



(19)
Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) **DE 699 10 360 T2** 2004.06.24

(12)

Übersetzung der europäischen Patentschrift

(97) **EP 1 062 816 B1**

(51) Int Cl.⁷: **H04N 7/52**

(21) Deutsches Aktenzeichen: **699 10 360.6**

(86) PCT-Aktenzeichen: **PCT/US99/05249**

(96) Europäisches Aktenzeichen: **99 913 844.9**

(87) PCT-Veröffentlichungs-Nr.: **WO 99/046938**

(86) PCT-Anmeldetag: **11.03.1999**

(87) Veröffentlichungstag

der PCT-Anmeldung: **16.09.1999**

(97) Erstveröffentlichung durch das EPA: **27.12.2000**

(97) Veröffentlichungstag

der Patenterteilung beim EPA: **13.08.2003**

(47) Veröffentlichungstag im Patentblatt: **24.06.2004**

(30) Unionspriorität:

42367 13.03.1998 US

(74) Vertreter:

**Hoffmann, E., Dipl.-Ing., Pat.-Anw., 82166
Gräfelfing**

(73) Patentinhaber:

**Dolby Laboratories Licensing Corp., San
Francisco, Calif., US**

(84) Benannte Vertragsstaaten:

AT, BE, CH, DE, DK, ES, FR, GB, IT, LI, NL, SE

(72) Erfinder:

TODD, Campbell, Craig, San Francisco, US

(54) Bezeichnung: **Audioinformationsverarbeitungsverfahren und -vorrichtung unter Verwendung von zeitangepassten kodierten Audioinformationsblöcken in Audio/Videoanwendungen zum Erleichtern von Tonumschaltung**

Anmerkung: Innerhalb von neun Monaten nach der Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents kann jedermann beim Europäischen Patentamt gegen das erteilte europäische Patent Einspruch einlegen. Der Einspruch ist schriftlich einzureichen und zu begründen. Er gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist (Art. 99 (1) Europäisches Patentübereinkommen).

Die Übersetzung ist gemäß Artikel II § 3 Abs. 1 IntPatÜG 1991 vom Patentinhaber eingereicht worden. Sie wurde vom Deutschen Patent- und Markenamt inhaltlich nicht geprüft.

Beschreibung**TECHNISCHES GEBIET**

[0001] Die vorliegende Erfindung bezieht sich ganz allgemein auf die Tonsignalverarbeitung bei Video/Audioanwendungen. Insbesondere betrifft die vorliegende Erfindung Blockkodierverfahren zum Kodieren von durch Lücken oder Schutzbänder getrennten Blockfolgen von Toninformation, so daß normale Schwankungen im Signalverarbeitungsverzug die Synchronisierung von Bild- und Toninformationen nicht aufheben.

STAND DER TECHNIK

[0002] Es sind schon mehrere internationale Normen entwickelt worden, in denen verschiedene Aspekte des Einbettens digitaler Toninformation in Rahmen von Bildinformation festgelegt sind. In der von der Society of Motion Picture and Television Engineers (SMPTE) veröffentlichten Norm SMPTE 259M ist beispielsweise eine serielle digitale Schnittstelle (Serial Digital Interface SDI) definiert, bei der bis zu vier Kanäle digitaler Toninformation in serielle digitale Teil- und zusammengesetzte Bildsignale eingefügt werden können. Eine vollständige Begriffsdefinition dessen, wie digitale Toninformation in Datenzusatzräume innerhalb von Rahmen von Videodaten einzubetten ist, findet sich in der Norm SMPTE 272M.

[0003] Die serielle Übertragung digitaler Audiodaten selbst ist Gegenstand verschiedener internationaler Normen. Beispielsweise ist in der von der Audio Engineering Society (AES) veröffentlichten Norm AES3 (ANSI S4.40) die serielle Übertragung von digitaler Zweikanal-Toninformation festgelegt, die in Form einer linearen Impulsmodulation (pulse code modulation (PCM) dargestellt ist. Gemäß dieser Norm werden PCM-Abtastwerte für die zwei Kanäle verschachtelt und paarweise weitergeleitet.

[0004] Nahezu allen Aufzeichnungs- und Rundfunkanwendungsfällen ist das Editieren und Zerschneiden eingebetteter Video/Audiodatenströme und das Spleißen der abgetrennten Datenströme zu einem einzigen neuen Strom gemeinsam. Durch ähnliches Vorgehen wird ein Datenstrom durch das Verschmelzen mehrerer Datenströme oder durch Umschalten zwischen mehreren Strömen erzeugt. Die Bildinformation ist normalerweise die primäre Bezugsquelle für das Synchronisieren, so daß eine Editier- oder Trennstelle normalerweise mit einem Bildrahmen fluchtet.

[0005] In Normen, wie AES11, werden Praktiken zum Synchronisieren digitaler Audiogeräte im Studiobetrieb empfohlen. Die AES11 ist darauf ausgerichtet, durch Flattern oder Verarbeitungsverzug verursachte Ungewißheiten der Zeitsteuerung unter Kontrolle zu bekommen und bietet an, Videorahmeninformation mit den zwei Abtastwerte umfassenden Rahmen von digitalen AES3 Audiodatenströmen auszu-

richten. Wenn sich Geräte und Verfahren an diese Norm halten, kann gewährleistet werden, daß synchronisierte Signale die gleiche Anzahl Rahmen über eine gegebene Zeitspanne hinweg haben und Abtastwerte enthalten, die eine gemeinsame Zeitsteuerung besitzen. Leider gibt es gegenwärtig keine Normen oder Praktiken, die eine Synchronisation zwischen Bildinformation und größeren Intervallen von Toninformation festlegen. Folglich gibt es bei Geräten unterschiedlicher Hersteller, und selbst vom gleichen Hersteller, Schwankungen in der Zeitsteuerung und im Verarbeitungsverzug, durch die ein beträchtliches Maß an Ungewißheit in die relative Synchronisierung von Ton- und Bildinformationen eingeführt wird.

[0006] Diese Ungewißheit der Synchronisation hat nur geringe Konsequenzen in Anwendungsfällen, bei denen lineare Wiedergaben von Toninformation verwendet werden, wie beispielsweise in der AES3 Norm festgelegt. Da es eine Eingrenzung der Editierstellen gibt, gemäß der sie zwischen den Rahmen mit zwei Abtastwerten der Toninformation auftreten, hat eine mögliche Ungewißheit in der Video/Audiosynchronisation keinen Verlust an Toninformation zur Folge. Sie beeinflusst nur die relative Zeitsteuerung von Ton und Bild so wie sie einer Person vorgestellt werden, und da ist es unwahrscheinlich, daß dies wahrnehmbar ist.

[0007] Es gibt allerdings eine wachsende Zahl von Anwendungen, bei denen Kodiertechniken mit einer Verringerung der Bitrate herangezogen werden, um größere Zahlen von Tonkanälen in einen Video/Audiodatenstrom einzufügen. Diese Kodiertechniken werden häufig zum Abtasten von Blöcken mit 128 oder mehr Tonabtastwerten benutzt, um Blöcke kodierter Information zu erzeugen. Diese Abtastwertblöcke geben typischerweise Toninformationen wieder, die über ein Intervall von 3 bis 12 ms reichen. Jeder Block der mit diesen Kodierv Verfahren kodierten Daten gibt die kleinste Informationseinheit wieder, aus der eine hinreichend genaue Wiedergabe eines Segments der ursprünglichen Toninformation wiederhergestellt werden kann. Mit Teilbandkodierverfahren werden Bitraten dadurch reduziert, daß Frequenzteilbandwiedergaben eines Tonsignals einem Kodieren auf psychoakustischer Basis unterzogen werden. Die Frequenzteilbandwiedergaben können durch Anwenden einer Vielzahl von Bandpaßfiltern oder einer oder mehr Transformationen erzeugt werden. Um die Beschreibung zu vereinfachen, werden diese Teilbandkodierverfahren hier als das Anlegen einer Filterbank zum Erzeugen von Teilbandsignalen beschrieben.

[0008] Die oben erwähnte Ungewißheit hinsichtlich der Synchronisierung hat bei diesen Blockkodieranwendungen große Bedeutung, denn eine zwischen die Grenzen eines kodierten Blocks fallende Editierstelle führt dazu, daß ein Teil dieses Blocks vom restlichen Signal abgetrennt wird. Der teilweise Verlust eines kodierten Blocks manifestiert sich als Verlust einer Dauer von typischerweise 3 ms oder mehr in dem wiederhergestellten Signal. Es ist wahrschein-

lich, daß ein solcher Verlust für das menschliche Hörsystem erkennbar wäre.

[0009] Diese Schwierigkeit läßt sich vermeiden, wenn man ein Nachverarbeitungsverfahren anwendet, bei dem eine PCM-Wiedergabe der ursprünglichen Tonsignale dadurch wiederhergestellt wird, daß auf das kodierte Tonsignal ein Dekodierprozeß angewandt wird, die wiederhergestellte PCM-Wiedergabe nach Bedarf editiert wird und durch Anwenden eines Kodierprozesses auf die editierte PCM-Toninformation eine neue kodierte Wiedergabe erzeugt wird. Wegen der zusätzlichen Kosten und der Verschlechterung der Tonqualität aufgrund der Dekodier/Neukodierverfahren ist diese Lösung uninteressant. Außerdem ist die Nachverarbeitung aus Gründen, die nach dem Lesen der folgenden Beschreibung besser verständlich sind, deshalb wenig reizvoll, weil durch die Dekodier/Neukodierverfahren zusätzliche Verzögerungen in den Audiodatenstrom gelangen.

OFFENBARUNG DER ERFINDUNG

[0010] Es ist eine Aufgabe der vorliegenden Erfindung, das Verarbeiten von eingebetteten Video/Audiodatenströmen so zu gestalten, daß Tätigkeiten wie Editieren und Umschalten möglich sind und dennoch die obigen Schwierigkeiten vermieden werden.

[0011] Gemäß den Lehren eines Aspektes der vorliegenden Erfindung wird bei einem Verfahren ein Toninformation übermittelndes Eingabetonsignal empfangen, und es werden Videorahmenreferenzen empfangen, die Zeitreferenzen für eine Folge von Videorahmen anzeigen, Blöcke kodierter Toninformation aus der Toninformation in einer Form mit reduzierter Bitrate erzeugt, indem das Eingabetonsignal einem Blockkodierverfahren unterzogen wird und die Blöcke kodierter Toninformation zeitkomprimiert werden, und es werden die zeitkomprimierten Blöcke zu einem kodierten Audiodatenstrom zusammengefügt, der eine Vielzahl von Folgen der zeitkomprimierten Blöcken aufweist, so daß ein Anfangsblock in einer jeweiligen Folge von einem Endblock in einer vorhergehenden Folge durch eine Lücke getrennt ist, die mit einem jeweiligen Videorahmenverweis zeitsynchronisiert ist.

[0012] Die verschiedenen Merkmale der vorliegenden Erfindung und die bevorzugten Ausführungsbeispiele derselben sind unter Hinweis auf die folgende Beschreibung und die beigefügten Zeichnungen, in denen gleiche Bezugszeichen gleiche Elemente in den verschiedenen Figuren kennzeichnen, besser verständlich. Der Inhalt der folgenden Beschreibung und der Zeichnungen wird lediglich als Beispiel angeführt und ist nicht als den Umfang der vorliegenden Erfindung einschränkend zu verstehen.

KURZBESCHREIBUNG DER ZEICHNUNGEN

[0013] **Fig. 1** ist ein Funktionsblockschaltbild eines Ausführungsbeispiels eines Systems zum Aufzeich-

nen und zur Leitwegbestimmung mehrerer Video/Audiodatenströme.

[0014] **Fig. 2A bis 2C** sind graphische Darstellungen hypothetischer Tonsignale, die verschiedene Ausrichtungen mit Videorahmenverweisen haben.

[0015] **Fig. 3** ist ein Funktionsblockschaltbild eines Ausführungsbeispiels eines Gerätes zum Verarbeiten von Bildsignalen mit eingebetteter Toninformation.

[0016] **Fig. 4** ist ein Funktionsblockschaltbild eines Ausführungsbeispiels eines Tonsignalprozessors zum Kodieren gemäß verschiedenen Aspekten der vorliegenden Erfindung.

[0017] **Fig. 5A bis 5C** sind graphische Darstellungen hypothetischer Tonsignale, die gemäß der vorliegenden Erfindung verarbeitet wurden und verschiedene Ausrichtungen mit Videorahmenverweisen haben.

[0018] **Fig. 6** ist eine schematische graphische Darstellung überlappender Toninformationsblöcke, die mittels Fensterfunktionen gewichtet sind.

ART UND WEISE DER AUSFÜHRUNG DER ERFINDUNG

Systemüberblick

[0019] **Fig. 1** veranschaulicht ein Ausführungsbeispiel eines Systems zum Aufzeichnen und zur Leitwegbestimmung mehrerer Video/Audiodatenströme und stellt ein Beispiel eines Systems dar, bei dem Aspekte der vorliegenden Erfindung mit Vorteil angewandt werden können. Aus Gründen der Klarheit sind weder in dieser Figur, noch in einer der anderen, Signalwege dargestellt, auf denen Haupttaktsignale zur Gerätesynchronisation geführt werden. Es wird hier davon ausgegangen, daß Signale, beispielsweise die längs Wegen **21**, **22**, **23** und **24** erzeugt werden, den Normen SMPTE 259M und SMPTE 272M entsprechen. Zum Ausführen der vorliegenden Erfindung ist allerdings weder eine bestimmte Norm, noch ein bestimmtes SignalfORMAT wesentlich. Bei einem alternativen Ausführungsbeispiel des Systems werden beispielsweise getrennte Signale längs der Wege **21** bis **24** erzeugt, von denen jedes jeweilige Bildinformation und Toninformation übermittelt, und ein Router **31** umfaßt eine Schaltkreisanordnung, um Bild- und Toninformationen auf getrennte Leitwege zu schicken. In einem derartigen Ausführungsbeispiel ist zwischen einen SDI Einbeter **12** und den Router **31** ein SDI Ausbeter geschaltet. Dieses alternative Ausführungsbeispiel wird hier erwähnt, um zu zeigen, daß zum Ausführen der vorliegenden Erfindung kein bestimmtes SignalfORMAT wesentlich ist.

[0020] Ein Videobandrekorder VTR **16** empfängt von einem Weg **1** Bildinformation und von einem Weg **2** Toninformation und zeichnet diese Bild/Toninformation auf ein Band auf. Anschließend liest der VTR **16** die auf dem Band aufgezeichnete Bild/Toninformation und erzeugt längs des Weges **21** ein Wiedergabe-

signal, welches die Bildinformation mit eingebetteter Toninformation übermittelt. Auf ähnliche Weise zeichnet ein VTR **17** von Wegen **3** und **4** empfangene Bild- bzw. Toninformation auf und erzeugt anschließend längs eines Weges **22** ein Wiedergabesignal, welches die Bildinformation mit eingebetteter Toninformation übermittelt.

[0021] VTR **16**, VTR **17** und ein VTR **18** umfassen Schaltkreisanordnungen, beispielsweise einen SDI Einbutter zum Einfügen von Toninformation in die Bildinformation während der Wiedergabe.

[0022] Der SDI Einbutter **11** empfängt von Wegen **5** und **6** Bild- bzw. Toninformation und erzeugt längs eines Weges **14** ein Signal, welches digitale Bildinformation mit eingebetteter digitaler Toninformation übermittelt. Der VTR **18**, der eine Schaltkreisanordnung wie einen SDI Ausbutter umfaßt, extrahiert aus dem Video/Audiodatensignal Toninformation und zeichnet die getrennte Bild- und Toninformation auf Band auf. Anschließend stellt der VTR **18** aus dem Band die Bild- und Toninformation wieder her und benutzt eine Schaltkreisanordnung, beispielsweise einen SDI Einbutter, um längs eines Weges **23** ein Wiedergabesignal zu erzeugen, welches die Bild- und eingebettete Toninformation übermittelt. Bei Ersatz des VTR **18** durch ein digitales Datenaufzeichnungsgerät ist jedoch im Aufzeichnungsgerät weder eine Einbettungs- noch eine Ausbettungsschaltung erforderlich, weil der Video/Audiodatenstrom selbst aufgezeichnet und wiedergegeben werden kann.

[0023] Der SDI Einbutter **12** empfängt von Wegen **7** und **8** Bild- bzw. Toninformation und erzeugt längs eines Weges **24** ein Signal, welches digitale Bildinformation mit eingebetteter digitaler Toninformation übermittelt.

[0024] Der SDI Router **31** empfängt Bild/Tonsignale von den Wegen **21**, **22**, **23** und **24** und lenkt oder schaltet diese Signale wahlweise längs eines Weges **34** zu einem Wiedergabe/Aufzeichnungsgerät **41**. Die Anzahl der vom SDI Router **31** empfangenen Signale hat keine nennenswerte Bedeutung. Das Wiedergabe/Aufzeichnungsgerät **41** stellt ein willkürliches Gerät dar, welches das längs des Weges **34** gesandte Signal benutzt. Es kann sich beispielsweise um ein Aufzeichnungsgerät, wie einen VTR oder ein Wiedergabegerät, wie ein Fernsehgerät handeln. Ferner kann das Wiedergabe/Aufzeichnungsgerät **41** vom SDI Einbutter **31** entfernt angeordnet sein, und in diesem Fall stellt der Weg **34** einen Nachrichtenübermittlungs- oder Rundfunksendekanal dar.

Verschiebungen in der Bild/Tonsynchronisation

[0025] Die relative Ausrichtung zwischen Bildinformation und Toninformation kann durch Schaltkreisverzögerungen in den VTR **16**, **17**, **18** und SDI Einbuttern **11** und **12** Änderungen erfahren. Demzufolge kann beispielsweise die Synchronisation von Bild/Toninformation im Wiedergabesignal **21** gegenüber der Synchronisation zwischen Videoinformation

und Toninformation, wie sie von den Wegen **1** bzw. **2** empfangen wird, verschoben werden. Der Grad der Änderung in der Ausrichtung unterscheidet sich unter Geräten unterschiedlicher Hersteller, kann sich unter verschiedenen Geräteteilen des gleichen Herstellers unterscheiden und kann sogar innerhalb eines gegebenen Geräteteils, zum Beispiel als Funktion des Initialisierungszustands von Puffern, unterschiedlich sein.

[0026] In Fig. 2A stellt ein Signal **111** Toninformation dar, die eine bestimmte Ausrichtung mit Videorahmenverweisen **101** und **102** hat. Jeder dieser Videorahmenreferenzen zeigt einen bestimmten Bezugspunkt in einem jeweiligen Videorahmen an. Ein üblicher Bezugspunkt für NTSC Bildinformation fällt beispielsweise mit der Bildinformation für Zeile **10** in jedem Rahmen zusammen. Ein üblicher Bezugspunkt für PAL Bildinformation fällt mit Zeile **1** in jedem Rahmen zusammen. Für die Durchführung der vorliegenden Erfindung hat keine bestimmte Ausrichtung kritische Bedeutung.

[0027] In Fig. 2B stellt ein Signal **121** die gleiche Toninformation dar, wie die vom Signal **111** übermittelte, die allerdings gegenüber dem Signal **111** verzögert ist. Infolgedessen ist die Synchronisation zwischen dem Signal **121** und den Videorahmenreferenzen im Verhältnis zur Synchronisation für das Signal **111** verschoben. In Fig. 2C stellt ein Signal **131** die gleiche Toninformation dar wie die vom Signal **111** übermittelte, ist aber gegenüber dem Signal **111** vorausgeeilt, und folglich ist die Synchronisation zwischen dem Signal **131** und den Videorahmenreferenzen verschoben, aber entgegengesetzt zur Verschiebung der Synchronisation beim Signal **121**.

[0028] Unter Hinweis auf Fig. 1 sei angenommen, daß die in Fig. 2A dargestellte Toninformation und Synchronisation über die Wege $1/2$, $3/4$, $5/6$ und $7/8$ übermittelt wird. Es ist wahrscheinlich, daß in den längs der Wege **21** bis **24** erzeugten Signalen unterschiedliche Verschiebungen in der Synchronisation, wie den in Fig. 2A bis 2C gezeigten, bestehen. Ferner sei angenommen, daß die in Fig. 2A bis 2C dargestellten Synchronisationen in den jeweils längs der Wege **21** bis **23** erzeugten Signalen bestehen. Wenn der SDI Router **31** zwischen den aus diesen drei Wegen empfangenen Signalen umschaltet, tritt eine kleine Diskontinuität in der Toninformation ein, die in das längs des Weges **34** weitergeleitete Signal eingebettet ist. Wenn die Toninformation in einer linearen Form, beispielsweise PCM, wiedergegeben ist, kann ein menschlicher Zuhörer diese Diskontinuität wahrscheinlich nicht bemerken, weil die Diskontinuität nur über wenige Abtastwerte besteht. Besonders schwierig wäre es, eine Diskontinuität zwischen zwei Signalen zu bemerken, die einen unterschiedlichen Toninhalt haben.

Auswirkungen des Kodierens

[0029] Wie oben schon gesagt, besteht allerdings

zunehmendes Interesse daran, Tonkanäle in größerer Zahl in einen Video/Audiodatenstrom einzubetten. Wenn die Informationskapazität in dieser größeren Anzahl von Tonkanälen die für Toninformation verfügbare räumliche Kapazität überschreitet, wird irgendeine Art Komprimierung der Bandbreite oder Bitrate angewandt. Ein Beispiel einer solchen Komprimierung ist das Tonkodieren auf der Grundlage psychoakustischer Grundsätze.

[0030] Diese Kodiertechniken werden oft auf Blöcke von Tonabstastwerten angewandt, um Blöcke kodierter Information zu erzeugen. Diese Abstastwertblöcke stellen typischerweise Toninformation dar, die über ein Intervall von 3 bis 12 ms reicht. Jeder der mittels dieser Kodierverfahren erzeugten Blöcke kodierter Information stellt die kleinste Informationseinheit dar, aus der eine hinreichend genaue Wiedergabe eines Segments der ursprünglichen Toninformation wiederhergestellt werden kann.

[0031] Eine Folge kodierter Informationsblöcke **112** wird als Impulszug in **Fig. 2A** dargestellt. Die von diesen Blöcken übermittelte Information ist eine kodierte Wiedergabe der Toninformation im Signal **111**. Die Gestalt und Größe dieser Impulse ist nicht entscheidend. Der Impulszug soll lediglich eine Folge von Blöcken angeben, welche kodierte Information entsprechend Blöcken von Tonabstastwerten übermitteln, die aneinander angrenzen oder vorzugsweise einander überlappen können. Bei dem in **Fig. 2A** gezeigten Beispiel ist die das Intervall zwischen einander benachbarten Videorahmenreferenzen überbrückende Toninformation von sechs Blöcken kodierter Information wiedergegeben. Eine Offenbarung verschiedener Überlegungen zur Qualitätsverbesserung der Tonkodierung bei Bild/Tonanwendungen findet sich in WO-A 99/21187.

[0032] Wenn im System gemäß **Fig. 1** Blockkodiertechniken angewandt werden, enthalten die Signale, die der SDI Router **31** von den Wegen **21** bis **24** empfängt, in Blöcken kodierte Toninformation. Wie oben gesagt, können zwischen den kodierten Informationsblöcken und den Videorahmenreferenzen Verschiebungen in der Ausrichtung auftreten. Dies ist in den **Fig. 2A, 2B** und **2C** durch unterschiedliche Synchronisation zwischen der Videorahmenreferenz **101** und beispielsweise den Blöcken **112, 122** bzw. **132** veranschaulicht.

[0033] Wie oben gesagt, sei beispielsweise angenommen, daß die in **Fig. 2A** bis **2C** gezeigten Synchronisationen in den jeweils längs der Wege **21** bis **23** erzeugten Signalen bestehen. Wenn der SDI Router **31** an der Videorahmenreferenz **101** von dem über den Weg **22** empfangenen Signal, in **Fig. 2B** gezeigt, auf das über den Weg **23** empfangene Signal, in **Fig. 2C** gezeigt, umschaltet, kann eine signifikante Menge der Toninformation am Umschaltpunkt aus dem längs des Weges **23** weitergegebenen Signal nicht wiederhergestellt werden. Die im Block **123** vor dem Umschaltpunkt übermittelte Toninformation kann nicht wiedergewonnen werden, weil einerseits

der gesamte Block zur Wiederherstellung der Toninformation nötig ist und andererseits der Teil des Blocks nach dem Umschaltpunkt fehlt. Ähnlich kann die im Block **133** nach dem Umschaltpunkt übermittelte Toninformation nicht wiedergewonnen werden, weil der Teil des Blocks **133** vor dem Umschaltpunkt fehlt.

[0034] Hierbei handelt es sich nicht um ein einmaliges Problem für die Art von System, die in **Fig. 1** veranschaulicht ist. Das Problem tritt beispielsweise auch beim Bandeditieren oder Tonübeispielen für einen einzigen VTR auf.

[0035] Wie nachfolgend näher erläutert, wird dieses Problem mit der vorliegenden Erfindung dadurch überwunden, daß in dem kodierten Audiodatenstrom Schutzbänder oder Lücken gebildet werden, so daß eine beträchtliche Schwankungsbreite der Bild/Ton-synchronisation ohne Verlust an Toninformation hinnehmbar ist.

Kodier-Signalprozessor

[0036] **Fig. 3** zeigt einen Bild/Tonsignalprozessor, der auf verschiedenerelei Weise in ein System, wie das in **Fig. 1** dargestellte, eingebaut werden kann. Bei dem gezeigten Ausführungsbeispiel wird eine Vielzahl von Bildinformation mit eingebetteter Toninformation übermittelnden Signalen von Eingabesignalwegen **61-1, 61-2** und **61-3** empfangen. In der Figur sind drei Eingabesignalwege gezeigt; aber Ausführungsbeispiele der vorliegenden Erfindung können Signalwege für im wesentlichen jede beliebige Anzahl von Eingabesignalen haben. Ein Signalverteiler **62** steht für einen großen Bereich an Signalverteilungsprozessen, einschließlich Umschalten, Zusammenlegen, Editieren, Spleißen sowie Speichern/Wiedergewinnen. Aus Gründen der Einfachheit wird bei der Darstellung und Beschreibung hier davon ausgegangen, daß der Signalverteiler **62** eine Vielzahl von Bild/Tonsignalen empfängt und diese Signale auf die eine oder andere Weise verarbeitet und/oder verteilt, um längs eines Weges **63** ein einziges Signal zu erzeugen, welches Bildinformation mit eingebetteter Toninformation übermittelt. Ein Entformatierer **64** empfängt die Bild/Toninformation vom Weg **63**, extrahiert die eingebettete Toninformation und leitet sie längs eines Weges **65** weiter. Die Bildinformation kann längs eines Weges **69** weitergegeben werden. Ein Tonsignalprozessor empfängt die Toninformation vom Weg **65** und unterzieht die Toninformation einem Blockkodierverfahren, um längs eines Weges **67** Blöcke kodierter Information zu erzeugen. Ein Formatierer **68** empfängt die Blöcke kodierter Information vom Weg **67** und erzeugt längs eines Weges **70** ein Ausgabesignal, welches eine Vielzahl von Blockfolgen kodierter Information mit Lücken oder Schutzbändern zwischen einem Anfangsblock in einer Folge und einem Endblock in einer vorhergehenden Folge aufweist. Mit Hilfe eines Bezugssignals, beispielsweise eines Haupttaktsignals, wird die Lücke oder das

Schutzband mit der Bildinformation zeitlich synchronisiert.

[0037] Wie vorstehend erwähnt, zeigen die Figuren keine Signalwege, die Haupttaktsignale für die Gerätesynchronisation führen. Bei einem bevorzugten Ausführungsbeispiel bildet ein Tonsignalprozessor **66** Tonsignalblöcke, die mit dem Haupttaktsignal synchronisiert sind. Diese Synchronisierung ist in **Fig. 2A** dargestellt, wo Grenzen zwischen einander benachbarten Abtastwertblöcken mit den Videorahmenreferenzen **101** und **102** zusammenfallen. Es können aber auch andere Arten der Ausrichtung benutzt werden.

[0038] Wie aus **Fig. 5A** hervorgeht, übermittelt eine Blockfolge **112-2** ein kodierte Information wiedergebendes Signalsegment **111-2**, bei dem es sich um eine hypothetische, zeitkomprimierte Wiedergabe des zwischen den Videorahmenreferenzen **101** und **102** liegenden Teils des Signals **111** handelt. Ähnlich übermittelt eine Blockfolge **112-1** kodierte Information, die ein Signalsegment **111-1** wiedergibt, und eine Blockfolge **112-3** übermittelt kodierte Information, die ein Signalsegment **111-3** wiedergibt. Vom Tonsignalprozessor **66** und Formatierer **68** werden Blocksequenzen erzeugt, die eine kodierte Wiedergabe der Toninformation übermitteln, in der beispielsweise ein Schutzband oder eine Lücke zwischen dem Endblock in der Folge **112-1** und dem Anfangsblock **112-2** gebildet wurde.

[0039] Die in **Fig. 2A** bis **2C** gezeigten Verschiebungen in der Ausrichtung sind auch in den **Fig. 5A** bis **5C** dargestellt. In diesen Figuren wird von der kodierten Information in den Folgen **122-1**, **122-2**, **122-3**, **132-1**, **132-2** und **132-3** kodierte Information übermittelt, die Signalsegmente **121-1**, **121-2**, **121-3**, **131-1**, **131-2** bzw. **131-3** wiedergibt. Wie aus den **Fig. 5B** und **5C** zu entnehmen ist, kommt es nicht zu einem Verlust an Toninformation infolge von Verschiebungen der Ausrichtung, weil die Potentialumschaltunkte an den Videorahmenreferenzen **101** und **102** innerhalb eines Schutzbandes liegen.

[0040] Der in **Fig. 3** gezeigte Signalprozessor kann zum Beispiel in einen SDI Router eingebaut werden, um Videosignale zu verarbeiten, die eingebettete AES3 oder PCM Toninformation enthalten. Ein Ausführungsbeispiel unter Weglassung eines Signalverteilers **62** kann in einen VTR oder SDI Einbatter eingebaut werden. Ein weiteres Ausführungsbeispiel, bei dem auch der Entformatierer **64** weggelassen ist, kann in einen VTR oder in die Eingangsschaltkreise eines SDI Einbatters eingebaut werden.

[0041] **Fig. 4** zeigt ein zum Einbau in das in **Fig. 3** gezeigte Ausführungsbeispiel geeignetes Ausführungsbeispiel eines Kodier-Tonsignalprozessors, der auch noch eine getrennte weitere Verwendung hat. Gemäß diesem Ausführungsbeispiel weist der Tonsignalprozessor **66** eine Vielzahl von Filterbänken **71**, **72** und **73** auf. Die Filterbank **71** erzeugt eine Vielzahl von Frequenzteilbandsignalen längs Wegen **75-1** bis **75-3** in Abhängigkeit von dem vom Weg **65-1** emp-

fangenen Signal. Die Filterbank **72** erzeugt eine Vielzahl von Frequenzteilbandsignalen längs Wegen **76-1** bis **76-3** in Abhängigkeit von dem vom Weg **65-2** empfangenen Signal. Die Filterbank **73** erzeugt eine Vielzahl von Frequenzteilbandsignalen längs Wegen **77-1** bis **77-3** in Abhängigkeit von dem vom Weg **65-3** empfangenen Signal. Die Filterbänke **71**, **72** und **73** können auf verschiedene Art und Weise, einschließlich einer Bandpaßfilterbank, eines Satzes von Bandaufteilungsfiltern in Kaskadenanordnung und einer oder mehr Transformationen von der Zeitdomäne in die Frequenzdomäne verwirklicht werden. Es sind nur drei Filterbänke gezeigt und für jede Filterbank nur drei Teilbandsignale; aber ein Ausführungsbeispiel kann viel mehr Filterbänke umfassen, die je vierundzwanzig oder mehr Teilbandsignale erzeugen, welche je Frequenzteilbänder darstellen, deren Bandbreiten den kritischen Bandbreiten des menschlichen Hörsystems entsprechen oder kleiner sind als diese. In einem Kodierer **79** werden die Teilbandsignale einem Blockkodierverfahren unterzogen, und längs eines Weges **67** wird eine Blockfolge erzeugt, die in einer kodierten Form die über die Wege **65-1**, **65-2** und **65-3** empfangene Toninformation wiedergibt.

[0042] Das Teilbandkodieren ist nicht wesentlich zum Ausführen der vorliegenden Erfindung, sondern es können auch andere Formen der Kodierung, beispielsweise blockkomprimierte PCM oder Delta-modulation angewandt werden.

[0043] Bei einem Ausführungsbeispiel in der Praxis empfängt ein Kodier-Tonsignalprozessor acht Kanäle Toninformation in linearer PCM-Form oder, gemäß einer Alternative, vier AES3 Datenströme und arbeitet mit acht Filterbänken und einem Kodierer, der ein Blockkodierverfahren anwendet, um Blöcke kodierter Information mit Schutzbändern zu erzeugen, die in einem Raum oder einer Bandbreite übermittelt werden können, welche nicht größer ist als die zum Übermitteln von zwei Kanälen von Toninformation in linearer PCM-Form oder, alternativ, eines einzigen AES3 Datenstroms benötigte.

Überlappende Blöcke und Fensterfunktionen

[0044] Der in den Figuren zur Darstellung von Informationsblöcken verwendete Impulszug läßt darauf schließen, daß einander benachbarte Blöcke zwar aneinandergrenzen, sich aber nicht überlappen. Es ist zwar keine bestimmte Anordnung von Blöcken für die Ausübung der vorliegenden Erfindung von kritischer Bedeutung, aber bevorzugte Ausführungsbeispiele verarbeiten Blöcke, die einander überlappen. Insgesamt werden überlappende Blöcke von Toninformation mittels einer Fensterfunktion gewichtet oder moduliert, so daß die Summe überlappender Abtastwerte in einander benachbarten Blöcken im wesentlichen eine Konstante ist.

[0045] **Fig. 6** veranschaulicht eine Blockfolge. Ein Anfangsblock **141** in der Folge überlappt den be-

nachbarten Block **142**. Alle Blöcke in der Folge sind durch eine Hüllkurve dargestellt, welche die Gestalt einer Fensterfunktion hat, die zum Gewichten entsprechender Toninformation in der Zeitdomäne benutzt wird. Ein Endblock **146** in der Folge überlappt den vorhergehenden Block sowie einen in der Figur nicht gezeigten, nachfolgenden Block. Das Ausmaß der Überlappung und die Wahl der Fensterfunktion kann signifikante Auswirkungen auf die Kodierleistung haben; aber für die vorliegende Erfindung ist keine bestimmte Fensterfunktion und kein bestimmtes Ausmaß der Überlappung von kritischer Bedeutung. Bei bevorzugten Ausführungsbeispielen beträgt das Ausmaß der Überlappung einer Hälfte der Blocklänge, und die Fensterfunktion ist von der Kaiser-Bessel-Funktion abgeleitet.

[0046] Wie oben erwähnt, erzeugt ein Tonsignalprozessor **86** Toninformation, die mit Videorahmenreferenzen synchronisiert ist. In Ausführungsbeispielen, mit denen Blockfolgen von Toninformation erzeugt werden, kann die Ausrichtung so gewählt sein, daß eine Videorahmenreferenz mit im wesentlichen jedem beliebigen Punkt in einem Block der Folge zusammenfällt. Bei dem in **Fig. 6** gezeigten Beispiel fällt der Anfang des Anfangsblocks **141** mit der Videorahmenreferenz **100** zusammen.

[0047] In manchen Anwendungsfällen kann sich der exakte Punkt des Zusammenfallens von Videorahmen zu Videorahmen unterscheiden. In Anwendungsfällen, in denen beispielsweise digitale Toninformation mit NTSC Bildinformation kombiniert wird, können aufeinanderfolgende Videorahmen unterschiedliche Anzahlen von Tonabtastraten enthalten, weil die Tonabtastrate kein ganzes Vielfaches der Videorahmenrate ist.

[0048] Verschiedene Überlegungen zur Blocklänge, Fensterfunktion sowie Bild/Tonsynchronisation sind in der oben schon erwähnten WO-A 99/21187 erörtert.

Patentansprüche

1. Verfahren zum Verarbeiten von Toninformation, aufweisend:
Empfangen eines Eingabetonsignals, welches die Toninformation übermittelt,
Empfangen von Videorahmenreferenzen, welche Zeitreferenzen für eine Folge von Videorahmen anzeigen,
Erzeugen von Blöcken kodierter Toninformation aus der Toninformation in einer Form mit reduzierter Bitrate durch Anwenden eines Blockkodierverfahrens auf das Eingabetonsignal und Zeitkomprimieren der Blöcke kodierter Toninformation, und
Zusammenfügen der zeitkomprimierten Blöcke zu einem kodierten Audiodatenstrom, der eine Vielzahl von Folgen der zeitkomprimierten Blöcke aufweist, so daß ein Anfangsblock in einer jeweiligen Folge von einem Endblock in einer vorhergehenden Folge durch eine Lücke getrennt ist, die mit einer jeweiligen

Videorahmenreferenz zeitlich synchronisiert ist.

2. Verfahren nach Anspruch 1, bei dem das Blockkodierverfahren die Toninformation zu einer Form kodiert, die weniger Redundanz und/oder weniger perzeptuelle Irrelevanz enthält.

3. Verfahren nach Anspruch 2, bei dem das Blockkodierverfahren aufweist: entweder eine Bandpaßfilterbank oder eine oder mehrere Transformationen auf das Eingabetonsignal anzuwenden, um eine Vielzahl von Frequenzteilbanddarstellungen des Eingabetonsignals zu erzeugen, und die Blöcke kodierter Toninformation durch adaptive Zuteilung von Bits zu den Frequenzteilbanddarstellungen entsprechend psychoakustischen Grundsätzen zu erzeugen.

4. Verfahren nach einem der Ansprüche 1 bis 3, bei dem die Videorahmen in einer Vielzahl von Bildzeilen angeordnete Bildinformation aufweisen und jede der Lücken eine Dauer hat, die länger ist als die von drei Bildzeilen.

5. Verfahren nach einem der Ansprüche 1 bis 3, bei dem jeder der Videorahmen eine Dauer hat, die einem Rahmenintervall gleich ist und die Dauer der Lücken innerhalb eines Bereichs von etwa einem Prozent des Rahmenintervalls bis zu etwa zwanzig Prozent des Rahmenintervalls liegt.

6. Verfahren nach einem der Ansprüche 1 bis 5, bei dem das Blockkodierverfahren auf überlappende Blöcke von Toninformation angewandt wird.

7. Vorrichtung zum Verarbeiten von Toninformation, aufweisend:
eine Einrichtung zum Empfangen eines Eingabetonsignals, welches die Toninformation übermittelt,
eine Einrichtung zum Empfangen von Videorahmenreferenzen, welche Zeitreferenzen für eine Folge von Videorahmen anzeigen,
eine Einrichtung zum Erzeugen von Blöcken kodierter Toninformation aus der Toninformation in einer Form mit reduzierter Bitrate durch Anwenden eines Blockkodierverfahrens auf das Eingabetonsignal und Zeitkomprimieren der Blöcke kodierter Toninformation, und
eine Einrichtung zum Zusammenfügen der zeitkomprimierten Blöcke zu einem kodierten Audiodatenstrom, der eine Vielzahl von Folgen der zeitkomprimierten Blöcke aufweist, so daß ein Anfangsblock in einer jeweiligen Folge von einem Endblock in einer vorhergehenden Folge durch eine Lücke getrennt ist, die mit einer jeweiligen Videorahmenreferenz zeitlich synchronisiert ist.

8. Vorrichtung nach Anspruch 7, bei der im Blockkodierverfahren die Toninformation zu einer Form kodiert wird, die weniger Redundanz und/oder weniger perzeptuelle Irrelevanz enthält.

9. Vorrichtung nach Anspruch 8, bei der das Blockkodierverfahren aufweist: eine Einrichtung, mit der entweder eine Bandpaßfilterbank oder eine oder mehrere Transformationen auf das Eingabetonsignal angewandt werden, um eine Vielzahl von Frequenzteilbanddarstellungen des Eingabetonsignals zu erzeugen, und eine Einrichtung, mit der die Blöcke kodierter Toninformation durch adaptive Zuteilung von Bits zu den Frequenzteilbanddarstellungen entsprechend psychoakustischen Grundsätzen erzeugt werden.

10. Vorrichtung nach einem der Ansprüche 7 bis 9, bei der die Videorahmen in einer Vielzahl von Bildzeilen angeordnete Bildinformation aufweisen und jede der Lücken eine Dauer hat, die länger ist als die von drei Bildzeilen.

11. Vorrichtung nach einem der Ansprüche 7 bis 9, bei der jeder der Videorahmen eine Dauer gleich einem Rahmenintervall hat und die Dauer der Lücken innerhalb eines Bereichs von etwa einem Prozent des Rahmenintervalls bis zu etwa zwanzig Prozent des Rahmenintervalls liegt.

12. Vorrichtung nach einem der Ansprüche 7 bis 11, bei der die Einrichtung, mit der die Blöcke kodierter Toninformation erzeugt werden, das Blockkodierverfahren auf überlappende Blöcke der Toninformation anwendet.

Es folgen 5 Blatt Zeichnungen

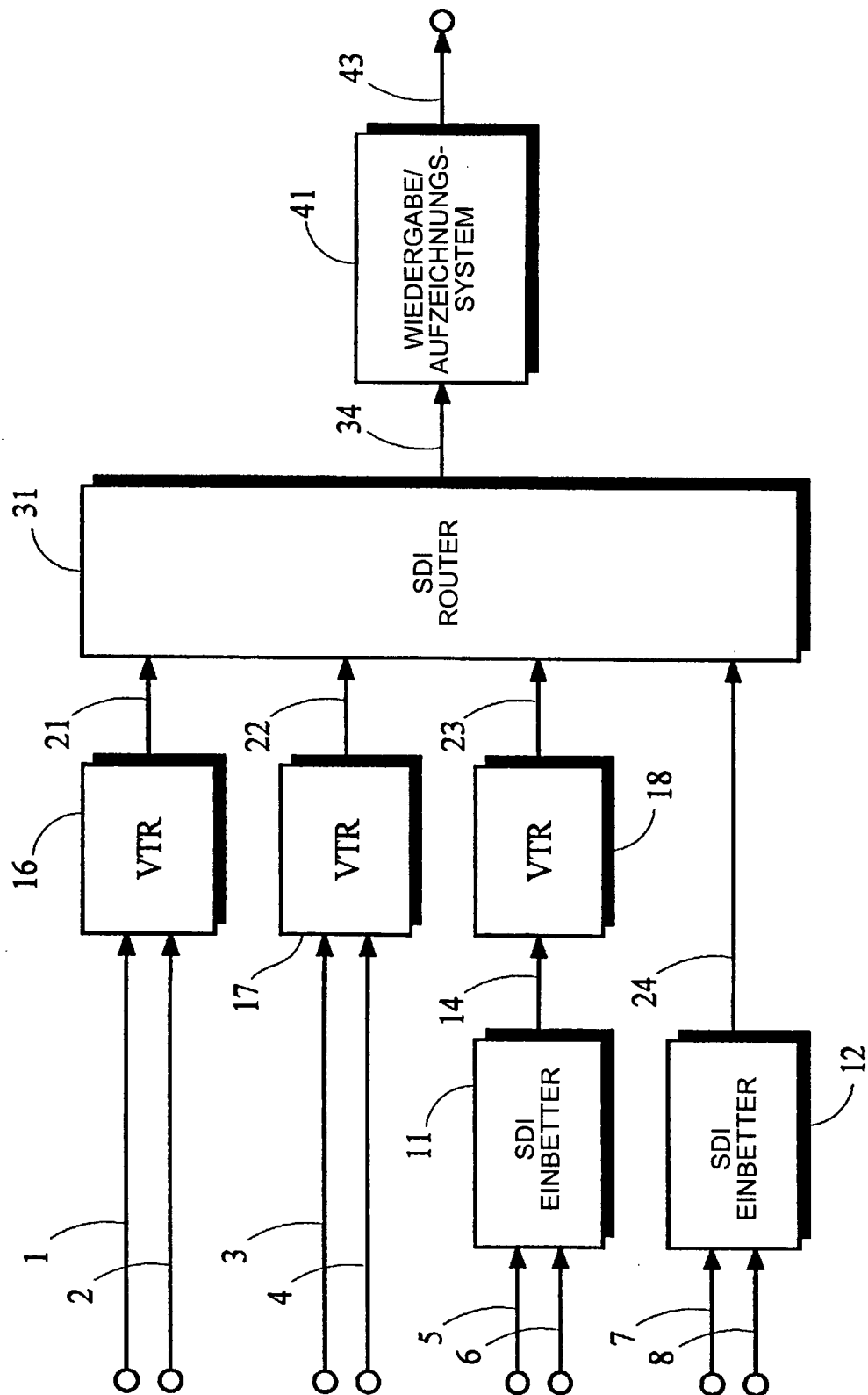


Fig. 1

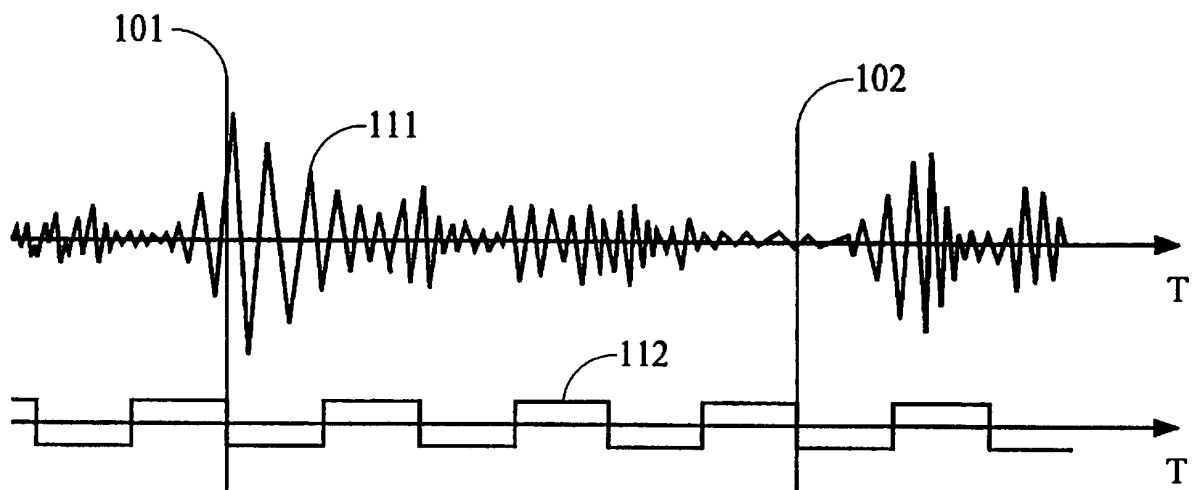


Fig. 2A

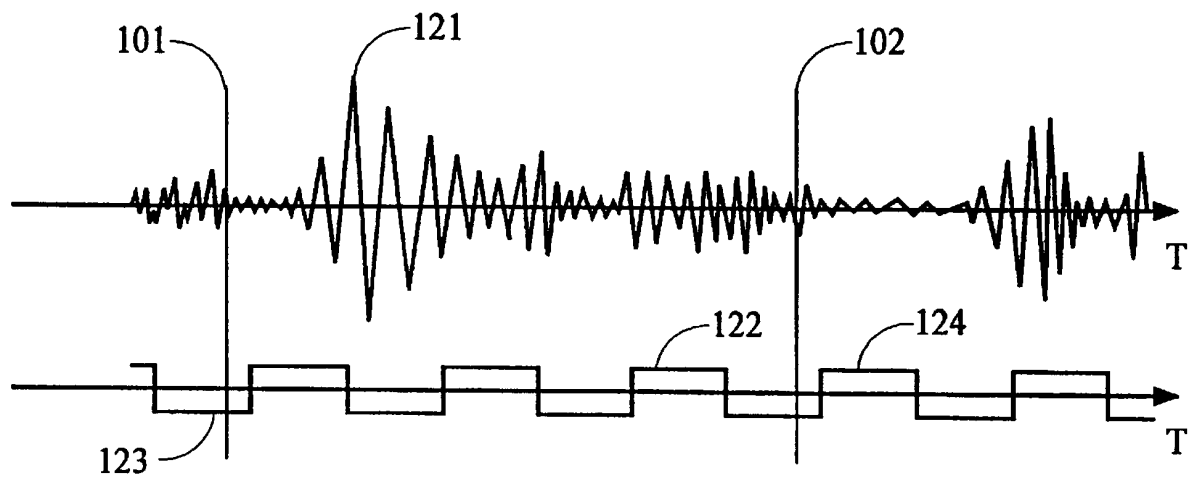


Fig. 2B

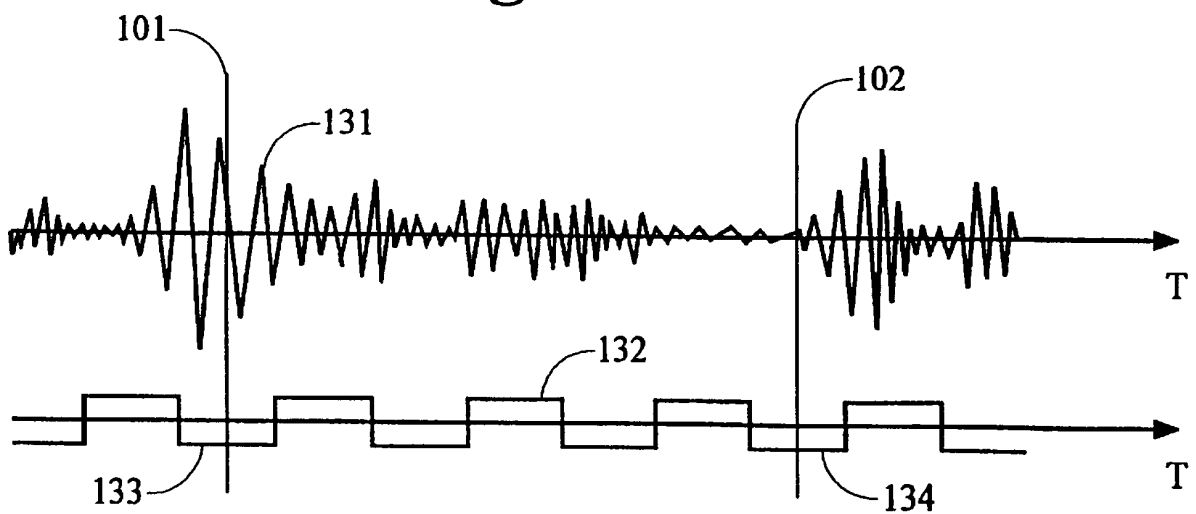


Fig. 2C

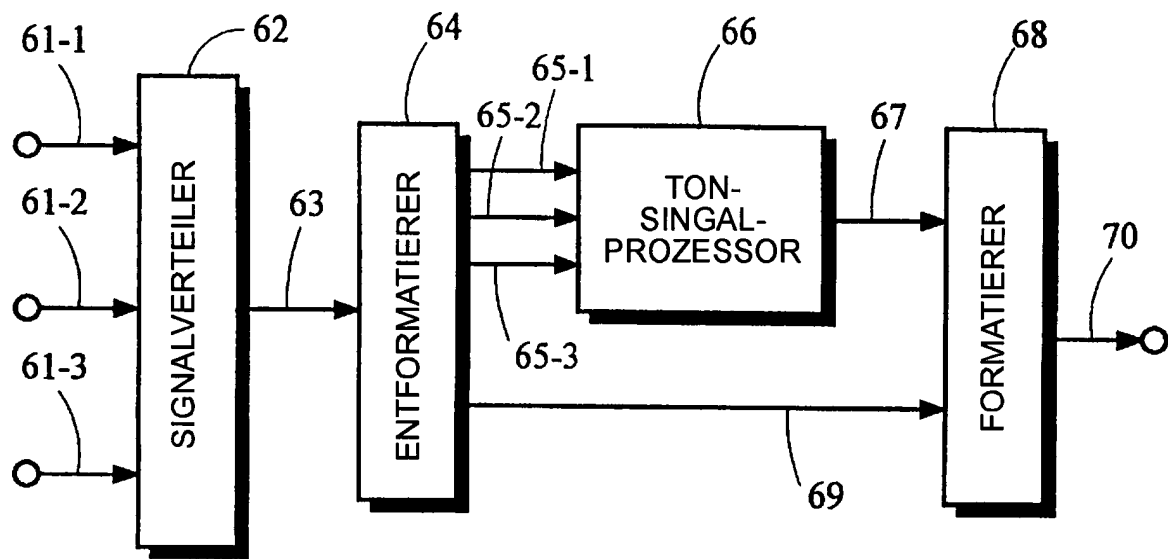


Fig. 3

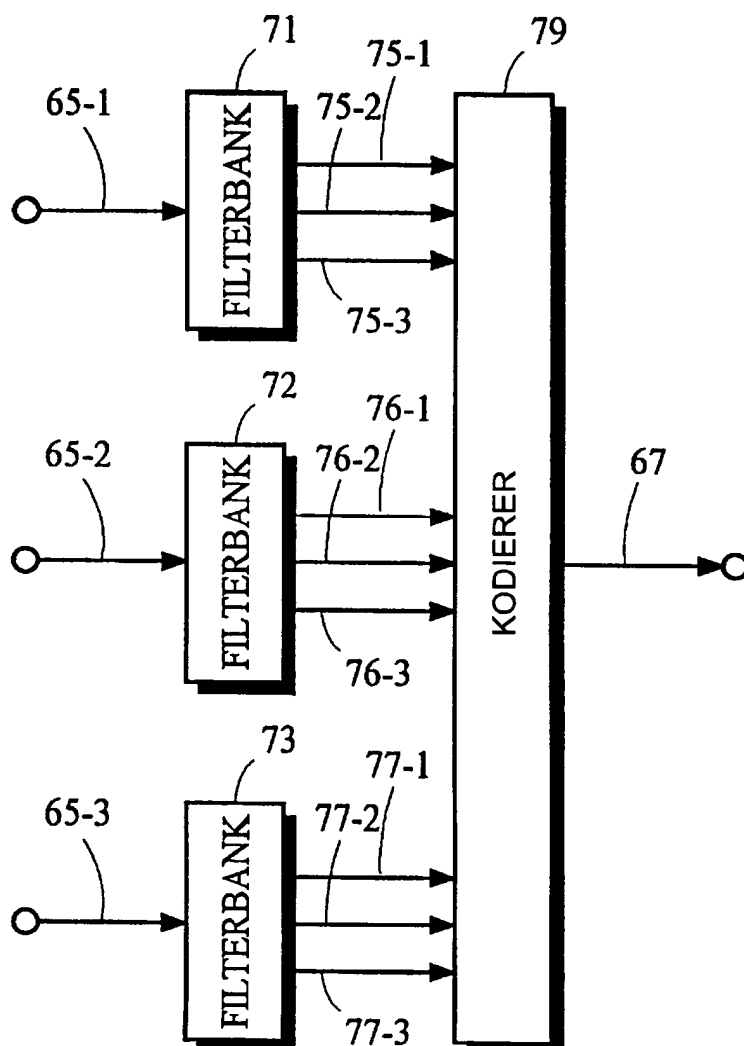


Fig. 4

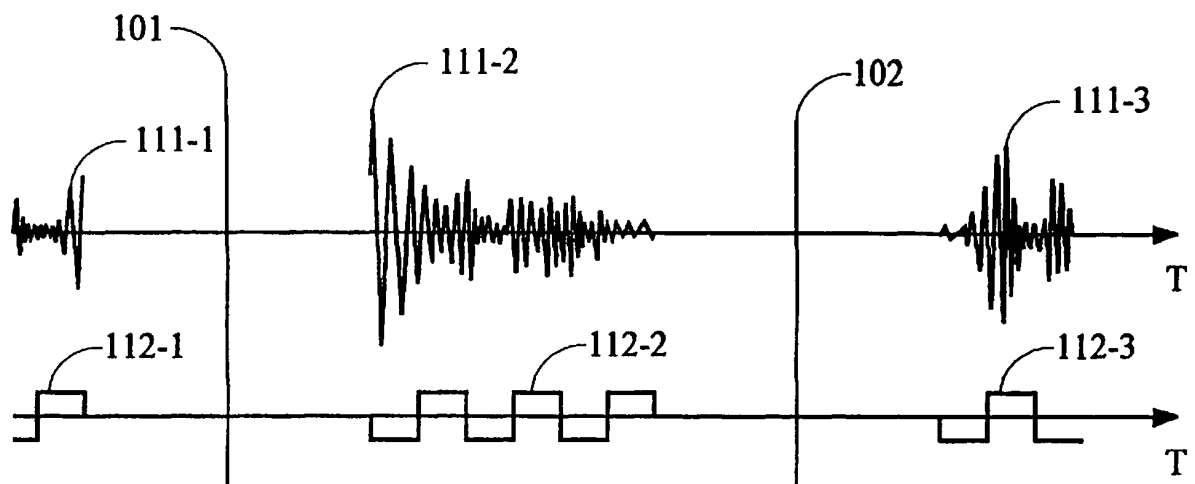


Fig. 5A

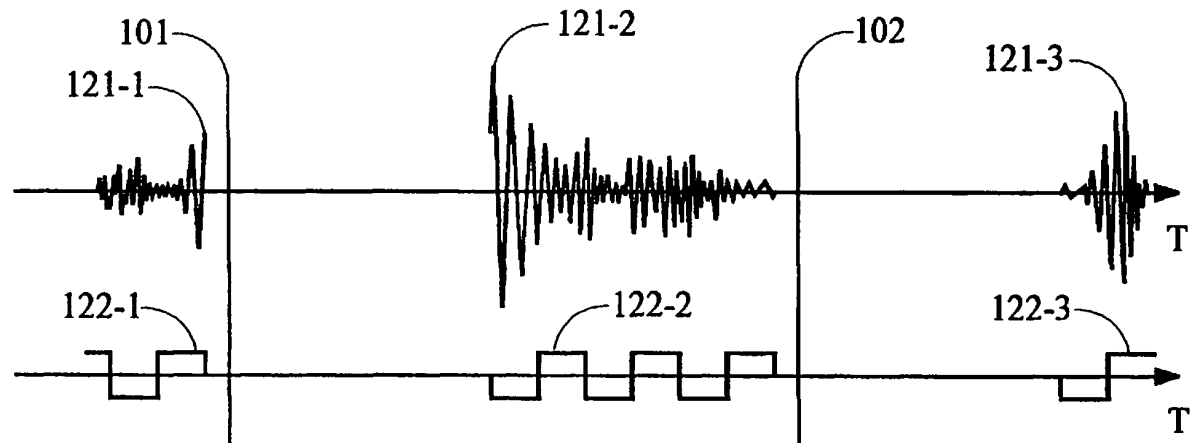


Fig. 5B

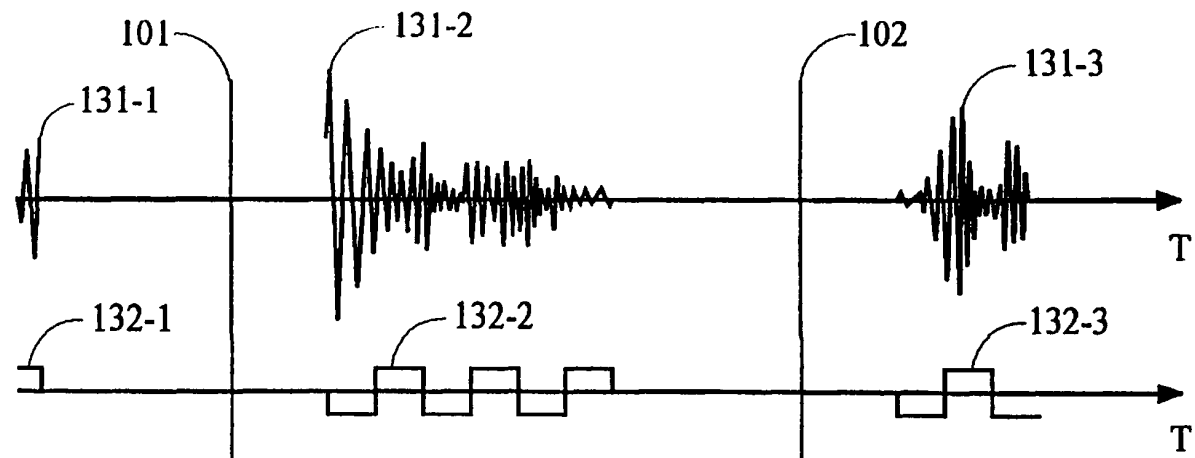


Fig. 5C

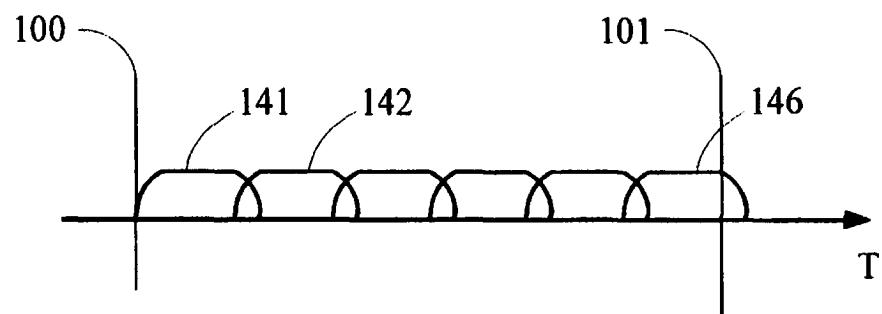


Fig. 6