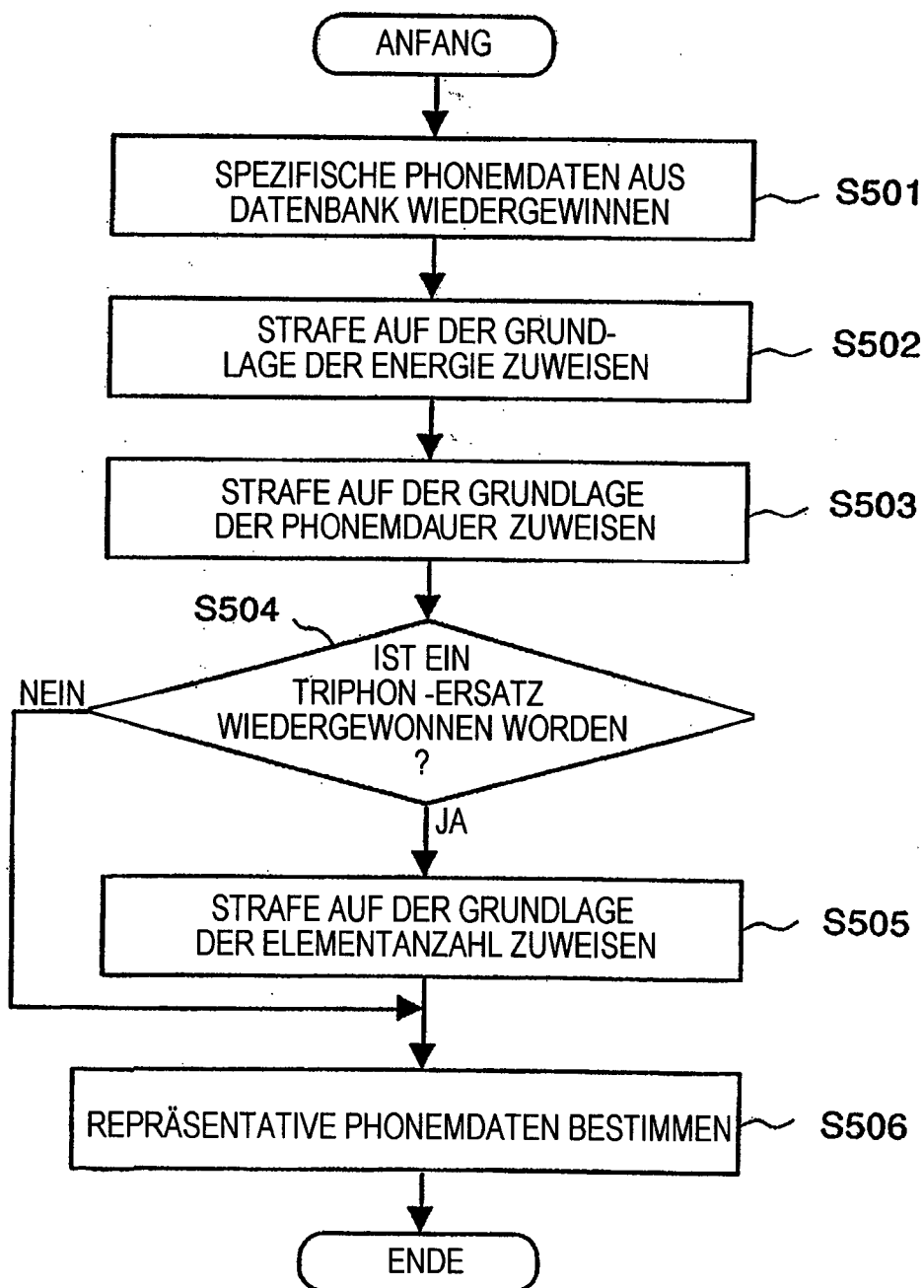


**FIG. 3**



**FIG. 4**

**FIG. 5**



**FIG. 6**

