



BUNDESPATENTGERICHT

IM NAMEN DES VOLKES

URTEIL

4 Ni 10/21 (EP)

(Aktenzeichen)

Verkündet am
6. Dezember 2021

...

In der Patentnichtigkeitssache

...

betreffend das europäische Patent EP 2 162 880
(DE 60 2008 036 032)

hat der 4. Senat (Nichtigkeitssenat) des Bundespatentgerichts auf Grund der mündlichen Verhandlung vom 6. Dezember 2021 durch die Vorsitzende Richterin Grote-Bittner sowie die Richter Dipl.-Phys. Univ. Dipl.-Wirt.-Phys. Arnoldi, Dipl.-Ing. Matter, Dr. Söchtig und Dipl.-Phys. Univ. Dr. Haupt

für Recht erkannt:

- I. Das europäische Patent 2 162 880 wird mit Wirkung für das Hoheitsgebiet der Bundesrepublik Deutschland dadurch teilweise für nichtig erklärt, dass seine Ansprüche 1, 2, 5, 19, 20 und 21 die nachfolgende Fassung erhalten:

1. A method for estimating a tonality of a sound signal, the method comprising:

- calculating a current residual spectrum of the sound signal by subtracting a spectral floor from a spectrum of the sound signal in a current frame;

- detecting peaks in the current residual spectrum;

- calculating a correlation map between the current residual spectrum and a previous residual spectrum for each detected peak; and

- calculating a long-term correlation map based on the calculated correlation map, the long-term correlation map being indicative of a tonality in the sound signal.

2. A method as defined in preceding claim 1, wherein calculating the current residual spectrum comprises:

searching for minima in the spectrum of the sound signal in the current frame;

estimating the spectral floor by connecting the minima with each other;

5. A method as defined in any of the preceding claims, wherein calculating the long-term correlation map comprises:

filtering the correlation map through a one-pole filter on a frequency bin by frequency bin basis; and

summing the filtered correlation map over the frequency bins so as to produce a summed long-term correlation map.

19. A device for estimating a tonality of a sound signal, the device comprising:

a calculator for calculating a current residual spectrum of the sound signal by subtracting a spectral floor from a spectrum of the sound signal in a current frame;

a detector for detecting peaks in the current residual spectrum;

a calculator for calculating a correlation map between the current residual spectrum and a previous residual spectrum for each detected peak; and

a calculator for calculating a long-term correlation map based on the calculated correlation map, the long-term correlation map being indicative of a tonality in the sound signal.

20. A device as defined in preceding claim 19, wherein the calculator of the current residual spectrum comprises:

a locator of minima in the spectrum of the sound signal in the current frame;

an estimator of the spectral floor which connects the minima with each other; and

a subtractor of the estimated spectral floor from the spectrum so as to produce the current residual spectrum.

21. A device as defined in preceding claim 19 or 20, wherein the calculator of the long-term correlation map comprises:

a filter for filtering the correlation map on a frequency bin by frequency bin basis; and

an adder for summing the filtered correlation map over the frequency bins so as to produce a summed long-term correlation map.

- II. Im Übrigen wird die Klage abgewiesen.
- III. Von den Kosten des Rechtsstreits haben die Klägerin 90% und die Beklagte 10 % zu tragen.
- IV. Das Urteil ist im Kostenpunkt gegen Sicherheitsleistung in Höhe von 120 % des jeweils zu vollstreckenden Betrages vorläufig vollstreckbar.

Tatbestand

Die Beklagte ist eingetragene Inhaberin des u. a. für die Bundesrepublik Deutschland erteilten europäischen Patents 2 162 880 (im Folgenden: Streitpatent), das aus der PCT-Anmeldung PCT/CA2008/001184 (offengelegt als WO 2009/000073 A1) hervorgegangen ist, am 20. Juni 2008 unter Inanspruchnahme der Priorität der US-Patentanmeldung 929336 vom 22. Juni 2007 angemeldet und dessen Erteilung am 24. Dezember 2014 veröffentlicht worden ist. Im Patentregister des Deutschen Patent- und Markenamts wird das Streitpatent mit der Bezeichnung „Verfahren und Einrichtung zur Schätzung der Tonalität eines Schallsignals“ unter dem Aktenzeichen 60 2008 036 032 geführt.

Das Streitpatent umfasst in seiner erteilten Fassung 27 Ansprüche mit u. a. dem unabhängigen Anspruch 1, den auf diesen direkt bzw. indirekt rückbezogenen Unteransprüchen 2 und 5, dem Vorrichtungsanspruch 19 und den auf diesen rückbezogenen Unteransprüchen 20 und 21.

Die erteilten Ansprüche **1, 2, 5, 19, 20 und 21** lauten in der englischen Verfahrenssprache und ihrer deutschen Übersetzung unter Hinzufügung einer Merkmalsgliederung des Senats wie folgt:

- 1.1 Verfahren zum Schätzen der Tonalität eines Schallsignals, wobei das Verfahren umfasst:
A method for estimating a tonality of a sound signal, the method comprising:
- 1.2 Berechnen eines aktuellen Residualspektrums des Schallsignals;
calculating a current residual spectrum of the sound signal;
- 1.3 Erkennen von Spitzen im aktuellen Residualspektrum;
detecting peaks in the current residual spectrum;
- 1.4 Berechnen einer Korrelationskarte zwischen dem aktuellen Residualspektrum und einem vorherigen Residualspektrum für jede erkannte Spitze; und
calculating a correlation map between the current residual spectrum and a previous residual spectrum for each detected peak; and

- 1.5 Berechnen einer Langzeit-Korrelationskarte basierend auf der berechneten Korrelationskarte, wobei die Langzeit-Korrelationskarte eine Tonalität im Schallsignal anzeigt.

calculating a long-term correlation map based on the calculated correlation map, the long-term correlation map being indicative of a tonality in the sound signal.

- 2.2 Verfahren wie in Anspruch 1 definiert, wobei das Berechnen des aktuellen Residualspektrums umfasst:

A method as defined in claim 1, wherein calculating the current residual spectrum comprises:

- 2.2.1 Suchen nach Minima im Spektrum des Schallsignals in einem aktuellen Rahmen;

searching for minima in the spectrum of the sound signal in a current frame;

- 2.2.2 Schätzen eines spektralen Untergrunds durch Verbinden der Minima miteinander; und

*estimating a spectral floor by connecting the minima with each other;
and*

- 2.2.3 Subtrahieren des geschätzten spektralen Untergrunds vom Spektrum des Schallsignals im aktuellen Rahmen, um so das aktuelle Residualspektrum zu erzeugen.

subtracting the estimated spectral floor from the spectrum of the sound signal in the current frame so as to produce the current residual spectrum.

- 5.5 Verfahren wie in einem der vorstehenden Ansprüche definiert, wobei das Berechnen der Langzeit-Korrelationskarte umfasst:

A method as defined in any of the preceding claim [sic!], wherein calculating the long-term correlation map comprises:

- 5.5.1 Filtern der Korrelationskarte durch ein einpoliges Filter für jedes einzelne Frequenzbin; und

filtering the correlation map through an one-pole filter on a frequency bin by frequency bin basis; and

- 5.5.2 Summieren der gefilterten Korrelationskarte über die Frequenzbins, um eine summierte Langzeit-Korrelationskarte zu erzeugen.

summing the filtered correlation map over the frequency bins so as to produce a summed long-term correlation map.

- 19.1 Vorrichtung zum Schätzen einer Tonalität eines Schallsignals, wobei die Vorrichtung umfasst:

A device for estimating a tonality of a sound signal, the device comprising:

- 19.2 einen Berechner zum Berechnen eines aktuellen Residualspektrums des Schallsignals;

a calculator for calculating a current residual spectrum of the sound signal;

- 19.3 einen Detektor zum Erkennen von Spitzen im aktuellen Residualspektrum;

a detector for detecting peaks in the current residual spectrum;

- 19.4 einen Berechner zum Berechnen einer Korrelationskarte zwischen dem aktuellen Residualspektrum und einem vorherigen Residualspektrum für jede erkannte Spitze; und

a calculator for calculating a correlation map between the current residual spectrum and a previous residual spectrum for each detected peak; and

- 19.5 einen Berechner zum Berechnen einer Langzeit-Korrelationskarte basierend auf der berechneten Korrelationskarte, wobei die Langzeit-Korrelationskarte eine Tonalität im Schallsignal anzeigt.

a calculator for calculating a long-term correlation map based on the calculated correlation map, the long-term correlation map being indicative of a tonality in the sound signal.

- 20.2 Vorrichtung wie in Anspruch 19 definiert, wobei der Berechner des aktuellen Residualspektrums umfasst:

A device as defined in claim 19, wherein the calculator of the current residual spectrum comprises:

- 20.2.1 einen Lokalisierer von Minima im Spektrum des Schallsignals in einem aktuellen Rahmen;
a locator of minima in the spectrum of the sound signal in a current frame;
 - 20.2.2 einen Schätzer eines spektralen Untergrunds, der die Minima miteinander verbindet; und
an estimator of a spectral floor which connects the minima with each other; and
 - 20.2.3 einen Subtrahierer des geschätzten spektralen Untergrunds vom Spektrum, um ein aktuelles Residualspektrum zu erzeugen.
a subtractor of the estimated spectral floor from the spectrum so as to produce the current residual spectrum.
- 21.5 Vorrichtung wie in einem der Ansprüche 19 oder 20 definiert, wobei der Berechner der Langzeit-Korrelationskarte umfasst:
A device as defined in claim 19 or 20, wherein the calculator of the long-term correlation map comprises:
- 21.5.1 ein Filter zum Filtern der Korrelationskarte für jedes einzelne Frequenzbin; und
a filter for filtering the correlation map on a frequency bin by frequency bin basis; and
 - 21.5.2 einen Addierer zum Summieren der gefilterten Korrelationskarte über die Frequenzbins, um eine summierte Langzeit-Korrelationskarte zu erzeugen.
an adder for summing the filtered correlation map over the frequency bins so as to produce a summed long-term correlation map.

Die Klägerin greift das erteilte Streitpatent im Umfang der Ansprüche 1, 2, 5, 19 bis 21 – und folgend alle von der Beklagten eingereichten geänderten Fassungen nach den Hilfsanträgen – an und macht die Nichtigkeitsgründe der mangelnden ausführbaren Offenbarung und der fehlenden Patentfähigkeit geltend. Die Beklagte verteidigt das Streitpatent im angegriffenen Umfang in der erteilten Fassung sowie mit drei Hilfsanträgen in

geänderten Fassungen, weiter hilfsweise verteidigt sie die Unteransprüche 2, 5, 20 und 21 in der erteilten Fassung isoliert.

Nach **Hilfsantrag 1** weisen die Ansprüche 1 und 19 gegenüber der erteilten Fassung nach dem Merkmal 1.2/19.2 jeweils das zusätzliche Merkmal 2.2.3^{Hi1}/19.2.3^{Hi1} auf:

durch Subtrahieren eines spektralen Untergrunds von einem Spektrum des Schallsignals in einem aktuellen Rahmen;

by subtracting a spectral floor from a spectrum of the sound signal in a current frame;

Wegen der geänderten Anspruchsfassungen nach den Hilfsanträgen 2 und 3 wird auf den Schriftsatz der Beklagten vom 28. Mai 2021 verwiesen.

Die Klägerin, die die wirksame Inanspruchnahme der US-Prioritätsschrift u. a. deswegen verneint, weil in dieser der Aspekt der Tonalität nicht thematisiert werde, vertritt die Auffassung, dass das Streitpatent dem Fachmann keine Lehre liefere, wie sich durch die Merkmale der Ansprüche 1 und 19 eine Tonalität im Sinne einer Beziehung von Tonhöhen schätzen lasse. Der Beschreibung sei eine dahingehende Lehre nicht zu entnehmen, weil sie mit der tonalen Stabilität die Dauer der Tonhöhen betreffe. Die beiden Ausführungsbeispiele mit einem Verfahren zum Schätzen der tonalen Stabilität einerseits und einem Verfahren zum Schätzen der Tonalität andererseits stünden im Gegensatz zueinander, wobei das erste Ausführungsbeispiel nicht von den Ansprüchen 1 und 19 erfasst sei. Bei den Begriffen Tonalität und tonale Stabilität handle es sich nämlich um zwei Fachbegriffe mit unterschiedlicher Bedeutung. Die Merkmale 1.2 bis 1.5 bzw. 19.1 bis 19.5 der angegriffenen Ansprüche 1 und 19 mögen allenfalls dienlich sein, um die Dauer von Tönen festzustellen, was dem Fachmann jedoch nicht weiterhelfe, wenn die Tonalität eines Schallsignals zu schätzen sei. Darüber hinaus seien die Gegenstände der erteilten unabhängigen Ansprüche 1 und 19 ausgehend von der Schrift

auch nicht patentfähig, nämlich weder neu noch beruhten sie auf erfinderischer Tätigkeit. Als weiteren Stand der Technik nennt die Klägerin u. a. die Schrift:

K4 MINAMI, K. et. al.: Video Handling with Music and Speech Detection. IEEE MultiMedia, vol. 5, no. 3, S. 17 – 25, 1998

Der in der K3 gezeigte Musikdetektor unterscheide ebenso wie das Streitpatent Sprache und Musik auf der Grundlage eines Spektrogramms, welches die zeitliche Entwicklung der spektralen Information des Audiosignals trage, und basiere damit ebenfalls auf der Erkenntnis, dass bestimmte Musikstücke Töne enthielten, die über einen vergleichsweise langen Zeitraum eine konstante Frequenz aufwiesen.

Auch die abhängigen Ansprüche 2, 5, 20 und 21 könnten keine Patentfähigkeit begründen. So seien die Ansprüche 2 und 20 durch die Schrift K4 bzw. durch das Wissen des Fachmanns nahegelegt, da sie nur eine der trivialen Spielarten von Basislinienkorrektur beschrieben, die dem Fachmann unter der Bezeichnung Gummiband-Methode bekannt sei. Der Gegenstand der Ansprüche 5 und 21 sei nicht ausführbar offenbart, jedenfalls aus K3 bekannt und dem Fachmann durch K4 bzw. durch sein Fachwissen nahegelegt.

Der Hilfsantrag 1 sei unzulässig, weil das zusätzlich aufgenommene Merkmal in den Ansprüchen 1 und 19 nicht ursprünglich offenbart sei. Es vermöge aber auch nicht die Patentfähigkeit der Gegenstände nach Anspruch 1 und 19 zu begründen, weil es im üblichen Handeln des Fachmanns liege, die Grobstruktur des Schallsignalspektrums möglichst zu eliminieren, um die Feinstruktur freizustellen und es insoweit mithin an einer erfinderischen Tätigkeit fehle. Die Hilfsanträge 2 und 3 seien ebenfalls unzulässig und unbegründet.

Der Senat hat den Parteien einen qualifizierten Hinweis vom 17. Februar 2021 und am 3. Dezember 2021 einen weiteren rechtlichen Hinweis erteilt.

Die Klägerin beantragt,

das europäische Patent 2 162 880 im Umfang seiner Ansprüche 1, 2, 5, 19 bis 21 mit Wirkung für das Hoheitsgebiet der Bundesrepublik Deutschland für nichtig zu erklären.

Die Beklagte beantragt,

die Klage abzuweisen,

hilfsweise die Klage mit der Maßgabe abzuweisen, dass das Streitpatent eine der Fassungen der angegriffenen Ansprüche 1, 2, 5, 19 bis 21 gemäß den Hilfsanträgen 1 bis 3, eingereicht mit Schriftsatz vom 28. Mai 2021, erhält, weiter hilfsweise die Klage abzuweisen, soweit sie sich gegen die angegriffenen Ansprüche 2, 5, 20 und 21 in der erteilten Fassung richtet.

Sie tritt der Auffassung der Klägerin in allen Punkten entgegen und meint, dass die Gegenstände der Ansprüche 1 und 19 sowohl ausführbar offenbart als auch patentfähig seien. Wie dem Gesamtinhalt der Streitpatentschrift zu entnehmen sei, handele es sich bei den darin verwendeten Begriffen tonale Stabilität und Tonalität um Synonyme, so dass bei zutreffendem Verständnis der Ansprüche 1 und 19 deren Ausführbarkeit nicht in Frage stehe. Entgegen der klägerischen Ansicht sei die Druckschrift K3 auch nicht neuheitsschädlich, da es in dieser – anders als im Streitpatent – nicht um die Unterscheidung stimmloser Sprachsignale von Musiksignalen, sondern um die Unterscheidung stimmhafter Sprachsignale von Musiksignalen gehe. Auch stelle weder der Peak-Rahmen der K3 das erfindungsgemäße Residualspektrum nach Merkmal 1.2 dar, noch werde in der K3 eine Korrelationskarte und eine Langzeitkorrelationskarte gemäß den Merkmalen 1.4 und 1.5 berechnet.

Ihr Verständnis des in der K3 gezeigten Musikdetektors hat die Beklagte in der Anlage NB4 zum Ausdruck gebracht:

NB4 Funktionsweise der Funktion „music“ in HAWLEY, eine Seite

Des Weiteren seien die Merkmale der Ansprüche 2 und 20, die Details der Berechnung des Residualspektrums angeben, entgegen der Auffassung der Klägerin aus der K3 und K4 nicht bekannt und durch diese Druckschriften auch nicht nahegelegt. Der Fachmann habe schließlich auch keine Veranlassung, die Druckschriften K3 und K4 miteinander zu kombinieren. Die Gegenstände der

Ansprüche 5 und 21 seien ausführbar offenbart und sowohl neu als auch erfinderisch gegenüber K3 und K4.

Die Hilfsanträge 1 bis 3 seien zulässig, da die Änderungen beschränkend wirkten und in der ursprünglichen Anmeldung als zur Erfindung gehörend offenbart seien. Die Gegenstände der Ansprüche nach den Hilfsanträgen seien auch patentfähig, da sie neu seien und auch auf erfinderischer Tätigkeit beruhten.

Wegen der weiteren Einzelheiten des Sach- und Streitstandes wird auf die Schriftsätze der Parteien nebst Anlagen und den weiteren Inhalt der Akte Bezug genommen.

Entscheidungsgründe

Die Nichtigkeitsklage, mit der die Nichtigkeitsgründe der mangelnden ausführbaren Offenbarung und der fehlenden Patentfähigkeit geltend gemacht werden (Art II § 6 Abs. 1 S. 1 Nr. 1, 2 IntPatÜG i. V. m. Art. 138 Abs. 1 lit. a), b), Art. 54, Art 56 EPÜ), ist zulässig.

Sie ist insoweit begründet, als das Streitpatent im angegriffenen Umfang für nichtig zu erklären ist, soweit es über die von der Beklagten beschränkt verteidigten Fassung nach Hilfsantrag 1 hinausgeht. Das Streitpatent erweist sich nämlich in der erteilten Fassung als nicht patentfähig. Dagegen ist das Streitpatent in der Fassung nach dem zulässigen Hilfsantrag 1 sowohl ausführbar offenbart als auch patentfähig, nämlich neu und zudem auf erfinderischer Tätigkeit beruhend. Die Klage ist daher insoweit unbegründet. Auf die weiteren Hilfsanträge kam es daher nicht mehr an.

I.

1. Das Streitpatent beschäftigt sich mit dem effizienten Codieren von Schallsignalen. Die Codierungstechnik CELP (*Code-Excited Linear Prediction*) sei eine der besten Techniken, um einen guten Kompromiss zwischen subjektiver Qualität

und Bitrate zu erzielen. Die quellengesteuerte bitratenvariable Sprachcodierung (VBR, *Variable Bit Rate*) erhöhe die Systemkapazität signifikant, da in Abhängigkeit von der Art des Eingangssignals (stimmhaft, stimmlos, Übergang, Hintergrundrauschen) ein Signalklassifizierungsmodul für jeden Sprachrahmen das zugehörige optimierte Codierermodell bzw. Bitrate wähle. Die bei VBR verwendeten Techniken Spracherkennungsdetektion (VAD, *Voice Activity Detection*), diskontinuierliche Übertragung (DTX, *Discontinuous Transmission*) und Komfortrauschenzeugung (CNG, *Comfort Noise Generation*) reduzierten die durchschnittliche Bitrate deutlich (Streitpatentschrift, Abs. 3 – 6).

Der mit Sprachsignalen gut funktionierende VAD-Algorithmus könne gravierende Probleme bereiten, wenn Musikabschnitte des Schallsignals versehentlich als stimmlose Signale oder als stabiles Hintergrundrauschen klassifiziert würden. Daher wäre es vorteilhaft, den VAD-Algorithmus so zu erweitern, dass er Musiksignale besser von anderen Signalen unterscheiden könne. Das Streitpatent bezeichnet diese Erweiterung als Schallsignalerkennungsalgorithmus (SAD, *sound activity detection*), wobei Schall sowohl Sprache als auch Musik oder jedes andere brauchbare Signal umfasse. Das Detektieren bzw. Schätzen der Tonalität / der tonalen Stabilität erhöhe die Leistungsfähigkeit des SAD-Algorithmus und könne auch in einem Super-Breitband Codierer-Decodierer (*codec*) verwendet werden, um für Signale oberhalb von 7 kHz das richtige Codierermodell auszuwählen (Abs. 2, 6, 7-9, 97, 149).

Die bessere Differenzierbarkeit von Musik- und anderen Signalen soll erreicht werden mit den Gegenständen

- der unabhängigen Ansprüche 1 und 19, d. h. einem **Verfahren und einer Vorrichtung zum Schätzen der Tonalität eines Schallsignals**,
- der von den Ansprüchen 1 bzw. 19 abhängigen und nebengeordneten Ansprüchen 6 und 22, d. h. mit einem Verfahren und einer Vorrichtung zum Erkennen von Schallaktivität in einem Schallsignal, wobei die Tonalitätsschätzung zur Unterscheidung eines Musiksignals von Hintergrundrauschen dient,

- der Ansprüche 11 und 23, d. h. mit einem Verfahren und einer Vorrichtung zum Klassifizieren eines Schallsignals mit dem Ziel, die Codierung des Schallsignals zu optimieren, wobei die Tonalitätsschätzung eine Klassifizierung von Musiksignalen als stimmlose Sprachsignale verhindern soll, und
- der Ansprüche 14 und 25, d. h. mit einem Verfahren und einer Vorrichtung zum Codieren eines höheren Bandes eines Schallsignals anhand einer Klassifizierung des Schallsignals, wobei die Tonalitätsschätzung zum Einstufen des Schallsignals als tonales oder nicht tonales Schallsignal dient.

Die Figur 5 des Streitpatents zeigt die Einbettung des Verfahrens zum Schätzen der Tonalität in den Prozess der Signalklassifizierung bei der Unterscheidung zwischen stimmloser Sprache und Musik:

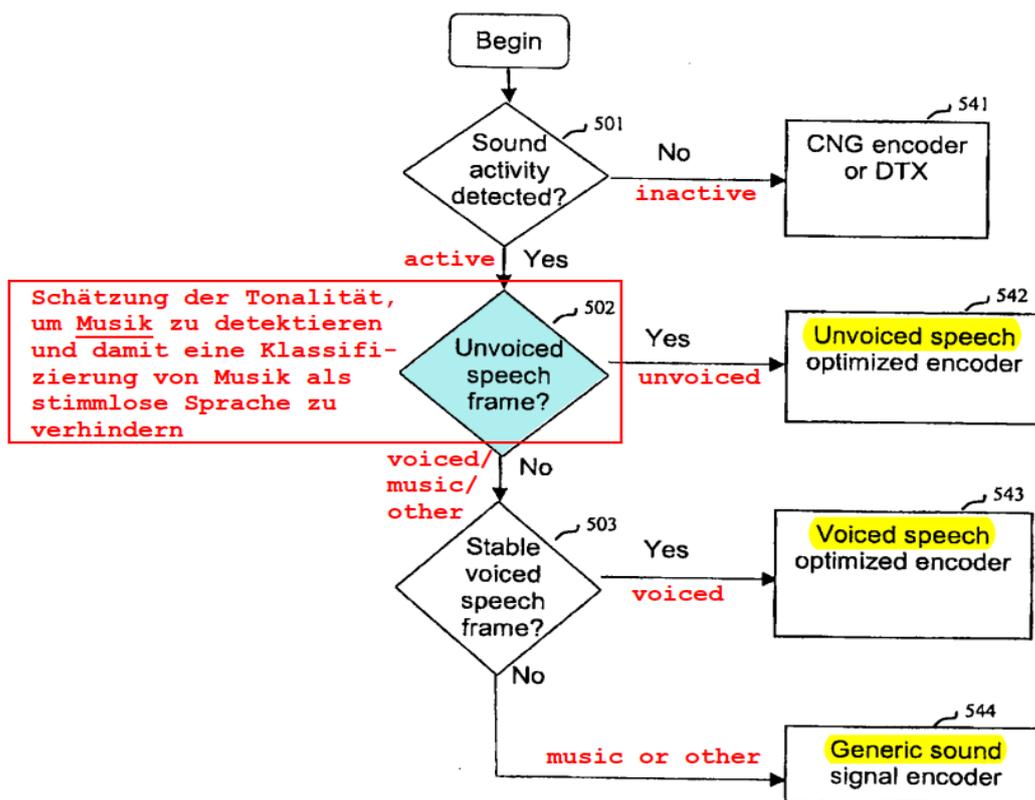
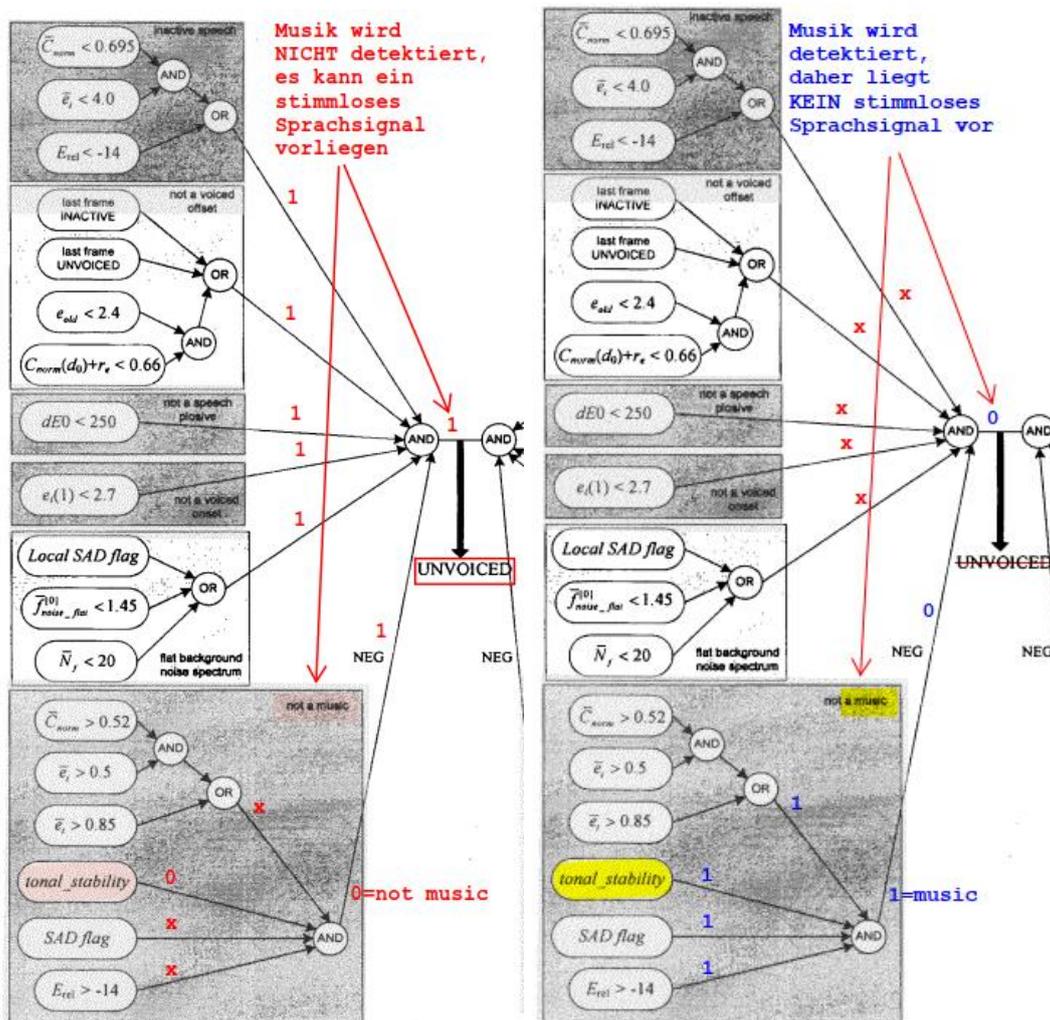


Fig. 5, mit Kommentierung und Kolorierung durch den Senat

Dabei ist die Tonalitätsschätzung nur ein Parameter von mehreren, die darüber entscheiden, ob ein Sprachrahmen mit stimmloser Sprache oder mit Musik vorliegt, wie aus der nachfolgend wiedergegebenen Figur 6 ersichtlich:



zweimal linker Ausschnitt aus Fig. 6, jeweils mit Kommentierung und Kolorierung durch den Senat

2. Als zuständigen Fachmann sieht der Senat einen Ingenieur der Elektro-, Nachrichten- oder Informationstechnik mit einem universitären Master oder Diplom an. Er hat mehrere Jahre Berufserfahrung auf dem Gebiet der Audiocodierung unter Berücksichtigung von Sprache und Musik in Schallsignalen.

3. Einige Merkmale bedürfen der Erläuterung:

a) Das **Spektrum eines Schallsignals** (Merkmal 2.2.3 des erteilten Anspruchs 2) hat in Abhängigkeit von dem Schallsignaltyp (Sprache (stimmhaft, stimmlos), Musik, sonstige Geräusche) unterschiedliche Gestalt. Musiksignale weisen regelmäßig Spektren auf, in denen bestimmte Signalanteile, die Töne, ihre

Frequenz (*position*) und Form (*shape*) über eine gewisse Zeit beibehalten (Streitpatentschrift, Abs. 97; K3, S. 79 – 81, Abschnitt *Measuring harmonic entropy to find pitched notes*; K4, Fig. 2, 3).

In dem Ausführungsbeispiel des Streitpatents werden Schmalband- bzw. Breitband-Schallsignale mit 8 bzw. 16 kHz abgetastet, auf 12,8 kHz aufwärtsabgetastet bzw. dezimiert und quantisiert. Damit enthält ein als Rahmen (*frame*; Merkmal 2.2.1 des erteilten Anspruchs 2) bezeichneter, 20 ms langer, Zeitabschnitt 256 quantisierte Abtastwerte (*samples*). Zwei um 128 Abtastwerte verschobene und gefenserte Abschnitte (Fig. 2) werden jeweils mittels einer schnellen Fouriertransformation (FFT, *Fast Fourier Transform*) in den Frequenzbereich transformiert und liefern zwei Spektren ($X^{(1),(2)}(k)$) mit jeweils 256 komplexwertigen Spektralwerten (*spectral parameters*) pro Rahmen (Abs. 17, 21, 22, 26, 36 - 42):

$$X^{(1)}(k) = \sum_{n=0}^{N-1} x_w^{(1)}(n) e^{-j2\pi \frac{kn}{N}} \quad X^{(2)}(k) = \sum_{n=0}^{N-1} x_w^{(2)}(n) e^{-j2\pi \frac{kn}{N}} \quad k = 0, \dots, L_{FFT} - 1$$

Aus den Real- und Imaginärteilen ($X_{R,I}(k)$) der Spektralwerte werden u. a. die normierte Energie für jede diskrete Frequenz ($E_{BIN}(k)$) sowie das über jeweils zwei Spektralanalysen gemittelte logarithmische Energiespektrum des Schallsignals eines Rahmens (*average log-energy spectrum*; $E_{dB}(k)$) im Sinne des Merkmals 2.2.3 des erteilten Anspruchs 2 bestimmt (Abs. 44, Gl. (3), (4)):

$$E_{BIN}(k) = \frac{4}{L_{FFT}^2} (X_R^2(k) + X_I^2(k)), \quad E_{dB}(k) = 10 \log \left[\frac{1}{2} (E_{BIN}^{(1)}(k) + E_{BIN}^{(2)}(k)) \right], \quad k = 1, \dots, 127$$

b) Unter einem **Residualspektrum** des Schallsignals nach Merkmal 1.2 versteht der Fachmann ein Spektrum, welches aus dem Spektrum des Schallsignals gewonnen ist und dessen lokal dominierende Frequenzanteile stärker hervortreten lässt, um deren zeitliche Stabilität besser beurteilen zu können. Dabei muss sich die Größe und Form der lokal dominierenden Signalanteile des ursprünglichen Schallsignalspektrums nicht exakt im Residualspektrum widerspiegeln. Beispielsweise ist auch ein nur zweiwertiges Spektrum – „1“ (= Spitze) bei lokalen Maxima im Spektrum des Schallsignals, „0“ sonst – als ein Residualspektrum gemäß Merkmal 1.2 anzusehen. Auch ein solchermaßen vereinfachtes Spektrum liefert eine

hinreichend genaue Aussage über die „Töne“, d. h. die lokal dominierenden Frequenzanteile des ursprünglichen Signalspektrums.

Aus der Eigenschaft „aktuell“ entnimmt der Fachmann, dass das Residualspektrum durch Untersuchung eines aktuell untersuchten Zeitabschnitts des Schallsignals gewonnen wurde.

Nach dem insoweit nicht einschränkenden Ausführungsbeispiel des Streitpatents werden zur Berechnung des Residualspektrums in dem gemittelten logarithmischen Energiespektrum ($E_{dB}(k)$) des Schallsignals eines Rahmens zunächst die lokalen Minima gesucht (Merkmal **2.2.1** des erteilten Anspruchs 2) und deren Indizes in einem Puffer (i_{\min}) gespeichert (Abs. 99, 100; Gl. (30)):

$$i_{\min} = (\forall i : (E_{dB}(i-1) > E_{dB}(i)) \wedge (E_{dB}(i) < E_{dB}(i+1))) \quad i = 1, \dots, N_{SPEC} - 2 \quad (30)$$

Die Verbindung der lokalen Minima im Spektrum des Schallsignals mittels einer stückweise linearen Funktion liefert einen spektralen Untergrund (*spectral floor*, $sp_floor(j)$; Merkmal **2.2.2** des erteilten Anspruchs 2), der von dem Energiespektrum des Schallsignals ($E_{dB}(k)$) subtrahiert wird (Merkmal **2.2.3** des erteilten Anspruchs 2). Das Ergebnis ist das Residualspektrum (*residual spectrum*, $E_{dB, res}$; Abs. 101 – 103; Gl. (32); Fig. 3):

$$E_{dB, res}(j) = E_{dB}(j) - sp_floor(j) \quad j = 0, \dots, N_{SPEC} - 1$$

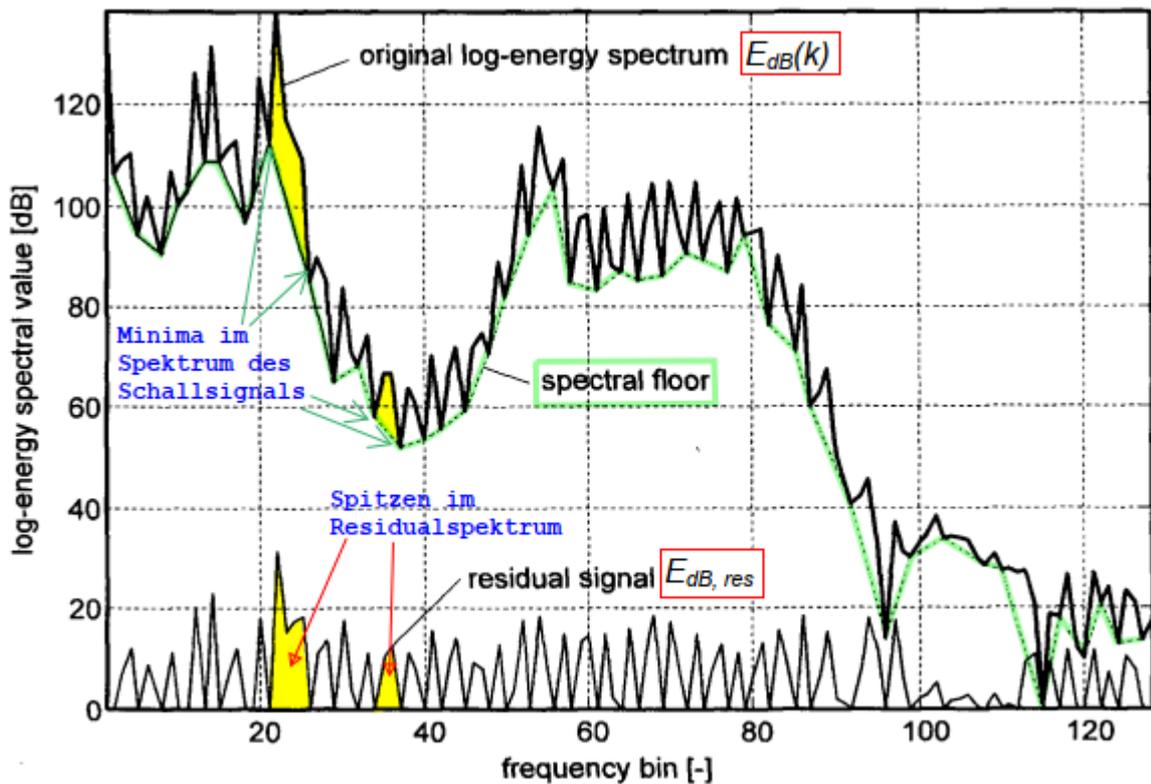


Fig. 3 mit Kolorierung und Kommentierung durch den Senat

c) Große Werte im berechneten Residualspektrum entsprechen lokal dominierenden Signalanteilen im Spektrum des Schallsignals und werden als „**Spitzen**“ (*peaks*; Merkmal 1.3) **erkannt**. Sie sind im Ausführungsbeispiel jeweils von zwei Minima begrenzt, die nicht zu den Spitzen gehören (Abs. 104). Eine Spitze des Residualspektrums kann somit mehrere benachbarte Spektralwerte umfassen.

d) Zum Schätzen der Tonalität des Schallsignals werden fortlaufend Residualspektren berechnet. Ein Vergleich, im Sinne einer Korrelation, der erkannten Spitzen bzw. der zu den einzelnen Spitzen gehörenden Spektralwerte des aktuellen Residualspektrums mit den entsprechenden Spektralwerten eines vorherigen Residualspektrums liefert eine „**Korrelationskarte**“ (*correlation map*) gemäß Merkmal 1.4. Vor dem Hintergrund der Gesamtoffenbarung des Streitpatents ergibt sich, dass nicht etwa für jede erkannte Spitze eine eigene Korrelationskarte berechnet wird, sondern dass für jedes neue (aktuelle) Residualspektrum eine Korrelationskarte berechnet wird, die mindestens so viele Einträge (Werte) aufweist, wie Spitzen im aktuellen Residualspektrum erkannt wurden.

Merkmal 1.4 schließt nicht aus, dass die Korrelationskarte auch für Frequenzen, die nicht zu einer „Spitze“ gehören, einen Eintrag aufweist. Für das vorstehend skizzierte zweiwertige Residualspektrum kann eine gleichfalls zweiwertige Korrelationskarte so berechnet werden, dass sich bei einer Frequenz nur dann ein großer Korrelationswert (z. B. „1“) ergibt, wenn in zwei aufeinanderfolgenden Residualspektren der jeweilige Spektralwert gleich „1“ ist, also ein „Ton“ vorhanden ist. Bei allen anderen Kombinationen („1-0“, „0-1“, „0-0“) läge ein niedriger Korrelationswert vor (z. B. „0“).

Nach dem insoweit nicht einschränkenden Ausführungsbeispiel ist die Korrelationskarte (*cor_map*) eine Relation, die den zu einer Spitze gehörenden Frequenzen einen mittels des aktuellen ($E_{dB, res}(k)$) und des vorherigen ($E_{dB, res}^{(-1)}(k)$) Residualspektrums berechneten Korrelationswert zuordnet (N_{min} = Anzahl der Minima im Spektrum des Schallsignals; i_{min} = Puffer der Indizes der Minima, vgl. Gl. (30)):

$$cor_map(i_{min}(x) : i_{min}(x+1)) = \frac{\left(\sum_{j=i_{min}(x)}^{i_{min}(x+1)-1} E_{dB, res}(j) E_{dB, res}^{(-1)}(j) \right)^2}{\sum_{j=i_{min}(x)}^{i_{min}(x+1)-1} (E_{dB, res}(j))^2 \sum_{j=i_{min}(x)}^{i_{min}(x+1)} (E_{dB, res}^{(-1)}(j))^2},$$

$$x = 0, \dots, N_{min} - 2 \quad (33)$$

Der obere Teil der Figur 4 der Streitpatentschrift zeigt die Residualspektren eines aktuellen (*current frame*) und eines vorhergehenden Rahmens (*previous frame*), der untere Teil der Figur 4 die daraus errechnete Korrelationskarte:

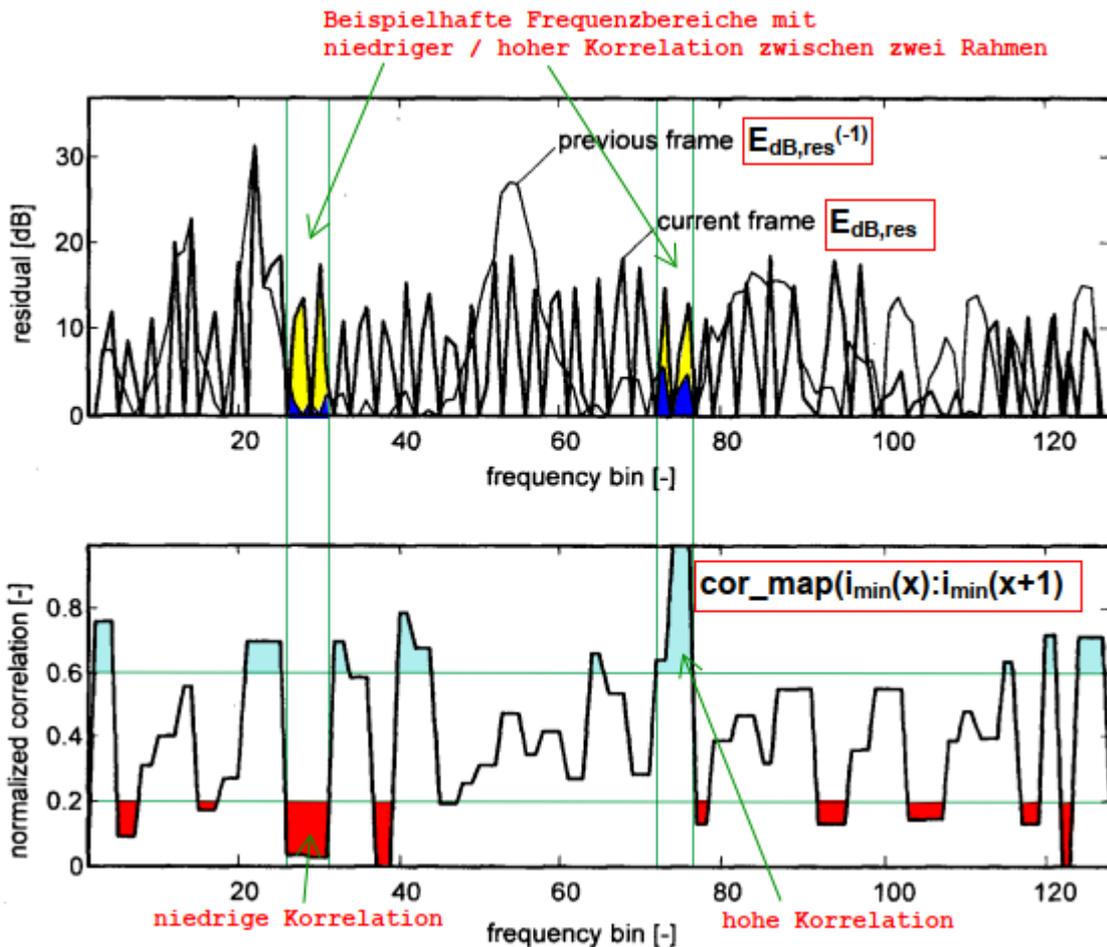


Fig. 4, mit Kolorierung und Kommentierung durch den Senat: Residualspektren zweier aufeinanderfolgender Rahmen und zugehörige Korrelationskarte

e) Gemäß Merkmal 1.5 wird, basierend auf der berechneten Korrelationskarte, eine **Langzeit-Korrelationskarte** berechnet, die eine **Tonalität** des Schallsignals (Merkmale 1.1, 1.5) anzeigt. Während die Korrelationskarte eine Aussage über die Korrelation der Spitzen zweier (unmittelbar) aufeinanderfolgender Residualspektren liefert, berücksichtigt die Langzeit-Korrelationskarte somit mehr als zwei Residualspektren, um eine (bessere) Aussage über die Dauer der einzelnen Töne, also über die **tonale Stabilität** treffen zu können. Die Werte der Langzeit-Korrelationskarte gemäß Merkmal 1.5 sind somit ein Maß für die zeitliche Stabilität der einzelnen spektralen Komponenten im Residualspektrum und damit auch im Spektrum des ursprünglichen Schallsignals. Der Fachmann liest mit, dass die Langzeit-Korrelationskarte mit jedem neu berechneten Residualspektrum bzw. jeder neu berechneten Korrelationskarte aktualisiert wird. Dabei ist dem Fachmann eine sogenannte exponentielle Glättung der Form $y_n = \alpha \cdot y_{n-1} + (1-\alpha) \cdot x_n$ bekannt, um aus den

aktuellen Werten x_n (der Korrelationskarte) und den vergangenen Werten $y_{n-1}, y_{n-2}, \dots, y_{n-m}$ (der Langzeit-Korrelationskarte) einen Schätzwert y_n (der Langzeit-Korrelationskarte) zu gewinnen. In dem skizzierten Fall zweiwertiger Residualspektren und Korrelationskarten ($\{0; 1\}$) ergibt sich mit exponentieller Glättung eine Langzeit-Korrelationskarte, deren Wertemenge (für $0 < \alpha < 1$) die reellen Zahlen zwischen Null und Eins sind.

Auch das Ausführungsbeispiel des Streitpatents verfährt so. Die Langzeit-Korrelationskarte (cor_map_LT) wird rekursiv wie folgt berechnet (Abs. 107, $cor_map_LT(k)$ wird für alle k mit „0“ initialisiert; $\alpha_{map} = 0,9$):

$$cor_map_LT(k) = \alpha_{map} cor_map_LT(k) + (1 - \alpha_{map}) cor_map(k),$$

$$k = 0, \dots, N_{SPEC} - 1, \quad (34)$$

Der Fachmann liest mit, dass die Gleichung (34) wie folgt um einen Laufindex für die Rahmennummer n zu ergänzen ist:

$$cor_map_LT^{(n)}(k) = \alpha_{map} cor_map_LT^{(n-1)}(k) + (1 - \alpha_{map}) cor_map^{(n)}(k) \quad (34^*)$$

Somit ergeben sich die einzelnen Werte der Langzeitkorrelationskarte durch fortlaufende Filterung der entsprechenden Werte der Korrelationskarten mittels eines einpoligen Filters mit der Übertragungsfunktion $H(z) = \frac{1 - \alpha_{map}}{1 - \alpha_{map} z^{-1}}$. (Merkmal 5.5.1 des

erteilten Anspruchs 5), welches für $0 < \alpha_{map} < 1$ stabil ist. Ein Marker (cor_strong) wird auf Eins gesetzt, wenn einer der Werte der Langzeit-Korrelationskarte größer als 0,95 ist (Abs. 109), d. h., wenn ein Ton besonders lange andauert.

f) Nach dem Ausführungsbeispiel wird mit jedem neuen Rahmen eine neue Langzeit-Korrelationskarte berechnet. Durch Aufsummierung ihrer Werte über alle Frequenzen wird rahmenweise ein Summenwert (cor_map_sum) gebildet (Abs. 108) (Merkmal 5.5.2 des erteilten Anspruchs 5: *Summieren der gefilterten Korrelationskarte über die Frequenzbins, um eine summierte Langzeit-Korrelationskarte zu erzeugen*):

$$cor_map_sum = \sum_{j=0}^{N_{SPEC}-1} cor_map_LT(j). \quad (35)$$

Die Entscheidung, ob Tonalität vorliegt oder nicht, wird anhand eines Vergleichs

des Summenwerts (*cor_map_sum*) mit einem adaptiven Schwellwert (*thr_tonal*) getroffen, der auf einen Anfangswert gesetzt, mit jedem Rahmen aktualisiert (Abs. 110) und auf einen bestimmten Wertebereich begrenzt wird (Abs. 111).

Ein Parameter (*tonal_stability*) wird auf „Eins“ gesetzt, wenn der Summenwert (*cor_map_sum*), d. h. die summierte Langzeit-Korrelationskarte, größer als der Schwellwert (*thr_tonal*) ist oder der Marker für besonders stabile (Einzel-)Töne gesetzt ist (*cor_strong* = 1); anderenfalls hat der Parameter den Wert „Null“ (Abs. 112):

```
if ((cor_map_sum > thr_tonal) OR (cor_strong = 1))
    tonal_stability = 1
else
    tonal_stability = 0
end.
```

Das Schallsignal wird somit als „tonal“ eingeschätzt, wenn die aufeinanderfolgenden Residualspektren des Schallsignals einen nicht unerheblichen Anteil von zeitlich hinreichend stabilen Spektralwerten (Tönen) aufweisen und/oder wenn ein (oder mehrere) Spektralwert(e) zeitlich sehr lange konstant bleiben (*cor_strong* = 1, vgl. Abs. 109).

g) Der Fachmann entnimmt dem Streitpatent, dass mit **Tonalität** (*tonality*; Ansprüche 1, 6, 11, 14 - 16, 19, 22, 23, 25 - 27; Abs. 2, 6, 9 - 12, 85, 148, 157) nichts Anderes gemeint ist als die **tonale Stabilität** eines Schallsignals (*tonal stability*; Anspruch 13; Abs. 70, 85, 92, 97, 98, 110, 112, 113, 115, 122, 139, 141, 143, 149).

Nach Ansicht der Klägerin handelt es sich dagegen um zwei Fachbegriffe mit unterschiedlicher Bedeutung. Unter Tonalität verstehe der Fachmann ein System hierarchischer Tonhöhenbeziehungen, während die tonale Stabilität ein Maß für die zeitliche Konstanz der Töne sei.

Die Klägerin verweist u. a. auf den Anspruch 13, der den Begriff *tonality* mit dem unbestimmten Artikel einführe, obwohl Anspruch 13 auf Anspruch 11 rückbezogen und in letzterem der Begriff *tonal stability* bereits eingeführt sei. Somit werde beim Gegenstand des Anspruchs 13 zweifelsfrei zwischen *tonality* und *tonal stability* im

Kontext des Merkmals „*classifying the active sound signal as an unvoiced speech signal*“ unterschieden. Denn anderenfalls, also bei einer Auslegung der beiden Begrifflichkeiten als Synonyme, wären die zusätzlichen Merkmale des Anspruchs 13 teilweise redundant.

Diese Ansicht trifft nicht zu. Die nicht angegriffenen Ansprüche 11 und 13 lauten in der erteilten Fassung:

11. A method for classifying a sound signal in order to optimize encoding of the sound signal using the classification of the sound signal, the method comprising:

detecting a sound activity in the sound signal;
classifying the sound signal as one of an inactive sound signal and an active sound signal according to the detected sound activity in the sound signal; and
in response to the classification of the sound signal as an active sound signal, further classifying the active sound signal as one of an unvoiced speech signal and a non-unvoiced speech signal;
wherein classifying the active sound signal as an unvoiced speech signal comprises estimating a tonality of the sound signal in order to prevent classifying music signals as unvoiced speech signals, wherein the tonality estimation is performed according to any one of claims 1 to 5.

13. A method as defined in claim 11 or 12, wherein classifying the active sound signal as an unvoiced speech signal comprises calculating a decision rule based on at least one of a voicing measure, an average spectral tilt measure, a maximum short-time energy increase at low level, a tonal stability and a relative frame energy.

Anspruch 11 beschäftigt sich demnach mit der Klassifizierung eines Schallsignals als inaktives oder aktives Schallsignale und differenziert letzteres in ein nicht-stimmloses und bzw. stimmloses Sprachsignal. Um eine fehlerhafte Klassifizierung eines Musiksignals als stimmloses Sprachsignal zu verhindern, umfasst die Klassifizierung eines aktiven Schallsignals als stimmloses Sprachsignal das Schätzen einer Tonalität des Schallsignals gemäß einem der Ansprüche 1 bis 5. Da diese sich unstreitig mit der Schätzung der tonalen Stabilität eines Schallsignals beschäftigen, lässt (auch) der Anspruch 11 klar erkennen, dass die beiden Begriffe im Streitpatent synonym verwendet werden.

Der Anspruch 13 macht Angaben dazu, welche Größen in die in Anspruch 11 genannte Klassifizierung eines aktiven Schallsignals als stimmloses Sprachsignal eingehen können. Hierzu wird eine Entscheidungsregel berechnet, die nicht nur auf einer tonalen Stabilität, sondern alternativ oder zusätzlich auf weiteren Größen basiert (gemäß Figur 6 und Abs. 139 – 144 gehen alle im Anspruch 13 genannten Größen und weitere ein), nämlich auf *voicing measure*, *average spectral tilt measure*, *maximum short-time energy increase at low level*, **tonal stability** und *relative frame energy*.

Während Anspruch 11 mit der Schätzung der Tonalität im Sinne der tonalen Stabilität nur einen Teilaspekt der Klassifizierung eines Signals als stimmloses Sprachsignal herausgreift, liefert Anspruch 13 die vorstehend wiedergegebenen weiteren Kriterien.

Damit lassen auch die Ansprüche 11 und 13, vor dem Hintergrund der zugehörigen Beschreibung, erkennen, dass es stets um die Schätzung der tonalen Stabilität geht und nicht um die Bestimmung einer Tonalität in dem Sinne eines Systems hierarchischer „Tonhöhenbeziehung“. Eine Grundlage für eine solche Auslegung des Begriffs der Tonalität kann der Fachmann dem Streitpatent nicht entnehmen. Das Streitpatent bildet hier sein eigenes Lexikon (BGH, Urteil vom 2. März 1999 – X ZR 85/96 – Spanschraube).

Die Klägerin ist weiter der Auffassung, nur das „erste“ Ausführungsbeispiel (Abs. 97 – 112) zeige die Bestimmung der tonalen Stabilität, während sich das „zweite“ Ausführungsbeispiel (Abs. 149 – 159) auf das Schätzen der Tonalität im Sinne einer tonalen Struktur, also einer strukturellen Beziehung zwischen Tonhöhen, beziehe.

Auch diese Ansicht trifft nicht zu. Das „zweite“ Ausführungsbeispiel möchte Rahmen mit starkem tonalen Inhalt bestimmen, um einen geeigneten Codierer auszuwählen und bedient sich dazu der Analyse der tonalen Stabilität gemäß dem ersten Ausführungsbeispiel (Abs. 149: *using the tonal stability analysis described earlier*). Dabei werden zwar einzelne Verfahrensschritte abgeändert, etwa das Bestimmen des spektralen Untergrunds (Abs. 150 – 153), jedoch werden alle im erteilten Anspruch 1 genannten Verfahrensschritte durchgeführt, insbesondere das Berechnen eines Residualspektrums, einer Korrelationskarte und einer Langzeit-Korrelationskarte (Abs. 154 – 156). Aus letzterer bzw. aus der summierten Langzeit-Korrelationskarte (*cor_map_sum*) wird – wie im ersten Ausführungsbeispiel – eine (binäre) Entscheidung über das Vorliegen einer Tonalität im Sinne der tonalen Stabilität getroffen (Abs. 157: *The decision about signal tonality ... is also the same as described earlier in the present disclosure*).

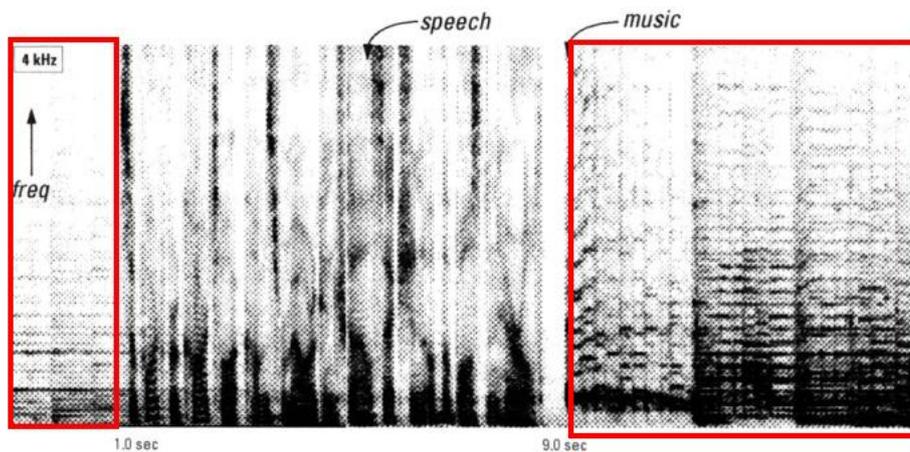
Mithin ist auch im „zweiten“ Beispiel keine Rede von der Bestimmung einer „Struktur von Tönen“ oder einer „strukturellen Beziehung zwischen Tonhöhen“. In diesem Zusammenhang spielt es auch keine Rolle, ob das Streitpatent die Priorität der US-Voranmeldung wirksam in Anspruch nehmen kann oder – wie die Klägerin meint – dies nicht der Fall sei, weil das genannte „zweite“ Ausführungsbeispiel in der Prioritätsanmeldung nicht enthalten sei.

II.

1. Hinsichtlich der angegriffenen Ansprüche 1 und 19 der erteilten Fassung des Streitpatents liegt der Nichtigkeitsgrund der fehlenden Patentfähigkeit vor (Art. 138 Abs. 1 lit. a) i. V. m. Art. 54 EPÜ i. V. m. Art. II § 6 Abs. 1 S. 1 Nr. 1 IntPatÜG).

Der Gegenstand des Anspruchs 1 erweist sich als nicht neu gegenüber dem aus der Dissertation *Structure out of Sound* von M. J. HAWLEY (K3) bekannten Verfahren.

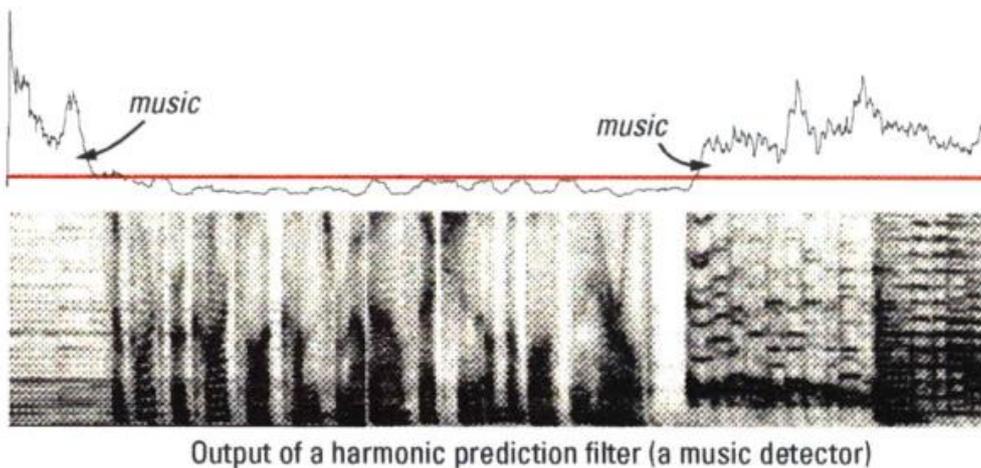
a) HAWLEY beschäftigt sich u. a. mit dem Erkennen von Musik in Audiosignalen (S. 23, Kap. 1.1, Abs. 3: *2: Listening to music [...] a program that finds segments of music in general soundstreams*) mittels eines **Musikdetektors** (S. 78ff, Kap. 2.3 *A music detector*). Dieser basiert auf der Erkenntnis, dass Musik, im Gegensatz zu Sprache, **Töne** enthält, die über einen vergleichsweise langen Zeitraum eine konstante Frequenz aufweisen (S. 79, Abs. 2: *music [...] has [...] notes with frequencies that remain relatively fixed for a finite period of time*; S. 80, Abs. 1: *but unlike most music, the harmonics of speech vary rather rapidly in pitch.*):



HAWLEY, S. 79: mit vom Senat vorgenommenen Markierungen der Musikabschnitte

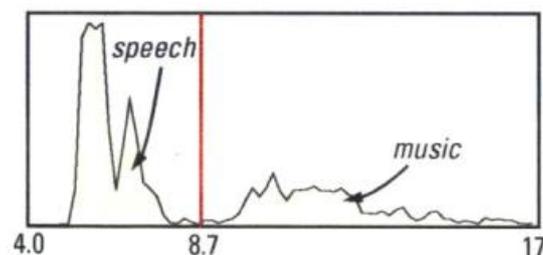
Die vorstehend wiedergegebene Figur zeigt Amplitudenspektren (Abszisse: Zeit, Ordinate: Frequenz, Schwärzung/Graustufen: Amplitude) eines 13 Sekunden langen Ausschnitts eines Audiosignals ($f_{\text{sample}} = 8 \text{ kHz}$; $N_{\text{Samples/Rahmen}} = 48$; $t_{\text{Rahmen}} = 6 \text{ ms}$; $N_{\text{Rahmen in 13 s}} = 2167$). In den rot umrahmten Zeitabschnitten zeigen die Spektren eine für Musik typische kammartige Struktur, d. h. einzelne (dominierende) Signalanteile weisen über mehrere Rahmen eine annähernd konstante Frequenz auf (S. 79, Abs. 2 – S. 80, Abs. 1).

Der Musikdetektor ist als Filter (*harmonic prediction filter*) implementiert, dessen mehrere Rahmen umfassendes Analysefenster über die zeitlich aufeinanderfolgenden (Rahmen-)Spektren gleitet und dabei die Spitzen, also die ausgeprägten und zeitlich stabilen Frequenzen (Töne), sammelt. Nach jedem Weiterrücken des Analysefensters, d. h. nach jedem neu eingelesenen Rahmen, gibt der Musikdetektor die durchschnittliche Dauer (in der Maßeinheit der Anzahl von Rahmen) der Frequenz-Spitzen an. Dieser gleitende Mittelwert ist hoch für Musik und niedrig für Sprache, so dass – wie im Streitpatent – durch Vergleich mit einem geeignet gewählten Schwellwert festgestellt werden kann, welche Signalart vorliegt (S. 80, Abs. 2):



HAWLEY, S. 80

HAWLEY kommt zu der Erkenntnis, dass sich Sprache und Musik in Audiosignalen in dieser Hinsicht deutlich unterscheiden, wie auch das nachfolgend eingeblendete Histogramm zeigt. Dabei wurde die Grenze zwischen Sprache und Musik bei einer durchschnittlichen Spitzen-Dauer von 8,7 Rahmen definiert, d. h. Musik hat eine durchschnittliche Tondauer vor mehr als 52 ms ($8,7 \cdot 6$ ms):



HAWLEY, S. 80: Histogramm; Abszisse: durchschnittliche Spitzen-Dauer in Rahmen

Damit zeigt HAWLEY ein Verfahren zum Schätzen der Tonalität, im Sinne der tonalen Stabilität, nach Merkmal 1.1, welches – wie das Verfahren nach Anspruch 1 des Streitpatents – dazu dient, Musik von Sprache zu unterscheiden.

b) HAWLEY realisiert den Musikdetektor in Form eines in der Programmiersprache C abgefassten Softwarepakets. Nachfolgend eingeblendet ist ein Ausschnitt (S. 176), der u. a. die Programme *findPeaks* und *music* zeigt:

```
frame
ove
box
008
009
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
```

```
#define HalfWin (WinSize/2)
#define N (WinSize*2)

int minRun = 3,
    maxRun = MaxPeaks/3,
    peakN = 0,
    peakT = 0,
    Peak[MaxPeaks][HalfWin],
    numPeaks = MaxPeaks,
    MinBin = 10, MaxBin = 70,
    CurFrame = 0;
float Offset = 0.0;

// shortest run counted as a held peak
// count really long runs early
// # of held runs in frame delay buffer
// total length of the runs
// the delay buffer (see P(...) macro)
// length of the delay buffer
// consider frequencies in this range
// the current frame number
// starting position in file

#define P(n) Peak[(CurFrame-(n))%numPeaks] // nth-delayed frame of peaks
#define Spectrum(m,h,s,n) WindowedHartley(h,n,s,n). sp_LogMag16(m,h,n)

findPeaks(int *p, int *m, int n){
    // 'm' is a magnitude spectrum; p[i] ← isPeak(m[i])? 1 : 0
    int i;
    bzero(p,n*sizeof(int));
    for (i=MinBin,m+=MinBin,p+=MinBin;i<MaxBin;)
        if (m[0]>=m[1] && m[0]>m[2] && m[0]>=m[-1] && m[0]>m[-2])
            *p = 1, i += 3, m += 3, p += 3;
        else
            i++, m++, p++;
}

music(int *s, int n){
    int m[N], h[N], *p3, *p2, *p1, *p0, *p1, *p0, n2=n/2;
    float f;

    // decrement peak accumulators:
    for (p0=P(0)+MinBin, p1=P(0)+MaxBin; p0 < p1; p0++)
        if (*p0 > minRun) peakN--, peakT -= *p0;

    // find a new frame of peaks
    Spectrum(m,h,s,n);
    findPeaks(P(0),m,n2);

    // 4-point smoother - emphasize sustained peaks: ...1101 => 1111
    p3=P(3)+MinBin, p2=P(2)+MinBin, p1=P(1)+MinBin, p0=P(0)+MinBin;
    for (p1=p0+MaxBin; p0<p1; p3++, p2++, p1++, p0++)
        if (*p3 && *p2 && !*p1) *p1 = *p0;

    // merge incoming columns of peaks and increment peak accumulator
    for (p2=P(4)+MinBin, p3=P(3)+MinBin, p1=P(3)+MaxBin; p3<p1; p2++, p3++)
        if (*p2 && *p3 && *p2 < maxRun)
            *p3 += *p2, *p2 = 0;
        else
            if (*p2 > minRun)
                peakN++, peakT += *p2;
    f = (float)peakT/(float)(peakN? peakN:1);
}
```

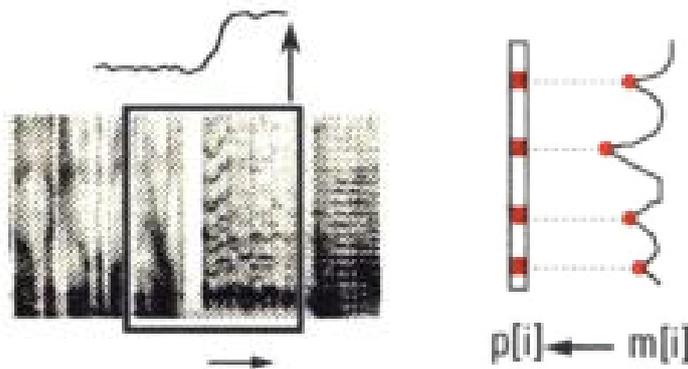
HAWLEY, S. 176

In dem Programm *music* wird durch den Aufruf des Programms *Spectrum(m, h, s, n)* aus den Abtastwerten *s* eines neuen Rahmens ein Amplitudenspektrum *m[i]* als logarithmierter Betrag der Fourierkoeffizienten *F[i]* gebildet (S. 81, unten):

$m[i] = \log|F[i]|$, where $F[i]$ is the discrete Fourier term
at frequency = $i \cdot \text{SampleRate} / \text{FrameSize}$

Danach werden mittels des Aufrufs des Programms *findPeaks(P(0),m,n2)* in dem Amplitudenspektrum $m[i]$ des neu eingelesenen Rahmens die Spitzen, also die lokalen Maxima, detektiert. Hierzu werden die Spektralwerte jeweils mit vier Nachbarwerten verglichen (S. 176: *findPeaks [...] m[0] >= m[1] && m[0] > m[2] && m[0] >= m[-1] && m[0] > m[-2]*) und liefern den sogenannten Spitzenrahmen (*peak frame p[i]*), dessen Einträge für die Frequenzen i , bei denen im Amplitudenspektrum $m[i]$ lokale Maxima vorliegen, den Wert „1“ (*findPeaks [...] *p = 1*) und sonst den Wert „0“ (*findPeaks [...] bzero(p,...)*) hat (S. 81, letzter Absatz):

$p[i] = 1$ if $m[i]$ is a local maximum (in i), else 0



HAWLEY, S. 81: Erzeugung von $p[i]$ aus $m[i]$

Der aus den Spitzenrahmen $p[i]$ gebildete Puffer $P(t)[i]$ (S. 176, *#define P(n)*) gibt an, bei welcher Frequenz i des Rahmens mit der Nr. t eine Spitze vorliegt. Während im ursprünglichen Amplitudenspektrum $m[i]$ des Schallsignals noch eine Vielzahl unterschiedlicher Amplitudenwerte vorhanden ist, ist der daraus abgeleiteten Puffers $P(t)[i]$ zweiwertig (S. 82, Abs. 1).

Der Puffer $P(0)[i]$ ist ein aktuelles **Residualspektrum** des Schallsignals gemäß Merkmal 1.2, denn er ist ein Spektrum (für jede Frequenz i liefert $P(0)[i]$ eine Aussage über die Amplitude, nämlich „1“ oder „0“), ist aus dem Spektrum $m[i]$ des Schallsignals s des aktuellen Rahmens berechnet und lässt dessen dominierende Frequenzanteile, die lokalen Maxima, stärker hervortreten.

c) Da das aus HAWLEY bekannte Residualspektrum $P(0)[i]$ nur zweiwertig ist, sind auch automatisch seine **Spitzen** ($P(0)[i] = 1$) **erkannt**, wie von Merkmal **1.3** gefordert.

d) Der nachfolgend eingeblendete Abschnitt des Programms *music* (S. 82) zeigt die Bildung des gleitenden Mittelwerts ($peakT/peakN$) der durchschnittlichen Spitzendauer mittels einer *for*-Schleife und darin enthaltener *if / else if*-Abfrage:

```
// merge incoming column of peaks
// and increment peak accumulator
for (p2=P(4)+MinBin, p3=P(3)+MinBin, p1=P(3)+MaxBin;
     p3<p1; p2++, p3++)
    if (*p2 && *p3 && *p2 < maxRun)
        *p3 += *p2, *p2 = 0;
    else if (*p2 > minRun)
        peakN++, peakT += *p2;

output((float)peakT/(float)(peakN? peakN:1));
CurFrame++;
```

Dabei wird ein Frequenzbereich von 150 - 1000 Hz ($MinBin = 10$ bis $MaxBin = 70$) ausgewertet und es werden nur Spitzen (als Töne) gezählt, die mehr als drei ($*p2 > minRun = 3$) und weniger als 42 aufeinanderfolgende Rahmen andauern ($*p2 < maxRun = MaxPeaks/3 = 128/3$) (S. 84, Abs. 2; S. 176).

In den Speicherzellen, auf die die Zeigervariable $*p2$ zeigt, ist vor Beginn der *for*-Schleife für jede Frequenz die über die vorherigen Rahmen summierte Anzahl von unmittelbar aufeinanderfolgenden Spitzen gespeichert, während in der Speicherzelle, auf die $*p3$ zeigt, im aktuellen Rahmen für jede Frequenz eine „1“ (Spitze) oder eine „0“ (keine Spitze) vorliegt. Mittels der über den auszuwertenden Frequenzbereich laufenden *for*-Schleife und der darin enthaltenen *if*-Abfrage wird zunächst für jede einzelne Frequenz überprüft, ob die drei Bedingungen $*p2 \ \&\& \ *p3 \ \&\& \ *p2 < maxRun$ erfüllt sind, wobei die *if*-Abfrage in der Programmiersprache C streng von links nach rechts ausgewertet wird, d. h. bei Nicht-Erfüllung der ersten Bedingung werden die weiteren Bedingungen nicht mehr geprüft. Die drei Bedingungen haben die folgende Bedeutung:

1. Lagen in dem vorherigen und ggfs. in noch älteren Rahmen Spitzen bei dieser Frequenz vor ($*p2$)? Diese Bedingung ist nur dann wahr, wenn $*p2 \neq 0$, z. B. $*p2 = 7$.

2. Liegt im aktuellen Rahmen eine Spitze bei der Frequenz vor ($*p3$)? Diese Bedingung ist nur dann wahr, wenn $*p3 \neq 0$, also nur für $*p3 = 1$.
3. Liegen in den vorherigen Rahmen weniger als *maxRun* aufeinanderfolgende Spitzen bei der Frequenz vor ($*p2 < MaxRun$)?

Sind alle drei Bedingungen erfüllt, wird – für jede Frequenz – in der Speicherzelle, auf die die Zeigervariable $*p3$ zeigt, der neue Summenwert der Spitzen gespeichert ($*p3 += *p2$; im Beispiel: $*p3 = *p3 + p2 = 1 + 7 = 8$) und die Speicherzelle, auf die die Zeigervariable $p2$ zeigt, auf „0“ gesetzt ($*p2 = 0$).

Falls eine (oder mehrere) der Bedingungen nicht erfüllt ist (sind), da z. B. im aktuellen Rahmen bei der betrachteten Frequenz keine Spitze vorliegt ($*p3 = 0$), wird im *else-if*-Zweig geprüft, ob die Anzahl der Spitzen in den vorigen Rahmen größer als der Wert *minRun* ist ($*p2 > minRun = 3$). Ist dies der Fall (im Beispiel: $*p2 = 7$), wird der Zähler *peakN*, der die Anzahl der Spitzen-Läufe angibt, um Eins ($peakN++$) und der Zähler *peakT*, der die Gesamtzahl der Spitzen angibt, um den Wert der bis dahin aufgelaufenen Spitzen erhöht ($peakT += *p2$).

Sind weder die drei Bedingungen des *if*-Zweigs, noch die Bedingung des *else-if*-Zweigs erfüllt, z. B. weil bei der betrachteten Frequenz im aktuellen Rahmen keine Spitze vorliegt ($*p3 = 0$) und in den vorigen Rahmen nur drei Spitzen aufeinander folgten ($*p2 = 3$), so werden weder die Speicherbereiche, auf die die Zeigervariablen $*p2$ und $*p3$ zeigen, noch die Zähler *peakN* und *peakT* verändert.

Nach dem Ende der *for*-Schleife, also der Abarbeitung eines Rahmens, werden die in der Zeigervariablen $*p3$ gespeicherten Werte auf die Zeigervariable $*p2$ umgespeichert, so dass bei einem neu eingelesenen Rahmen die Zeigervariable $*p3$ auf das aktuelle Residualspektrum verweist. Die vorstehend skizzierte Funktionsweise des Musikdetektors nach HAWLEY ergibt sich auch durch die von der Beklagten in der mündlichen Verhandlung überreichten Anlage NB4 (vgl. dort insbesondere die Tabelle im mittleren Bereich der einzigen Seite).

e) Das Ergebnis der Überprüfung $*p2 \ \&\& \ *p3$ ist bei einer Frequenz nur dann gleich „1“ (WAHR), wenn sowohl in dem vorherigen als auch im aktuellen Residualspektrum bei dieser Frequenz eine Spitze vorlag bzw. vorliegt. Anderenfalls ergibt sich eine „0“ (FALSCH). Damit liegt aber für jede Frequenz, also auch für die

Frequenzen, bei denen im aktuellen Rahmen eine Spitze erkannt wurde, ein Korrelationswert vor („1“ oder „0“), so dass eine **Korrelationskarte** gemäß Merkmal 1.4 berechnet wird. Dabei hat es auf das Ergebnis der Überprüfung $*p2$ & $*p3$, also auf das Berechnen der Korrelationskarte, keinen Einfluss, wie viele Spitzen in den Vorgängerrahmen vorlagen (z. B.: $*p2 = 7$ oder $*p2 = 1$), maßgeblich ist alleine, dass im (unmittelbar) vorherigen Residualspektrum eine Spitze vorlag, dass also $*p2 \neq 0$ gilt.

f) Basierend auf dem Ergebnis der Überprüfung $*p2$ & $*p3$, d. h. basierend auf der berechneten Korrelationskarte, wird mit der Anweisung $*p3 += *p2$ eine **Langzeitkorrelationskarte** gemäß Merkmal 1.5 berechnet, denn für jede Frequenz gibt der Speicher, auf den die Zeigervariable $*p3$, bzw. $*p2$ nach erfolgtem Umspeichern, zeigt (vgl. NB4, Tabelle, Spalte 2 zeigt Inhalt von $*p2$ nach Übernahme des Werte von $*p3$, Spalte 3 zeigt Inhalt von $*p3$ vor Übergabe nach $*p2$), an, über wie viele Rahmen eine Spitze vorhanden ist, was zweifelsfrei eine Tonalität anzeigt, wie von Merkmal 1.5 gefordert. Dies gilt auch in den Fällen, in denen bei einer Frequenz zwar in den vorherigen Rahmen keine ($*p2 = 0$), im aktuellen Rahmen jedoch eine Spitze ($*p3 = 1$) vorliegt. Denn auch dann gibt der Speicher, auf den die Zeigervariable $*p3$ bzw. $*p2$ zeigt, an, für wie viele Rahmen eine Spitze vorliegt – in diesem Fall ein Rahmen.

Nach alledem zeigt HAWLEY ein Verfahren zum Schätzen der Tonalität gemäß Anspruch 1 des Streitpatents. Entsprechendes gilt für den Vorrichtungsanspruch 19.

2. Der Hilfsantrag 1 ist zulässig. Die Fassung des Streitpatents im angegriffenen Umfang nach Hilfsantrag 1 vom 28. Mai 2021 verteidigt die Beklagte erfolgreich.

Hilfsantrag 1 ergänzt im Anspruch 1 ein Merkmal aus dem erteilten Anspruch 2 (Unterschiede zwischen Merkmal 2.2.3^{Hi1} und Merkmal 2.2.3 sind durch Streichungen bzw. Unterstreichungen gekennzeichnet):

- 1.2 Berechnen eines aktuellen Residualspektrums des Schallsignals;
calculating a current residual spectrum of the sound signal;
- 2.2.3^{Hi1} ~~durch~~ Subtrahieren eines des geschätzten spektralen Untergrunds von einem ~~vom~~ Spektrum des Schallsignals in einem ~~im~~ aktuellen Rahmen; ~~um so das aktuelle Residualspektrum zu erzeugen~~
by subtracting a the estimated spectral floor from a the spectrum of the sound signal in a the current frame; ~~so as to produce the current residual spectrum.~~

a) Nach Merkmal 2.2.3^{Hi1} wird das aktuelle Residualspektrum des Schallsignals (Merkmal 1.2) also dadurch gebildet, dass ein **spektraler Untergrund** vom dem Spektrum des Schallsignals in einem aktuellen Rahmen **subtrahiert** wird. Der Untergrund (*floor*) bzw. Untergrenze eines Spektrums ist – jedenfalls lokal – kleiner als der kleinste Spektralwert, da anderenfalls durch die Subtraktion negative Werte im Residualspektrum auftreten würden. Der spektrale Untergrund muss nicht – wie im Ausführungsbeispiel nach Figur 3 bzw. nach dem erteilten Anspruch 2 – lokal an das Spektrum angepasst sein. Vielmehr könnte er im betrachteten Frequenzbereich, nach Art eines Rauschbodens, konstant sein. Dem Fachmann ist bewusst, dass die Anzahl und die Frequenzen der Spektralwerte des Schallsignals und des spektralen Untergrunds gleich sein müssen, um eine Subtraktion der einzelnen Spektralwerte voneinander zu ermöglichen.

b) Der Anspruch 1 nach Hilfsantrag 1 geht nicht über den Inhalt der Anmeldung in der ursprünglichen Fassung hinaus. Das zusätzliche Merkmal 2.2.3^{Hi1} ist wie folgt ursprünglich offenbart (WO 2009/000073 A1 = Anlage K2b; Unterstreichungen hinzugefügt):

- *calculating a spectral floor (through a spectral floor estimator for example) and subtracting it from the spectrum (via a suitable subtractor for example). (S. 32, Z. 14 – 16)*
- *the spectral floor is subtracted from the spectrum using the following relation: $E_{dB,res}(j) = E_{dB}(j) - sp_floor(j)$ $j = 0, \dots, N_{SPEC} - 1$ (32) and the result is called the residual spectrum (S. 33, Z. 4 – 9)*

Der Fachmann hat der Anmeldung als zur Erfindung gehörend entnommen, dass

ein Residualspektrum des Schallsignals durch Subtrahieren eines spektralen Untergrunds von dem Spektrum des Schallsignals gewonnen wird. Die exakte Vorgehensweise zur Berechnung des spektralen Untergrunds ist von untergeordneter Bedeutung. Die Gewinnung des spektralen Untergrunds nach allen Merkmalen des erteilten Anspruchs 2, also auch *searching for minima ...* und *estimating a spectral floor by connecting the minima* ist für den Fachmann nur eine von mehreren denkbaren Ausgestaltungen. Darüber hinaus lehrt auch das Streitpatent eine andere Möglichkeit zur Bestimmung des spektralen Untergrunds (Abs. 150 – 153; Anspruch 15). Insofern trifft die Auffassung der Klägerin nicht zu, es sei nur das Subtrahieren eines solchen spektralen Untergrunds (*floor*) ursprünglich offenbart, der durch die abschnittsweise Verbindung benachbarter Minima durch Linien gebildet wird.

c) Die Ansprüche nach Hilfsantrag 1 sind deutlich und knapp gefasst und von der Beschreibung gestützt (Art. 84 EPÜ).

In der mündlichen Verhandlung hat die Klägerin erstmals geltend gemacht, der Gegenstand des Anspruchs 1 nach Hilfsantrag 1 erfülle nicht die Anforderungen nach Art. 84 EPÜ, weil gemäß Merkmal 2.2.3^{Hi1} das Spektrum des Schallsignals in einem aktuellen Rahmen zu berechnen sei, wohingegen nach dem Ausführungsbeispiel keine rahmenweise Berechnung stattfindet.

Abgesehen davon, dass das Merkmal 2.2.3^{Hi1} im Wesentlichen dem Merkmal 2.2.3 des erteilten Anspruchs 2 entspricht und daher eine Prüfung auf „Klarheit“ nach Art. 84 nicht möglich ist (BGH, Urteil vom 27. Oktober 2015 – X ZR 11/13 – Fugenband), trifft die Behauptung der Klägerin nicht zu. Gemäß den Angaben in der Beschreibung werden zwar je Rahmen zwei Spektralanalysen mit einer (zeitlichen) Überlappung von 50 % durchgeführt. Diese werden jedoch geeignet kombiniert, um ein Spektrum ($E_{dB}(k)$) des Schallsignals in einem aktuellen Rahmen zu berechnen (Abs. 27, 37 – 39, 44 - 46, Fig. 2 und insbesondere Abs. 44, Gl. (4)), von dem dann der spektrale Untergrund subtrahiert wird, um ein aktuelles Residualspektrum des Schallsignals zu erhalten (Abs. 103, Gl. (32)).

d) Das Streitpatent offenbart die Erfindung so deutlich und vollständig, dass ein Fachmann sie ausführen kann (Art. II § 6 Abs. 1 Nr. 2 IntPatÜG, Art. 138 Abs. 1 lit b EPÜ).

Nach Ansicht der Klägerin offenbart das Streitpatent nicht, wie die Tonalität eines Schallsignals zu schätzen sei. Vielmehr beschäftigt es sich ausschließlich mit dem Schätzen der tonalen Stabilität. Letzteres trifft – wie zur Auslegung dargelegt – zwar zu. Da das Streitpatent die beiden Begriffe jedoch synonym i. S. d. tonalen Stabilität verwendet, offenbart das Streitpatent die Erfindung, also das Schätzen der Tonalität, so deutlich und vollständig, dass ein Fachmann sie ausführen kann. Die hierzu erforderlichen einzelnen Verfahrensschritte sind in dem Streitpatent detailliert beschrieben und ermöglichen dem Fachmann so die Nacharbeitung der Erfindung.

e) In der Fassung nach Hilfsantrag 1 erweist sich das Streitpatent als patentfähig (Art. 138 Abs. 1 Buchstabe a) i. V. m. Art. 54 EPÜ i. V. m. Art. II § 6 Abs. 1 S. 1 Nr. 1 IntPatÜG).

(i) Der Anspruch 1 nach Hilfsantrag 1 ist neu gegenüber HAWLEY (K3).

Die Merkmale 1.1 bis 1.5 sind zwar – wie zum Hauptantrag dargelegt – aus HAWLEY bekannt. Das Berechnen eines aktuellen Residualspektrums des Schallsignals (Merkmal 1.2) findet bei HAWLEY jedoch **nicht** durch das Subtrahieren eines spektralen Untergrunds von einem Spektrum des Schallsignals in einem aktuellen Rahmen statt (nicht Merkmal 2.2.3^{Hi1}).

HAWLEY berechnet aus den Abtastwerten s des Schallsignals eines aktuellen Rahmens die Spektralwerte $F[i]$ und bestimmt daraus mittels der Beziehung $m[i] = \log|F[i]|$ ein logarithmisches Amplitudenspektrum (S. 81, letzter Absatz; S. 176: Programme $Spectrum(m, h, s, n)$, $WindowedHartley(h, n, s, n)$). Mittels des

Programms *findPeaks* werden sodann die lokalen Maxima in dem Amplitudenspektrum $m[i]$ bestimmt und in dem „Spitzenrahmen“ (*peak frame*) $p[i]$ wird gespeichert, bei welchen Frequenzen ein lokales Maximum vorliegt (S. 81: $p[i]=1$ if $m[i]$ is a local maximum (in i), else 0). Damit ist – wie zum Hauptantrag dargelegt – das aktuelle Residualspektrum des Schallsignals berechnet (Merkmal **1.2**) und dessen Spitzen ($p[i] = 1$) erkannt (Merkmal **1.3**), hierfür wurde jedoch kein spektraler Untergrund von dem Spektrum des Schallsignals subtrahiert. Merkmal 2.2.3^{Hi1} ist somit aus HAWLEY nicht bekannt.

(ii) Der Anspruch 1 nach Hilfsantrag 1 ist neu gegenüber MINAMI (K4).

MINAMI schlägt vor, Videomaterial anhand von Audioinformationen zu indexieren, um einem Benutzer einen schnellen Zugang zu bestimmten Szenen zu ermöglichen (S. 17, re. Sp., Abs. 2; S. 19, re. Sp., Abs. 3). Eine Audiosignalanalyse unterscheidet zwischen Musik und Sprache (S. 17, re. Sp., Abs. 2, letzter Satz), basiert auf der Dissertation von HAWLEY (S. 19, re. Sp., Abs. 4; S. 25, li. Sp., Quelle Nr. 24) und wandelt diese insofern ab, als dass ein Kantendetektions-Algorithmus zur Ermittlung der Spitzen, also der lokalen Maxima, im Spektrum verwendet wird (S. 19, re. Sp., Abs. 4).

MINAMI berechnet rahmenweise (alle 64 ms) Spektren des Schallsignals (S. 20, li. Sp., Abs. 1) und behandelt diese als Graustufenbilder, in welchen mittels eines Kantendetektionsoperators anhand der Helligkeit der Pixel die Spitzen verortet werden (S. 20, li. Sp., letzter Abs.; Merkmale **1.2** und **1.3**). Die Kanten-Intensität der Spitzen wird durch Vergleich der Intensitäten der in Frequenzrichtung benachbarten Pixel ermittelt (S. 20, re. Sp., Satz 1). Die Kanten-Intensitäten bei einer diskreten Frequenz werden über eine Vielzahl von Rahmen (z. B. für eine oder eine halbe Sekunde) summiert. Die nachfolgende Summation dieser Werte in der Frequenzrichtung liefert die Gesamt-Kanten-Intensität, die bei Musiksignalen hoch ist (S. 20, re. Sp., Abs. 1, 2). Damit offenbart MINAMI zwar ein Verfahren zum Schätzen der Tonalität eines Schallsignals (Merkmal **1.1**) und berechnet eine Größe (Gesamt-Kanten-Intensität), die eine Tonalität im Schallsignal anzeigt (Teil des Merk-

mals 1.5). Die Summationen in Zeit- und Frequenzrichtung realisieren jedoch weder das Berechnen einer Korrelationskarte noch einer Langzeitkorrelationskarte (**nicht** Merkmale 1.4 und 1.5).

MINAMI zeigt auch **nicht** das Subtrahieren eines spektralen Untergrunds von einem Spektrum des Schallsignals in einem aktuellen Rahmen gemäß Merkmal 2.2.3^{Hi1}.

(iii) Der Anspruch 1 nach Hilfsantrag 1 beruht auch auf einer erfinderischen Tätigkeit.

Nach Ansicht der Klägerin lese der Fachmann bei HAWLEY mit bzw. es ergebe sich für ihn in naheliegender Weise, dass von dem Spektrum $m[i]$ des Schallsignals ein spektraler Untergrund abgezogen werde. Denn anderenfalls würde der Hintergrund in denen das Spektrum darstellenden Graustufenbildern (K3, S. 81 – 83) einen Helligkeitsverlauf zeigen, der den Kontrast der Spitzen bei weitem übersteige. Dann wären die hohen Frequenzen nicht zu erkennen, da bei Sprach- und Musiksignalen die unteren Frequenzbereiche regelmäßig wesentlich größere Amplituden aufwiesen als die oberen. Ohne Subtraktion eines frequenzabhängigen spektralen Untergrunds müsste der Hintergrund bei unteren Frequenzen um ein Mehrfaches dunkler sein als die Spitzen bei hohen Frequenzen.

Dies trifft aus mehreren Gründen nicht zu. Zum einen zeigen die Graustufenbilder eine deutliche Zunahme der Schwärzung des Hintergrunds zu tieferen Frequenzen (S. 79), was dagegen spricht, dass ein spektraler Untergrund abgezogen wurde. Zum anderen dienen die Graustufenbilder in HAWLEY lediglich der Illustration der spektralen Unterschiede zwischen Musik- und Sprachsignalen. Der Fachmann kann HAWLEY nicht entnehmen, dass die Graustufenbilder in einem Zusammenhang mit der den Musikdetektor realisierenden Software stehen. Deren Aufbau ist auf den Seiten 82 und 176 angegeben. Dort finden sich keine Programme oder Programmteile, die eine Subtraktion eines spektralen Untergrunds von dem Spektrum des Schallsignals realisieren würden.

Hinzu kommt, dass HAWLEY die spektrale Analyse in dem Musikfilter auf den Frequenzbereich zwischen 150 und 1000 Hz beschränkt (S. 84, Abs. 2, Satz 2), so dass die von der Klägerin angeführten großen Dynamikunterschiede zwischen unterem und oberem Frequenzbereich des Schallsignalspektrums ohnehin nicht vorhanden sind.

Weiter ist zu beachten, dass HAWLEY das Residualspektrum $p[i]$ aus dem logarithmischen Amplitudenspektrum $m[i]$ des Schallsignals durch Suche der lokalen Maxima gewinnt. Die *for*-Schleife der Funktion *findPeaks* (K3, S. 176) durchläuft den Frequenzbereich von *MinBin* bis *MaxBin*, und die *if*-Abfrage prüft, ob ein Spektralwert größer oder gleich als seine unmittelbaren und größer als seine übernächsten Nachbarwerte ist. Ist dies der Fall, wird ein lokales Maximum festgestellt ($*p = 1$) und drei Frequenzen weiter gesprungen ($i += 3, m += 3, p += 3$). Wird kein lokales Maximum festgestellt, wird die Prüfung mit dem unmittelbar benachbarten Spektralwert fortgesetzt (*else*-Zweig: $i++, m++, p++$).

Der Fachmann erkennt, dass die Anzahl der Vergleiche mit Nachbarwerten zum einen so groß ist (zwei Vergleiche wäre das Minimum, vier Vergleiche werden durchgeführt), dass kleine rauschartige Spitzen nicht versehentlich als lokale Maxima erkannt werden, zum anderen aber so klein ist, dass lokale Maxima auch dann noch erkannt werden, wenn die Grobstruktur des Spektrums in dem relevanten Bereich relativ steil ansteigt oder abfällt.

Damit präsentiert HAWLEY eine in sich geschlossene und gut funktionierende (K3, S. 86, vorletzter Abs.: *2 errors per hour*) Lösung für das Berechnen eines aktuellen Residualspektrums des Schallsignals und liefert dem Fachmann somit keinen Anlass, eine „Begradigung“ des Schallsignalspektrums durch Subtrahieren eines spektralen Untergrunds in Betracht zu ziehen.

Auch MINAMI vermag dem Fachmann keine Hinweise in dieser Richtung zu geben, da aus MINAMI das Subtrahieren eines spektralen Untergrunds ebenfalls nicht bekannt ist.

(iv) Die vorstehenden Ausführungen zu Anspruch 1 des Hilfsantrags 1 gelten in entsprechender Weise für den Vorrichtungsanspruch 19, der sich damit ebenfalls als neu und als auf einer erfinderischen Tätigkeit beruhend erweist. Auch die übrigen angegriffenen Ansprüche nach Hilfsantrag 1 erfüllen die an sie zu stellenden Anforderungen.

III.

Die Kostenentscheidung beruht auf § 84 Abs. 2 PatG i. V. m. § 92 Abs. 1 ZPO. Die ausgerichtete Kostenquote entspricht dem Anteil des Obsiegens und Unterliegens der Parteien. Da der wirtschaftliche Wert, der dem Streitpatent aufgrund des nach Hilfsantrag 1 als schutzfähig verbleibenden Patentgegenstands gegenüber der erteilten Fassung zukommt, nur um einen relativ geringen Teil reduziert ist, ist das Unterliegen der Klägerin mit 90 % und dementsprechend das der Beklagten mit 10 % zu bewerten.

Die Entscheidung über die vorläufige Vollstreckbarkeit folgt aus § 99 Abs. 1 PatG i. V. m. § 709 S. 1 und S. 2 ZPO.

Rechtsmittelbelehrung

Gegen dieses Urteil ist das Rechtsmittel der Berufung gegeben.

Die Berufung ist innerhalb eines Monats nach Zustellung des in vollständiger Form abgefassten Urteils, spätestens aber innerhalb eines Monats nach Ablauf von fünf Monaten nach Verkündung, durch einen in der Bundesrepublik Deutschland zugelassenen Rechtsanwalt oder Patentanwalt als Bevollmächtigten schriftlich oder in elektronischer Form beim Bundesgerichtshof, Herrenstr. 45 a, 76133 Karlsruhe, einzulegen.

Grote-Bittner

Arnoldi

Matter

Söchtig

Dr. Haupt