

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4639073号
(P4639073)

(45) 発行日 平成23年2月23日(2011.2.23)

(24) 登録日 平成22年12月3日(2010.12.3)

(51) Int.Cl.

F I

G 1 O L 19/02 (2006.01)

G 1 O L 19/02 1 4 2 B

G 1 O L 19/02 1 4 2 D

請求項の数 10 (全 29 頁)

(21) 出願番号	特願2004-335005 (P2004-335005)	(73) 特許権者	000001007
(22) 出願日	平成16年11月18日(2004.11.18)		キヤノン株式会社
(65) 公開番号	特開2006-145782 (P2006-145782A)		東京都大田区下丸子3丁目30番2号
(43) 公開日	平成18年6月8日(2006.6.8)	(74) 代理人	100076428
審査請求日	平成19年11月19日(2007.11.19)		弁理士 大塚 康德
		(74) 代理人	100112508
			弁理士 高柳 司郎
		(74) 代理人	100115071
			弁理士 大塚 康弘
		(74) 代理人	100116894
			弁理士 木村 秀二
		(74) 代理人	100130409
			弁理士 下山 治
		(74) 代理人	100134175
			弁理士 永川 行光

最終頁に続く

(54) 【発明の名称】 オーディオ信号符号化装置および方法

(57) 【特許請求の範囲】

【請求項1】

オーディオ入力信号をチャンネルごとに処理単位フレームに分割するフレーム分割部と、
オーディオ入力信号を分析し、変換ブロック長の決定および聴覚マスキングの計算を行う聴覚心理演算部と、

前記聴覚心理演算部で決定された前記変換ブロック長に従って、処理対象フレームをブロック化し、フレーム中の時間領域信号を1または2以上の周波数スペクトルの組に変換するフィルタバンク部と、

前記フィルタバンク部より出力された周波数スペクトルを、複数の周波数帯域に分割して、前記聴覚心理演算部の演算結果に基づいて、各周波数帯域のスペクトルを重み付けするためのスケールファクタを算出するスケールファクタ算出部と、

前記スケールファクタ算出部によって算出された前記スケールファクタで重み付けされた量子化前のスペクトル全体が持つ聴覚情報量から量子化後のスペクトル全体の情報量を減じ、量子化粗さの刻み幅から得られる係数を積算することによって、フレーム全体の量子化ステップをスペクトル量子化前に決定する量子化ステップ決定部と、

前記スケールファクタと前記量子化ステップとを利用して前記周波数スペクトルを量子化するスペクトル量子化部と、

前記スペクトル量子化部から出力される量子化スペクトルを規定のフォーマットに従って整形したビットストリームを作成出力するビット整形部と、

を備え、

10

20

前記量子化ステップ決定部は、符号化対象であるフレームに割り当てられたビット量に基づいて前記量子化スペクトル全体の情報量を予測する量子化スペクトル情報量予測部を含むことを特徴とするオーディオ信号符号化装置。

【請求項 2】

オーディオ入力信号をチャンネルごとに処理単位フレームに分割するフレーム分割部と、オーディオ入力信号を分析し、変換ブロック長の決定および聴覚マスキングの計算を行う聴覚心理演算部と、

前記聴覚心理演算部で決定された変換ブロック長に従って、処理対象フレームをブロック化し、フレーム中の時間領域信号を 1 または 2 以上の周波数スペクトルの組に変換するフィルタバンク部と、

前記フィルタバンク部より出力された周波数スペクトルを、複数の周波数帯域に分割して、前記聴覚心理演算部の演算結果に基づいて、各周波数帯域のスペクトルを重み付けするためのスケールファクタを算出するスケールファクタ算出部と、

符号化対象であるフレームに割り当てられたビット量に基づいて前記量子化スペクトル全体の情報量を量子化する前に予測する量子化スペクトル情報量予測部と、

前記スケールファクタ算出部によって算出された前記スケールファクタで重み付けされた量子化前のスペクトル全体が持つ聴覚情報量から前記量子化後のスペクトル全体の情報量を減じ、量子化粗さの刻み幅から得られる係数を積算することによって、フレーム全体の量子化ステップをスペクトル量子化前に決定する量子化ステップ決定部と、

前記スケールファクタと前記量子化ステップとを利用して前記周波数スペクトルを量子化するスペクトル量子化部と、

前記スペクトル量子化部から出力される量子化スペクトルを規定のフォーマットに従って整形したビットストリームを作成出力するビット整形部と、

を備え、

前記量子化スペクトル情報量予測部は、固定ビットレート符号化時に、入力信号の予測符号量が平均フレーム割当ビットに満たない場合は、聴覚エントロピーに基づいて量子化スペクトル情報量を予測することを特徴とするオーディオ信号符号化装置。

【請求項 3】

オーディオ入力信号をチャンネルごとに処理単位フレームに分割するフレーム分割部と、オーディオ入力信号を分析し、変換ブロック長の決定および聴覚マスキングの計算を行う聴覚心理演算部と、

前記聴覚心理演算部で決定された変換ブロック長に従って、処理対象フレームをブロック化し、フレーム中の時間領域信号を 1 または 2 以上の周波数スペクトルの組に変換するフィルタバンク部と、

前記フィルタバンク部より出力された周波数スペクトルを、複数の周波数帯域に分割して、前記聴覚心理演算部の演算結果に基づいて、各周波数帯域のスペクトルを重み付けするためのスケールファクタを算出するスケールファクタ算出部と、

符号化対象であるフレームに割り当てられたビット量に基づいて前記量子化スペクトル全体の情報量を量子化する前に予測する量子化スペクトル情報量予測部と、

前記スケールファクタ算出部によって算出された前記スケールファクタで重み付けされた量子化前のスペクトル全体が持つ聴覚情報量から前記量子化後のスペクトル全体の情報量を減じ、量子化粗さの刻み幅から得られる係数を積算することによって、フレーム全体の量子化ステップをスペクトル量子化前に決定する量子化ステップ決定部と、

前記スケールファクタと前記量子化ステップを利用して前記周波数スペクトルを量子化するスペクトル量子化部と、

前記スペクトル量子化部から出力される量子化スペクトルを規定のフォーマットに従って整形したビットストリームを作成出力するビット整形部と、

を備え、

前記スペクトル量子化部は、量子化スペクトルに使用される符号量が割り当てられた符号量を超えた場合に、前記量子化ステップを調整してスペクトルの再量子化を行うことを

10

20

30

40

50

特徴とするオーディオ信号符号化装置。

【請求項 4】

符号化形式がMPEG-1 Audio Layer IIIであることを特徴とする請求項 1 乃至 3 のいずれか 1 項に記載のオーディオ信号符号化装置。

【請求項 5】

符号化形式がMPEG-2/4 AACであることを特徴とする請求項 1 乃至 3 のいずれか 1 項に記載のオーディオ信号符号化装置。

【請求項 6】

オーディオ入力信号をチャンネルごとに処理単位フレームに分割するフレーム分割ステップと、

オーディオ入力信号を分析し、変換ブロック長の決定および聴覚マスキングの計算を行う聴覚心理演算ステップと、

前記聴覚心理演算ステップで決定された前記変換ブロック長に従って、処理対象フレームをブロック化し、フレーム中の時間領域信号を 1 または 2 以上の周波数スペクトルの組に変換するフィルタバンク処理ステップと、

前記フィルタバンク処理ステップで得られた周波数スペクトルを、複数の周波数帯域に分割して、前記聴覚心理演算ステップでの演算結果に基づいて、各周波数帯域のスペクトルを重み付けするためのスケールファクタを算出するスケールファクタ算出ステップと、

前記スケールファクタ算出ステップによって算出された前記スケールファクタで重み付けされた量子化前のスペクトル全体が持つ情報量から量子化後のスペクトル全体の情報量を減じ、量子化粗さの刻み幅から得られる係数を積算することによって、フレーム全体の量子化ステップをスペクトル量子化前に決定する量子化ステップ決定ステップと、

前記スケールファクタと前記量子化ステップとを利用して前記周波数スペクトルを量子化するスペクトル量子化ステップと、

前記スペクトル量子化ステップで得られた量子化スペクトルを規定のフォーマットに従って整形したビットストリームを作成出力するビット整形ステップと、

を有し、

前記量子化ステップ決定ステップは、符号化対象であるフレームに割り当てられた情報量に基づいて前記量子化スペクトル全体の情報量を予測する量子化スペクトル総量予測ステップを含むことを特徴とするオーディオ信号符号化方法。

【請求項 7】

オーディオ入力信号をチャンネルごとに処理単位フレームに分割するフレーム分割ステップと、

オーディオ入力信号を分析し、変換ブロック長の決定および聴覚マスキングの計算を行う聴覚心理演算ステップと、

前記聴覚心理演算ステップで決定された前記変換ブロック長に従って、処理対象フレームをブロック化し、フレーム中の時間領域信号を 1 または 2 以上の周波数スペクトルの組に変換するフィルタバンク処理ステップと、

前記フィルタバンク処理ステップで得られた周波数スペクトルを、複数の周波数帯域に分割して、前記聴覚心理演算ステップでの演算結果に基づいて、各周波数帯域のスペクトルを重み付けするためのスケールファクタを算出するスケールファクタ算出ステップと、

符号化対象であるフレームに割り当てられたビット量に基づいて前記量子化スペクトル全体の情報量を量子化する前に予測する量子化スペクトル情報量予測ステップと、

前記スケールファクタ算出ステップによって算出された前記スケールファクタで重み付けされた量子化前のスペクトル全体が持つ聴覚情報量から前記量子化後のスペクトル全体の情報量を減じ、量子化粗さの刻み幅から得られる係数を積算することによって、フレーム全体の量子化ステップをスペクトル量子化前に決定する量子化ステップ決定ステップと、

前記スケールファクタと前記量子化ステップとを利用して前記周波数スペクトルを量子化するスペクトル量子化ステップと、

10

20

30

40

50

前記スペクトル量子化で得られた量子化スペクトルを規定のフォーマットに従って整形したビットストリームを作成出力するビット整形ステップと、

を有し、

前記量子化スペクトル情報量予測ステップでは、固定ビットレート符号化時に、入力信号の予測符号量が平均フレーム割当ビットに満たない場合は、聴覚エントロピーに基づいて量子化スペクトル情報量を予測することを特徴とするオーディオ信号符号化方法。

【請求項 8】

オーディオ入力信号をチャンネルごとに処理単位フレームに分割するフレーム分割ステップと、

オーディオ入力信号を分析し、変換ブロック長の決定および聴覚マスキング計算を行う聴覚心理演算ステップと、

前記聴覚心理演算ステップで決定された前記変換ブロック長に従って、処理対象フレームをブロック化し、フレーム中の時間領域信号を 1 または 2 以上の周波数スペクトルの組に変換するフィルタバンク処理ステップと、

前記フィルタバンク処理ステップで得られた周波数スペクトルを、複数の周波数帯域に分割して、前記聴覚心理演算ステップでの演算結果に基づいて、各周波数帯域のスペクトルを重み付けするためのスケールファクタを算出するスケールファクタ算出ステップと、

符号化対象であるフレームに割り当てられたビット量に基づいて前記量子化スペクトル全体の情報量を量子化する前に予測する量子化スペクトル情報量予測ステップと、

前記スケールファクタ算出ステップによって算出された前記スケールファクタで重み付けされた量子化前のスペクトル全体が持つ聴覚情報量から前記量子化後のスペクトル全体の情報量を減じ、量子化粗さの刻み幅から得られる係数を積算することによって、フレーム全体の量子化ステップをスペクトル量子化前に決定する量子化ステップ決定ステップと、

前記スケールファクタと前記量子化ステップとを利用して前記周波数スペクトルを量子化するスペクトル量子化ステップと、

前記スペクトル量子化ステップで得られた量子化スペクトルを規定のフォーマットに従って整形したビットストリームを作成出力するビット整形ステップと、

を有し、

前記スペクトル量子化ステップでは、量子化スペクトルに使用される符号量が割り当てられた符号量を超えた場合に、前記量子化ステップを調整してスペクトルの再量子化を行うことを特徴とするオーディオ信号符号化方法。

【請求項 9】

請求項 6 乃至 8 のいずれか 1 項に記載のオーディオ信号符号化方法をコンピュータに実行させるためのプログラム。

【請求項 10】

請求項 9 に記載のプログラムを格納したコンピュータ読み取り可能な記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、オーディオ信号の符号化装置および方法に関する。

【背景技術】

【0002】

近年、高音質かつ高効率なオーディオ信号符号化技術が、DVD-Videoの音声トラック、半導体メモリやHDDなどを利用した携帯オーディオプレイヤー、インターネットを介した音楽配信、家庭内LANにおけるホームサーバへの楽曲蓄積などに広く利用され、幅広く普及するとともにその重要性も増している。

【0003】

このようなオーディオ信号符号化技術の多くは、変換符号化技術を利用して時間周波数変換を行なっている。例えば、MPEG-2 AACやDolby Digital(AC-3)などでは、MDCTなどの

10

20

30

40

50

直交変換単体でフィルタバンクを構成しており、MPEG-1 Audio Layer III(MP3)やATRAC(MDに利用されている符号化方式)では、QMFなどのサブバンド分割フィルタと直交変換を多段接続してフィルタバンクを構成している。

【 0 0 0 4 】

これらの変換符号化技術では、人間の聴覚特性を利用したマスキング分析を行うことによって、マスキングされると判断したスペクトル成分を取り除く、あるいはマスクされる量子化誤差を許容することにより、スペクトルを表現するための情報量を削減し、圧縮効率を高めている。

【 0 0 0 5 】

また、これらの変換符号化技術では、その多くが、スペクトル成分を非線形量子化することにより、スペクトルが持つ情報量を圧縮している。例えば、MP3やAACでは、各スペクトル成分を0.75乗することにより情報量を圧縮している。

10

【 0 0 0 6 】

また、これらの変換符号化技術では、フィルタバンクによって周波数成分に変換された入力信号を、人間の聴覚の周波数分解能に基づいて設定される分割周波数帯域ごとにまとめ、量子化時に各分割周波数帯域毎の正規化係数を聴覚分析の結果より決定し、正規化係数と量子化スペクトルの組み合わせで周波数成分を表現することで情報量を削減している。この正規化係数は、実際には分割帯域毎の量子化粗さの調整を行う変数であり、正規化係数が1変化することによって、量子化粗さは1ステップ分変化することになる。MPEG-2 AACでは、この分割周波数帯域をスケールファクタバンド(SFB)と呼び、正規化係数をスケールファクタと呼称する。

20

【 0 0 0 7 】

また、これらの変換符号化方式では、符号化単位である1フレーム全体の量子化粗さを制御することによって符号量を制御している。多くの変換符号化方式では、量子化粗さは、ある基数の整数乗幅でステップ状に制御されており、この整数を量子化ステップと呼ぶ。MPEGオーディオ規格では、この、フレーム全体の量子化粗さを設定する量子化ステップを「グローバルゲイン」もしくは「コモンスケールファクタ」と呼称している。また、前述のスケールファクタは量子化ステップとの相対値で表現することによって、これらの変数の符号に必要な情報量を削減している。

【 0 0 0 8 】

例えば、MP3やAACではこれらの変数が1変化することによって、実際の量子化粗さは2の3/16乗分変化する。

30

【 0 0 0 9 】

変換符号化方式の量子化処理では、スケールファクタを制御して聴覚演算の結果を反映して量子化誤差がマスクされるように量子化歪みを制御すると同時に、量子化ステップを制御してフレーム全体の量子化粗さを適宜調整することによってフレーム全体の符号量制御を行わなければならない。これらの量子化粗さを決める二種類の数値は、符号化品質に重大な影響を及ぼすため、慎重かつ正確に、この二つの制御を同時に効率よく行うことが求められる。

【 0 0 1 0 】

MPEG-1 Audio Layer III(MP3)の規格書(ISO/IEC 11172-3)やMPEG-2 AACの規格書(ISO/IEC 13818-7)では、量子化時にスケールファクタとグローバルゲインを適宜制御する方法として、歪み制御ループ(アウターループ)と符号量制御ループ(インナーループ)の二重ループによって繰り返し処理を行う方法が紹介されている。以下、この方法を図面を用いて説明する。なお、便宜上、MPEG-2 AACの場合を例にとって説明を行う。

40

【 0 0 1 1 】

図13は、ISO/IEC規格書に記載されている量子化処理を簡単なフローチャートにしたものである。

【 0 0 1 2 】

まず、ステップS501では全てのSFBのスケールファクタと、グローバルゲインが0

50

に初期化され、歪み制御ループ（アウターループ）に入る。

【 0 0 1 3 】

歪み制御ループでは、まず、符号量制御ループ（インナーループ）が実行される。

【 0 0 1 4 】

符号量制御ループでは、まず、ステップ S 5 0 2 において、1 フレーム分、すなわち、1 0 2 4 個のスペクトル成分が、下記の量子化式に従って量子化される。

【 0 0 1 5 】

【 数 1 】

$$X_q = \text{Int} \left[\left[|x_i| \cdot 2^{-\frac{1}{4} \cdot (\text{global_gain} - \text{scalefac})} \right]^{\frac{3}{4}} + 0.4054 \right] \quad (1)$$

10

【 0 0 1 6 】

ただし、(1) 式において X_q は量子化スペクトル、 x_i は量子化前のスペクトル (MDCT 係数)、 global_gain はグローバルゲイン、 scalefac はこのスペクトル成分が含まれる SFB のスケールファクタである。

【 0 0 1 7 】

次に、ステップ S 5 0 3 において、これらの量子化スペクトルをハフマン符号化した時の 1 フレーム分の使用ビット数が計算され、S 5 0 4 でフレームに割り当てられたビット数と比較する。割り当てられたビット数より使用ビット数が大きい場合は、S 5 0 5 においてグローバルゲインを 1 増加して、量子化粗さを粗くして、再び S 5 0 2 のスペクトル量子化に戻る。この繰り返しは量子化後に必要なビット数が割り当てられたビット数より少なくなるまで行われ、この時点でのグローバルゲインが決定されて、符号量制御ループが終了する。

20

【 0 0 1 8 】

ステップ S 5 0 6 では、符号量制御ループによって量子化されたスペクトルを逆量子化して、量子化前のスペクトルとの差分を取ることによって量子化誤差を算出する。この量子化誤差は、SFB 毎にまとめられる。

【 0 0 1 9 】

ステップ S 5 0 7 では、全ての SFB でスケールファクタが 0 より大きくなったか、もしくは、量子化誤差が許容誤差範囲内に納まっているかどうかを調べる。このいずれの条件も満たさない SFB がある場合は、ステップ S 5 0 8 に進み、許容誤差範囲内に量子化誤差が納まっていない SFB のスケールファクタを 1 増やし、再び歪み制御ループ処理を繰り返す。なお、SFB 毎の許容誤差は聴覚演算によって量子化処理の前に求められている。

30

【 0 0 2 0 】

以上説明したように、I S O 規格書に記載されている量子化処理方法は二重ループで構成されており、しかも、グローバルゲインとスケールファクタは 1 刻みの制御しか行われないために、この処理が収束するまでに、スペクトル量子化とビット計算が幾度となく延々と繰り返されることになる。

40

【 0 0 2 1 】

ここで、例えば MPEG-2 AAC の場合では、スペクトル量子化は 1 回処理するたびに式 (1) の計算を 1024 回行うことになるため、計算量が多い処理である。また、ビット計算時に検索されるハフマン符号表が 1 1 種も存在するため、ハフマン符号表を全探索するとビット計算もやはり計算量が必然的に多くなる。

【 0 0 2 2 】

さらに、歪み制御ループでは逆量子化後に量子化誤差の計算を行っているが、この処理にも計算量がかかってしまう。そのため、この二重ループが収束するまでには膨大な処理量がかかってしまう。

【 0 0 2 3 】

50

この問題を解決するために、二重ループの繰り返し回数を削減することによって、処理量を削減しようとする様々な試みがなされている。

【 0 0 2 4 】

例えば、特開2003-271199号公報（特許文献 1）には、ハフマン符号表の特性に応じて決定したステップ数によって、コモンスケールファクタやスケールファクタを 1 刻みではなく飛び飛びに制御することにより、二重ループそれぞれのループ回数を減らし、処理量を削減する方法が開示されている。

【 0 0 2 5 】

また、特開2001-184091号公報（特許文献 2）には、最初に量子化ステップの推定値を算出した後、スケールファクタをMNRに応じて計算後、通常のインナーループを実行する方法が開示されている。

10

【 0 0 2 6 】

また、A.D.Duenes、R.Perez、B.Rivasらの論文“A robust and efficient implementation of MPEG-2/4 AAC Natural Audio Coders”（AES 112th Convention Paper, 2002）（非特許文献 1）には、式（ 1 ）を変形した式と、聴覚分析によって求められるSFB毎の許容誤差エネルギーを用いることによって、スケールファクタをスペクトル量子化に先行して適宜計算することにより、二重ループの外側の歪み制御ループを取り除き、処理量を削減する方法が紹介されている。

【 0 0 2 7 】

これらの従来技術を用いることによって、量子化処理の二重ループの収束を早め、量子化処理の処理量のある程度までは削減することができる。

20

【 0 0 2 8 】

【特許文献 1】特開2003-271199号公報

【特許文献 2】特開2001-184091号公報

【非特許文献 1】A.D.Duenes、R.Perez、B.Rivas 等, “A robust and efficient implementation of MPEG-2/4 AAC Natural Audio Coders”, AES 112th Convention Paper (2002)

【発明の開示】

【発明が解決しようとする課題】

【 0 0 2 9 】

30

しかしながら、従来の技術では、ISO規格書に記載されている二重ループを完全に繰り返さないようにすることはできないため、依然としてスペクトル量子化を数回から数十回繰り返さないと量子化処理を終えることができず、符号化処理全体に占める量子化処理の処理量は依然として大きかった。

【 0 0 3 0 】

特に、二重ループのうち、聴覚演算結果を利用して事前にスケールファクタを計算することによって、外側の歪み制御ループを無くすことは可能であるが、量子化ステップを量子化前に計算することは従来の技術では不可能であった。

【 0 0 3 1 】

そのため、従来の技術では符号量制御ループにおけるスペクトル量子化とビット計算をやはり繰り返し行っており、処理量を浪費しているという課題がある。

40

【 0 0 3 2 】

本発明は、上記問題点に鑑みてなされたものであり、オーディオ信号の符号化における量子化処理の処理量を大幅に削減することを目的とする。

【課題を解決するための手段】

【 0 0 3 3 】

本発明は、基本的には量子化前の情報量を量子化後の情報量で割ることによって、全体の量子化粗さを求めることができるという考えに基づき、量子化ステップを実際の量子化前に求めようとするものである。ここで、量子化粗さは一般的に基数を量子化ステップ乗したものであるため、量子化ステップを求めるために底をこの基数にした対数をとると、

50

情報量の除算は情報量の差分に変化する。この差分に、量子化の刻み幅によって決定される係数を積算すると正確な量子化ステップを求めることができる。さらに、実際の量子化後の情報量は量子化後でないと求めることができないが、フレームに割り当てられた符号量から予測することができるため、この予測を利用して量子化前に正確な量子化ステップを求めるものである。

【 0 0 3 4 】

例えば本発明の一側面に係るオーディオ信号符号化装置は、オーディオ入力信号をチャネルごとに処理単位フレームに分割するフレーム分割部と、オーディオ入力信号を分析し、変換ブロック長の決定および聴覚マスキングの計算を行う聴覚心理演算部と、前記聴覚心理演算部で決定された前記変換ブロック長に従って、処理対象フレームをブロック化し、フレーム中の時間領域信号を1または2以上の周波数スペクトルの組に変換するフィルタバンク部と、前記フィルタバンク部より出力された周波数スペクトルを、複数の周波数帯域に分割して、前記聴覚心理演算部の演算結果に基づいて、各周波数帯域のスペクトルを重み付けするためのスケールファクタを算出するスケールファクタ算出部と、前記スケールファクタ算出部によって算出された前記スケールファクタで重み付けされた量子化前のスペクトル全体が持つ聴覚情報量から量子化後のスペクトル全体の情報量を減じ、量子化粗さの刻み幅から得られる係数を積算することによって、フレーム全体の量子化ステップをスペクトル量子化前に決定する量子化ステップ決定部と、前記スケールファクタと前記量子化ステップとを利用して前記周波数スペクトル列を量子化するスペクトル量子化部と、前記スペクトル量子化部から出力される量子化スペクトルを規定のフォーマットに従って整形したビットストリームを作成出力するビット整形部とを備え、前記量子化ステップ決定部は、符号化対象であるフレームに割り当てられたビット量に基づいて前記量子化スペクトル全体の情報量を予測する量子化スペクトル情報量予測部を含むことを特徴とする。

【 0 0 3 5 】

本発明においては、最初にスケールファクタを算出、確定した後に、その値を使用した計算で量子化ステップをほぼ正確に算出することができるので、ほぼ一回のスペクトル量子化とビット計算で量子化を終了することが可能になる。

【発明の効果】

【 0 0 3 6 】

本発明によれば、オーディオ信号の符号化における量子化処理の処理量を大幅に低減することができる。

【発明を実施するための最良の形態】

【 0 0 3 7 】

以下、図面を参照して本発明の好適な実施形態について詳細に説明する。

【 0 0 3 8 】

(第1の実施形態)

図1は、本実施形態におけるオーディオ信号符号化装置の一構成例を示す図である。同図において、太線はデータ信号、細線は制御信号を示す。

【 0 0 3 9 】

図示の構成において、1はオーディオ入力信号を処理単位であるフレームに分割するフレーム分割器である。ここでフレーム単位に分割されたオーディオ入力信号は後述する聴覚心理演算器2とフィルタバンク3へ送出される。

【 0 0 4 0 】

2は聴覚心理演算器であり、オーディオ入力信号をフレーム単位に分析し、SFBよりも詳細な分割周波数帯域でマスキング計算を行う。この演算の結果、ブロックタイプがフィルタバンク3へ、また、SFB毎の信号対マスク比(SMR)がスケールファクタ計算器4へ出力される。

【 0 0 4 1 】

3はフィルタバンクであり、フレーム分割器1から入力された時間信号に対して、聴覚

心理演算器 2 によって指定されたブロックタイプのウィンドウ掛けを行った後、指定されたブロック長で時間周波数変換を行い、周波数スペクトルに変換する。

【 0 0 4 2 】

4 はスケールファクタ計算器であり、SFB毎のSMR（信号対マスク比）と、周波数スペクトルより、SFB毎の許容誤差エネルギーを算出し、それに基づいて全てのSFBのスケールファクタを決定する。

【 0 0 4 3 】

5 はスペクトル割当ビット計算器であり、量子化スペクトル符号に割り当てられるビット数を計算する。

【 0 0 4 4 】

6 は量子化スペクトル総量予測器であり、スペクトル割当ビット数に基づいて、量子化後のスペクトル総量を予測する。

【 0 0 4 5 】

7 は量子化ステップ計算器であり、量子化前のスペクトルが持つ聴覚情報量を計算し、量子化後のスペクトル総量から求めた量子化後のスペクトル情報量を減じることによって量子化ステップを求める。

【 0 0 4 6 】

8 はスペクトル量子化器であり、各周波数スペクトルを量子化する。

【 0 0 4 7 】

9 はビット整形器であり、スケールファクタと量子化スペクトルを適宜規定のフォーマットに整形してビットストリームを作成し、出力する。

【 0 0 4 8 】

上記構成によるオーディオ信号符号化装置におけるオーディオ信号の処理動作を以下に説明する。

【 0 0 4 9 】

なお、本実施形態では、説明の便宜のために符号化方式としてMPEG-2 AACを例にとって説明するが、同様な量子化手法を適用可能なその他の符号化方式についても全く同様な方法で実現可能である。

【 0 0 5 0 】

まず、処理に先立ち、各部の初期化を行う。初期化によって、量子化ステップと全てのスケールファクタの値は 0 にセットされる。

【 0 0 5 1 】

オーディオPCM信号などのオーディオ入力信号はフレーム分割器 1 によってフレーム単位に分割され、聴覚心理演算器 2 とフィルタバンク 3 に送出される。MPEG-2 AAC LC(Low-Complexity)プロファイルの場合、1 フレームは1024サンプルのPCM信号で構成され、この信号が送出される。

【 0 0 5 2 】

聴覚心理演算器 2 では、フレーム分割器 1 から送出された入力信号を適宜分析し、聴覚マスキング分析を行い、ブロックタイプをフィルタバンク 3 に、また、SFB毎の信号対マスク比(SMR)をスケールファクタ計算器 4 に各々出力する。なお、聴覚心理演算器 2 で行われる分析やマスキング計算は当分野において公知であるため、これらの詳細な説明は行わない。

【 0 0 5 3 】

フィルタバンク 3 では、聴覚心理演算器 2 が出力するブロックタイプに従って、フレーム分割器 1 から送出される 1 フレーム分の現入力信号と、前回の変換時に受け取った前フレームの入力信号を合わせ、2フレーム分、2048サンプルの時間信号を周波数成分へ変換する。なお、本実施形態において、前フレームの入力信号はフィルタバンク 3 内のバッファに保持されている。ここで、ブロックタイプが長いブロック長を用いる場合は、入力信号の2048サンプルを一つのブロックとして、ブロックタイプに応じた形の窓掛けを実行後、MDCTを行い、1024個の周波数スペクトルを出力する。短いブロック長を用いる場合には

10

20

30

40

50

、入力信号の2048サンプルのうち、448番目のサンプルを先頭として256サンプルを一つのブロックとして窓掛け後にMDCTを行い、128個の周波数成分を出力する変換を、入力信号を128サンプルずつずらしながら都合8回を行い、8組の周波数スペクトルを得る。

【0054】

スケールファクタ計算器4は、フィルタバンク3から出力されるスペクトル成分と聴覚心理演算器2から出力されるSFB毎のSMR値からSFB毎の許容誤差エネルギーを算出し、これに基づき、SFB毎のスケールファクタを計算する。許容誤差エネルギーに基づくスケールファクタの算出方法は、当分野では公知であるため、ここでは詳細は述べないが、例えば、前述した非特許文献1に記載された手法をとるならば、MPEG-2 AACにおいて、SFB bにおけるスケールファクタscalefac[b]は次式で求めることができる。

10

【0055】

【数2】

$$scalefac[b] = \text{Int} \left[-\frac{16}{3} \cdot \left[\frac{1}{2} \cdot \log_2 xmin[b] + \log_2 \left(\frac{3}{4} \right) - \frac{1}{4} \cdot \log_2 x_{avg} \right] \right] \quad (2)$$

【0056】

ただし、式(2)において、 x_{avg} はSFB bに含まれるスペクトル成分の平均レベルである。また、xmin[b]はSFB bの許容誤差エネルギーであり、SFB bのスペクトルエネルギーをenergy[b]、信号対マスク比をSMR[b]、含まれるスペクトル本数をsfb_width[b]とすると、このxmin[b]は次式によって求められる。

20

【0057】

【数3】

$$xmin[b] = \frac{energy[b]}{SMR[b]} / sfb_width[b] \quad (3)$$

【0058】

スペクトル割当ビット計算器5では、スケールファクタ計算器4から出力されたスケールファクタをハフマン符号化したときのビット数を計算し、指定されたフレームビット数から減じることによって、量子化スペクトルに割り当てられるビット数を計算し、量子化スペクトル総量予測器6へ出力する。

30

【0059】

量子化スペクトル総量予測器6は、スペクトル割当ビット計算器5から出力されたビット数に基づいて量子化スペクトル総量の予測計算を行う。本実施形態において、この計算は、従来の量子化器によって量子化した際の、スペクトル割当ビット数と量子化スペクトル総量との関係を実際に測定し、その結果に基づいて作成した近似式によって計算する。例えば、この近似式をF(x)として、スペクトル割当ビットをspectrum_bitsとすると、量子化後スペクトル予測総量は以下の式によって求めることができる。

【0060】

40

【数4】

$$\sum_i X_q \approx F(spectrum_bits) \quad (4)$$

【0061】

量子化ステップ計算器7では、まず、フィルタバンク3から出力された各周波数スペクトルに、スケールファクタによって聴覚上の重み付けをした値の総計を取り、これに基づいて量子化前の周波数スペクトルが持つ聴覚情報量を計算する。

次に、量子化スペクトル総量予測器6から出力された量子化スペクトル総量に基づいて量子化スペクトルが持つ情報量を計算する。

50

【 0 0 6 2 】

最後に、量子化前スペクトルの聴覚情報量から量子化スペクトルの情報量を減じ、量子化粗さの刻み幅から得られる係数を掛けることによって、フレーム全体の量子化粗さである量子化ステップを計算する。

【 0 0 6 3 】

具体的には、MPEG-2 AACの場合は、量子化ステップの予測値は次式を計算することによって得られる。

【 0 0 6 4 】

【 数 5 】

$$global_gain = Int \left[\frac{16}{3} \cdot \left[\log_2 \sum_i \left[|x_i|^{\frac{3}{4}} \cdot 2^{\frac{3}{16} \cdot scalefac} \right] - \log_2 \sum_i X_q \right] \right] \quad (5)$$

10

【 0 0 6 5 】

ただし、式(5)において、 X_q は量子化スペクトル、 x_i は量子化前のスペクトル、 $global_gain$ はグローバルゲイン(量子化ステップ)、 $scalefac$ はこのスペクトル成分が含まれるSFBのスケールファクタである。また、総計を取る i の範囲は1フレーム分、すなわち0 i 1023である。

ここで、式(5)において、以下に示す右辺の第1項

$$\log_2 \sum_i \left[|x_i|^{\frac{3}{4}} \cdot 2^{\frac{3}{16} \cdot scalefac} \right]$$

20

が、量子化前のスペクトル全体が持つ聴覚情報量であり、各スペクトルに、スケールファクタによって聴覚上の重み付けがなされた値の総計である。また、右辺の第2項である $\log_2 \sum_i X_q$ が、量子化後のスペクトルが持つ情報量であり、このうち、 $\sum_i X_q$ は量子化スペクトルの総計であり、量子化スペクトル総量予測器6によって予測された値である。この値は前述したように例えば近似式(4)を計算することによって得られる。

【 0 0 6 6 】

30

なお、式(5)はスペクトル量子化式(1)を適宜変形することによって得ることができる。

【 0 0 6 7 】

スペクトル量子化器8は、スケールファクタ計算器4が出力したスケールファクタと量子化ステップ計算器7が出力した量子化ステップに従って、1024本の周波数スペクトルを量子化する。具体的には、例えば、MPEG-2 AACの場合では式(1)によって量子化スペクトルを算出し、フレーム全体で消費されるビット数をカウントする。

【 0 0 6 8 】

ここで、使用ビット数がスペクトル割当ビット数を超過してしまった場合には、使用ビット数がスペクトル割当ビット数に収まるまで量子化ステップを増加して再度スペクトル量子化を行う。しかしながら、量子化ステップ計算器7の計算が正確であるため、多くの場合1回のみの量子化スペクトル計算とビット計算が行われる。

40

【 0 0 6 9 】

各SFBのスケールファクタと量子化スペクトルはビット整形器9によって定められた書式に従ってビットストリームに整形されて、出力される。

【 0 0 7 0 】

以上説明したように、本実施形態におけるオーディオ信号符号化装置は、フレームに割り当てられたビット量から、量子化後のスペクトル総量を予測し、これを用いて量子化前後のスペクトル全体が持つ情報量の差分を計算することによって、スペクトル量子化の前に量子化ステップをほぼ正確に予測する。これにより、量子化ステップの調整のための繰

50

り返しを行うことが減るために、迅速に量子化処理を終了することができる。

【 0 0 7 1 】

(第 2 の実施形態)

本発明は、パーソナルコンピュータ (P C) 等の汎用的な計算機上で動作するソフトウェアプログラムとして実施することも可能である。以下、この場合について図面を用いて説明する。

【 0 0 7 2 】

図 5 は、本実施形態におけるオーディオ信号符号化装置の構成例を示す図である。

【 0 0 7 3 】

図示の構成において、100はCPUであり、オーディオ信号符号化処理のための演算、論理判断等を行い、102のバスを介して各構成要素を制御する。 10

【 0 0 7 4 】

101はメモリであり、本実施形態の構成例における基本I/Oプログラムや、実行しているプログラムコード、プログラム処理時に必要なデータなどを格納する。

【 0 0 7 5 】

102はバスであり、CPU100の制御の対象とする構成要素を指示するアドレス信号を転送し、CPU100の制御の対象とする各構成要素のコントロール信号を転送し、各構成機器相互間のデータ転送を行う。

【 0 0 7 6 】

103は端末であり、装置の起動、各種条件や入力信号の設定、符号化開始の指示を行う。 20

【 0 0 7 7 】

104はデータやプログラム等を記憶するための外部記憶領域を提供する外部記憶装置であり、例えばハードディスク装置などによって実現される。ここに、OSをはじめとするプログラムやデータ等が保管され、また、保管されたデータやプログラムは必要な時にCPU100によって呼び出される。また、後述するように、オーディオ信号符号化処理プログラムもこの外部記憶装置104にインストールされることになる。

【 0 0 7 8 】

105はメディアドライブである。記録媒体 (例えば、C D - R O M) に記録されているプログラムやデータ、デジタルオーディオ信号などはこのメディアドライブ105が読み取ることにより本オーディオ信号符号化装置にロードされる。また、外部記憶装置104に蓄えられた各種データや実行プログラムを、記録媒体に書き込むこともできる。 30

【 0 0 7 9 】

106はマイクロフォンであり、実際の音を集音してオーディオ信号に変換する。107はスピーカーであり、任意のオーディオ信号データを実際の音にして出力することができる。

【 0 0 8 0 】

108は通信網であり、LAN、公衆回線、無線回線、放送電波などで構成されている。

109は通信インタフェースであり、通信網108に接続されている。本実施形態におけるオーディオ信号符号化装置はこの通信インタフェース109を介して通信網108を経由し、外部機器と通信を行い、データやプログラムを送受信することができる。 40

【 0 0 8 1 】

かかる構成を備えるオーディオ信号符号化装置は、端末103からの各種の入力に応じて作動する。端末103からの入力 that 供給されると、インタラプト信号がCPU100に送られることによって、CPU100がメモリ101内に記憶してある各種の制御信号を読み出し、それらの制御信号に従って、各種の制御が行われる。

【 0 0 8 2 】

本実施形態のオーディオ信号符号化装置は、CPU100が、メモリ101に格納されている基本I/Oプログラムを実行し、これより外部記憶装置104に記憶されているOSをメモリ101にロードしてこれを実行することによって、動作する。具体的には、本 50

装置の電源がONにされると、基本I/Oプログラム中のIPL（イニシャルプログラムローディング）機能により外部記憶装置104からOSがメモリ101に読み込まれ、OSの動作が開始される。

【0083】

オーディオ信号符号化処理プログラムは、図2に示されるオーディオ信号符号化処理手順のフローチャートに基づいてプログラムコード化されたものである。

【0084】

図6は、オーディオ信号符号化処理プログラムおよび関連データを記録媒体に記録したときの内容構成例を示す図である。本実施形態において、オーディオ信号符号化処理プログラムおよびその関連データは記録媒体に記録されている。図示したように記録媒体の先頭領域には、この記録媒体のディレクトリ情報が記録されており、その後この記録媒体のコンテンツであるオーディオ信号符号化処理プログラムと、オーディオ信号符号化処理関連データがファイルとして記録されている。

【0085】

図7は、オーディオ信号符号化処理プログラムのオーディオ信号符号化装置（PC）への導入を示す模式図である。記録媒体に記録されたオーディオ信号符号化処理プログラムおよびその関連データは、図7に示したようにメディアドライブ105を通じて本装置にロードすることができる。この記録媒体110をメディアドライブ105にセットすると、OS及び基本I/Oプログラムの制御のもとにオーディオ信号符号化処理プログラムおよびその関連データが記録媒体110から読み出され、外部記憶装置104に格納される。その後、再起動時にこれらの情報がメモリ101にロードされて動作可能となる。

【0086】

図8は、本実施形態におけるオーディオ信号符号化処理プログラムがメモリ101にロードされ実行可能となった状態のメモリマップを示す図である。図示のように、メモリ101のワークエリアには例えば、基準ビットレート、基準サンプリングレート、ビットレート、サンプリングレート、割当ビット上限値、平均割当ビット、PEビット、使用ビット、スケールファクタビット、スペクトル割当ビット、量子化前スペクトル聴覚情報量、量子化後スペクトル予測情報量、許容誤差エネルギー、スペクトルバッファ、量子化スペクトル、入力信号バッファ、スケールファクタ、量子化ステップ、ブロックタイプ、SMR、PE、リザーブビット量が格納されている。

【0087】

図9は、本実施形態におけるオーディオ信号符号化装置における入力信号バッファの構成例を示す図である。図示の構成において、バッファサイズは1024×3サンプルであり、説明の便宜上1024サンプル毎に縦線で区切っている。入力信号は右から1フレーム分の1024サンプルずつ入力されて、左から逐次処理される。なお、図示の構成は1チャンネル分の入力信号バッファを模式的に示したものであり、本実施形態では入力信号のチャンネル分だけ同様なバッファが用意される。

【0088】

以下、本実施形態においてCPU100で実行されるオーディオ信号符号化処理をフローチャートを用いて説明する。

【0089】

図2は、本実施形態におけるオーディオ信号符号化処理のフローチャートである。このフローチャートに対応するプログラムはオーディオ信号符号化処理プログラムに含まれ、上記のとおりメモリ101にロードされCPU100によって実行される。

【0090】

まず、ステップS1は、符号化する入力オーディオ信号をユーザが端末103を用いて指定する処理である。本実施形態において、符号化するオーディオ信号は、外部記憶装置104に格納されているオーディオPCMファイルでも良いし、マイク106で捉えたりリアルタイムの音声信号をアナログ・デジタル変換した信号でも良い。この処理を終えると、ステップS2へ進む。

【 0 0 9 1 】

ステップ S 2 は、符号化する入力オーディオ信号が終了したかどうかを判定する処理である。入力信号が終了している場合は、ステップ S 1 1 へ処理が進む。未終了の場合は、ステップ S 3 へ処理が進む。

【 0 0 9 2 】

ステップ S 3 は、図 9 に示した入力信号バッファにおいて、右から 2 フレーム分、すなわち 2048 サンプルの時間信号を 1 フレーム分左にシフトするとともに、新たに 1 フレーム分、すなわち 1024 サンプルを右側に読み込む入力信号シフト処理である。この処理は入力信号に含まれる全てのチャンネルに対して行われる。処理を終えると、ステップ S 4 へ処理が進む。

10

【 0 0 9 3 】

ステップ S 4 は、入力信号バッファに格納されている時間信号を分析し、現行フレームの聴覚心理演算を行う処理である。この演算の結果、現行フレームのブロックタイプと、聴覚エントロピー (PE) と、SFB ごとの SMR 値が算出され、メモリ 1 0 1 上のワークエリアに格納される。ここで、SMR 値は、現行フレームのブロック長が短い場合はショートブロック時の 8 組が、若しくは、ブロックタイプがそれ以外の場合はロングブロック時の 1 組が算出される。このような聴覚演算は当分野において公知であるため、詳細な説明は行わない。処理を終えると、ステップ S 5 へ処理が進む。

【 0 0 9 4 】

ステップ S 5 では、ステップ S 4 で得られたブロックタイプに従って、現行フレームの時間信号、すなわち、図 1 0 の現行フレーム先頭ポイントから右に 2048 サンプル (2 フレーム分) の信号に対して窓掛け後、時間周波数変換を行う。この結果、MPEG- 2 AAC の場合、変換ブロック長が短いときは、128 の周波数成分に分割されたスペクトルの組が 8 組得られる。それ以外のブロック長が長いブロックタイプの場合は、1024 の周波数成分に分割されたスペクトルの組が 1 組得られる。双方の場合とも、算出された計 1024 本のスペクトルは、メモリ 1 0 1 上のワークエリアにあるスペクトルバッファに格納される。処理を終えると、ステップ S 6 に処理が進む。

20

【 0 0 9 5 】

ステップ S 6 は、ステップ S 5 で得られた周波数スペクトルとステップ S 4 で得られた SFB 毎の SMR から許容誤差エネルギーを計算したのち、これを利用して各 SFB 毎のスケールファクタを算出する処理である。例えば、MPEG- 2 AAC の場合は前述の実施形態 1 の式 (2) によってスケールファクタを算出する。この処理において算出した SFB 毎の許容誤差エネルギーとスケールファクタはメモリ 1 0 1 上のワークエリア上に格納される。処理を終えると、ステップ S 7 へ進む。

30

【 0 0 9 6 】

ステップ S 7 は、量子化前のスペクトルが持つ聴覚情報量と量子化後のスペクトルが持つ情報量との差分から量子化ステップを計算する処理である。この処理の詳細は図 3 を用いて後述する。処理を終えると、ステップ S 8 へ進む。

【 0 0 9 7 】

ステップ S 8 は、ステップ S 6 で求めたスケールファクタとステップ S 7 で求めた量子化ステップに従って、1024 本の周波数スペクトルを量子化して、使用ビットを計算し、使用ビットがメモリ 1 0 1 上のワークエリアに格納されている割当ビットを超えた場合のみ、量子化ステップの増加と再量子化を行う処理である。この処理の詳細は図 4 を用いて後述する。処理を終えると、ステップ S 9 へ処理が進む。

40

【 0 0 9 8 】

ステップ S 9 は、ステップ S 8 で算出された量子化スペクトルと、スケールファクタを、符号化方式によって定められたフォーマットに従って整形し、ビットストリームとして出力する処理である。本実施形態において、この処理によって出力されるビットストリームは、外部記憶装置 1 0 4 に格納されても良いし、あるいは、通信インタフェース 1 0 9 を介して通信網 1 0 8 に繋がっている外部機器に出力されても良い。処理を終えると、ス

50

テップ S 1 0 へ処理が進む。

【 0 0 9 9 】

ステップ S 1 0 は、ステップ S 9 で出力されたビットストリームに使用されたビット量と符号化ビットレートから、メモリ 1 0 1 上に格納されている蓄積ビット数の補正を行う処理である。処理を終えると、処理はステップ S 2 へと戻る。

【 0 1 0 0 】

ステップ S 1 1 は、聴覚心理演算や直交変換などで生じる遅延によってまだ出力されていない量子化スペクトルがメモリ上に残っているため、それらをビットストリームに整形して出力する処理である。処理を終えると、オーディオ信号符号化処理を終了する。

【 0 1 0 1 】

図 3 は、上記したステップ S 7 の量子化ステップ予測処理の詳細を示すフローチャートである。

【 0 1 0 2 】

まず、ステップ S 1 0 1 は、メモリ 1 0 1 上のワークエリアに保存されているスケールファクタを符号化形式によって定められているフォーマットに従って符号化したときに使用するビット数を算出する処理である。算出されたビット数は、メモリ 1 0 1 上のワークエリアに保存される。処理を終えると、ステップ S 1 0 2 へ進む。

【 0 1 0 3 】

ステップ S 1 0 2 は、フレームに割り当てられたビット数からメモリ 1 0 1 上に格納されたスケールファクタビット数を引いて、スペクトル符号に割り当てられるビット数を算出する処理である。算出されたスペクトル割当ビット数はメモリ 1 0 1 上のワークエリアに保存される。処理を終えると、ステップ S 1 0 3 へ進む。

【 0 1 0 4 】

ステップ S 1 0 3 は、メモリ 1 0 1 上のスペクトル割当ビット数を用いて、量子化スペクトル総量の予測計算を行う処理である。この予測計算は、予め実験を実施することによって求めた近似式によって行う。例えば、この近似式を $F(x)$ として、スペクトル割当ビットを $spectrum_bits$ とすると、量子化後スペクトル予測総量は以下の式によって求めることができる。

【 0 1 0 5 】

【数 6】

$$\sum_i X_q \approx F(spectrum_bits) \quad (4)$$

【 0 1 0 6 】

算出された量子化スペクトル予測総量はメモリ 1 0 1 上のワークエリアに格納される。処理を終えると、ステップ S 1 0 4 へ進む。

【 0 1 0 7 】

ステップ S 1 0 4 は、量子化前のスペクトルが持つ聴覚情報量を算出する処理である。量子化前のスペクトルの聴覚情報量は、各スペクトル成分に、そのスペクトル成分が含まれる SFB のスケールファクタによる量子化粗さの減少分を積算し、1 フレーム分の総量を求め、その対数を算出することによって求められる。例えば、MPEG-2 AAC の場合、量子化前のスペクトルが持つ聴覚情報量は以下の式を計算することによって求めることができる。

【 0 1 0 8 】

【数 7】

$$\log_2 \sum_i \left[|x_i|^{\frac{3}{4}} \cdot 2^{\frac{3}{16} \cdot scalefac} \right] \quad (6)$$

【 0 1 0 9 】

算出された量子化前スペクトルの聴覚情報量はメモリ 1 0 1 上のワークエリアに保存さ

10

20

30

40

50

れる。処理を終えると、ステップ S 1 0 5 へ進む。

【 0 1 1 0 】

ステップ S 1 0 5 は、ステップ S 1 0 3 で求めた量子化スペクトルの予測総量の対数を計算し、量子化スペクトルの予測情報量を算出する処理である。例えば、MPEG-2 AAC の場合は以下の式を計算することによって算出することができる。

【 0 1 1 1 】

【 数 8 】

$$\log_2 \sum_i X_q \quad (7)$$

10

【 0 1 1 2 】

すなわち、ステップ S 1 0 3 で得られた量子化スペクトル総量の対数を計算することによって、量子化スペクトル予測情報量が得られる。この処理によって算出された量子化後のスペクトル情報量はメモリ 1 0 1 上のワークエリアに保存される。処理を終えると、ステップ S 1 0 6 へ進む。

【 0 1 1 3 】

ステップ S 1 0 6 では、ステップ S 1 0 4 で求めた量子化前スペクトルの聴覚情報量から、ステップ S 1 0 5 で求めた量子化スペクトル予測情報量を引き、その結果にステップ S 1 0 7 で、量子化粗さの刻み幅によって決定される係数をかけ、グローバルゲイン、すなわち量子化ステップの予測値が算出される。MPEG-2 AAC の場合は、この予測値は結局実施形態 1 と同じく式 (5) を計算したことになる。

20

【 0 1 1 4 】

【 数 9 】

$$global_gain = Int \left[\frac{16}{3} \cdot \left[\log_2 \sum_i \left[|x_i|^{\frac{3}{4}} \cdot 2^{\frac{3}{16} \cdot scalefac} \right] - \log_2 \sum_i X_q \right] \right] \quad (5)$$

【 0 1 1 5 】

算出された量子化ステップ予測値は、メモリ 1 0 1 上のワークエリアに量子化ステップとして格納される。処理を終えると、量子化ステップ予測処理を終了し、リターンする。

30

【 0 1 1 6 】

図 4 は、上記したステップ S 8 のスペクトル量子化処理を詳細化したフローチャートである。

【 0 1 1 7 】

ステップ S 2 0 1 は、メモリ 1 0 1 上に格納されている量子化ステップとスケールファクタに従って、スペクトルバッファに格納されている 1024 本のスペクトル成分を量子化する処理である。MPEG-2 AAC の場合は、前出の式 (1) に従って量子化スペクトルが計算される。処理を終えると、ステップ S 2 0 2 へ進む。

【 0 1 1 8 】

ステップ S 2 0 2 は、ステップ S 2 0 1 で計算された量子化スペクトル全てを符号化した時に使用されるビット数を計算する処理である。例えば、MPEG-2 AAC の場合は、量子化スペクトルは複数個をまとめた上でハフマン符号化されるため、この処理においてハフマンコード表の探索が行われ、符号化ビット数の総計が計算される。計算された使用ビット数はメモリ 1 0 1 上のワークエリアに格納される。処理を終えると、ステップ S 2 0 3 へ進む。

40

【 0 1 1 9 】

ステップ S 2 0 3 は、メモリ 1 0 1 上のスペクトル割当ビットと使用ビットの大きさを比較する処理である。この比較の結果、使用ビットが割り当てられたビットよりも大きい場合は、ステップ S 2 0 4 へ進み、符号量を削減するためにメモリ 1 0 1 に格納されてい

50

る量子化ステップを増加した後、ステップS 2 0 1に戻り再度スペクトルの量子化を行うが、前述の量子化ステップ予測処理によってほぼ正確な量子化ステップが予測されるため、ステップS 2 0 4が実際に実行されることは少ない。

【0 1 2 0】

ステップS 2 0 3の比較において使用ビットが割り当てられたビットよりも小さい場合は、スペクトル量子化処理を終了してリターンする。

【0 1 2 1】

以上説明したように、本実施形態におけるオーディオ信号符号化処理では、スペクトル符号に割り当てられたビット数から量子化後のスペクトルが持つ情報量を予測し、さらに、量子化前の聴覚情報量との差分を取ることによって量子化ステップを実際の量子化を行う前にほぼ正確に予測することによって、量子化ステップの調整を極力避けることが可能になるため、量子化処理にかかる処理量を大幅に削減することができる。

【0 1 2 2】

(第3の実施形態)

固定ビットレートで符号化するとき、ビットリザーバに蓄積されている蓄積ビットを入力信号の特性によって各フレームに適宜配分する場合でも、本発明の技術を適用することが可能である。本実施形態ではこの場合について図面を用いて説明する。

【0 1 2 3】

図10は、本実施形態におけるオーディオ信号符号化装置の一構成例を示す図である。第1の実施形態に係る図1と同じく、図中の太線はデータの流れを、細線は制御信号の流れを示す。また、図10においては、図1と同様の機能を持つ構成要素には同じ番号を振っている。

【0 1 2 4】

図示の構成において、1はフレーム分割器、2は聴覚心理演算器、3はフィルタバンク、4はスケールファクタ計算器、7は量子化ステップ計算器、8はスペクトル量子化器、9はビット整形器である。

【0 1 2 5】

11はPEビット計算器であり、フレームの聴覚エン트로ピー(PE)に基づいてフレームの予測発生符号量であるPEビットを計算する。

【0 1 2 6】

12はスペクトル割当ビット計算器であり、ビットレートやPEビット、蓄積ビット量、スケールファクタなどに基づいてスペクトル符号に割り当てられるビット数を計算する。

【0 1 2 7】

13はビットリザーバであり、符号化方式に従って規定されている蓄積ビット量を逐次管理する。

【0 1 2 8】

14は量子化スペクトル総量予測器であり、条件により、フレーム割当ビット、もしくはPEビットに基づいて量子化スペクトル総量を予測する。

【0 1 2 9】

以上のような構成のオーディオ信号符号化装置における処理動作を以下に説明する。なお、本実施形態においても説明の便宜のために符号化方式としてMPEG-2 AACを例にとって説明するが、非線形量子化を行うその他の符号化方式においても全く同様な方法で実現可能である。

【0 1 3 0】

まず、処理に先立ち、各部の初期化が行われる。初期化によって、量子化ステップと全てのスケールファクタは0にセットされる。

【0 1 3 1】

オーディオ入力信号はフレーム分割器1によってフレーム単位に分割され、聴覚心理演算器2とフィルタバンク3に出力される。

【0 1 3 2】

聴覚心理演算器 2 では、フレーム分割器 1 から出力された入力信号に対して適宜聴覚マスキング分析を行い、ブロックタイプとSFB毎のSMR、PEを出力する。

【 0 1 3 3 】

フィルタバンク 3 はフレーム分割器 1 から出力された入力信号1フレームとフィルタバンク 3 内に保持している前フレームの1フレームを合わせた 2 フレーム分の入力信号を、聴覚心理演算器 2 から出力されたブロックタイプに従って時間周波数変換を行い、周波数スペクトルに変換する。

【 0 1 3 4 】

スケールファクタ計算器 4 は、フィルタバンク 3 から出力される周波数スペクトルと聴覚心理演算器 2 から出力されるSFB毎のSMR値に基づいて、スケールファクタを実施形態1

10

と同様に適宜算出する。

【 0 1 3 5 】

PEビット計算器 1 1 は聴覚心理演算器 2 が出力されるPEから、PEビットを計算する。すなわち、処理中のフレームの入力信号が持つ聴覚的な情報量を聴覚上完全に符号化した場合の予想符号量に変換する。MPEG-2 AACの場合、ISO規格書に記載されているPEビットの計算式は次のようになる。

【 0 1 3 6 】

ブロック長がロングのとき：

【数 1 0】

$$pe_bits = 0.3 \cdot PE + 6.0 \cdot \sqrt{PE} \quad (8)$$

20

ブロック長がショートのとき：

【数 1 1】

$$pe_bits = 0.6 \cdot PE + 24.0 \cdot \sqrt{PE} \quad (9)$$

【 0 1 3 7 】

本実施形態では、この計算式をこのまま用いて、ブロックタイプのブロック長に応じてPEビットを算出する。

30

【 0 1 3 8 】

スペクトル割当ビット計算器 1 2 では、まず、スケールファクタ計算器 4 から出力されるスケールファクタを符号化するために必要なビット数を計算し、次に、ビットレートに基づく 1 フレーム・チャンネル当りの平均ビット量との差分を求めて平均スペクトル割当ビットを計算する。

【 0 1 3 9 】

次に、この値とPEビット計算器 1 1 が出力するPEビットを比較し、PEビットが大きい場合はPEビットをビットリザーバ 1 3 に蓄積されている蓄積ビット量によって決定される最大値まで割り当てる。PEビットが小さい場合は平均スペクトル割当ビットをそのまま割り当てる。

40

【 0 1 4 0 】

すなわち、本実施形態において、スペクトル割当ビットは具体的には以下の手順で計算される。

【 0 1 4 1 】

1. 蓄積ビット使用許容量を蓄積ビット量から計算する。

ブロック長がロングのとき：蓄積ビット量の10%、

ブロック長がショートの場合：蓄積ビット量の25%、

を蓄積ビット使用許容量とする。これをusable_bitsとする。

【 0 1 4 2 】

2. 平均スペクトル割当ビット量をaverage_bitsとすると、スペクトル割当ビット量、s

50

spectrum_bitsは以下の要領で決定される。

pe_bits > (average_bits+usable_bits) のとき、

spectrum_bits=average_bits+usable_bits;

pe_bits < average_bits のとき、

spectrum_bits=average_bits;

それ以外、average_bits pe_bits (average_bits+usable_bits) のとき、

spectrum_bits = pe_bits;

【 0 1 4 3 】

次に、スペクトル割当ビット計算器 1 2 は、PEビットが平均スペクトル割当ビット量よりも少ない場合は、PEビットを量子化スペクトル総量予測器 1 4 に出力し、PEビットが平均スペクトル割当ビット以上の場合は上記手順で計算したスペクトル割当ビットを量子化スペクトル総量予測器 1 4 に出力する。このとき、どちらのビット数を量子化スペクトル総量予測器 1 4 に出力したかを示すフラグであるビット選択情報（以下、単に「選択情報」という。）を同時に出力する。

【 0 1 4 4 】

量子化スペクトル総量予測器 1 4 は、入力された選択情報とビット数に基づいて量子化スペクトル総量を予測する。この予測計算は、第 1 の実施形態で示した方法と同様に、実験によって求めた近似式によって行うが、本実施形態における量子化スペクトル総量予測器 1 4 では、この近似式を選択情報によって切り換えて予測計算を行う。例えば、スペクトル割当ビットによる量子化スペクトル総量の近似式をF(x)、PEビットによる量子化スペクトル総量の近似式をG(x)とすると、スペクトル予測総量は以下の式で求められる。

【 0 1 4 5 】

選択情報がスペクトル割当ビットの選択を示している場合：

【数 1 2】

$$\sum_i X_q \approx F(\text{spectrum_bits}) \quad (4)$$

選択情報がPEビットの選択を示している場合：

【数 1 3】

$$\sum_i X_q \approx \frac{\text{bit_rate}}{\text{base_bit_rate}} \cdot \frac{\text{base_sampling_rate}}{\text{sampling_rate}} \cdot G(\text{pe_bits}) \quad (10)$$

【 0 1 4 6 】

ただし、式 (1 0) において、bit_rateは処理中の入力信号のビットレート、sampling_rateは処理中の入力信号のサンプリングレート、base_bit_rateは基準ビットレート、base_sampling_rateは基準サンプリングレートである。基準ビットレートと基準サンプリングレートは、PEビットによる量子化スペクトル総量予測式G(x)を実験によって求めたときの入力信号のビットレートとサンプリングレートであり、本実施形態におけるオーディオ信号符号化装置において予め定められている値である。

【 0 1 4 7 】

ここで、本実施形態において上記のような量子化スペクトルの予測方法を取る理由を以下説明する。

【 0 1 4 8 】

本実施形態では、スペクトル割当ビット計算器 1 2 においてPEビットを基準にしたビットの割当が行われる。従って、スペクトル割当ビットには通常PEビットの大きさ、すなわち、処理中のフレームにおける入力信号が持つ聴覚上の発生符号量が反映される。ところが、固定ビットレート制御において、PEビットの大きさが平均スペクトル割当ビットを下回る場合には、スペクトル割当ビットには平均スペクトル割当ビットがそのまま割り当てられる。したがって、この場合はスペクトル割当ビットには入力信号の聴覚上の発生符号量が反映されないため、スペクトル割当ビットを用いて量子化スペクトル総量を予測する

10

20

30

40

50

と予測誤差が大きくなってしまいます。そこで、この場合にはPEビットを用いて量子化スペクトル総量を予測することで、より正確な量子化スペクトル総量を予測することができる。

【0149】

また、スペクトル割当ビットはビットレートやサンプリングレートの制約を考慮して計算されるため、ビットレートやサンプリングレートの変化に追従する特性を持っている。一方、PEビットは、サンプリングレートの変化によって元となるPE自体の値は変化するものの、ビットレートやサンプリングレートが変化しても式(8)、(9)自体は変化しない。そこで、PEビットによる予測を行う場合は式(10)に示すように、基準となるビットレートやサンプリングレートからの変化率の影響を考慮した予測を行なっている。

【0150】

このようにすることで、一つの近似式 $G(x)$ をあらゆるビットレートやサンプリングレートに適用することが可能になる。

【0151】

図10の説明に戻る。量子化ステップ計算器7は第1の実施形態と同様に、フィルタバンク3が出力する周波数スペクトルにスケールファクタ計算器4から出力されたスケールファクタによって重み付けをした値の総量を計算し、さらにその対数を計算して量子化前のスペクトルが持つ聴覚情報量を算出する。次に、量子化スペクトル総量予測器14によって予測された量子化スペクトル総量の対数を計算して量子化後のスペクトル情報量を算出する。さらに、この差分を取り、量子化粗さの刻み幅によって決定される係数を掛けることによって量子化ステップを計算する。具体的には、前出の式(5)の計算を行う。

【0152】

第1の実施形態と同様に、スペクトル量子化器8はスケールファクタ計算器4が出力するスケールファクタと量子化ステップ計算器7が出力する量子化ステップを用いてフィルタバンク3が出力する周波数スペクトルを量子化し、必要ビット数をカウントし、スペクトル割当ビット計算器12が出力するスペクトル割当ビットと比較する。ここで、必要ビット数がスペクトル割当ビットを超えてしまった場合は、量子化ステップを適宜増加して再度量子化を行うが、前述したように量子化ステップ計算器7による量子化ステップの予測値はほぼ正確であるため、この再量子化が行われることは少ない。

【0153】

スペクトル量子化器8によって最終的に出力される量子化スペクトルとスケールファクタ、量子化ステップは、ビット整形器9によってエントロピー符号化後、符号化方式が定めるビットストリーム形式に適宜整形され、出力される。

【0154】

このとき、ビットリザーバ13に実際に符号に使用したビット数が通知され、ビットリザーバ13はフレームビットとの差分を計算し、増減分を蓄積ビット量に加減することで適宜蓄積ビット量を調整する。

【0155】

以上説明したように、本実施形態のように固定ビットレートにおいて、入力信号に応じてビットリザーバに蓄積されている蓄積ビットを適宜フレームに割り当てる場合においても、量子化前に正確に量子化スペクトル総量を予測することによって、量子化前に正確に量子化ステップを決定することが可能になり、スペクトル量子化とビット計算の繰り返しを避けて、効率的に量子化を行うことができる。

【0156】

(第4の実施形態)

第3の実施形態で説明したオーディオ符号化装置もまた、PC等の汎用的なコンピュータ上で動作するソフトウェアプログラムとして実施することが可能である。以下、この場合について図面を用いて説明する。

【0157】

本実施形態におけるオーディオ信号符号化装置の構成およびオーディオ信号符号化処理プログラムの処理内容等については、第2の実施形態と概ね共通しているので、本実施形

10

20

30

40

50

態では、第 2 の実施形態で説明した図 5、図 2、図 6 ~ 9 を援用することとし、それらの詳細な説明は省略する。第 2 の実施形態と相違する点は、ステップ S 7 の量子化ステップ予測処理の内容にある。そこで以下では、このステップ S 7 の量子化ステップ予測処理の説明のみを行う。

【 0 1 5 8 】

図 1 1 は、本実施形態におけるステップ S 7 の量子化ステップ予測処理の詳細を示すフローチャートである。

【 0 1 5 9 】

まず、ステップ S 3 0 1 は、ステップ S 4 の聴覚心理演算処理によって得られたメモリ 1 0 1 上の PE とブロックタイプより、PE ビットを算出する処理である。具体的には、第 3 の実施形態と同様にブロックタイプによって、前出の式 (9) もしくは式 (1 0) を選択して PE ビットを算出する。算出した PE ビットはメモリ 1 0 1 上のワークエリアに格納される。処理を終えるとステップ S 3 0 2 へ進む。

10

【 0 1 6 0 】

ステップ S 3 0 2 は、メモリ 1 0 1 上のワークエリアに保存されているスケールファクタを符号化方式規定のフォーマットに符号化したときに使用するビット数を算出する処理である。この処理によって算出されたスケールファクタビット数は、メモリ 1 0 1 上のワークエリアに保存される。処理を終えると、ステップ S 3 0 3 へ進む。

【 0 1 6 1 】

ステップ S 3 0 3 は、フレームに割り当てられる平均ビット数からメモリ 1 0 1 上に格納されたスケールファクタビット数を引いて、スペクトル符号に割り当てられるビット数、すなわち平均スペクトル割当ビット数 (平均割当ビット) を算出する処理である。算出された平均割当ビット数はメモリ 1 0 1 上のワークエリアに保存される。処理を終えると、ステップ S 3 0 4 へ進む。

20

【 0 1 6 2 】

ステップ S 3 0 4 は、メモリ 1 0 1 上の平均割当ビット数と PE ビット数の大きさを比較する処理である。この比較の結果、PE ビット数の方が大きい場合はステップ S 3 0 5 へ進む。それ以外の場合は、ステップ S 3 0 7 へ進む。

【 0 1 6 3 】

ステップ S 3 0 5 は、メモリ 1 0 1 上の PE ビットと平均割当ビット、蓄積ビット量よりスペクトル割当ビットを算出する処理である。この処理の詳細は図 1 2 を用いて後述する。処理を終えると、ステップ S 3 0 6 へ進む。

30

【 0 1 6 4 】

ステップ S 3 0 6 は、メモリ 1 0 1 上のスペクトル割当ビット数を用いて、量子化スペクトル総量の予測計算を行う処理である。この予測計算は、予め実験を実施することによって求めた近似式によって行う。例えば、この近似式を $F(x)$ として、スペクトル割当ビットを $spectrum_bits$ とすると、量子化後スペクトル予測総量は第 2 の実施形態と同様に式 (4) によって求めることができる。

【 0 1 6 5 】

【数 1 4】

40

$$\sum_i X_q \approx F(spectrum_bits) \quad (4)$$

【 0 1 6 6 】

算出された量子化スペクトル予測総量はメモリ 1 0 1 上のワークエリアに格納される。処理を終えると、ステップ S 3 0 9 へ進む。

【 0 1 6 7 】

一方のステップ S 3 0 7 は、メモリ 1 0 1 上の平均割当ビットをスペクトル割当ビットとしてメモリ 1 0 1 に格納する処理である。すなわち、平均割当ビットの値をスペクトル割当ビットにコピーする。処理を終えると、ステップ S 3 0 8 へ進む。

50

【 0 1 6 8 】

ステップ S 3 0 8 は、メモリ 1 0 1 上の PE ビット数を用いて、量子化スペクトル総量の予測計算を行う処理である。この予測計算も、予め実験を実施することによって求めた近似式によって行う。この近似式を $G(x)$ として、PE ビットを pe_bits とすると、量子化後スペクトル予測総量は第 3 の実施形態と同様に式 (1 0) によって求めることができる。

【 0 1 6 9 】

【 数 1 5 】

$$\sum_i X_q \approx \frac{bit_rate}{base_bit_rate} \cdot \frac{base_sampling_rate}{sampling_rate} \cdot G(pe_bits) \quad (10)$$

10

【 0 1 7 0 】

算出された量子化スペクトル予測総量はメモリ 1 0 1 上のワークエリアに格納される。処理を終えると、ステップ S 3 0 9 へ進む。

【 0 1 7 1 】

ステップ S 3 0 9 は、量子化前のスペクトルが持つ聴覚情報量を算出する処理である。第 2 の実施形態と同様に、量子化前スペクトルの聴覚情報量は、各スペクトル成分に、そのスペクトル成分が含まれる SFB のスケールファクタによる量子化粗さの減少分を積算し、1 フレーム分の総量を求め、その対数を算出することによって求められる。例えば、MPEG-2 AAC の場合、量子化前のスペクトルが持つ聴覚情報量は以下の式を計算することによって求めることができる。

20

【 0 1 7 2 】

【 数 1 6 】

$$\log_2 \sum_i \left[|x_i|^{\frac{3}{4}} \cdot 2^{\frac{3}{16} \cdot scalefac} \right] \quad (6)$$

【 0 1 7 3 】

算出された量子化前スペクトルの聴覚情報量はメモリ 1 0 1 上のワークエリアに保存される。処理を終えると、ステップ S 3 1 0 へ進む。

【 0 1 7 4 】

30

ステップ S 3 1 0 は、ステップ S 3 0 6 もしくはステップ S 3 0 8 で求めた量子化スペクトルの予測総量の対数を計算し、量子化スペクトルの予測情報量を算出する処理である。例えば、MPEG-2 AAC の場合は以下の式を計算することによって算出することができる。

【 0 1 7 5 】

【 数 1 7 】

$$\log_2 \sum_i X_q \quad (7)$$

【 0 1 7 6 】

40

この処理によって算出された量子化後のスペクトル予測情報量はメモリ 1 0 1 上のワークエリアに保存される。処理を終えると、ステップ S 3 1 1 へ進む。

【 0 1 7 7 】

ステップ S 3 1 1 では、ステップ S 3 0 9 で求めた量子化前スペクトルの聴覚情報量から、ステップ S 3 1 0 で求めた量子化スペクトル予測情報量を引き、ステップ S 3 1 2 で、その結果に量子化粗さの刻み幅によって決定される係数を掛け、グローバルゲイン、すなわち量子化ステップの予測値が算出される。MPEG-2 AAC の場合は、この予測値は結局第 1 の実施形態と同じく式 (5) を計算したことになる。

【 0 1 7 8 】

【数 18】

$$global_gain = \text{Int} \left[\frac{16}{3} \cdot \left[\log_2 \sum_i \left[|x_i|^{\frac{3}{4}} \cdot 2^{\frac{3}{16} \cdot scalefac} \right] - \log_2 \sum_i X_q \right] \right] \quad (5)$$

【0179】

算出された量子化ステップ予測値は、メモリ101上のワークエリアに量子化ステップとして格納される。処理を終えると、量子化ステップ予測処理を終了し、リターンする。

【0180】

図12は、本実施形態におけるステップS305のスペクトル割当ビット算出処理の詳細を示すフローチャートである。

10

【0181】

ステップS401は、このフレームに割り当てることができる蓄積ビット数を、メモリ101上の蓄積ビット量とブロックタイプに従って計算し、この値を平均割当ビットに足すことで、スペクトル割当ビットの上限値を計算する処理である。

本実施形態において、蓄積ビット数は第3の実施形態と同じく以下の要領で決定される。

【0182】

ブロック長がロングのとき：蓄積ビット量の10%、

ブロック長がショートのとき：蓄積ビット量の25%

【0183】

上記手順で求められた値をメモリ101上の平均割当ビットに加えることで、スペクトル割当ビット上限値が得られる。

20

【0184】

この計算によって得られたスペクトル割当ビット上限値はメモリ101に格納される。処理を終えると、ステップS402へ進む。

【0185】

ステップS402は、メモリ101上のPEビットとスペクトル割当ビット上限値を比較する処理である。この比較の結果、PEビット数がスペクトル割当ビット上限値よりも少ない場合はステップS403へ進む。そうでない場合は、ステップS404へ進む。

【0186】

ステップS403は、メモリ101上のPEビットをスペクトル割当ビットとして格納する処理である。すなわち、PEビットの値をスペクトル割当ビットにコピーする。処理を終えると、スペクトル割当ビット算出処理を終えて、リターンする。

30

【0187】

ステップS404は、メモリ101上のスペクトル割当ビット上限値をスペクトル割当ビットとして格納する処理である。すなわち、スペクトル割当ビット上限値をスペクトル割当ビットにコピーする。処理を終えると、スペクトル割当ビット算出処理を終えて、リターンする。

【0188】

本処理では、上述したようにPEビットによって割り当てられるビット数に上限値を設定することによって、蓄積ビットが枯渇してビットリザーバが破綻することを防止する効果がある。

40

【0189】

以上説明したように、本実施形態によれば、固定ビットレートにおいて、入力信号の特性に応じてビットリザーバに蓄積されている蓄積ビットを適宜フレームに割り当てる場合においても、量子化前に正確に量子化スペクトル総量を予測することによって、量子化前に正確に量子化ステップを決定することが可能になり、スペクトル量子化とビット計算の繰り返しを避けて、効率的に量子化を行うことができる。

【0190】

以上説明したように、本発明のオーディオ信号符号化処理では、フレームに割り当てら

50

れたビット量から、量子化後のスペクトル総量を予測することにより、量子化前後のスペクトル全体が持つ情報量の差分を計算し、スペクトル量子化の前にフレーム全体の量子化ステップをほぼ正確に予測することで、ほぼ一回のスペクトル量子化処理を行うだけで量子化処理を終了することが可能になる。これにより、従来の技術と同等の符号化品質を保ちながら、従来の技術に比べて量子化処理にかかる処理量を大幅に削減することができる。

【0191】

(他の実施形態)

本発明は上述した実施形態に限定されるものではない。

【0192】

上述の実施形態では、オーディオ符号化装置および方法としてブロックスイッチングが行われる符号化方式の場合を取り扱っているが、ブロックスイッチングを行わない符号化方式においても同様に実施することが可能である。

【0193】

その他、本発明はその要旨を逸脱しない範囲で種々変形して実施することができる。

【0194】

以上、本発明の実施形態を詳述したが、本発明は、複数の機器から構成されるシステムに適用してもよいし、また、一つの機器からなる装置に適用してもよい。

【0195】

なお、本発明は、前述したとおり、実施形態の機能を実現するオーディオ信号符号化処理プログラムを、システムあるいは装置に直接あるいは遠隔から供給し、そのシステムあるいは装置のコンピュータがその供給されたプログラムコードを読み出して実行することによっても達成される。その場合、プログラムの機能を有していれば、その形態はプログラムである必要はない。

【0196】

従って、本発明の機能処理をコンピュータで実現するために、そのコンピュータにインストールされるプログラムコード自体およびそのプログラムを格納した記憶媒体も本発明を構成することになる。つまり、本発明の特許請求の範囲には、本発明の機能処理を実現するためのコンピュータプログラム自体、およびそのプログラムを格納した記憶媒体も含まれる。

【0197】

その場合、プログラムの機能を有していれば、オブジェクトコード、インタプリタにより実行されるプログラム、OSに供給するスクリプトデータ等、プログラムの形態を問わない。

【0198】

プログラムを供給するための記憶媒体としては、例えば、フレキシブルディスク、ハードディスク、光ディスク、光磁気ディスク、MO、CD-ROM、CD-R、CD-RW、磁気テープ、不揮発性のメモ리카ード、ROM、DVD(DVD-ROM, DVD-R)などがある。

【0199】

その他、プログラムの供給方法としては、クライアントコンピュータのブラウザを用いてインターネットのホームページに接続し、そのホームページから本発明のコンピュータプログラムそのもの、もしくは圧縮され自動インストール機能を含むファイルをハードディスク等の記憶媒体にダウンロードすることによっても供給できる。また、本発明のプログラムを構成するプログラムコードを複数のファイルに分割し、それぞれのファイルを異なるホームページからダウンロードすることによっても実現可能である。つまり、本発明の機能処理をコンピュータで実現するためのプログラムファイルを複数のユーザに対してダウンロードさせるWWWサーバも、本発明の範囲に含まれるものである。

【0200】

また、本発明のプログラムを暗号化してCD-ROM等の記憶媒体に格納してユーザに

10

20

30

40

50

配布し、所定の条件をクリアしたユーザに対し、インターネットを介してホームページから暗号化を解く鍵情報をダウンロードさせ、その鍵情報を使用することにより暗号化されたプログラムを実行してコンピュータにインストールさせて実現することも可能である。

【0201】

また、コンピュータが、読み出したプログラムを実行することによって、前述した実施形態の機能が実現される他、そのプログラムの指示に基づき、コンピュータ上で稼動しているOSなどが、実際の処理の一部または全部を行い、その処理によっても前述した実施形態の機能が実現され得る。

【0202】

さらに、記憶媒体から読み出されたプログラムが、コンピュータに挿入された機能拡張ボードやコンピュータに接続された機能拡張ユニットに備わるメモリに書き込まれた後、そのプログラムの指示に基づき、その機能拡張ボードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行い、その処理によっても前述した実施形態の機能が実現される。

【図面の簡単な説明】

【0203】

【図1】図1は、本発明の第1の実施形態におけるオーディオ信号符号化装置の一構成例を示す図である。

【図2】図2は、本発明の第2の実施形態におけるオーディオ信号符号化処理のフローチャートである。

【図3】図3は、本発明の第2の実施形態における量子化ステップ予測処理のフローチャートである。

【図4】図4は、本発明の第2の実施形態におけるスペクトル量子化処理のフローチャート、

【図5】図5は、本発明の第2の実施形態におけるオーディオ信号符号化装置の一構成例を示す図である。

【図6】図6は、本発明の第2の実施形態におけるオーディオ信号符号化処理プログラムを格納した記憶媒体の内容構成例を示す図である。

【図7】図7は、本発明の第2の実施形態におけるオーディオ信号符号化処理プログラムのPCへの導入を示す模式図である。

【図8】図8は、本発明の第2の実施形態におけるメモリマップの例を示す図である。

【図9】図9は、本発明の第2の実施形態における入力信号バッファの構成例を示す図である。

【図10】図10は、本発明の第3の実施形態におけるオーディオ信号符号化装置の一構成例を示す図である。

【図11】図11は、本発明の第4の実施形態における量子化ステップ予測処理のフローチャートである。

【図12】図12は、本発明の第4の実施形態におけるスペクトル割当ビット算出処理のフローチャートである。

【図13】図13は、ISO規格書に従う量子化処理のフローチャートである。

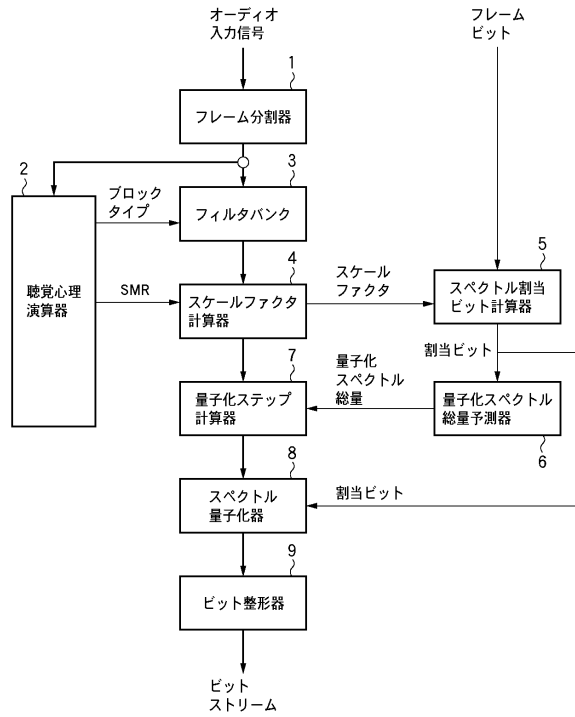
10

20

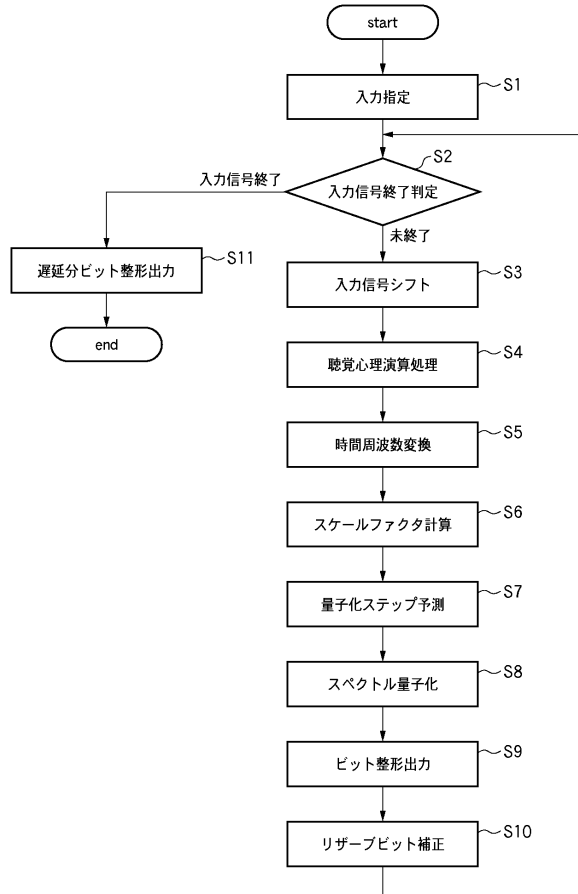
30

40

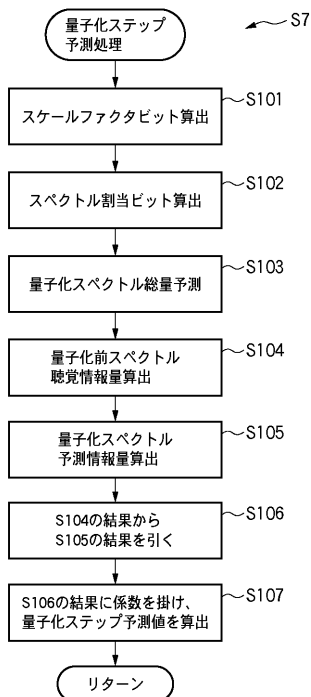
【図 1】



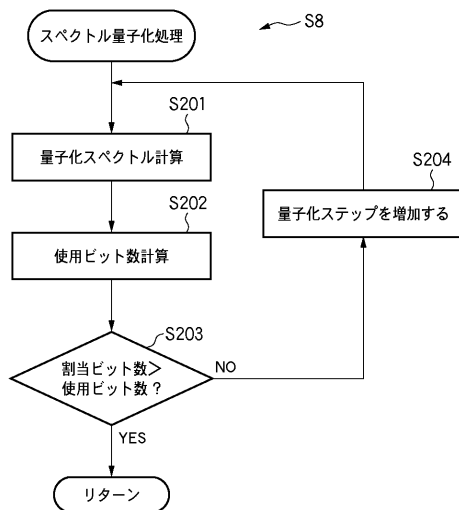
【図 2】



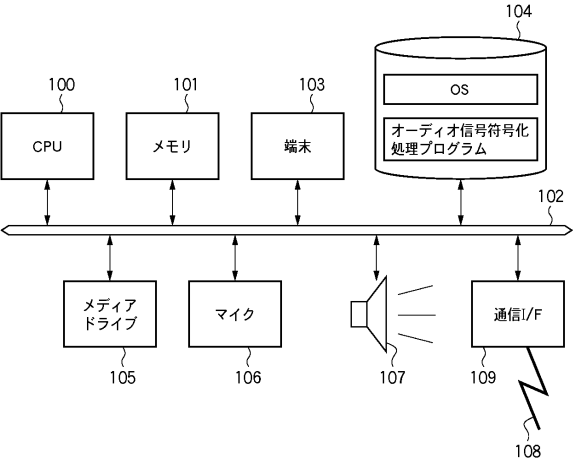
【図 3】



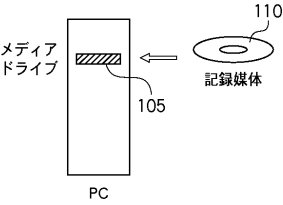
【図 4】



【図 5】



【図 7】



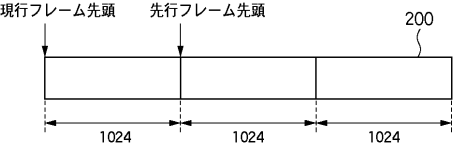
【図 6】

ディレクトリ情報
⋮
オーディオ信号符号化処理プログラム 実行ファイル
⋮
オーディオ信号符号化処理プログラム 関連データファイル

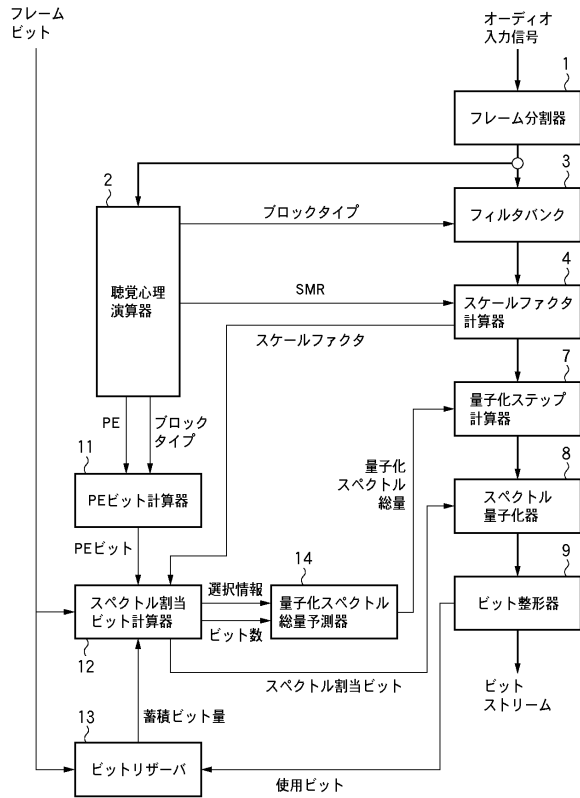
【図 8】

上位アドレス ↓ 下位アドレス	基本I/Oプログラム	
	OS	
	オーディオ信号符号化処理プログラム	
	関連データ	
	ワークエリア	
	基準 ビットレート	基準 サンプリングレート
	ビットレート	サンプリングレート
	割当ビット上限値	平均割当ビット
	PEビット	使用ビット
	スケールファクタ ビット	スペクトル 割当ビット
	量子化前スペクトル 聴覚情報量	量子化後スペクトル 予測情報量
	許容誤差エネルギー	スペクトルバッファ
	量子化スペクトル	入力信号バッファ
	スケールファクタ	量子化ステップ
	ブロックタイプ	SMR
	PE	リザーブビット量

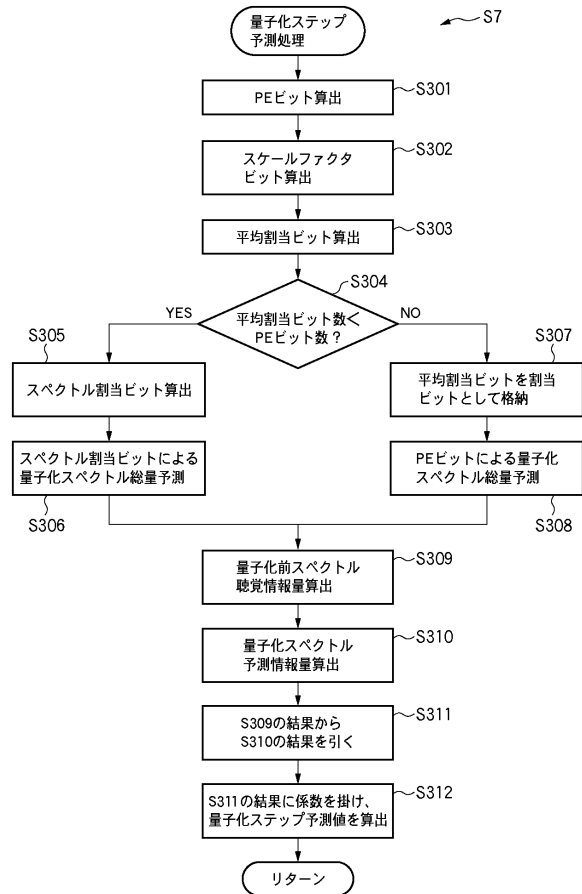
【図 9】



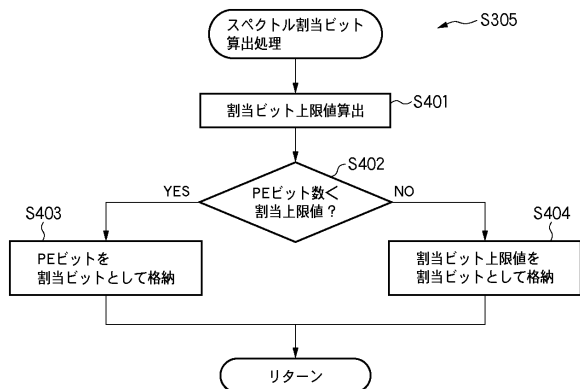
【図 10】



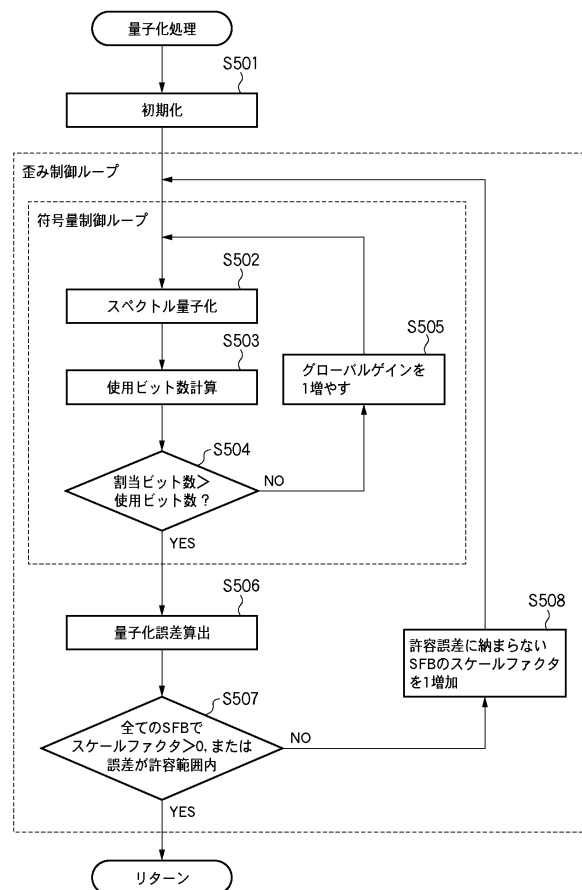
【図 11】



【図 12】



【図 13】



フロントページの続き

(72)発明者 船越 正伸
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

審査官 山下 剛史

(56)参考文献 特開2002-182695(JP,A)
特開2001-242895(JP,A)
特開2004-309921(JP,A)
特開2002-91498(JP,A)
特開2001-148632(JP,A)
特開2001-249699(JP,A)

(58)調査した分野(Int.Cl., DB名)
G10L 19/00 - 19/14
H04B 14/04
H03M 7/30