



BUNDESGERICHTSHOF

IM NAMEN DES VOLKES

URTEIL

X ZR 14/22

Verkündet am:
16. Januar 2024
Anderer
Justizangestellte
als Urkundsbeamtin
der Geschäftsstelle

in der Patentnichtigkeitssache

Der X. Zivilsenat des Bundesgerichtshofs hat auf die mündliche Verhandlung vom 16. Januar 2024 durch den Vorsitzenden Richter Dr. Bacher, den Richter Hoffmann und die Richterinnen Dr. Kober-Dehm, Dr. Marx und Dr. Rombach

für Recht erkannt:

Die Berufung gegen das Urteil des 4. Senats (Nichtigkeitssenats) des Bundespatentgerichts vom 6. Dezember 2021 wird auf Kosten der Klägerin zurückgewiesen.

Von Rechts wegen

Tatbestand:

1 Die Beklagte ist Inhaberin des mit Wirkung für die Bundesrepublik Deutschland erteilten europäischen Patents 2 162 880 (Streitpatents), das am 20. Juni 2008 unter Inanspruchnahme einer US-amerikanischen Priorität vom 22. Juni 2007 angemeldet wurde und die Schätzung der Tonalität eines Schallsignals betrifft.

2 Patentanspruch 1, auf den vier Ansprüche zurückbezogen sind, lautet in der Verfahrenssprache:

A method for estimating a tonality of a sound signal, the method comprising:
calculating a current residual spectrum of the sound signal;
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.

3 Patentanspruch 19, auf den zwei Ansprüche zurückbezogen sind, betrifft sinngemäß eine Vorrichtung zum Ausführen eines solchen Verfahrens.

4 Die Klägerin hat das Streitpatent im Umfang der Patentansprüche 1, 2, 5 und 19 bis 21 angegriffen und geltend gemacht, der angegriffene Gegenstand sei nicht patentfähig und nicht so offenbart, dass ein Fachmann die Erfindung ausführen könne. Die Beklagte hat die angegriffenen Ansprüche in der erteilten Fassung und hilfsweise in drei geänderten Fassungen verteidigt; höchst hilfsweise hat sie die Ansprüche 2, 5, 20 und 21 gesondert verteidigt.

5 Das Patentgericht hat das Streitpatent für nichtig erklärt, soweit der Gegenstand der angegriffenen Ansprüche über die Fassung nach Hilfsantrag 1 hinausgeht, und die Klage im Übrigen abgewiesen.

6 Dagegen richtet sich die Berufung der Klägerin, die ihr erstinstanzliches Klageziel weiterverfolgt. Die Beklagte tritt dem Rechtsmittel mit ihren erstinstanzlichen Anträgen entgegen.

Entscheidungsgründe:

7 Die zulässige Berufung hat in der Sache keinen Erfolg.

8 I. Das Streitpatent betrifft die Schätzung der Tonalität eines Schallsignals.

9 1. Die Streitpatentschrift beschreibt im Stand der Technik bekannte Methoden zur Umwandlung eines Schallsignals in einen digitalen Bitstrom.

10 Das Schallsignal wird hierbei abgetastet und quantisiert. Der abbildbare Frequenzbereich und damit die subjektive Qualität der Aufnahme hängt von der Abtastfrequenz ab (Abs. 3). Die digitalen Abtastwerte werden einer Kodierung unterzogen, um die Anzahl der benötigten Bits zu reduzieren (Abs. 4). Auch damit kann ein Qualitätsverlust einhergehen.

11 Nach der Beschreibung des Streitpatents ermöglicht die Kodiermethode CELP (Code-Excited Linear Prediction) einen guten Kompromiss zwischen subjektiver Qualität und Bitrate. Dabei werde das abgetastete Sprachsignal in aufeinanderfolgenden Blöcken von mehreren Abtastwerten verarbeitet, die in der Regel als Rahmen bezeichnet würden (Abs. 5).

12 Zur weiteren Verbesserung der Systemkapazität diene die quellengesteuerte Sprachkodierung mit variabler Bitrate (VBR). Bei dieser werde ein optimiertes Kodierungsmodell auf der Grundlage der Art des jeweiligen Sprach-Rahmens (z.B. stimmhaft, stimmlos, transient, Hintergrund-Rauschen) verwendet. Der hierbei eingesetzte Algorithmus zur Spracherkennung (Voice Activity Detection, VAD) könne jedoch dazu führen, dass Musik-Signale als stimmlose Signale oder als Hintergrund-Rauschen eingestuft werden, was die Qualität bzw. die Performance des Algorithmus beeinträchtige (Abs. 6).

13 2. Das Streitpatent betrifft vor diesem Hintergrund das technische Problem, die Erkennung von Musik-Signalen zu verbessern.

- 14 3. Zur Lösung schlägt das Streitpatent in der im Berufungsverfahren in erster Linie verteidigten und vom Patentgericht als rechtsbeständig angesehene Fassung von Patentanspruch 1 ein Verfahren vor, dessen Merkmale sich wie folgt gliedern lassen (Änderungen gegenüber der erteilten Fassung sind unterstrichen, abweichende Gliederungsziffern des Patentgerichts in Klammern):

15

1.1	A method for estimating a tonality of a sound signal, the method comprising:	Verfahren zum Schätzen der Tonalität eines Schallsignals, das folgende Schritte umfasst:
1.2	calculating a current residual spectrum of the sound signal	Berechnen eines aktuellen Residualspektrums des Schallsignals
1.2.1 [2.2.3 ^{HI}]	<u>by subtracting a spectral floor from a spectrum of the sound signal in a current frame;</u>	<u>durch Subtrahieren eines spektralen Untergrunds von einem Spektrum des Schallsignals in einem aktuellen Rahmen;</u>
1.3	detecting peaks in the current residual spectrum;	Erkennen von Spitzen im aktuellen Residualspektrum;
1.4	calculating a correlation map between the current residual spectrum and a previous residual spectrum for each detected peak; and	Berechnen einer Korrelationskarte zwischen dem aktuellen Residualspektrum und einem vorherigen Residualspektrum für jede erkannte Spitze; und
1.5	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.	Berechnen einer Langzeit-Korrelationskarte basierend auf der berechneten Korrelationskarte, wobei die Langzeit-Korrelationskarte eine Tonalität im Schallsignal anzeigt.

16 4. Patentanspruch 19 schützt sinngemäß eine Vorrichtung mit Kom-
ponenten, die diese Verfahrensschritte ausführen. Dieser Gegenstand unterliegt
derselben Beurteilung wie Patentanspruch 1.

17 5. Einige Merkmale bedürfen der näheren Erläuterung.

18 a) Das geschützte Verfahren dient dem Zweck, Musiksignale zu er-
kennen, und beurteilt dazu die Tonalität (tonality) eines Schallsignals anhand der
tonalen Stabilität (tonal stability).

19 aa) Mit dem Begriff "Tonalität" bezeichnet das Streitpatent eine Eigen-
schaft von Schallsignalen, die - neben anderen Merkmalen - darauf hindeutet,
dass es sich um ein Musiksignal handelt. Das geschützte Verfahren dient nach
Merkmal 1.1 der Abschätzung, inwieweit ein untersuchtes Signal diese Eigen-
schaft aufweist.

20 Patentanspruch 1 schützt damit ein Verfahren, mit dem Musik von ande-
ren Arten von Schallsignalen, insbesondere von Sprache unterschieden werden,
also als Musik erkannt werden soll. Nicht erfasst sind demgegenüber Verfahren
zur Spracherkennung oder zur Abgrenzung verschiedener Arten von Sprache
wie zum Beispiel stimmhaft und stimmlos.

21 bb) Als Kriterium für diese Schätzung der Tonalität eines Schallsignals
sieht Merkmal 1.5 eine Langzeit-Korrelationskarte auf der Grundlage der in den
Merkmalen 1.2 bis 1.4 festgelegten Berechnungen vor.

22 Eine solche Karte enthält Angaben darüber, inwieweit bestimmte Werte
des Signals über mehrere untersuchte Rahmen hinweg korrelieren. Dies ermög-
licht eine Beurteilung der tonalen Stabilität (tonal stability). Diese wiederum ist
nach der Beschreibung des Streitpatents ein Merkmal für die tonale Natur eines
Musiksignals (Abs. 97 Z. 26 f.: Detection of tonal stability exploits the tonal nature
of music signals).

23 cc) Vor diesem Hintergrund kann dahingestellt bleiben, ob das Streit-
patent die Begriffe "Tonalität" und "tonale Stabilität" als Synonyme verwendet.

24 Selbst wenn dies zu verneinen wäre, ergäbe sich aus Merkmal 1.1 nicht,
dass es über die in den Merkmalen 1.2 bis 1.5 vorgesehenen Berechnungen hin-
aus weiterer Maßnahmen bedarf, um die Tonalität zu beurteilen. Wie Merk-
mal 1.5 ausdrücklich klarstellt, dient vielmehr die Langzeit-Korrelationskarte - die
Auskunft über die tonale Stabilität gibt - als einziges Mittel zur Beurteilung der
Tonalität.

25 Dies schließt nicht aus, dass zusätzliche Kriterien herangezogen werden,
um die Tonalität sicher beurteilen zu können. Zwingend erforderlich sind solche
zusätzlichen Maßnahmen aber nicht.

26 dd) Entgegen der Auffassung der Berufung erfasst Patentanspruch 1
damit beide in der Beschreibung des Streitpatents geschilderten Ausführungs-
beispiele.

27 (1) Die Ausführungen zum ersten Ausführungsbeispiel sind zwar mit
"Tonal stability" überschrieben. Wie bereits oben dargelegt wurde, wird jedoch in
den einleitenden Bemerkungen zu diesem Beispiel darauf hingewiesen, dass
sich die Detektion der tonalen Stabilität die tonale Natur eines Musiksignals
zunutze mache (Abs. 97 Z. 26 f.).

28 Dementsprechend wird in den mit "Estimation of tonality in super wideband
content" überschriebenen Ausführungen zum zweiten Ausführungsbeispiel dar-
gelegt, zum Auffinden von Rahmen mit starkem tonalem Inhalt werde die zuvor
beschriebene Analyse der tonalen Stabilität eingesetzt (Abs. 149 Z. 4-6). Hierbei
werden zwar einige Modifikationen vorgenommen, um den Besonderheiten von
Signalen mit hohen Abstraten Rechnung zu tragen. Dennoch geht es auch in
diesem Zusammenhang darum, die Tonalität anhand der tonalen Stabilität zu be-
urteilen.

29 (2) Der von der Berufung mehrfach hervorgehobene Umstand, dass die US-Anmeldung 60/929336 (K2c), deren Priorität das Streitpatent in Anspruch nimmt, nur das erste Ausführungsbeispiel offenbart und lediglich den Begriff "tonal stability" verwendet, ist in diesem Zusammenhang nicht erheblich.

30 Das Prioritätsdokument könnte für die Auslegung des Streitpatents allenfalls dann von Bedeutung sein, wenn dem Streitpatent Hinweise darauf zu entnehmen wären, dass die der früheren Anmeldung zugrunde liegende Begriffsbildung weiterhin maßgeblich sein soll. Im Streitfall ergibt sich jedoch aus dem bereits erwähnten Umstand, dass das Streitpatent die tonale Stabilität schon im Zusammenhang mit dem ersten Ausführungsbeispiel als Indikator für die tonale Natur eines Signals anführt, dass auch dieses Ausführungsbeispiel das Schätzen der Tonalität im Sinne von Merkmal 1.1 betrifft.

31 (3) Aus demselben Grund ist unerheblich, wie die im Prioritätsdokument und im Streitpatent benannten Erfinder die Begriffe "Tonalität" und "tonale Stabilität" in anderen Veröffentlichungen verwendet haben.

32 b) Ein Residualspektrum im Sinne von Merkmal 1.2 ist ein Spektrum, das aus einem Ausgangsspektrum abgeleitete Werte enthält.

33 aa) Ein Spektrum gibt die Zusammensetzung eines Schallsignals an.

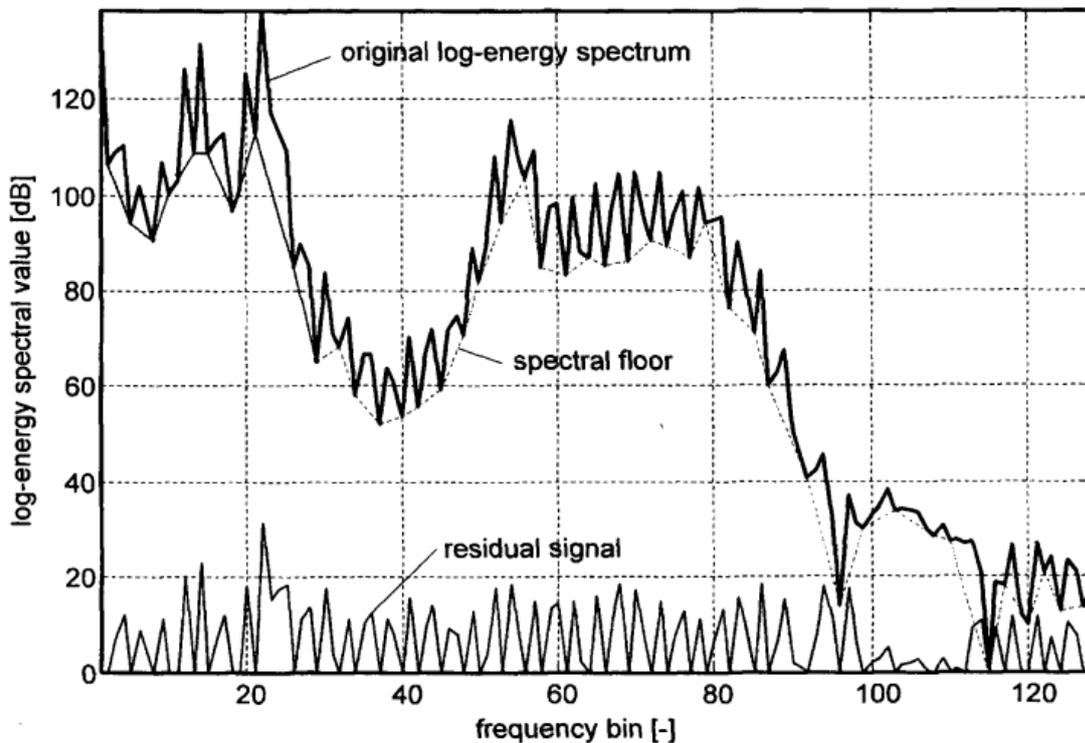
34 Bei dem in der Beschreibung des Streitpatents geschilderten Ausführungsbeispiel wird zur Erzeugung des Ausgangsspektrums jeweils eine Folge von Abtastwerten in den Frequenzbereich transformiert. Aus diesen Werten wird für jede analysierte Frequenz (k) unter anderem ein durchschnittliches log-Energie-Spektrum $E_{db}(k)$ berechnet (Abs. 44 mit Gleichung 4). Dieses wird bei beiden Ausführungsbeispielen als Ausgangsspektrum herangezogen (Abs. 100 und Abs. 150).

35 Patentanspruch 1 lässt offen, in welcher Weise das Ausgangsspektrum erstellt wird.

36 bb) Das Residualspektrum wird gemäß Merkmal 1.2.1 durch Subtrahieren eines spektralen Untergrunds von dem Ausgangsspektrum erzeugt.

37 (1) Beim ersten Ausführungsbeispiel wird der spektrale Untergrund durch eine stückweise lineare Funktion definiert, die durch die ermittelten lokalen Minima des Ausgangsspektrums verläuft.

38 Ein Beispiel für diese Vorgehensweise ist in der nachfolgend wiedergegebenen Figur 3 dargestellt.



39 Die Funktionskurve des spektralen Untergrunds besteht aus geraden Linien, die die Tiefpunkte der Funktionskurve des log-Energie-Spektrums miteinander verbinden. Das Residualspektrum entspricht der Differenz zwischen den Werten des log-Energie-Spektrums und des spektralen Untergrunds.

40 Wie die Berufung zu Recht und unwidersprochen ausführt, kann es bei dieser Vorgehensweise je nach der Gestalt des Ausgangsspektrums vorkommen, dass einzelne Bereiche des Ausgangsspektrums unterhalb des spektralen

Untergrunds verlaufen, so dass sich für das Residualspektrum an den entsprechenden Stellen negative Werte ergeben.

41 (2) Beim zweiten Ausführungsbeispiel wird der spektrale Untergrund durch Filterung des Ausgangsspektrums mit einem gleitenden Mittelwert (moving average filter) erzeugt (Abs. 150-153).

42 Wie die Berufung ebenfalls zutreffend und unwidersprochen darlegt, gibt der so definierte spektrale Untergrund Durchschnittswerte wieder, die teils oberhalb, teils unterhalb der jeweiligen Werte des Ausgangsspektrums liegen. Das durch Subtraktion und nachfolgende Glättung (Abs. 154 f.) ermittelte Residualspektrum weist folglich Bereiche mit negativen Werten auf.

43 (3) Merkmal 1.2.1 lässt offen, in welcher Weise der spektrale Untergrund gebildet wird. Aus der Funktion, die dem Residualspektrum nach Patentanspruch 1 zukommt, ergeben sich insoweit allenfalls rudimentäre Anforderungen.

44 Das Residualspektrum lässt nach den Feststellungen des Patentgerichts die dominierenden Frequenzanteile stärker hervortreten. Daraus ergeben sich besser geeignete Ausgangswerte für die Berechnung der Korrelationswerte. Patentanspruch 1 enthält aber keine Festlegungen hinsichtlich Art oder Umfang solcher Verbesserungen und hinsichtlich der Art und Weise, in der diese gegebenenfalls herbeigeführt werden.

45 cc) Subtrahieren im Sinne von Merkmal 1.2.1 ist eine mathematische Operation, mit der die Werte des Ausgangsspektrums durch die Differenz aus einem Ausgangswert und dem korrespondierenden (Frequenzbin-)Wert des spektralen Untergrunds ersetzt werden.

46 (1) Dies ergibt sich aus der allgemeinen Bedeutung des Begriffs "Subtrahieren" und den damit in Einklang stehenden Ausführungen in der Beschrei-

bung des Streitpatents, laut denen bei beiden Ausführungsbeispielen die jeweiligen Werte der beiden Spektren im mathematischen Sinne voneinander subtrahiert werden (Abs. 103 mit Gleichung 32; Abs. 154).

47 (2) Aus der von der Berufung herangezogenen Veröffentlichung der beiden Erfinder Jelinek und Salami (Noise Reduction Method for Wideband Speech Coding, Proc. Eusipco, 2004, B4) zu der als spectral subtraction bekannten Methode, die im Zusammenhang mit Rauschminderung und spektraler Analyse auch in der Beschreibung des Streitpatents erwähnt wird (Abs. 30 und Abs. 37), ergibt sich kein abweichendes Verständnis.

48 Nach den Ausführungen in B4 kann eine spectral subtraction zwar auch dadurch ausgeführt werden, dass eine spezielle Funktion definiert und mit dem Ausgangsspektrum multipliziert wird. Daraus ergibt sich jedoch lediglich, dass das Ergebnis einer Subtraktion auch auf anderem mathematischem Wege erzielt werden kann, nicht aber die Schlussfolgerung, dass jede beliebige mathematische Funktion als Subtraktion im Sinne von Merkmal 1.2.1 anzusehen ist.

49 Aus der Vorgabe, dass das Berechnen des Residualspektrums durch Subtrahieren zu erfolgen hat, ergibt sich vielmehr, dass die Werte des Residualspektrums der Differenz zwischen dem Wert des Ausgangsspektrums und einem anhand einer geeigneten Funktion bestimmten Operanden entsprechen müssen. Merkmal 1.2.1 mag hierbei je nach Einzelfall auch dann verwirklicht sein, wenn dieselbe Differenz alternativ auch auf anderem mathematischem Wege berechnet werden kann. Nicht ausreichend ist es jedoch, wenn die Operanden, die von den Werten des Residualspektrums subtrahiert werden, dadurch definiert werden, dass zunächst auf anderem Wege die angestrebten Zielwerte definiert und die Operanden dann so bestimmt werden, dass die Subtraktion zu diesen Zielwerten führt.

50 c) Die nach Merkmal 1.3 zu detektierenden Spitzen im Residualspektrum sind gemäß der darauf bezogenen Definition in der Beschreibung die Bereiche zwischen zwei Minima (Abs. 104).

51 aa) Sowohl die Spitzen als auch die sie definierenden Minima müssen
im Residualspektrum vorhanden sein.

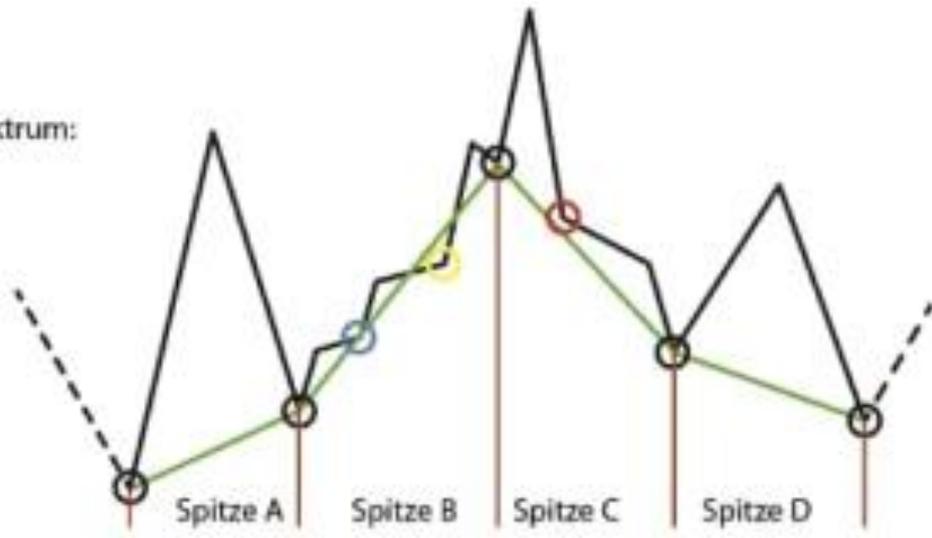
52 Dies ergibt sich ebenfalls aus der bereits erwähnten Definition in der Be-
schreibung. Danach werden Spitzen durch die sie umgebenden Minima im Resi-
dualspektrum definiert (Abs. 104). In Einklang damit werden die Minima auch bei
dem zweiten Ausführungsbeispiel im Residualspektrum gesucht (Abs. 156).

53 bb) Entgegen der Auffassung der Berufung ergibt sich daraus nicht,
dass die Minima im Residualspektrum nicht mit den Minima im Ausgangsspek-
trum übereinstimmen dürfen, wie dies bei dem in Figur 3 des Streitpatents dar-
gestellten Beispiel der Fall ist.

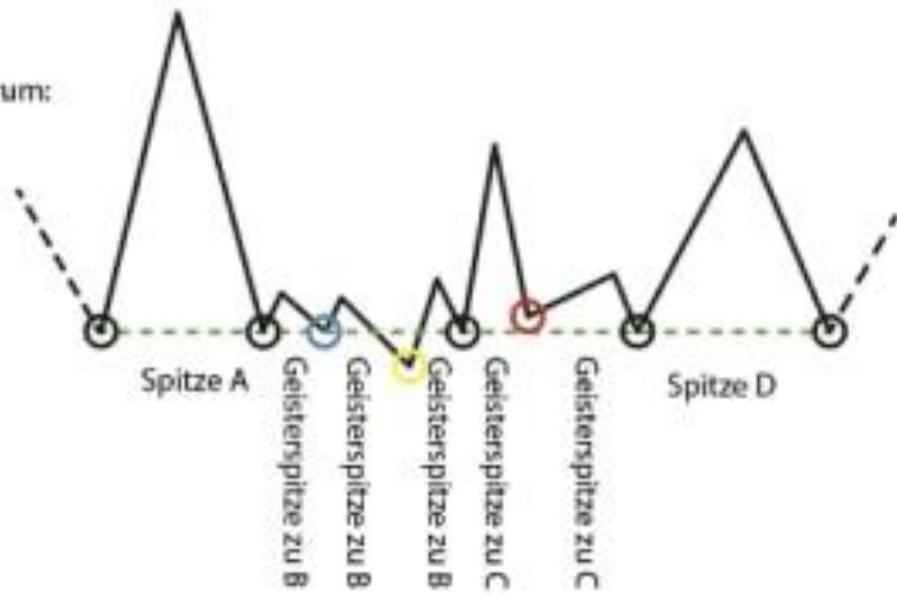
54 Bei einer solchen Ausgestaltung bedarf es zwar keiner erneuten Suche
nach Minima im Residualspektrum, weil schon die zum Berechnen des spektra-
len Untergrunds durchgeführte Suche nach Minima im Ausgangsspektrum das
gewünschte Ergebnis liefert. Dies steht einer Verwirklichung von Merkmal 1.3
aber nicht entgegen, weil dieses Merkmal nicht vorgibt, in welcher Weise und zu
welchem Zeitpunkt die Detektion zu erfolgen hat. Eine Suche nach Minima im
Ausgangssignal reicht danach zur Verwirklichung von Merkmal 1.3 aus, wenn der
spektrale Untergrund so bestimmt wird, dass zu erwarten ist, dass sich die Mi-
nima im Residualspektrum im Wesentlichen an denselben Stellen befinden.

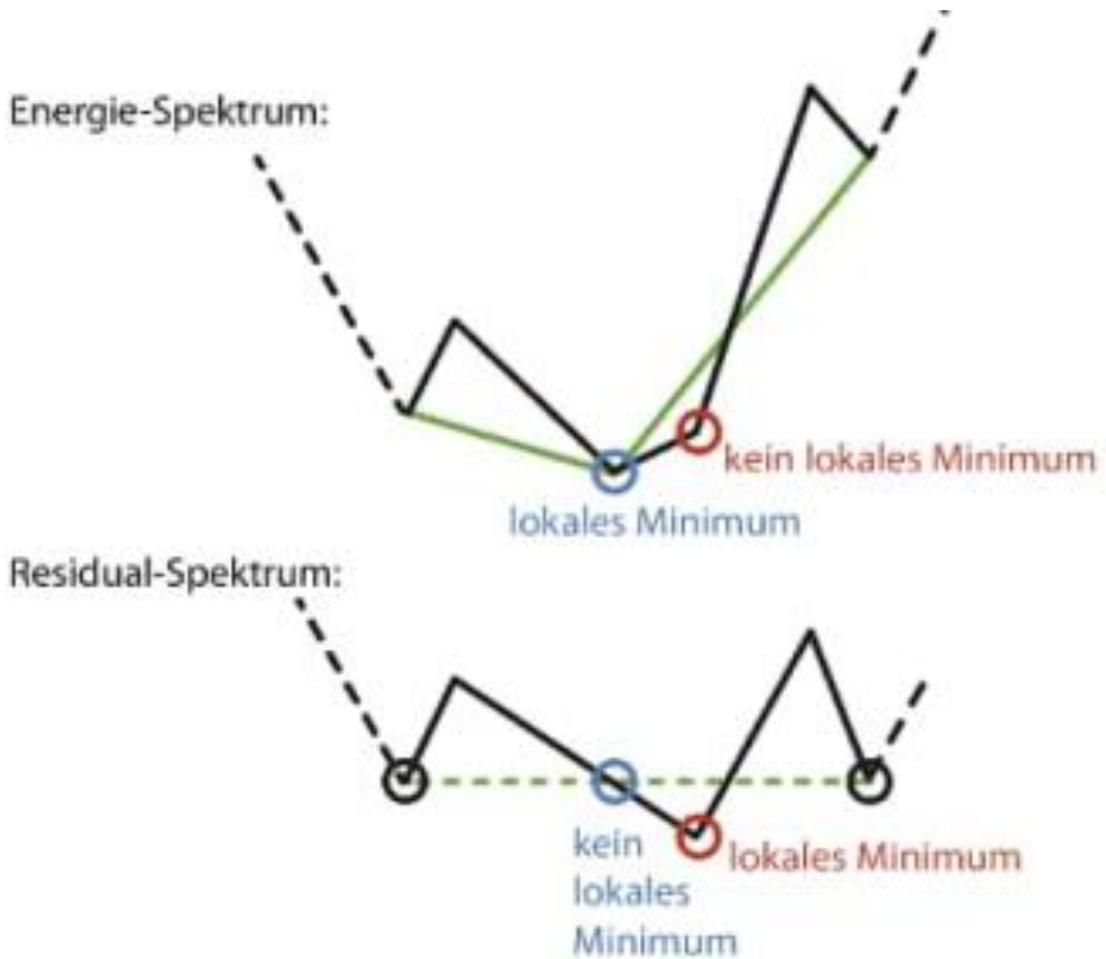
55 Ob die zuletzt genannte Voraussetzung auch dann erfüllt ist, wenn es bei
der Vorgehensweise gemäß dem ersten Ausführungsbeispiel des Streitpatents
aufgrund der besonderen Ausgestaltung des Ausgangssignals in erheblichem
Umfang zu so genannten Geisterspitzen oder zu verschwundenen Minima
kommt, wie dies die Berufung anhand der beiden nachfolgend wiedergegebenen
Skizzen geltend macht, bedarf keiner abschließenden Entscheidung. Die aus-
schließliche Suche im Ausgangsspektrum ist jedenfalls dann ausreichend, wenn
solche Effekte nicht oder allenfalls in zu vernachlässigendem Umfang zu erwar-
ten sind.

Spitzen im Energie-Spektrum:



Spitzen im Residual-Spektrum:





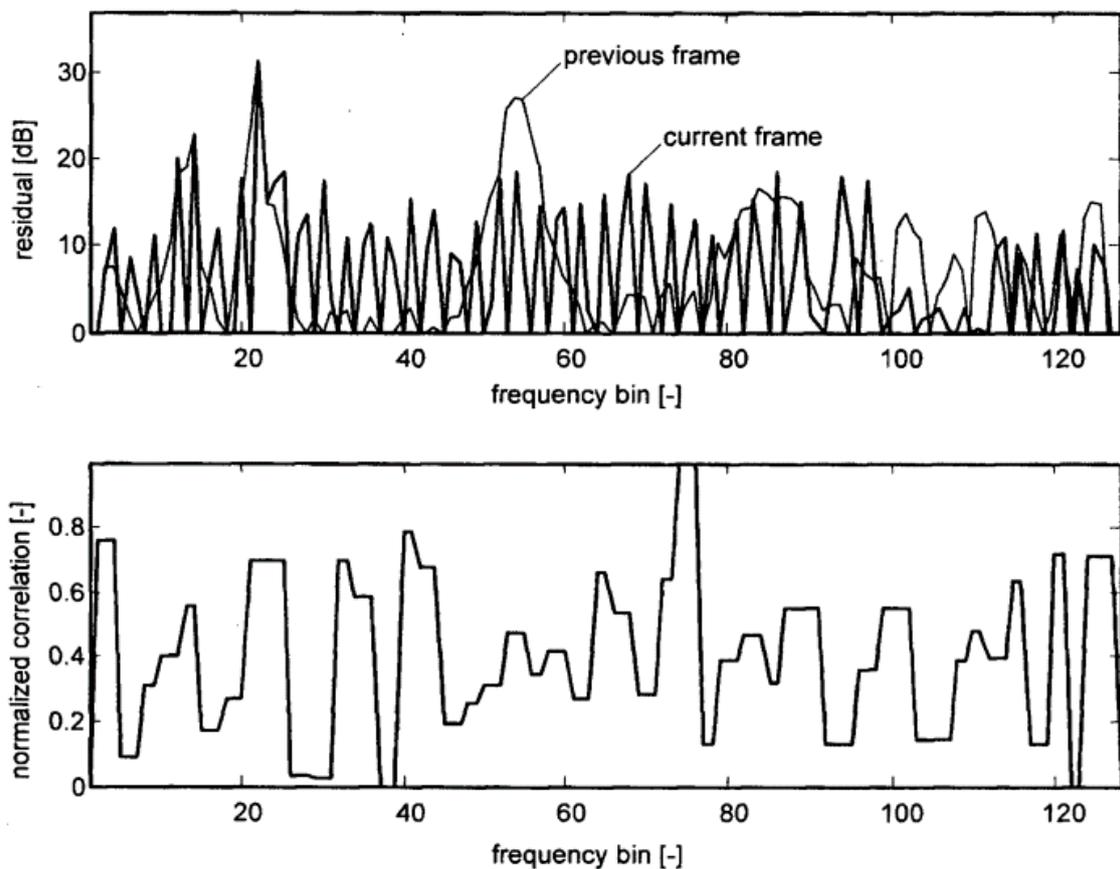
56 cc) Dass die Vorgehensweisen nach den beiden im Streitpatent geschilderten Ausführungsbeispielen zu Unterschieden in Anzahl und Position der im Residualspektrum erkannten Spitzen führen können, führt schon deshalb nicht zu einem anderen Ergebnis, weil Patentanspruch 1 eine diesbezügliche Übereinstimmung weder vorschreibt noch verbietet.

57 Wie bereits im Zusammenhang mit Merkmal 1.2.1 dargelegt wurde, enthält Patentanspruch 1 keine konkreten Mindestanforderungen bezüglich von Art oder Ausmaß der mit der Erzeugung des Residualspektrums erzielten Verbesserung. Angesichts dessen kann ein Residualspektrum nicht schon deshalb als ungeeignet angesehen werden, weil es eine hohe oder geringe Übereinstimmung mit dem Ausgangsspektrum aufweist.

58 d) Eine Korrelationskarte im Sinne von Merkmal 1.4 ist eine Sammlung von Daten, die Auskunft darüber gibt, inwieweit die nach Merkmal 1.3 detektierten Spitzen im Residualspektrum in zwei aufeinanderfolgenden Rahmen übereinstimmen.

59 aa) Nach den Ausführungen im Streitpatent liegt dem die Erkenntnis zugrunde, dass sich bei einem stabilen Signal die Spitzen von Rahmen zu Rahmen nicht wesentlich verschieben und deshalb sowohl ihre Position als auch ihre Form ungefähr gleich sein sollten (Abs. 105).

60 Ein Beispiel für das Residualspektrum von zwei aufeinanderfolgenden Rahmen und eine daraus abgeleitete Korrelationskarte ist in der nachfolgend wiedergegebenen Figur 4 dargestellt.



61 bb) In diesem Zusammenhang bedarf es keiner Entscheidung, ob Patentanspruch 1 festlegt, in welcher Weise die Korrelationswerte errechnet werden, und ob gegebenenfalls eine Korrelation nur in Bezug auf die Position der Spitzen ausreichend ist. Auch auf der Grundlage der weiteren Auslegung erweist sich das Streitpatent aus den unten aufgezeigten Gründen als rechtsbeständig.

62 e) Eine Langzeit-Korrelationskarte im Sinne von Merkmal 1.5 ist eine Sammlung von Daten, die aus den einzelnen Korrelationskarten abgeleitet ist und deshalb Auskunft über die Ähnlichkeit der Spitzen über mehrere aufeinanderfolgende Rahmen hinweg gibt.

63 Bei dem im Streitpatent geschilderten Ausführungsbeispiel enthält die Langzeit-Korrelationskarte für jeden Frequenzbereich einen Langzeit-Wert, der nach der Auswertung jedes Rahmens aktualisiert wird (Abs. 107). Zur Beurteilung der Tonalität wird rahmenweise ein Summenwert gebildet (Abs. 108). Auf dieser Grundlage wird ein Signal als tonal beurteilt, wenn einer der Langzeit-Werte einen fest definierten Schwellenwert überschreitet oder wenn der Summenwert einen laufend angepassten Schwellenwert übersteigt (Abs. 109-112).

64 Entgegen der Auffassung der Beklagten ist danach nicht erforderlich, dass die Korrelationswerte für jeden Rahmen dauerhaft vorgehalten werden. Vielmehr reicht es aus, wenn festgehalten wird, über wie viele Rahmen hinweg sich eine festgestellte Korrelation fortsetzt. Dies geschieht auch bei dem geschilderten Ausführungsbeispiel nicht durch dauerhaftes Vorhalten aller Werte, sondern durch die laufende Aktualisierung des als Langzeit-Korrelationskarte dienenden Summenwerts.

65 II. Das Patentgericht hat seine Entscheidung, soweit für das Berufungsverfahren von Interesse, im Wesentlichen wie folgt begründet:

66 Der Fachmann, ein Ingenieur der Elektro-, Nachrichten- oder Informationstechnik mit einem universitären Master oder Diplom und mehreren Jahren Berufserfahrung auf dem Gebiet der Audiocodierung unter Berücksichtigung von

Sprache und Musik in Schallsignalen, entnehme dem Streitpatent, dass mit Tonalität nichts Anderes gemeint sei als die tonale Stabilität eines Schallsignals. Entgegen der Auffassung der Klägerin zeigten beide Ausführungsbeispiele die Bestimmung der tonalen Stabilität. Dabei würden zwar einzelne Verfahrensschritte geändert, etwa das Bestimmen des spektralen Untergrunds, jedoch würden alle im erteilten Anspruch 1 genannten Verfahrensschritte durchgeführt.

67 Der Gegenstand des erteilten Patentanspruchs 1 sei durch die Dissertation von Hawley (Structure out of Sound, Massachusetts Institute of Technology, 1993, in Auszügen vorgelegt als K3, vom Patentgericht berücksichtigte vollständige Fassung abrufbar unter <http://hdl.handle.net/1721.1/29068> und zitiert als K3') vorweggenommen. Die mit Hilfsantrag 1 verteidigte Fassung erweise sich demgegenüber als rechtsbeständig.

68 Der vom Spektrum des Schallsignals subtrahierte spektrale Untergrund müsse nicht - wie im Ausführungsbeispiel nach Figur 3 bzw. dem erteilten Anspruch 2 - lokal an das Spektrum angepasst sein. Dem Fachmann sei bewusst, dass die Anzahl und die Frequenzen der Spektralwerte des Schallsignals und des spektralen Untergrunds gleich sein müssten, um eine Subtraktion der einzelnen Spektralwerte voneinander zu ermöglichen.

69 Der mit Hilfsantrag 1 verteidigte Gegenstand gehe nicht über den Inhalt der Anmeldung hinaus. Der Fachmann entnehme der Anmeldung als zur Erfindung gehörend, dass ein Residualspektrum durch Subtrahieren eines spektralen Untergrundes von dem Spektrum des Schallsignals gewonnen werde. Die exakte Vorgehensweise zur Berechnung des spektralen Untergrundes sei von untergeordneter Bedeutung. Die Gewinnung des spektralen Untergrundes nach allen Merkmalen des erteilten Anspruchs 2 sei für den Fachmann nur eine von mehreren denkbaren Ausgestaltungen.

70 Die Ansprüche nach Hilfsantrag 1 seien deutlich und knapp gefasst und von der Beschreibung gestützt.

- 71 Es treffe zwar zu, dass das Streitpatent nicht offenbare, wie die Tonalität eines Schallsignals zu schätzen sei. Vielmehr beschäftige es sich ausschließlich mit dem Schätzen der tonalen Stabilität. Da das Streitpatent die beiden Begriffe jedoch synonym im Sinne der tonalen Stabilität verwende, seien die hierzu erforderlichen Verfahrensschritte detailliert beschrieben und ermöglichten dem Fachmann die Nacharbeitung der Erfindung.
- 72 Der mit Hilfsantrag 1 verteidigte Gegenstand sei neu gegenüber K3'. Das Berechnen eines aktuellen Residualspektrums des Schallsignals finde in K3' nicht durch das Subtrahieren eines spektralen Untergrundes von einem Spektrum des Schallsignals in einem aktuellen Rahmen statt.
- 73 Der mit Hilfsantrag 1 verteidigte Gegenstand sei auch durch die Veröffentlichung von Minami et al. (Video Handling with Music and Speech Detection, IEEE MultiMedia, vol. 5, no. 3, S. 17-25, 1998, K4) nicht vorweggenommen. Bei dem in K4 offenbarten Verfahren werde zwar eine Größe (Gesamt-Kanten-Intensität) berechnet, die eine Tonalität im Schallsignal anzeige. Hierbei würden aber weder eine Korrelationskarte noch eine Langzeit-Korrelationskarte eingesetzt. Ebenfalls nicht offenbart sei das Subtrahieren eines spektralen Untergrundes von einem Spektrum des Schallsignals.
- 74 Der mit Hilfsantrag 1 verteidigte Gegenstand beruhe auch auf einer erfinderschen Tätigkeit. Entgegen der Ansicht der Klägerin ergebe sich aus K3' nicht, dass von dem Spektrum des Schallsignals ein spektraler Untergrund abgezogen werde. Die in K3' wiedergegebenen Graustufenbilder zeigten eine deutliche Zunahme der Schwärzung des Hintergrundes zu tieferen Frequenzen. Dies spreche dagegen, dass ein spektraler Untergrund abgezogen worden sei. Zudem dienten diese Bilder lediglich der Illustration der spektralen Unterschiede zwischen Musik- und Sprachsignalen. Der Fachmann könne K3' nicht entnehmen, dass die Bilder in einem Zusammenhang mit der den Musikdetektor realisierenden Soft-

ware stehen. Bei deren Darstellung fänden sich keine Programme oder Programmteile, die das Subtrahieren eines spektralen Untergrunds von dem Spektrum des Schallsignals realisieren würden.

75 Die von der Klägerin angeführten großen Dynamikunterschiede zwischen unterem und oberem Frequenzbereich des Schallsignalspektrums seien ohnehin nicht vorhanden, weil K3' die spektrale Analyse in dem Musikfilter auf den Frequenzbereich zwischen 150 und 1000 Hz beschränke. Ferner erkenne der Fachmann, dass die Anzahl der in K3' durchgeführten Vergleiche mit Nachbarwerten so groß sei, dass kleine rauschartige Spitzen nicht versehentlich als lokale Maxima erkannt würden, und zugleich so klein, dass lokale Maxima auch dann noch erkannt würden, wenn die Grobstruktur des Spektrums in dem relevanten Bereich relativ steil ansteige oder abfalle.

76 K3' präsentiere damit eine in sich geschlossene und gut funktionierende Lösung für das Berechnen eines aktuellen Residualspektrums des Schallsignals und liefere dem Fachmann somit keinen Anlass, eine "Begradigung" des Schallsignalspektrums durch Subtrahieren eines spektralen Untergrundes in Betracht zu ziehen.

77 Auch K4 vermöge dem Fachmann keine Hinweise in diese Richtung zu geben.

78 III. Diese Beurteilung hält der Überprüfung im Berufungsrechtszug stand.

79 1. Der mit Hilfsantrag 1 verteidigte Gegenstand ist in den ursprünglich eingereichten Unterlagen als zur Erfindung gehörend offenbart.

80 a) Entgegen der Auffassung der Berufung ergibt sich aus der Anmeldung kein zwingender technischer Zusammenhang zwischen dem Subtrahieren im Sinne von Merkmal 1.2.1 und den im ersten Ausführungsbeispiel sowie in Patentanspruch 2 zusätzlich vorgesehenen Merkmalen der Suche nach Minima

im Ausgangsspektrum und der Bestimmung des spektralen Untergrunds durch Verbinden dieser Minima.

81 In den ursprünglichen Unterlagen, deren Inhalt mit der Offenlegungsschrift (WO 2009/000073 A1, K2b) übereinstimmt, wird zwar ebenso wie in der Patentschrift ein erstes Ausführungsbeispiel geschildert, bei dem der subtrahierte spektrale Untergrund durch Verbindung der Minima des Ausgangsspektrums gebildet wird (S. 32 f.).

82 Ebenfalls wie in der Patentschrift sieht die Anmeldung das Subtrahieren eines spektralen Untergrunds aber auch bei einem zweiten Ausführungsbeispiel vor, bei dem der spektrale Untergrund auf andere Weise berechnet wird (S. 48 f.).

83 Damit ergibt sich bereits aus der Anmeldung, dass das Berechnen des Residualspektrums durch Subtrahieren eines spektralen Untergrunds vom Ausgangsspektrum nicht von der Art und Weise abhängt, in der der spektrale Untergrund bestimmt worden ist.

84 b) Entgegen der Auffassung der Berufung ist der Anmeldung vor diesem Hintergrund auch nicht zu entnehmen, dass ein Subtrahieren im Sinne von Merkmal 1.2.1 zwingend voraussetzt, dass der spektrale Untergrund nach einer der beiden in der Beschreibung dargestellten Methoden bestimmt wird.

85 Aus der Schilderung der beiden Ausführungsbeispiele und dem in den Erläuterungen zum zweiten Ausführungsbeispiel enthaltenen Hinweis, dass der spektrale Untergrund, der vom log-Energie-Spektrum abgezogen wird, unter Umständen an Besonderheiten des zu kodierenden Signals angepasst werden muss (S. 48 Z. 10-15), ergibt sich vielmehr hinreichend deutlich, dass ein Subtrahieren grundsätzlich mit jedem geeigneten spektralen Untergrund in Frage kommt.

86 2. Im Ergebnis zutreffend hat das Patentgericht entschieden, dass die Erfindung so offenbart ist, dass ein Fachmann sie ausführen kann.

87 Die dagegen gerichteten Angriffe der Berufung beruhen auf der Prämisse,
dass die Begriffe "Tonalität" und "tonale Stabilität" unterschiedliche Bedeutung
haben und die Ausführungen zum ersten Ausführungsbeispiel nicht erkennen
lassen, wie neben der tonalen Stabilität auch die Tonalität beurteilt werden soll.

88 Diese Prämisse ist unzutreffend.

89 Wie bereits oben dargelegt wurde, kann dabei offenbleiben, ob das Streit-
patent die beiden genannten Begriffe als synonym behandelt. Ausschlaggebend
ist auch insoweit der Umstand, dass die Tonalität gemäß Patentanspruch 1 je-
denfalls allein anhand der in den Merkmalen 1.2 bis 1.5 vorgesehenen Maßnah-
men zur Beurteilung der tonalen Stabilität geschätzt werden kann. Diese Maß-
nahmen sind ausführbar offenbart.

90 3. Zu Recht hat das Patentgericht die mit Hilfsantrag 1 verteidigte An-
spruchsfassung als hinreichend klar angesehen.

91 Entgegen der Auffassung der Berufung ist Merkmal 1.2.1 nicht deshalb
unklar, weil es unabhängig vom Kontext des ersten Ausführungsbeispiels und
des Patentanspruchs 2 beansprucht wird und es damit dem Fachmann überlas-
sen bleibt, wie er den spektralen Untergrund bestimmt.

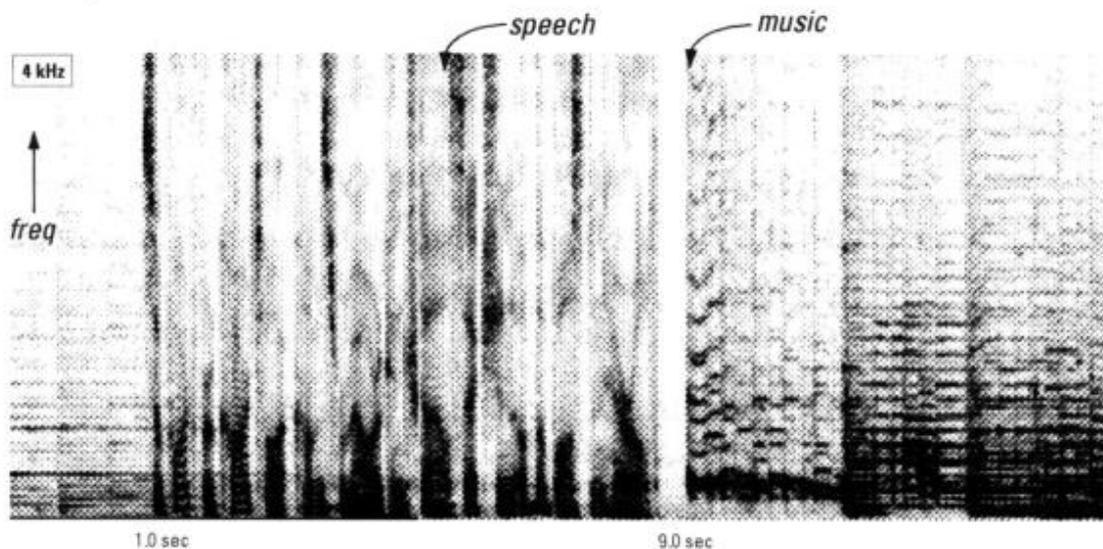
92 Für die Frage der Klarheit ist allein entscheidend, ob auch unabhängig von
dem genannten Kontext beurteilt werden kann, ob ein Subtrahieren im Sinne von
Merkmal 1.2.1 vorliegt. Diese Frage ist aus den bereits im Zusammenhang mit
der Auslegung dieses Merkmals angeführten Gründen zu bejahen. Dass die Be-
urteilung, ob dieses Merkmal verwirklicht ist, in Bezug auf bestimmte Ausfüh-
rungsformen mit rechtlichen Schwierigkeiten verbunden sein kann, macht den
Anspruch nicht unklar.

93 4. Der mit Hilfsantrag 1 verteidigte Gegenstand ist patentfähig.

94 a) K3' nimmt diesen Gegenstand nicht vollständig vorweg.

95 aa) K3' befasst sich mit dem maschinellen Erkennen von Strukturen in Daten, die Schallsignale repräsentieren. Wünschenswert sei hierfür insbesondere ein Detektor, der Musik von anderen Schallsignalen unterscheiden kann. K3' schlägt hierzu vor, nach akustischen Eigenschaften zu suchen, die für Musik charakteristisch sind (S. 78).

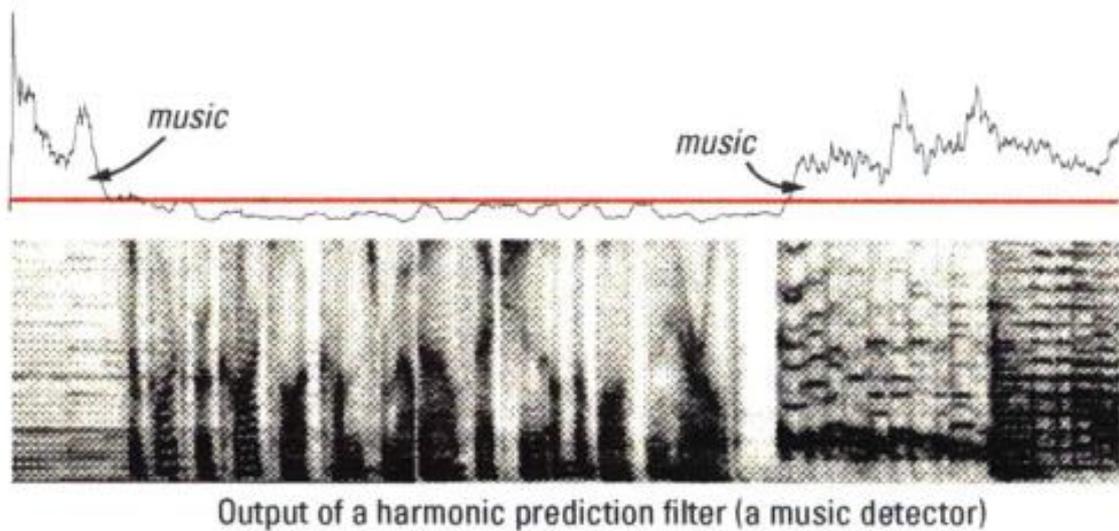
96 (1) Zur Vereinfachung des Problems könne die Betrachtung auf Musik beschränkt werden, die Tonhöhen (pitches) habe, nämlich Noten mit Frequenzen, die für eine bestimmte Zeitspanne relativ stabil blieben. Dies könne anhand eines Spektrums beurteilt werden, wie es in der nachfolgend wiedergegebenen Abbildung dargestellt ist (S. 79).



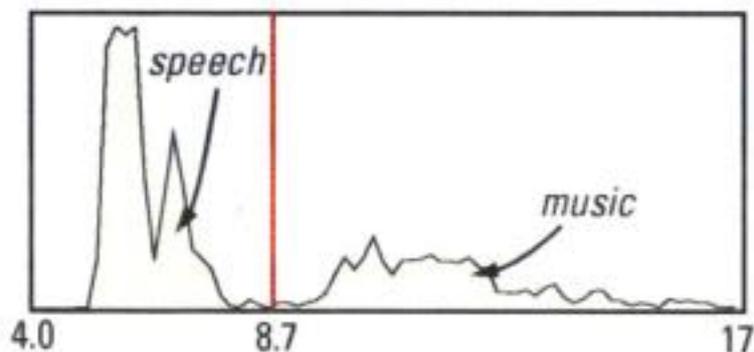
97 Wenn Musik vorhanden sei, seien Käme von Obertönen mit relativ festen Frequenzen zu sehen, die dem Spektrum ein gestreiftes Aussehen verliehen. Sprache enthalte ebenfalls Käme von Obertönen; deren Tonhöhe ändere sich aber relativ schnell (S. 80 Abs. 1).

98 Eine Möglichkeit, dies zu messen, sei ein Filter, der die "Streifenbildung" erkenne, indem er ein Analysefenster über das Spektrum schiebe, die Spitzen in

der Frequenzmagnitude erfasse und einen laufend aktualisierten Zähler für die durchschnittliche Dauer einer stationären Frequenzspitze führe. Dies spiegele die Vorhersehbarkeit einer Spitze oder umgekehrt deren Entropie. Wenn die Vorhersehbarkeit der Spitze einen Schwellenwert überschreite, sei es wahrscheinlich, dass es sich um Musik handle. Dieser Zusammenhang wird in der nachfolgend wiedergegebenen Abbildung dargestellt (S. 80 Abs. 2).



99 Die durchgeführten Studien hätten ergeben, dass die Unterscheidung zwischen Musik und Sprache häufig so deutlich sei, dass ein Histogramm des Prädiktionsfilters die beiden Zustände klar offenbare. Dies wird anhand der nachfolgend wiedergegebenen Abbildung verdeutlicht (S. 80 Abs. 3).



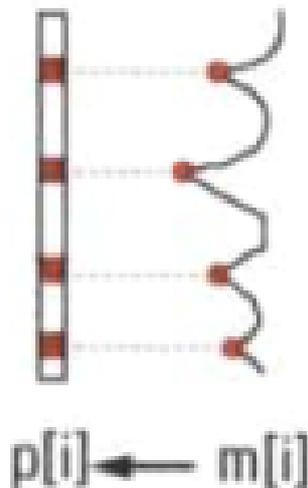
100 Die horizontale Achse dieser Abbildung zeige die durchschnittliche Länge einer Frequenzspitze in Rahmen. Der dargestellte Schwellenwert von

8,7 Rahmen entspreche einem Zeitraum von 0,05 Sekunden bei einer Abtastrate von 8 kHz und einer Rahmenlänge von 48 Abtastwerten (S. 80/81).

101 (2) Auf dieser Grundlage schlägt K3' einen Detektor mit einem Prädiktionsfilter vor, der auf sequentielle Rahmen angewendet werde, wie sie aus der Fast-Fourier-Analyse bekannt seien. Dies könne mit einfachen Mitteln realisiert werden.

102 Der Filter wirke wie ein rechteckiges Fenster, das über das Abbild des Spektrums gleite und einen gleitenden Durchschnittswert für die Dauer einer Spitze festhalte. Mit Erreichen eines neuen Rahmens würden der älteste Rahmen entfernt, der neue Rahmen eingefügt, die Spitzenwerte aktualisiert und die neue Schätzung der Spitzenwertvorhersehbarkeit ausgegeben.

103 Um einen neuen Rahmen einzufügen, werde sein anhand der diskreten Fourier-Terme für die einzelnen Frequenzen i aus Einzelwerten $m[i]$ bestehendes Amplitudenspektrum berechnet. In diesem Spektrum würden die Spitzen lokalisiert, um einen Spitzenrahmen zu erhalten. Dieser enthalte für jede Frequenz eine Variable $p[i]$, die den Wert 1 aufweise, wenn es sich um ein lokales Maximum handle, und anderenfalls den Wert 0. Letzteres ist in der nachfolgend wiedergegebenen Abbildung dargestellt (S. 81).



- 104 Die Spitzenrahmen würden in einem Analysefenster beibehalten. Dieses sei als modulo-indizierter Verzögerungspuffer P implementiert, in dem $P(0)$ den aktuellen Spitzenrahmen angebe, $P(t)$ den t -ten vorhergehenden Spitzenrahmen, und $P(t)[i]$ wahr sei, wenn das i -te Bin des t -ten vorhergehenden Rahmens eine Spitze enthalten habe. Dadurch werde das ursprüngliche Amplitudenfenster auf ein binäres Spitzenfenster reduziert. Auf dieser Grundlage werde die durchschnittliche Länge einer aufeinanderfolgenden Reihe von Spitzen ermittelt, wobei einige zusätzliche Vereinfachungen vorgenommen würden (S. 82).
- 105 (3) Ein in der Programmiersprache C erstellter Quellcode für einen solchen Musikfilter ist in dem nachfolgend wiedergegebenen Listing dargestellt (S. 82).

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

    // 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); // m=magnitude data; h=hartley data
    findPeaks(P(0),m,n2); // P(0)[i] = 1 if m[i] is a peak

    // 4-pt smoother; emphasize sustained peaks: 1101=>1111
    smoothPeaks();

    // 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++;
}
```

106 Nach den ergänzenden Feststellungen des Patentgerichts ist dem in einem Anhang (S. 176) wiedergegebenen Programmcode zu entnehmen, dass die in diesem Code aufgerufene Funktion findPeaks ein lokales Maximum dann annimmt, wenn der untersuchte Spektralwert größer ist als die beiden jeweils vorangehenden und nachfolgenden Nachbarwerte.

107 Der nachfolgende, durch die Kommentarzeile "//merge incoming column of peaks" eingeleitete Code dient der Aktualisierung des Puffers P. Dieser Code vergleicht für jede untersuchte Frequenz die Puffereinträge für den aktuellen Rahmen P(3) mit den kumulierten Puffereinträgen der vorangegangenen Rahmen P(4). Wenn beide Rahmen einen von 0 verschiedenen Wert aufweisen und

eine vordefinierte Höchstgrenze (maxRun) noch nicht überschritten ist, werden die beiden Werte addiert und im Puffereintrag für den aktuellen Rahmen gespeichert. Wenn diese Bedingung nicht erfüllt ist, die zuvor gefundene Anzahl von aufeinander folgenden Spitzen aber einen vordefinierten Mindestwert (minRun) überschritten hat, werden die Zahl der gefundenen Spitzenläufe (peakN) um eins und die Gesamtdauer der gefundenen Spitzen (peakT) um die aktuell gefundene Dauer erhöht. Nach dem Ende dieser Schleife wird der gleitende Mittelwert (peakT/peakN) angepasst.

108 bb) Damit ist Merkmal 1.1 unmittelbar und eindeutig offenbart.

109 cc) Zugunsten der Berufung kann unterstellt werden, dass die Merkmale 1.2 bis 1.4 und das Merkmal 1.5 aus den vom Patentgericht genannten Gründen ebenfalls offenbart sind, dass also die in K3' offenbarten Spitzenrahmen, die pro Frequenz nur zwei unterschiedliche Werte (1 oder 0) enthalten können, ein Residualspektrum im Sinne dieses Merkmals darstellen und dass es ausreicht, anhand dieser Werte die Position dieser Spitzen und darauf bezogene Korrelationswerte zu ermitteln.

110 dd) Zu Recht hat das Patentgericht entschieden, dass das Subtrahieren eines lokalen Untergrundes im Sinne von Merkmal 1.2.1 in K3' nicht offenbart ist.

111 (1) Entgegen der Auffassung der Berufung erfolgt die in K3' offenbarte Berechnung von Spitzenrahmen (findPeaks) nicht durch Subtrahieren.

112 Wie bereits dargelegt wurde, kann zugunsten der Berufung unterstellt werden, dass die in K3' offenbarten Spitzenrahmen ein Residualspektrum im Sinne von Merkmal 1.2 darstellen. Dieses wird aber jedenfalls nicht durch Subtrahieren eines spektralen Untergrundes ermittelt, sondern durch einen Vergleich von Werten benachbarter Rahmen.

113 Dass dasselbe Ergebnis bei bekannten Zielwerten auch durch Subtrahieren erzielt werden könnte, indem man den zu subtrahierenden Wert anhand des auf anderem Wege gefundenen Zielwerts bestimmt, reicht, wie bei der Auslegung von Merkmal 1.2.1 aufgezeigt wurde, nicht aus.

114 (2) Entgegen der Auffassung der Berufung offenbaren die in K3' wiedergegebenen Graustufenbilder ebenfalls kein Subtrahieren eines spektralen Untergrunds.

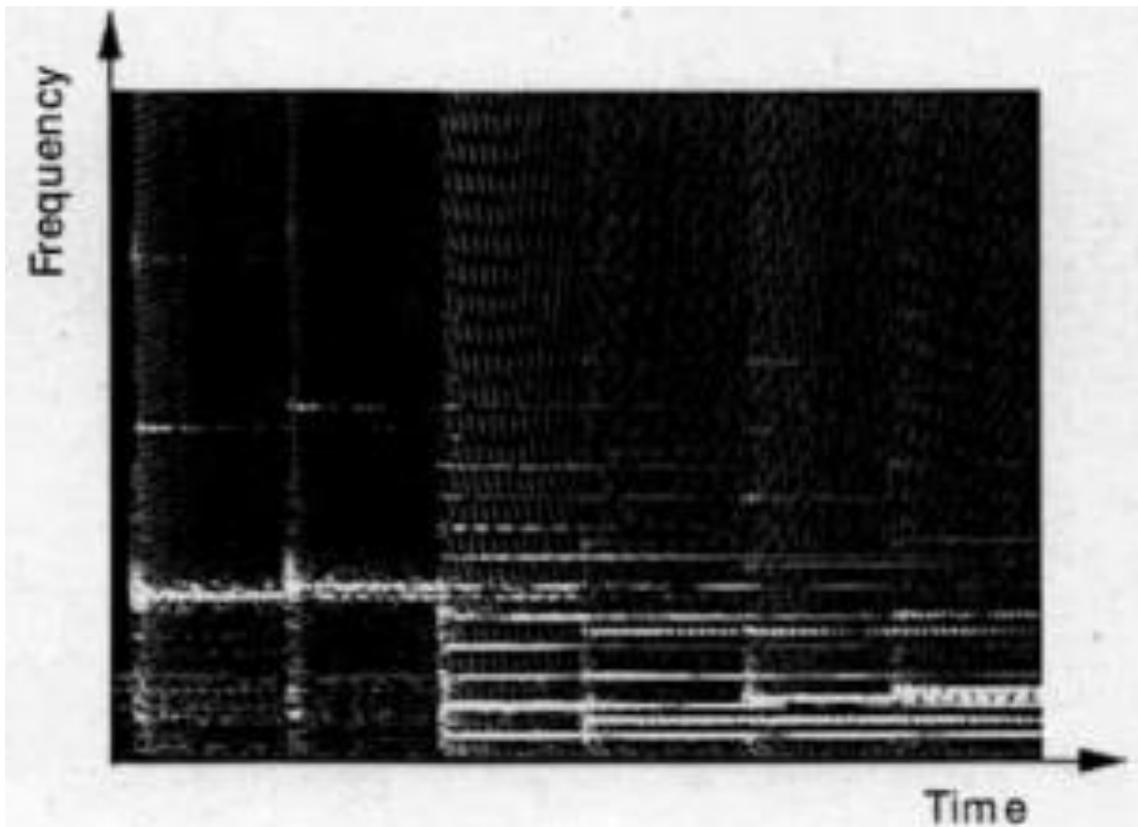
115 In diesem Zusammenhang kann dahingestellt bleiben, ob die Abbildungen darauf hindeuten, dass die Spektren in irgendeiner Weise vorbehandelt worden sind, weil sie eine von der Darstellung in Figur 3 des Streitpatents abweichende Grobstruktur zeigen. Selbst wenn dies zuträfe, ergäbe sich dieses Erkenntnis nicht unmittelbar und eindeutig aus der Darstellung in K3', sondern allenfalls aufgrund ergänzender fachlicher Überlegungen. Unabhängig davon ergäbe sich aus dem Erkenntnis, dass eine Vorbehandlung stattgefunden hat, nicht unmittelbar und eindeutig, dass diese durch Subtrahieren eines spektralen Untergrunds erfolgt ist.

116 b) K4 nimmt den mit Hilfsantrag 1 verteidigten Gegenstand ebenfalls nicht vorweg.

117 aa) K4 befasst sich mit der Strukturierung von Videomaterial, um dem Benutzer einen schnelleren Zugang zu bestimmten Szenen zu ermöglichen.

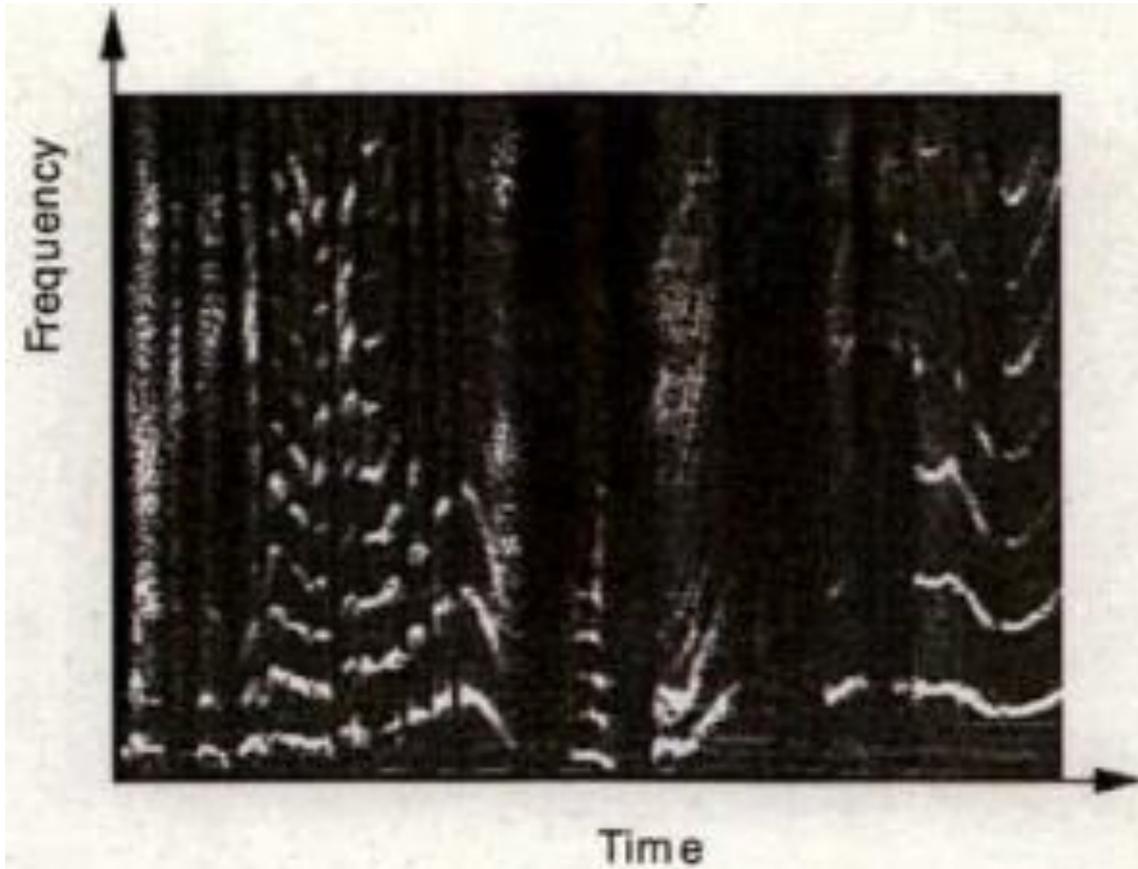
118 (1) Nach den Ausführungen in K4 haben bekannte Vorgehensweisen hierzu vor allem die Bildinformationen herangezogen. K4 schlägt vor, ergänzend die Audioinformationen zu nutzen (S. 17 rechts Abs. 2). Da hierbei mit gemischten Schallsignalen umzugehen sei, seien bestehende Algorithmen zur Spracherkennung nicht ohne weiteres geeignet (S. 18 links Abs. 2).

- 119 Aus der Sicht der Videoproduktion ließen sich Audioinformationen in drei Klassen einteilen: Musik, Sprache und Soundeffekte (S. 19 links Abs. 2). Deshalb sei es notwendig, zwischen diesen verschiedenen Audioelementen zu unterscheiden (S. 19 rechts Abs. 3).
- 120 (2) Als Grundidee für die Erkennung von Musik zieht K4 die in K3' vorgeschlagene Methode heran, bei der stabile Spitzen im Leistungsspektrum als Musikmerkmal betrachtet würden.
- 121 K4 schlägt ergänzend vor, zur Detektion der Spitzen einen Algorithmus zur Kantenerkennung einzusetzen. Damit könnten Musiksegmente innerhalb eines Soundtracks und Musikkomponenten innerhalb eines Tonspektrums lokalisiert werden. Die Bestimmung der Position der Musikkomponenten könne auch dazu beitragen, eine irrtümliche Spracherkennung zu verhindern.
- 122 Ein Beispiel für das Tonspektrum eines typischen Musiksegments sei in der nachfolgend wiedergegebenen Figur 2 dargestellt. Darin stelle die Helligkeit des Bildes die Leistung des Spektrums dar (S. 19 rechts unten).



123 Aus der Abbildung ergebe sich, dass die Spitzen des Spektrums in Frequenzrichtung stabil seien, wenn Musik vorhanden sei. Daher könne man Musik erkennen, indem die stabilen Spitzen extrahiert würden und deren Dauer berechnet werde (S. 19/20).

124 Die nachfolgend wiedergegebene Figur 3 zeige ein Schallspektrum von Sprache.



125 Das auffälligste Merkmal, die gleichmäßig verteilten Streifen, ergebe sich aus der harmonischen Struktur des stimmhaften Klangs. Eine Reihe von Streifen in Zeitrichtung schein eindeutig mit der Äußerung von Vokalen übereinzustimmen, so dass durch die Erkennung dieser Streifen festgestellt werden könne, ob das Schallsegment Sprache enthalte (S. 20 links Abs. 1 f.).

126 Der vorgeschlagene Algorithmus zur Musikerkennung berechne zunächst das Leistungsspektrum mit Hilfe der schnellen Fourier-Transformation (FFT) (S. 20 links Abs. 1). Er behandle das Schallspekrogramm als Graustufenbild und berechne die Lage der Peaks durch Anwendung eines Operators zur Kantenerkennung (S. 20 links Abs. 3).

127 Die Randintensität der Spektralspitzen könne erhalten werden, indem man einfach die Differenz der Intensität benachbarter Pixel in der Frequenzrichtung nehme. Die Kantenintensität für einen diskreten Frequenzwert werde entlang der

Zeitachse für eine Sekunde aufsummiert. Ihr absoluter Wert, aufaddiert in der Frequenzrichtung, ergebe die gesamte Kantenintensität (S. 20 rechts Abs. 1).

128 Bei Musik neigten die Spitzen des Spektrums dazu, sich bei bestimmten Frequenzen einzupendeln und für eine gewisse Zeit anzuhalten, so dass die Gesamtkantenintensität hoch werde. Wenn sie einen Schwellenwert überschreite, werde angenommen, dass der Erkennungsrahmen Musik enthalte.

129 In der nachfolgend wiedergegebenen Figur 4 zeige der Abschnitt c ein Beispiel einer Gesamtkantenintensität (ED), das mit dieser Methode berechnet worden sei (S. 20 rechts Abs. 1 f.).

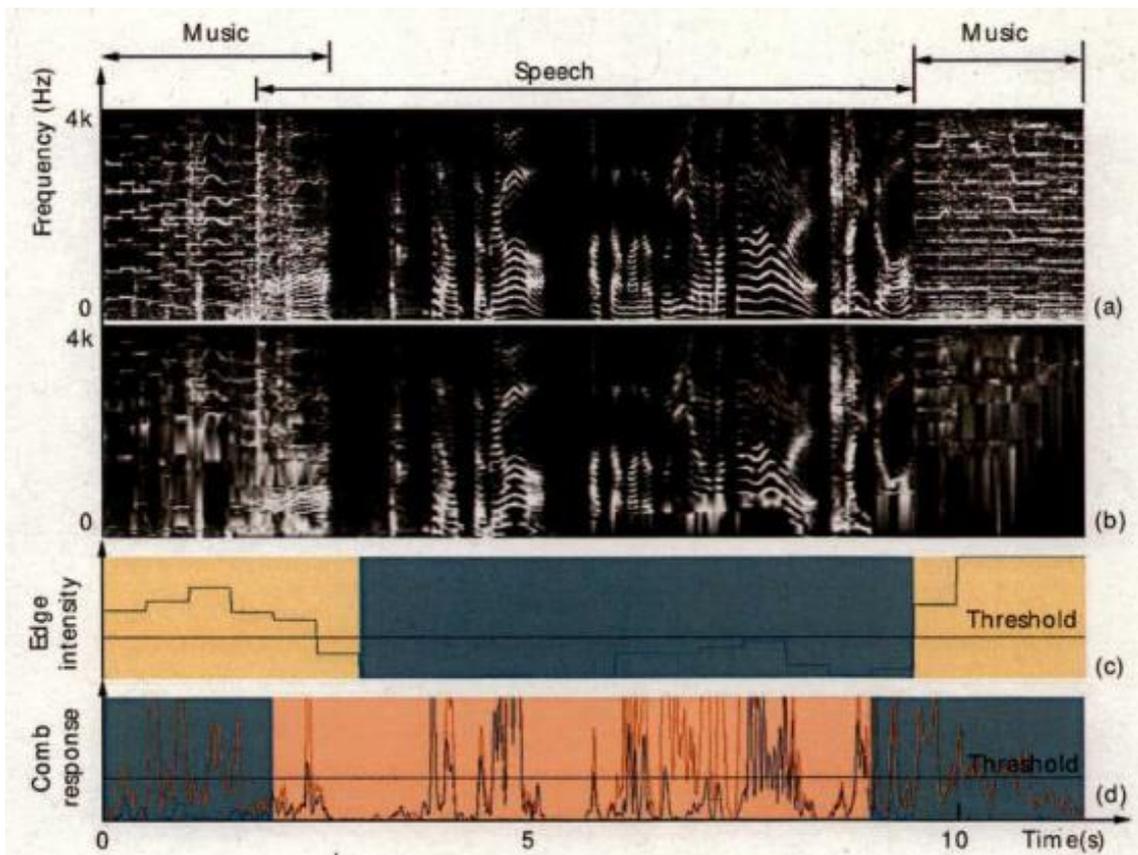


Figure 4. Detection results: (a) sound spectrogram, (b) after the removal of music components, (c) result of music detection, and (d) result of speech detection.

130 Dieses Beispiel, das aus einem Fernsehspiel stamme, enthalte sowohl Musik als auch Sprache. Mehrere Abschnitte des Diagramms überschritten die Schwellenwerte nicht, insbesondere wenn die Musik durch Sprache überdeckt werde. Die vereinzelt Momente, in denen die Gesamtkantenintensität (ED) unter den Schwellenwert falle, könnten jedoch leicht durch die Werte vor und nach diesen Instanzen interpoliert werden (S. 20 rechts Abs. 2).

131 Vor der Erkennung von Sprache entferne die vorgeschlagene Methode Musikkomponenten. Obwohl sich Dauer und Stabilität der Spektralspitzen in Frequenzrichtung von denen des stimmhaften Klangs unterschieden, habe auch Musik eine harmonische Struktur, die oft fälschlicherweise als Sprache erkannt werde. Musikkomponenten würden mit der oben beschriebenen Methode lokalisiert. Wenn die summierte Kantenintensität eines diskreten Frequenzwerts einen bestimmten Schwellwert überschreite, würden die Pixel bei dieser Frequenz innerhalb eines Erkennungsrahmens auf 0 gesetzt. Um zu verhindern, dass im Spektrogramm Lücken entstehen, interpoliere die vorgeschlagene Methode die entfernten Teile mit Hilfe der benachbarten Spektralleistung in Frequenzrichtung linear. Anschließend werde eine Glättung in der Zeitrichtung durchgeführt (S. 20 rechts Abs. 3).

132 Abschnitt b in Figur 4 zeige das Schallspektrogramm aus Abschnitt a nach dem Entfernen der Musikkomponente. Ein Kammfilter erkenne die Obertöne. Die gleichmäßig verteilten Kammzinken dehnten und verschöben sich in Frequenzrichtung und der Filter summiere die Spektralwerte an jeder Zinke (S. 20 rechts

unten). Die Kammreaktion eines einzelnen Linienspektrums sei durch den Maximalwert an einem bestimmten Ort und einer bestimmten Position des Kammes gegeben. Wenn die Leistung des Spektrums auf halber Strecke zwischen den Zinken subtrahiert werde, werde eine übermäßige Kammreaktion auf das Breitbandrauschen vermieden. Wenn die Kammreaktion den Schwellenwert überschreite, werde das Klangsegment als stimmhafter Klang erkannt (S. 21 links oben).

133 bb) Damit ist Merkmal 1.1 offenbart.

134 cc) Ebenfalls offenbart sind die Merkmale 1.2 und 1.3.

135 Nach den Feststellungen des Patentgerichts wird die in K4 mittels eines Kantendetektionsoperators erfolgende Suche nach Spitzen in einem Residualspektrum durchgeführt.

136 dd) Nicht offenbart ist das Subtrahieren eines spektralen Untergrundes gemäß Merkmal 1.2.1.

137 (1) Den in K4 wiedergegebenen Graustufenbildern lässt sich aus den bereits im Zusammenhang mit K3' dargelegten Gründen nicht unmittelbar und eindeutig entnehmen, dass die darin gezeigten Spektren durch Subtrahieren eines spektralen Untergrundes vorbehandelt worden sind. Dem Text von K4 lassen sich diesbezüglich ebenfalls keine Hinweise entnehmen.

138 (2) Die von der Berufung angeführten Passagen zur Spracherkennung sind, wie die Berufungserwiderung zu Recht geltend macht, insoweit schon deshalb nicht von Bedeutung, weil sie sich nicht auf die Schätzung von Tonalität im Sinne von Merkmal 1.1 beziehen.

139 K4 befasst sich zwar - ebenso wie K3' - mit der Erkennung von Musik und von verschiedenen Formen von Sprache. K4 und K3' sehen für die beiden Aufgaben aber unterschiedliche Vorgehensweisen vor.

140 Wie bereits oben aufgezeigt wurde, werden in K4 die in einem ersten Schritt als Musik erkannten Passagen vor der anschließenden Spracherkennung entfernt. Der Algorithmus zur Spracherkennung dient damit nicht der Erkennung von Musik und folglich nicht der Schätzung von Tonalität im Sinne von Merkmal 1.1. Entsprechendes gilt für die ähnlichen Passagen in K3', die sich mit einem gesplitteten Kamm befassen (S. 113), auf die sich der (keine Seitenangabe enthaltende) Verweis in K4 bezieht.

141 Unmittelbare und eindeutige Hinweise darauf, dass die von der Berufung angeführten Merkmale des Verfahrens zur Spracherkennung auch zur Erkennung von Musik eingesetzt werden können, finden sich in K4 und K3' nicht.

142 ee) Die Merkmale 1.4 und 1.5 sind ebenfalls nicht offenbart.

143 (1) Wie das Patentgericht zutreffend ausgeführt hat, führt die im Rahmen der Musikerkennung durchgeführte Summation der Kanten-Intensitäten nicht zum Berechnen einer Korrelationskarte oder einer Langzeit-Korrelationskarte.

144 (2) Die - ohnehin nicht relevanten - Ausführungen zur Spracherkennung offenbaren entgegen der Auffassung der Berufung ebenfalls nicht das Berechnen solcher Karten.

145 Nach den Ausführungen in K4 werden die Spektralwerte aller Zinken für einen Zeitausschnitt summiert. Damit ist, wie die Berufungserwiderung zu Recht geltend macht, nur eine Summierung über die Frequenz offenbart.

146 Die Formeln, anhand derer die Berufung darzulegen versucht, dass auch eine Summierung über die Zeit hinweg erfolgt, finden sich in K4 nicht.

147 Aus den Ausführungen in K4, wonach ein Klangsegment als stimmhafter Klang erkannt werde, wenn die Kammreaktion einen Schwellenwert überschreite, ergibt sich ebenfalls kein Hinweis auf eine Summierung über die Zeit hinweg.

148 c) Der mit dem Hilfsantrag 1 verteidigte Gegenstand ist durch den
Stand der Technik nicht nahegelegt.

149 Weder aus K3' noch aus K4 ergaben sich Anregungen, ein Residualspektrum durch Subtrahieren eines spektralen Untergrunds zu berechnen.

150 aa) Wie bereits oben dargelegt wurde, wird bei dem in K3' offenbarten Verfahren ein rudimentäres Residualspektrum erzeugt, das nur aus binären Werten besteht.

151 Damit zeigt K3' bereits eine Methode auf, ein Ausgangssignal für die Erkennung tonaler Stabilität zu optimieren. Diese Vorgehensweise bietet nach K3' den Vorteil, dass sie mit einem geringen Umfang an Daten auskommt und dennoch gute Ergebnisse liefert. Vor diesem Hintergrund war eine Abkehr von diesem Konzept zugunsten des Subtrahierens eines spektralen Untergrunds nicht veranlasst.

152 Dass das Subtrahieren eines spektralen Untergrunds insbesondere die Möglichkeit eröffnet hätte, nicht nur die Position der Spitzen, sondern auch deren Form zu ermitteln und bei der Korrelationsanalyse zu berücksichtigen, führt nicht zu einer abweichenden Beurteilung. Auch ein solcher Ansatz hätte eine Abkehr von dem in K3' offenbarten Konzept bedeutet, das sich auf eine Analyse der Position beschränkt.

153 bb) Ausgehend von K3' bestand auch kein Anlass, die dort offenbarte Methode durch das vorhergehende Subtrahieren eines spektralen Untergrunds zu ergänzen.

154 Nach dem Vorbringen der Berufung könnte eine solche Vorgehensweise zwar den Vorteil bieten, dass die Position von Spitzen auch bei steilem Verlauf des Untergrunds zuverlässig ermittelt werden kann. Aus K3' ergaben sich aber keine Hinweise darauf, dass eine solche Maßnahme erforderlich oder hilfreich

sein könnte, um das dort offenbarte, als einfach und dennoch effektiv bezeichnete Verfahren zu verbessern.

155 cc) Ob sich aus den in K3' wiedergegebenen Abbildungen bei ergänzender Heranziehung von Fachwissen Hinweise auf eine vorherige Aufbereitung des Amplitudenspektrums ergaben, bedarf keiner Entscheidung.

156 Selbst wenn die Frage zu bejahen wäre, ergäbe sich daraus keine hinreichende Anregung, diese Aufbereitung durch Subtrahieren eines spektralen Untergrunds vorzunehmen.

157 Dass spectral subtraction in anderen Zusammenhängen als geläufige Methode für Rauschminderung und spektrale Analyse bekannt war, führt nicht zu einer abweichenden Beurteilung. Aus diesem Umstand ergab sich nicht die Erkenntnis, dass diese Methode auch im Zusammenhang mit der in K3' offenbarten Methode von Vorteil sein kann.

158 dd) Aus den in K3' enthaltenen Ausführungen zu einem gesplitteten Kammfilter ergaben sich für den vorliegenden Zusammenhang schon deshalb keine weiteren Anregungen, weil K3' auf diese Vorgehensweise gerade nicht für die Erkennung von tonaler Stabilität zurückgreift, sondern nur im Zusammenhang mit der Erkennung von Sprache.

159 ee) Aus K4 ergab sich hinsichtlich keines der aufgeführten Aspekte eine weitergehende Anregung. Dies gilt insbesondere auch für eine Kombination der Elemente von Musik- und Spracherkennung.

160 IV. Die Kostenentscheidung beruht auf § 121 Abs. 2 PatG und § 97 Abs. 1 ZPO.

Bacher

Hoffmann

Kober-Dehm

Marx

Rombach

Vorinstanz:

Bundespatentgericht, Entscheidung vom 06.12.2021 - 4 Ni 10/21 (EP) -