

(19)



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11)

EP 1 033 702 A2

(12)

EUROPÄISCHE PATENTANMELDUNG

(43) Veröffentlichungstag:
06.09.2000 Patentblatt 2000/36

(51) Int. Cl.⁷: **G10L 21/02**

(21) Anmeldenummer: **00104304.1**

(22) Anmeldetag: **02.03.2000**

(84) Benannte Vertragsstaaten:
**AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE**
Benannte Erstreckungsstaaten:
AL LT LV MK RO SI

(72) Erfinder: **Schröder, Hehrwart
88662 Überlingen (DE)**

(30) Priorität: **02.03.1999 DE 19909150**

(74) Vertreter:
**Patentanwälte
Eisele, Otten, Roth & Dobler
Karlstrasse 8
88212 Ravensburg (DE)**

(71) Anmelder: **Schröder, Hehrwart
88662 Überlingen (DE)**

(54) **Verfahren zur Trennung von Sprache und Geräuschen**

(57) Die Erfindung betrifft ein Verfahren zur Trennung von Sprache und Geräuschen, mit folgenden Verfahrensschritten:

- Trennen von Sprache und Nichtsprache mit Hilfe der Symmetriegrenze.

- Zerlegung des Sprachsignals in Einzelklänge.
- Transformation des Sprachsignals von der Darstellung Intensität über die Zeit in eine Darstellung Amplitude über die Frequenz und Frequenz über die Zeit.
- Im Frequenzraum wird ein Band um die Intensitäten gelegt.
- Mittels zahlentheoretischer Funktionen werden Intervalle der Sprache berechnet und ein Raum der Äquivalenzklassen der Sprachtöne erzeugt.

Die eigentliche Trennung von Sprache und Geräuschen erfolgt anschließend durch eines oder mehreren der folgenden Verfahrensschritte :

- Die Frequenzen innerhalb eines Lautes (Hüllkurve) werden durch geeignete Schleifenverfahren miteinander paarweise verglichen.
- Trennen von Sprache und Nichtsprache mit Hilfe eines Amplitudenbandes im Frequenzraum.
- Aus dem Verhältnis der Tonhöhendifferenz wird durch eine logarithmische Beziehung zwischen den Amplituden eine neue Beziehung gebildet und für die Trennung Geräusch/Sprache genutzt.

EP 1 033 702 A2

Beschreibung

[0001] Die Erfindung betrifft ein Verfahren zur Trennung von Sprache und Geräuschen.

5 [0002] Aufgabe der Erfindung ist es, ein Verfahren zu schaffen, mit der eine zuverlässige Trennung von Sprache und Geräuschen, möglichst in Echtzeit, erreicht wird.

[0003] Diese Aufgabe wird mit dem Verfahren nach Anspruch 1 gelöst.

[0004] Der diesen Unterlagen beigeheftete Bericht "Trennung Sprache von Nichtsprache auf gruppentheoretischer Basis", Stand vom Dienstag, den 29.02.2000, mit den zugehörigen Testdiagrammen ist Bestandteil der Beschreibung der Erfindung und dient der Ergänzung des Offenbarungsgehalts.

10 [0005] Das erfindungsgemäße Verfahren umfaßt die folgenden Schritte:

Vorbereitende Rechenschritte

1. Schritt:

15 [0006] Das von Störungen zu befreiende digitalisierte Sprachsignal wird in seine Einzelklänge mit Hilfe einer Hüllkurvenmethode zerlegt.

2. Schritt:

20 [0007] Mittels Integraltransformationen wird das Sprachsignal von der Darstellung Intensität über die Zeit in eine Darstellung Amplitude über die Frequenz und Frequenz über die Zeit transformiert.

3. Schritt:

25 [0008] Im Frequenzraum wird ein Band um die Intensitäten gelegt. Hierdurch wird ein Vergleich aller übrigen Intensitäten nur innerhalb dieses Bandes zugelassen. Die Größe des Bandes hängt unter anderem von der Intensität beziehungsweise von der Lautstärke ab.

30 4. Schritt:

[0009] Mittels zahlentheoretischer Funktionen werden Intervalle der Sprache berechnet und ein Raum der Äquivalenzklassen der Sprachtöne erzeugt.

35 **Trennung Sprache von Nichtsprache**

5. Schritt:

40 [0010] Die Frequenzen innerhalb eines Lautes (Hüllkurve) werden durch geeignete Schleifenverfahren miteinander paarweise verglichen. Es entstehen dann für jedes Frequenzpaar aus der Teilung Intervalle. Diese Intervalle werden mit den berechneten Intervallen der Äquivalenzklassen verglichen. Als Ergebnis dieser Überprüfung ergibt sich, ob die betrachteten Frequenzen miteinander harmonische Beziehungen haben oder nicht. Diese harmonischen Beziehungen haben Sprachfrequenzen mit allen anderen Sprachfrequenzen innerhalb einer Hüllkurve oder eines Lautes. Nur dann, wenn im Raum der Äquivalenzklassen ein Bezugsintervall gefunden wird, wird die entsprechende Frequenz in die Datei Sprache überwiesen. Wenn es kein Bezugsintervall gibt, wird die Frequenz in die Datei Nichtsprache überwiesen.

6. Schritt

50 Trennen von Sprache und Nichtsprache mit Hilfe eines Amplitudenbandes im Frequenzraum.

[0011] Wenn das Band um die Amplituden sehr viel kleiner als 1 gemacht wird, dann wechselwirken die Intervalle der Äquivalenzklassen die ihrerseits kleiner als 1 sind mit den zu vergleichenden Frequenzen des zu untersuchenden Sprachsignals. Dieser Effekt trennt einen Teil der Nichtsprache von Sprache.

55 7. Schritt:

[0012] Aus dem Verhältnis der Tonhöhendifferenz wird durch eine logarithmische Beziehung zwischen den Ampli-

tuden eine neue Beziehung gebildet und für die Trennung Geräusch/Sprache genutzt. Für das Ergebnis dieser Beziehung wird die Prozedur wiederholt, in dem Raum der Äquivalenzklassen ein Bezugsintervall zu suchen.

8. Schritt:

5

Trennen von Sprache und Nichtsprache mit Hilfe der Symmetriegrenze

[0013] Wenn das Frequenzverhältnis größer als 8 wird, dann nähert man sich der Symmetriegrenze für Wohlklang oder Konsonanz. Wenn man Frequenzen entfernt, deren Intervalle größer als ein bestimmter Faktor werden, dann handelt es sich um Nichtsprache.

10

[0014] Die einzelnen Trennverfahren (Schritte 5 bis 8) können durch geeignete gruppentheoretische Verfahren miteinander verbunden und optimiert werden. Sie können auch einzeln eingesetzt werden.

[0015] Vorteilhafte Ausführungen des erfindungsgemäßen Verfahrens werden im folgenden beschrieben.

15

[0016] Die stimmhaften Sprachlaute unseres Sprechorgans sind die Harmonische Klänge. Die harmonischen Verwandtschaftsbeziehungen zwischen den Teiltönen stellen die unveränderlichen (invarianten) Eigenschaften der Klänge dar. Bei dem Lernprozeß, welcher von frühestem Alter an die Sprachlauterkennung ermöglicht, erlernt der Mensch mit seinem Gehör die invarianten Klangeigenschaften. Denn dadurch wird es in die Lage versetzt, informationstragende von nichtinformationstragenden Klangmerkmalen zu unterscheiden.

20

Berechnen der einhüllenden Kurve eines Klanges (Hüllkurve)

[0017] Um die Sprachsignale in Echtzeit verarbeiten zu können, ist es erforderlich, daß das Signal Klang für Klang einzeln der Reihe nach abgearbeitet werden kann. Um ein willkürliches Abschneiden des Signales für die Verarbeitung zu verhindern, wird ein individuelles flexibles Verfahren gebraucht. Dazu wird die Hüllkurve berechnet, denn sie liefert Anfangs-, Endpunkt und die Klangdauer.

25

[0018] Aus dem digitalisierten Sprachsignal werden mit Hilfe der Differentialrechnung zunächst die Stützpunkte der Hüllkurven gewonnen, die alle einzelne Klänge einschließen, gleichgültig wie stark gestört das Signal ist. Alle Einschnürungen im Signal werden als Anfangs- und Endpunkte der Hüllkurve definiert. Damit kann dann das Signal Klang für Klang verarbeitet werden.

30

[0019] Das Differenzieren der Hüllkurve führt zu den Minima der Hüllkurve. Die Minima der Hüllkurve bilden Anfangs- und Endpunkte der Hüllkurven. Die Hüllkurven der Klänge enthalten unterschiedlich viele Frequenzen und sind deshalb zeitlich verschieden lang.

Transformation des Sprachsignals mit Integraltransformationen

35

[0020]

Fouriertransformation:

Darstellung des Sprachsignales im Frequenzraum: Amplitude als Ordinate und Frequenz als Abszisse

40

Hilberttransformation:

Darstellung des Sprachsignales im Frequenzraum: Frequenz als Ordinate und Zeit als Abszisse

[0021] Jedes Wellenpaket (jeder Klang) enthält n Summanden. Jeder n-te Summand ist die n-te Partialschwingung von dem Wellenpaket.

45

[0022] Das Ursprungssignal wird mit Integraltransformationen so transformiert, daß es als Amplitude über die Frequenz, und als Frequenz über die Zeit für den Datensatz jeder einzelnen Hüllkurve zur weiteren Verarbeitung zur Verfügung steht.

[0023] Im Fourierraum beispielsweise sind die Amplituden beziehungsweise die Intensitäten jeweils für eine Hüllkurve über die Frequenz aufgetragen. Hier wird mit einem Amplitudenband innerhalb des Frequenzraumes um die Intensitäten ein Vergleich aller übrigen Intensitäten nur innerhalb dieses Bandes zugelassen. Die Größe der Amplitudenbandes hängt von der Intensität beziehungsweise der Lautstärke ab.

50

[0024] Alle zu den Intensitäten gehörenden Frequenzen sind im Funktionenraum der Hilberttransformation als Frequenzen über die Zeit vorhanden.

55

Entfaltung des Signals mit Integraltransformationen

[0025] In dem digitalisierten Sprachsignal als Ausgangssignal ist die Information als Intensität über die Zeit in Form

von Frequenzen in einem zweidimensionalen Feld kodiert.

[0026] Die Integraltransformationen werden als lineare Operatoren der Gruppentheorie definiert und zur Entfaltung des Signals eingesetzt. Beispielsweise werden Fourier- und Hilberttransformation eingesetzt. Sie transformieren das Signal in einen für die nachfolgende Verarbeitung erforderlichen Zustand in Funktionenräume, beispielsweise in den Frequenzraum.

Töne und Klänge

[0027] Ein Tonsystem ist ein geordnetes Paar von Elementen. Dabei ist das Frequenzintervall eine Gruppe und die Intervalle sind eine injektive Abbildung des Tonsystems in die Menge aller positiven reellen Zahlen. Eine endliche Teilmenge aus den Frequenzen heißt Klang. Ein geordnetes n-Tupel von Tönen wird geordneter n-Klang genannt

Ordnen des Sprachsignales durch Kongruenzgruppen

[0028]

Zahlentheoretische Funktionen aus Primzahlen und Zweierpotenzen bilden Felder für Äquivalenzklassen. Die Äquivalenzklassen werden zu dem Sprachsignal in Form von Äquivalenzrelationen in Beziehung gebracht. Algebraische Kongruenzgruppen ordnen das Sprachsignal in Sprache und Nichtsprache in verschiedene Dateien. Rückgewinnung des Sprachsignales im Zeitbereich durch Integralrücktransformationen

Äquivalenzrelationen

[0029] Elementvertauschungen werden durch eine zwei- oder mehrstellige Relation auf der Menge aller Klänge eines Tonsystems präzisiert. Dabei gilt für ein Element aus der Menge des Sprachsignals die Zuordnung zu Sprache oder Nichtsprache genau dann, wenn zu den Intervallen des zu prüfenden Signals ein Bezugston aus der Äquivalenzklasse der zahlentheoretischen Funktionen des Primzahlkörpers existiert.

Berechnung der Intervalle aus den Frequenzen des Sprachsignals

[0030] Alle Frequenzen des Frequenzraumes des transformierten Sprachsignales müssen überprüft werden, ob sie Symmetriebeziehungen mit anderen Frequenzen der gleichen Hüllkurve haben. In den Symmetriebeziehungen der Elemente liegen invariante Eigenschaften der Sprache verschlüsselt vor.

[0031] Von dem zu verarbeitendem Signal werden deshalb innerhalb der Hüllkurve alle Frequenzen paarweise der Reihe nach mit allen anderen Frequenzen dividiert, um Intervalle zu bilden. Dies geschieht solange, bis jede Frequenz in der fortlaufenden Reihenfolge nach und nach durch alle anderen Frequenzen in der gleichen Hüllkurve geteilt wurde, um vorübergehend zu einer Verhältniszahl (Intervall) dieses Frequenzpaares zu kommen.

[0032] Die erste Frequenz wird durch die zweite, der dritten und so fort geteilt, bis zum Ende der Hüllkurve. Dann wird die zweite durch die dritte, der vierten und so fort geteilt. In einer Hüllkurve können beispielsweise jeweils tausende Frequenzen sein. Dabei darf nie die Reihenfolge der Frequenzen vertauscht werden.

[0033] Die Trennung von Sprache und Nichtsprache wird mit einer Folge von mehreren Symmetrieoperationen durchgeführt.

1. Äquivalenzklassen mit Bezugsintervall als Entscheidungskriterium

[0034] Nach jeder Intervallbildung innerhalb des Frequenzraumes des zu überprüfenden Sprachsignals muß ein Bezugsintervall aus den Äquivalenzklassen der berechneten Intervalle gefunden werden.

[0035] Die Äquivalenzklassen wiederum werden durch verschiedene zahlentheoretische Funktionen berechnet. Die Trennwirkung wird deutlich besser, wenn man die Erfüllung eines Bezugsintervalles für jede zahlentheoretische Funktion fordert. Wenn dies n Funktionen sind, (wobei n = natürliche ganze Zahlen sind) dann spannen diese zahlentheoretischen Funktionen einen n-dimensionalen metrischen Raum auf.

Wenn das aus dem zu prüfenden Signal gewonnene Intervall in allen zahlentheoretischen Funktionen genau eine Bezugsordinate findet, dann sind alle zahlentheoretischen Funktionen mindestens einmal erfüllt. Die Koordinaten führen zu genau einem Bezugsintervall aus dem Raum der Äquivalenzklassen.

[0036] Über diesen Bezugston können genau dann weitere Sprachparameter wie beispielsweise Anklingzeit und Abklingzeit eines Lautes abgefragt werden, wenn diesem n-dimensionalen Raum eine Metrik aufgeprägt wird. Dies

kann mit Hilfe der Gruppentheorie auf der Basis von Matrizen-Operationen durchgeführt werden.

[0037] Dieses Bezugsintervall kann als Bezugston aufgefaßt werden. Dann können viele weitere Parameter für die Spracherkennung herangezogen werden, wie beispielsweise die Anklingzeit und die Abklingzeit eines Tones.

5 **[0038]** Gibt es ein Bezugsintervall, dann muß sich das Rechenverfahren diese Elemente so lange merken, bis alle Elemente miteinander verglichen wurden. Erst nach Abschluß dieses Sortierungsprozesses weiß das Verfahren, welche Elemente mit anderen Intervallbeziehungen haben.

[0039] Alle die Frequenzen, die über die Intervalle Beziehungen miteinander haben, stehen schließlich mit allen anderen Frequenzen innerhalb der gleichen Hüllkurve und damit des gleichen Lautes der Sprache in Beziehung. Das Verfahren merkt sich all diese Elemente. Am Ende des Rechenprozesses werden alle solche Elemente in den Speicher
10 für Sprache abgelegt. Alle anderen Frequenzen, die sich nicht in dieses Beziehungssystem einbinden lassen, sind beziehungslose Elemente und kommen in den Speicher für Nichtsprache. Ein Teil der Nichtsprache wird hierdurch bereits von der Sprache getrennt.

[0040] Elemente, die miteinander durch symmetrische Beziehungen dieser Intervalle verbunden sind, sind Elemente der Sprache. Elemente die in keinem symmetrischen Zusammenhang mit allen anderen Elementen innerhalb
15 der Hüllkurve stehen sind keine Sprache. Sie übertragen keine Information.

2. Verbessertes Trennen von Sprache und Nichtsprache durch eine Verknüpfung der Tonhöhendifferenz mit den Intensitäten

20 **[0041]** Es wird eine zweistellige Relation der Intervalle so definiert, daß ein Bezugston innerhalb von mehreren Koordinaten des Äquivalenzklassenraumes dem Element aus dem zu untersuchenden Signal genau entspricht.

[0042] Die Elemente der existierenden Faktormenge der Frequenzintervalle aus den Äquivalenzklassen werden nun als Harmonien definiert.

25 **[0043]** In der folgenden Formel wird die Tonhöhendifferenz mit dem logarithmischen Verhältnis der Amplituden in Beziehung gebracht.

$$V = \exp(\text{abs}((\text{Faktor} \cdot (\ln(f1/f2))))/(\text{Faktor} \cdot \ln(A1/A2)));$$

$$f1/f2 = \text{Integerdivision}$$

30

es bedeuten:

V = Intervall aus den Äquivalenzklassen aller zahlentheoretischen Funktionen;

f1/f2 = Verhältnis der Frequenzen = Tonhöhenabstand, (A1/A2) = Verhältnis der Intensitäten

35

[0044] Mit Hilfe dieser Beziehung wird ein Bezugston definiert.

[0045] Dann wiederholt sich das vorhergehende Verfahren.

[0046] Diese Gruppenoperation verbessert die Qualität der Trennung Sprache von Nichtsprache.

40 3. Trennen von Sprache und Nichtsprache mit Hilfe einer Amplitudenbandbeziehung im Frequenzraum

[0047] Wenn das Band um die Amplituden sehr viel kleiner als 1 gemacht wird, dann wechselwirken die Intervalle der Äquivalenzklassen die ihrerseits kleiner als 1 sind mit den zu vergleichenden Frequenzen des zu untersuchenden Sprachsignals.

45 Beispielsweise sei das Maximum des ersten Wellenpaketes bei 500 Einheiten. Um alle zu vergleichenden Amplituden sei das Amplitudenband mit 50 Einheiten Bandbreite durch den ganzen Frequenzraum als Eingangsband für das Verarbeitungsprogramm eingestellt.

[0048] Es entsteht ein neuer Effekt, wenn diesem Verarbeitungsprogramm ein Vorprogramm vorgeschaltet wird, bei dem nur die Bandbreite des Amplitudenbandes viel kleiner als 1 eingestellt, beispielsweise auf 0,05 statt 50.

50 **[0049]** Dann trennt dieses Verfahren Sprache von Nichtsprache mit Hilfe der besprochenen Methode des Äquivalenzklassen genau umgekehrt. Man muß dann die Speicherzuweisung umkehren, da sich dieses Verfahren umgekehrt verhält.

Dieser Effekt trennt einen Teil der Nichtsprache von Sprache. Weil das Verfahren sehr schnell ist, bietet es sich an, damit zu beginnen. Da jedes Verfahren, das Nichtsprache entfernt, die Arbeit nachfolgender Verfahren beschleunigt,
55 muß man mit dem schnellsten Verfahren beginnen.

4. Trennen von Sprache und Nichtsprache mit Hilfe der Symmetriegrenze

[0050] Wenn das Frequenzverhältnis zu groß wird, beispielsweise größer als 8, dann nähert man sich der Symmetriegrenze für Wohlklang oder Konsonanz. Wenn man Frequenzen entfernt, deren Intervalle größer als ein bestimmter Faktor werden, dann handelt es sich um Nichtsprache.

Qualität der Trennung

[0051] Je mehr richtig berechnete Intervalle in einer Gruppe sind, um so besser wird der Trennprozeß.

[0052] Wenn in einer Gruppe Intervalle doppelt oder mehrfach sind, dann verschlechtert sich der Trennprozeß.

[0053] Je mehr verschiedene zahlentheoretische Funktionen so eingesetzt werden, daß jede sich in einer eigenen Gruppe befindet, um so besser wird der Trennprozeß.

[0054] Wenn verschiedene zahlentheoretische Funktionen in einer Gruppe zusammengefaßt werden, dann verschlechtert sich der Trennprozeß.

[0055] Wenn die Intervalle beispielsweise durch nicht geeignete zahlentheoretische Funktionen berechnet werden, dann verschlechtert sich der Trennprozeß.

Existenz der Intervalle

[0056] Aus der Musiktheorie kennt man seit Jahrhunderten die Existenz von Intervallen. Schon Kepler suchte nach allgemeinen Harmoniegesetzen. Er suchte nach Weltharmoniegesetzen. Neben vielen ist aber besonders Leonhard Euler zu nennen, der den Zusammenhang zwischen der Zerlegung der natürlichen Zahlen in Produkte von Primzahlpotenzen, den Fermatschen Sätzen und der Zahlentheorie mit zahlentheoretischen Funktionen erkannte.

[0057] Der Mathematiker Leonhard Euler stellte eine komplizierte zahlentheoretische Funktion (Eulersche Gradusfunktion) zur Berechnung solcher Intervalle auf und berechnete sie für die Zahlen 1 bis 10. Das sind dann 92 Intervalle für die Musiklänge.

[0058] Leibnitz und Herder sagten, daß unsere Seele unbewußt zähle.

[0059] Für das erfindungsgemäße Verfahren zur Spracherkennung kann die Eulersche Gradusfunktion besonders vorteilhaft als zahlentheoretische Funktion angewandt werden.

[0060] Es gibt weiter zahlentheoretische Funktionen. Dahinter stehen komplizierte Strukturen und Sachverhalte der Zahlentheorie und der Algebra, wie beispielsweise die Gammafunktion.

[0061] Die Intervalle werden mit zahlentheoretischen Funktionen berechnet. Es entsteht aus n zahlentheoretischen Funktionen ein n-dimensionaler metrischer Raum, der aus den Feldern der Äquivalenzklassen der Symmetriebeziehungen besteht. Es werden n zahlentheoretische Funktionen, beispielsweise 10, zur Berechnung der Intervalle eingesetzt. Diese n zahlentheoretischen Funktionen erzeugen Intervalle die sich zu einem n-dimensionalen metrischen Raum der Äquivalenzklassen aufspannen lassen. Möglicherweise erzeugen diese Funktionen alle Intervalle für alle menschlichen Sprachen.

[0062] Die Beziehungen zwischen den Elementen wird wichtiger als die Elemente selbst.

Die Intervallgruppen der Sprache können z.B. durch folgende zahlentheoretische Funktionen erzeugt werden:

1. Funktion ist die Eulersche Gradusfunktion

[0063] Jede natürliche Zahl n wird aus Primzahlen P_1, P_2, \dots, P_n zusammengesetzt. Dann wird eine Zahl n zu:

$$n = P_1^{a_1} \cdot P_2^{a_2} \cdot \dots \cdot P_m^{a_m}$$

a_1, a_2, a_m müssen ganze Zahlen sein.

Beispiel: $1440 = 2^5 \cdot 3^2 \cdot 5^1$

Dabei ist $P_1 = 2, P_2 = 3, P_3 = 5$

$A_1 = 5; a_2 = 2; a_3 = 1$

Euler definierte den Gradus als

$$G(n) = a_1 P_1 + a_2 P_2 + a_m P_m - (a_1 + a_2 + \dots + a_m - 1)$$

EP 1 033 702 A2

Beispiel mit Zahlen:

$$\begin{aligned} G(1440) &= 5 \cdot 2 + 2 \cdot 3 + 1 \cdot 5 - (5 + 2 + 1 - 1) \\ &= 10 + 6 + 5 - 7 \\ &= 14 \text{ als Konsonanzgrad für 1440} \end{aligned}$$

5

Für die Intervallmaße muß einem beliebigen Intervall $a : b$ eine Zahl n so zugeordnet werden, daß n sinngemäß als Grad von $a : b$ betrachtet werden kann. Entsprechend muß verfahren werden, wenn eine Verhältniskette $a : b : c$ oder $a_1 : a_2 : \dots : a_n$ vorliegt, also ein Mehrklang.

10 **[0064]** n muß somit selbst eine Funktion des Systems (a_1, a_2, \dots, a_n) sein. Dieser Wert einer Konsonanz ist der Quotient aus dem kleinsten gemeinschaftlichen Vielfachen und dem größten gemeinschaftlichen Teiler der einen Klang beschreibenden natürlichen Zahlen.

[0065] Die Gradusfunktion für gekürzte Brüche a/b ergibt sich aus

15

$$G(a/b) = G(a \cdot b)$$

2. Funktion

[0066]

20

$$\begin{aligned} &2^n / (3^x \cdot 5^y \cdot 7^z) \\ &(3^x \cdot 5^y \cdot 7^z) / 2^n \end{aligned}$$

25 3. Funktion

[0067]

30

$$\begin{aligned} &(n + 2)/n \\ &n / (n + 2) \end{aligned}$$

4. Funktion

35 **[0068]**

$$\begin{aligned} &(2n + 2)/n \\ &n / (2n + 2) \end{aligned}$$

40

5. Funktion

[0069]

45

$$\begin{aligned} &3^x / 5^y \\ &5^y / 3^x \end{aligned}$$

6. Funktion

50

[0070]

$$\begin{aligned} &2^n / 3^x \\ &3^x / 2^n \end{aligned}$$

55

7. Funktion

[0071]

5

$$2^n / 5^y$$

$$5^y / 2^n$$

8. Funktion

10

[0072]

$$3^x / 7^z$$

15

$$7^z / 3^x$$

9. Funktion

[0073]

20

$$2^n / 7^z$$

$$7^z / 2^n$$

25

10. Funktion

[0074]

30

$$7^z / 5^y$$

$$5^y / 7^z$$

35

**„Trennung Sprache von Nichtsprache auf
gruppentheoretischer Basis“**

40

45

50

Bericht

Stand vom Dienstag den 29. 02. 2000

55

Inhalt

5

1 ZUSAMMENFASSUNG

10

2 FUNKTIONSPRINZIP ZUR TRENNUNG DER SPRACHE VON GERÄUSCHEN 3

- 2.1 Vorbereitende Rechenschritte 4
- 2.2 Trennung Sprache von Nichtsprache 4

15

3 BESCHREIBUNG DES VERFAHRENS 5

20

3.1 Berechnen der einhüllenden Kurve eines Klanges (Hüllkurve) 5

3.2 Transformation des Sprachsignals mit Integraltransformationen 6

3.3 Entfaltung des Signals mit Integraltransformationen 6

25

3.4 Töne und Klänge 7

- 3.4.1 Ordnen des Sprachsignals durch Kongruenzgruppen 7
- 3.4.2 Äquivalenzrelationen 7
- 3.4.3 Berechnung der Intervalle aus den Frequenzen des Sprachsignals 7

30

3.5 Die Trennung von Sprache und Nichtsprache wird mit einer Folge von mehreren Symmetrieoperationen durchgeführt. 8

- 2.5.1 Äquivalenzklassen mit Bezugsintervall als Entscheidungskriterium 8
- 2.5.2 Verbessertes Trennen von Sprache und Nichtsprache durch eine Verknüpfung der Tonhöhendifferenz mit den Intensitäten 9
- 2.5.3 Trennen von Sprache und Nichtsprache mit Hilfe einer Amplitudenband-beziehung im Frequenzraum 10
- 2.5.4 Trennen von Sprache und Nichtsprache mit Hilfe der Symmetriegrenze 10
- 2.5.5 Über die Qualität der Trennung 11
- 2.5.6 Über die Existenz der Intervalle 11

40

3.6 Zahlentheoretische Funktionen erzeugen die Intervallgruppen der Sprache. 12

- 2.6.1 Basisfunktion ist die Eulersche Gradusfunktion 12
- 2.6.2 2. Funktion 13
- 2.6.3 3. Funktion 13
- 2.6.4 4. Funktion 13
- 2.6.5 5. Funktion 13
- 2.6.6 6. Funktion 13
- 2.6.7 7. Funktion 14
- 2.6.8 8. Funktion 14
- 2.6.9 9. Funktion 14
- 2.6.10 10. Funktion 14

50

4. Ergebnisse

55

1. Zusammenfassung

5 Harmonie in der Musik hat ganz wesentlich mit harmonischen Intervallen zwischen benachbarten Tönen zu tun. Töne klingen offenbar dann harmonisch, wenn die Intervalle genau bestimmten Größen entsprechen, die wir beispielsweise als Oktave, Quinte oder Terz kennen.

10 Auch die Elemente des Sprachsignals stehen in definierten Beziehungen zueinander. Wie in der Musik gibt es in der Sprache harmonische Intervalle zwischen den benachbarten Sprachfrequenzen. Es zeigt sich, daß die Intervallbeziehungen zwischen den Sprachelementen zur Trennung von Sprache und Nichtsprache herangezogen werden können.

15 Es stellt sich heraus, daß die Intervallgrößen ganz bestimmte Elemente des Primzahlkörpers sind, und daß sie das Ergebnis komplizierter zahlentheoretischer Funktionen sind.

20 Wenn die Intervalle benachbarter Töne nicht exakt diesen zahlentheoretischen Funktionswerten entsprechen, dann ist einer der Tonelemente Nichtsprache. Es zeigte sich, daß für die Trennung Sprache von Nichtsprache weitere mathematische Operationen eingesetzt werden müssen.

25 Die erste ist eine kombinatorische Aufgabe. Durch geschickte kombinatorische Vergleichsoperationen müssen systematisch alle Töne eines Klanges, also alle Frequenzen eines Wellenpaketes, solange mit ihren benachbarten Frequenzen durcheinander geteilt werden, bis alle Elemente (Frequenzen eines Wellenpaketes) miteinander verglichen wurden. Jeder dieser Vergleiche wiederum muß mit den berechneten Intervallen verglichen werden, die ja Äquivalenzklassen der Töne sind.

30 Die aus den Wellenpaketen der Sprachprobe durch Kombinatorik und Vergleich gewonnenen Intervalle gehen dann in mathematische Formeln ein, die unter anderem die Besonderheiten in der Physik des Hörens berücksichtigen.

35 Der Trennungsprozeß erfolgt dann durch das Gegenüberstellen einer speziellen Formelstruktur, die genau ein aus der Sprachprobe gewonnenes Intervall enthält einerseits, - mit einem sogenannten Zielintervall andererseits, - welches aus den zahlentheoretischen Funktionen des Primzahlkörpers berechnet wurde. Damit können gruppentheoretische Strukturen in Ansatz gebracht werden, hinter denen die Symmetriegesetze stehen. Die Trennung kann offenbar dann erfolgen, wenn die Töne einer Sprachprobe genau den berechneten Kongruenzgruppen der Töne in Form von
40 Intervallen entsprechen.
45

50

55

2 Funktionsprinzip zur Trennung der Sprache von Geräuschen

5

2.1 Vorbereitende Rechenschritte

1. Schritt:

10

Das von Störungen zu befreiende digitalisierte Sprachsignal wird in seine Einzelklänge mit Hilfe einer Hüllkurvenmethode zerlegt.

2. Schritt:

15

Integraltransformationen transformieren das Sprachsignal von der Darstellung Intensität über die Zeit in eine Darstellung Amplitude über die Frequenz und Frequenz über die Zeit.

3. Schritt:

20

Im Frequenzraum wird ein Band um die Intensitäten gelegt. Hierdurch wird ein Vergleich aller übrigen Intensitäten nur innerhalb dieses Bandes zugelassen. Die Größe des Bandes hängt unter anderem von der Intensität beziehungsweise von der Lautstärke ab.

4. Schritt:

25

Eine Anzahl zahlentheoretischer Funktionen berechnet Intervalle der Sprache und erzeugt den Raum der Äquivalenzklassen der Sprachtöne.

30

2.2 Trennung Sprache von Nichtsprache

5. Schritt:

35

Alle Frequenzen innerhalb eines Lautes (Hüllkurve) werden durch geeignete Schleifenverfahren mit einander paarweise verglichen. Es entstehen dann für jedes Frequenzpaar aus der Teilung Intervalle. Diese Intervalle werden mit den berechneten Intervallen der Äquivalenzklassen verglichen. Während dieser Überprüfung stellt sich heraus, ob die betrachteten Frequenzen miteinander harmonische Beziehungen haben.

40

Diese harmonischen Beziehungen haben Sprachfrequenzen mit allen anderen Sprachfrequenzen innerhalb einer Hüllkurve oder eines Lautes. Nur dann, wenn im Raum der Äquivalenzklassen ein Bezugsintervall gefunden wird, wird die entsprechende Frequenz in die Datei Sprache überwiesen. Wenn es kein Bezugsintervall gibt wird die Frequenz in die Datei Nichtsprache überwiesen.

45

6. Schritt

Trennen von Sprache und Nichtsprache mit Hilfe eines Amplitudenbandes im Frequenzraum

50

Wenn das Band um die Amplituden sehr viel kleiner als 1 gemacht wird, dann wechselwirken die Intervalle der Äquivalenzklassen die ihrerseits kleiner als 1 sind mit

55

den zu vergleichenden Frequenzen des zu untersuchenden Sprachsignals. Dieser Effekt trennt einen Teil der Nichtsprache von Sprache.

5

7. Schritt:

Es läßt sich aus dem Verhältnis der Tonhöhendifferenz durch eine logarithmische Beziehung zwischen den Amplituden eine neue Beziehung bilden und für das Trennverfahren nutzen. Für das Ergebnis dieser Beziehung wird die Prozedur wiederholt, in dem Raum der Äquivalenzklassen ein Bezugsintervall zu suchen.

10

8. Schritt:

Trennen von Sprache und Nichtsprache mit Hilfe der Symmetriegrenze.

15

Wenn Frequenzintervalle eine bestimmte Größe erreichen, dann nähert man sich der Symmetriegrenze für Wohlklang oder Konsonanz. Dieses kann für die Trennung von Sprache und Nichtsprache genutzt werden.

20

Die einzelnen Trennverfahren können durch geeignete gruppentheoretische Verfahren miteinander verbunden und optimiert werden. Sie können auch einzeln eingesetzt werden.

25

3. Beschreibung des Verfahrens

30

Die stimmhaften Sprachlaute unseres Sprechorgans sind die Harmonische Klänge. Die harmonischen Verwandtschaftsbeziehungen zwischen den Teiltönen stellen die unveränderlichen (invarianten) Eigenschaften der Klänge dar. Bei dem Lernprozeß, welcher von frühestem Alter an die Sprachlauterkennung ermöglicht, erlernt der Mensch mit seinem Gehör die invarianten Klangeigenschaften. Denn dadurch wird es in die Lage versetzt, informationstragende von nichtinformationstragenden Klangmerkmalen zu unterscheiden.

35

3.1 Berechnen der einhüllenden Kurve eines Klanges (Hüllkurve)

40

Um die Sprachsignale in Echtzeit verarbeiten zu können, ist es erforderlich, daß das Signal Klang für Klang einzeln der Reihe nach abgearbeitet werden kann. Um ein willkürliches Abschneiden des Signales für die Verarbeitung zu verhindern, wird ein individuelles flexibles Verfahren gebraucht. Dazu wird die Hüllkurve berechnet, denn sie liefert Anfangs-, Endpunkt und die Klangdauer.

45

Aus dem digitalisierten Sprachsignal werden mit Hilfe der Differentialrechnung zunächst die Stützpunkte der Hüllkurven gewonnen, die alle einzelne Klänge einschließen, gleichgültig wie stark gestört das Signal ist. Alle Einschnürungen im Signal werden als Anfangs- und Endpunkte der Hüllkurve definiert. Damit kann dann das Signal Klang für Klang verarbeitet werden.

50

55

Das Differenzieren der Hüllkurve führt zu den Minima der Hüllkurve. Die Minima der Hüllkurve bilden Anfangs- und Endpunkte der Hüllkurven. Die Hüllkurven der Klänge enthalten unterschiedlich viele Frequenzen und sind deshalb zeitlich verschieden lang.

3.2 Transformation des Sprachsignals mit Integraltransformationen

Fouriertransformation:

Darstellung des Sprachsignales im Frequenzraum: Amplitude als Ordinate und Frequenz als Abszisse

Hilberttransformation:

Darstellung des Sprachsignales im Frequenzraum: Frequenz als Ordinate und Zeit als Abszisse

Jedes Wellenpaket (jeder Klang) enthält n Summanden. Jeder n -te Summand ist die n -te Partialschwingung von dem Wellenpaket.

Das Ursprungssignal wird mit Integraltransformationen so transformiert, daß es als Amplitude über die Frequenz, und als Frequenz über die Zeit für den Datensatz jeder einzelnen Hüllkurve zur weiteren Verarbeitung zur Verfügung steht.

Im Fourierraum beispielsweise sind die Amplituden beziehungsweise die Intensitäten jeweils für eine Hüllkurve über die Frequenz aufgetragen. Hier wird mit einem Amplitudenband innerhalb des Frequenzraumes um die Intensitäten ein Vergleich aller übrigen Intensitäten nur innerhalb dieses Bandes zugelassen. Die Größe des Amplitudenbandes hängt von der Intensität beziehungsweise der Lautstärke ab.

Alle zu den Intensitäten gehörenden Frequenzen sind im Funktionenraum der Hilberttransformation als Frequenzen über die Zeit vorhanden.

3.3 Entfaltung des Signals mit Integraltransformationen

In dem digitalisierten Sprachsignal als Ausgangssignal ist die Information als Intensität über die Zeit in Form von Frequenzen in einem zweidimensionalen Feld kodiert.

Die Integraltransformationen werden als lineare Operatoren der Gruppentheorie definiert und zur Entfaltung des Signals eingesetzt. Beispielsweise werden Fourier- und Hilberttransformation eingesetzt. Sie transformieren das Signal in einen für die nachfolgende Verarbeitung erforderlichen Zustand in Funktionenräume. (Beispielsweise Frequenzraum)

3.4 Töne und Klänge

5 Ein Tonsystem ist ein geordnetes Paar von Elementen. Dabei ist das
Frequenzintervall eine Gruppe und die Intervalle sind eine injektive Abbildung des
Tonsystems in die Menge aller positiven reellen Zahlen. Eine endliche Teilmenge aus
den Frequenzen heißt Klang. Ein geordnetes n-Tupel von Tönen wird geordneter n-
Klang genannt

3.4.1 Ordnen des Sprachsignales durch Kongruenzgruppen

15 Zahlentheoretische Funktionen aus Primzahlen und Zweierpotenzen bilden Felder für
Äquivalenzklassen

Die Äquivalenzklassen werden zu dem Sprachsignal in Form von Äquivalenzrelationen
in Beziehung gebracht.

18 Algebraische Kongruenzgruppen ordnen das Sprachsignal in Sprache und
Nichtsprache in verschiedene Dateien.

20 Rückgewinnung des Sprachsignales im Zeitbereich durch Integralrück-
transformationen

3.4.2 Äquivalenzrelationen

25 Elementvertauschungen werden durch eine zwei- oder mehrstellige Relation auf der
Menge aller Klänge eines Tonsystems präzisiert.

30 Dabei gilt für ein Element aus der Menge des Sprachsignals die Zuordnung zu
Sprache oder Nichtsprache genau dann, wenn zu den Intervallen des zu prüfenden
Signals ein Bezugston aus der Äquivalenzklasse der zahlentheoretischen Funktionen
des Primzahlkörpers existiert.

3.4.3 Berechnung der Intervalle aus den Frequenzen des Sprachsignals

35 Alle Frequenzen des Frequenzraumes des transformierten Sprachsignales müssen
überprüft werden, ob sie Symmetriebeziehungen mit anderen Frequenzen der
40 gleichen Hüllkurve haben. In den Symmetriebeziehungen der Elemente liegen
invariante Eigenschaften der Sprache verschlüsselt vor.

45 Von dem zu verarbeitenden Signal werden deshalb innerhalb der Hüllkurve alle
Frequenzen paarweise der Reihe nach mit allen anderen Frequenzen dividiert, um
Intervalle zu bilden. Dies geschieht solange, bis jede Frequenz in der fortlaufenden
Reihenfolge nach und nach durch alle anderen Frequenzen in der gleichen Hüllkurve
geteilt wurde, um vorübergehend zu einer Verhältniszahl (Intervall) dieses
Frequenzpaares zu kommen.

50 Die erste Frequenz wird durch die zweite, der dritten und so fort geteilt, bis zum Ende
der Hüllkurve. Dann wird die zweite durch die dritte, der vierten und so fort geteilt. In

einer Hüllkurve können beispielsweise jeweils tausende Frequenzen sein. Dabei darf nie die Reihenfolge der Frequenzen vertauscht werden.

5

3.5 Die Trennung von Sprache und Nichtsprache wird mit einer Folge von mehreren Symmetrieoperationen durchgeführt.

10

3.5.1 Äquivalenzklassen mit Bezugsintervall als Entscheidungskriterium

Nach jeder Intervallbildung innerhalb des Frequenzraumes des zu überprüfenden Sprachsignals muß ein Bezugsintervall aus den Äquivalenzklassen der berechneten Intervalle gefunden werden.

15

Die Äquivalenzklassen wiederum werden durch verschiedene zahlentheoretische Funktionen berechnet. Die Trennwirkung wird deutlich besser, wenn man die Erfüllung eines Bezugsintervalles für jede zahlentheoretische Funktion fordert. Wenn dies n Funktionen sind, (wobei n = natürliche ganze Zahlen sind) dann spannen diese zahlentheoretischen Funktionen einen n -dimensionalen metrischen Raum auf.

20

Wenn das aus dem zu prüfenden Signal gewonnene Intervall in allen zahlentheoretischen Funktionen genau eine Bezugsordinate findet, dann sind alle zahlentheoretischen Funktionen mindestens einmal erfüllt. Die Koordinaten führen zu genau einem Bezugsintervall aus dem Raum der Äquivalenzklassen.

25

Über diesen Bezugston können genau dann weitere Sprachparameter wie beispielsweise Anklingzeit und Abklingzeit eines Lautes abgefragt werden, wenn diesem n -dimensionalen Raum eine Metrik aufgeprägt wird. Dies kann mit Hilfe der Gruppentheorie auf der Basis von Matrizen-Operationen durchgeführt werden.

30

Dieses Bezugsintervall kann als Bezugston aufgefaßt werden. Dann können viele weitere Parameter für die Spracherkennung herangezogen werden, wie beispielsweise die Anklingzeit und die Abklingzeit eines Tones.

35

Gibt es ein Bezugsintervall, dann muß sich das Rechenverfahren diese Elemente so lange merken, bis alle Elemente miteinander verglichen wurden. Erst nach Abschluß dieses Sortierungsprozesses weiß das Verfahren, welche Elemente mit anderen Intervallbeziehungen haben.

40

Alle die Frequenzen, die über die Intervalle Beziehungen miteinander haben, stehen schließlich mit allen anderen Frequenzen innerhalb der gleichen Hüllkurve und damit des gleichen Lautes der Sprache in Beziehung. Das Verfahren merkt sich all diese Elemente. Am Ende des Rechenprozesses werden alle solche Elemente in den Speicher für Sprache abgelegt. Alle anderen Frequenzen, die sich nicht in dieses Beziehungssystem einbinden lassen, sind beziehungslose Elemente und kommen in den Speicher für Nichtsprache. Ein Teil der Nichtsprache wird hierdurch bereits von der Sprache getrennt.

45

50

55

5 Elemente, die miteinander durch symmetrische Beziehungen dieser Intervalle verbunden sind, sind Elemente der Sprache. Elemente die in keinem symmetrischen Zusammenhang mit allen anderen Elementen innerhalb der Hüllkurve stehen sind keine Sprache. Sie übertragen keine Information.

10 3.5.2 Verbessertes Trennen von Sprache und Nichtsprache durch eine Verknüpfung der Tonhöhendifferenz mit den Intensitäten

15 Es wird eine zweistellige Relation der Intervalle so definiert, daß ein Bezugston innerhalb von mehreren Koordinaten des Äquivalenzklassenraumes dem Element aus dem zu untersuchenden Signal genau entspricht.

20 In dem folgenden Formel-Paar werden vor allem das logarithmische Verhältnis der Amplituden zur Tonhöhendifferenz in Beziehung gebracht. Eine Konstante berücksichtigt, daß hohe Frequenzen lauter als niedrige Frequenzen gehört werden. Die Rückkoppelungsgröße erfaßt die aktuellen Frequenz und Amplitudenverhältnisse für Anpassungen.

Die beiden Formeln sind zueinander symmetrisch.

$$25 \quad V = \exp \left(\frac{\text{abs} \left(\text{Faktor} \cdot \left(\ln \left(\frac{f_1}{f_2} \right) \right) \right)}{2^{\left(\text{Faktor} \cdot \ln \left(\frac{A_1}{A_2} \right) - \text{Rückkoppelungsgröße} \cdot \text{Konstante} \right) / 10}} \right)$$

besonders gegen tiefe Störfrequenzen

$$30 \quad V = \exp \left(\frac{2^{\left(\text{Faktor} \cdot \ln \left(\frac{A_1}{A_2} \right) - \text{Rückkoppelungsgröße} \cdot \text{Konstante} \right) / 10}}{\text{abs} \left(\text{Faktor} \cdot \left(\ln \left(\frac{f_1}{f_2} \right) \right) \right)} \right)$$

besonders gegen hohe Störfrequenzen

35 es bedeuten:

40 V = Intervall aus den Äquivalenzklassen aller zahlentheoretischen Funktionen
 f_1 / f_2 = Verhältnis der Frequenzen = Tonhöhenabstand,

A_1 / A_2 = Verhältnis der Intensitäten

Rückkoppelungsgröße =

$$45 \quad \text{fabs} (\text{Korrektur1} \cdot \text{Korrektur2}) \cdot (1./\log_{10}(2.)) \cdot \log_{10}((A) \cdot f) + \text{fFaktor} \cdot \log((A) \cdot f)$$

fFaktor = Korrekturgröße, zur Berücksichtigung höherer Frequenzen

50 Korrektur1 = absolut –Wert vom Verhältnis der Amplituden der zu prüfenden Signalfrequenz

Korrektur2 = absolut – Wert vom Verhältnis der Frequenzen der zu prüfenden Signalfrequenz

5

Fallunterscheidungen:

wenn Korrektur1 < 1 dann ist Korrektur1 = 1./Korrektur1;

wenn Korrektur2 < 1 dann ist Korrektur2 = 1./Korrektur2;

10

Mit Hilfe dieser Formeln wird der Signalton mit einem Modell der Sprache geprüft und mit Bezugintervallen verglichen.

15

Dann wiederholt sich das vorhergehende Verfahren.

20

Diese Gruppenoperation verbessert die Qualität der Trennung Sprache von Nichtsprache.

3.5.3 Trennen von Sprache und Nichtsprache mit Hilfe einer Amplitudenbandbeziehung im Frequenzraum

25

Wenn das Band um die Amplituden sehr viel kleiner als 1 gemacht wird, dann wechselwirken die Intervalle der Äquivalenzklassen die ihrerseits kleiner als 1 sind mit den zu vergleichenden Frequenzen des zu untersuchenden Sprachsignals.

30

Beispielsweise sei das Maximum des ersten Wellenpaketes bei 500 Einheiten. Um alle zu vergleichenden Amplituden sei das Amplitudenband mit 50 Einheiten Bandbreite durch den ganzen Frequenzraum als Eingangsband für das Verarbeitungsprogramm eingestellt.

35

Es entsteht ein neuer Effekt, wenn diesem Verarbeitungsprogramm ein Vorprogramm vorgeschaltet wird, bei dem nur die Bandbreite des Amplitudenbandes viel kleiner als 1 eingestellt, beispielsweise auf 0.05 statt 50.

40

Dann trennt dieses Verfahren Sprache von Nichtsprache mit Hilfe der besprochenen Methode des Äquivalenzklassen genau umgekehrt. Man muß dann die Speicherzuweisung umkehren, da sich dieses Verfahren umgekehrt verhält.

45

Dieser Effekt trennt einen Teil der Nichtsprache von Sprache. Weil das Verfahren sehr schnell ist, bietet es sich an, damit zu beginnen. Da jedes Verfahren, das Nichtsprache entfernt, die Arbeit nachfolgender Verfahren beschleunigt, muß man mit dem schnellsten Verfahren beginnen.

3.5.4 Trennen von Sprache und Nichtsprache mit Hilfe der Symmetriegrenze

50

Wenn das Frequenzverhältnis zu groß wird, beispielsweise größer als 8, dann nähert man sich der Symmetriegrenze für Wohlklang oder Konsonanz. Wenn man Frequenzen entfernt, deren Intervalle größer als ein bestimmter Faktor werden, dann handelt es sich um Nichtsprache.

55

3.5.5 Über die Qualität der Trennung

5 Je mehr richtig berechnete Intervalle in einer Gruppe sind, um so besser wird der Trennprozeß.

10 Wenn in einer Gruppe Intervalle doppelt oder mehrfach sind, dann verschlechtert sich der Trennprozeß.

15 Je mehr verschiedene zahlentheoretische Funktionen so eingesetzt werden, daß jede sich in einer eigenen Gruppe befindet, um so besser wird der Trennprozeß.

20 Wenn verschiedene zahlentheoretische Funktionen in einer Gruppe zusammengefaßt werden, dann verschlechtert sich der Trennprozeß.

25 Wenn die Intervalle beispielsweise durch nicht geeignete zahlentheoretische Funktionen berechnet werden, dann verschlechtert sich der Trennprozeß.

3.5.6 Über die Existenz der Intervalle

30 Aus der Musiktheorie kennt man seit Jahrhunderten die Existenz von Intervallen. Schon Kepler suchte nach allgemeinen Harmoniegesetzen. Er suchte nach Weltharmoniegesetzen. Neben vielen ist aber besonders Leonhard Euler zu nennen, der den Zusammenhang zwischen der Zerlegung der natürlichen Zahlen in Produkte von Primzahlpotenzen, den Fermatschen Sätzen und der Zahlentheorie mit zahlentheoretischen Funktionen erkannte.

35 Der Mathematiker Leonhard Euler stellte eine komplizierte zahlentheoretische Funktion zur Berechnung solcher Intervalle auf und berechnete sie für die Zahlen 1 bis 10. Das sind dann 92 Intervalle für die Musikklänge. Eulers Zeitgenossen standen dieser Gradustabelle faßt alle verständnislos gegenüber.

40 Hilfe bekam er durch Leibnitz und Herder, die sagten, daß unsere Seele unbewußt zähle.

45 Mit dem hier vorgestellten neuen Verfahren zur Spracherkennung ist die Eulersche Gradusfunktion eine der wichtigsten zahlentheoretischen Funktionen. Bis einschließlich Gradus 34 sind dies bereits trotz einer Auswahl weit über zehntausend Intervalle.

50 Es gibt weiter zahlentheoretische Funktionen. Dahinter stehen komplizierte Strukturen und Sachverhalte der Zahlentheorie und der Algebra, wie beispielsweise die Gammafunktion.

55

55

Die Intervalle werden mit zahlentheoretischen Funktionen berechnet. Es entsteht aus
 n zahlentheoretischen Funktionen ein n-dimensionaler metrischer Raum, der aus den
 5 Feldern der Äquivalenzklassen der Symmetriebeziehungen besteht.

Es werden n zahlentheoretische Funktionen, beispielsweise 10, zur Berechnung der
 Intervalle eingesetzt. Diese n zahlentheoretischen Funktionen erzeugen Intervalle die
 10 sich zu einem n-dimensionalen metrischen Raum der Äquivalenzklassen aufspannen
 lassen. Möglicherweise erzeugen diese Funktionen alle Intervalle für alle
 menschlichen Sprachen.

Die Beziehungen zwischen den Elementen wird wichtiger als die Elemente selbst.

3.6 Zahlentheoretische Funktionen erzeugen die Intervallgruppen der Sprache.

*Die in den folgenden Kapiteln dargestellten Funktionen sind noch vorläufig. Es
 gibt Verbesserungsvorschläge für weiterführende Arbeiten.*

3.6.1 Basisfunktion ist die Eulersche Gradusfunktion

Jede natürliche Zahl n wird aus Primzahlen P_1, P_2, \dots, P_m zusammengesetzt. Dann wird
 eine Zahl n zu:

$$n = P_1^{a_1} \cdot P_2^{a_2} \cdot \dots \cdot P_m^{a_m}$$

a_1, a_2, a_m müssen ganze Zahlen sein.

Beispiel: $1440 = 2^5 \cdot 3^2 \cdot 5^1$

Dabei ist $P_1 = 2, P_2 = 3, P_3 = 5$

$A_1 = 5; a_2 = 2; a_3 = 1$

Euler definierte den Gradus als

$$G(n) = a_1 P_1 + a_2 P_2 + a_m P_m - (a_1 + a_2 + \dots + a_m - 1)$$

Beispiel mit Zahlen:

$$G(1440) = 5 \cdot 2 + 2 \cdot 3 + 1 \cdot 5 - (5 + 2 + 1 - 1)$$

$$= 10 + 6 + 5 - 7$$

= 14 als Konsonanzgrad für 1440

Für die Intervallmaße muß einem beliebigen Intervall $a : b$ eine Zahl n so zugeordnet
 werden, daß n sinngemäß als Grad von $a : b$ betrachtet werden kann. Entsprechend
 50 muß verfahren werden, wenn eine Verhältniskette $a : b : c$ oder $a_1 : a_2 : \dots : a_n$ vorliegt,
 also ein Mehrklang.

n muß somit selbst eine Funktion des Systems ($a_1, a_2, \dots a_n$) sein. Dieser Wert einer Konsonanz ist der Quotient aus dem kleinsten gemeinschaftlichen Vielfachen und dem größten gemeinschaftlichen Teiler der einen Klang beschreibenden natürlichen Zahlen.

Die Gradusfunktion für gekürzte Brüche a/b ergibt sich aus

$$G(a/b) = G(a \cdot b)$$

3.6.2 2. Funktion

$$2^n / (3^x \cdot 5^y \cdot 7^z)$$

$$(3^x \cdot 5^y \cdot 7^z) / 2^n$$

3.6.3 3. Funktion

$$(n + 2) / n$$

$$n / (n+2)$$

3.6.4 4. Funktion

$$(2n + 2) / n$$

$$n / (2n + 2)$$

3.6.5 5. Funktion

$$3^x / 5^y$$

$$5^y / 3^x$$

3.6.6 6. Funktion

$$2^n / 3^x$$

$$3^x / 2^n$$

3.6.7 7. Funktion

5 $2^n / 5^y$

$5^y / 2^n$

10 **3.6.8 8. Funktion**

$3^x / 7^z$

15 $7^z / 3^x$

20 **3.6.9 9. Funktion**

$2^n / 7^z$

25 $7^z / 2^n$

3.6.10 10. Funktion

30 $7^z / 5^y$

$5^y / 7^z$

35 **4. Ergebnisse**40

An zwei Beispielen konnte die neue Methode demonstriert werden. Es handelt sich um die Sprachproben Test6 und Test7, die von freundlicherweise von der Firma MEDAV aus Erlangen zur Verfügung gestellt wurden.

45

Die Störungen in beiden Proben konnten deutlich hörbar gemindert werden. In der Anlage befinden sich Abbildungen für beide Sprachproben.

50

In den Abbildungen zu Sprachprobe Test 6 und Test 7 sind auf der Ordinate die Zeit und auf der Abszisse die Frequenz dargestellt. Die Farbe weiß entspricht der geringsten Intensität und die Farbe blau entspricht der maximalen Intensität:
Die Bilder stellen die

- Originaldarstellung der Signalprobe
- Verbesserte Sprachprobe und
- Rauschen und übrige Nichtsprache aus der Sprachprobe

55

dar.

5 In der Sprachprobe **test6** sieht man im Original neben dem reinen Sprachsignal ein Hintergrundrauschen und eine zusätzliche Störung durch Funkrauschen. Nach der Behandlung sind deutliche Verbesserungen sichtbar (siehe hierzu die Abbildung „Verbesserte Sprachprobe von Test6“).

10 Es zeigt sich, daß in der **test7** Sprachprobe zwei Störsignale vorhanden sind, ein in der Frequenz zunehmendes Signal und eine feste Störfrequenz bei 1kHz. Nach der Behandlung mit unserem Rechenprogramm sind die Störungen zum größten Teil verschwunden (siehe hierzu die Abbildung "Verbesserte Sprachprobe von Test7").

15 Am 26. April 1999 konnten bei einer Vorführung in Bonn im Hause der Firma Elekluft GmbH gezeigt werden, daß diese Ergebnisse in der Qualität mit den Ergebnissen vom "Stand der Technik" vergleichbar sind.

20 Patentansprüche

25 1. Verfahren zur Trennung von Sprache und Geräuschen, mit folgenden Verfahrensschritten:

- Das von Störungen zu befreiende digitalisierte Sprachsignal wird in seine Einzelklänge mit Hilfe einer Hüllkurvenmethode zerlegt.
- 30 - Mittels Integraltransformationen wird das Sprachsignal von der Darstellung Intensität über die Zeit in eine Darstellung Amplitude über die Frequenz und Frequenz über die Zeit transformiert.
- Im Frequenzraum wird ein Band um die Intensitäten gelegt. Hierdurch wird ein Vergleich aller übrigen Intensitäten nur innerhalb dieses Bandes zugelassen. Die Größe des Bandes hängt unter anderem von der Intensität beziehungsweise von der Lautstärke ab.
- 35 - Mittels zahlentheoretischer Funktionen werden Intervalle der Sprache berechnet und ein Raum der Äquivalenzklassen der Sprachtöne erzeugt.
Die eigentliche Trennung von Sprache und Geräuschen erfolgt anschließend durch eines oder mehreren der folgenden Verfahrensschritte:
- Die Frequenzen innerhalb eines Lautes (Hüllkurve) werden durch geeignete Schleifenverfahren miteinander paarweise verglichen. Es entstehen dann für jedes Frequenzpaar aus der Teilung Intervalle. Diese Intervalle werden mit den berechneten Intervallen der Äquivalenzklassen verglichen. Als Ergebnis dieser Überprüfung ergibt sich, ob die betrachteten Frequenzen miteinander harmonische Beziehungen haben oder nicht. Diese harmonischen Beziehungen haben Sprachfrequenzen mit allen anderen Sprachfrequenzen innerhalb einer Hüllkurve oder eines Lautes. Nur dann, wenn im Raum der Äquivalenzklassen ein Bezugsintervall gefunden wird, wird die entsprechende Frequenz in die Datei Sprache überwiesen. Wenn es kein Bezugsintervall gibt, wird die Frequenz in die Datei Nichtsprache überwiesen.
- 45 - Trennen von Sprache und Nichtsprache mit Hilfe eines Amplitudenbandes im Frequenzraum:
Wenn das Band um die Amplituden sehr viel kleiner als 1 gemacht wird, dann wechselwirken die Intervalle der Äquivalenzklassen die ihrerseits kleiner als 1 sind mit den zu vergleichenden Frequenzen des zu untersuchenden Sprachsignals. Dieser Effekt trennt einen Teil der Nichtsprache von Sprache.
- 50 - Aus dem Verhältnis der Tonhöhendifferenz wird durch eine logarithmische Beziehung zwischen den Amplituden eine neue Beziehung gebildet und für die Trennung Geräusch/Sprache genutzt. Für das Ergebnis dieser

EP 1 033 702 A2

Beziehung wird die Prozedur wiederholt, in dem Raum der Äquivalenzklassen ein Bezugsintervall zu suchen.

- Trennen von Sprache und Nichtsprache mit Hilfe der Symmetriegrenze:
Wenn das Frequenzverhältnis größer als 8 wird, dann nähert man sich der Symmetriegrenze für Wohlklang oder Konsonanz. Es werden diejenigen Frequenzen entfernt, deren Intervalle größer als ein bestimmter Faktor sind.

5

10

15

20

25

30

35

40

45

50

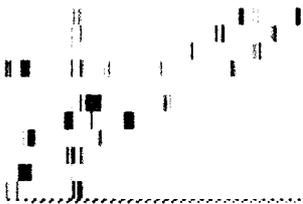
55



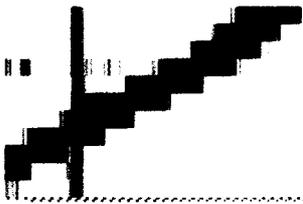
test6 (2).bmp



test6ausgabe (2).bmp



test7ausgabe (2).bmp



test7orig (2).bmp



test7rauschen (2).bmp

