



(11) **EP 2 402 942 A2**

(12) **EUROPÄISCHE PATENTANMELDUNG**

(43) Veröffentlichungstag:  
**04.01.2012 Patentblatt 2012/01**

(51) Int Cl.:  
**G10L 19/02<sup>(2006.01)</sup> H04S 5/00<sup>(2006.01)</sup>**

(21) Anmeldenummer: **11182960.2**

(22) Anmeldetag: **30.01.2007**

(84) Benannte Vertragsstaaten:  
**AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LI LT LU LV MC NL PL PT RO SE SI SK TR**

- **Herre, Jürgen**  
**91054 Buckenhof (DE)**
- **Geyersberger, Stephan**  
**97076 Würzburg (DE)**
- **Walther, Andreas**  
**96502 Bamberg (DE)**
- **Janssen, Christiaan**  
**90409 Nürnberg (DE)**

(30) Priorität: **12.04.2006 DE 102006017280**  
**12.04.2006 US 744718 P**

(62) Dokumentnummer(n) der früheren Anmeldung(en) nach Art. 76 EPÜ:  
**07703145.8 / 2 005 421**

(74) Vertreter: **Zinkler, Franz et al**  
**Schoppe, Zimmermann, Stöckeler & Zinkler**  
**Patentanwälte**  
**Postfach 246**  
**82043 Pullach bei München (DE)**

(71) Anmelder: **Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.**  
**80686 München (DE)**

Bemerkungen:

Diese Anmeldung ist am 27-09-2011 als Teilanmeldung zu der unter INID-Code 62 erwähnten Anmeldung eingereicht worden.

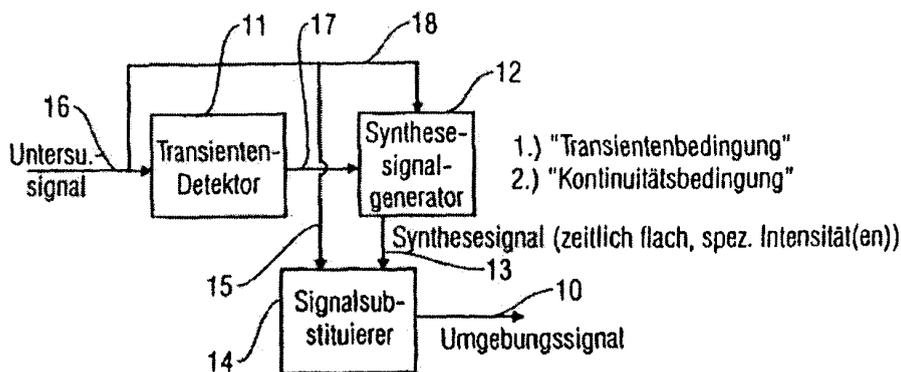
(72) Erfinder:  
• **Hellmuth, Oliver**  
**91052 Erlangen (DE)**

(54) **Vorrichtung und Verfahren zum Erzeugen eines Umgebungssignals**

(57) Zum Erzeugen eines Umgebungssignals, das zur Ausstrahlung über Lautsprecher geeignet ist, für die kein eigenes Lautsprechersignal existiert, also beispielsweise für Surround-Kanäle, ist ein Transienten-Detektor (11) vorgesehen, um einen Transientenzeitraum zu detektieren. Ein Synthesesignalgenerator (12) erzeugt ein

Synthesesignal, das einerseits die Transientenbedingung und andererseits die Kontinuitätsbedingung für das Synthesesignal erfüllt. Ein Signalsubstituierer (14) ersetzt dann einen Abschnitt des Untersuchungssignals durch das Synthesesignal, um ein Umgebungssignal für die Surround-Kanäle zu erhalten.

**FIG 1**



**EP 2 402 942 A2**

## Beschreibung

**[0001]** Die vorliegende Erfindung bezieht sich auf die Audiosignalverarbeitung und insbesondere auf Konzepte zum Erzeugen von Umgebungssignalen (Ambience-Signalen) für Lautsprecher in einem Multikanal-Szenario, für die kein eigenes Lautsprechersignal übertragen worden ist.

**[0002]** Multikanal-Audiomaterial wird mehr und mehr populär. Dies hat dazu geführt, dass inzwischen auch viele Endbenutzer Multikanal-Wiedergabesysteme besitzen. Dies ist hauptsächlich darauf zurückzuführen, dass DVDs immer populärer werden, und dass daher auch viele Benutzer von DVDs inzwischen 5.1-Multikanal-Ausrüstungen haben. Solche Wiedergabesysteme bestehen im allgemeinen aus drei Lautsprechern L (Links), C (Center) und R (Rechts), die typischerweise vor dem Benutzer angeordnet sind, und zwei Lautsprechern Ls und Rs, die hinter dem Benutzer angeordnet sind, und typischerweise noch aus einem LFE-Kanal, der auch Niederfrequenz-Effekt-Kanal oder Subwoofer genannt wird. Ein solches Kanal-Szenario ist in Fig. 10 und in Fig. 11 angedeutet. Während die Positionierung der Lautsprecher L, C, R, Ls, Rs, bezüglich des Benutzers wie in den Figuren 10 und 11 gezeichnet vorgenommen werden sollte, damit der Benutzer einen möglichst guten Höreindruck bekommt, ist die Positionierung des LFE-Kanals (in Figs. 10 und 11 nicht gezeigt) nicht so entscheidend, da das Ohr bei derart niedrigen Frequenzen keine Lokalisierung vornehmen kann und der LFE-Kanal somit irgendwo, wo er aufgrund seiner beträchtlichen Größe nicht stört, angeordnet werden kann.

**[0003]** Ein solches Mehrkanalsystem erzeugt mehrere Vorteile gegenüber einer typischen Stereo-Reproduktion, die eine Zweikanal-Reproduktion ist, wie sie z.B. in Fig. 9 gezeigt ist. Auch außerhalb der optimalen mittigen Hörposition ergibt sich eine verbesserte Stabilität des vorderen Höreindrucks, der auch als "Front Image" bezeichnet wird, und zwar aufgrund des Mitten-Kanals. Es ergibt sich somit ein größerer "Sweet-Spot", wobei "Sweet Spot" für die optimale Hörposition steht.

**[0004]** Ferner hat der Zuhörer ein besseres Gefühl des "Eintauchens" in die Audiaszene aufgrund der beiden hinteren Lautsprecher Ls und Rs.

**[0005]** Dennoch existiert eine riesige Anzahl an im Besitz des Benutzers befindlichen oder allgemein verfügbaren Audiomaterials, das nur als Stereomaterial existiert, das also nur zwei Kanäle hat, nämlich den linken Kanal und den rechten Kanal. Typische Tonträger für solche Stereostücke sind Kompakt-Disks.

**[0006]** Um ein solches Stereomaterial über eine 5.1-Multikanal-Audioanlage abzuspielen, hat man zwei Optionen, die gemäß der ITU empfohlen werden.

**[0007]** Die erste Option besteht darin, den linken und den rechten Kanal über den linken und den rechten Lautsprecher des Multikanal-Wiedergabesystems abzuspielen. Nachteilig an dieser Lösung ist jedoch, dass man die Vielzahl der bereits bestehenden Lautsprecher nicht

ausnutzt, dass man also das Vorhandensein des Center-Lautsprechers und der beiden hinteren Lautsprecher nicht vorteilhaft ausnutzt.

**[0008]** Eine andere Option besteht darin, die zwei Kanäle in ein Multikanalsignal umzuwandeln. Dies kann während der Wiedergabe oder durch eine spezielle Vorverarbeitung geschehen, welche alle sechs Lautsprecher des beispielsweise vorhandenen 5.1-Wiedergabesystems vorteilhaft ausnutzt und damit zu einem verbesserten Höreindruck führt, wenn das Hochmischen oder der "Upmix" von zwei Kanälen auf 5 bzw. 6 Kanäle fehlerfrei durchgeführt wird.

**[0009]** Nur dann hat die zweite Option, also die Verwendung sämtlicher Lautsprecher des Mehrkanalsystems einen Vorteil gegenüber der ersten Lösung, wenn man also keine Upmix-Fehler begeht. Solche Upmix-Fehler können insbesondere störend sein, wenn Signale für die hinteren Lautsprecher, welche auch als Ambience-Signale oder Umgebungssignale bekannt sind, nicht fehlerfrei erzeugt werden.

**[0010]** Eine Möglichkeit, diesen sogenannten Upmix-Prozess durchzuführen, ist unter dem Stichwort "Direct Ambience-Konzept" bekannt. Die Direktschallquellen werden durch die drei vorderen Kanäle derart wiedergegeben, dass sie von dem Benutzer an der gleichen Position wie in der ursprünglichen Zweikanalversion wahrgenommen werden. Die ursprüngliche Zweikanalversion ist in Fig. 9 schematisch dargestellt, und zwar am Beispiel verschiedener Drum-Instrumente.

**[0011]** Fig. 10 zeigt eine hochgemischte Version des Konzepts, bei der alle ursprünglichen Schallquellen, also die Drum-Instrumente wieder von den drei vorderen Lautsprechern L, C und R wiedergegeben werden, wobei zusätzlich von den beiden hinteren Lautsprechern spezielle Umgebungssignale ausgegeben werden. Der Ausdruck "Direkt-Schallquelle" wird somit dazu verwendet, um einen Ton zu beschreiben, der nur und direkt von einer diskreten Schallquelle wie beispielsweise einem Drum-Instrument oder einem anderen Instrument oder allgemein einem speziellen Audioobjekt kommt, wie es schematisch z.B. in Fig. 9 anhand eines Drum-Instruments dargestellt ist. Irgendwelche zusätzlichen Töne, wie beispielsweise aufgrund von Wandreflexionen etc. sind in einer solchen Direktschallquelle nicht vorhanden. In diesem Szenario bestehen die Schallsignale, die von den beiden hinteren Lautsprechern Ls, Rs in Fig. 10 abgegeben werden, nur aus Umgebungssignalen, die in der ursprünglichen Aufzeichnung vorhanden sind oder nicht. Solche Umgebungssignale oder "Ambience"-Signale gehören nicht zu einer einzigen Schallquelle, sondern tragen zur Reproduktion der Raumakustik einer Aufzeichnung bei und führen somit zu dem sogenannten "Eintauch"-Gefühl des Zuhörers.

**[0012]** Ein weiteres Alternativkonzept, das als "In-the-Band"-Konzept bezeichnet ist, ist in Fig. 11 schematisch dargestellt. Jeder Schalltyp, also Direktschallquellen und umgebungs-artige Töne werden alle um den Zuhörer herum positioniert. Die Position eines Tons ist unabhän-

gig von seiner Charakteristik (Direktschallquellen oder umgebungs-artige Töne) und hängt nur von dem spezifischen Entwurf des Algorithmus ab, wie es z.B. in Fig. 11 dargestellt ist. So wurde in Fig. 11 durch den Upmix-Algorithmus bestimmt, dass die beiden Instrumente 1100 und 1102 seitlich bezüglich des Zuhörers positioniert werden, während die beiden Instrumente 1104 und 1106 vor dem Benutzer positioniert werden. Dies führt dazu, dass die beiden hinteren Lautsprecher Ls, Rs nunmehr auch Anteile der beiden Instrumente 1100 und 1102 enthalten und nicht mehr nur umgebungs-artige Töne, wie es bei Fig. 10 noch der Fall war, wo dieselben Instrumente alle vor dem Benutzer positioniert worden sind.

**[0013]** Die Fachveröffentlichung "C. Avendano und J.M. Jot: "Ambience Extraction and Synthesis from Stereo Signals for Multichannel Audio Upmix", IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 02, Orlando, FL, May 2002" offenbart eine Frequenzbereichstechnik, um Umgebungsinformationen in Stereo-Audiosignalen zu identifizieren und zu extrahieren. Dieses Konzept basiert auf der Berechnung einer Inter-Kanal-Kohärenz und einer nichtlinearen Abbildungsfunktion, die es erlauben soll, Zeit-Frequenz-Regionen in dem Stereosignal zu bestimmen, die hauptsächlich aus Umgebungs-Komponenten bestehen. Umgebungssignale werden dann synthetisiert und verwendet, um die hinteren Kanäle oder "Surround"-Kanäle Ls, Rs (Figs. 10 und 11) eines Multikanal-Wiedergabesystems zu speichern.

**[0014]** In der Fachveröffentlichung "R. Irwan und Ronald M. Aarts: "A method to convert stereo to multi-channel sound", The proceedings of the AES 19th International Conference, Schloss Elmau, Deutschland, Juni 21-24, Seiten 139-143, 2001" wird ein Verfahren präsentiert, um ein Stereosignal in ein Multikanalsignal umzuwandeln. Das Signal für die Surround-Kanäle wird unter Verwendung einer Kreuzkorrelationstechnik berechnet. Eine Hauptkomponentenanalyse (PCA; PCA = Principle Component Analysis) wird verwendet, um einen Vektor zu berechnen, der eine Richtung des dominanten Signals anzeigt. Dieser Vektor wird dann von einer Zwei-Kanal-Darstellung auf eine Drei-Kanal-Darstellung abgebildet, um die drei vorderen Kanäle zu erzeugen.

**[0015]** Die Fachveröffentlichung "G. Soulodre, "Ambience-Based Upmixing", Workshop "Spatial Coding Of Surround Sound: A Progress Report", 117th AES Convention, San Francisco, CA, USA, 2004" offenbart ein System, das ein Multikanalsignal aus einem Stereosignal erzeugt. Das Signal wird in sogenannte individuelle Quellenströme und Umgebungsströme zerlegt. Basierend auf diesen Strömen synthetisiert ein sogenannter "Ästhetik-Prozessor" das Multikanal-Ausgangssignal.

**[0016]** Alle bekannten Techniken versuchen auf verschiedene Arten und Weisen die Ambience-Signale bzw. Umgebungssignale aus dem ursprünglichen Stereosignal zu extrahieren oder sogar aus Rauschen bzw. weiteren Informationen zu synthetisieren, wobei zur Synthese der Ambience-Signale auch Informationen, welche

nicht im Stereosignal sind, verwendet werden können. Letztendlich geht es jedoch immer darum, Informationen aus dem Stereosignal zu extrahieren bzw. Informationen in ein Wiedergabe-Szenario einzuspeisen, die nicht explizit vorliegen, da typischerweise nur ein Zweikanal-Stereosignal und gegebenenfalls irgendwelche Zusatzinformationen bzw. Metainformationen zur Verfügung stehen.

**[0017]** Insofern ist die Extraktion oder teilweise Extraktion und teilweise Synthetisierung von solchen Umgebungssignalen eine riskante Angelegenheit, da es ein Benutzer als störend empfinden würde, wenn in den Umgebungs Kanälen Informationen von Schallquellen enthalten sind, die der Benutzer als direkt von vorne kommend, also vom linken Kanal, mittleren Kanal und rechten Kanal kommend identifiziert. Aus diesem Grund würde man eine Erzeugung von Umgebungssignalen sehr "defensiv" machen, um ganz sicher zu gehen, dass keine vom Benutzer als störend empfundenen Artefakte erzeugt werden. Der andere Extremfall, wenn man zu defensiv bei der Erzeugung der Umgebungssignale agiert, besteht darin, dass ein sehr leises, bzw. kaum mehr wahrnehmbares Umgebungssignal extrahiert wird, oder dass das Umgebungssignal nur noch Rauschen hat, jedoch keine besonderen Informationen mehr hat, so dass das Umgebungssignale sehr wenig zum Hörerlebnis beiträgt und eigentlich in diesem Fall ganz weggelassen werden könnte.

**[0018]** Problematisch bei der Erzeugung des Umgebungssignals ist somit, dass man einerseits ein Umgebungssignal erzeugt, das Informationen umfasst, die über normales Rauschen hinausgehen, das jedoch das Umgebungssignal nicht zu hörbaren Artefakten führt, dass also ein richtiges Maß zwischen Hörbarkeit und Informationsgehalt beibehalten wird.

**[0019]** Die Aufgabe der vorliegenden Erfindung besteht darin, ein Konzept zum Erzeugen eines Umgebungssignals zu schaffen, in dem hörbare Artefakte reduziert sind.

**[0020]** Diese Aufgabe wird durch eine Vorrichtung zum Erzeugen eines Umgebungssignals gemäß Patentanspruch 1, ein Verfahren zum Erzeugen eines Umgebungssignals gemäß Patentanspruch 20 oder ein Computer-Programm gemäß Patentanspruch 21 gelöst.

**[0021]** Der vorliegenden Erfindung liegt die Erkenntnis zugrunde, dass die Artefakte, die von Zuhörern bei Umgebungssignalen am negativsten empfunden werden, Artefakte sind, die dazu führen, dass der Zuhörer meint, dass im hinteren Lautsprecher eine direkte Schallquelle ist, obgleich er diese Schallquelle von vorne wahrnimmt. Charakteristika für das Empfinden von direkten Schallquellen sind transiente Vorgänge, also Signal-Feinstrukturen im Zeitsignal, die eine (schnelle) Änderung über einer Änderungsschwelle von einem leisen Zustand in einen lauten Zustand oder von einem lauten Zustand in einen leisen Zustand betreffen, bzw. die einen (starken) Energieanstieg über einer Änderungsschwelle in speziellen Bändern und insbesondere in den oberen Bändern innerhalb einer bestimmten Zeit betreffen.

**[0022]** Solche transienten Vorgänge sind beispielsweise das Einsetzen eines Instruments oder der Anschlag eines Schlaginstruments, oder das Ende eines Tons, der nicht langsam ausklingt, sondern der abrupt beendet wird. Ein Zuhörer nimmt solche transienten Vorgänge als Kennzeichen von Direktschallquellen wahr, die erfindungsgemäß aus einem Umgebungssignal eliminiert werden, so dass den Umgebungslautsprechern ein erfindungsgemäß erzeugtes Umgebungssignal geliefert wird, das keine oder nur stark gedämpfte Transienten umfasst.

**[0023]** Erfindungsgemäß wird ferner sichergestellt, dass die Unterdrückung eines Transienten in dem Umgebungssignal nicht zu einer zu großen Amplitudenmodulation führt. Erfindungsgemäß wurde nämlich ferner herausgefunden, dass Variationen der Amplitude, also der Lautstärke, auch wenn sie nicht transient sind, also unter der Transientenschwelle liegen, wenn sie jedoch oberhalb einer bestimmten Variationsschwelle sind, vom Benutzer als störend und damit dann, wenn solche Amplitudenvariationen aufgrund einer einfachen Eliminierung eines Transienten in einem Umgebungssignal entstehen würden, vom Zuhörer als Artefakt oder Fehler erkannt werden würden.

**[0024]** Erfindungsgemäß wird somit in einem Untersuchungssignal ein Transientenzeitraum detektiert, in dem ein transienter Bereich im Untersuchungssignal vorhanden ist. Hierauf wird mit Hilfe eines Synthesesignalgenerators ein Synthesesignal für den Transientenzeitraum erzeugt, der ausgebildet ist, um das Synthesesignal so zu erzeugen, dass es einen flacheren zeitlichen Verlauf als das Untersuchungssignal in dem transienten Bereich aufweist, wobei ferner der Synthesesignalgenerator ausgebildet ist, um das Synthesesignal so zu erzeugen, dass es sich im Hinblick auf seine Intensität eines vorangehenden oder nachfolgenden Abschnitts des Untersuchungssignals um weniger als eine vorbestimmte Schwelle unterscheidet. Dieses erzeugte Synthesesignal wird dann mittels eines Signalsubstituierers anstatt des Untersuchungssignals im Transientenzeitraum verwendet, um das Umgebungssignal zu erhalten.

**[0025]** Erfindungsgemäß wird somit die Extraktion eines Umgebungssignal-artigen Signals aus einem Zwei-Kanal-Stereoeingangssignal verbessert, oder es wird eine Nachverarbeitung eines existierenden Signals, das z.B. bereits ein extrahiertes Roh-Umgebungssignal ist, vorgenommen. Im ersten Fall ist das Untersuchungssignal das tatsächliche Zwei-Kanal-Stereosignal bzw. jeweils ein Kanal des Zwei-Kanal-Signals, während im zweiten Fall das Untersuchungssignal bereits ein extrahiertes Umgebungssignal oder ein vorsynthetisiertes Umgebungssignal ist. Damit ist das erfindungsgemäße Konzept besonders nützlich für das Upmix-Konzept, das auch als "Direct-Ambience-Konzept" dargestellt worden ist. Auch für das "In-the-Band"-Konzept kann das erfindungsgemäße Konzept vorteilhaft sein, da es auch dort zu einem verbesserten Umgebungssignal führt, das einerseits keine störenden Artefakte mehr hat, das jedoch

andererseits noch genug Informationen umfasst, damit ein Benutzer einen Nutzen vom Umgebungssignal hat.

**[0026]** Die erfindungsgemäße Umgebungssignalerzeugung führt dazu, dass das Umgebungssignal keine relevanten Teile von Direktschallquellen hat, wobei insbesondere keine Transienten enthalten sind bzw. Transienten nur in sehr stark gedämpfter Form enthalten sind. Anderenfalls würde der Zuhörer Direktschallquellen hinter sich wahrnehmen, was im Konflikt zu der Erfahrung des Benutzers sein würde, der typischerweise Schallquellen nur von vorne wahrnimmt.

**[0027]** Ferner stellt das erfindungsgemäße Konzept sicher, dass das Umgebungssignal ein durchgehendes ununterbrochenes diffuses Tonsignal ist, da ein unterbrochener umgebungs-artiger Ton, der beispielsweise erhalten wird, wenn Transienten einfach komplett eliminiert werden würden, vom Benutzer als unangenehm oder sogar als Fehler im Hochmisch-Prozess wahrgenommen werden würde.

**[0028]** Bei einem bevorzugten Ausführungsbeispiel der vorliegenden Erfindung wird, um einen Direkt-Ambience-Typ eines Upmix-Prozesses zu erreichen, ein ambience-artiges Signal für die hinteren Kanäle aus dem Stereosignal extrahiert. Um dies zu erreichen, wird beispielsweise nur auf die unkorrelierten Signalkomponenten oder als einfache Lösung einfach auf die Differenz zwischen dem ursprünglichen rechten und linken Kanal zurückgegriffen. Wenn die hinteren Kanäle derart erzeugt werden, haben sie oft transienten-artige Komponenten von Direktschallquellen. Diese Transienten können Töne sein, wie beispielsweise Notenanfänge oder Teile von perkussiven Instrumenten. Ein Transient, der hinter dem Zuhörer wahrgenommen wird, während eine Direktschallquelle (zu der der Transient typischerweise gehört) vor dem Zuhörer positioniert ist, hat eine negative Auswirkung auf die Lokalisierung der Direktschallquelle. Die Direktschallquelle erscheint somit entweder breiter als das Original oder wird, was noch schädlicher ist, als unabhängige Direktschallquelle hinter dem Benutzer wahrgenommen, wobei beide Effekte insbesondere für das Direkt-Ambience-Konzept sehr unerwünscht sind.

**[0029]** Erfindungsgemäß wird diese Problematik dadurch angegangen, dass Transienten im umgebungsartigen Signal unterdrückt werden, und dass die Auswirkung dieser Unterdrückung auf das restliche Signal minimiert wird, d.h. dass die Kontinuität des Signals bewahrt wird, indem nur begrenzte Intensitäts-Variationen für den Transientenzeitraum zugelassen werden.

**[0030]** Beim bevorzugten Ausführungsbeispiel der vorliegenden Erfindung wird das für den Transientenzeitraum erzeugte Signal, bevor es durch den Signalsubstituierer verwendet wird, mit dem ursprünglich im Transientenzeitraum vorhandenen Signal gemischt, was beispielsweise durch eine überlappende Verarbeitung erreicht wird. Alternativ oder zusätzlich kann, um die Diskontinuitäten an den Rändern des Transientenzeitraums zu unterdrücken bzw. wenigstens zu reduzieren, ein Überblenden vorgenommen werden, um in einem Über-

blendbereich langsam von dem Signal vor dem Transientenzeitraum in das Signal im Transientenzeitraum überzublenzen oder um aus dem Transientenzeitraum wieder langsam auszublenden.

**[0031]** Insbesondere die Ausblendung vom Transientenzeitraum in das ursprüngliche Signal, wenn kein Transient mehr detektiert wird, ist für einen artefaktfreien Höreindruck bevorzugt, da sichergestellt werden soll, dass dann, wenn ein Nicht-Artefakt-behaftetes Untersuchungssignal vorliegt, durch den Übergang vom Synthesesignal in das ursprüngliche Untersuchungssignal kein Knacken oder etwas Ähnliches entsteht.

**[0032]** Bei weiteren bevorzugten Ausführungsbeispielen der vorliegenden Erfindung wird eine Manipulation des Signals im Transientenzeitraum im Frequenzbereich vorgenommen, indem Vorzeichen von Spektralwerten oder allgemeiner gesagt Phasen von Spektralwerten randomisiert werden, was unweigerlich zu einer Glättung der zeitlichen Feinstruktur dieses im Frequenzbereich manipulierten Signals führt. Eine weitere Spektralverarbeitung besteht darin, eine Prädiktion über der Frequenz der Spektralwerte durchzuführen und dann die Prädiktionspektralwerte als Spektralwerte des Synthesesignals zu verwenden, da die Prädiktion über der Frequenz zu einer Glättung des korrespondierenden Zeitsignals führt. Zur Unterdrückung von Transienten bei gleichzeitiger Beibehaltung oder nur geringer Beeinflussung, es wird bevorzugt, die Intensität des Transientenzeitraums um höchstens plus oder minus 50% zu ändern, bestehen darin, die Änderung der Spektralwerte von einem Block zu einem nächsten zu limitieren, wobei diese Limitierung global, also für alle Spektralwerte gleich oder selektiv, also nur für bestimmte Spektralwerte, die eine besonders große Änderung haben, erfolgen kann.

**[0033]** Bevorzugte Ausführungsbeispiele der vorliegenden Erfindung werden nachfolgend bezugnehmend auf die beiliegenden Zeichnungen detailliert erläutert. Es zeigen:

- Fig. 1 ein Blockschaltbild der erfindungsgemäßen Vorrichtung zum Erzeugen eines Umgebungssignals;
- Fig. 2a eine schematische Darstellung der Blockverarbeitung bei nicht-überlappenden Blöcken, jedoch mit Überblendbereich;
- Fig. 2b eine schematische Darstellung der Synthesesignalerzeugung bei überlappenden Blöcken;
- Fig. 3 eine spezielle Implementierung einer Überblendung mit einer Einblendfunktion und einer Ausblendfunktion, die für Fig. 2a oder Fig. 2b eingesetzt werden kann;
- Fig. 4 ein Blockschaltbild einer bevorzugten Implementierung mit einer Verarbeitung im Frequenzbereich;

- Fig. 5a eine alternative Implementierung der Frequenzbereichsverarbeitung;
- Fig. 5b eine wieder alternative Frequenzbereichsverarbeitung;
- Fig. 5c eine bevorzugte Implementierung einer Intensitäts-basierten Verarbeitung;
- Fig. 6 eine Implementierung zur Erhaltung tonaler Bereiche im Synthesesignal;
- Fig. 7 ein Blockschaltbild einer bevorzugten Ausführungsform basierend auf dem Hochfrequenzgehalt HFC;
- Fig. 8 eine bevorzugte Implementierung der erfindungsgemäßen Vorrichtung mit zusätzlicher Funktionalität zum Erzeugen der Direktschallkanäle L, R, C;
- Fig. 9 ein Stereo-Wiedergabe-Szenario;
- Fig. 10 ein Multikanal-Wiedergabe-Szenario, bei dem alle Direktschallquellen durch die vorderen Kanäle wiedergegeben werden; und
- Fig. 11 ein Multikanal-Niedergabe-Szenario, bei dem Schallquellen auch durch hintere Kanäle wiedergegeben werden können.

**[0034]** Fig. 1 zeigt eine erfindungsgemäße Vorrichtung zum Erzeugen eines Umgebungssignals 10, das zur Ausstrahlung über Lautsprecher geeignet ist, für die kein eigenes Lautsprechersignal übertragen worden ist. Solche Lautsprecher sind typischerweise die hinteren Lautsprecher oder Surround-Lautsprecher, wie sie bei Ls, Rs in Fig. 10 und Fig. 11 beispielsweise gezeigt sind.

**[0035]** Die in Fig. 1 gezeigte Vorrichtung umfasst einen Transienten-Detektor 11 zum Detektieren eines Transientenzeitraums (bei 20 in Fig. 2 gezeigt), in dem ein Untersuchungssignal einen transienten Bereich aufweist. Obgleich hier einige Implementierungen des Transienten-Detektors beschrieben sind, sei darauf hingewiesen, dass auch beliebige andere Methoden zur Transientendetektion eingesetzt werden können, wie sie beispielsweise in einem MPEG4-Audiocodierer anzutreffen sind, bei dem abhängig von einer Transientendetektion von kurzen auf lange Fenster umgeschaltet wird. Auch in anderen Gebieten der Audiosignalverarbeitung werden Transienten-Detektoren eingesetzt, die schnelle und starke Änderungen der Hüllkurve eines Zeitsignals detektieren können. Beispielhafte zu detektierende Größenordnungen sind Änderungen der Hüllkurve, die in einem Zeitraum von 1 ms Änderungen von gleich oder über 100 % der Amplitude der Hüllkurve betreffen.

**[0036]** Der Transienten-Detektor 11 ist mit einem Synthesesignalgenerator 12 gekoppelt, der ausgebildet ist,

um ein Synthesesignal 13 zu erzeugen, das die beiden Bedingungen erfüllt, nämlich einerseits die Transientenbedingung und andererseits die Kontinuitätsbedingung. Die Transientenbedingung besteht darin, dass das Synthesesignal einen flacheren zeitlichen Verlauf als das Untersuchungssignal in dem transienten Bereich aufweist, während die Kontinuitätsbedingung darin besteht, dass die Intensität des Synthesesignals im transienten Bereich von einer Intensität eines vorhergehenden oder nachfolgenden Abschnitts des Untersuchungssignals um weniger als eine voreingestellte Schwelle abweicht. Vorzugsweise ist die Schwelle eine relative Schwelle und liegt bei einem Wert = 2,5, wobei Werte = 1,5 sogar bevorzugt werden. Dies bedeutet, dass die Intensität des Signals im transienten Bereich höchstens das 1,5-fache oder das 0,66-fache der Intensität eines vorausgehenden nicht-transienten Abschnitts oder nachfolgenden nicht-transienten Abschnitts des Untersuchungssignals ist. Damit wird sichergestellt, dass eine Transientenunterdrückung nicht zu einer störenden Amplituden-Variation bzw. Intensitätsvariation führen wird.

**[0037]** Die Schwelle kann auch durch ein Vertrauensintervall von 80% oder weniger, das anhand der Vergangenheitswerte bestimmt wird, realisiert werden.

**[0038]** Intensitätsmaße, die für die vorliegende Erfindung eingesetzt werden können, umfassen die Energie, die durch Addition der Abtastwert-Quadrate oder Spektralwert-Quadrate eines Blocks erhalten wird, oder ein Leistungsmaß, das unter Berücksichtigung der zeitlichen Blocklänge erhalten werden kann, oder auch ein Maß, das Beträge von Spektralwerten in einem Band gewichtet oder ungewichtet aufaddiert, wobei dieses spezielle Maß, das ebenfalls eine Intensität darstellt, auch als Hochfrequenzgehalt bezeichnet wird, wenn das Band, in dem aufaddiert wird, das obere Frequenzband des Untersuchungssignals ist oder allgemein höhere Frequenzen gegenüber niedrigeren Frequenzen stärker gewichtet werden bzw. einen stärkeren Einfluss auf das Endergebnis haben.

**[0039]** Der Synthesesignalgenerator erzeugt dann ein Synthesesignal, das von einem Signalsubstituierer 14 verwendet wird, um das Synthesesignal statt des entsprechenden Bereichs des ursprünglichen Untersuchungssignals zu verwenden, um schließlich das Umgebungssignal 10 zu liefern. Der Signalsubstituierer 14 erhält somit neben dem Synthesesignal über die Leitung 13 auch das Untersuchungssignal über eine Leitung 15, wie es in Fig. 1 angedeutet ist. Der Transienten-Detektor 11 erhält über eine Eingangsleitung 16 das Untersuchungssignal und liefert über eine Ausgangsleitung 17 eine Transienten-Information zum Synthesesignalgenerator 12, damit dieser unter Verwendung des Untersuchungssignals, das ihm über eine Leitung 18 bereitgestellt wird, das Synthesesignal erzeugt.

**[0040]** Bei speziellen Ausführungsbeispielen der vorliegenden Erfindung wird eine nicht-überlappende Blockverarbeitung, wie in Fig. 2a dargestellt oder eine überlappende Blockverarbeitung wie in Fig. 2b dargestellt ein-

gesetzt. Bei der nicht-überlappenden Blockverarbeitung in Fig. 2a wird ein Untersuchungssignal 21 in vorzugsweise gleich lange Blöcke einer speziellen Blocklänge eingeteilt. Der Transienten-Detektor erfasst dann einen Transienten 22 in dem Transientenzeitraum 20. Der Transient 22 liegt somit im Transientenzeitraum 20 von Fig. 2a, was dazu führt, dass der Transienten-Detektor 11 ein Ausgangssignal über seine Ausgangsleitung 17 liefert, die dem Synthesesignalgenerator 12 mitteilt, dass er nun mit einer Signalsynthese beginnen muss. Während die dem Transientenzeitraum 20 vorausgehenden und nachfolgenden Blöcke bis auf eine Überblendung in einem Überblendbereich 23 unmittelbar die entsprechenden Teile des Umgebungssignals 10 darstellen, wird nunmehr der Block des Untersuchungssignals, der dem Transientenzeitraum 20 entspricht, durch den Synthesesignalgenerator synthetisiert und dann durch den Signalsubstituierer 14 anstatt des ursprünglichen Blocks des Untersuchungssignals im Umgebungssignal verwendet.

**[0041]** Bei den bevorzugten Ausführungsbeispielen wird, wie es später noch dargestellt wird, eine Verarbeitung des Blocks des Untersuchungssignals vorgenommen, die im Frequenzbereich stattfindet. Dies führt dazu, dass an einer Blockgrenze das Synthesesignal einen Abtastwert hat, der sich von einem Abtastwert, der der letzte Abtastwert des vorausgehenden Blocks im Untersuchungssignal ist, deutlich unterscheiden kann. Um solche Blockgrenzen-Artefakte, die auftreten können, zu eliminieren, wird es bei dem in Fig. 2a gezeigten Ausführungsbeispiel bevorzugt, von einem Block vor einem Transientenzeitraum in das Synthesesignal im Transientenzeitraum überzublendern, indem beispielsweise der erste Abtastwert des generierten Synthesesignals zu den z.B. letzten 10 Abtastwerte des vorhergehenden Blocks, die gemäß der Ausblendfunktion gewichtet sind, hinzuaddiert wird, und zwar beispielsweise gemäß der Einblendfunktion in Fig. 3. Gleichzeitig wird der letzte Abtastwert des vorhergehenden Blocks noch gemäß der Ausblendfunktion in Fig. 3 zu dem ersten bzw. den dem ersten Abtastwert folgenden Abtastwerten des synthetisierten Blocks die gemäß der Einblendfunktion gewichtet sind, im Transientenzeitraum hinzuaddiert, um eine Überblendung zu schaffen. Entsprechend kann im hinteren Überblendbereich, wenn also vom Transientenzeitraum wieder zurück in den nicht-Transienten-behafteten Block des Umgebungssignals übergegangen wird, vorgegangen werden.

**[0042]** Um solche Blockgrenzen-Artefakte noch weiter zu reduzieren, wird, wie es in Fig. 2b gezeigt ist, eine überlappende Verarbeitung bevorzugt. Der Transienten-Detektor detektiert dann bei dem in Fig. 2b gezeigten Ausführungsbeispiel Blockbereiche, die mit eingeringelten Zahlen (1), (2), (3), (4), (5), (6) dargestellt sind. Ein Transient wird bei 22 detektiert. Dies führt dazu, dass es im Vergleich zu Fig. 2a einen größeren Transienten-Zeitraum 20 gibt, da der Transient an der Position 22 sowohl im Block 4 als auch im Block 5 detektiert worden ist. Da-

her wird der Synthesesignalgenerator 12 von Fig. 1 Synthesesignale sowohl für den Block 4 als auch für den Block 5 erzeugen. Während für die den drei Transientenzeitraum-Bereichen A, B, C vorausgehenden Blöcke das Untersuchungssignal keine Transienten hat und somit unmittelbar in das Umgebungssignal übernommen wird, werden die Bereiche A, B, C durch den Signalsubstituierer 14 von Fig. 1 ersetzt, und zwar durch die von den Synthesesignalgeneratoren erzeugten Abschnitte A, B, C. Der Abschnitt A wird durch die Addition der zweiten Hälfte des nicht-Transienten-behafteten Blocks 3 des Untersuchungssignals mit der ersten Hälfte des für den Block 4 erzeugten Synthesesignals erzeugt. Der zweite Teil B des Transientenzeitraums 20 wird durch Addition der zweiten Hälfte des Synthesesignals, das für den Block 4 erzeugt worden ist, mit der ersten Hälfte des Synthesesignals, das für den Block 5 erzeugt worden ist, geliefert und vom Signalsubstituierer als entsprechender Abschnitt des Umgebungssignals 10 substituiert. Der dritte Teil C des Transientenzeitraums 20 wird durch Addition der zweiten Hälfte des vom Synthesesignalgenerator erzeugten Blocks Nr. 5 mit der ersten Hälfte des Blocks Nr. 6, der nicht mehr Transienten-behaftet ist, erzeugt und vom Signalsubstituierer 14 in das Umgebungssignal geschrieben.

**[0043]** Die in Fig. 3 gezeigte Ausblendfunktion wird nachfolgend detaillierter erläutert. So kann diese Ausblendfunktion dazu verwendet werden, um bei der Blockverarbeitung mit nicht-überlappenden Blöcken einen weichen Blockübergang von einem nicht-synthetisierten Block zu einem synthetisierten Block zu schaffen und ferner einen weichen Übergang von einem synthetisierten Block wieder in einen nicht-synthetisierten Block zu schaffen. Alternativ kann eine entsprechende Überblendfunktion auch dazu verwendet werden, um insbesondere dann, wenn durch eine bestimmte spezifizierte Anzahl von Blöcken ein Synthesesignal erzeugt worden ist, wieder zurück zum ursprünglichen Untersuchungssignal überzublenden. Nachdem die Wahrscheinlichkeit gegeben ist, dass das Synthesesignal sich aufgrund der Extrapolation relativ weit von dem Untersuchungssignal entfernt hat, würde ein abruptes Zurückkehren zum Untersuchungssignal in bestimmten Fällen zu hörbaren Artefakten führen. Daher wird es bevorzugt, langsam gemäß der Einblend-/Ausblend-Funktion von Fig. 3 überzublenden, indem für einen Block, in dem bereits kein Transient mehr detektiert worden ist, ein Synthesesignal erzeugt wird, das zu 90% aus dem letzten synthetisierten Block und zu 10% aus dem aktuellen Untersuchungsblock besteht. Im nächsten Block könnte das Verhältnis dann auf 80%:20% umgeändert werden, bis dann nach einer bestimmten Anzahl von Blöcken das Synthesesignal vollständig ausgeblendet ist und das aktuelle nicht-Transientenbehaftete Untersuchungssignal wieder vollständig eingeblendet ist.

**[0044]** Nachfolgend wird eine bevorzugte Implementierung eines Teils des Synthesesignalgenerators 12 anhand von Fig. 4 gestellt. Hierzu wird das Zeitsignal, das

einen Block des Untersuchungssignals darstellt, in eine Frequenzbereichsdarstellung oder eine Subband-Darstellung durch einen Umsetzer 40 umgesetzt, der eine Transformation oder eine Analyse-Filterbank umfassen kann. Die spektrale Darstellung in Form von Spektralkoeffizienten oder die Subband-Signale werden dann, wie es bei 41 dargestellt ist, gegebenenfalls um Informationen über eine extrapolierte spektrale Darstellung bzw. extrapolierte Subband-Signale ersetzt, wenn es sich um einen Block des Zeitsignals handelt, in dem ein Transient detektiert worden ist. Hierauf wird die spektrale Darstellung gegebenenfalls unter Verwendung von zusätzlichen Informationen aufgrund einer Extrapolation einem Glätter 42 zugeführt, der die Spektralwerte derart beeinflusst, dass der zeitliche Verlauf des zugrunde liegenden Signals geglättet wird. Im Falle einer Filterbank wird der Glätter 42 die Subband-Signale so beeinflussen, dass der zeitliche Verlauf des den Subband-Signalen zugrunde liegenden Signals glatter als vor der Glättung ist. Dann, in einem Block 43 findet eine inverse Umsetzung in den Zeitbereich statt, wobei entweder eine Rücktransformation oder eine Synthese-Filterbank eingesetzt wird, um schließlich zu einem Zeitsignal 44 zu kommen, das einen glatteren Verlauf hat als das Zeitsignal am Eingang der Stufe 40, das jedoch eine Energie hat, die durch die Glättung nicht erheblich beeinflusst worden ist. Ferner ist die Glättung so vorgenommen worden, dass die Energie des geglätteten Zeitsignals 44 sich von der Energie des vorherigen Zeitsignals nicht mehr als die Schwelle unterscheidet.

**[0045]** So kann bei der vorliegenden Erfindung eine insgesamt E-nergiemanipulation der Energie des Zeitsignals zwar stattfinden. So wird jedoch lediglich der Transient gedämpft, während die tonalen Anteile weiterlaufen bzw. aus der Vergangenheit synthetisiert werden, indem das Signal im Transientenzeitraum durch eine Prädiktion unter Verwendung eines nicht transienten Signals aus der Vergangenheit synthetisiert wird.

**[0046]** Wenn die Energie - wie bei der Randomisierung oder spektralen Prädiktion - jedoch nicht angetastet wird, hat die Glättung dazu geführt, dass die Energie über dem Block gleichmäßiger verteilt wird, so dass ein glatterer zeitlicher Verlauf erzeugt worden ist, ohne jedoch die Energie des Blocks von Abtastwerten des Untersuchungssignals erheblich zu ändern. Dies ist in den meisten Fällen ausreichend und stellt sicher, dass der Benutzer ein Untersuchungssignal hört, das die Kontinuitätsbedingung immer erfüllt. Erst wenn die Transiente zu einer erheblichen Energiezunahme auf den ganzen Block betrachtet führt, wird die Glättung allein, also die gleichmäßigere Verteilung der Energie über dem Block, nicht mehr ausreichend sein und es kann dann eine gesteuerte Signalbegrenzung vorgenommen werden.

**[0047]** Bekannte Verfahren, die darin bestehen, eine Lokalisierung von Direktschallquellen in den hinteren Kanälen zu vermeiden, bestehen darin, die hinteren Kanäle ein paar Millisekunden zu verzögern. Diese Lösung führt zu keiner Transientenunterdrückung, versucht jedoch

die Transienten durch Verwenden des Präzedenz-Effekts zu "maskieren". Der Präzedenz-Effekt besteht darin, dass das Ohr dort eine Schallquelle vermutet, wo es zuerst etwas von dieser Schallquelle hört, wobei das, was man dann von dieser Schallquelle hört, durchaus lauter oder von einer anderen Richtung kommen kann. Nachteilhaft an dieser Lösung ist jedoch, dass sehr kurze Schallereignisse mit scharfen Transienten oft immer noch hörbar sind und dann zweimal wahrgenommen werden, und zwar durch einen vorderen Lautsprecher und einige Millisekunden später durch die hinteren Kanäle, was einen unangenehmen Höreindruck bewirkt.

**[0048]** Kommerziell verfügbare Matrix-Decodierer, wie beispielsweise Dolby Pro Logic II oder Logic 7 haben die Fähigkeit, nicht vorverarbeitete 2-Kanal-Stereo-Dateien in Multikanal-Surround-Dateien hochzumischen (upmixen), obwohl sie unmittelbar nicht für diese Aufgabe entworfen worden sind. Diese Matrix-Decodierer sind oft nicht in der Lage, transiente Töne in den hinteren Kanälen zu unterdrücken, was in einem Signal resultiert, das die Anforderungen nach Transientenfreiheit und Amplituden- bzw. Intensitäts-Kontinuität nicht erfüllt.

**[0049]** Erfindungsgemäß werden dagegen Kanalbereiche, wo Transienten auftreten, erfasst und gedämpft. Eine einfache Dämpfung des gesamten Signals zu diesen Zeitbereichen würde jedoch in einer Amplitudenmodulation des Umgebungssignals resultieren und es würde als unangenehme oder sogar als Artefakt wahrgenommen werden. Dies würde daher das Qualitätsempfinden des extrahierten oder verarbeiteten Umgebungssignals verschlechtern. Um diesen unangenehmen Amplitudenmodulationseffekt zu überwinden, wird erfindungsgemäß eine Transientenunterdrückung ohne Beeinträchtigung der Kontinuität des Synthesesignals bzw. Umgebungssignals erzeugt. Hierzu wird ein Eingangssignal, das z.B. ein hochgemischtes Signal, wie es durch einen Matrix-Hochmischer erhalten wird, für die hinteren Kanäle, oder es wird ein Signal mit ähnlichen Charakteristika und einem ähnlichen Anwendungsbereich analysiert, um zu erfassen, ob eine Transiente vorhanden ist.

**[0050]** Wenn eine Transiente erfasst wird, wird der gegenwärtig verarbeitete Block durch ein Substitutionssignal ersetzt, das eine flache (nicht-transiente) zeitliche Hüllkurve hat. Dieses Substitutionssignal wird entweder von vorhergehenden Signalabschnitten, in denen keine Transiente vorhanden war, erzeugt oder wird von dem gegenwärtig verarbeiteten Block durch einen Verarbeitungsschritt erzeugt, der die zeitliche Hüllkurve bzw. Feinstruktur des Signals flacher macht oder wird von einer Kombination beider Verfahren erzeugt.

**[0051]** Das Substitutionssignal, das von vorhergehenden Abschnitten erzeugt wird, wird beispielsweise durch Extrapolation vorheriger Energiepegel des Signals oder durch Kopieren/Wiederholen von vorherigen Signalabschnitten ohne einen transienten Bereich des Signals erzeugt.

**[0052]** Ein Abflachen oder "Flattening" der zeitlichen Feinstruktur oder des feinen Zeitsignals auf der Basis

des gegenwärtig verarbeiteten Blocks kann beispielsweise derart durchgeführt werden, wie es nachfolgend bezugnehmend auf die Figuren 5a, 5b oder 5c dargestellt ist.

5 **[0053]** So können die Absolutwerte der Spektralkoeffizienten innerhalb eines begrenzten Bereichs, der sich um die extrapolierten Spektralkoeffizienten oder Beträge derselben erstreckt, randomisiert werden, wie es später noch in Verbindung mit Fig. 5c dargestellt wird.

10 **[0054]** Alternativ oder zusätzlich können die Phasen bzw. Vorzeichen der Spektralkoeffizienten des verarbeiteten Blocks, in dem sich der Transient befindet, durch einen Randomisierer 50 randomisiert werden. Hierzu wird ein Kurzzeitspektrum des betrachteten Blocks des Untersuchungssignals erzeugt, und die dort erhaltenen komplexen Spektralwerte werden nach Betrag und Phase berechnet, um dann die Phasen der Spektralwerte zu randomisieren. Wird eine Transformation eingesetzt, die nur Phasen von  $\pm 180^\circ$  auflösen kann, die also Spektralwerte nur mit positiven und negativen Vorzeichen liefern kann, so können auch die Vorzeichen randomisiert werden, um ein Kurzzeitspektrum mit randomisierten Phasen/Vorzeichen zu erhalten, das einen flacheren zeitlichen Verlauf des korrespondierenden Zeitsignals hat.

25 **[0055]** Dieser Ansatz basiert darauf, dass eine schnelle Änderung in einem Zeitsignal nur dann möglich ist, wenn die Phasen der diesem transienten Bereich zugrunde liegenden Grundwelle und zugehörigen Oberwellen in einem ganz bestimmten Verhältnis sind. Wird eine Randomisierung der Phasen erreicht, so führt dies dazu, dass der transiente Bereich geglättet wird, da das genaue Zusammenspiel der Phasen der einzelnen Sinusschwingungen, die durch die Spektralwerte abgebildet werden, nicht mehr vorhanden ist.

30 **[0056]** Eine alternative Implementierung ist in Fig. 5b anhand eines Prädiktors 51 dargestellt, der ausgebildet ist, um eine Prädiktion des Kurzzeitspektrums über der Frequenz durchzuführen. Ein solcher Prädiktor ist in J. Herre, J.D. Johnston: "Exploiting Both Time and Frequency Structure in a System that Uses an Analysis/Synthesis Filterbank with High Frequency Resolution", 103rd AES Convention, New York 1997, Preprint 4519 dargestellt.

40 **[0057]** Wieder wird ein Kurzzeitspektrum erzeugt, das in seinem zugeordneten Zeitsignal einen transienten Verlauf hat. Typischerweise unter Verwendung eines Open-Loop-Prädiktors wird ein aktueller Spektralwert des Kurzzeitspektrums mittels eines vorherigen oder einer Mehrzahl von vorherigen Spektralwerten vorhergesagt, wobei dann der vorhergesagte Spektralwert von dem tatsächlichen Spektralwert subtrahiert werden könnte, um einen spektralen Restwert zu erhalten. Während der spektrale Restwert einer typischen Prädiktion über der Frequenz den Wert darstellt, der von Interesse ist und zusammen mit Koeffizienten eines Prädiktionsfilters informationstragend ist, wird erfindungsgemäß ein bestimmtes Prädiktionsfilter voreingestellt, und werden

die Spektralwerte des Kurzzeitspektrums durch die unter Verwendung dieses Prädiktionsfilters prädierten Spektralwerte ersetzt, während das Prädiktionsfehlersignal nicht weiter verwendet wird.

**[0058]** Die damit erhaltenen eigentlich fehlerhaften Prädiktions-Spektralwerte haben jedoch nunmehr einen flacheren zeitlichen Verlauf als das ursprüngliche Kurzzeitspektrum, haben jedoch noch annähernd soviel Energie, so dass sowohl die Transientenbedingung als auch die Kontinuitätsbedingung, wie sie in Verbindung mit dem Synthesesignalgenerator 12 von Fig. 1 dargestellt worden ist, erfüllt ist. Eine bevorzugte einfache Ausgestaltung des Prädiktionsfilters besteht einfach darin, dass ein Wert einer im Index niedrigeren Spektrallinie als Prädiktionswert für eine aktuelle Spektrallinie verwendet wird.

**[0059]** Allgemein kann das extrapolierte Signal mit dem ursprünglichen Signal nach einer spezifizierten Zeitdauer überblendet werden, anstatt einer harten Umschaltung, um Langzeit-Extrapolationsartefakte zu vermeiden.

**[0060]** Ferner wird es bevorzugt, wie es anhand von Fig. 6 dargestellt ist, tonale Anteile/Bänder durch einen Detektor 60 zu delektieren und durch den Synthesesignalgenerator nicht zu beeinflussen, sondern in einem Mischer/Kombinierer 61 mit Synthesesignalen für transiente Bänder zu kombinieren, um dann, nach einer Transformation bzw. Umsetzung in den Zeitbereich, die im Block 61 stattfinden könnte, ein Zeitsignal mit flacherem zeitlichem Verlauf zu erhalten, das jedoch noch die tonalen Bänder, also Anteile, die nicht transient waren, in unveränderter Gestalt umfasst.

**[0061]** Somit werden stationäre/tonale Frequenzkomponenten im Eingangssignal, die während der Zeitdauer des Transienten beispielsweise in lediglich Teilen des Spektrums vorhanden waren, erfasst und es wird ein Substitutionssignal erzeugt, das aus einer Extrapolation der vergangenen stationären/tonalen Signalkomponenten und den im aktuellen Block erfassten stationären/tonalen Frequenzkomponenten besteht.

**[0062]** Nachfolgend wird anhand von Fig. 5c eine Implementierung der vorliegenden Erfindung, die durch einen impliziten und nicht mehr expliziten Transienten-Detektor auskommt, dargestellt. Hierzu ist eine Einrichtung 53 zum Berechnen der Intensität eines Blocks und eines vorherigen Blocks in Fig. 5c gezeigt. Ein Maß für die Intensität eines verarbeiteten Signalblocks ist beispielsweise die Energie oder der Hochfrequenzgehalt (HFC) oder ein anderes Maß, das auf der Basis der Spektralwerte, zeitlichen Abtastwerte, der Energie, der Leistung oder einem anderen Amplituden-bezogenen Maß des Signals basiert. Hierauf wird in einer Einrichtung 54 festgestellt, ob eine Intensität von einem Block zum nächsten über einer Schwelle ansteigt. Wenn dies der Fall ist, werden die Spektralwerte des verarbeiteten Blocks so begrenzt, dass ihre Intensität die Intensität des vorhergehenden Signalblocks nicht mehr als um die bestimmte relative oder absolute Schwelle überschreitet, derart,

dass zumindest die insgesamt Dominanz von Transienten reduziert wird. Diese Begrenzung findet in einer Einrichtung 55 statt, die ausgebildet ist, um dann, wenn ein Bedarf nach einer Begrenzung erfasst worden ist, wenn also implizit ein Transient detektiert worden ist, entweder Spektralwerte individuell oder global zu begrenzen. Eine individuelle Begrenzung würde darin bestehen, dass für Spektralwerte oder für Bänder ein Energieanstieg berechnet wird und die Spektralwerte bzw. die Energiebänder nur bis zu einem maximalen Energieanstieg ansteigen und darüber hinausgehende Werte abgeschnitten werden.

**[0063]** Die Einrichtung 55 zum Begrenzen der Spektralwerte begrenzt also individuell oder global die Spektralwerte, wobei eine individuelle Begrenzung darin besteht, dass nur die Spektralwerte, die oberhalb einer Schwelle ansteigen, begrenzt und vorzugsweise auf diese Schwelle begrenzt werden, während die anderen Spektralwerte, die nicht so stark ansteigen, nicht beeinflusst werden. Alternativ wird es jedoch in bestimmten Fällen günstiger sein und im Hinblick auf den Rechenaufwand einfacher sein, dann, wenn eine zu starke Zunahme festgestellt worden ist, sämtliche Spektralwerte um dasselbe absolute oder relative Maß zu begrenzen.

**[0064]** Hierauf wird es noch bevorzugt, eine Nachverarbeitung der begrenzten Spektralwerte mittels einer Einrichtung 56 zum Nachverarbeiten vorzunehmen, wobei diese Nachverarbeitung in einer Randomisierung, wie es in Fig. 5a beschrieben worden ist, oder auch in einer Prädiktion, wie es in Fig. 5b beschrieben worden ist, bestehen kann. Die Reihenfolge der Verarbeitung durch die Einrichtungen 55 und 56 kann auch umgekehrt werden, derart, dass zunächst eine Randomisierung bzw. eine Prädiktionsverarbeitung mit einem Block durchgeführt wird, für den eine Transiente delektiert worden ist, wobei erst dann eine Intensitäts-Limitierung bzw. Begrenzung gemäß der Verarbeitung im Block 55 vorgenommen wird. Im Hinblick auf Fig. 5c sei noch darauf hingewiesen, dass der Block Z/F eine Zeit/Frequenzbereich-Umsetzung 57 darstellt, wobei eine Umsetzung vom Zeit- in den Frequenzbereich auch eine Filterung mittels einer Analyse-Filterbank sein kann, derart, dass in diesem Fall die Spektraldarstellung aus Subbandsignalen und nicht einzelnen Spektralkomponenten besteht.

**[0065]** Nachfolgend wird ein speziell bevorzugtes Ausführungsbeispiel der vorliegenden Erfindung anhand von Fig. 7 dargelegt. Der Transientendetektor, wie er bei 11 in Fig. 1 gezeigt worden ist, umfasst bei diesem Ausführungsbeispiel eine Einrichtung 71 zum Berechnen des Hochfrequenz-Inhalts (HFC) für jeden Block, der eine Einrichtung zum Berechnen des Langzeit-HFC-72 nachgeschaltet ist. Ein Komparator 73 erfasst dann, ob es einen Transienten gibt, bzw. ob es einen Transienten-Zeitraum gibt, in dem ein Transient vorhanden ist. Insbesondere ist die Einrichtung 71 ausgebildet, um den gewichteten Hochfrequenz-Inhalt (HFC) für jeden Block des ursprünglichen linken Signals und des ursprüngli-

chen rechten Signals zu berechnen. Alternativ kann auch ein HFC für jeden Kanal für sich berechnet werden. Der HFC ist die gewichtete Summe der Absolutwerte aller Frequenzlinien in einem Block, mit zunehmenden Gewichtungsfaktoren von niedrigeren zu höheren Frequenzen. Der HFC berechnet sich folgendermaßen:

$$\text{HFC} = \text{Summe} (X(f) \cdot w(f)),$$

wobei  $X(f)$  die Spektralkoeffizienten für bestimmte Frequenzen sind, und wobei  $w(f)$  Gewichtungsfaktoren für bestimmte Frequenzen sind.

**[0066]** Dadurch, dass die Gewichtungsfaktoren von niedrigen zu höheren Frequenzen ansteigen, wird sichergestellt, dass im HFC-Wert die Energie in den höheren Frequenzkomponenten im Vergleich zu der Energie in den niedrigeren Frequenzkomponenten gewichtet wird. Eine Energie in höheren Spektralkomponenten ist nämlich ein besseres Indiz für einen Transient als eine Energie in niedrigeren Spektralkomponenten. In der Implementierung kann zur Berechnung des HFC auf sämtliche Spektralkomponenten zurückgegriffen werden. Alternativ kann die Berechnung des HFC auch ausgehend von einem Grenzwert durchgeführt werden, der etwa im mittleren Bereich des Spektrums liegt, so dass die niedrigen Spektralkoeffizienten in der Berechnung des HFC keine Rolle spielen.

**[0067]** Ferner wird ein Langzeit-HFC-Mittelwert, der als HFC' bezeichnet wird, über wenigstens drei und vorzugsweise fünf vorausgehende Blöcke berechnet. Wird dann in der Einrichtung 73 bestimmt, dass der HFC im aktuellen Block um einen Faktor mehr als einen konstanten Faktor  $c$  von dem Langzeit-Mittelwert HFC' abweicht, wobei als konstanter Faktor  $c$  eine Zahl  $>$  oder  $= 1,0$  verwendet wird, so wird ein Transient detektiert. Die Schwelle hängt von der Art des gleitenden Mittelwerts ab. Ist der gleitende Mittelwert eine Mittelwert, in dem die Vergangenheit gegenüber dem aktuelleren Block stärker gewichtet ist, also ein langsamer Mittelwert, so ist die Schwelle näher bei Eins als in dem Fall, in dem die Vergangenheit weniger stark in den gleitenden Mittelwert eingeht. Hier würde die Schwelle weiter weg von Eins liegen.

**[0068]** Wenn ein Transient detektiert wird, wie es einer Einrichtung 74 zum Berechnen des Mittelwerts von der Einrichtung 73 signalisiert wird, wird der Mittelwert der vergangenen Absolutwerte jeder Frequenzlinie (Spektralkoeffizient) über einem definierten Zeitintervall, wie beispielsweise fünf Blöcken, berechnet. Zusätzlich wird ein Prädiktions-Zulässigkeits-Intervall  $\Delta_{\text{max}}$  für die extrapolierten Absolutwerte berechnet. Die extrapolierten Absolutwerte variieren zufällig innerhalb dieses Intervalls  $\Delta_{\text{max}}$ . Um dies zu erreichen, wird eine Berechnung gemäß einer Gleichung durchgeführt, wie sie bei der Einrichtung 75 in Fig. 7 gezeigt ist. RN steht für Zufallszahl,  $\Delta_{\text{max}}$  stellt das Zulässigkeitsintervall dar, SW ist ein Spek-

tralwert, wie er durch die Einrichtung 75 zum Berechnen berechnet wird, und  $SW_m$  ist der Spektralwert, der sich als Mittelwert aus mehreren vergangenen Blöcken ergibt, wie er durch den Block 74 berechnet worden ist. Die Einrichtung 75 ist daher ausgebildet, um folgende Gleichung auszuwerten:

$$SW = SW_m + RN \cdot \Delta_{\text{max}}.$$

**[0069]** Um Wiederholungseffekte zu vermeiden, die auftreten können, wenn ein detektierter Transient zu lang ist, werden die extrapolierten Werte mit den ursprünglichen Werten überblendet, und zwar dann, wenn ein festes Zeitintervall verstrichen ist, wie beispielsweise, dass drei Blöcke von Synthesignalen vorhanden waren, von denen dann wieder auf das ursprüngliche Signal zurückgegangen werden muss. Wird der Transientenzeitraum jedoch kürzer als drei Blöcke sein, so wird es bevorzugt, das Überblenden nicht durchzuführen, da dann davon ausgegangen wird, dass sich die extrapolierten Signale noch nicht so stark von den ursprünglichen Signalen entfernt haben. Das Überblenden kann entweder vor einer Umsetzung in den Zeitbereich oder vorzugsweise nach einer Umsetzung in den Zeitbereich stattfinden, wie es bei 76 in Fig. 7 dargestellt ist, um das Synthesignal zu erhalten.

**[0070]** In einer Implementierung kann das erfindungsgemäße Konzept in einen Extraktionsprozess eines Umgebungssignals integriert werden oder als separater Nachverarbeitungsschritt unter Verwendung eines existierenden Umgebungssignals verwendet werden, das jedoch immer noch vor der erfindungsgemäßen Verarbeitung unerwünschte Transienten umfasst.

**[0071]** Die erfindungsgemäßen Verarbeitungsschritte können im Frequenzbereich pro Frequenzlinie oder in Subbändern durchgeführt werden. Sie können jedoch auch nur teilweise im Frequenzbereich typischerweise oberhalb eines bestimmten Frequenzlimits vorgenommen werden, oder aber auch im Zeitbereich ausschließlich oder in einer Kombination von Zeit- und Frequenzbereich durchgeführt werden.

**[0072]** Fig. 8 zeigt ein bevorzugtes Ausführungsbeispiel der vorliegenden Erfindung, bei dem die Vorrichtung zum Erzeugen eines Umgebungssignals nicht nur ausgebildet ist, um Umgebungssignale für einen Ausgang 80 für einen linken Umgebungskanal und einen Ausgang 81 für einen rechten Umgebungskanal zu erzeugen. Die erfindungsgemäße Vorrichtung umfasst ferner zusätzlich einen Hochmischer 82 zum Erzeugen von Signalen für den linken Kanal L, den rechten Kanal R, den Mitten-Kanal C und vorzugsweise auch noch für den LFE-Kanal, wie es in Fig. 8 gezeigt ist. Sowohl die Kombination aus Transienten-Detektor 12, Synthesegenerator 14 und Signalsubstituierer 16, also der Hochmischer 82 werden von einem Decodierer 84 gespeist. Der De-

codierer 84 ist ausgebildet, um einen Bitstrom 85 zu empfangen und zu verarbeiten, um ausgangsseitig ein Mono-Signal oder ein Stereo-Signal 86 zu liefern. Der Bitstrom kann ein MP3-Bitstrom oder eine MP3-Datei sein, oder er kann eine AAC-Datei sein, oder kann auch eine Darstellung eines parametrisch codierten Multikanalsignals sein. So könnte der Bitstrom 85 beispielsweise eine parametrische Darstellung des linken Kanals, des rechten Kanals und des Center-Kanals sein, wobei ein Übertragungskanal und mehrere Cues für den zweiten und den dritten Kanal enthalten sind, wobei diese Verarbeitung aus der BCC-Multikanalverarbeitung bekannt ist. Dann wäre der Decodierer 84 ein BCC-Decodierer, der nicht nur ein Mono- oder ein Stereo-Signal liefert, sondern der sogar ein 3-Kanal-Signal liefert, das jedoch noch keine Daten über die beiden Surround-Kanäle Ls, Rs umfasst. In einer Implementierung wird das Untersuchungssignal somit in diesem Fall ein Mono-Signal, ein Stereo-Signal oder sogar ein Multikanal-Signal sein, das jedoch keine eigenen Lautsprecher-Signale für die Surround-Kanäle Ls, Rs umfasst.

**[0073]** Es sei darauf hingewiesen, dass man entweder dasselbe Ambiente-Signal für beide Surround-Kanäle oder für jeden Surround-Kanal ein eigenes Signal berechnen kann. Im ersten Fall wird z. B. das Untersuchungssignal bzw. das Umgebungssignal von einer Summe aus linkem und rechten Kanal abgeleitet. Im anderen Fall wird z. B. vom linken Kanal das Umgebungssignal für den linken Surround-Kanal berechnet und wird vom rechten Kanal das Umgebungssignal für den rechten Kanal berechnet.

**[0074]** Ein Aspekt der vorliegenden Erfindung bezieht sich auf eine Vorrichtung zum Erzeugen eines Umgebungssignals, das zur Ausstrahlung über Lautsprecher Ls, Rs geeignet ist, für die kein geeignetes Lautsprecher-Signal vorhanden ist, mit folgenden Merkmalen: einem Transienten-Detektor 11 zum Detektieren eines Transientenzeitraums 20, in dem ein Untersuchungssignal einen transienten Bereich 22 aufweist; einem Synthesesignalgenerator 12 zum Erzeugen eines Synthesesignals für den Transientenzeitraum 20, wobei der Synthesesignalgenerator 12 ausgebildet ist, um ein Synthesesignal zu erzeugen, das einen flacheren zeitlichen Verlauf als das Untersuchungssignal in dem Transientenzeitraum 20 aufweist, und dessen Intensität von einer Intensität eines vorangehenden oder nachfolgenden Abschnitts des Untersuchungssignals um weniger als eine vorbestimmte Schwelle abweicht; und einem Signalsubstituierer 14 zum Substituieren des Untersuchungssignals im Transientenzeitraum durch das Synthesesignal, um das Umgebungssignal zu erhalten.

**[0075]** Bei einem Ausführungsbeispiel ist die Vorrichtung für eine Blockverarbeitung ausgebildet, um aufeinander folgende Blöcke von zeitdiskreten Abtastwerten überlappend oder nicht-überlappend zu verarbeiten.

**[0076]** Bei einem Ausführungsbeispiel ist der der Transienten-Detektor 11 ausgebildet ist, um für aufeinander folgende Blöcke Intensitätswerte zu berechnen, und um

einen Transientenzeitraum 20 zu detektieren, wenn ein Intensitätswert eines Blocks von einem vorausgehenden oder nachfolgenden intensitätswert um mehr als eine vorbestimmte Transientenschwelle unterschiedlich ist.

**[0077]** Bei einem Ausführungsbeispiel ist der Synthesesignalgenerator 12 ausgebildet, um für einen Block im Transientenzeitraum 20 eine Mehrzahl von Spektralwerten, die ein Kurzzeitspektrum des Blocks darstellen, so zu begrenzen, dass ihre Intensität sich von der Intensität eines vorausgehenden oder nachfolgenden Blocks oder Transient um weniger als die vorbestimmte Schwelle unterschiedlich ist.

**[0078]** Bei einem Ausführungsbeispiel ist der Synthesesignalgenerator 12 ausgebildet, um komplexe Spektralwerte, die ein Kurzzeitspektrum des Blocks, der den Transientenzeitraum 20 umfasst, darstellen, im Hinblick auf ihre Phase oder ihr Vorzeichen zu randomisieren.

**[0079]** Bei einem Ausführungsbeispiel ist der Synthesesignalgenerator 12 ausgebildet, um eine Prädiktionsverarbeitung 51 über der Frequenz durchzuführen, um ein Prädiktions-Spektrum zu erhalten, dessen zugeordnetes Zeitsignal einen flacheren zeitlichen Verlauf als ein Zeitsignal hat, das einem Spektrum vor der Prädiktionsverarbeitung über der Frequenz zugeordnet ist.

**[0080]** Bei einem Ausführungsbeispiel ist der Transienten-Detektor 11 ausgebildet, um einen Hochfrequenz-Gehalt für einen Block des Untersuchungssignals zu berechnen 61, wobei der der Transienten-Detektor 11 ausgebildet ist, um den gewichteten HF-Gehalt mit einem gleitenden Mittelwert über einer Mehrzahl von vorausgehenden oder nachfolgenden Blöcken ohne Transiente zu vergleichen 73,

wobei der Transienten-Detektor 11 ausgebildet ist, um eine Transiente für einen Block zu detektieren, wenn der HF-Gehalt eines aktuellen Blocks den gleitenden Mittelwert um mehr als eine Schwelle  $c$  überschreitet.

**[0081]** Bei einem Ausführungsbeispiel ist der Transienten-Detektor ausgebildet, um eine Schwelle zu verwenden, die abhängig von der Art der Berechnung des gleitenden Mittelwerts gewählt ist und näher bei Eins liegt, wenn in den gleitenden Mittelwert die Vergangenheit stärker eingeht und weiter weg von Eins liegt, wenn in den gleitenden Mittelwert die Vergangenheit vergleichsweise schwächer eingeht.

**[0082]** Bei einem Ausführungsbeispiel ist der Synthesesignalgenerator ausgebildet, um für jeden Spektralwert eines Kurzzeitspektrums einer Mehrzahl von Blöcken einen Mittelwert unter Verwendung entsprechender Spektralwerte der Mehrzahl von Blöcken zu berechnen 74, um ein Mittelwert-Spektrum zu erhalten, um für Spektralwerte Abweichungen zu berechnen, die für Spektralwerte unterschiedlich sind und kleiner als eine maximale Abweichung  $\Delta_{\max}$  sind, und um die Abweichungen und die Mittelwert-Spektralwerte zu addieren, um ein verarbeitetes Spektrum zu erhalten.

**[0083]** Bei einem Ausführungsbeispiel ist der Synthesesignalgenerator 12 ausgebildet, um das Synthesesignal aus Signalabschnitten des Untersuchungssignals

vor oder nach dem Transientenzeitraum, aus dem Untersuchungssignal im Transientenzeitraum nach Glättung des zeitlichen Verlaufs desselben oder aus einer Kombination der Signalabschnitte des Untersuchungssignals und dem Untersuchungssignals nach einer Glättung zu berechnen.

**[0084]** Bei einem Ausführungsbeispiel ist der Synthesesignalgenerator 12 ausgebildet, um Signalabschnitte des Untersuchungssignals vor oder nach dem Transientenzeitraum zu kopieren.

**[0085]** Bei einem Ausführungsbeispiel ist der Synthesesignalgenerator 12 ausgebildet, um extrapolierte Spektralwerte, die aus dem Untersuchungssignal außerhalb des Transientenzeitraums abgeleitet sind, in einem vorbestimmten Bereich zu randomisieren.

**[0086]** Bei einem Ausführungsbeispiel ist der Synthesesignalgenerator 12 ausgebildet, um dann, wenn der Transientenzeitraum länger als eine vorbestimmte Zeit dauert, für Zeiten, die später als der bestimmte Zeitraum sind, Synthesesignal-Werte mit Signal-Werten des Untersuchungssignals zu mischen.

**[0087]** Bei einem Ausführungsbeispiel ist der Signal-substituierer 14 ausgebildet, um von einem Abschnitt vor dem Transientenzeitraum in den Transientenzeitraum gemäß einer Überblendfunktion überzublenden, oder um von dem Transientenzeitraum in einen Abschnitt nach dem Transientenzeitraum gemäß einer Überblendfunktion überzublenden.

**[0088]** Bei einem Ausführungsbeispiel ist der Synthesesignalgenerator 12 ausgebildet, um ein Kurzzeitspektrum des Synthesesignals mit Spektralwerten zu berechnen 40, 41, 42, um das Kurzzeitspektrum in eine zeitliche Darstellung umzusetzen 43, die das Synthesesignal 44 darstellt.

**[0089]** Bei einem Ausführungsbeispiel ist der Synthesesignalgenerator 12 ausgebildet, um ein Kurzzeitspektrum des Synthesesignals mit Subband-Signalen zu berechnen 40, 41, 42, und um das Kurzzeitspektrum mit Subband-Signalen in eine zeitliche Darstellung umzusetzen, die das Synthesesignal darstellt 43.

**[0090]** Bei einem Ausführungsbeispiel ist der Synthesesignalgenerator 12 ausgebildet, um das Synthesesignal so zu erzeugen, dass die vorbestimmte Schwelle kleiner oder gleich einem Faktor 2 ist.

**[0091]** Bei einem Ausführungsbeispiel ist der Synthesesignalgenerator 12 ausgebildet, um eine bandselektive voreingestellte Schwelle oder eine einzige Schwelle für das gesamte Spektrum zu verwenden.

**[0092]** Bei einem Ausführungsbeispiel weist die Vorrichtung ferner eine Extraktionseinrichtung zum Verarbeiten eines linken Kanal-Signals und eines rechten Kanal-Signals, um das Untersuchungssignal zu extrahieren, auf.

**[0093]** Bei einem Ausführungsbeispiel weist die Vorrichtung einen 2-zu-3-Mischer 82 zum Erzeugen eines linken Kanals, eines rechten Kanals und eines Mittel-Kanals aus einem übertragenen Stereo- oder Mono-Signal auf, wobei der Synthesesignalgenerator 12 ausgebildet

ist, um das gleiche Umgebungssignal für den hinteren linken oder hinteren rechten Kanal zu liefern, oder um das Untersuchungssignal zu skalieren, so dass der hintere linke Kanal und der hintere rechte Kanal unterschiedlich skalierte Version des Umgebungssignals erhalten können, oder um für zwei Surround-Kanäle zwei eigene Umgebungssignale zu berechnen.

**[0094]** Die vorliegende Erfindung betrifft ferner ein Verfahren zum Erzeugen eines Umgebungssignals, das zur Ausstrahlung über Lautsprecher Ls, Rs geeignet ist, für die kein geeignetes Lautsprechersignal vorhanden ist, mit folgenden Schritten: Detektieren 11 eines Transientenzeitraums 20, in dem ein Untersuchungssignal einen transienten Bereich 22 aufweist; Erzeugen 12 eines Synthesesignals für den Transientenzeitraum 20, wobei der Synthesesignalgenerator 12 ausgebildet ist, um ein Synthesesignal zu erzeugen, das einen flacheren zeitlichen Verlauf als das Untersuchungssignal in dem Transientenzeitraum 20 aufweist, und dessen Intensität von einer Intensität eines vorangehenden oder nachfolgenden Abschnitts des Untersuchungssignals um weniger als eine vorbestimmte Schwelle abweicht; und Substituieren 14 des Untersuchungssignals im Transientenzeitraum 20 durch das Synthesesignal, um das Umgebungssignal zu erhalten.

**[0095]** Abhängig von der Gegebenheit, kann das erfindungsgemäße Verfahren in Hardware oder in Software implementiert werden. Die Implementierung kann auf einem digitalen Speichermedium, insbesondere einer Diskette oder CD mit elektronisch auslesbaren Steuersignalen erfolgen, die so mit einem programmierbaren Computersystem zusammenwirken können, dass das Verfahren ausgeführt wird. Allgemein besteht die Erfindung somit auch in einem Computer-Programm-Produkt mit einem auf einem maschinenlesbaren Träger gespeicherten Programmcode zur Durchführung des erfindungsgemäßen Verfahrens, wenn das Computer-Programm-Produkt auf einem Rechner abläuft. In anderen Worten ausgedrückt, kann die Erfindung somit als ein Computer-Programm mit einem Programmcode zur Durchführung des Verfahrens realisiert werden, wenn das Computer-Programm auf einem Computer abläuft.

#### 45 Patentansprüche

1. Vorrichtung zum Erzeugen eines Umgebungssignals, das zur Ausstrahlung über Lautsprecher (Ls, Rs) geeignet ist, für die kein geeignetes Lautsprechersignal vorhanden ist, mit folgenden Merkmalen:

einem Transienten-Detektor (11) zum Detektieren eines Transientenzeitraums (20), in dem ein Untersuchungssignal einen transienten Bereich (22) aufweist;  
einem Synthesesignalgenerator (12) zum Erzeugen eines Synthesesignals für den Transientenzeitraum (20), wobei der Synthesesi-

- gnalgenerator (12) ausgebildet ist, um ein Synthesesignal zu erzeugen, das einen flacheren zeitlichen Verlauf als das Untersuchungssignal in dem Transientenzeitraum (20) aufweist, und dessen Intensität von einer Intensität eines vorangehenden oder nachfolgenden Abschnitts des Untersuchungssignals um weniger als eine vorbestimmte Schwelle abweicht; und einem Signalsubstituierer (14) zum Substituieren des Untersuchungssignals im Transientenzeitraum durch das Synthesesignal, um das Umgebungssignal zu erhalten, wobei der Signalsubstituierer (14) ausgebildet ist, um von einem Abschnitt vor dem Transientenzeitraum in den Transientenzeitraum gemäß einer Überblendfunktion überzublenden, oder um von dem Transientenzeitraum in einen Abschnitt nach dem Transientenzeitraum gemäß einer Überblendfunktion überzublenden.
2. Vorrichtung nach Anspruch 1, die für eine Blockverarbeitung ausgebildet ist, um aufeinander folgende Blöcke von zeitdiskreten Abtastwerten überlappend oder nicht-überlappend zu verarbeiten.
  3. Vorrichtung nach Anspruch 2, bei der der Transienten-Detektor (11) ausgebildet ist, um für aufeinander folgende Blöcke Intensitätswerte zu berechnen, und um einen Transientenzeitraum (20) zu detektieren, wenn ein Intensitätswert eines Blocks von einem vorausgehenden oder nachfolgenden Intensitätswert um mehr als eine vorbestimmte Transientenschwelle unterschiedlich ist.
  4. Vorrichtung nach Anspruch 3, bei der der Synthesesignalgenerator (12) ausgebildet ist, um für einen Block im Transientenzeitraum (20) eine Mehrzahl von Spektralwerten, die ein Kurzzeitspektrum des Blocks darstellen, so zu begrenzen, dass ihre Intensität sich von der Intensität eines vorausgehenden oder nachfolgenden Blocks oder Transient um weniger als die vorbestimmte Schwelle unterschiedlich ist.
  5. Vorrichtung nach Anspruch 3 oder 4, bei der der Synthesesignalgenerator (12) ausgebildet ist, um komplexe Spektralwerte, die ein Kurzzeitspektrum des Blocks, der den Transientenzeitraum (20) umfasst, darstellen, im Hinblick auf ihre Phase oder ihr Vorzeichen zu randomisieren.
  6. Vorrichtung nach Anspruch 3 oder 4, bei der der Synthesesignalgenerator (12) ausgebildet ist, um eine Prädiktionsverarbeitung (51) über der Frequenz durchzuführen, um ein Prädiktions-Spektrum zu erhalten, dessen zugeordnetes Zeitsignal einen flacheren zeitlichen Verlauf als ein Zeitsignal hat, das einem Spektrum vor der Prädiktionsverarbeitung über der Frequenz zugeordnet ist.
  7. Vorrichtung nach einem der vorhergehenden Ansprüche, bei der der Transienten-Detektor (11) ausgebildet ist, um einen Hochfrequenz-Gehalt für einen Block des Untersuchungssignals zu berechnen (61); bei der der Transienten-Detektor (11) ausgebildet ist, um den gewichteten HF-Gehalt mit einem gleitenden Mittelwert über einer Mehrzahl von vorausgehenden oder nachfolgenden Blöcken ohne Transiente zu vergleichen (73), wobei der Transienten-Detektor (11) ausgebildet ist, um eine Transiente für einen Block zu detektieren, wenn der HF-Gehalt eines aktuellen Blocks den gleitenden Mittelwert um mehr als eine Schwelle (c) überschreitet.
  8. Vorrichtung nach Anspruch 7, bei der der Transienten-Detektor ausgebildet ist, um eine Schwelle zu verwenden, die abhängig von der Art der Berechnung des gleitenden Mittelwerts gewählt ist und näher bei Eins liegt, wenn in den gleitenden Mittelwert die Vergangenheit stärker eingeht und weiter weg von Eins liegt, wenn in den gleitenden Mittelwert die Vergangenheit vergleichsweise schwächer eingeht.
  9. Vorrichtung nach Anspruch 7 oder 8, bei der der Synthesesignalgenerator ausgebildet ist, um für jeden Spektralwert eines Kurzzeitspektrums einer Mehrzahl von Blöcken einen Mittelwert unter Verwendung entsprechender Spektralwerte der Mehrzahl von Blöcken zu berechnen (74), um ein Mittelwert-Spektrum zu erhalten, um für Spektralwerte Abweichungen zu berechnen, die für Spektralwerte unterschiedlich sind und kleiner als eine maximale Abweichung ( $\Delta_{\max}$ ) sind, und um die Abweichungen und die Mittelwert-Spektralwerte zu addieren, um ein verarbeitetes Spektrum zu erhalten.
  10. Vorrichtung nach einem der vorhergehenden Ansprüche, bei der der Synthesesignalgenerator (12) ausgebildet ist, um das Synthesesignal aus Signalabschnitten des Untersuchungssignals vor oder nach dem Transientenzeitraum, aus dem Untersuchungssignal im Transientenzeitraum nach Glättung des zeitlichen Verlaufs desselben oder aus einer Kombination der Signalabschnitte des Untersuchungssignals und dem Untersuchungssignals nach einer Glättung zu berechnen.
  11. Vorrichtung nach Anspruch 10, bei der der Synthesesignalgenerator (12) ausgebildet ist, um Signalabschnitte des Untersuchungssignals vor oder nach dem Transientenzeitraum zu kopieren.

12. Vorrichtung nach Anspruch 10,  
bei der der Synthesesignalgenerator (12) ausgebildet ist, um extrapolierte Spektralwerte, die aus dem Untersuchungssignal außerhalb des Transientenzeitraums abgeleitet sind, in einem vorbestimmten Bereich zu randomisieren. 5
13. Vorrichtung nach einem der vorhergehenden Ansprüche,  
bei der der Synthesesignalgenerator (12) ausgebildet ist, um dann, wenn der Transientenzeitraum länger als eine vorbestimmte Zeit dauert, für Zeiten, die später als der vorbestimmte Zeitraum sind, Synthesesignal-Werte mit Signal-Werten des Untersuchungssignals zu mischen. 10 15
14. Vorrichtung nach einem der vorhergehenden Ansprüche,  
bei der der Synthesesignalgenerator (12) ausgebildet ist, um ein Kurzzeitspektrum des Synthesesignals mit Spektralwerten zu berechnen (40, 41, 42), um das Kurzzeitspektrum in eine zeitliche Darstellung umzusetzen (43), die das Synthesesignal (44) darstellt. 20 25
15. Vorrichtung nach einem der vorhergehenden Ansprüche,  
bei der der Synthesesignalgenerator (12) ausgebildet ist, um ein Kurzzeitspektrum des Synthesesignals mit Subband-Signalen zu berechnen (40, 41, 42), und um das Kurzzeitspektrum mit Subband-Signalen in eine zeitliche Darstellung umzusetzen, die das Synthesesignal darstellt (43). 30 35
16. Vorrichtung nach einem der vorhergehenden Ansprüche,  
bei der der Synthesesignalgenerator (12) ausgebildet ist, um das Synthesesignal so zu erzeugen, dass die vorbestimmte Schwelle kleiner oder gleich einem Faktor 2 ist. 40
17. Vorrichtung nach einem der vorhergehenden Ansprüche,  
bei der der Synthesesignalgenerator (12) ausgebildet ist, um eine bandselektive voreingestellte Schwelle oder eine einzige Schwelle für das gesamte Spektrum zu verwenden. 45
18. Vorrichtung nach einem der vorhergehenden Ansprüche, die ferner folgende Merkmale aufweist: 50
- eine Extraktionseinrichtung zum Verarbeiten eines linken Kanal-Signals und eines rechten Kanal-Signals, um das Untersuchungssignal zu extrahieren. 55
19. Vorrichtung nach einem der vorhergehenden An-

sprüche, die ferner folgendes Merkmal aufweist:

einen 2-zu-3-Mischer (82) zum Erzeugen eines linken Kanals, eines rechten Kanals und eines Mittel-Kanals aus einem übertragenen Stereo- oder Mono-Signal; und

bei der der Synthesesignalgenerator (12) ausgebildet ist, um das gleiche Umgebungssignal für den hinteren linken oder hinteren rechten Kanal zu liefern, oder um das Untersuchungssignal zu skalieren, so dass der hintere linke Kanal und der hintere rechte Kanal unterschiedlich skalierte Version des Umgebungssignals erhalten können, oder um für zwei Surround-Kanäle zwei eigene Umgebungssignale zu berechnen.

20. Verfahren zum Erzeugen eines Umgebungssignals, das zur Ausstrahlung über Lautsprecher (Ls, Rs) geeignet ist, für die kein geeignetes Lautsprechersignal vorhanden ist, mit folgenden Schritten:

Detektieren (11) eines Transientenzeitraums (20), in dem ein Untersuchungssignal einen transienten Bereich (22) aufweist;

Erzeugen (12) eines Synthesesignals für den Transientenzeitraum (20), wobei der Synthesesignalgenerator (12) ausgebildet ist, um ein Synthesesignal zu erzeugen, das einen flacheren zeitlichen Verlauf als das Untersuchungssignal in dem Transientenzeitraum (20) aufweist, und dessen Intensität von einer Intensität eines vorangehenden oder nachfolgenden Abschnitts des Untersuchungssignals um weniger als eine vorbestimmte Schwelle abweicht; und

Substituieren (14) des Untersuchungssignals im Transientenzeitraum (20) durch das Synthesesignal, um das Umgebungssignal zu erhalten, wobei im Schritt des Substituierens (14) von einem Abschnitt vor dem Transientenzeitraum in den Transientenzeitraum gemäß einer Überblendfunktion übergeblendet wird, oder von dem Transientenzeitraum in einen Abschnitt nach dem Transientenzeitraum gemäß einer Überblendfunktion übergeblendet wird.

21. Computerprogramm zum Ausführen eines Verfahrens gemäß Patentanspruch 20, wenn das Verfahren auf einem Computer läuft.

FIG 1

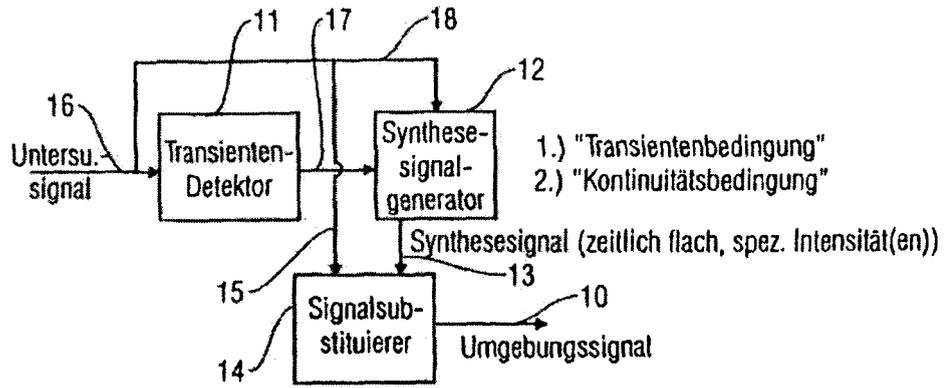


FIG 2a

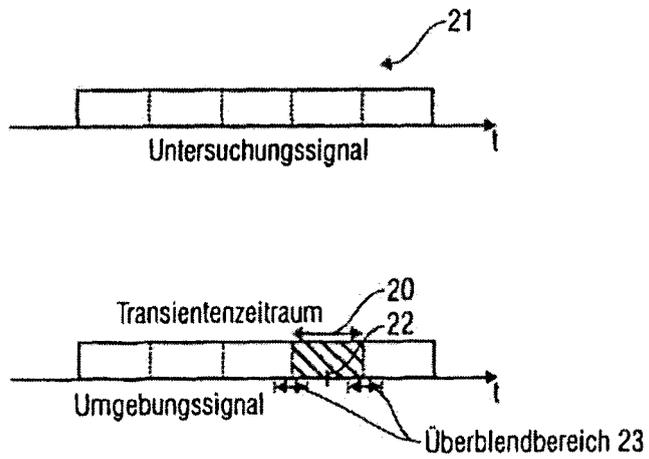
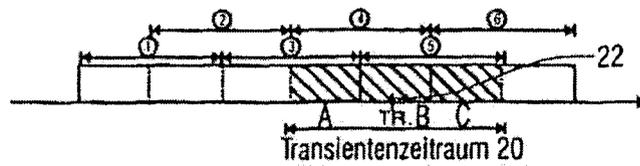


FIG 2b



- Transient wird in ④ und ⑤ detektiert
- Synthesewerte werden für ④ und ⑤ erzeugt
- Synthesesignalgenerator erzeugt:
  - Teil A durch Addition von NR. 3, 2. Hälfte und NR. 4, 1. Hälfte
  - Teil B durch Addition von NR. 4, 2. Hälfte und NR. 5, 1. Hälfte
  - Teil C durch Addition von NR. 5, 2. Hälfte und NR. 6, 1. Hälfte

FIG 3

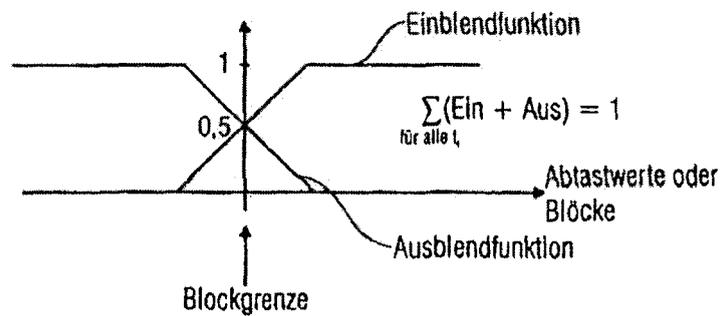


FIG 4

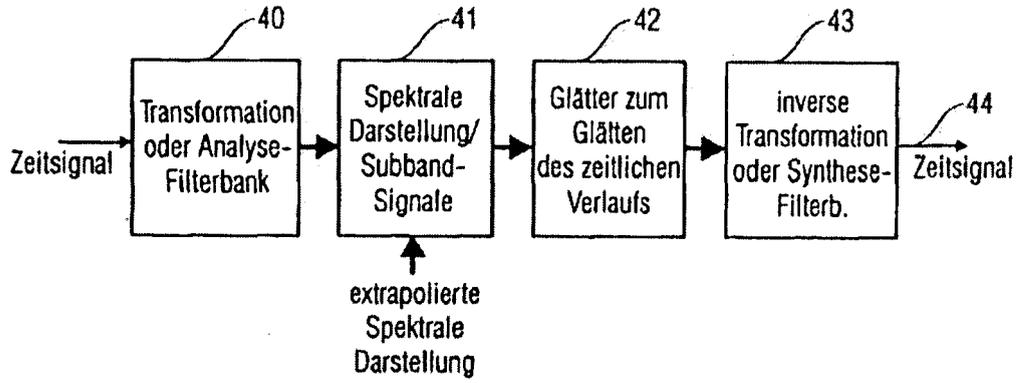


FIG 5a

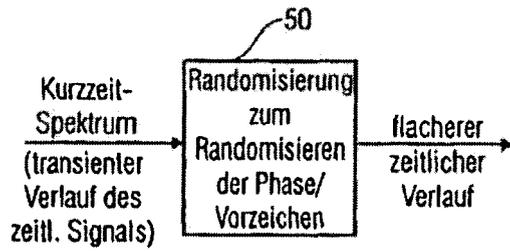


FIG 5b

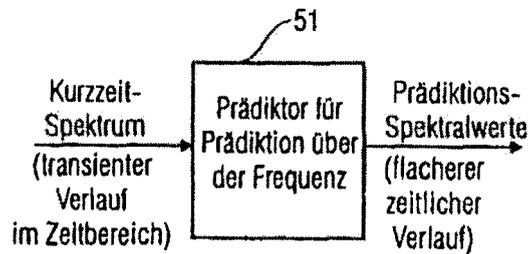


FIG 5c

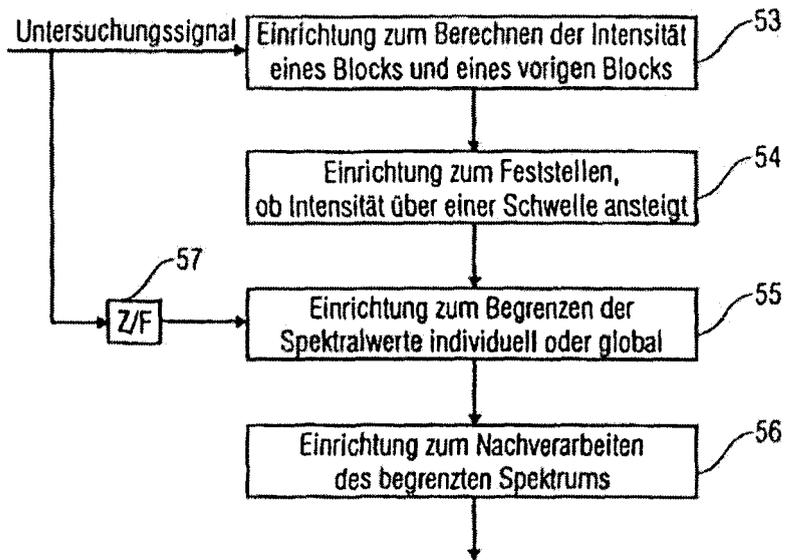


FIG 6

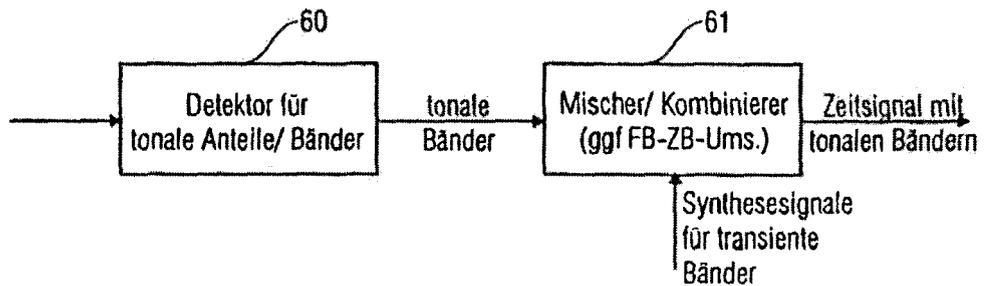


FIG 7

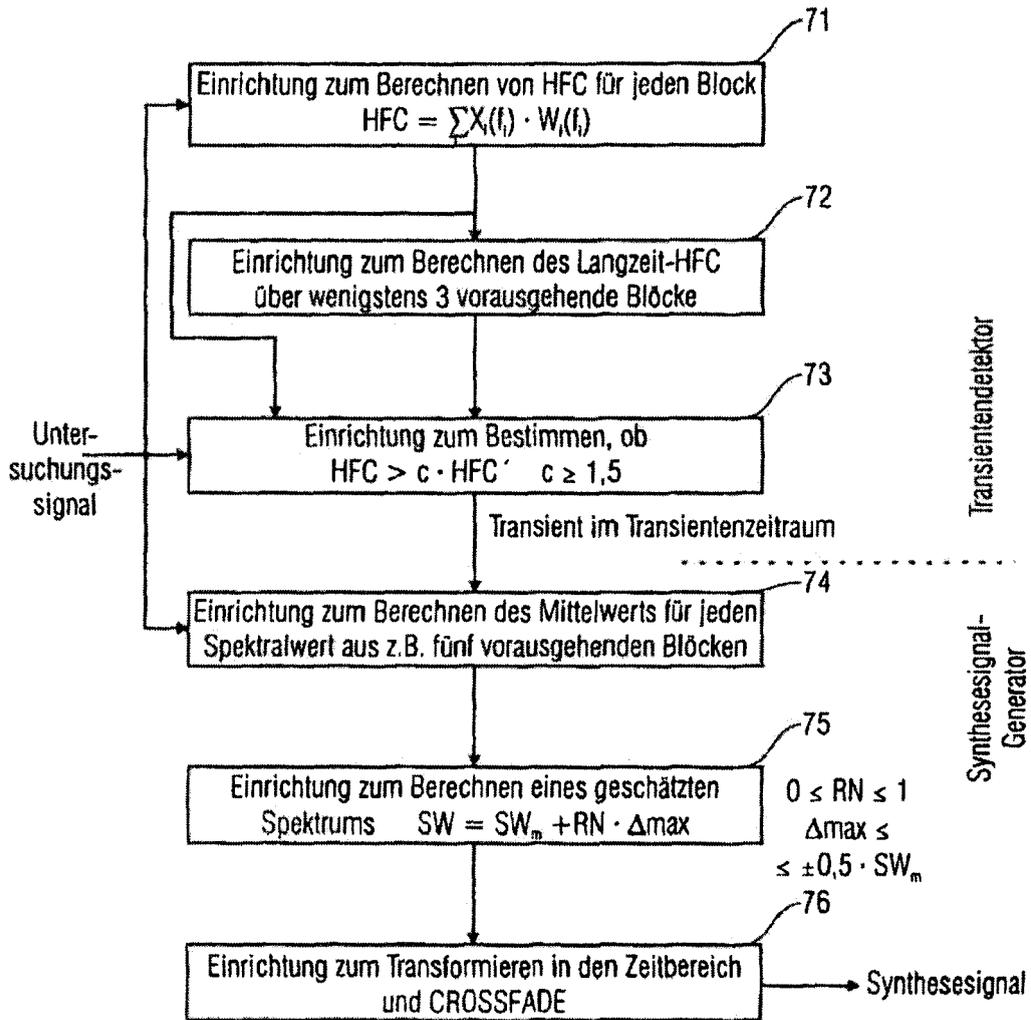


FIG 8

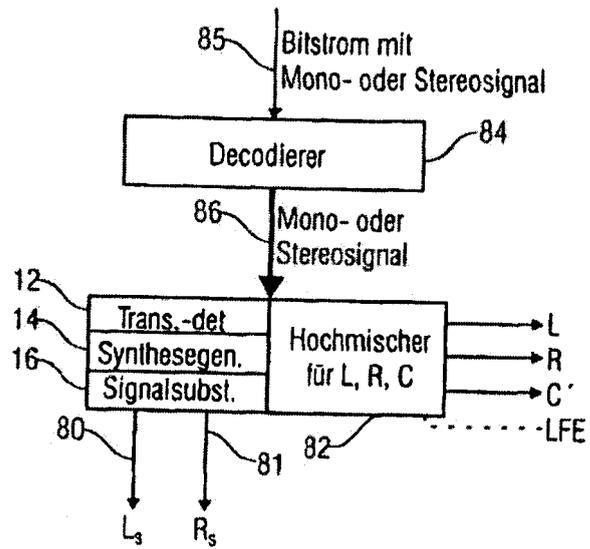


FIG 9

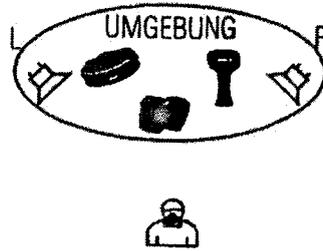


FIG 10

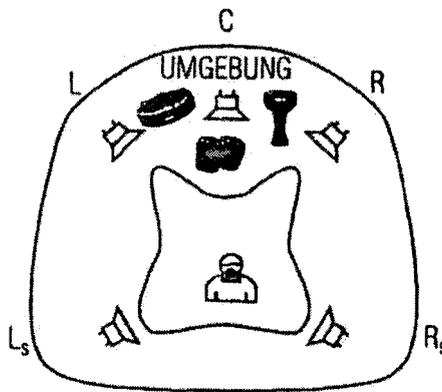
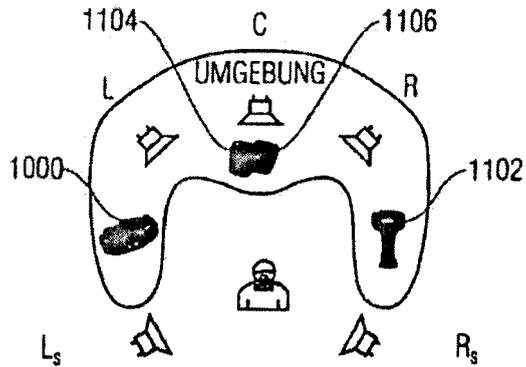


FIG 11



**IN DER BESCHREIBUNG AUFGEFÜHRTE DOKUMENTE**

*Diese Liste der vom Anmelder aufgeführten Dokumente wurde ausschließlich zur Information des Lesers aufgenommen und ist nicht Bestandteil des europäischen Patentdokumentes. Sie wurde mit größter Sorgfalt zusammengestellt; das EPA übernimmt jedoch keinerlei Haftung für etwaige Fehler oder Auslassungen.*

**In der Beschreibung aufgeführte Nicht-Patentliteratur**

- **C. AVENDANO ; J.M. JOT.** Ambience Extraction and Synthesis from Stereo Signals for Multichannel Audio Upmix. *IEEE International Conference on Acoustics, Speech and Signal Processing*, Mai 2002 **[0013]**
- R. Irwan und Ronald M. Aarts: "A method to convert stereo to multi-channel sound. *The proceedings of the AES 19th International Conference*, 21. Juni 2001, 139-143 **[0014]**
- **G. SOULODRE.** Ambience-Based Upmixing", Workshop "Spatial Coding Of Surround Sound: A Progress Report. *117th AES Convention*, 2004 **[0015]**
- **J. HERRE ; J.D. JOHNSTON.** Exploiting Both Time and Frequency Structure in a System that Uses an Analysis/Synthesis Filterbank with High Frequency Resolution. *103rd AES Convention*, 1997 **[0056]**