Vorrichtung und Verfahren für objektbasiertes, räumliches Audio-Mastering

Beschreibung

Die Anmeldung betrifft Audioobjektverarbeitung, Audioobjekt-Enkodierung und Audioobjekt-Dekodierung und, im Besonderen, Audio-Mastering für Audioobjekte.

Objektbasiertes räumliches Audio ist ein Ansatz für interaktive dreidimensionale Audioreproduktion. Dieses Konzept ändert nicht nur die Art, wie Inhaltsersteller bzw. Autoren mit dem Audio interagieren können, sondern auch, wie dasselbe gespeichert und übertragen wird. Um dies zu ermöglichen, muss ein neuer Prozess in der Reproduktionskette, „Rendering" genannt, etabliert werden. Der Renderingprozess erzeugt Lautsprechersignale aus einer objektbasierten Szenenbeschreibung heraus. Obwohl Aufzeichnen und Mischen in den vergangenen Jahren erforscht wurden, fehlen Konzepte für ein objektbasiertes Mastering nahezu. Der Hauptunterschied im Vergleich zu kanalbasiertem Audiomastering besteht darin, dass anstelle eines Anpassens der Audiokanäle die Audioobjekte geändert werden müssen. Dies erfordert ein grundsätzlich neues Konzept für das Mastering. In der Abhandlung wird ein neues Verfahren für das Mastering von objektbasiertem Audio vorgestellt.

In den letzten Jahren hat der objektbasierte Audioansatz viel Interesse hervorgerufen. Im Vergleich zu kanalbasiertem Audio, bei dem Lautsprechersignale als Ergebnis der Räumliches-Audio-Produktion gespeichert werden, wird die Audioszene durch Audioobjekte beschrieben. Ein Audioobjekt kann als eine virtuelle Schallquelle betrachtet werden, die aus einem Audiosignal mit zusätzlichen Metadaten, z. B. Position und Verstärkung, besteht. Um Audioobjekte zu reproduzieren, ist ein so genannter Audiorenderer erforderlich. Das Audiorendering ist der Prozess eines Erzeugens von Lautsprecher- oder Kopfhörersignalen auf der Basis weiterer Informationen, beispielsweise der Position von Lautsprechern oder der Position des Hörers in der virtuellen Szene.

Der Prozess der Audioinhaltserstellung kann in drei Hauptteile gegliedert werden: Aufzeichnen, Mischen und Mastering. Während alle drei Schritte in den vergangenen Jahrzehnten für kanalbasiertes Audio ausgiebig behandelt wurden, erfordert objektbasiertes Audio bei zukünftigen Anwendungen neue Arbeitsabläufe. Bisher muss der Aufzeichnungsschritt allgemein noch nicht geändert werden, selbst wenn zukünftige Techniken neue Möglichkeiten [1], [2] mit sich bringen könnten. Bei dem Mischprozess verhält es

sich etwas anders, da der Sound Engineer nicht länger einen räumlichen Mix durch Schwenken von Signalen an dedizierte Lautsprecher erzeugt. Stattdessen werden alle Positionen von Audioobjekten durch ein Räumliches-Authoring-Tool erzeugt, das es ermöglicht, den Metadatenteil jedes Audioobjekts zu definieren. Ein vollständiger Masteringprozess für Audioobjekte ist bis dato noch nicht etabliert worden [3].

Herkömmliche Audiomixes leiten mehrere Audiospuren zu einer bestimmten Anzahl von Ausgabekanälen. Dies macht es erforderlich, individuelle Mixes für unterschiedliche Wiedergabekonfigurationen zu erstellen, ermöglicht jedoch eine effiziente Behandlung der Ausgabekanäle beim Mastering [4]. Bei Verwendung des objektbasierten Audioansatzes ist der Audiorenderer für das Erstellen aller Lautsprechersignale in Echtzeit verantwortlich. Die Anordnung einer großen Anzahl von Audioobjekten in Rahmen eines kreativen Mischprozesses führt zu komplexen Audioszenen. Da der Renderer jedoch die Audioszene in mehreren unterschiedlichen Lautsprecher-Einrichtungen reproduzieren kann, ist es während der Produktion nicht möglich, die Ausgabekanäle direkt zu adressieren. Das Masteringkonzept kann deshalb lediglich auf einem individuellen Modifizieren von Audioobjekten basieren.

Bis heute richtet sich die herkömmliche Audioproduktion an äußerst spezifische Hör-Einrichtungen und deren Kanalkonfiguration, beispielsweise Stereo- oder Surroundwiedergabe. Die Entscheidung, für welche Wiedergabe-Einrichtung(en) der Inhalt konzipiert ist, muss deshalb zu Beginn dessen Produktion getroffen werden. Der Produktionsprozess selbst besteht dann aus Aufzeichnen, Mischen und Mastering. Der Masteringprozess optimiert den endgültigen Mix, um sicherzustellen, dass derselbe auf allen Verbrauchersystemen mit unterschiedlichen Lautsprechercharakteristiken in zufriedenstellender Qualität wiedergegeben wird. Da das gewünschte Ausgabeformat eines Mixes fest ist, kann der Mastering Engineer (ME) einen optimierten Master für diese Wiedergabekonfiguration erstellen.

Die Masteringphase macht es für Ersteller sinnvoll, Audio in suboptimalen akustischen Umgebungen zu produzieren, da sie sich auf eine finale Prüfung ihres Mixes beim Mastering verlassen können. Dies senkt die Zugangsbarrieren für das Produzieren von professionalem Inhalt. Auf der anderen Seite ist den MEs selbst über die Jahres ein breites Spektrum von Masteringtools angeboten worden, das deren Möglichkeiten für Korrekturen und Verbesserung drastisch verbessert hat. Nichtsdestotrotz ist der finale Inhalt normalerweise auf die Wiedergabe-Einrichtung beschränkt, für die er konzipiert wurde.

Diese Einschränkung wird durch Objektbasierte Räumliche Audio Produktion (englisch: Object-Based Spatial Audio Production, OBAP) grundsätzlich überwunden. Im Gegensatz zu kanalbasiertem Audio basiert OBAP auf individuellen Audioobjekten mit Metadaten, die deren Position in einer künstlichen Umgebung umfassen, die auch als „Szene" bezeichnet wird. Lediglich an dem finalen Hör-Ausgang berechnet eine dedizierte Renderingeinheit, der Renderer, die finalen Lautsprechersignale in Echtzeit basierend auf der Lautsprecher-Einrichtung des Hörers.

Obwohl OBAP jedes Audioobjekt und dessen Metadaten dem Renderer individuell bereitstellt, sind während der Produktion keine direkten kanalbasierten Anpassungen möglich, und somit können keine vorhandenen Masteringtools für herkömmliche Wiedergabe-Einrichtungen verwendet werden. Währenddessen erfordert OBAP, dass alle finalen Anpassungen in dem Mix erfolgen. Während die Erfordernis, Gesamt-Schallanpassungen durch manuelles Behandeln jedes individuellen Audioobjekts zu realisieren, nicht nur höchst ineffizient ist, führt dieser Umstand auch zu hohen Anforderungen an die Überwachungs-Einrichtung jedes Erstellers und schränkt die Schallqualität von objektbasiertem 3D-Audioinhalt strikt auf die akustischen Eigenschaften der Umgebung ein, in der derselbe erstellt wurde.

Ultimativ könnte ein Entwickeln von Tools zum Ermöglichen eines ähnlich leistungsstarken Masteringprozesses für OBAP auf Erstellerseite die Akzeptanz für ein Produzieren von 3D-Audioinhalt verbessern, indem Produktionsbarrieren gesenkt werden und neuer Raum für Klangästhetik und Schallqualität eröffnet wird.

Während erste Gedanken über ein räumliches Mastering der Öffentlichkeit zugänglich gemacht wurden [5], stellt diese Abhandlung neue Ansätze vor, wie herkömmliche Masteringtools angepasst werden können und welche Typen neuer Tools beim Mastering für ein objektbasiertes räumliches Audio als hilfreich erachtet werden können. So ist in [5] ist eine Basissequenz beschrieben, wie Metadaten verwendet werden können, um objektspezifische Parameter von globalen Eigenschaften abzuleiten. Ferner ist in [6] ein Konzept eines interessierenden Bereichs mit einem umgebenden Übergangsbereich in Zusammenhang mit OBAP-Anwendungen beschrieben.

Es ist daher wünschenswert, verbesserte objektbasierte Audio-Mastering Konzepte bereitzustellen.

Eine Vorrichtung nach Anspruch 1, ein Enkoder nach Anspruch 14, ein Dekoder nach Anspruch 15, ein System nach Anspruch 17, ein Verfahren nach Anspruch 18 und ein Computerprogramm nach Anspruch 19 werden bereitgestellt.

Eine Vorrichtung zur Erzeugung eines verarbeiteten Signals unter Verwendung einer Mehrzahl von Audioobjekten gemäß einer Ausführungsform wird bereitgestellt, wobei jedes Audioobjekt der Mehrzahl von Audioobjekten ein Audioobjektsignal und Audioobjekt-Metadaten umfasst, wobei die Audioobjekt-Metadaten eine Position des Audioobjekts und einen Verstärkungsparameter des Audioobjekts umfassen. Die Vorrichtung umfasst: eine Schnittstelle zum Spezifizieren wenigstens eines Effekt-Parameters einer Verarbeitungsobjekt-Gruppe von Audioobjekten durch einen Benutzer, wobei die Verarbeitungsobjekt-Gruppe von Audioobjekten zwei oder mehrere Audioobjekte der Mehrzahl von Audioobjekten umfasst. Ferner umfasst die Vorrichtung eine Prozessoreinheit, die ausgebildet ist, das verarbeitete Signal so zu erzeugen, dass der wenigstens eine Effekt-Parameter, der mittels der Schnittstelle spezifiziert wurde, auf das Audioobjektsignal oder auf die Audioobjekt-Metadaten jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe von Audioobjekten angewendet wird. Ein oder mehrere Audioobjekte der Mehrzahl von Audioobjekten gehören dabei nicht der Verarbeitungsobjekt-Gruppe von Audioobjekten an.

Ferner wird ein Verfahren zur Erzeugung eines verarbeiteten Signals unter Verwendung einer Mehrzahl von Audioobjekten bereitgestellt, wobei jedes Audioobjekt der Mehrzahl von Audioobjekten ein Audioobjektsignal und Audioobjekt-Metadaten umfasst, wobei die Audioobjekt-Metadaten eine Position des Audioobjekts und einen Verstärkungsparameter des Audioobjekts umfassen. Das Verfahren umfasst:

- Spezifizieren von wenigstens einem Effekt-Parameter einer Verarbeitungsobjekt-Gruppe von Audioobjekten durch einen Benutzer mittels einer Schnittstelle (110), wobei die Verarbeitungsobjekt-Gruppe von Audioobjekten zwei oder mehrere Audioobjekte der Mehrzahl von Audioobjekten umfasst. Und:

- Erzeugen des verarbeiteten Signals durch eine Prozessoreinheit (120) derart, dass der wenigstens eine Effekt-Parameter, der mittels der Schnittstellte spezifiziert wurde, auf das Audioobjektsignal oder auf die Audioobjekt-Metadaten jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe von Audioobjekten angewendet wird.

Des Weiteren wird ein Computerprogramm mit einem Programmcode zur Durchführung des oben beschriebenen Verfahrens bereitgestellt.

Das bereitgestellte Audio-Mastering basiert auf einem Mastering von Audioobjekten. Diese können in Ausführungsformen in einer Szene an beliebiger Stelle und frei in Echtzeit positioniert sein. In Ausführungsformen werden beispielsweise die Eigenschaften allgemeiner Audioobjekte beeinflusst. In ihrer Funktion als künstliche Container können sie jeweils eine willkürliche Anzahl von Audioobjekten enthalten. Jede Anpassung an einem Masteringobjekt wird in Echtzeit in individuelle Anpassungen an Audioobjekten desselben umgewandelt.

Solche Masteringobjekte werden als auch als Verarbeitungsobjekte bezeichnet.

Somit kann der Benutzer, anstelle eines separaten Anpassens zahlreicher Audioobjekte, ein Masteringobjekt verwenden, um wechselseitige Anpassungen gleichzeitig an mehreren Audioobjekten durchzuführen.

Der Satz von Zielaudioobjekten für ein Masteringobjekt kann gemäß Ausführungsformen beispielsweise auf zahlreiche Arten definiert werden. Aus einer räumlichen Perspektive kann der Benutzer einen benutzerdefiniert gebildeten Gültigkeitsbereich um die Position des Masteringobjekts herum festlegen. Alternativ ist es möglich, individuell ausgewählte Audioobjekte unabhängig von ihrer Position mit dem Masteringobjekt zu verknüpfen. Das Masteringobjekt berücksichtigt auch potenzielle Änderungen an der Position von Audioobjekten im Lauf der Zeit.

Eine zweite Eigenschaft von Masteringobjekten gemäß Ausführungsformen kann beispielsweise deren Fähigkeit sein, auf der Basis von Interaktionsmodellen zu berechnen, wie jedes Audioobjekt individuell beeinflusst wird. Ähnlich wie bei einem Kanalzug kann ein Masteringobjekt beispielsweise einen beliebigen, allgemeinen Masteringeffekt übernehmen, beispielsweise Equalizer und Kompressoren. Effekt-Plug-Ins stellen dem Benutzer üblicherweise zahlreiche Parameter bereit, z. B. für Frequenz oder Verstärkungssteuerung. Wenn ein neuer Masteringeffekt einem Masteringobjekt hinzugefügt wird, wird dieser automatisch in alle Audioobjekte des Zielsatzes desselben kopiert. Es werden jedoch nicht alle Effektparameterwerte unverändert übertragen. Je nach Berechnungsverfahren für den Zielsatz können einige Parameter des Masteringeffekts gewichtet werden, bevor sie auf ein bestimmtes Audioobjekt angewendet werden. Die Gewichtung kann auf beliebigen Metadaten oder einer Schallcharakteristik des Audioobjekts basieren.

Nachfolgend werden bevorzugte Ausführungsformen der Erfindung unter Bezugnahme auf die Zeichnungen beschrieben.

In den Zeichnungen ist dargestellt:

Fig. 1 zeigt eine Vorrichtung zur Erzeugung eines verarbeiteten Signals unter Verwendung einer Mehrzahl von Audioobjekten gemäß einer Ausführungsform.

Fig. 2 zeigt Vorrichtung gemäß einer weiteren Ausführungsform, wobei die Vorrichtung ein Enkoder ist.

Fig. 3 zeigt Vorrichtung gemäß einer weiteren Ausführungsform, wobei die Vorrichtung ein Dekoder ist.

Fig. 4 zeigt ein System gemäß einer Ausführungsform.

Fig. 5 zeigt ein Verarbeitungsobjekt mit dem Bereich A und dem Fading-Bereich A_f gemäß einer Ausführungsform.

Fig. 6 zeigt ein Verarbeitungsobjekt mit dem Bereich A und Objektradien gemäß einer Ausführungsform.

Fig. 7 zeigt einen relativen Winkel von Audioobjekten zu dem Verarbeitungsobjekt gemäß einer Ausführungsform.

Fig. 8 zeigt ein Equalizer-Objekt mit neuem radialem Umkreis nach einer Ausführungsform.

Fig. 9 zeigt einen Signalfluss einer Kompression der Signal von n Quellen gemäß einer Ausführungsform.

Fig. 10 zeigt eine Szenentransformation unter Verwendung einer Steuertafel M nach einer Ausführungsform.

Fig. 11 zeigt den Zusammenhang eines Verarbeitungsobjekts, mit dem Audiosignal-Effekte und Metadaten-Effekte bewirkt werden, gemäß einer Ausführungsform.

Fig. 12 zeigt die Veränderung von Audioobjekten und Audiosignalen auf eine Eingabe eines Benutzer hin gemäß einer Ausführungsform.

Fig. 13 zeigt ein Verarbeitungsobjekt PO_4 mit Rechteck M zur Verzerrung der Ecken C_1, C_2, C_3 und C_4 durch den Benutzer gemäß einer Ausführungsform.

Fig. 14 zeigt Verarbeitungsobjekte PO_1 und PO_2 mit ihren jeweiligen, sich überlappenden zweidimensionalen Einzugsbereichen A und B nach einer Ausführungsform.

Fig. 15 zeigt Verarbeitungsobjekt PO_3 mit rechteckigem, zweidimensionalen Einzugsbereich C und den Winkeln zwischen PO_3 und den zugeordneten Quellen S_1, S_2 und S_3 gemäß einer Ausführungsform.

Fig. 16 zeigt mögliche schematische Implementation eines Equalizer-Effekts, der auf ein Verarbeitungsobjekt angewandt wurde nach einer Ausführungsform.

Fig. 17 zeigt das Verarbeitungsobjekt PO_5 mit einem dreidimensionalen Einzugsbereich D und den jeweiligen Abständen d_{S_1}, d_{S_2} und d_{S_3} zu den über den Einzugsbereich zugeordneten Quellen S_1, S_2 und S_3 gemäß einer Ausführungsform.

Fig. 18 zeigt prototypische Implementierung eines Verarbeitungsobjekts, auf das ein Equalizer angewandt wurde nach einer Ausführungsform.

Fig. 19 zeigt ein Verarbeitungsobjekt wie in Fig. 18, nur an einer anderen Position und ohne Übergangsfläche gemäß einer Ausführungsform.

Fig. 20 zeigt ein Verarbeitungsobjekt mit einer über seinen Azimut definierten Fläche als Einzugsbereich, sodass die Quellen Src22 und Src4 dem Verarbeitungsobjekt zugeordnet werden nach einer Ausführungsform.

Fig. 21 zeigt ein Verarbeitungsobjekt wie in Fig. 20, jedoch mit zusätzlichem Übergangsbereich, der vom Benutzer über den "Feather"-Schieberegler kontrolliert werden kann gemäß einer Ausführungsform.

Fig. 22 zeigt mehrere Verarbeitungsobjekts in der Szene, mit unterschiedlichen Einzugsbereichen nach einer Ausführungsform.

Fig. 23 zeigt das rote Quadrat auf der rechten Bildseite zeigt ein Verarbeitungsobjekt zur horizontalen Verzerrung der Position von Audioobjekten gemäß einer Ausführungsform.

Fig. 24 zeigt die Szene, nachdem der Benutzer die Ecken des Verarbeitungsobjekts verzogen hat. Die Position aller Quellen hat sich entsprechend der Verzerrung verändert nach einer Ausführungsform.

Fig. 25 zeigt eine mögliche Visualisierung der Zuordnung einzelner Audioobjekte zu einem Verarbeitungsobjekt gemäß einer Ausführungsform.

Fig. 1 zeigt eine Vorrichtung zur Erzeugung eines verarbeiteten Signals unter Verwendung einer Mehrzahl von Audioobjekten gemäß einer Ausführungsform, wobei jedes Audioobjekt der Mehrzahl von Audioobjekten ein Audioobjektsignal und Audioobjekt-Metadaten umfasst, wobei die Audioobjekt-Metadaten eine Position des Audioobjekts und einen Verstärkungsparameter des Audioobjekts umfassen.

Die Vorrichtung umfasst: eine Schnittstelle 110 zum Spezifizieren wenigstens eines Effekt-Parameters einer Verarbeitungsobjekt-Gruppe von Audioobjekten durch einen Benutzer, wobei die Verarbeitungsobjekt-Gruppe von Audioobjekten zwei oder mehrere Audioobjekte der Mehrzahl von Audioobjekten umfasst.

Ferner umfasst die Vorrichtung eine Prozessoreinheit 120, die ausgebildet ist, das verarbeitete Signal so zu erzeugen, dass der wenigstens eine Effekt-Parameter, der mittels der Schnittstelle 110 spezifiziert wurde, auf das Audioobjektsignal oder auf die Audioobjekt-Metadaten jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe von Audioobjekten angewendet wird.

Ein oder mehrere Audioobjekte der Mehrzahl von Audioobjekten gehören dabei nicht der Verarbeitungsobjekt-Gruppe von Audioobjekten an.

Die oben beschriebene Vorrichtung der Fig. 1 realisiert eine effiziente Form des Audio-Masterings für Audioobjekte.

Bei Audioobjekten stellt sich das Problem, das in einer Audioszenerie oftmals eine Vielzahl von Audioobjekten existieren. Sollen diese modifiziert werden, so würde es einen erheblichen Aufwand darstellen, jedes Audioobjekt einzeln zu spezifizieren.

Erfindungsgemäß werden nun eine Gruppe von zwei oder mehreren Audioobjekten in einer Gruppe von Audioobjekten organisiert, die als Verarbeitungsobjekt-Gruppe bezeichnet wird. Eine Verarbeitungsobjekt-Gruppe ist also eine Gruppe von Audioobjekten, die in dieser speziellen Gruppe, der Verarbeitungsobjekt-Gruppe organisiert sind.

Ein Benutzer hat erfindungsgemäß nun die Möglichkeit einen oder mehrere (wenigstens einen) Effekt-Parameter mittels der Schnittstelle 110 zu spezifizieren. Die Prozessoreinheit 120 sorgt dann dafür, dass der Effekt-Parameter durch eine einzelne Eingabe des Effekt-Parameters auf alle zwei oder mehrere Audioobjekte der Verarbeitungsobjekt-Gruppe angewandt wird.

Eine solche Anwendung des Effekt-Parameters kann nun beispielsweise darin bestehen, dass der Effekt-Parameter z.B. einen bestimmten Frequenzbereich des Audioobjektsignals jedes der Audio Objekte des Verarbeitungsobjekt-Gruppe modifiziert.

Oder, der Verstärkungsparameter der Audioobjekt-Metadaten jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe kann beispielsweise in Abhängigkeit von dem Effekt-Parameter entsprechend vergrößert oder verkleinert werden.

Oder, die Position der Audioobjekt-Metadaten jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe kann beispielsweise in Abhängigkeit von dem Effekt-Parameter entsprechend verändert werden. Beispielsweise ist denkbar, dass alle Audioobjekte der Verarbeitungsobjekt-Gruppe um +2 entlang einer x-Koordinatenachse, um -3 entlang einer y-Koordinatenachse und um +4 entlang einer z-Koordinatenachse verschoben werden.

Auch ist denkbar, dass sich die Anwendung eines Effekt-Parameters auf die Audioobjekte der Verarbeitungsobjekt-Gruppe für jedes Audioobjekt der Verarbeitungsobjekt-Gruppe unterschiedlich auswirkt. Beispielsweise kann eine als Effekt-Parameter eine Achse defi-

niert werden, an der die Position aller Audioobjekte der Verarbeitungsobjekt-Gruppe gespiegelt wird. Die Positionsveränderung der Audioobjekte der Verarbeitungsobjekt-Gruppe wirkt sich dann für jedes Audioobjekt der Verarbeitungsobjekt-Gruppe unterschiedlich aus.

In einer Ausführungsform kann die Prozessoreinheit 120 z.B. ausgebildet sein, den wenigstens einen Effekt-Parameter, der mittels der Schnittstellte spezifiziert wurde, auf kein Audioobjektsignal und keine Audioobjekt-Metadaten der ein oder mehreren Audioobjekte anzuwenden, die der Verarbeitungsobjekt-Gruppe von Audioobjekten nicht angehören.

Für eine solche Ausführungsform ist festgelegt, dass der Effekt-Parameter auf Audioobjekte gerade nicht angewandt wird, die nicht der Verarbeitungsobjekt-Gruppe angehören.

Grundsätzlich kann das Audioobjekt-Mastering entweder zentral auf der Enkoder-Seite durchgeführt werden. Oder, Dekoder-seitig kann der Endnutzer als Empfänger der Audioobjekt-Szenerie die Audioobjekte selber erfindungsgemäß modifizieren.

Eine Ausführungsform, die Audioobjekt-Mastering erfindungsgemäß auf der Enkoderseite realisiert, ist in Fig. 2 gezeigt.

Eine Ausführungsform, die Audioobjekt-Mastering erfindungsgemäß auf der Dekoderseite realisiert, ist in Fig. 3 gezeigt.

Fig. 2 zeigt Vorrichtung gemäß einer weiteren Ausführungsform, wobei die Vorrichtung ein Enkoder ist.

In Fig. 2 ist die Prozessoreinheit 120 ausgebildet, unter Verwendung der Audioobjektsignale der Mehrzahl von Audioobjekten ein Downmix-Signal zu erzeugen. Dabei ist die Prozessoreinheit 120 ausgebildet, unter Verwendung der Audioobjekt-Metadaten der Mehrzahl von Audioobjekten ein Metadatensignal-Signal zu erzeugen.

Des Weiteren ist die Prozessoreinheit 120 in Fig. 2 ausgebildet, als das verarbeitete Signal das Downmix-Signal zu erzeugen, wobei in dem Downmix-Signal zumindest ein modifiziertes Objektsignal für jedes Audioobjekt der Verarbeitungsobjekt-Gruppe von Audioobjekten gemischt ist, wobei die Prozessoreinheit 120 ausgebildet ist, für jedes Audioobjekt der Verarbeitungsobjekt-Gruppe von Audioobjekten das modifizierte Objektsignal dieses Audioobjekts mittels der Anwendung des wenigstens einen Effekt-

Parameters, der mittels der Schnittstelle 110 spezifiziert wurde, auf das Audioobjektsignal dieses Audioobjekts zu erzeugen.

Oder, die Prozessoreinheit 120 der Fig. 2 ist ausgebildet, als das verarbeitete Signal das Metadaten-Signal zu erzeugen, wobei das Metadaten-Signal zumindest eine modifizierte Position für jedes Audioobjekt der Verarbeitungsobjekt-Gruppe von Audioobjekten umfasst, wobei die Prozessoreinheit 120 ausgebildet ist, für jedes Audioobjekt der Verarbeitungsobjekt-Gruppe von Audioobjekten die modifizierte Position dieses Audioobjekts mittels der Anwendung des wenigstens einen Effekt-Parameters, der mittels der Schnittstelle 110 spezifiziert wurde, auf die Position dieses Audioobjekts zu erzeugen.

Oder, die Prozessoreinheit 120 der Fig. 2 ist ausgebildet, als das verarbeitete Signal das Metadaten-Signal zu erzeugen, wobei das Metadaten-Signal zumindest einen modifizierten Verstärkungsparameter für jedes Audioobjekt der Verarbeitungsobjekt-Gruppe von Audioobjekten umfasst, wobei die Prozessoreinheit 120 ausgebildet ist, für jedes Audioobjekt der Verarbeitungsobjekt-Gruppe von Audioobjekten der modifizierte Verstärkungsparameter dieses Audioobjekts mittels der Anwendung des wenigstens einen Effekt-Parameters, der mittels der Schnittstelle 110 spezifiziert wurde, auf den Verstärkungsparameter dieses Audioobjekts zu erzeugen.

Fig. 3 zeigt Vorrichtung gemäß einer weiteren Ausführungsform, wobei die Vorrichtung ein Dekoder ist. Die Vorrichtung der Fig. 3 ist zum Empfang eines Downmix-Signals ausgebildet, in dem die Mehrzahl von Audioobjektsignalen der Mehrzahl von Audioobjekten gemischt sind. Des Weiteren ist die Vorrichtung der Fig. 3 zum Empfang eines Metadatensignals ausgebildet, wobei das Metadatensignal für jedes Audioobjekt der Mehrzahl von Audioobjekten die Audioobjekt-Metadaten dieses Audioobjekts umfasst.

Die Prozessoreinheit 120 der Fig. 3 ist ausgebildet, die Mehrzahl von Audioobjektsignalen der Mehrzahl von Audioobjekten basierend aus einem Downmix-Signal zu rekonstruieren.

Ferner ist die Prozessoreinheit 120 der Fig. 3 ausgebildet, als das verarbeitete Signal ein Audioausgangssignal umfassend ein oder mehrere Audioausgangskanäle zu erzeugen.

Des Weiteren ist die Prozessoreinheit 120 der Fig. 3 ausgebildet, zum Erzeugen des verarbeiteten Signals den wenigstens einen Effekt-Parameter, der mittels der Schnittstelle 110 spezifiziert wurde, auf das Audioobjektsignal jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe von Audioobjekten anzuwenden, oder zum Erzeugen des verarbeiteten Signals den wenigstens einen Effekt-Parameter, der mittels der Schnittstelle

110 spezifiziert wurde, auf die Position oder auf den Verstärkungsparameter der Audioobjekt-Metadaten jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe von Audioobjekten anzuwenden.

In Audioobjekt-Dekodierung ist das Rendering auf Dekoderseite dem Fachmann wohlbekannt, beispielsweise aus dem SAOC Standard (Spatial Audio Object Coding; deutsch: räumliche Audioobjekt-Kodierung), siehe [8].

Dekoderseitig können dabei durch eine Nutzereingabe über die Schnittstelle 110 beispielsweise ein oder mehrere Rendering-Parameter spezifiziert werden.

So kann in einer Ausführungsform die Schnittstelle 110 der Fig. 3 beispielsweise des Weiteren zum Spezifizieren von ein oder mehreren Rendering-Parametern durch den Benutzer ausgebildet sein. Dabei kann die Prozessoreinheit 120 der Fig. 3 beispielsweise ausgebildet sein, das verarbeitete Signal unter Verwendung der ein oder mehreren Rendering-Parameter in Abhängigkeit von der Position jedes Audioobjekts der Verarbeitungsobjekt-Gruppe von Audioobjekten zu erzeugen.

Fig. 4 zeigt ein System gemäß einer Ausführungsform umfassend einen Enkoder 200 und einen Dekoder 300.

Der Enkoder 200 der Fig. 4 ist dabei zur Erzeugung eines Downmix-Signals basierend auf Audioobjektsignalen einer Mehrzahl von Audioobjekten und zur Erzeugung eines Metadaten-Signals basierend auf Audioobjekt-Metadaten der Mehrzahl von Audioobjekten ausgebildet, wobei die Audioobjekt-Metadaten eine Position des Audioobjekts und einen Verstärkungsparameter des Audioobjekts umfassen.

Der Dekoder 400 der Fig. 4 ist zur Erzeugung eines Audioausgangssignals umfassend ein oder mehrere Audioausgangskanäle basierend auf dem Downmix-Signal und basierend auf dem Metadaten-Signal ausgebildet.

Der Enkoder 200 des Systems der Fig. 4 kann eine Vorrichtung gemäß Fig. 2 sein.

Oder, der Dekoder 300 des Systems der Fig. 4 ist kann eine Vorrichtung gemäß Fig. 3 sein.

Oder, der Enkoder 200 des Systems der Fig. 4 kann eine Vorrichtung gemäß Fig. 2 sein, und der Dekoder 300 des Systems der Fig. 4 kann eine Vorrichtung der Fig. 3 sein.

Die nachfolgenden Ausführungsformen sind gleichermaßen in einer Vorrichtung der Fig. 1 und in einer Vorrichtung der Fig. 2 und in einer Vorrichtung der Fig. 3 implementierbar. Ebenso sind sie in einem Enkoder 200 des Systems der Fig. 4 realisierbar, sowie in einem Dekoder 300 des Systems der Fig. 4.

Gemäß einer Ausführungsform kann die Prozessoreinheit 120 z.B. ausgebildet sein, das verarbeitete Signal so zu erzeugen, dass der wenigstens eine Effekt-Parameter, der mittels der Schnittstelle 110 spezifiziert wurde, auf das Audioobjektsignal jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe von Audioobjekten angewendet wird. Dabei kann die Prozessoreinheit 120 beispielsweise ausgebildet sein, den wenigstens einen Effekt-Parameter, der mittels der Schnittstellte spezifiziert wurde, auf kein Audioobjektsignal der ein oder mehreren Audioobjekte der Mehrzahl von Audioobjekten anzuwenden, die der Verarbeitungsobjekt-Gruppe von Audioobjekten nicht angehören.

Eine solche Anwendung des Effekt-Parameters kann nun beispielsweise darin bestehen, dass die Anwendung des Effekt-Parameters auf das Audioobjektsignal jedes Audioobjekts der Verarbeitungsobjekt-Gruppe z.B. einen bestimmten Frequenzbereich des Audioobjektsignals jedes der Audio Objekte des Verarbeitungsobjekt-Gruppe modifiziert.

In einer Ausführungsform kann die Prozessoreinheit 120 z.B. ausgebildet sein, das verarbeitete Signal so zu erzeugen, dass der wenigstens eine Effekt-Parameter, der mittels der Schnittstelle 110 spezifiziert wurde, auf den Verstärkungsparameter der Metadaten jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe von Audioobjekten angewendet wird. Dabei kann die Prozessoreinheit 120 beispielsweise ausgebildet sein, den wenigstens einen Effekt-Parameter, der mittels der Schnittstellte spezifiziert wurde, auf keinen Verstärkungsparameter der Audioobjekt-Metadaten der ein oder mehreren Audioobjekte der Mehrzahl von Audioobjekten anzuwenden, die der Verarbeitungsobjekt-Gruppe von Audioobjekten nicht angehören.

Wie bereits beschrieben kann in einer solchen Ausführungsform, der Verstärkungsparameter der Audioobjekt-Metadaten jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe kann beispielsweise in Abhängigkeit von dem Effekt-Parameter entsprechend vergrößert, (z.B. um +3dB erhöht) oder verkleinert werden.

Gemäß einer Ausführungsform kann die Prozessoreinheit 120 z.B. ausgebildet sein, das verarbeitete Signal so zu erzeugen, dass der wenigstens eine Effekt-Parameter, der mit-

tels der Schnittstelle 110 spezifiziert wurde, auf die Position der Metadaten jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe von Audioobjekten angewendet wird. Dabei kann die Prozessoreinheit 120 beispielsweise ausgebildet sein, den wenigstens einen Effekt-Parameter, der mittels der Schnittstellte spezifiziert wurde, auf keine Position der Audioobjekt-Metadaten der ein oder mehreren Audioobjekte der Mehrzahl von Audioobjekten anzuwenden, die der Verarbeitungsobjekt-Gruppe von Audioobjekten nicht angehören.

Wie bereits beschrieben, kann in einer solchen Ausführungsform die Position der Audioobjekt-Metadaten jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe kann beispielsweise in Abhängigkeit von dem Effekt-Parameter entsprechend verändert werden. Dies kann z.B. durch Spezifikation der entsprechenden x-, y-, und z-Koordinatenwerte erfolgen, um die die Position jedes der Audioobjekte verschoben werden soll. Oder, es kann beispielsweise eine Verschiebung um einen bestimmten Winkel, gedreht um eine definierten Mittelpunkt, zum Beispiel um eine Nutzerposition, spezifiziert werden, Oder, aber, es kann beispielsweise eine Verdoppelung (oder beispielsweise eine Halbierung) des Abstands zu einem bestimmten Punkt als Effekt-Parameter für die Position jedes Audioobjekts der Verarbeitungsobjekt-Gruppe vorgesehen sein.

In einer Ausführungsform kann die Schnittstelle 110 beispielsweise zum Spezifizieren wenigstens eines Definitions-Parameters der Verarbeitungsobjekt-Gruppe von Audioobjekten durch den Benutzer ausgebildet sein. Dabei kann die Prozessoreinheit 120 beispielsweise ausgebildet sein, in Abhängigkeit von dem wenigstens einen Definitions-Parameter der Verarbeitungsobjekt-Gruppe von Audioobjekten, der mittels der Schnittstelle 110 spezifiziert wurde, zu bestimmen, welche Audioobjekte der Mehrzahl von Audioobjekten der Verarbeitungsobjekt-Gruppe von Audioobjekten angehören.

So kann gemäß einer Ausführungsform dabei der wenigstens eine Definitions-Parameter der Verarbeitungsobjekt-Gruppe von Audioobjekten beispielsweise zumindest eine Position eines interessierenden Bereichs umfassen (wobei die Position des interessierenden Bereichs beispielsweise der Mittelpunkt oder Schwerpunkt des interessierenden Bereichs ist). Dabei kann der interessierende Bereich der Verarbeitungsobjekt-Gruppe von Audioobjekten zugeordnet sein. Die Prozessoreinheit 120 kann dabei z.B. ausgebildet sein, für jedes Audioobjekt der Mehrzahl von Audioobjekten in Abhängigkeit von der Position der Audioobjekt-Metadaten dieses Audioobjekts und in Abhängigkeit von der Position des interessierenden Bereichs zu

bestimmen, ob dieses Audioobjekt der Verarbeitungsobjekt-Gruppe von Audioobjekten angehört.

In einer Ausführungsform kann der wenigstens eine Definitions-Parameter der Verarbeitungsobjekt-Gruppe von Audioobjekten z.B. des Weiteren einen Radius des interessierenden Bereichs umfasst, der der Verarbeitungsobjekt-Gruppe von Audioobjekten zugeordnet sein. Dabei kann die Prozessoreinheit 120 beispielsweise ausgebildet sein, für jedes Audioobjekt der Mehrzahl von Audioobjekten in Abhängigkeit von der Position der Audioobjekt-Metadaten dieses Audioobjekts und in Abhängigkeit von der Position des interessierenden Bereichs und in Abhängigkeit von dem Radius des interessierenden Bereichs zu entscheiden, ob dieses Audioobjekt der Verarbeitungsobjekt-Gruppe von Audioobjekten angehört.

Beispielsweise kann ein Nutzer eine Position der Verarbeitungsobjekt-Gruppe und einen Radius der Verarbeitungsobjekt-Gruppe spezifizieren. Die Position der Verarbeitungsobjekt-Gruppe kann dabei einen räumlichen Mittelpunkt spezifizieren, und der Radius der Verarbeitungsobjekt-Gruppe definiert dann zusammen mit dem Mittelpunkt der Verarbeitungsobjekt-Gruppe einen Kreis. Alle Audioobjekte mit einer Position innerhalb des Kreises oder auf der Kreislinie können dann als Audioobjekte dieser Verarbeitungsobjekt-Gruppe definiert sein; alle Audioobjekte mit einer Position außerhalb des Kreises sind dann nicht von der Verarbeitungsobjekt-Gruppe umfasst. Der Bereich auf innerhalb der Kreislinie und auf der Kreislinie kann dann als ein „interessierender Bereich" verstanden werden.

Gemäß einer Ausführungsform kann die Prozessoreinheit 120 z.B. ausgebildet sein, für jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe von Audioobjekten einen Gewichtungsfaktor in Abhängigkeit von einem Abstand zwischen der Position der Audioobjekt-Metadaten dieses Audioobjekts und der Position des interessierenden Bereichs zu bestimmen. Dabei kann die Prozessoreinheit 120 beispielsweise ausgebildet sein, für jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe von Audioobjekten, den Gewichtungsfaktor dieses Audioobjekts zusammen mit dem wenigstens eine Effekt-Parameter, der mittels der Schnittstelle 110 spezifiziert wurde, auf das Audioobjektsignal oder auf den Verstärkungsparameter der Audioobjekt-Metadaten dieses Audioobjekts anzuwenden.

In solch einer Ausführungsform wird Einfluss des Effekt-Parameters auf die einzelnen Audioobjekte der Verarbeitungsobjekt-Gruppe dadurch für jedes Audioobjekt individualisiert, dass zusätzlich zu Effekt-Parameter ein für jedes Audioobjekt individueller Gewichtungsfaktor bestimmt wird, der auf das Audioobjekt angewandt wird.

In einer Ausführungsform kann der wenigstens eine Definitions-Parameter der Verarbeitungsobjekt-Gruppe von Audioobjekten beispielsweise zumindest einen Winkel umfassen, der eine Richtung von einer definierten Benutzerposition aus spezifiziert, in der sich ein interessierender Bereich befindet, der der Verarbeitungsobjekt-Gruppe von Audioobjekten zugeordnet ist. Dabei kann die Prozessoreinheit 120 z.B. ausgebildet sein, für jedes Audioobjekt der Mehrzahl von Audioobjekten in Abhängigkeit von der Position der Metadaten dieses Audioobjekts und in Abhängigkeit von dem Winkel, der die Richtung von der definierten Benutzerposition aus spezifiziert, in der sich der interessierende Bereich befindet, zu bestimmen, ob dieses Audioobjekt der Verarbeitungsobjekt-Gruppe von Audioobjekten angehört.

Gemäß einer Ausführungsform kann die Prozessoreinheit 120 z.B. ausgebildet sein, für jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe von Audioobjekten einen Gewichtungsfaktor zu bestimmen, der von einer Differenz eines ersten Winkels und eines weiteren Winkels abhängt, wobei der erste Winkel der Winkel ist, der die Richtung von der definierten Benutzerposition aus spezifiziert, in der sich der interessierende Bereich befindet, und wobei der weitere Winkel, von der definierten Benutzerposition und von der Position der Metadaten dieses Audioobjekts abhängt. Dabei kann die Prozessoreinheit 120 beispielsweise ausgebildet sein, für jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe von Audioobjekten, den Gewichtungsfaktor dieses Audioobjekts zusammen mit dem wenigstens eine Effekt-Parameter, der mittels der Schnittstelle 110 spezifiziert wurde, auf das Audioobjektsignal oder auf den Verstärkungsparameter der Audioobjekt-Metadaten dieses Audioobjekts anzuwenden.

In einer Ausführungsform kann die Verarbeitungsobjekt-Gruppe von Audioobjekten beispielsweise eine erste Verarbeitungsobjekt-Gruppe von Audioobjekten sein, wobei z.B. zudem ein oder mehrere weitere Verarbeitungsobjekt-Gruppen von Audioobjekten existieren können.

Dabei kann jede Verarbeitungsobjekt-Gruppe der ein oder mehrere weiteren Verarbeitungsobjekt-Gruppen von Audioobjekten ein oder mehrere Audioobjekte der Mehrzahl von Audioobjekten umfassen, wobei wenigstens ein Audioobjekt einer Verarbeitungsobjekt-Gruppe der ein oder mehrere weitere Verarbeitungsobjekt-Gruppen von Audioobjekten kein Audioobjekt der ersten Verarbeitungsobjekt-Gruppe von Audioobjekten ist.

Hierbei kann die Schnittstelle 110 für jede Verarbeitungsobjekt-Gruppe der ein oder mehreren weiteren Verarbeitungsobjekt-Gruppen von Audioobjekten zum Spezifizieren we-

nigstens eines weiteren Effekt-Parameters für diese Verarbeitungsobjekt-Gruppe von Audioobjekten durch den Benutzer ausgebildet ist.

Dabei kann die Prozessoreinheit 120 ausgebildet sein, das verarbeitete Signal so zu erzeugen, dass für jede Verarbeitungsobjekt-Gruppe der ein oder mehreren weiteren Verarbeitungsobjekt-Gruppen von Audioobjekten der wenigstens eine weitere Effekt-Parameter dieser Verarbeitungsobjekt-Gruppe, der mittels der Schnittstelle 110 spezifiziert wurde, auf das Audioobjektsignal oder auf die Audioobjekt-Metadaten jedes der ein oder mehreren Audioobjekte dieser Verarbeitungsobjekt-Gruppe angewendet wird, wobei ein oder mehrere Audioobjekte der Mehrzahl von Audioobjekten nicht dieser Verarbeitungsobjekt-Gruppe angehören.

Hierbei kann die Prozessoreinheit 120 beispielsweise ausgebildet sein, den wenigstens einen weiteren Effekt-Parameter dieser Verarbeitungsobjekt-Gruppe, der mittels der Schnittstellte spezifiziert wurde, auf kein Audioobjektsignal und keine Audioobjekt-Metadaten der ein oder mehreren Audioobjekte anzuwenden, die dieser Verarbeitungsobjekt-Gruppe nicht angehören.

Es können in solchen Ausführungsformen also mehr als eine Verarbeitungsobjekt-Gruppe existieren. Für jede der Verarbeitungsobjektgruppen werden ein oder mehrere eigene Effekt-Parameter bestimmt.

Gemäß einer Ausführungsform kann die Schnittstelle 110 zusätzlich zu der ersten Verarbeitungsobjekt-Gruppe von Audioobjekten beispielsweise zum Spezifizieren der ein oder mehreren weiteren Verarbeitungsobjekt-Gruppen von ein oder mehreren Audioobjekten durch den Benutzer ausgebildet sein, indem die Schnittstelle 110 für jede Verarbeitungsobjekt-Gruppe der ein oder mehreren weiteren Verarbeitungsobjekt-Gruppen von ein oder mehreren Audioobjekten zum Spezifizieren von wenigstens einem Definitions-Parameter dieser Verarbeitungsobjekt-Gruppe durch den Benutzer ausgebildet ist.

Dabei kann die Prozessoreinheit 120 beispielsweise ausgebildet sein, für jede Verarbeitungsobjekt-Gruppe der ein oder mehreren weiteren Verarbeitungsobjekt-Gruppen von ein oder mehreren Audioobjekten in Abhängigkeit von dem wenigstens einen Definitions-Parameter dieser Verarbeitungsobjekt-Gruppe, der mittels der Schnittstelle 110 spezifiziert wurde, zu bestimmen, welche Audioobjekte der Mehrzahl von Audioobjekten dieser Verarbeitungsobjekt-Gruppe angehören.

Im Folgenden werden Konzepte von Ausführungsformen der Erfindung und bevorzugte Ausführungsformen dargestellt.

In Ausführungsformen werden jegliche Arten globaler Anpassungen in OBAP möglich gemacht, indem globale Anpassungen in individuelle Änderungen der betroffenen Audioobjekte umgewandelt werden (z.B. durch die Prozessoreinheit 120).

Räumliches Mastering für objektbasierte Audioproduktion kann dabei beispielsweise wie folgt realisiert werden, indem erfindungsgemäße Verarbeitungsobjekte realisiert werden.

Die vorgeschlagene Implementierung von Gesamtanpassungen wird über Verarbeitungsobjekte (englisch: Processing Objects, POs) umgesetzt. Diese können genauso wie gewöhnliche Audioobjekte an beliebiger Stelle in einer Szene und frei in Echtzeit positioniert werden. Der Benutzer kann eine beliebige Signalverarbeitung auf das Verarbeitungsobjekt (auf die Verarbeitungsobjekt-Gruppe) anwenden, beispielsweise Equalizer (EQ) oder Kompression. Für jedes dieser Verarbeitungstools können die Parametereinstellungen des Verarbeitungsobjekts in objektspezifische Einstellungen umgewandelt werden. Für diese Berechnung werden verschiedene Verfahren vorgestellt.

Nachfolgend wird ein interessierender Bereich betrachtet.

Fig. 5 zeigt ein Verarbeitungsobjekt mit dem Bereich A und dem Fading-Bereich A_f gemäß einer Ausführungsform.

Wie in Fig. 5 gezeigt ist, definiert der Benutzer einen Bereich A und einen Ausblendungsbereich A_f um das Verarbeitungsobjekt herum. Die Verarbeitungsparameter des Verarbeitungsobjekts sind in konstante Parameter und gewichtete Parameter aufgeteilt. Werte von konstanten Parametern werden unverändert durch alle Audioobjekte innerhalb von A und A_f vererbt. Gewichtete Parameterwerte werden lediglich durch Audioobjekte innerhalb von A vererbt. Audioobjekte innerhalb von A_f werden mit einem Abstandsfaktor gewichtet. Die Entscheidung, welche Parameter gewichtet werden und welche nicht, hängt von dem Parametertyp ab.

Den benutzerdefinierten Wert p_M eines derartigen gewichteten Parameters für das Verarbeitungsobjekt gegeben, wird für jedes Audioobjekt S_i, die Parameterfunktion p_i wie folgt definiert:

$$p_i(t) = \left\{ \begin{array}{ll} p_M(t), & \text{for } S_i \in A \\ p_M(t) * f_i(t), & \text{for } S_i \in A_f \\ 0, & \text{else.} \end{array} \right\}, \quad (1)$$

wobei der Faktor f_i wie folgt gegeben ist:

$$f_i(t) = \frac{r_{A_f} - r_{S_i}}{r_{A_f} - r_A}. \quad (2)$$

Folglich gilt, falls der Benutzer $r_A = 0$ festlegt, dass kein Gültigkeitsbereich vorhanden ist, innerhalb dessen gewichtete Parameter konstant gehalten werden.

Im Folgenden wird eine Berechnung inverser Parameter gemäß einer Ausführungsform beschrieben.

Fig. 6 zeigt ein Verarbeitungsobjekt mit dem Bereich A und Objektradien gemäß einer Ausführungsform.

Benutzeranpassungen an dem Verarbeitungsobjekt, die über die Gleichung (1) umgewandelt werden, führen ggf. nicht immer schnell genug zu den gewünschten Ergebnissen, da die genaue Position von Audioobjekten nicht berücksichtigt wird. Wenn beispielsweise der Bereich um das Verarbeitungsobjekt herum sehr groß ist und die enthaltenen Audioobjekte von der Verarbeitungsobjekt-Position weit entfernt sind, kann der Effekt berechneter Anpassungen unter Umständen nicht einmal an der Verarbeitungsobjekt-Position hörbar sein.

Für Verstärkungsparameter ist ein anderes Berechnungsverfahren auf der Basis der Abklingrate jedes Objekts vorstellbar. Erneut wird innerhalb eines benutzerdefinierten interessierenden Bereichs, der in Fig. 6 dargestellt ist, der individuelle Parameter p_i für jedes Audioobjekt dann wie folgt berechnet.

$$p_i(t) = \left\{ \begin{array}{ll} h_i(t), & \text{for } S_i \in A \\ 0, & \text{else.} \end{array} \right\}, \quad (3)$$

wobei h_i wie folgt definiert sein könnte

$$h_i(t) = \operatorname{sgn} g_e(t) * (|g_e(t)| + |10 * \log_{10}(\frac{a_i}{d_i(t)})^2|). \quad (4)$$

a_i ist eine Konstante für den nächstmöglichen Abstand zu einem Audioobjekt, und $d_i(t)$ ist der Abstand von dem Audioobjekt zu dem EQ-Objekt. Abgeleitet von dem Abstandsgesetz wurde die Funktion geändert, um mögliche positive oder negative EQ-Verstärkungsänderungen korrekt zu handhaben.

In der nachfolgenden modifizierten Ausführungsform erfolgt eine winkelbasierte Berechnung.

Die vorherigen Berechnungen basieren auf dem Abstand zwischen Audioobjekten und dem Verarbeitungsobjekt. Aus einer Benutzerperspektive kann jedoch der Winkel zwischen dem Verarbeitungsobjekt und den umgebenden Audioobjekten deren Höreindruck gelegentlich akkurater darstellen. [5] schlägt die globale Steuerung eines beliebigen Audio-Plugin-Parameters über den Azimut von Audioobjekten vor. Dieser Ansatz kann übernommen werden, indem die Differenz bezüglich des Winkels α_i zwischen dem Verarbeitungsobjekt mit Versatzwinkel α_{eq} und Audioobjekten S_i in dessen Umkreis berechnet wird, wie in Fig. 7 dargestellt ist.

So zeigt Fig. 7 einen relativen Winkel von Audioobjekten zu dem Verarbeitungsobjekt gemäß einer Ausführungsform.

Der benutzerdefinierte, interessierende Bereich, der oben angesprochen ist, könnte entsprechend unter Verwendung der Winkel α_A und α_{A_f}geändert werden, was in Fig. 8 gezeigt ist.

So zeigt Fig. 8 ein Equalizer-Objekt mit neuem radialem Umkreis nach einer Ausführungsform.

Bezüglich des Ausblendungsbereichs müsste A_f, f_i wie folgt neu definiert werden:

$$f_i(t) = \frac{\alpha_{A_f} - \alpha_{S_i}}{\alpha_{A_f} - \alpha_A}. \quad (5)$$

Obwohl für den geänderten Ansatz, der in oben vorgestellt ist, der Abstand d_i in diesem Zusammenhang einfach als der Winkel zwischen dem Audioobjekt und dem EQ-Objekt interpretiert werden könnte, würde dies nicht länger ein Anwenden des Abstandsgesetzes rechtfertigen. Deshalb wird lediglich der benutzerdefinierte Bereich geändert, während die Verstärkungsberechnung wie zuvor beibehalten wird.

In einer Ausführungsform wird als Anwendung Entzerrung realisiert.

Entzerrung kann beim Mastering als das wichtigste Werkzeug betrachtet werden, da das Frequenzansprechverhalten eines Mixes der kritischste Faktor für eine gute Übersetzung über Wiedergabesysteme hinweg ist.

Die vorgeschlagene Implementierung einer Entzerrung wird über EQ-Objekte realisiert. Da alle weiteren Parameter nicht abstandsabhängig sind, ist lediglich der Verstärkungsparameter von besonderem Interesse.

In einer weiteren Ausführungsform wird als Anwendung dynamische Steuerung realisiert.

Bei herkömmlichem Mastering wird dynamische Kompression verwendet, um dynamische Abweichungen in einem Mix über die Zeit zu steuern. Abhängig von den Kompressionseinstellungen ändert dies die empfundene Dichte und das Transientenansprechverhalten eines Mixes. Im Fall einer festen Kompression wird die empfundene Änderung der Dichte als ‚glue' (engl. für „kleben, zusammenkleben") bezeichnet, während stärkere Kompressionseinstellungen für Pump- oder Seitenketteneffekte auf sog. Beat-Heavy-Mixes verwendet werden können.

Bei OBAP könnte der Benutzer ohne Weiteres identische Kompressionseinstellungen für mehrere benachbarte Objekte festlegen, um eine Mehrkanalkompression zu realisieren. Jedoch wäre die summierte Kompression auf Gruppen von Audioobjekten nicht nur für zeitkritische Arbeitsabläufe vorteilhaft, sondern es wäre auch wahrscheinlicher, dass der psychoakustische Eindruck von sog. „glued" Signalen erfüllt würde.

Fig. 9 zeigt einen Signalfluss einer Kompression der Signal von *n* Quellen gemäß einer Ausführungsform.

Gemäß einer weiteren Ausführungsform wird als Anwendung Szenentransformation realisiert.

Beim Stereomastering ist ein Mitten-/Seitenverarbeiten eine häufig verwendete Technik zum Erweitern oder Stabilisieren des Stereobilds eines Mixes. Für räumliche Audiomixes kann eine ähnliche Option hilfreich sein, falls der Mix in einer akustisch kritischen Umgebung mit potentiell asymmetrischen Raum- oder Lautsprechereigenschaften erzeugt wurde. Es könnten auch neue kreative Möglichkeiten für den ME bereitgestellt werden, um die Auswirkungen eines Mixes zu verbessern.

Fig. 10 zeigt eine Szenentransformation unter Verwendung einer Steuertafel *M* nach einer Ausführungsform. Speziell zeigt Fig. 10 eine schematische Umsetzung unter Verwendung eines Verzerrungsbereichs mit benutzerziehbaren Rändern C_1 bis C_4.

Eine zweidimensionale Transformation einer Szene in der horizontalen Ebene kann unter Verwendung einer Homographie-Transformationsmatrix *H* realisiert werden, die jedes Audioobjekt an Position p_i auf eine neue Position p'_i abbildet, siehe auch [7]:

$$H := \begin{pmatrix} h1 & h2 & h3 \\ h4 & h5 & h6 \\ h7 & h8 & h9 \end{pmatrix}, \; p'_i = H\,p_i. \tag{6}$$

Wenn der Benutzer mit einem Steuerfeld *M* zu *M'* unter Verwendung der vier ziehbaren Ecken C_{1-4} verzerrt (siehe Figur 6), können deren 2D-Koordinaten $\begin{bmatrix} x_{1-4} \\ y_{1-4} \end{bmatrix}$ für ein lineares System von Gleichungen verwendet werden (7), um die Koeffizienten von *H* zu erhalten [7].

$$\begin{pmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x'_1x_1 & -x'_1y_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -y'_1x_1 & -y'_1y_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x'_2x_2 & -x'_2y_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -y'_2x_2 & -y'_2y_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -x'_3x_3 & -x'_3y_3 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -y'_3x_3 & -y'_3y_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x'_4x_4 & -x'_4y_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -y'_4x_4 & -y'_4y_4 \end{pmatrix} * \begin{pmatrix} h1 \\ h2 \\ h3 \\ h4 \\ h5 \\ h6 \\ h7 \\ h8 \end{pmatrix} = \begin{pmatrix} x'_1 \\ y'_1 \\ x'_2 \\ y'_2 \\ x'_3 \\ y'_3 \\ x'_4 \\ y'_4 \end{pmatrix} \quad (7)$$

Da Audioobjektpositionen über die Zeit variieren können, können die Koordinatenpositionen als zeitabhängige Funktionen interpretiert werden.

In Ausführungsformen sind dynamische Equalizer realisiert. Andere Ausführungsformen realisieren Mehrbandkompression.

Objektbasierte Schallanpassungen sind nicht auf die eingeführten Equalizer-Anwendungen beschränkt.

Die obige Beschreibung wird im Folgenden noch einmal durch eine generellere Beschreibung von Ausführungsbeispielen ergänzt.

Objektbasierte dreidimensionale Audioproduktion verfolgt den Ansatz, dass über einen Rendering-Prozess Audioszenen in Echtzeit für weitestgehend beliebige Lautsprecherkonfigurationen berechnet und wiedergegeben werden. Audioszenen beschreiben zeitabhängig die Anordnung von Audioobjekten. Audioobjekte bestehen aus Audiosignalen und Metadaten. Zu diesen Metadaten gehören u.a. Position im Raum und Lautstärke. Um die Szene zu bearbeiten, muss der Benutzer bisher alle Audioobjekte einer Szene einzeln ändern.

Wenn im Folgenden einerseits von Verarbeitungsobjekt-Gruppe und andererseits von einem Verarbeitungsobjekt (englisch: Processing Object) die Rede ist, so ist festzustellen, dass für jedes Verarbeitungsobjekt immer eine Verarbeitungsobjekt-Gruppe definiert ist, die Audioobjekte umfasst. Die Verarbeitungsobjekt-Gruppe wird dabei beispielsweise auch als Container des Verarbeitungsobjekts bezeichnet. Für jedes Verarbeitungsobjekt ist also eine Gruppe von Audioobjekten aus der Mehrzahl von Audioobjekten definiert. Die

entsprechende Verarbeitungsobjekt-Gruppe umfasst die so spezifizierte Gruppe von Audioobjekten. Eine Verarbeitungsobjekt-Gruppe ist also eine Gruppe von Audioobjekten.

Verarbeitungsobjekte (englisch: Processing Objects) können als Objekte definiert werden, die die Eigenschaften anderer Audioobjekte verändern können. Verarbeitungsobjekte sind künstliche Container, denen beliebige Audioobjekte zugeordnet werden können, d.h. über den Container werden alle seine zugeordneten Audioobjekte angesprochen. Über eine beliebige Anzahl an Effekten werden die zugeordneten Audioobjekte beeinflusst. Somit bieten Verarbeitungsobjekte dem Benutzer die Möglichkeit, mehrere Audioobjekte simultan zu bearbeiten.

Ein Verarbeitungsobjekt weist beispielweise Position, Zuordnungsverfahren, Container, Gewichtungsverfahren, Audiosignalverarbeitungseffekte und Metadateneffekte auf.

Die Position ist dabei eine Position des Verarbeitungsobjekts in einer virtuellen Szene.

Das Zuordnungsverfahren ordnet Audioobjekte dem Verarbeitungsobjekt zu (ggf. unter Verwendung von deren Position).

Der Container (bzw. Connections) ist die Menge aller dem Verarbeitungsobjekt zugeordneten Audioobjekte (oder ggf. zusätzlicher anderer Verarbeitungsobjekte).

Gewichtungsverfahren sind die Algorithmen zur Berechnung der individuellen Effektparameterwerte für die zugeordneten Audioobjekte.

Audiosignalverarbeitungseffekte verändern die Audiokomponente von Audioobjekten (z.B. Equalizer, Dynamics).

Metadateneffekte verändern die Metadaten von Audioobjekten und/oder Verarbeitungsobjekten (z.B. Positionsverzerrung).

Gleichermaßen können der Verarbeitungsobjekt-Gruppe die oben-beschriebene Position, das Zuordnungsverfahren, der Container, Gewichtungsverfahren, Audiosignalverarbeitungseffekte und Metadateneffekte zugewiesen sein. Dabei sind die Audioobjekte des Containers des Verarbeitungsobjekts die Audioobjekte der Verarbeitungsobjekt-Gruppe.

Fig. 11 zeigt den Zusammenhang eines Verarbeitungsobjekts, mit dem Audiosignal-Effekte und Metadaten-Effekte bewirkt werden, gemäß einer Ausführungsform.

Im Folgenden werden Eigenschaften von Verarbeitungsobjekten gemäß speziellen Ausführungsformen beschrieben:

Verarbeitungsobjekte können vom Benutzer beliebig in einer Szene platziert werden, die Position kann über die Zeit konstant oder zeitabhängig gesetzt werden.

Verarbeitungsobjekte können vom Benutzer mit Effekten belegt werden, die das Audiosignal und/oder die Metadaten von Audioobjekten verändern. Beispiele für Effekte sind Entzerrung des Tonsignals, Bearbeitung der Dynamik des Tonsignals, oder Veränderung der Positionskoordinaten von Audioobjekten.

Verarbeitungsobjekte können mit beliebig vielen Effekten in beliebiger Reihenfolge belegt werden.

Effekte verändern das Audiosignal und/oder die Meta-Daten der zugeordneten Menge von Audioobjekten, jeweils über die Zeit konstant oder zeitabhängig.

Effekte haben Parameter zur Steuerung der Signal- und/oder Metadaten-Bearbeitung. Diese Parameter werden vom Benutzer, oder festgelegt je nach Typ, in konstante und gewichtete Parameter aufgeteilt.

Die Effekte eines Verarbeitungsobjekts werden auf seine zugeordneten Audioobjekte kopiert und angewandt. Die Werte konstanter Parameter werden dabei von jedem Audioobjekt unverändert übernommen. Die Werte gewichteter Parameter werden für jedes Audioobjekt individuell nach verschiedenen Gewichtungsverfahren berechnet. Der Benutzer kann für jeden Effekt ein Gewichtungsverfahren wählen, bzw. dieses für einzelne Audio-Quellen aktivieren oder deaktivieren.

Die Gewichtungsverfahren berücksichtigen individuelle Metadaten und/oder Signalcharakteristiken einzelner Audioobjekte. Dies entspricht zum Beispiel dem Abstand eines Audioobjektes zum Verarbeitungsobjekt oder dem Frequenzspektrum eines Audioobjekts. Die Gewichtungsverfahren können auch die Hörposition des Hörers berücksichtigen. Ferner können die genannten Eigenschaften von Audioobjekten für die Gewichtungsverfahren auch miteinander kombiniert werden, um daraus individuelle Parameterwerte abzuleiten. Beispielsweise können die Schallpegel von Audioobjekten im Rahmen einer Dynamikbearbeitung addiert werden, um daraus für jedes Audioobjekt individuell eine Veränderung der Lautstärke abzuleiten.

Effekt-Parameter können über die Zeit konstant oder zeitabhängig gesetzt werden. Die Gewichtungsverfahren berücksichtigen solche zeitlichen Änderungen.

Gewichtungsverfahren können auch Informationen, welche der Audio Renderer aus der Szene analysiert, verarbeiten.

Die Reihenfolge der Belegung des Verarbeitungsobjekts mit Effekten entspricht der Abfolge der Bearbeitung von Signalen und/oder Metadaten jedes Audioobjekts, d. h. die von einem vorhergehenden Effekt veränderten Daten werden vom nächsten Effekt als Basis für seine Berechnung verwendet. Der erste Effekt arbeitet auf den noch unveränderten Daten eines Audioobjekts.

Einzelne Effekte können deaktiviert werden. Dann werden die berechneten Daten des vorherigen Effekts, sofern einer existiert, an den Effekt nach dem deaktivierten Effekt weitergeleitet.

Ein explizit neu entwickelter Effekt ist die Veränderung der Position von Audioobjekten mittels Homographie („Verzerreffekt"). Dem Benutzer wird hierbei an der Position des Verarbeitungsobjekts ein Rechteck mit individuell verschiebbaren Ecken angezeigt. Verschiebt der Benutzer eine Ecke, wird aus dem vorherigen Zustand des Rechtecks und dem neu verzerrten Zustand eine Transformations-Matrix für diese Verzerrung berechnet. Die Matrix wird anschließend auf alle Positionskoordinaten der dem Verarbeitungsobjekt zugeordneten Audioobjekte angewandt, sodass sich deren Position entsprechend der Verzerrung ändert.

Effekte, die nur Meta-Daten verändern, können auch auf andere Verarbeitungsobjekte angewandt werden (u.a. „Verzerreffekt").

Die Zuordnung von Audio-Quellen zu den Verarbeitungsobjekten kann auf verschiedene Weise erfolgen. Die Menge an zugeordneten Audioobjekten kann sich je nach Art der Zuordnung auch über die Zeit ändern. Diese Änderung wird von allen Berechnungen berücksichtigt.

Es kann ein Einzugsbereich um die Position von Verarbeitungsobjekten definiert werden.

Alle Audioobjekte, die innerhalb des Einzugsbereichs positioniert sind, bilden die zugeordnete Menge an Audioobjekten, auf die die Effekte des Verarbeitungsobjekts angewandt werden.

Der Einzugsbereich kann ein beliebiger Körper (dreidimensional) oder eine beliebige Form (zweidimensional) sein, der bzw. die vom Benutzer definiert wird.

Der Mittelpunkt des Einzugsbereichs kann, aber musst nicht der Position des Verarbeitungsobjekt entsprechen. Der Benutzer trifft diese Festlegung.

Innerhalb eines dreidimensionalen Einzugsbereichs liegt ein Audioobjekt, wenn seine Position innerhalb des dreidimensionalen Körpers liegt.

Innerhalb eines zweidimensionalen Einzugsbereichs liegt ein Audioobjekt, wenn seine auf die horizontale Ebene projizierten Position innerhalb der zweidimensionalen Form liegt.

Der Einzugsbereich kann eine nicht weiter spezifizierte allumfassende Größe annehmen, sodass sich alle Audioobjekte einer Szene im Einzugsbereich befinden.

Die Einzugsbereiche passen sich ggf. Veränderung der Szeneneigenschaften (z.B. Szenenskalierung) an.

Unabhängig vom Einzugsbereich können Verarbeitungsobjekte mit einer beliebigen Auswahl von Audioobjekten einer Szene gekoppelt werden.

Die Kopplung kann vom Benutzer so definiert werden, dass alle ausgewählten Audioobjekte eine Menge an Audioobjekten bilden, auf die die Effekte des Verarbeitungsobjekts angewandt werden.

Die Kopplung kann alternativ vom Benutzer so definiert werden, dass das Verarbeitungsobjekt seine Position zeitabhängig nach der Position der ausgewählten Audioobjekte anpasst. Diese Anpassung der Position kann die Hörposition des Hörers berücksichtigen. Dabei müssen die Effekte des Verarbeitungsobjekts nicht zwangsläufig auf die gekoppelten Audioobjekte angewandt werden.

Die Zuordnung kann automatisch anhand vom Benutzer definierter Kriterien erfolgen. Dabei werden kontinuierlich alle Audioobjekte einer Szene auf das bzw. die definierten Kriterien untersucht und bei Erfüllen des bzw. der Kriterien dem Verarbeitungsobjekt zugeord-

net. Die Dauer der Zuordnung kann auf die Zeit der Erfüllung des bzw. der Kriterien beschränkt werden, oder es können Übergangszeiträume definiert werden. Die Übergangszeiträume bestimmen, wie lange ein bzw. mehrere Kriterien kontinuierlich vom Audioobjekt erfüllt sein müssen, damit es dem Verarbeitungsobjekt zugeordnet wird bzw. wie lange ein bzw. mehrere Kriterien kontinuierlich verletzt sein müssen, damit die Zuordnung zum Verarbeitungsobjekt wieder aufgelöst wird.

Verarbeitungsobjekte können vom Benutzer deaktiviert werden, sodass deren Eigenschaften erhalten bleiben und dem Benutzer weiterhin angezeigt werden, aber keine Beeinflussung von Audioobjekten durch das Verarbeitungsobjekt stattfindet.

Beliebig viele Eigenschaften eines Verarbeitungsobjekts können vom Benutzer mit gleichartigen Eigenschaften beliebig vieler anderer Verarbeitungsobjekte gekoppelt werden. Diese Eigenschaften schließen Parameter von Effekten mit ein. Die Kopplung kann vom Benutzer absolut oder relativ gewählt werden. Bei konstanter Kopplung wird der geänderte Eigenschaftswert eines Verarbeitungsobjekts von allen gekoppelten Verarbeitungsobjekten exakt übernommen. Bei relativer Kopplung wird der Wert der Veränderung mit den Eigenschaftswerten gekoppelter Verarbeitungsobjekte verrechnet.

Verarbeitungsobjekte können dupliziert werden. Dabei wird ein zweites Verarbeitungsobjekt mit identischen Eigenschaften des ursprünglichen Verarbeitungsobjekte erzeugt. Die Eigenschaften der Verarbeitungsobjekte sind anschließend unabhängig voneinander.

Eigenschaften von Verarbeitungsobjekte können z.B. beim Kopieren dauerhaft vererbt werden, so dass Änderungen bei den Eltern automatisch bei den Kindern übernommen werden.

Fig. 12 zeigt die Veränderung von Audioobjekten und Audiosignalen auf eine Eingabe eines Benutzer hin gemäß einer Ausführungsform.

Eine weitere neue Anwendung von Verarbeitungsobjekten ist die intelligente Parameter-Berechnung mittels einer Szenenanalyse. Der Benutzer definiert über das Verarbeitungsobjekt an einer bestimmten Position Effektparameter. Der Audio Renderer macht eine vorausschauende Szenenanalyse, um zu detektieren, welche Audio-Quellen auf die Position des Verarbeitungsobjekts Einfluss haben. Daraufhin werden auf die selektierten Audioquellen unter Berücksichtigung der Szenenanalyse Effekte so angewandt, dass die

vom Benutzer definierten Effekteinstellungen bestmöglich an der Position des Verarbeitungsobjekts erreicht werden.

Im Folgenden werden weitere Ausführungsbeispiele der Erfindung, die mittels der Fig. 13 - Fig. 25 visuell dargestellt sind, beschrieben.

So zeigt Fig. 13 Verarbeitungsobjekt PO_4 mit Rechteck M zur Verzerrung der Ecken C_1, C_2, C_3 und C_4 durch den Benutzer. So zeigt Fig. 13 schematisch eine mögliche Verzerrung hin zu M' mit den Ecken C_1', C_2', C_3' und C_4', sowie die entsprechende Auswirkung auf die Quellen S_1, S_2, S_3 und S_4 mit ihren neuen Positionen S_1', S_2', S_3' und S_4'.

Fig. 14 zeigt Verarbeitungsobjekte PO_1 und PO_2 mit ihren jeweiligen, sich überlappenden zweidimensionalen Einzugsbereichen A und B, sowie die Abstände a_{S_1}, a_{S_2} und a_{S_3} bzw. b_{S_3}, b_{S_4} und b_{S_6} vom jeweiligen Verarbeitungsobjekt zu den durch die Einzugsbereiche zugeordneten Quellen S_1, S_2, S_3, S_4 und S_6.

Fig. 15 zeigt Verarbeitungsobjekt PO_3 mit rechteckigem, zweidimensionalen Einzugsbereich C und den Winkeln zwischen PO_3 und den zugeordneten Quellen S_1, S_2 und S_3 für eine mögliche Gewichtung von Parametern, die die Hörposition des Hörers mit einbezieht. Die Winkel können durch die Differenz des Azimut der einzelnen Quellen und dem Azimut α_{po} von PO_3 bestimmt werden.

Fig. 16 zeigt mögliche schematische Implementation eines Equalizer-Effekts, der auf ein Verarbeitungsobjekt angewandt wurde. Über Knöpfe wie w neben jedem Parameter kann die Gewichtung für den jeweiligen Parameter aktiviert werden. m_1, m_2 und m_3 bieten für die genannten gewichteten Parameter Optionen für das Gewichtungsverfahren.

Fig. 17 zeigt das Verarbeitungsobjekt PO_5 mit einem dreidimensionalen Einzugsbereich D und den jeweiligen Abständen d_{S_1}, d_{S_2} und d_{S_3} zu den über den Einzugsbereich zugeordneten Quellen S_1, S_2 und S_3.

Fig. 18 zeigt prototypische Implementierung eines Verarbeitungsobjekts, auf das ein Equalizer angewandt wurde. Das türkise Objekt mit dem Wellensymbol auf der rechten Bildseite zeigt das Verarbeitungsobjekt in der Audio-Szene, das der Benutzer frei mit der Maus bewegen kann. Innerhalb der türkisen, transparenten homogenen Fläche um das Verarbeitungsobjekt werden die Equalizer-Parameter wie auf der linken Bildseite definiert auf die Audioobjekte Src1, Src2 und Src3 unverändert angewandt. Um die homogene

Kreisfläche zeigt die ins transparente verlaufende Schattierung den Bereich an, in dem alle Parameter bis auf die Gain-Parameter von den Quellen unverändert übernommen werden. Die Gain-Parameter des Equalizers werden hingegen je nach Abstand der Quellen zum Verarbeitungsobjekt gewichtet. Da sich nur Quelle Src4 und Quelle Src24 in diesem Bereich befinden, findet In diesem Fall nur für deren Parameter eine Gewichtung statt. Quelle Src22 wird vom Verarbeitungsobjekt nicht beeinflusst. Über den „Area"-Schieberegler kontrolliert der Benutzer die Größe des Radius der Kreisfläche um das Verarbeitungsobjekt. Über den „Feather"-Schieberegler kontrolliert er die Größe des Radius der umliegenden Übergangsfläche.

Fig. 19 zeigt ein Verarbeitungsobjekt wie in Fig. 18, nur an einer anderen Position und ohne Übergangsfläche. Auf die Quellen Src22 und Src4 werden alle Parameter des Equalizers unverändert übernommen. Die Quellen Src3, Src2, Src1 und Src24 werden nicht vom Verarbeitungsobjekt beeinflusst.

Fig. 20 zeigt ein Verarbeitungsobjekt mit einer über seinen Azimut definierten Fläche als Einzugsbereich, sodass die Quellen Src22 und Src4 dem Verarbeitungsobjekt zugeordnet werden. Die Spitze der Einzugsfläche in der Mitte der rechten Bildseite entspricht der Position des Hörers/Benutzers. Beim Bewegen des Verarbeitungsobjekts wird die Fläche entsprechend des Azimut mitbewegt. Über den „Area"-Schieberegler bestimmt der Benutzer die Größe des Winkels der Einzugsfläche. Die Änderung von einer kreisförmigen zur winkelbasierten Einzugsfläche erreicht der Benutzer über das untere Auswahlfeld über den „Area"-/"Feather"-Schiebereglern, das nun „radius" anzeigt.

Fig. 21 zeigt ein Verarbeitungsobjekt wie in Fig. 20, jedoch mit zusätzlichem Übergangsbereich, der vom Benutzer über den "Feather"-Schieberegler kontrolliert werden kann.

Fig. 22 zeigt mehrere Verarbeitungsobjekts in der Szene, mit unterschiedlichen Einzugsbereichen. Die grauen Verarbeitungsobjekts sind vom Benutzer deaktiviert worden, d. h. sie beeinflussen die Audioobjekte in ihrem Einzugsbereich nicht. Auf der linken Bildseite werden stets die Equalizer-Parameter des aktuell ausgewählten Verarbeitungsobjekts angezeigt. Die Auswahl wird durch eine dünne, helle türkise Linie um das Objekt angezeigt.

Fig. 23 zeigt das rote Quadrat auf der rechten Bildseite zeigt ein Verarbeitungsobjekt zur horizontalen Verzerrung der Position von Audioobjekten. Der Benutzer kann die Ecken mit der Maus in beliebige Richtung ziehen, um eine Verzerrung der Szene zu erreichen.

Fig. 24 zeigt die Szene, nachdem der Benutzer die Ecken des Verarbeitungsobjekts verzogen hat. Die Position aller Quellen hat sich entsprechend der Verzerrung verändert.

Fig. 25 zeigt eine mögliche Visualisierung der Zuordnung einzelner Audioobjekte zu einem Verarbeitungsobjekt.

Obwohl manche Aspekte im Zusammenhang mit einer Vorrichtung beschrieben wurden, versteht es sich, dass diese Aspekte auch eine Beschreibung des entsprechenden Verfahrens darstellen, sodass ein Block oder ein Bauelement einer Vorrichtung auch als ein entsprechender Verfahrensschritt oder als ein Merkmal eines Verfahrensschrittes zu verstehen ist. Analog dazu stellen Aspekte, die im Zusammenhang mit einem oder als ein Verfahrensschritt beschrieben wurden, auch eine Beschreibung eines entsprechenden Blocks oder Details oder Merkmals einer entsprechenden Vorrichtung dar. Einige oder alle der Verfahrensschritte können durch einen Hardware-Apparat (oder unter Verwendung eines Hardware-Apparats), wie zum Beispiel einen Mikroprozessor, einen programmierbaren Computer oder einer elektronischen Schaltung durchgeführt werden. Bei einigen Ausführungsbeispielen können einige oder mehrere der wichtigsten Verfahrensschritte durch einen solchen Apparat ausgeführt werden.

Je nach bestimmten Implementierungsanforderungen können Ausführungsbeispiele der Erfindung in Hardware oder in Software oder zumindest teilweise in Hardware oder zumindest teilweise in Software implementiert sein. Die Implementierung kann unter Verwendung eines digitalen Speichermediums, beispielsweise einer Floppy-Disk, einer DVD, einer BluRay Disc, einer CD, eines ROM, eines PROM, eines EPROM, eines EEPROM oder eines FLASH-Speichers, einer Festplatte oder eines anderen magnetischen oder optischen Speichers durchgeführt werden, auf dem elektronisch lesbare Steuersignale gespeichert sind, die mit einem programmierbaren Computersystem derart zusammenwirken können oder zusammenwirken, dass das jeweilige Verfahren durchgeführt wird. Deshalb kann das digitale Speichermedium computerlesbar sein.

Manche Ausführungsbeispiele gemäß der Erfindung umfassen also einen Datenträger, der elektronisch lesbare Steuersignale aufweist, die in der Lage sind, mit einem programmierbaren Computersystem derart zusammenzuwirken, dass eines der hierin beschriebenen Verfahren durchgeführt wird.

Allgemein können Ausführungsbeispiele der vorliegenden Erfindung als Computerprogrammprodukt mit einem Programmcode implementiert sein, wobei der Programmcode

dahin gehend wirksam ist, eines der Verfahren durchzuführen, wenn das Computerprogrammprodukt auf einem Computer abläuft.

Der Programmcode kann beispielsweise auch auf einem maschinenlesbaren Träger gespeichert sein.

Andere Ausführungsbeispiele umfassen das Computerprogramm zum Durchführen eines der hierin beschriebenen Verfahren, wobei das Computerprogramm auf einem maschinen-lesbaren Träger gespeichert ist. Mit anderen Worten ist ein Ausführungsbeispiel des erfindungsgemäßen Verfahrens somit ein Computerprogramm, das einen Programmcode zum Durchführen eines der hierin beschriebenen Verfahren aufweist, wenn das Computerprogramm auf einem Computer abläuft.

Ein weiteres Ausführungsbeispiel der erfindungsgemäßen Verfahren ist somit ein Datenträger (oder ein digitales Speichermedium oder ein computerlesbares Medium), auf dem das Computerprogramm zum Durchführen eines der hierin beschriebenen Verfahren aufgezeichnet ist. Der Datenträger oder das digitale Speichermedium oder das computerlesbare Medium sind typischerweise greifbar und/oder nicht flüchtig.

Ein weiteres Ausführungsbeispiel des erfindungsgemäßen Verfahrens ist somit ein Datenstrom oder eine Sequenz von Signalen, der bzw. die das Computerprogramm zum Durchführen eines der hierin beschriebenen Verfahren darstellt bzw. darstellen. Der Datenstrom oder die Sequenz von Signalen kann bzw. können beispielsweise dahin gehend konfiguriert sein, über eine Datenkommunikationsverbindung, beispielsweise über das Internet, transferiert zu werden.

Ein weiteres Ausführungsbeispiel umfasst eine Verarbeitungseinrichtung, beispielsweise einen Computer oder ein programmierbares Logikbauelement, die dahin gehend konfiguriert oder angepasst ist, eines der hierin beschriebenen Verfahren durchzuführen.

Ein weiteres Ausführungsbeispiel umfasst einen Computer, auf dem das Computerprogramm zum Durchführen eines der hierin beschriebenen Verfahren installiert ist.

Ein weiteres Ausführungsbeispiel gemäß der Erfindung umfasst eine Vorrichtung oder ein System, die bzw. das ausgelegt ist, um ein Computerprogramm zur Durchführung zumindest eines der hierin beschriebenen Verfahren zu einem Empfänger zu übertragen. Die Übertragung kann beispielsweise elektronisch oder optisch erfolgen. Der Empfänger kann beispielsweise ein Computer, ein Mobilgerät, ein Speichergerät oder eine ähnliche Vor-

richtung sein. Die Vorrichtung oder das System kann beispielsweise einen Datei-Server zur Übertragung des Computerprogramms zu dem Empfänger umfassen.

Bei manchen Ausführungsbeispielen kann ein programmierbares Logikbauelement (beispielsweise ein feldprogrammierbares Gatterarray, ein FPGA) dazu verwendet werden, manche oder alle Funktionalitäten der hierin beschriebenen Verfahren durchzuführen. Bei manchen Ausführungsbeispielen kann ein feldprogrammierbares Gatterarray mit einem Mikroprozessor zusammenwirken, um eines der hierin beschriebenen Verfahren durchzuführen. Allgemein werden die Verfahren bei einigen Ausführungsbeispielen seitens einer beliebigen Hardwarevorrichtung durchgeführt. Diese kann eine universell einsetzbare Hardware wie ein Computerprozessor (CPU) sein oder für das Verfahren spezifische Hardware, wie beispielsweise ein ASIC.

Die oben beschriebenen Ausführungsbeispiele stellen lediglich eine Veranschaulichung der Prinzipien der vorliegenden Erfindung dar. Es versteht sich, dass Modifikationen und Variationen der hierin beschriebenen Anordnungen und Einzelheiten anderen Fachleuten einleuchten werden. Deshalb ist beabsichtigt, dass die Erfindung lediglich durch den Schutzumfang der nachstehenden Patentansprüche und nicht durch die spezifischen Einzelheiten, die anhand der Beschreibung und der Erläuterung der Ausführungsbeispiele hierin präsentiert wurden, beschränkt sei.

Referenzen

[1] Coleman, P., Franck, A., Francombe, J., Liu, Q., Campos, T. D., Hughes, R., Menzies, D., Galvez, M. S., Tang, Y., Woodcock, J., Jackson, P., Melchior, F., Pike, C., Fazi, F., Cox, T., and Hilton, A., "An Audio-Visual System for Object-Based Audio: From Recording to Listening," IEEE Transactions on Multimedia, PP(99), pp. 1–1, 2018, ISSN 1520- 9210, doi:10.1109/TMM.2018.2794780.

[2] Gasull Ruiz, A., Sladeczek, C., and Sporer, T., "A Description of an Object-Based Audio Workflow for Media Productions," in Audio Engineering Society Conference: 57th International Conference: The Future of Audio Entertainment Technology, Cinema, Television and the Internet, 2015.

[3] Melchior, F., Michaelis, U., and Steffens, R., "Spatial Mastering - a new concept for spatial sound design in object-based audio scenes," in Proceedings of the International Computer Music Conference 2011, 2011.

[4] Katz, B. and Katz, R. A., Mastering Audio: The Art and the Science, Butterworth-Heinemann, Newton, MA, USA, 2003, ISBN 0240805453. AES Conference on Spatial Reproduction, Tokyo, Japan, 2018 August 6 – 9, Page 2

[5] Melchior, F., Michaelis, U., and Steffens, R., "Spatial Mastering - A New Concept for Spatial Sound Design in Object-based Audio Scenes," Proceedings of the International Computer Music Conference 2011, University of Huddersfield, UK, 2011.

[6] Sladeczek, C., Neidhardt, A., Böhme, M., Seeber, M., and Ruiz, A. G., "An Approach for Fast and Intuitive Monitoring of Microphone Signals Using a Virtual Listener," Proceedings, International Conference on Spatial Audio (ICSA), 21.2. - 23.2.2014, Erlangen, 2014

[7] Dubrofsky, E., Homography Estimation, Master's thesis, University of British Columbia, 2009.

[8] ISO/IEC 23003-2:2010 Information technology - MPEG audio technologies - Part 2: Spatial Audio Object Coding (SAOC); 2010

Patentansprüche

1. Vorrichtung zur Erzeugung eines verarbeiteten Signals unter Verwendung einer Mehrzahl von Audioobjekten, wobei jedes Audioobjekt der Mehrzahl von Audioobjekten ein Audioobjektsignal und Audioobjekt-Metadaten umfasst, wobei die Audioobjekt-Metadaten eine Position des Audioobjekts und einen Verstärkungsparameter des Audioobjekts umfassen, wobei die Vorrichtung umfasst:

eine Schnittstelle (110) zum Spezifizieren wenigstens eines Effekt-Parameters einer Verarbeitungsobjekt-Gruppe von Audioobjekten durch einen Benutzer, wobei die Verarbeitungsobjekt-Gruppe von Audioobjekten zwei oder mehrere Audioobjekte der Mehrzahl von Audioobjekten umfasst, und

eine Prozessoreinheit (120), die ausgebildet ist, das verarbeitete Signal so zu erzeugen, dass der wenigstens eine Effekt-Parameter, der mittels der Schnittstelle (110) spezifiziert wurde, auf das Audioobjektsignal oder auf die Audioobjekt-Metadaten jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe von Audioobjekten angewendet wird.

2. Vorrichtung nach Anspruch 1,

wobei ein oder mehrere Audioobjekte der Mehrzahl von Audioobjekten nicht der Verarbeitungsobjekt-Gruppe von Audioobjekten angehören, und wobei die Prozessoreinheit (120) ausgebildet ist, den wenigstens einen Effekt-Parameter, der mittels der Schnittstellte spezifiziert wurde, auf kein Audioobjektsignal und keine Audioobjekt-Metadaten der ein oder mehreren Audioobjekte anzuwenden, die der Verarbeitungsobjekt-Gruppe von Audioobjekten nicht angehören.

3. Vorrichtung nach Anspruch 2,

wobei die Prozessoreinheit (120) ausgebildet ist, das verarbeitete Signal so zu erzeugen, dass der wenigstens eine Effekt-Parameter, der mittels der Schnittstelle (110) spezifiziert wurde, auf das Audioobjektsignal jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe von Audioobjekten angewendet wird,

wobei die Prozessoreinheit (120) ausgebildet ist, den wenigstens einen Effekt-Parameter, der mittels der Schnittstellte spezifiziert wurde, auf kein Audioobjekt-

signal der ein oder mehreren Audioobjekte der Mehrzahl von Audioobjekten anzuwenden, die der Verarbeitungsobjekt-Gruppe von Audioobjekten nicht angehören.

4. Vorrichtung nach Anspruch 2 oder 3,

wobei die Prozessoreinheit (120) ausgebildet ist, das verarbeitete Signal so zu erzeugen, dass der wenigstens eine Effekt-Parameter, der mittels der Schnittstelle (110) spezifiziert wurde, auf den Verstärkungsparameter der Metadaten jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe von Audioobjekten angewendet wird,

wobei die Prozessoreinheit (120) ausgebildet ist, den wenigstens einen Effekt-Parameter, der mittels der Schnittstellte spezifiziert wurde, auf keinen Verstärkungsparameter der Audioobjekt-Metadaten der ein oder mehreren Audioobjekte der Mehrzahl von Audioobjekten anzuwenden, die der Verarbeitungsobjekt-Gruppe von Audioobjekten nicht angehören.

5. Vorrichtung nach einem der Ansprüche 2 bis 4,

wobei die Prozessoreinheit (120) ausgebildet ist, das verarbeitete Signal so zu erzeugen, dass der wenigstens eine Effekt-Parameter, der mittels der Schnittstelle (110) spezifiziert wurde, auf die Position der Metadaten jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe von Audioobjekten angewendet wird,

wobei die Prozessoreinheit (120) ausgebildet ist, den wenigstens einen Effekt-Parameter, der mittels der Schnittstellte spezifiziert wurde, auf keine Position der Audioobjekt-Metadaten der ein oder mehreren Audioobjekte der Mehrzahl von Audioobjekten anzuwenden, die der Verarbeitungsobjekt-Gruppe von Audioobjekten nicht angehören.

6. Vorrichtung nach einem der vorherigen Ansprüche,

wobei die Schnittstelle (110) zum Spezifizieren wenigstens eines Definitions-Parameters der Verarbeitungsobjekt-Gruppe von Audioobjekten durch den Benutzer ausgebildet ist,

wobei die Prozessoreinheit (120) ausgebildet ist in Abhängigkeit von dem wenigstens einen Definitions-Parameter der Verarbeitungsobjekt-Gruppe von Audioobjek-

ten, der mittels der Schnittstelle (110) spezifiziert wurde, zu bestimmen, welche Audioobjekte der Mehrzahl von Audioobjekten der Verarbeitungsobjekt-Gruppe von Audioobjekten angehören.

7. Vorrichtung nach Anspruch 6,

wobei der wenigstens eine Definitions-Parameter der Verarbeitungsobjekt-Gruppe von Audioobjekten zumindest eine Position eines interessierenden Bereichs umfasst, der der Verarbeitungsobjekt-Gruppe von Audioobjekten zugeordnet ist, und

wobei die Prozessoreinheit (120) ausgebildet ist, für jedes Audioobjekt der Mehrzahl von Audioobjekten in Abhängigkeit von der Position der Audioobjekt-Metadaten dieses Audioobjekts und in Abhängigkeit von der Position des interessierenden Bereichs zu bestimmen, ob dieses Audioobjekt der Verarbeitungsobjekt-Gruppe von Audioobjekten angehört.

8. Vorrichtung nach Anspruch 7,

wobei der wenigstens eine Definitions-Parameter der Verarbeitungsobjekt-Gruppe von Audioobjekten des Weiteren einen Radius des interessierenden Bereichs umfasst, der der Verarbeitungsobjekt-Gruppe von Audioobjekten zugeordnet ist, und

wobei die Prozessoreinheit (120) ausgebildet ist, für jedes Audioobjekt der Mehrzahl von Audioobjekten in Abhängigkeit von der Position der Audioobjekt-Metadaten dieses Audioobjekts und in Abhängigkeit von der Position des interessierenden Bereichs und in Abhängigkeit von dem Radius des interessierenden Bereichs zu entscheiden, ob dieses Audioobjekt der Verarbeitungsobjekt-Gruppe von Audioobjekten angehört.

9. Vorrichtung nach Anspruch 7 oder 8,

wobei die Prozessoreinheit (120) ausgebildet ist, für jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe von Audioobjekten einen Gewichtungsfaktor in Abhängigkeit von einem Abstand zwischen der Position der Audioobjekt-Metadaten dieses Audioobjekts und der Position des interessierenden Bereichs zu bestimmen, und

wobei die Prozessoreinheit (120) ausgebildet ist, für jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe von Audioobjekten, den Gewichtungsfaktor dieses Audioobjekts zusammen mit dem wenigstens eine Effekt-Parameter, der mittels der Schnittstelle (110) spezifiziert wurde, auf das Audioobjektsignal oder auf den Verstärkungsparameter der Audioobjekt-Metadaten dieses Audioobjekts anzuwenden.

10. Vorrichtung nach Anspruch 6,

wobei der wenigstens eine Definitions-Parameter der Verarbeitungsobjekt-Gruppe von Audioobjekten zumindest einen Winkel umfasst, der eine Richtung von einer definierten Benutzerposition aus spezifiziert, in der sich ein interessierender Bereich befindet, der der Verarbeitungsobjekt-Gruppe von Audioobjekten zugeordnet ist, und

wobei die Prozessoreinheit (120) ausgebildet ist, für jedes Audioobjekt der Mehrzahl von Audioobjekten in Abhängigkeit von der Position der Metadaten dieses Audioobjekts und in Abhängigkeit von dem Winkel, der die Richtung von der definierten Benutzerposition aus spezifiziert, in der sich der interessierende Bereich befindet, zu bestimmen, ob dieses Audioobjekt der Verarbeitungsobjekt-Gruppe von Audioobjekten angehört.

11. Vorrichtung nach Anspruch 10,

wobei die Prozessoreinheit (120) ausgebildet ist, für jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe von Audioobjekten einen Gewichtungsfaktor zu bestimmen, der von einer Differenz eines ersten Winkels und eines weiteren Winkels abhängt, wobei der erste Winkel der Winkel ist, der die Richtung von der definierten Benutzerposition aus spezifiziert, in der sich der interessierende Bereich befindet, und wobei der weitere Winkel, von der definierten Benutzerposition und von der Position der Metadaten dieses Audioobjekts abhängt,

wobei die Prozessoreinheit (120) ausgebildet ist, für jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe von Audioobjekten, den Gewichtungsfaktor dieses Audioobjekts zusammen mit dem wenigstens eine Effekt-Parameter, der mittels der Schnittstelle (110) spezifiziert wurde, auf das Audioobjektsignal oder auf den Verstärkungsparameter der Audioobjekt-Metadaten dieses Audioobjekts anzuwenden.

12. Vorrichtung nach einem der vorherigen Ansprüche,

wobei die Verarbeitungsobjekt-Gruppe von Audioobjekten eine erste Verarbeitungsobjekt-Gruppe von Audioobjekten ist, wobei zudem ein oder mehrere weitere Verarbeitungsobjekt-Gruppen von Audioobjekten existieren, wobei jede Verarbeitungsobjekt-Gruppe der ein oder mehrere weiteren Verarbeitungsobjekt-Gruppen von Audioobjekten ein oder mehrere Audioobjekte der Mehrzahl von Audioobjekten umfasst, wobei wenigstens ein Audioobjekt einer Verarbeitungsobjekt-Gruppe der ein oder mehrere weitere Verarbeitungsobjekt-Gruppen von Audioobjekten kein Audioobjekt der ersten Verarbeitungsobjekt-Gruppe von Audioobjekten ist,

wobei die Schnittstelle (110) für jede Verarbeitungsobjekt-Gruppe der ein oder mehreren weiteren Verarbeitungsobjekt-Gruppen von Audioobjekten zum Spezifizieren wenigstens eines weiteren Effekt-Parameters für diese Verarbeitungsobjekt-Gruppe von Audioobjekten durch den Benutzer ausgebildet ist,

wobei die Prozessoreinheit (120) ausgebildet ist, das verarbeitete Signal so zu erzeugen, dass für jede Verarbeitungsobjekt-Gruppe der ein oder mehreren weiteren Verarbeitungsobjekt-Gruppen von Audioobjekten der wenigstens eine weitere Effekt-Parameter dieser Verarbeitungsobjekt-Gruppe, der mittels der Schnittstelle (110) spezifiziert wurde, auf das Audioobjektsignal oder auf die Audioobjekt-Metadaten jedes der ein oder mehreren Audioobjekte dieser Verarbeitungsobjekt-Gruppe angewendet wird, wobei ein oder mehrere Audioobjekte der Mehrzahl von Audioobjekten nicht dieser Verarbeitungsobjekt-Gruppe angehören, und wobei die Prozessoreinheit (120) ausgebildet ist, den wenigstens einen weiteren Effekt-Parameter dieser Verarbeitungsobjekt-Gruppe, der mittels der Schnittstellte spezifiziert wurde, auf kein Audioobjektsignal und keine Audioobjekt-Metadaten der ein oder mehreren Audioobjekte anzuwenden, die dieser Verarbeitungsobjekt-Gruppe nicht angehören.

13. Vorrichtung nach Anspruch 12,

wobei die Schnittstelle (110) zusätzlich zu der ersten Verarbeitungsobjekt-Gruppe von Audioobjekten zum Spezifizieren der ein oder mehreren weiteren Verarbeitungsobjekt-Gruppen von ein oder mehreren Audioobjekten durch den Benutzer ausgebildet ist, indem die Schnittstelle (110) für jede Verarbeitungsobjekt-Gruppe der ein oder mehreren weiteren Verarbeitungsobjekt-Gruppen von ein oder mehre-

ren Audioobjekten zum Spezifizieren von wenigstens einem Definitions-Parameter dieser Verarbeitungsobjekt-Gruppe durch den Benutzer ausgebildet ist,

wobei die Prozessoreinheit (120) ausgebildet ist, für jede Verarbeitungsobjekt-Gruppe der ein oder mehreren weiteren Verarbeitungsobjekt-Gruppen von ein oder mehreren Audioobjekten in Abhängigkeit von dem wenigstens einen Definitions-Parameter dieser Verarbeitungsobjekt-Gruppe, der mittels der Schnittstelle (110) spezifiziert wurde, zu bestimmen, welche Audioobjekte der Mehrzahl von Audioobjekten dieser Verarbeitungsobjekt-Gruppe angehören.

14. Vorrichtung nach einem der vorherigen Ansprüche,

wobei die Vorrichtung ein Enkoder ist, wobei die Prozessoreinheit (120) ausgebildet ist, unter Verwendung der Audioobjektsignale der Mehrzahl von Audioobjekten ein Downmix-Signal zu erzeugen, und wobei die Prozessoreinheit (120) ausgebildet ist, unter Verwendung der Audioobjekt-Metadaten der Mehrzahl von Audioobjekten ein Metadatensignal-Signal zu erzeugen,

wobei die Prozessoreinheit (120) ausgebildet ist, als das verarbeitete Signal das Downmix-Signal zu erzeugen, wobei in dem Downmix-Signal zumindest ein modifiziertes Objektsignal für jedes Audioobjekt der Verarbeitungsobjekt-Gruppe von Audioobjekten gemischt ist, wobei die Prozessoreinheit (120) ausgebildet ist, für jedes Audioobjekt der Verarbeitungsobjekt-Gruppe von Audioobjekten das modifizierte Objektsignal dieses Audioobjekts mittels der Anwendung des wenigstens einen Effekt-Parameters, der mittels der Schnittstelle (110) spezifiziert wurde, auf das Audioobjektsignal dieses Audioobjekts zu erzeugen, oder

wobei die Prozessoreinheit (120) ausgebildet ist, als das verarbeitete Signal das Metadaten-Signal zu erzeugen, wobei das Metadaten-Signal zumindest eine modifizierte Position für jedes Audioobjekt der Verarbeitungsobjekt-Gruppe von Audioobjekten umfasst, wobei die Prozessoreinheit (120) ausgebildet ist, für jedes Audioobjekt der Verarbeitungsobjekt-Gruppe von Audioobjekten die modifizierte Position dieses Audioobjekts mittels der Anwendung des wenigstens einen Effekt-Parameters, der mittels der Schnittstelle (110) spezifiziert wurde, auf die Position dieses Audioobjekts zu erzeugen, oder

wobei die Prozessoreinheit (120) ausgebildet ist, als das verarbeitete Signal das Metadaten-Signal zu erzeugen, wobei das Metadaten-Signal zumindest einen

modifizierten Verstärkungsparameter für jedes Audioobjekt der Verarbeitungsobjekt-Gruppe von Audioobjekten umfasst, wobei die Prozessoreinheit (120) ausgebildet ist, für jedes Audioobjekt der Verarbeitungsobjekt-Gruppe von Audioobjekten der modifizierte Verstärkungsparameter dieses Audioobjekts mittels der Anwendung des wenigstens einen Effekt-Parameters, der mittels der Schnittstelle (110) spezifiziert wurde, auf den Verstärkungsparameter dieses Audioobjekts zu erzeugen.

15. Vorrichtung nach einem der Ansprüche 1 bis 13,

wobei die Vorrichtung ein Dekoder ist, wobei die Vorrichtung zum Empfang eines Downmix-Signals ausgebildet ist, in dem die Mehrzahl von Audioobjektsignalen der Mehrzahl von Audioobjekten gemischt sind, wobei die Vorrichtung des Weiteren zum Empfang eines Metadatensignals ausgebildet ist, wobei das Metadatensignal für jedes Audioobjekt der Mehrzahl von Audioobjekten die Audioobjekt-Metadaten dieses Audioobjekts umfasst,

wobei die Prozessoreinheit (120) ausgebildet ist, die Mehrzahl von Audioobjektsignalen der Mehrzahl von Audioobjekten basierend aus einem Downmix-Signal zu rekonstruieren,

wobei die Prozessoreinheit (120) ausgebildet ist, als das verarbeitete Signal ein Audioausgangssignal umfassend ein oder mehrere Audioausgangskanäle zu erzeugen,

wobei die Prozessoreinheit (120) ausgebildet ist, zum Erzeugen des verarbeiteten Signals den wenigstens einen Effekt-Parameter, der mittels der Schnittstelle (110) spezifiziert wurde, auf das Audioobjektsignal jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe von Audioobjekten anzuwenden, oder zum Erzeugen des verarbeiteten Signals den wenigstens einen Effekt-Parameter, der mittels der Schnittstelle (110) spezifiziert wurde, auf die Position oder auf den Verstärkungsparameter der Audioobjekt-Metadaten jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe von Audioobjekten anzuwenden.

16. Vorrichtung nach Anspruch 15,

wobei die Schnittstelle (110) des Weiteren zum Spezifizieren von ein oder mehreren Rendering-Parametern durch den Benutzer ausgebildet ist,

und wobei die Prozessoreinheit (120) ausgebildet ist, das verarbeitete Signal unter Verwendung der ein oder mehreren Rendering-Parameter in Abhängigkeit von der Position jedes Audioobjekts der Verarbeitungsobjekt-Gruppe von Audioobjekten zu erzeugen.

17. System, umfassend,

einen Enkoder (200) zur Erzeugung eines Downmix-Signals basierend auf Audioobjektsignalen einer Mehrzahl von Audioobjekten und zur Erzeugung eines Metadaten-Signals basierend auf Audioobjekt-Metadaten der Mehrzahl von Audioobjekten, wobei die Audioobjekt-Metadaten eine Position des Audioobjekts und einen Verstärkungsparameter des Audioobjekts umfassen, und

einen Dekoder (300) zur Erzeugung eines Audioausgangssignals umfassend ein oder mehrere Audioausgangskanäle basierend auf dem Downmix-Signal und basierend auf dem Metadaten-Signal,

wobei der Enkoder (200) eine Vorrichtung nach Anspruch 14 ist, oder

wobei der Dekoder (300) eine Vorrichtung nach Anspruch 15 oder 16 ist, oder

wobei der Enkoder (200) eine Vorrichtung nach Anspruch 14 und der Dekoder (300) eine Vorrichtung nach Anspruch 15 oder 16 ist.

18. Verfahren zur Erzeugung eines verarbeiteten Signals unter Verwendung einer Mehrzahl von Audioobjekten, wobei jedes Audioobjekt der Mehrzahl von Audioobjekten ein Audioobjektsignal und Audioobjekt-Metadaten umfasst, wobei die Audioobjekt-Metadaten eine Position des Audioobjekts und einen Verstärkungsparameter des Audioobjekts umfassen, wobei das Verfahren umfasst:

Spezifizieren von wenigstens einem Effekt-Parameter einer Verarbeitungsobjekt-Gruppe von Audioobjekten durch einen Benutzer mittels einer Schnittstelle (110), wobei die Verarbeitungsobjekt-Gruppe von Audioobjekten zwei oder mehrere Audioobjekte der Mehrzahl von Audioobjekten umfasst, und

Erzeugen des verarbeiteten Signals durch eine Prozessoreinheit (120) derart, dass der wenigstens eine Effekt-Parameter, der mittels der Schnittstellte spezifiziert

wurde, auf das Audioobjektsignal oder auf die Audioobjekt-Metadaten jedes der Audioobjekte der Verarbeitungsobjekt-Gruppe von Audioobjekten angewendet wird.

19. Computerprogramm mit einem Programmcode zur Durchführung des Verfahrens nach Anspruch 18.