

【公報種別】特許法第17条の2の規定による補正の掲載

【部門区分】第6部門第2区分

【発行日】平成28年8月4日(2016.8.4)

【公表番号】特表2015-519615(P2015-519615A)

【公表日】平成27年7月9日(2015.7.9)

【年通号数】公開・登録公報2015-044

【出願番号】特願2015-516683(P2015-516683)

【国際特許分類】

G 10 L 19/02 (2013.01)

【F I】

G 10 L 19/02 150

【手続補正書】

【提出日】平成28年6月10日(2016.6.10)

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【発明の詳細な説明】

【発明の名称】二重の互換性を持つ損失のないオーディオ帯域幅拡張

【技術分野】

【0001】

本発明は、デジタルオーディオ信号に関し、特に標準PCM再生との互換性を持つ損失のない帯域幅拡張スキームに関する。

【背景技術】

【0002】

多くのオーディオ愛好家やミュージシャンが、現行媒体の44.1kHz又は48kHzよりもかなり高い平均周波数でサンプリングされ、かつ16ビットよりも多いビット数の分解能で量子化されたオーディオであると通常は理解される、「高分解能」デジタルオーディオを要望している。

【0003】

損失のある圧縮がなされたオーディオは、需要者市場でありふれたものであるが、経験は、「透明」であると主張されるシステムの経験でさえも、損失のある圧縮がなされたオーディオに対する疑問へと、多くの人々を導いてきた。1つの例外は、一定のビット深度への、単純な、非適応的な、ノイズシェーピングされ、ディザリングされた、再量子化である。適切な注意を払えば、これは、(入力と出力との間の差違の1次及び2次の統計に従い、)一定のノイズを付加することに等しい(J. Vanderkooy and S. P. Lipshitz, "Digital Dither: Signal Processing with Resolution Far below the Least Significant Bit" in Proc. AES 7th Int. Conf. on Audio in Digital Times (Toronto, Ont., Canada, 1989), pp. 87-96. 参照)。そのようなノイズは、アナログ媒体及びデジタル媒体の双方における数十年にわたる経験の結果、「良性」であると考えられている。

【0004】

2つの音楽配布媒体が大量消費市場で優勢である。1つはコンパクトディスク(CD)であり、これは、44.1kHzのサンプリング周波数と、16ビットのビット深度を持つ。他の1つはインターネットダウンロードであり、これは、一般にコンピュータ又はパーソナルプレーヤを通じて聞かれる。大抵のダウンロードは損失のある圧縮がなされているが、コンピュータ又はプレーヤは、ほぼ必ず、44.1kHz及び48kHzのサンプリング周波数で非圧縮のPCM(Pulse Code Modulation)信号を扱うことができる。

多くは 24 ビットのビット深度を扱うことができるが、いくつかのパーソナルプレーヤは 16 ビットに制限されている。

【0005】

オーディオ愛好家バージョン（一般に 96 kHz のサンプリング周波数を持つ）と、大量消費市場プレーヤで再生可能なフォーマットとの双方でオーディオ録音を発行することは、商業的には魅力がない。標準大量消費市場プレーヤで再生可能な録音でありながら、特別なデコーダが追加の帯域幅を再生することを可能にする隠れた情報をも含む録音を発行する可能性が、過去に何度も探られてきた（MITSUYA KOMAMURA “Wide-Band and Wide-Dynamic-Range Recording and Reproduction of Digital Audio” J.Audio Eng. soc. Vol.43, No.1/2, 1995 January/February を含む）。しかしながら、今までのところ、オリジナルの高サンプリングレート信号の、損失のない再生への要求に注意を向けつつも、いずれも標準 PCM 再生互換性を提供していない。また、2つの異なるビット深度（例えば 16 ビットプレーヤと 24 ビットプレーヤとの双方）で、デコーダがリスナに如何にして最適な経験を提供するかという問題を、いずれも考えてこなかった。

【0006】

国際公開第 2007 / 128662 号には、エンコードされた損失のあるデータストリームと、損失のない拡張データストリームとを用いた、ソース信号の損失のないエンコードのための方法及び装置が開示されている。他の文献（M van der Veen et al: “High Capacity Reversible Watermarking for Audio”, Security and Watermarking of Multimedia Contents V, Proceedings of SPIE-IS&T Electronic Imaging, SPIE vol. 5020, 1 January 2003, pages 1-11）には、入力信号のダイナミックレンジが制限され、かつウォーターマークビットのエンコードに不使用部分のビットが用いられる、デジタルオーディオ信号のための可逆的なウォーターマーク技術が記載されている。

【発明の概要】

【0007】

本発明の第 1 の観点によれば、損失のないオーディオエンコーダは、第 1 のサンプリングレートで入力デジタルオーディオ信号を受け取り、かつ、それから、複数のサンプルを有しあつ第 1 のサンプリングレートよりも低い第 2 のサンプリングレートを持つ PCM デジタルオーディオ出力を生成するように構成され、

複数のサンプルの各々は、上位部分と下位部分とを持ち、

上位部分と下位部分とは、いずれも、第 1 のデコーダが損失なしで入力デジタルオーディオ信号を再生できるようにする情報を含み、

標準 PCM ストリームとして解釈されたとき、上位部分は、低減された帯域幅を持つ入力デジタルオーディオ信号のバージョンの損失のある表現を提供し、

上位部分は、前記低減された帯域幅よりも大きい帯域幅を持つ入力デジタルオーディオ信号の損失のある表現を第 2 のデコーダが再生できるようにする情報を含み、

入力デジタルオーディオ信号は、高周波数出力と低周波数出力とを持つ損失のない帯域分割器に結合され、高周波数出力は、損失のある圧縮出力と復元出力とを持つ圧縮ユニットに結合され、

上位部分は、低周波数出力及び損失のある圧縮出力に応じて引き出され、下位部分は、復元出力に応じて引き出される。

【0008】

本発明での使用のために設計されたものでない標準『遺産』PCM 再生装置は、一般に、ここでは『上位部分』と称される、一般に 44.1 kHz 又は 48 kHz の第 2 のサンプリングレートでサンプリングされたオーディオストリームの各サンプルの上位 16 ビットのみを受け取り又は再生し、およそ 0 ~ 20 kHz の帯域幅で損失のある表現をリスナへ提供する。第 2 のデコーダは、同じ 16 ビットの 44.1 kHz 又は 48 kHz ストリームから、拡張された帯域幅が再生されることを可能にする。第 1 のデコーダは、一般に、24 ビットストリームを受け取り、かつ、各サンプルの『下位部分』へ、すなわち 16 番目を超えるビットへもアクセスすることを見込んでいる。この追加情報は、88 kHz

又は 96 kHz のような、第 1 の、より高いサンプリングレートで、そして 0 ~ 40 kHz のような、より広いオーディオ帯域幅を持つ、入力オーディオ信号の損失のない再生が提供されることを可能にする。

【0009】

以下の説明では、復元出力又は復元信号が、修整出力又は修整信号として、より口語的な表現として参照されることがある。

【0010】

好みしくは、第 1 の損失のある表現は、時不变のフィルタリング効果と、サンプリングレートの低減効果と、時不变のノイズフロアを課す再量子化効果とを除いて、入力デジタルオーディオ信号の正確な表現である。もしサンプリングレートの低減を伴う量子化を含む全ての量子化が、一定のビット深度に対してかつ適切なディザを伴って実行されるならば、『損失のある』表現は、CD品質と同等の標準的な表現であり得て、数年前にも『オーディオ愛好家』向けの再生と考えられたであろう。これは、スペクトル的なノイズフロアと時折入力信号に応じた帯域幅とを動的に採用する、伝統的な『損失のあるコーデック』と対照的である。

【0011】

好みしくは、圧縮ユニットは、損失のある圧縮出力に結合された出力を持つ、損失のある圧縮ユニットを備える。

【0012】

損失のない帯域分割器は、オリジナル信号スペクトラムの一般には二半部の取り扱いを分割し、下半部は PCM として伝達され、上半部は圧縮されたフォーマットで伝達されるための鍵である。

【0013】

ある実施形態では、各上位部分は、16 バイナリビットからなる。ある実施形態では、各下位部分は、8 バイナリビットからなる。

【0014】

ある実施形態では、第 2 のサンプリングレートは、第 1 のサンプリングレートの半分である。特に好みしい第 2 のサンプリングレートは、48 kHz 及び 44.1 kHz を含む。

【0015】

本発明のエンコーダでは、第 2 のデコーダは、第 1 のサンプリングレートに対応するナイキスト周波数に等しいオーディオ帯域幅を再生し得る。あるいは、第 2 のデコーダは、第 1 のサンプリングレートに対応するナイキスト周波数の 4 分の 3 に等しい帯域幅を再生し得る。

【0016】

「ナイキスト周波数」という用語は、デジタルシステムのサンプリングレートの半分を意味するものと、通常は理解されている。したがって、一般に、第 1 のサンプリングレートは 96 kHz であり、第 2 のサンプリングレートは 48 kHz であり、第 1 のサンプリングレートに対応するナイキスト周波数もまた 48 kHz であって、第 2 のデコーダは、48 kHz であるナイキスト周波数までの信号の損失のある再生を提供する。他の構成は、第 2 のデコーダが 36 kHz までの損失のある再生を提供することを可能にする。その利点は、0 ~ 24 kHz の範囲でノイズフロアが若干低い点にある。

【0017】

ある実施形態では、圧縮ユニットは、損失のある圧縮ユニットに結合された入力と、復元出力に結合された出力とを持つ、損失のない圧縮ユニットを更に備える。損失のない圧縮器は、最下位ユニットにおけるビットの使用を最適化する。あるいは、もし修整出力が既に圧縮された形式又は『バックされた』形式であれば、別個の損失のない圧縮器は不要である。

【0018】

下位部分は、帯域分割器の低周波数出力に応じても引き出され得る。これは、帯域分割

器の低周波数出力の全体が上位部分中に伝達された場合よりも精密に量子化されたオリジナル信号を、第1のデコーダが損失なしで再生することを可能にする。

【0019】

好ましくは、損失のない帯域分割器の低周波数出力は、上位部分に結合された第1の出力と、下位部分に結合された第2の出力とを持つ分割器に結合されている。好ましくは、分割器は、ノイズシェーピングフィルタを備える。分割器は、帯域分割器のL.F.出力の量子化されかつ好ましくはノイズシェーピングされた表現を上位部分へ提供する。一方、その第2の出力は、量子化により除去された情報を第1のデコーダが再生することを可能にする。

【0020】

ある実施形態では、上位部分の中の複数のビットは、損失のない帯域分割器の低周波数出力に結合された第1の入力と、損失のある圧縮ユニットの圧縮出力に結合された第2の入力とを持つ減算器の出力に応じて引き出されることが好ましい。上位部分は、第2のデコーダの動作をサポートするために、圧縮出力を含まなければならない。しかしながら、圧縮出力は、オーディオ信号ではなくデータ信号であって、減算器の目的は、遺産装置により再生されたオーディオ信号への、このデータ信号の効果を補償することにある。

【0021】

本発明の第2の観点によれば、第1の観点による損失のないオーディオエンコーダに結合されたノイズシェーバを備えた装置が提供される。通常、このノイズシェーバは、96 kHzで動作して、48 kHzのサンプリング周波数にて24ビット出力ワードの制約下で、入力信号が損失なしで伝達されるように、エンコーダへの入力信号のワード幅を低減する。

【0022】

本発明の第3の観点によれば、ウォーターマーク出力を供給する損失なしの可逆的なウォーターマークエンコーダに結合された、第1の観点による損失のないオーディオエンコーダを備えた装置が提供される。しかも、当該装置は、コンフィギュレーションパラメータに応じてエンコーディングを実行し、かつ、ウォーターマークエンコーダは、デコーダによる使用のためにウォーターマーク出力の中にコンフィギュレーションパラメータを埋め込む。

【0023】

当該装置は、損失のないオーディオエンコーダの入力に量子化された信号を供給するノイズシェーバを更に備え得る。当該ノイズシェーバは、あるビット深度まで量子化し、コンフィギュレーションパラメータは、ビット深度を含む。また、当該装置は、下位部分の情報保持容量を超えないようにオーディオ品質を最大化するために、量子化のビット深度を選択する選択ユニットを更に備え得る。

【0024】

このように、本発明は、高品質・広帯域幅信号がベースバンドPCM伝送チャンネルを介して伝達され得て、伝送チャンネルがトップ16ビットを伝達するのみであっても良好な動作をし、エンコードされたストリームが、信号をベースバンドPCMと解釈する遺産装置によりデコードされた際に、帯域制限されたオーディオの妥当な抽出を更に提供するシステムを提供する。

【0025】

本発明の第4の観点によれば、対応する第1の観点のオーディオエンコーダにより第2のサンプリングレートで生成された複数の入力サンプルを有するPCM入力デジタルオーディオ信号を受け取るように構成されたオーディオデコーダが提供される。当該オーディオデコーダは、PCM入力デジタルオーディオ信号から、第2のサンプリングレートよりも高い第1のサンプリングレートを持つ出力デジタルオーディオ信号を生成するように更に構成され、

出力デジタルオーディオ信号と比較信号との間の差違は、0~5 kHzの周波数範囲で、静的な統計でスペクトル的にシェーピングされたノイズであって、比較信号は、入力デ

ジタルオーディオ信号から、フィルタリングの動作及び第1のサンプリングレートへのリサンプリングの動作により生成され、

出力デジタルオーディオ信号と第2の出力信号との間の差違は、0～5 kHzの周波数範囲で、静的な統計でスペクトル的にシェーピングされたノイズであって、第2の出力信号は、デコーダに送られるとき、各サンプルから下位部分を除去する以外はPCM入力デジタルオーディオ信号と一致する信号から作られ、

出力デジタルオーディオ信号は、エンコーダに提供されたデジタルオーディオ入力信号の正確なレプリカである。

【0026】

したがって、第4の観点のデコーダは、第1の観点による対応するエンコーダとともに使用されることが意図されている。当該エンコーダの出力は、単純なPCM信号と解釈されたときに、スペクトル的にシェーピングされてはいるが時間とともに変化することのないノイズフロアのような、オーディオ愛好家の基準を満足し得る。デコーダは、出力信号を生成するために、フィルタリング動作、リサンプリング動作及び量子化動作を実行する。比較信号は、デコーダのフィルタリング動作及びリサンプリング動作を模倣することにより、ただしデコーダの量子化なしで高精度で、生成され得る。その結果、出力デジタル信号と比較信号との間の差違は、デコーダにより持ち込まれた量子化歪みを抜き出す。デコーダへの入力は、好ましくはオーディオ愛好家の基準を満足する信号であるので、比較信号もまたオーディオ愛好家の基準を満足し、よって比較信号と出力信号との間の差違は、オーディオ愛好家の基準を満足し、かつ、したがって静的な統計でスペクトル的にシェーピングされたノイズに等価である量子化歪みのみを含むことになる。このことは、聴取により、又はスペクトラムアナライザの使用により、テストされ得る。

【0027】

本発明の第5の観点によれば、第2のサンプリングレートで複数の入力サンプルを有するPCM入力デジタルオーディオ信号を受け取り、かつ、それから、第2のサンプリングレートよりも高い第1のサンプリングレートを持つ出力デジタルオーディオ信号を生成するように構成されたオーディオデコーダが提供される。当該デコーダは、

高周波数入力と低周波数入力とを持ち、出力デジタルオーディオ信号を供給する、損失のない帯域合成器と、

損失のある入力と、復元入力と、出力とを持ち、出力は損失のない帯域合成器の高周波数入力に結合された伸張ユニットとを備え、

各入力サンプルは、上位部分と下位部分とを有し、

帯域合成器の低周波数入力は、上位部分に応じて引き出され、

伸張ユニットの損失のある入力は、上位部分に応じて、かつ下位部分から独立して引き出され、

伸張ユニットの復元入力は、下位部分に応じて、かつ上位部分から独立して引き出される。

【0028】

帯域合成器及び伸張ユニットは、対応するエンコーダで実行された帯域分割動作及び圧縮動作を逆転するために必要とされる。完全な損失のない再生は、完全な入力サンプルがデコーダに提供されることを要求するが、下位部分がない場合には損失のある再生をサポートすることも要求される。この理由から、伸張への損失のある入力はストリームの上位部分から送られ、帯域合成器への低周波数入力は実質的に上位部分から取られ、下位部分への依存は低周波数信号の分解能を改善するためのみであることもまた、要求される。

【0029】

好ましくは、帯域合成器の低周波数入力は、上位部分の中に含まれる全ビットに応じて引き出される。上位部分は、損失のない帯域合成器へ高周波数入力を提供する伸張ユニットへ送られるビットを含む。したがって、低周波数入力を引き出す際には、これらのビットを除外するのが自然であろう。これらのビットは、上位部分を標準PCMデコーダでデコードする遺産リスナの耳に入る信号に影響を及ぼす。しかしながら、これらのビットが

低周波数入力に寄与することを可能にすることが好ましい。エンコーダは、本発明のデコーダと標準PCMデコーダとの一致を結果としてもたらす様で、『負の埋め込みデータ』の原理に従って他のビットを調整することにより、これらのビットを補償することができる。

【0030】

好ましくは、帯域合成器の低周波数入力もまた、下位部分に応じている。このことは、下位部分がデコーダで使用可能である場合に改善されるように、その分解能の信号が帯域合成器の低周波数入力に提供されることを可能にする。

【0031】

更に、出力デジタルオーディオ信号と比較信号との間の差違は、0～5kHzの周波数範囲で、静的な統計でスペクトル的にシェーピングされたノイズであって、比較信号は、PCM入力デジタルオーディオ信号から、フィルタリングの動作及び第1のサンプリングレートへのリサンプリングの動作により生成されることが好ましい。したがって、本発明の第4の観点に関して上記した利点の1つは、本発明の第5の観点により提供される利点と結合され得る。

【0032】

好ましくは、オーディオデコーダは、対応するオーディオエンコーダにより生成された信号を受け取るように構成され、出力デジタルオーディオ信号は、対応するオーディオエンコーダに提供されたデジタルオーディオ入力信号の正確なレプリカである。

【0033】

このように、第4の観点に関して上記した他の利点は、本発明の第5の観点により提供される利点と結合され得る。

【0034】

当業者により評価されるように、本発明の損失のないオーディオエンコーダの他の採用も可能である。更に、他の観点によれば、対応するデコーダは、エンコーダとデコーダとを備えた通信システムと考えられる。

【0035】

本発明の実施例を、添付の図面を参照しながら詳細に説明する。

【図面の簡単な説明】

【0036】

【図1】(a)は単純な損失のある帯域幅拡張を持つ従来のエンコーダを示す図であり、(b)は対応するデコーダを示す図である。

【図2】(a)は改善された損失のある帯域幅拡張を持つエンコーダを示す図であり、(b)は対応するデコーダを示す図である。

【図3】(a)は単純な損失のある帯域幅拡張を持つノイズシェーバ及びエンコーダを示す図であり、(b)は対応するデコーダを示す図である。

【図4】(a)はリフティングを用いた損失のない帯域分割を示す図であり、(b)は対応する帯域合成を示す図である。

【図5】(a)は単純な二重の互換性を持つ損失のない帯域幅拡張を有するノイズシェーバ及びエンコーダを示す図であり、(b)は対応するデコーダを示す図である。

【図6】(a)は改善された二重の互換性を持つ損失のない帯域幅拡張を有するノイズシェーバ及びエンコーダを示す図であり、(b)は対応するデコーダを示す図である。

【図7】(a)はノイズシェーピングされた分割器を用いた、二重の互換性を持つ損失のない帯域幅拡張を有するノイズシェーバ及びエンコーダを示す図であり、(b)はノイズシェーピングされた合成器を用いた、対応するデコーダを示す図である。

【図8】(a)はノイズシェーピングされた分割器を示す図であり、(b)は対応する合成器を示す図である。

【図9】図7(a)のエンコーダの部分の他の構成と、ノイズシェーピングされた分割器とを示す図である。

【発明を実施するための形態】

【0037】

損失のある帯域幅拡張

需要者オーディオのための商業的な「スケーラブル」伝送システムが、米国特許第6226616号明細書 (You et. al.: "Sound Quality of Established Low Bit-Rate Audio Coding Systems without loss of Decoder Compatibility") に掲載された。損失のある圧縮がなされたオーディオ信号を表すデータストリームを、標準S P D I F デジタルオーディオインターフェイスで転送され得る16ビットワードへパッケージ化する確立したシステムから始めて、強化されたシステムは、より高いオーディオ品質を可能にするために、オリジナルシステムのために設計されたデコーダと互換性を持つ様で、更に「拡張ストリーム」を同じフォーマットへパックするというオプションを提供する。しかしながら、P C Mストリームを伝達するのにS P D I Fがしばしば用いられるが、この場合の『互換性』は、所有者デコーダの確立したインフラストラクチャに関するものであって、特別なデコーダなしにP C Mストリームを再生するために採用された装置に関するものではなく、このことが本発明の目的である。

【0038】

図1 (a) 及び図1 (b) は、上記文献でK o m a m u r a が提案したスキームと同様の、P C M互換性を持つ帯域幅拡張スキームを示している。図1 (a) のエンコーダでは、帯域分割器3が、例えば96 k H zのレートでサンプリングされて潜在的に周波数範囲0 ~ 48 k H zの情報を伝えるオリジナル信号2を受け取る。帯域分割器は、直交ミラーフィルタ (Quadrature Mirror Filter) のような既知の方法を使用して、各々低周波数0 ~ 24 k H zの情報と高周波数24 ~ 48 k H zの情報を伝える低周波数 (L F) 信号15と高周波数 (H F) 信号28とに信号2を分割する。L F信号とH F信号との各々は、48 k H zで、すなわちオリジナルサンプリングレートの半分のレートでサンプリングされる。H Fストリームは、既知の方法を用いた損失のある圧縮4がなされて、少数ビット、例えば1、2又は3ビットを持つデータストリーム7となる。一方、L Fストリームは、切り捨てられ又はノイズシェーピング5がなされて、多数ビット、例えば15、14又は13ビットを持つ信号6となる。図1 (a) は、データストリーム7が3ビットを持ち、信号6が13ビットを持つ例を示している。そして、図1 (a) に示すように、2つのストリームからのサンプルを、16ビットサンプル、つまりビットB₁ ~ B₁₆を持つ単一の複合出力ストリーム8へパックするのが直接的である。16ビット出力ストリームは、低いレート、例えば48 k H zのサンプルを含み、サンプルストリーム8の再生も可能な標準需要者装置を用いて転送されかつ格納され得る。

【0039】

K o m a m u r a の提案は、損失のある圧縮の基礎としてA D P C M (Adaptive Differential Pulse Code Modulation) を用いる。K o m a m u r a は、24 k H zのレートでH Fストリームの表現を提供するために、A D P C Mユニットの前にダウンサンプラーを置く。この表現は次に1サンプルあたり2ビットに圧縮され、当該2ビットはシリアル化されて48 k H zの1ビットストリームとなる。したがって、H F情報は最終16ビット出力のうちの1ビットのみを占め、15ビットのL F分解能を可能にする。ダウンサンプリング自体は損失のある処理であるので、K o m a m u r a のダウンサンプラー及びA D P C Mユニットは、損失のある圧縮ユニット4であると考え得る。ダウンサンプリングの結果、デコーダは、48 k H zまでの、又は36 k H zまでの周波数の明瞭な再生を提供することができる。

【0040】

図1 (b) は、図1 (a) に対応するデコーダを示している。ここでは、ストリーム6及び7が、各々転送されたストリーム8のトップ13ビットB₁ ~ B₁₃及びボトム3ビットB₁₄ ~ B₁₆からアンパックされる。伸張ユニット9は、帯域分割器3によって作られたL F信号15及びH F信号28に実質的に同様のL F信号及びH F信号が帯域合成器10に送られるように、圧縮ユニット4の動作を逆転させる。帯域合成器10は、周波数範囲0 ~ 24 k H zにおけるオーディオ品質が主としてノイズシェーパ5により制限さ

れ、かつ超音波範囲 24 ~ 48 kHz では圧縮ユニット 4 と伸張ユニット 9 との結合動作により歪みが持ち込まれる、出力信号 11 を作るために、これら 2 つの信号を再結合させる。

【0041】

デコーダを持たずにストリーム 8 を PCM オーディオとして再生する『遺産』リスナは、主として帯域分割器からのノイズシェーピングされた（又は切り捨てられた）LF 出力を耳にするが、これは、オリジナル信号 2 のダウンサンプリングされた低品質バージョンとして受け容れ可能であろう。しかしながら、圧縮された HF 信号 7 を含むストリーム 8 の最下位 3 ビットは、遺産リスナのプレーヤのオーディオ出力にも寄与する。理想圧縮器の出力はノイズ様であって、そのほか冗長を含み、原理的には改善された圧縮を与えるよう除去され得るものである。実用上は、音の歪みを除去して圧縮器の出力を真にノイズ様にするように、明示的スクランブルを提供することが必要であり得る。本明細書では、もし必要なら圧縮器 4 がそのようなスクランブルを内部に含むものとする。これにより、その出力が統計的に独立なバイナリビットからなることを保証する。

【0042】

本明細書を通じた他の仮定は、圧縮及び伸張といった処理が瞬時的であることである。実用上は、これらの処理は信号ディレイを招く。したがって、並列信号バスにディレイ補償が持ち込まれなければならない。明瞭性のために、処理ユニットの正常な動作にとって便利であり又は必要である場合には、そのようなディレイ補償は図面から省略され、同様に図面はブロック中に信号サンプリングを組織化することを妨げない。

【0043】

負の埋め込みデータを用いた帯域幅拡張

図 2 (a) では、損失のある圧縮器 4 の出力はデータ信号であるが、図 1 (a) について説明したように、それはオーディオ信号として遺産リスナの耳にも入る。このような二重解釈は、図 2 (a) にて認識される。つまり、ユニット 12 は、実用上は存在しなくてもよいが、信号 7 がデータ信号と PCM オーディオ信号との二重解釈を持つことを強調するために含められており、信号 7 は、オーディオ信号として解釈される場合には、右詰めされたものと考えられて、16 ビットワードのボトム 3 ビット、すなわちビット B₁₄ ~ B₁₆ を占め、当該ワードの他のビットはゼロである。

【0044】

したがって、オーディオ信号として解釈された信号 7 は、減算器 13 へ送られる結果、ノイズシェーパ 5 は、出力ワード 8 のトップ 13 ビット B₁ ~ B₁₃ に配置される、修飾された 13 ビット信号 6' を作るために、LF 信号とは逆位相で信号 7 を受け取る。遺産リスナは、信号 6' 及び 7 の和である、PCM オーディオ信号として解釈された出力ワード 8 の全体を耳にする。したがって、遺産リスナは、完全なワード 8 のボトム 3 ビットを介して直接に、またワード 8 のトップ 13 ビットにてノイズシェーパを介した逆位相で、圧縮器信号 7 を耳にする。したがって、これらの圧縮器信号 7 の 2 つの表現は、打ち消し合う。これは、M. A. Gerzon and P. G. Craven, "A High-Rate Buried Data Channel for Audio CD," J. Audio Eng. Soc. Volume 43 Issue 1/2 pp. 3-22; February 1995 に記載された『負の埋め込みデータ』の例である。

【0045】

内部的には、ノイズシェーパ 5 は、13 ビット量子化器とノイズシェーピングフィルタとを含む。圧縮器からのノイズを打ち消すのと同様に、負の埋め込みデータは、13 ビット量子化器のための負のディザを提供する。加法性ノイズ以外の量子化歪みは、いまや 13 ビットレベルではなく、16 ビットレベルである。13 ビットレベルにおける加法性ノイズは、ノイズシェーピングフィルタによりシェーピングされて、潜在的に 2 ビット以上の知覚的利点を提供する一方、負のディザは、従来の TPDF ディザよりも 4.77 dB だけ小さいノイズを持ち込む。ゆえに、知覚されるパフォーマンスは、TPDF ディザを用いる 16 ビットシステムのパフォーマンスと等価である。

【0046】

対応するデコーダは、図2(b)に示されている。このデコーダは、帯域合成器10へのLF入力が、トップ13ビットのみではなくて16ビット複合信号の全体から送られることを除いて、図1(b)のデコーダと同様である。したがって、このLF信号は、信号6'及び7の結合であり、遺産リスナの耳に入るものと同様であって、リスナは負のディザと同様の利点を享受する。

【0047】

Gerzon及びCravenによる上記参考文献は、非整数ビットの他のデータをPCM信号のボトムビットの中に「埋め込む」方法をも記載している。特に、2チャネル(ステレオ)ストリームの各チャネルに半整数ビットを埋め込むのが直接的である。簡単のために、本明細書では整数を仮定するが、ここに記載した着想が非整数ビットの圧縮データとともに使用可能であることは明らかであろう。

【0048】

損失のない帯域幅拡張 - 一般的検討

図3(a)及び図3(b)は、各々単純な損失のない帯域幅拡張システムのためのエンコーダ及びデコーダを示している。図3(a)及び図3(b)と図1(a)及び図1(b)との間の構造的類似性は明白であるが、損失のない再生のための要求は、追加の制約を課し、かつ、損失のある場合には生じることのなかった、量子化の観点への注意深い配慮を必要とする。

【0049】

損失のないシステムは情報を捨てることが許されないので、伝送チャネルは、少なくとも伝達されるべき信号中の情報と同じ大きさの情報保持容量を持たねばならない。損失のない圧縮での経験は、16ビット以上の分解能を持つ96kHzオーディオ信号における冗長が一般に約8ビットであることを示唆している。したがって、16ビットの96kHz信号は、1サンプルあたり8ビットのデータレートに圧縮され得る。そして、24ビットの96kHz信号は、16ビットに圧縮され得る。したがって、16ビットの96kHz信号は、通常、16ビットの48kHzチャネルを通じて転送され得る。しかしながら、それは互換性を持たない。なぜなら、最適に圧縮された信号は、PCM信号として解釈された場合には、フルスケールのホワイトノイズとして現れるからである。PCM互換性の要求は、PCM信号中への冗長を課し、したがってより大きいワード幅を必要とする。

【0050】

したがって、一般に、16ビットの96kHz信号を16ビットの48kHzチャネルへ損失なしでかつPCM互換性をもってパックすることは不可能である。また、一般に、24ビットの96kHz信号を24ビットの48kHzチャネルへ損失なしでかつPCM互換性をもってパックすることも不可能である。しかしながら、16ビットの96kHz信号の、24ビットの48kHzチャネルへの、PCM互換性を持つ損失のないパッキングは、通常は実行可能である。

【0051】

現在、『96/24』(すなわち、サンプリングレート96kHz、ビット深度24ビット)が、CD(Compact Disc)の『44/16』からの次ステップとして広く認識されている。しかしながら、Gerzonによって1995年に、ノイズシェーピングにとって96kHzサンプリングが大いに有利であって、CDで広く用いられてきた44.1kHzシェーパよりも、高周波数ノイズスペクトラムにおける緩やかな上昇をもって大きい知覚的改善を可能にすることが現実化された。Gerzonの96kHzシェーパのための係数は、ほぼ5ビットの知覚的改善を提供するものであって、Acoustic Renaissance for Audio, "A Proposal for High-Quality Application of High-Density CD Carriers" private publication (1995 April); reprinted in Stereophile (1995 Aug.); in Japanese in J. Japan Audio Soc., vol. 35 (1995 Oct.); available for download at www.meridian-audio.com/araに与えられている。Stuartは、人の聴覚能力を考慮した注意深い分析を提供しており("Coding for High-Resolution Audio Systems" J. Au

dio Eng. Soc., Vol. 52, No. 3, 2004 March, 特に図 16 参照)、これから、T P D F ディザをもって(ただし、ノイズシェーピングなしで) 20.5 ビットに適切に量子化された 44.1 kHz サンプリングのデジタルシステムは、配布媒体としての十分なダイナミックレンジを常に提供する、ということが結論付けられよう。96 kHz サンプリングが採用される場合には、ノイズシェーピングされないときのノイズスペクトル密度が、更に 3.4 dB だけ低減される。適切なノイズシェーピングを伴う 16 ビットの 96 kHz チャンネルは、配布フォーマットとして全面的に適切であって、ある余裕をもってオーディオ愛好家の要求を満足する、との結論付けが可能である。

【0052】

したがって、情報理論の議論とともに音響心理学の議論を考慮すれば、24 ビットのような大きいビット深度を持ち得る 96 kHz 入力信号を、16 ビットのような小さいビット深度へ再量子化することは、必要であり、かつ許容される。ゆえに、図 3 (a) には 96 kHz ノイズシェーパ 1 が示され、『A』で識別される量子化された信号 2 を供給するように、不特定の分解能を持つ 96 kHz 入力信号を、例えば 17 ビットに再量子化する。帯域分割器 3 は、損失のない分割器であって、また 17 ビットの低周波数出力 15 と、18 ビットと表示された分解能を持つ高周波数出力 28 とを作る。ただし、18 ビットの全てを与えることは、現実のオーディオ信号ではまれであろう。したがって、24 ビット出力ワード 16 を仮定したとき、低周波数出力は出力ワード 16 のうちの 17 ビット $B_1 \sim B_{17}$ を占め、残る 7 ビット $B_{18} \sim B_{24}$ は、損失のない圧縮器 14 で作られた、高周波数信号 28 の損失なしで圧縮されたバージョンである。

【0053】

図 3 (b) のデコーダでは、損失のない伸張ユニット 9 が、高周波数信号 28 のレプリカとして信号 28a を再生する。そして、損失のない帯域合成器 10 は、損失のない帯域分割器 3 により作られた信号 15 及び 28 と同様の信号を受け取り、その結果、信号 2 の損失のないレプリカとして、出力信号 11 を再生することができる。したがって、信号 11 もまた『A』で識別される。

【0054】

量子化は損失のある処理なので、図 3 (a) 及び図 3 (b) に示された全処理が損失のない処理ではあり得ず、損失がないのは、エンコーダ中の信号 2 からデコーダの出力 11 までのパスである。したがって、図 3 (a) 及び図 3 (b) のエンコーダ及びデコーダにより提供される処理は、全体として、入力信号のノイズシェーピングされたバージョンを届ける。ここで、ノイズシェーピング 1 は、ディザを含みかつ一定のビット深度を持つオーディオ愛好家の基準を満足するように選択され得る。

【0055】

「リフティング」を用いた損失のない帯域分割器及び帯域合成器

図 3 (a) 及び図 3 (b) の構成は、損失のない帯域分割器 3 及び帯域合成器 10 を必要とする。ここで、『損失のない』とは、処理の中の量子化誤差を考慮に入れた、ビット的に正確な再生のことである。そのような損失のない帯域分割器及び帯域合成器を構成する方法はいくつかある。図 4 (a) 及び図 4 (b) に示されたものは、「リフティング」原理に基づくものである (Calderbank, Daubechies, Sweldone and Yeo: "Wavelet Transforms That Map Integers to Integers" Applied and Computational Harmonic Analysis, vol. 5, pp 332-369 (1998), 特にその図 4 及び図 5 参照)。

【0056】

図 4 (a) の帯域分割器では、96 kHz のような『 $2 \times$ 』サンプリングレートでサンプリングされた入力ストリームが、各々 48 kHz のような『 $1 \times$ 』サンプリングレートで奇数番サンプルと偶数番サンプルとの別個のストリームを作るためにデインターリーブされる。これら 2 つのストリームは、ほとんど又は全く共起関係にはなく、 $2 \times$ ストリーム中のオリジナルの低周波数信号は、偶数ストリームに対する奇数ストリームにて、 $1 \times$ サンプルの半分だけ遅れて又は進んで現れる。

【0057】

ここで、2つのリフティングステップが適用される。1つのリフティングステップは、ある信号の関数を他の信号に加算する。つまり、

$$X' = X + f(Y)$$

$$Y' = Y$$

であって、これを単純に逆転させれば、

$$X = X' - f(Y')$$

$$Y = Y'$$

となる。これら2つの場合の間で関数 f が正確に一致（状態変数の量子化又は初期化を含む。）するという条件では、損失がない。

【0058】

図4(a)の第1のリフティングステップでは、『X』は奇数番サンプルのストリームで識別され、『Y』は偶数番サンプルのストリームで識別される。奇数ストリームから偶数ストリームを引き算すると実質的に低周波数が打ち消されるが、最善の打ち消しのためには、半サンプルシフトの修正が必要である。したがって、偶数番サンプルに半サンプルディレイを適用したい。これは、偶数のタップを持つ対称 FIR フィルタによって近似され得るが、それは因果的でなく、フィルタ『f』は、実際には、ある n に対して ($n + 1 / 2$) サンプルディレイを実行し、奇数パスには n サンプルの補償ディレイが存在する。次式は、 $n = 2$ で 2.5 サンプルのディレイを持つようなフィルタの例である。

【0059】

【数1】

$$\frac{3}{256} - \frac{25}{256}z^{-1} + \frac{75}{128}z^{-2} + \frac{75}{128}z^{-3} - \frac{25}{256}z^{-4} + \frac{3}{256}z^{-5}$$

【0060】

10～20タップの長さのフィルタが、『HF』ストリームを供給するのに妥当であり得る。このフィルタは、オリジナルスペクトラム、すなわち 24kHz よりもかなり低い周波数のうちのボトム側の半分のほとんどを、良好に阻止する。

【0061】

$2\times$ ストリームが 96kHz でサンプリングされるものと再び仮定すれば、オリジナルスペクトラムのトップ側の半分は、ともにデインターリーブユニットから逆位相で現れる偶数及び奇数ストリームの双方にて、 $0\sim24\text{kHz}$ にエイリアスを生じる。したがって、 $24\sim48\text{kHz}$ の範囲のオリジナル信号は、第1のリフティング動作により振幅が2倍になり、 $1\times$ のHF出力は、潜在的に $2\times$ 入力の振幅の2倍の振幅を持つ。これは、図3(a)にてHF出力28が17ビットではなくて18ビットを持つものとして示されている理由である。

【0062】

図4(a)中の第1のリフティングステップは、奇数サンプルストリームに影響を及ぼさず、その結果、奇数サンプルストリームは、オリジナルの $2\times$ スペクトラムのトップ側及びボトム側の両半部からの信号を、同程度に伝える。第2のリフティングステップの目的は、偶数ストリームからHF出力を引き算することにより、オリジナルの高周波数情報を除去することにある。再び『半サンプルディレイ』フィルタ（実際には $n - 1 / 2$ サンプル）が時間合わせのために必要とされるとともに、HF出力の2倍になった振幅を補償するために 0.5 をかける乗算が必要とされる。

【0063】

図4(b)は、図4(a)のリフティングステップの順序が逆転されることを、右から左への信号の流れで強調した、対応する帯域合成器を示す。 $1\times$ サンプリングレートでの『奇数』及び『偶数』の結果は、 $2\times$ サンプリングレートでオリジナルストリームを損失なしに再生するために、インターリーブされる。

【0064】

2つのリフティング動作は、1対のストリーム（LF, HF）を供給する。このストリームでは、クロスオーバー近傍のLFストリームの正確な応答は理想的でなく、カットオフ前に若干上昇し得る。これが問題であると考えられる場合には、調整されたフィルタ構造とともに3つのリフティング動作を使用することで、問題を回避し得る。

【0065】

各量子化 Q_1, Q_2 は、もし帯域分割器への入力が-1から+1までの信号範囲を占める17ビット信号である場合には、オリジナルのステップサイズ、例えば $2^{-1/6}$ に一致すべきである。図4(a)における帯域分割器のLF出力及びHF出力も、そのようなオリジナルのステップサイズに量子化される。

【0066】

損失のない再生のため、デコーダ中の各量子化 Q_1, Q_2 は、エンコーダ中の対応プロックと同様の動作、例えば両者とも切り上げ、又は両者とも切り下げという動作をしなければならない。

【0067】

損失のない帯域幅拡張 - 単一の互換性

図3(a)に戻って、損失のない帯域分割器3に17ビット入力が与えられるとき、その合計出力ビット数（半分のサンプリングレートで）は、 $17 + 18 = 35$ ビットであって、これは、要求される24ビット出力ワードに明らかに適合しない。

【0068】

HF信号が潜在的に18ビットの情報を含む一方、実用上、そのピークレベルは、「力強い」商業的録音であっても、理論最大値より35dB以上も低い。損失のない圧縮がビット数を低減する手段として示されていることは、明らかである。損失のない圧縮器は、元来、可変のデータレートを作り出し、これは、実用上は、例えば FIFO (First In First Out) バッファを用いたバッファリングにより平滑化される必要がある。一般に、帯域分割により作られたHF信号は、標準オーディオ信号よりも「バースト的」に現れるので、バッファリングはより重要でさえある。明瞭性のために、図面にはここに必要なバッファが示されていないが、MLP圧縮システムの場合と同様に、損失のない圧縮器及び伸張器の各々にそのようなバッファが組み込まれていることを仮定している。もちろん、FIFOバッファリングはディレイを持ち込むので、時間合わせを維持するように、(LF信号パスのような)並列信号パスにて一定のディレイを加えることが必要である。ただし、再び明瞭性のために、そのような一定のディレイは図面から省略されている。

【0069】

商業的な96kHz録音の970例からなるコーパスに基づくテストは、0.3秒のFIFOバッファで、複合LFと損失なし圧縮HFとの情報は、もし15ビットと18ビットとの間のビット深度に量子化されるならば、97.6%のケースで24ビットに適合することを示している。

【0070】

したがって、一般には、異なる量子化深度でのエンコーディング試行は、各アイテムがエンコードされる際に使用されるべき最大の量子化深度を確立するのに用いられ得る。96kHz量子化を粗くすることは、複合情報により必要とされるビット幅を2つの方法で低減する、すなわち、

- ・ 直接には、LF信号が粗く量子化されるので、
 - ・ 間接には、HF信号が、粗く量子化される結果、少ないビットからなるので、
- ということが判る。

【0071】

しかしながら、粗い量子化は、またHF信号の中のシェーピングされたノイズを増加させる。これが重大な効果をもたらすかどうかは、HFパスの信号中でノイズが、つまり時間とともに変化し得るものが、そしていかなる時点でも損失のないエンコーダのFIFOバッファ中に格納されたデータに寄与する各瞬間で異なるものが優勢であるかどうかにかかっている。経験的に、96kHz量子化を1ビットだけ粗くすることは、48kHzで

の複合情報を 1 . 5 ビットだけ低減し得ることが知られている。

【 0 0 7 2 】

16 ビットのオリジナル素材の場合、複合情報は、しばしば 24 ビットに直接適合する。この場合には、図 3 (a) に示されている前段の量子化器は除去され得る。

【 0 0 7 3 】

既に示されているように、図 3 (b) の『 A 』で示されているデコーダの出力 11 は、同じく『 A 』で示されているエンコーダ中の信号 2 の、損失のないレプリカである。したがって、デコードされた出力 11 を耳にしたリスナは、単に 16 ビットのみへの量子化であっても 20 ビット量子化又は 21 ビット量子化と等価なノイズ密度を 0 ~ 7 kHz の範囲で提供し得る、 96 kHz ノイズシェーバの利益を享受する。

【 0 0 7 4 】

デコーダを持たない『遺産』リスナは、PCM 信号として解釈されたエンコーダ出力を耳にする。したがって、主として帯域分割器の LF 出力が、しかし潜在的には損失のない圧縮器の出力もが、24 ビットワードのうちのボトムビットにて PCM 信号として解釈される。既に言及したように、この出力は、それが未だノイズ様の信号でない場合には、ランダム化されなければならない。

【 0 0 7 5 】

また、遺産リスナは、図 4 (a) 中の量子化器 Q_1 及び Q_2 によって作られる量子化歪みを体験する。これらが帯域分割器の LF 出力に結合されているからである。これらの歪みは、ディザの使用により良性にされ、かつノイズシェーピングにより知覚的に低減され得る。しかしながら、損失のない再生を維持するため、図 4 (b) のデコーダは、その量子化器 Q_1 及び Q_2 にて同様のノイズシェーピングと同様の同期したディザとを用いなければならない。更に、ノイズシェーバが状態変数を有する場合には、デコーダ及びエンコーダで同様に、これらの変数を初期化することが必要であり得る。

【 0 0 7 6 】

二重の互換性：単純なアプローチ

図 5 (a) は、図 3 (a) 及び図 1 (a) に示された着想を結合したエンコーダを示しており、次の 3 つの聴取オプションを提供する。すなわち、

・ 遺産リスナは、1 × サンプリングレートで信号の 13 ビット表現を耳にするが、図 2 (a) 及び図 2 (b) の、ノイズシェーピングの利益と負のディザの利点とは得られない、

・ 複合信号のトップ 16 ビットのみにアクセスするリスナは、13 ビット表現の損失のある帯域幅拡張を享受するために、図 1 (b) のデコーダを使用し得る、

・ 全 24 ビットにアクセスするリスナは、『 A 』点にて、すなわち 96 kHz シェーバの結果として臨界的な周波数範囲 0 ~ 7 kHz で 17 ビット又は 18 ビットの分解能をもって、13 ビット信号の完全な帯域幅を持つ損失のない再生を享受するために、図 5 (b) のデコーダを使用し得る、

というオプションである。

【 0 0 7 7 】

信号『 A 』が 13 ビットに量子化されているので、帯域分割器 3 は、出力ワード 16 のトップ 13 ビット $B_{1 \sim B_{13}}$ に直接に適合する 13 ビットの LF 出力 15 を作るためにも構成され得る。HF 出力 28 は、損失のある圧縮 4 がなされ、かつ出力ワード 16 の第 14 ~ 第 16 ビット $B_{14 \sim B_{16}}$ へと桁合わせ 12 がなされる。したがって、16 ビットリスナのために、出力ワード 16 の上位部分 8 は、上記 2 つの黒丸により与えられるように、図 1 (a) にて 16 ビットワード 8 が提供したのと同じデコーディングオプションを提供する。

【 0 0 7 8 】

24 ビットリスナのための損失のないエンコーディングをサポートするために、図 5 (a) のエンコーダと同様のエンコーダは、圧縮された信号 27 を供給するように、HF 信号 28 の損失のない圧縮 14 を提供し得る。そして、圧縮された信号 27 は、出力ストリ

ーム 1 6 の下位部分 1 7、すなわち $B_{17} \sim B_{24}$ に位置する。しかしながら、エンコーダのために改善がなされて、図 5 (b) のデコーダ中で使用される損失のある伸張ユニット 9 のレプリカ 9' を組み入れ、損失のない圧縮ユニット 14 に送られる『修整』信号を形成するようにユニット 9' の出力を非圧縮の HF 信号 28 から減算 18 する。損失のある圧縮及び伸張の適切な設計があれば、減算 18 は、損失のある圧縮された信号 7 が消耗するデータレートにほぼ等しい大きさだけ、圧縮された修整信号 27 のデータレートを低減し得る。

【 0 0 7 9 】

図 5 (b) のデコーダは、修整信号のレプリカを供給するように、圧縮されたストリーム 27 を伸張 19 し、その修整信号のレプリカは、エンコーダ中の減算 18 を補償するために、損失のある伸張器 9 の出力に加算 20 されて、帯域分割器出力 28 のレプリカ 28a を供給する。したがって、帯域合成器 10 は、帯域分割器 3 からの信号 15 及び 28 と同じ信号 15 及び 28a の供給を受け、信号 2 の正確なレプリカ『 A 』である出力 11 を供給することができる。

【 0 0 8 0 】

一般に、図 5 (a) に示された伸張、減算、及び損失のない圧縮は、データレートが不十分であって、修整信号を直接に提供するための損失のある圧縮器を採用することにより、修整信号のよりコンパクトな表現が通常は引き出され得る。例えば、Yuらは、損失のある MPEG 4 コーデックが、MPEG - SLS として損失のない動作に、いかにして効率良く拡張されるかを示している (Yu, Geiger, Rahardja, Herre, Lin, and Huang: "MPEG-4 Scalable to Lossless Audio Coding", Audio Eng. Soc. 117th Convention 2004 October 28-31 San Francisco, AES preprint # 6183)。

【 0 0 8 1 】

したがって、図 6 (a) では、これら全ての処理が単一の圧縮ユニット 21 の内部で実行されることが仮定され、圧縮ユニット 21 が効率良くパックされた修整信号を生成するので、別個の損失のない圧縮器の要求は生じない。逆の処理が図 6 (b) 中の伸張ユニット 22 の内部で実行されることが同様に仮定され、伸張ユニット 22 は、標準的な損失のある圧縮された信号 7 と修整信号とを入力として受け取る。

【 0 0 8 2 】

したがって、ある好ましくない実施形態では、圧縮ユニット 21 は、図 5 (a) 中の破線で囲まれた枠内に示された内部サブユニットを含み、同様に伸張ユニット 22 は、図 5 (b) 中の破線で囲まれた枠内の内部サブユニットを含み得るが、これは次善の構成である。

【 0 0 8 3 】

図 6 (a) 及び図 6 (b) は、HF 信号及び LF 信号の各々の量子化深度の間の異なる関係をも示している。96 kHz 量子化は 15 ビットへの量子化であるが、損失のない帯域分割器の LF 出力 15 は、単に 13 ビットのみへ量子化され、一方の HF 出力は、18 ビットへ量子化される。このような量子化深度の不均等は、図 5 (a) の帯域分割器の LF 出力から最下位 2 ビットを取り、それらのビットを HF ワードのボトムへ加えることにより達成され得る。より洗練された方法のためには、上記 Calderbank らの文献の 2.3 章 "Different Expansion Factors for the High and Low Pass Channels" を参照されたい。この変更は 16 ビットリスナの助けにはならないが、長い HF ワードから引き出された修整信号が依然として 8 ビットに適合するように十分に圧縮されているならば、24 ビットリスナは、追加 2 ビットの分解能の利益が得られる。

【 0 0 8 4 】

本明細書及び図面では、13 ビット及び 15 ビットのような 96 kHz 量子化ビット深度が示されているが、それは単なる例示であって、それに限定することを意図したものではない。同じことは、96 kHz の周波数自体にも言える。同様に、損失のある圧縮出力として示された 3 ビットは例であって、より少ないビット数への圧縮が実用上は使用可能である。

【0085】

改良された二重の互換性

図6(a)及び図6(b)のスキームは、24ビットリスナにとって優れたパフォーマンスを提供するが、遺産リスナにとって、またデコーダを持った16ビットリスナにとっては、図2(a)のエンコーダを用いた場合よりもパフォーマンスが悪くなる。図6(a)及び図6(b)は、図2(a)及び図2(b)のスキームにより提供される、LF信号をノイズシェーピングする利点と、LF信号のための負のディザとして、圧縮されたHF信号を使用する利点とを失うからである。図7(a)のエンコーダは、これらの利点を回復して、複合ワード16の次の3つの聴取可能性を提供するように設計されている。すなわち、

- ・ 遺産リスナにより、そのプレーヤが上位部分8を標準16ビットPCM信号として解釈する、
- ・ リスナにより16ビット上位部分のみが受け取られて、図2(b)のデコーダが使用される、
- ・ リスナにより全24ビットが受け取られて、図7(b)のデコーダが使用される、という可能性である。

【0086】

出力ワードの下位部分17とそれを供給する信号バスとを削除し、ノイズシェーピングされた分割器5'をノイズシェーパ5に置き換えるならば、図7(a)のエンコーダは、図2(a)のエンコーダと等価になるという点に注意すべきである。したがって、図2(a)及び図2(b)のスキームを参照して既になされた説明は、遺産であるか図2(b)のデコーダを使用するかにかかわらず16ビットリスナにも通用するので、これら2つの場合の正常なデコーディングが保証される。そこで、ここでは、リスナは複合ワードの全24ビットを受け取るとの仮定のもとに、図7(b)とともに図7(a)の動作に説明を集中することとする。

【0087】

図7(a)の新規な特徴は、ノイズシェーピングされた出力6'、ノイズシェーピング処理で除去された情報を含む『LSB』信号23とを提供するノイズシェーピングされた分割器5'である。信号23は、出力ワード16の下位部分17のうちビットB₁₇~B₂₀のいくつかへ導かれる。その結果、図7(b)のデコーダでは、信号6'及び23の双方が、ノイズシェーピングされた合成器24にて使用可能となり、当該合成器24は、信号26のレプリカとして信号26aを再生する。信号7は、図7(a)のエンコーダ中の信号15のレプリカとしてLF信号15aを供給するために、信号26aに加算25される。

【0088】

図7(b)中の伸張器22は、図6(b)の場合と同様に機能して、HF信号28の損失のない再生であるHF信号28aを提供する。したがって、損失なしで再生されたLF信号及びHF信号を受け、帯域合成器10は、信号2の損失のないレプリカとして出力信号11を再生することができる。

【0089】

エンコーダは、複合ワードの上位部分8と下位部分17との間でLF信号15の情報を分けるので、当該エンコーダは、図6(a)のエンコーダの場合よりも高精度の96kHz信号2を扱うことができる。図7(a)及び図7(b)は、17ビットを持つ信号2に対してシステムがどのように構成されるかを示している。信号2が16ビットの場合には、信号26も16ビットを持ち、信号23は3ビットを持ち、したがって『修整(パックされた)』信号27は5ビットになる。信号2が18ビットの場合には、信号26も18ビットを持ち、信号23は5ビットを持ち、したがって『修整(パックされた)』信号27は3ビットになる。

【0090】

各々ノイズシェーピングされた分割器5'及び合成器24は、種々の方法で実装され得

る。図 8 (a) 及び図 8 (b) は、各々の例を提供する。

【 0 0 9 1 】

図 8 (a) では、13ビット量子化器 3 1 は、インパルス応答がゼロ遅延項を持たず、かつ伝達関数が $H(z) - 1$ であるフィルタ 3 3 を用いて、ノイズシェーピングされる。関数 H の最適化は、文献にて広く議論されてきた。 $H(z) = 1 - 0.886z^{-1} + 0.391z^{-2}$ なる $H(z)$ が 1 つの選択であるが、更に多くの『アグレッシブな』シェーパが、2ビット以上の知覚的改善をもたらすものとして知られている。出力 6' を作るためのサブユニット 3 0, 3 1, 3 2 及び 3 3 の動作もまた、広く議論されてきた。

【 0 0 9 2 】

標準的な実務では、フィルタ 3 3 の出力が入力信号から直接に引き算される。しかしながら、ここでは、ノイズシェーピングが損失のある処理であるので、24ビットデコーダがシェーパの効果を『元に戻す』ことができなければならない。図 7 (b) を参照すれば、合成器 2 4 は、いずれもエンコーダの分割器 5' からの出力である、『MSB』6' と『LSB』2 3 との双方を受け取る。もしノイズシェーピングがなければ、合成器は、MSB と LSB とを（適切に桁合わせしたうえで）加算することによって、信号 2 6 を再生することができたであろう。もしノイズシェーピングからの信号修飾が LSB の決定関数であるなら、合成器は、入力 2 6 を再生することもできる。もし H が、量子化された係数を持つ有限インパルス応答フィルタであるなら、当該修飾が決定的であるように準備することは、最も容易である。更に、このフィルタ 3 3 の出力は、入力と同じビット幅に、すなわち図示のように 17ビットに量子化 3 6 がなされるべきである。さもなければ、LSB 出力のビット幅は増加するであろう。更にまた、17ビットレベルでのディザリングがなされない量子化歪みが遺産リスナ及び 16ビットリスナの耳に入る信号に持ち込まれることを回避するため、17ビットへの量子化は、ディザリング 3 5 がなされるべきである。このディザは決定的でなければならず、またディザ生成器 3 5, 3 5 a は、エンコーダとデコーダとの間で同期していなければならない。

【 0 0 9 3 】

これらの条件が与えられるなら、図 8 (b) の合成器は、ユニット 3 3 a, 3 4 a, 3 5 a, 3 6 a にて、『LSB』信号 2 3 から、図 8 (a) の分割器にてユニット 3 3, 3 4, 3 5 及び 3 6 により作られたノイズシェーピング修飾 3 8 のレプリカ 3 8 a を作ることができる。加算器 3 2 a は、量子化器 3' により信号 3 7 から取り出された下位ビット 2 3 を加算し、加算器 3 0 a は、減算器 3 0 の効果を補償し、これにより信号 2 6 のレプリカ 2 6 a を作る。

【 0 0 9 4 】

図 7 (a) 及び図 7 (b) に戻って、信号 2 のビット数が 16ビットよりも少ない場合には、システムは、次のようにして改善され得る。すなわち、ノイズシェーピングされた分割器 5' は 16ビット入力 2 6 を受け取るように構成され、その結果、その 16ビットのうちのボトムビットは、減算器 1 3 により持ち込まれる符号が逆転される点を除いて、対応するボトムビットの圧縮された信号 7 のみを含み得る。図 8 (a) では、これらのビットもまた、分割器を通して伝播し、ノイズシェーピング修飾 3 8 が引き算されている点を除いて、信号 2 3 の中に現れる。ゆえに、その信号 3 8 の情報を持つデコーダは、これらのビットを推論し得る。したがって、これらのビットは、信号 7 及び信号 2 3 の両者において二重に、複合ワードへ効果的に提供される。このため、エンコーダは、信号 2 3 から冗長なビットを除去するように変形され得る。このとき、デコーダは、除去されたビットを回復する。信号 2 が 15ビット信号の場合には、『LSB』信号 2 3 から最下位の 1 ビットのみがエンコーダにより除去され、その 1 ビットは、

- 図 8 (b) 中の信号 3 8 a の最下位ビット出力と、
- 図 7 (b) 中の信号 7 の最下位ビットと

の排他的 OR として回復され得る。

【 0 0 9 5 】

この処理は、再帰的である。なぜなら、このようにしてある特定のサンプリング時に引

き出された、回復された分割器 L S B は、ノイズシェーピングフィルタ 3 3 a を通した伝播のゆえに、次のサンプリング時に信号 3 8 a に影響を及ぼすからである。したがって、ノイズシェーピングフィルタ 3 3 及び 3 3 a の中の各々の状態変数が同じ値に初期化されることを、保証することが必要である。ストリームの開始時に、エンコーダ及びデコーダの双方にて、これらの変数をゼロに設定するのが自然であろう。

【0096】

エンコードされた複合ワードのうちの下位部分の割り付けは、実装者の裁量による。例えば、シェーパからの L S B とバックされた修整信号とは、全体動作にいかなる効果をも及ぼすことなく、交換され得る。図 9 は、破線で囲まれた枠内に示されたように、出力複合ワードの上位部分 8 を直接に提供する 16 ビット信号 2 9 を供給する分割器を組み入れた、エンコーダの関連部分を示している。解析により、図 7 (a) 中の対応する要素 1 2 , 1 3 及び 5 ' を図 9 に置き換えれば、複合ワード 1 6 に何の変更もないことが判る。当業者ならば、ステップサイズが必ずしも正確な 2 のべき乗によらない量子化に、量子化 1 及び 3 1 が置き換えられ得ることをも実現するであろう。この場合には、信号のうちのいくつかはバイナリでなく n 値であって、最高効率のために、これらの信号は、複合ワードの中にエントロピ符号化がなされ得る。しかしながら、PCM 互換性のため、『MSB』信号 6 ' は、標準バイナリフォーマットにて整数として表現され、かつエントロピ符号化がなされないべきである。

【0097】

ある状況にて、20 ビットオーディオなら伝達可能であるが、24 ビットオーディオは伝達不可能であるという事態を考慮すれば、三重の互換性を提供するとの要求もあり得る。すなわち、24 ビットリストナのための損失のない拡張された帯域幅再生を提供するだけでなく、遺産リストナと、デコーダを持つ 16 ビットリストナと、デコーダを持つ 20 ビットリストナとの間のバランスした利点をも提供するとの要求である。これは、24 ビット複合ワードのうちの下位部分の更なる細分と、既に説明した原理の更なる応用とによって、達成され得る。

【0098】

本明細書における 16 ビット及び 24 ビットへの参照は、単に現在の実務で一般的なワード幅を反映したものであって、本発明は、これより長い又は短いワード幅の、異なるワード幅の値であっても、等しく好適に適用され得る。

【0099】

以上を要約すれば、様々なデコーディングオプションとともに PCM 互換性を持つストリームを提供するシステムを説明してきた。オリジナルの高サンプリングレート信号の損失のない再生を実現するためにはデコーダが必要であり、したがって遺産リストナに提供される信号は「損失のある」信号であると説明したが、損失のある信号への減縮は、時不变のフィルタリング動作と、サンプリングレート低減動作と、時不变のノイズフロアを課す再量子化動作とを用いるだけで、オーディオ愛好家仲間で「良性」であると称される態様で実行される。

【手続補正 2】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項 1】

第 1 のサンプリングレートで入力デジタルオーディオ信号 (2) を受け取り、かつ、それから、複数のサンプル (16) を有しあつ前記第 1 のサンプリングレートよりも低い第 2 のサンプリングレートを持つ PCM デジタルオーディオ出力を生成するように構成された、損失のないオーディオエンコーダであって、

前記複数のサンプル (16) の各々は、上位部分 (8) と下位部分 (17) とを持ち、

前記上位部分(8)と前記下位部分(17)とは、いずれも、第1のデコーダが損失なしで前記入力デジタルオーディオ信号(2)を再生できるようにする情報を含み、

標準PCMストリームとして解釈されたとき、前記上位部分(8)は、低減された帯域幅を持つ前記入力デジタルオーディオ信号のバージョンの第1の損失のある表現を提供し、

前記上位部分(8)は、前記第1の損失のある表現の帯域幅よりも大きい帯域幅を持つ前記入力デジタルオーディオ信号の第2の損失のある表現を第2のデコーダが再生できるようにする情報を含み、

前記入力デジタルオーディオ信号(2)は、高周波数出力(28)と低周波数出力(15)とを持つ損失のない帯域分割器(3)に結合され、前記高周波数出力(28)は、損失のある圧縮出力(7)と復元出力(27)とを持つ圧縮ユニット(21)に結合され、

前記上位部分(8)は、前記低周波数出力(15)及び前記損失のある圧縮出力(7)に応じて引き出され、前記下位部分(17)は、前記復元出力(27)に応じて引き出されるオーディオエンコーダ。

【請求項2】

請求項1記載の損失のないオーディオエンコーダにおいて、

前記第1の損失のある表現は、時不变のフィルタリングと、サンプリングレートの低減と、時不变のノイズフロアを課す再量子化とのうちの1つ以上のみにかけられた、前記入力デジタルオーディオ信号(2)の正確な表現であるオーディオエンコーダ。

【請求項3】

請求項1又は2に記載の損失のないオーディオエンコーダにおいて、

各上位部分は、16バイナリビットからなるオーディオエンコーダ。

【請求項4】

先行する請求項のいずれか1項に記載の損失のないオーディオエンコーダにおいて、

各下位部分は、8バイナリビットからなるオーディオエンコーダ。

【請求項5】

先行する請求項のいずれか1項に記載の損失のないオーディオエンコーダにおいて、

前記第2のサンプリングレートは、前記第1のサンプリングレートの半分であるオーディオエンコーダ。

【請求項6】

請求項1～5のいずれか1項に記載の損失のないオーディオエンコーダにおいて、

前記第2のサンプリングレートは、48kHzであるオーディオエンコーダ。

【請求項7】

請求項1～5のいずれか1項に記載の損失のないオーディオエンコーダにおいて、

前記第2のサンプリングレートは、44.1kHzであるオーディオエンコーダ。

【請求項8】

請求項1～7のいずれか1項に記載の損失のないオーディオエンコーダにおいて、

前記第2のデコーダは、前記第1のサンプリングレートに対応するナイキスト周波数に等しいオーディオ帯域幅を再生するオーディオエンコーダ。

【請求項9】

請求項1～7のいずれか1項に記載の損失のないオーディオエンコーダにおいて、

前記第2のデコーダは、前記第1のサンプリングレートに対応するナイキスト周波数の4分の3に等しい帯域幅を再生するオーディオエンコーダ。

【請求項10】

先行する請求項のいずれか1項に記載の損失のないオーディオエンコーダにおいて、

前記圧縮ユニット(21)は、前記損失のある圧縮出力(7)に結合された出力を持つ、損失のある圧縮ユニット(4)を備えたオーディオエンコーダ。

【請求項11】

請求項10記載の損失のないオーディオエンコーダにおいて、

前記圧縮ユニット(21)は、前記損失のある圧縮ユニット(4)に結合された入力と

、前記復元出力(27)に結合された出力とを持つ、損失のない圧縮ユニット(14)を更に備えたオーディオエンコーダ。

【請求項12】

先行する請求項のいずれか1項に記載の損失のないオーディオエンコーダにおいて、前記下位部分(17)は、前記帯域分割器(3)の低周波数出力(15)に応じて引き出されるオーディオエンコーダ。

【請求項13】

先行する請求項のいずれか1項に記載の損失のないオーディオエンコーダにおいて、前記損失のない帯域分割器(3)の低周波数出力(15)は、前記上位部分(8)に結合された第1の出力(6')と、前記下位部分(17)に結合された第2の出力(23)とを持つ分割器(5')に結合されているオーディオエンコーダ。

【請求項14】

請求項13記載の損失のないオーディオエンコーダにおいて、前記分割器(5')は、ノイズシェーピングフィルタを備えたオーディオエンコーダ。

【請求項15】

先行する請求項のいずれか1項に記載の損失のないオーディオエンコーダにおいて、前記上位部分(8)の中の複数のビットは、前記損失のない帯域分割器(3)の前記低周波数出力(15)に結合された第1の入力と、前記圧縮出力(7)に結合された第2の入力とを持つ減算器(13)の出力に応じて引き出されるオーディオエンコーダ。

【請求項16】

先行する請求項のいずれか1項に記載の損失のないオーディオエンコーダに結合されたノイズシェーパ(1)を備えた装置。

【請求項17】

ウォーターマーク出力を供給する損失なしの可逆的なウォーターマークエンコーダに結合された、請求項1～15のいずれか1項に記載の損失のないオーディオエンコーダを備えた装置であって、

前記装置は、コンフィギュレーションパラメータに応じてエンコーディングを実行し、かつ、前記ウォーターマークエンコーダは、デコーダによる使用のために前記ウォーターマーク出力の中に前記コンフィギュレーションパラメータを埋め込む装置。

【請求項18】

請求項17記載の装置において、

前記損失のないオーディオエンコーダの入力に量子化された信号を供給するノイズシェーパを更に備え、

前記ノイズシェーパは、あるビット深度まで量子化し、前記コンフィギュレーションパラメータは、前記ビット深度を含む装置。

【請求項19】

請求項18記載の装置において、

前記下位部分の情報保持容量を超えないようにオーディオ品質を最大化するために、前記量子化のビット深度を選択する選択ユニットを更に備えた装置。

【請求項20】

対応する請求項1記載のオーディオエンコーダにより第2のサンプリングレートで生成された複数の入力サンプル(16)を有するPCM入力デジタルオーディオ信号を受け取るよう構成されたオーディオデコーダであって、

前記オーディオデコーダは、前記PCM入力デジタルオーディオ信号から、前記第2のサンプリングレートよりも高い第1のサンプリングレートを持つ出力デジタルオーディオ信号(11)を生成するよう更に構成され、

前記出力デジタルオーディオ信号と比較信号との間の差違は、0～5kHzの周波数範囲で、静的な統計でスペクトル的にシェーピングされたノイズであって、前記比較信号は、前記入力デジタルオーディオ信号から、フィルタリングの動作及び前記第1のサンプリングレートへのリサンプリングの動作により生成され、

前記出力デジタルオーディオ信号と第2の出力信号との間の差違は、0～5kHzの周波数範囲で、静的な統計でスペクトル的にシェーピングされたノイズであって、前記第2の出力信号は、前記デコーダに送られるとき、各サンプルから下位部分を除去する以外は前記PCM入力デジタルオーディオ信号と一致する信号から作られ、

前記出力デジタルオーディオ信号(11)は、前記エンコーダに提供されたデジタルオーディオ入力信号(2)の正確なレプリカであるオーディオデコーダ。

【請求項21】

第2のサンプリングレートで複数の入力サンプル(16)を有するPCM入力デジタルオーディオ信号を受け取り、かつ、それから、前記第2のサンプリングレートよりも高い第1のサンプリングレートを持つ出力デジタルオーディオ信号(11)を生成するように構成されたオーディオデコーダであって、

前記デコーダは、

高周波数入力(28a)と低周波数入力(15a)とを持ち、前記出力デジタルオーディオ信号(11)を供給する、損失のない帯域合成器(10)と、

損失のある入力(7)と、復元入力(27)と、出力(28a)とを持ち、前記出力は前記損失のない帯域合成器(10)の前記高周波数入力に結合された伸張ユニット(22)とを備え、

各入力サンプル(16)は、上位部分(8)と下位部分(17)とを有し、

前記帯域合成器(10)の前記低周波数入力(15)は、前記上位部分(8)に応じて引き出され、

前記伸張ユニット(22,9)の前記損失のある入力(7)は、前記上位部分(8)に応じて、かつ前記下位部分(17)から独立して引き出され、

前記伸張ユニット(22,19)の前記復元入力(27)は、前記下位部分(17)に応じて、かつ前記上位部分(8)から独立して引き出されるオーディオデコーダ。

【請求項22】

請求項21記載のオーディオデコーダにおいて、

前記帯域合成器(10)の前記低周波数入力(15)は、前記上位部分(8)の中に含まれる全ビットに応じて引き出されるオーディオデコーダ。

【請求項23】

請求項21又は22に記載のオーディオデコーダにおいて、

前記帯域合成器(10)の前記低周波数入力(15)もまた、前記下位部分(17)に応じて引き出されるオーディオデコーダ。

【請求項24】

請求項21～23のいずれか1項に記載のオーディオデコーダにおいて、

前記出力デジタルオーディオ信号(11)と比較信号との間の差違は、0～5kHzの周波数範囲で、静的な統計でスペクトル的にシェーピングされたノイズであって、前記比較信号は、前記PCM入力デジタルオーディオ信号から、フィルタリングの動作及び前記第1のサンプリングレートへのリサンプリングの動作により生成されるオーディオデコーダ。

【請求項25】

請求項21～24のいずれか1項に記載のオーディオデコーダにおいて、

対応するオーディオエンコーダにより生成された信号を受け取るように構成され、

前記出力デジタルオーディオ信号(11)は、前記対応するオーディオエンコーダに提供されたデジタルオーディオ入力信号(2)の正確なレプリカであるオーディオデコーダ。