



BUNDESGERICHTSHOF

IM NAMEN DES VOLKES

URTEIL

X ZR 119/20

Verkündet am:
15. November 2022
Schönthal
Justizangestellte
als Urkundsbeamtin
der Geschäftsstelle

in der Patentnichtigkeitssache

Der X. Zivilsenat des Bundesgerichtshofs hat auf die mündliche Verhandlung vom 15. November 2022 durch den Vorsitzenden Richter Dr. Bacher, die Richter Hoffmann und Dr. Deichfuß sowie die Richterinnen Dr. Kober-Dehm und Dr. Marx

für Recht erkannt:

Die Berufung gegen das Urteil des 6. Senats (Nichtigkeitssenats) des Bundespatentgerichts vom 18. November 2020 wird auf Kosten der Klägerin zurückgewiesen.

Von Rechts wegen

Tatbestand:

1 Der Beklagte ist Inhaber des mit Wirkung für die Bundesrepublik Deutschland erteilten europäischen Patents 954 909 (Streitpatents), das am 13. März 1998 unter Inanspruchnahme einer deutschen Priorität vom 14. Juli 1997 angemeldet wurde und Verfahren zum Codieren und Decodieren eines Audiosignals betrifft. Patentanspruch 2, auf den sechs weitere Ansprüche zurückbezogen sind, lautet:

Verfahren zum Decodieren eines codierten Audiosignals, mit folgenden Schritten:

Empfangen (212) eines codierten Audiosignals;

Erfassen (214) von Informationen in den Seiteninformationen, die sich auf eine Rauschsubstitution und auf Rauschbereiche der spektralen Restwerte beziehen;

Erzeugen (312) von spektralen Rausch-Restwerten aufgrund der erfassten Informationen in den Rauschbereichen;

Durchführen einer inversen Prädiktion (900) über der Frequenz, um aus den rauschsubstituierten spektralen Rausch-Restwerten Spektralwerte zu erhalten; und

Transformieren (218) der Spektralwerte in den Zeitbereich, um ein decodiertes Audiosignal zu erhalten.

2 Patentanspruch 1, auf den die gleichen sechs Ansprüche zurückbezogen sind, betrifft ein entsprechendes Codierverfahren.

3 Die Klägerin, die vom Beklagten aus Patentanspruch 2 gerichtlich in Anspruch genommen wird, hat geltend gemacht, der Gegenstand des Streitpatents sei nicht patentfähig. Der Beklagte hat das Schutzrecht in der erteilten Fassung verteidigt. Nach Erlöschen des Streitpatents haben die Parteien den Rechtsstreit hinsichtlich der Ansprüche 1 und 3 bis 8 übereinstimmend für erledigt erklärt.

4 Das Patentgericht hat die Klage abgewiesen. Hiergegen richtet sich die Berufung der Klägerin, die weiterhin die Nichtigklärung im Umfang von Patentanspruch 2 begehrt. Der Beklagte tritt dem Rechtsmittel entgegen und verteidigt das Streitpatent in der erteilten Fassung sowie mit einem Hilfsantrag.

Entscheidungsgründe:

5 Die zulässige Berufung ist unbegründet.

6 I. Das Streitpatent betrifft die Codierung und Decodierung von Audio-
signalen.

7 1. Nach der Beschreibung des Streitpatents sind solche Verfahren
etwa von der Standardisierungsorganisation ISO/IEC JTC1/SC29/WG11 entwi-
ckelt worden, die auch unter dem Namen Moving Pictures Expert Group (MPEG)
bekannt ist.

8 Bei dem bekannten Verfahren werde ein zeitkontinuierliches Audiosignal
abgetastet, um ein zeitdiskretes Signal zu erhalten (Abs. 5). Dieses werde mit
einer Funktion bearbeitet, um aufeinanderfolgende Blöcke oder Frames mit einer
bestimmten Anzahl gefensterter zeitdiskreter Abtastwerte zu erhalten. Jeder die-
ser Blöcke werde in den Frequenzbereich transformiert, etwa mittels einer modi-
fizierten diskreten Kosinustransformation (Abs. 6).

9 Die erhaltenen Spektralwerte würden so quantisiert, dass das Quantisie-
rungsrauschen durch die quantisierten Signale überlagert (maskiert) und damit
unhörbar werde. Dazu werde ein psychoakustisches Modell herangezogen, das
die Eigenschaften des menschlichen Gehörs berücksichtige (Abs. 6). Zur Quan-
tisierung würden Spektralwerte zu Skalenfaktorbändern gruppiert, die mit be-
stimmten Faktoren skaliert werden könnten. Die Information hierüber werde als
Seiteninformation an den Decodierer übertragen (Abs. 7).

10 Nach der Quantisierung würden die Spektralwerte redundanzcodiert.
Dazu würden sie in Abschnitte (sections) eingeteilt, um Bereiche mit gleicher Sig-
nalstatistik zu erhalten. Dabei seien Abschnittsgrenzen nur an Skalenfaktorband-
grenzen vorgesehen. Dies ermögliche es, einen Abschnitt mithilfe einer einzigen
Codiertabelle, etwa einer Huffman-Codiertabelle zu codieren. Aus den zur Verfü-
gung stehenden zwölf Codiertabellen werde diejenige gewählt, die den größten

Codiergewinn erbringe. Die Länge des Abschnitts in Skalenfaktorbändern und die Nummer der eingesetzten Codiertabelle würden als Seiteninformation an den Decodierer übertragen (Abs. 10).

11 Ebenfalls vorbekannt sei die Technik des "Temporal Noise Shaping" (TNS), die mittels einer prädiktiven Codierung (Vorhersagecodierung) der Spektralwerte die zeitliche Formung der Feinstruktur des Quantisierungsrauschens gestatte (Abs. 12). Das Quantisierungsrauschen könne hierdurch zeitlich unter das tatsächliche Signal gelegt und somit maskiert werden. Probleme der zeitlichen Maskierung von transienten Signalen oder Sprachsignalen ließen sich auf diese Weise vermeiden (Abs. 20 Z. 29-38). Beim TNS-Verfahren werde ein Eingangssignal mittels einer hochauflösenden Analysefilterbank in seine spektrale Darstellung transformiert. Anschließend werde eine lineare Prädiktion im Frequenzbereich ausgeführt, und zwar zwischen frequenzmäßig benachbarten Spektralwerten. Die ursprünglichen Spektralwerte würden durch die Prädiktionsfehler (sogenannte spektrale Restwerte) ersetzt. Diese Restwerte würden ebenso wie übliche Spektralwerte quantisiert und entropie- bzw. redundanzcodiert an den Decodierer übertragen, damit die Werte wieder decodiert, invers quantisiert und invers prädiziert werden könnten (Abs. 29).

12 Die Beschreibung des Streitpatents verweist des Weiteren auf Erkenntnisse aus der Psychoakustik, nach denen der Wahrnehmungseindruck von Rauschsignalen primär nicht von ihrer tatsächlichen Signalform bestimmt werde, sondern von ihrer spektralen Zusammensetzung. Dies ermögliche die Benutzung einer Rauschsubstitutionstechnik zur Datenreduktion von Audiosignalen (Abs. 31). Einer von Donald Schulz (Improving Audio Codecs by Noise Substitution, Journal of the Audio Eng. Soc. Bd. 44 [1996], Nr. 7/8, S. 593-598, D2) beschriebenen Technik der Rauschsubstitution liege zugrunde, dass das menschliche Gehör bei rauschhaften Signalen nicht in der Lage sei, den exakten Zeitverlauf zu erfassen. Ein Codieren der Wellenform erfordere hohe Bitraten für Informationen, die nicht hörbar seien. Gelingen es, rauschhaltige Komponenten von Signalen zu erfassen, könne man sich damit begnügen, Informationen über den

Rauschpegel, den Frequenzbereich oder den zeitlichen Ausdehnungsbereich zu codieren (Abs. 33).

13 Vor diesem Hintergrund bestehe beim Codieren eines Audiosignals die Aufgabe, in dem Spektrum des Audiosignals rauschartige oder rauschhafte Spektralwerte zu erkennen (Abs. 34 Z. 5-7). Diese seien dadurch definiert, dass sie ohne hörbare Unterschiede für das menschliche Gehör durch ein Rauscheretzungsverfahren rekonstruiert werden könnten. Rauschbereiche in Spektralwerten des Audiosignals könnten in unterschiedlicher Weise erfasst werden. Entsprechende Verfahren basierten auf den Spektralwerten, auf dem zeitdiskreten Audiosignal oder sowohl auf dem Audiosignal als auch auf den Spektralwerten (Abs. 49, 34, 35).

14 Eine als rauschhaft klassifizierte Gruppe von Spektralwerten werde nicht wie sonst üblich quantisiert und redundanzcodiert zum Empfänger übertragen. Stattdessen werde nur eine Kennung zur Anzeige der Rauschsubstitution und ein Maß für die Energie der rauschhaften Gruppe von Spektralwerten als Seiteninformation an den Decoder übertragen. Dort würden dann für die substituierten Werte Zufallswerte mit der übertragenen Energie eingesetzt (Abs. 36). Weil nur die Übertragung einer einzigen Energieinformation anstelle einer Gruppe von Codes erfolge, seien erhebliche Dateneinsparungen möglich (Abs. 37).

15 Mit der bekannten Rauschsubstitutionstechnik könne eine mit nicht hörbaren Qualitätseinbußen verbundene Decodierung erreicht werden, wenn das Eingangssignal eine gleichmäßige Rauschstruktur, das heißt ein ebenes oder flaches Spektrum habe. Bei transienten Signalen oder Sprachsignalen sei dies nicht der Fall, so dass hier auf die Rauschsubstitution verzichtet oder störende Verfälschungen in Kauf genommen werden müssten (Abs. 41).

16 2. Das Streitpatent betrifft vor diesem Hintergrund das technische Problem, ein Verfahren zum Codieren bzw. Decodieren von Audiosignalen zu schaffen, das eine hohe Codiereffizienz ermöglicht und möglichst keine hörbaren Signalverfälschungen zur Folge hat (Abs. 42).

17 3. Zur Lösung schlägt Patentanspruch 2 ein Verfahren vor, dessen Merkmale sich wie folgt gliedern lassen:

Verfahren zum Decodieren eines codierten Audiosignals, mit folgenden Schritten:

- a) Empfangen (212) des codierten Audiosignals;
- b) Erfassen (214) von Informationen in den Seiteninformationen, die sich auf eine Rauschsubstitution und auf Rauschbereiche der spektralen Restwerte beziehen;
- c) Erzeugen (312) von spektralen Rausch-Restwerten aufgrund der erfassten Informationen in den Rauschbereichen;
- d) Durchführen einer inversen Prädiktion (900) über der Frequenz, um aus den rauschsubstituierten spektralen Rausch-Restwerten Spektralwerte zu erhalten, und
- e) Transformieren (218) der Spektralwerte in den Zeitbereich, um ein decodiertes Audiosignal zu erhalten.

18 4. Der Anspruch bedarf näherer Erörterung.

19 a) Das Streitpatent befasst sich zwar in erster Linie mit dem zum maßgeblichen Zeitpunkt in der Entwicklung befindlichen Standard MPEG-2 AAC. Der Gegenstand von Patentanspruch 2 ist aber auf diesen nicht beschränkt.

20 Erfasst sind auch andere Verfahren zum Codieren und Decodieren von Audiosignalen mit Transformationscodierung, die im Hinblick auf die hier interessierenden Punkte ähnlich strukturiert sind.

21 b) Patentanspruch 2 befasst sich lediglich mit rauschhaften Gruppen von Spektralwerten, für die eine Rauschsubstitution erfolgt.

22 Für solche Gruppen sind gemäß Merkmal b - anders als üblicherweise für nicht-rauschhafte Gruppen - keine Signalwerte codiert, sondern Seiteninformationen, die die Rekonstruktion des Rauschsignals ermöglichen.

23 c) Nach Merkmal c bezieht sich diese Rauschsubstitution nicht auf die
ursprünglichen Spektralwerte, sondern auf spektrale Rausch-Restwerte, wie sie
etwa mit Hilfe der TNS-Technik erzeugt werden können.

24 d) Die Kombination von Rauschsubstitution und TNS-Technik ermög-
licht nach der Beschreibung des Streitpatents eine Erhöhung des Codiergewinns
ohne hörbare Signalverfälschungen.

25 Das Streitpatent führt hierzu aus, die spektralen Restwerte wiesen einen
wesentlich geringeren Energiegehalt auf als die ursprünglichen Spektralwerte.
Das zugehörige Signal habe im Vergleich zum ursprünglichen Signal einen fla-
cheren Verlauf. Durch Prädiktion der Spektralwerte über der Frequenz werde der
stark schwankende Verlauf der Hüllkurve von transienten Signalen extrahiert,
wodurch ein Signal mit flacher Hüllkurve zurückbleibe, auf das die Rauschsubsti-
tution angewendet werden könne, um auch bei transienten Signalen erhebliche
Biteinsparungen erreichen zu können (Abs. 44, 52).

26 II. Das Patentgericht hat seine Entscheidung - soweit im Berufungs-
verfahren noch von Interesse - im Wesentlichen wie folgt begründet:

27 Der Gegenstand von Patentanspruch 2 sei durch den Stand der Technik
nicht nahegelegt.

28 Der in der Beschreibung des Streitpatents erwähnte Aufsatz von Schulz
(D2) stelle einen geeigneten Ausgangspunkt zur Lösung des gestellten Problems
dar, weil er die Verbesserung von Verfahren zum Codieren und Decodieren von
Audiosignalen (Audio-Codecs) durch Rauschsubstitution betreffe.

29 Als rauschhaft erkannte Spektralwerte würden bei dem in D2 offenbarten
Verfahren nicht codiert, sondern durch Angaben zu den Rauschbereichen, zur
mittleren Leistung der Gruppen von rauschhaften Spektralwerten und zur
Rauschhüllkurve ersetzt. Derartige Informationen verstehe der Fachmann, ein
Diplom-Ingenieur der Fachrichtung Elektro- oder Nachrichtentechnik oder ein

Diplom-Informatiker mit mehrjähriger Berufserfahrung und einschlägigen Kenntnissen auf dem Gebiet der digitalen Signalverarbeitung, insbesondere der Codierung und Decodierung von Audiosignalen, als Seiteninformationen.

30 D2 offenbare nicht das Erfassen und Substituieren von Rauschbereichen in spektralen Restwerten. Daher könne D2 auch kein Decodierungsverfahren mit den entsprechenden Vorgaben der Merkmale b, c und d entnommen werden.

31 Derartige Maßnahmen lägen ausgehend von D2 auch nicht nahe. Keine der dort gelehrtten Möglichkeiten, zwischen rauschhaften und tonalen Signalkomponenten eines Audiosignals zu unterscheiden, verwende eine Prädiktion der Spektralwerte über der Frequenz, um spektrale Restwerte zu erhalten.

32 Der Fachmann möge allerdings Veranlassung gehabt haben, das in D1 beschriebene Codierverfahren durch das in D2 vorgestellte Verfahren zum Erfassen und Substituieren von Rauschen zu verbessern. Dies führe aber lediglich zu einem Codierverfahren, welches gemäß D1 eine Prädiktion von Spektralwerten über der Frequenz mit Hilfe der TNS-Technik durchführe und zusätzlich nach D2 rauschhafte Komponenten des Audiosignals erfasse und substituiere, und zwar unter Verwendung einer Prädiktion der Spektralwerte, des Audiosignals oder der einzelnen Teilbänder über der Zeit. Es fehle hingegen an einem Hinweis oder einer Anregung zum Erfassen von Informationen, die sich auf Rauschbereiche der spektralen Restwerte bezögen (Merkmal b), zum Erzeugen von spektralen Rausch-Restwerten aufgrund solcher Informationen (Merkmal c) und zum Durchführen einer inversen Prädiktion über der Frequenz, um aus solchen Restwerten Spektralwerte zu erhalten (Merkmal d).

33 Aus der Veröffentlichung von Herre und Johnston (Enhancing the Performance of Perceptual Audio Coders by Using Temporal Noise Shaping (TNS), In: An Audio Engineering Society preprint 4384 (N-3), Presented at the 101st Convention 1996 November 8-11 Los Angeles, California, D5/D11) und dem Entwurf des Standards MPEG-2 AAC (ISO/IEC 13818-7 First edition 1996(E). MPEG-2 Audio NBC Committee Draft. ISO/IEC JTC1/SC29/WG11 N1307, D9), ergebe

sich nichts anderes. Ihr Inhalt gehe in Bezug auf den Gegenstand des Anspruchs 2 nicht über den Inhalt von D1 hinaus.

34 Ob der Aufsatz von Edler (Very Low Bit Rate Audio Coding Development, in: Proc. 14th Audio Eng. Soc. Int. Conf., June 1997, D12) vorveröffentlicht sei, müsse nicht entschieden werden, da der Gegenstand von Patentanspruch 2 auch ausgehend hiervon nicht nahegelegt sei.

35 Bei dem in D12 beschriebenen PARA-Codec würden bei der Analyse des Audiosignals für Rahmen von Abtastwerten Parameter extrahiert, die sinusförmige Komponenten (z.B. Frequenz, Amplitude und Phase) und Rauschkomponenten (z.B. Spektralhüllkurve) beschrieben. Aufgrund des Wahrnehmungsmodells würden Phaseninformationen für die sinusförmige Komponente nicht übertragen. Weiterhin sei die Übertragung der Spektralhüllkurve für die Rauschkomponente ausreichend und es müsse kein Restsignal (residual signal) übertragen werden. Dieses Restsignal sei nicht als Differenz zwischen dem tatsächlichen Spektralwert des Audiosignals und dessen Prädiktion durch einen frequenzmäßig benachbarten Spektralwert und damit nicht als spektraler Restwert offenbart. D12 enthalte auch keinen sonstigen Hinweis auf die diesbezügliche Lehre des Streitpatents.

36 Die weiteren in das Verfahren eingeführten Druckschriften lägen weiter ab oder seien nachveröffentlicht.

37 III. Diese Beurteilung hält der Nachprüfung im Berufungsverfahren stand.

38 1. Das Patentgericht hat zu Recht entschieden, dass der Gegenstand von Patentanspruch 2 ausgehend von D1 und D2 nicht nahegelegt wurde.

39 a) D1 nimmt die Merkmale von Patentanspruch 2 nicht vollständig vorweg.

40 aa) D1 befasst sich in Abschnitt 8.1 mit der zeitlichen Rauschformung
(Temporal Noise Shaping).

41 Bei transienten und tonalen Eingangssignalen mit einem nicht flachen
Spektrum sei das Erreichen eines Maskierungseffekts im reproduzierten Audio-
signal aufgrund der zeitlichen Diskrepanz zwischen Maskierungsschwelle und
Quantisierungsrauschen problematisch. Dies sei bekannt als Vor-Echo-Problem.

42 Die TNS-Technik ermögliche es, die zeitliche Feinstruktur des Quantisie-
rungsrauschens auch innerhalb eines Filterbank-Fensters zu kontrollieren. Durch
Anwendung prädiktiver Codierungsverfahren auf Spektraldaten könne der Pegel
des Quantisierungsrauschens effektiv unter das tatsächliche Audiosignal ge-
bracht werden. Dies ermögliche eine effizientere Nutzung von Maskierungseffek-
ten, indem die zeitliche Feinstruktur des Quantisierungsrauschens an die des
maskierenden Signals angepasst werde (S. 23 Abs. 2).

43 Der prädiktive Codierungs- und Decodierungsprozess über die Frequenz
könne durch Hinzufügen eines Bausteins zur Standardstruktur eines Codierers
und Decodierers realisiert werden. Nach der Analyse-Filterbank werde, wie in der
nachfolgend wiedergegebenen Figur 8.1 gezeigt, eine TNS-Filterung eingefügt.

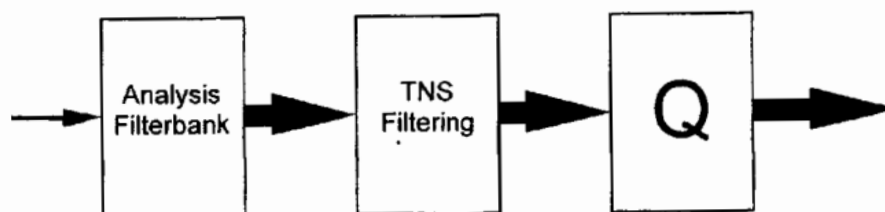


Figure 8.1 - TNS Processing for the MPEG-2 NBC Encoder.

44 Es erfolge eine Prädiktion von Spektralwerten des Audiosignals über der
Frequenz, um sie durch ihre spektralen Restwerte zu ersetzen (Abschnitt 8.1.1
erster Absatz). Diese Restwerte würden quantisiert und entropie- bzw. redun-
danzcodiert an den Decodierer übertragen.

45 In entsprechender Weise sei, wie durch die nachfolgend wiedergegebene
Figur 8.3 veranschaulicht, beim Decodieren eine inverse TNS-Filterung vor der
Synthese-Filterbank vorgesehen.

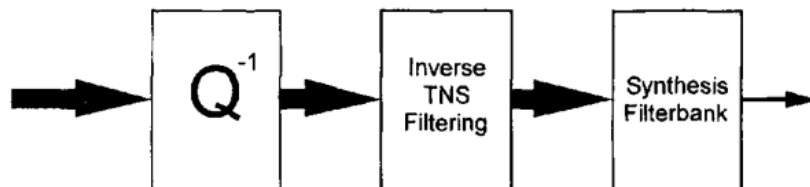


Figure 8.3 - TNS Processing for the MPEG-2 NBC Decoder.

46 bb) Damit ist, wie das Patentgericht zutreffend angenommen hat, ein
Verfahren zum Decodieren eines codierten Audiosignals offenbart, bei dem im
Sinne von Merkmal a ein codiertes Audiosignal empfangen (Block Q^{-1}) und eine
inverse Prädiktion über der Frequenz durchgeführt wird, um aus spektralen Rest-
werten Spektralwerte zu erhalten. Um ein decodiertes Audiosignal zu erhalten,
müssen diese Spektralwerte dann entsprechend Merkmal e in den Zeitbereich
transformiert werden.

47 cc) Nicht vollständig offenbart sind die Merkmale b bis d.

48 D1 sieht nicht vor, beim Codieren Rauschbereiche in den spektralen Rest-
werten zu erfassen, die spektralen Restwerte für diese Bereiche durch Informa-
tionen bezüglich der Rauschbereiche zu ersetzen und diese in die Seiteninfor-
mationen des codierten Audiosignals aufzunehmen. Dementsprechend fehlt es
auch an einer Offenbarung diesbezüglicher Decodierschritte.

49 b) D2 nimmt ebenfalls nicht sämtliche Merkmale von Patentan-
spruch 2 vorweg.

50 aa) D2 führt aus, die Verringerung der erforderlichen Daten bei der
Audiocodierung sei bislang vor allem durch Maskierungseffekte erzielt worden,
etwa dadurch, dass Komponenten, die durch andere Komponenten überdeckt

werden, nicht codiert werden. Die meisten Audiosignale enthielten aber auch Frequenzbereiche, die das menschliche Gehör nicht wahrnehme.

51 Zur korrekten Codierung solcher Bereiche würden bislang hohe Bitraten benötigt. Würden die erkannten rauschhaften Komponenten des Signals demgegenüber nur als Information über den Rauschpegel, den Frequenz- und den Zeitbereich codiert, könne eine korrekte Beschreibung der rauschhaften Signale mit erheblichen Einsparungen erreicht werden (S. 593 linke Spalte, Abschnitt 0).

52 Eine Signalkomponente sei rauschhaft, wenn sie durch den Pegel, den Frequenzbereich und den Zeitbereich so charakterisiert werden könne, dass eine Rekonstruktion möglich sei, ohne dass der Hörer einen Unterschied bemerke (S. 593 rechte Spalte, Abschnitt 1). Es gebe verschiedene Möglichkeiten, wie rauschhafte Komponenten von nicht-rauschhaften (tonalen) Komponenten unterschieden und erfasst werden könnten.

53 Nach D2 ist es vorteilhaft, die Datenkompression von Audiosignalen für rauschhafte Signale mit der herkömmlichen Kompression für nicht-rauschhafte Signale zu verbinden. Wie für die Rauschsubstitution sei es auch für die Maskierung erforderlich, das gesamte Frequenzband zu unterteilen. Dies könne durch adaptive Transformationscodierung geschehen, etwa durch die modifizierte diskrete Kosinustransformation (MDCT) oder durch Verwendung einer Teilband-Filterbank. Mit MDCT könne eine hohe Kompression erzielt werden, doch erfordere sie eine höhere Rechenleistung als die Teilbandkodierung(S. 595 rechte Spalte, Abschnitt 2).

54 Wie die nachfolgend wiedergegebene Figur 4 veranschaulicht, werden zur adaptiven Transformationscodierung mit Rauschsubstitution ein Audiosignal und eine Prädiktion dieses Audiosignals in den Frequenzbereich transformiert.

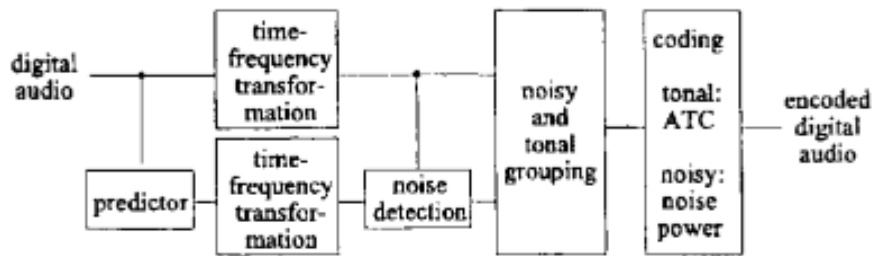


Fig. 4. Noise substitution in combination with adaptive transform coding.

55 Rauschhafte Spektralwerte würden erkannt, indem die Spektralwerte des Audiosignals mit denen des prädizierten Audiosignals verglichen werden (S. 596 linke Spalte, Absatz 1). Hierbei könne eine Prädiktion in entsprechender Weise verwendet werden, wie sie in der nachfolgend wiedergegebenen Figur 2 für eine Rauscherkennung in Subbändern dargestellt sei.

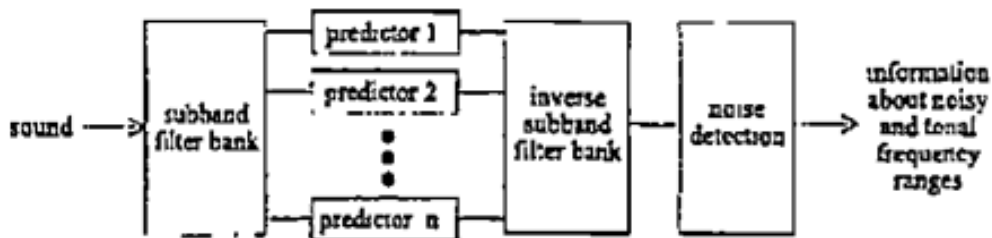


Fig. 2. Noise detection by prediction in subbands.

56 Das Audiosignal werde hier durch eine mehrphasige Quadraturfilterbank in Teilbänder aufgeteilt. In jedem Teilband werde eine separate Prädiktion über der Zeit durchgeführt. Die prädizierten Teilbandsignale würden unter Verwendung der inversen Filterbank wieder zusammengeführt und mittels einer schnellen Fourier-Transformation (FFT) in den Frequenzbereich transformiert. Die nachfolgend wiedergegebene Gleichung (3) könne zur Erfassung der Tonalität der Spektralwerte und damit zur Unterscheidung tonaler von rauschhaften Spektralwerten verwendet werden (S. 594 rechte Spalte, Abschnitt 1.4).

$$T^i(n) = \alpha \left| \frac{\hat{P}^i(n) - P^i(n)}{P^i(n)} \right| + \beta \left| \frac{\hat{\Theta}^i(n) - \Theta^i(n)}{\Theta^i(n)} \right| \quad (3)$$

57 Das Patentgericht hat insoweit unbeanstandet festgestellt, dass in das Tonalitätsmaß $T^i(n)$ die normierte Differenz eingeht zwischen dem Spektralwert $\hat{P}^i(n)$ der Frequenz n im i -ten Frame des aus den prädizierten Teilbandsignalen zusammengeführten Signals und dem Spektralwert $P^i(n)$ der Frequenz n im i -ten Frame des ursprünglichen Audiosignals.

58 Zur Erhöhung des Kompressionsfaktors würden benachbarte tonale und rauschhafte Frequenzwerte gruppiert. Ein rauschhafter Spektralwert (noisy frequency value), der keine rauschhaften Nachbarn habe, werde als tonal angesehen. Da das Gehör im Frequenzbereich bis 5 kHz Zeitverzögerungen zwischen rechtem und linkem Kanal wahrnehme, sei die Rauschsubstitution auf den Bereich über 5 kHz zu beschränken. Verwende man eine Transformation mit einer Frequenzauflösung von 40 Hz, bedürfe es zudem einer Einhüllenden (envelope) der Zeitstruktur rauschhafter Signalkomponenten. Sodann würden der mittlere Energielevel der verbliebenen Gruppe, ihr Frequenzbereich und ihre Einhüllenden an den Decoder übertragen (S. 596, Abschnitt 2.1, vorletzter Absatz).

59 bb) Das Patentgericht ist zutreffend davon ausgegangen, dass mit der Beschreibung des Codierverfahrens auch die notwendigen inversen Verfahrensschritte zum Decodieren des Signals offenbart sind. Folglich offenbart D2 die Merkmale a und e.

60 cc) Nicht offenbart sind die Merkmale b bis d.

61 D2 offenbart insoweit zwar das Erfassen von Informationen in den Seiteninformationen, die sich auf eine Rauschsubstitution und auf Rauschbereiche der

Spektralwerte beziehen, um Rausch-Spektralwerte in diesen Rauschbereichen zu erzeugen.

62 Nicht offenbart ist hingegen, dass sich die erfassten Informationen auf Rauschbereiche der spektralen Restwerte beziehen, wie dies die Merkmale b bis d vorsehen.

63 c) Ausgehend von D1 und D2 war der Gegenstand von Patentanspruch 2 nicht nahegelegt. Der Stand der Technik vermittelte keine Anregung, die Rauschsubstitution auch auf die durch Nutzung der TNS-Technik erlangten spektralen Restwerte anzuwenden.

64 aa) Wie das Patentgericht zu Recht angenommen hat, bestand ausgehend von D1 allerdings Anlass zu prüfen, ob sich die in D2 vorgeschlagene Rauschsubstitution in das für den Standard vorgesehene Codierverfahren einfügen lässt.

65 (1) Ausgangspunkt der in D1 beschriebenen TNS-Technik ist, eine Prädiktion von Spektralwerten des Audiosignals über der Frequenz etwa bei transienten Signalen durchzuführen, um die zeitliche Feinstruktur des Quantisierungsrauschens so zu formen, dass es von dem Audiosignal maskiert werden kann und insbesondere keine hörbaren Vor-Echos auftreten. Lediglich in diesem Zusammenhang werden die Spektralwerte durch ihre spektralen Restwerte (prediction residual) ersetzt und quantisiert und entropie- bzw. redundanzcodiert an den Decodierer übertragen.

66 Das in D2 vorgeschlagene Verfahren beruht demgegenüber auf dem Ansatz, rauschhafte Spektralwerte nicht in der üblichen Weise zu codieren, sondern lediglich Werte zu übermitteln, die es ermöglichen, die Spektralwerte rauschhafter Gruppen durch zufällig auf der Decodier-Seite generierte Spektralwerte zu ersetzen. In diesem Zusammenhang erfolgt zwar ebenfalls eine Prädiktion des Audiosignals. Diese dient aber nicht der Bildung oder Analyse von Restwerten, sondern der Erkennung rauschhafter Komponenten des Audiosignals.

67 Um zur Lehre des Streitpatents zu gelangen, bedurfte es der Erkenntnis, dass die in D2 vorgeschlagene Rauschsubstitution nicht nur in Bezug auf rauschhafte Bereiche des ursprünglichen Signals Vorteile bietet, sondern mit vergleichbarer oder sogar besserer Wirkung auch auf mittels Prädiktion ermittelte spektrale Restwerte angewendet werden kann. Hierfür ist, wie das Patentgericht zutreffend erkannt hat, aus D1 und D2 keine Anregung ersichtlich.

68 (2) Entgegen der Auffassung der Berufung genügte es nicht, die Verfahren nach D1 und D2 einfach in Reihe zu schalten. Eine solche Kombination wäre zwar bezüglich der einzelnen Verfahrensschritte möglich. Um zur Lehre des Streitpatents zu gelangen, müssen die in D2 vorgesehenen Verfahrensschritte aber an einem anderen Objekt durchgeführt werden, nämlich an den gemäß D1 ermittelten spektralen Restwerten. Diese Art der Kombination lag ohne Anregung nicht nahe.

69 (3) Entgegen der Ansicht der Berufung ergab sich für den Fachmann eine entsprechende Anregung nicht aus dem Hinweis der D2 auf die Zeitauflösung des menschlichen Gehörs und die daraus resultierende Problematik gemeinsam codierbarer Gruppen von nicht mehr als drei Subband-Samples (S. 596, Abschnitt 2.3, erster Absatz).

70 Diese nur die Subbandcodierung betreffenden Ausführungen bieten keinen Hinweis für eine Transformationscodierung nach D1. Da die Subbandcodierung eine höhere zeitliche Auflösung besitzt und die in D1 beschriebenen Ursachen für das Auftreten von Vor-Echos nicht vorhanden sind, ist auch nicht ersichtlich, weshalb der Fachmann bei der Subbandcodierung einen Zusammenhang zur Technik der zeitlichen Rauschumformung nach D1 herstellen sollte. Ferner fehlt es an Anhaltspunkten dafür, dass eine Prädiktion über die Frequenz bei einer Subbandcodierung überhaupt zu einer relevanten Verbesserung der zeitlichen Auflösung führen würde.

71 (4) Nichts anderes folgt aus der Offenbarung in D2, dass das Rauschen durch den Vergleich der Frequenzwerte des in den Frequenzbereich transformierten vorhergesagten Tons mit dem Originalton erkannt wird (S. 596 linke Spalte, erster Absatz; Fig. 4).

72 Das Patentgericht hat festgestellt, dass die insoweit vorgeschlagene Prädiktion entsprechend Figur 2 ebenfalls über die Zeit durchgeführt wird. Die Berufung zeigt keine Umstände auf, die konkrete Zweifel an der Vollständigkeit oder Richtigkeit dieser Feststellungen begründen.

73 bb) Aus einer Kombination von D2 mit D1, D5 oder D11 ergeben sich keine weitergehenden Anregungen.

74 Der Umstand, dass D2 die dort in Figur 2 dargestellte Prädiktion als ideal bezeichnet (S. 596, linke Spalte, erster Absatz), schließt Alternativen zwar nicht aus. Aus ihm ergibt sich aber noch keine Anregung, ein grundlegend abweichendes Prädiktionsmodell heranzuziehen.

75 Eine Anregung, außerhalb der Offenbarung der D2 nach Prädiktionsverfahren zu suchen, lässt sich auch nicht daraus herleiten, dass nach D2 die bei echten stationären Signalen gut funktionierende Gleichung (3) bei den häufig in Audiosignalen vorkommenden Sinuskurven mit sich im Zeitverlauf ändernden Frequenzen versagt und diese Methode daher nicht zusammen mit der Rauschsubstitution verwendet werden kann (S. 594, letzter Absatz vor Abschnitt 1.2). Denn D2 stellt hierfür in den Folgeabschnitten eigene Lösungsansätze vor.

76 Selbst wenn unterstellt wird, dass Anlass bestand, andere als die in D2 vorgestellten Prädiktionsverfahren in den Blick zu nehmen, fehlt es jedenfalls an einer Anregung dafür, anstelle einer Prädiktion über der Zeit eine Prädiktion über der Frequenz vorzunehmen.

77 Damit bestand entgegen der Ansicht der Berufung auch kein Anlass, das in D1, D5 oder D11 für die TNS-Technik offenbarte Prädiktionsmodell, bei dem

spektrale Restwerte in Verbindung mit der Formung der zeitlichen Feinstruktur des Quantisierungsrauschens erzeugt werden, in den Blick zu nehmen, und auf eine Verwendbarkeit im Zusammenhang mit der in D2 gelehrt Rauschsubstitution zu überprüfen.

78 cc) D5, D11 oder D9 gehen nach den nicht angegriffenen Feststellungen des Patentgerichts im Hinblick auf die Beschreibung spektraler Restwerte in Verbindung mit der Formung der zeitlichen Feinstruktur des Quantisierungsrauschens nach der TNS-Technik nicht in relevanter Weise über die Offenbarung der D1 hinaus. Damit ergeben sich aus diesen Veröffentlichungen keine weitergehenden Anregungen.

79 2. Schließlich hat das Patentgericht zu Recht entschieden, dass der Gegenstand von Patentanspruch 2 auch nicht durch eine Kombination des vorbekannten TNS-Verfahrens, wie es etwa in D11 offenbart ist, mit D12 nahegelegt war.

80 a) Nach D12 erlaubt die Unempfindlichkeit des menschlichen Ohrs gegenüber Veränderungen der Wellenform von Audiosignalen die Entwicklung von Codiertechniken, die auf einer effizienten Darstellung der spektralen Eigenschaften basieren, ohne eine Annäherung für die zeitliche Wellenform des Eingangssignals zu erzeugen. Das zugrundeliegende Modell für diese Codierungstechniken beruhe auf der Annahme, dass das Eingangssignal aus sinusförmigen Signalkomponenten und rauschähnlichen Komponenten mit relativ konstanten Eigenschaften für bestimmte Zeitintervalle bestehe.

81 Die entsprechende Analyse des Audiosignals müsse für Rahmen von Abtastwerten Parameter extrahieren, die sinusförmige Komponenten (z.B. Frequenz, Amplitude und Phase) und Rauschkomponenten (z.B. Spektralhüllkurve) beschrieben. Phaseninformationen für die sinusförmige Komponente müssten nicht übertragen werden. Weiterhin sei die Übertragung der Spektralhüllkurve für die Rauschkomponente ausreichend. Ein Restsignal (residual signal) müsse

nicht übertragen werden. Das Modell lasse sich im Hinblick auf Sprach- und Musikinhalte eines Eingangssignals verfeinern.

82 Eine auf dieser Technik basierende Vorgehensweise bezeichnet D12 als PARA-Codec, da sie eher auf einer parametrischen Darstellung der Signale als auf einer Annäherung der Wellenform beruhe (S. 3 Kapitel 3, erster Absatz).

83 b) Das Patentgericht hat zutreffend angenommen, dass damit keine spektralen Restwerte offenbart werden, die die Differenz zwischen dem Spektralwert des Audiosignals und der Prädiktion durch einen frequenzmäßig benachbarten Spektralwert beschreiben, und damit erst recht nicht die Verwendung von Rauschbereichen spektraler Restwerte bzw. von spektralen Rausch-Restwerten im Sinne der Merkmale b bis d.

84 Der Einwand der Berufung, der Begriff der Rauschsubstitution lasse sich als Ersetzen einer Wellenform-Darstellung des Rauschsignals durch seine parametrische Repräsentation definieren, führt zu keinem anderen Ergebnis. Denn auch eine solche Betrachtung führt nicht zwangsläufig zu der Annahme, dass im genannten Sinne spektrale Restwerte erzeugt und gemäß dem beanspruchten Verfahren erfasst und verarbeitet werden.

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

Bacher

Hoffmann

Deichfuß

Kober-Dehm

Marx

Vorinstanz:

Bundespatentgericht, Entscheidung vom 18.11.2020 - 6 Ni 30/17 (EP) -