

【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 20 年 1 月 17 日 (2008.1.17)

【公開番号】特開 2006-145782(P2006-145782A)
 【公開日】平成 18 年 6 月 8 日 (2006.6.8)
 【年通号数】公開・登録公報 2006-022
 【出願番号】特願 2004-335005(P2004-335005)
 【国際特許分類】

G 1 0 L 19/02 (2006.01)

H 0 3 M 7/30 (2006.01)

【F I】

G 1 0 L 19/02 1 4 2 B

G 1 0 L 19/02 1 4 2 C

H 0 3 M 7/30 A

【手続補正書】

【提出日】平成 19 年 11 月 19 日 (2007.11.19)

【手続補正 1】

【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】

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、

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、

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。

【手続補正2】

【補正対象書類名】明細書

【補正対象項目名】0034

【補正方法】変更

【補正の内容】

【0034】

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。

【手続補正 3】

【補正対象書類名】明細書

【補正対象項目名】0061

【補正方法】変更

【補正の内容】

【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。

【手続補正 4】

【補正対象書類名】明細書

【補正対象項目名】0069

【補正方法】変更

【補正の内容】

【0069】

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。

【手続補正 5】

【補正対象書類名】明細書

【補正対象項目名】0078

【補正方法】変更

【補正の内容】

【0078】

105 はメディアドライブである。記録媒体（例えば、CD-ROM）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。

【手続補正 6】

【補正対象書類名】明細書

【補正対象項目名】0082

【補正方法】変更

【補正の内容】

【0082】

本実施形態のオーディオ信号符号化装置は、CPU 100 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。

【手続補正 7】

【補正対象書類名】明細書

【補正対象項目名】0085

【補正方法】変更

【補正の内容】

【 0 0 8 5 】

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（ P C ）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 1 0 5 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 1 1 0 をメディアドライブ 1 0 5 にセットすると、 O S 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 1 1 0 から読み出され、外部記憶装置 1 0 4 に格納される。その後、再起動時にこれらの情報がメモリ 1 0 1 にロードされて動作可能となる。

【 手続補正 8 】

【 補正対象書類名 】 明細書

【 補正対象項目名 】 0 0 9 0

【 補正方法 】 変更

【 補正の内容 】

【 0 0 9 0 】

まず、ステップ S 1 は、符号化する入力オーディオ信号をユーザが端末 1 0 3 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 1 0 4 に格納されているオーディオ P C M ファイルでも良いし、マイク 1 0 6 で捉えたり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 S 2 へ進む。

【 手続補正 9 】

【 補正対象書類名 】 明細書

【 補正対象項目名 】 0 1 2 5

【 補正方法 】 変更

【 補正の内容 】

【 0 1 2 5 】

1 1 は P E ビット計算器であり、フレームの聴覚エントロピー (P E) に基づいてフレームの予測発生符号量である P E ビットを計算する。

【 手続補正 1 0 】

【 補正対象書類名 】 明細書

【 補正対象項目名 】 0 1 3 3

【 補正方法 】 変更

【 補正の内容 】

【 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。

【 手続補正 1 1 】

【 補正対象書類名 】 明細書

【 補正対象項目名 】 0 1 3 5

【 補正方法 】 変更

【 補正の内容 】

【 0 1 3 5 】

P E ビット計算器 1 1 は聴覚心理演算器 2 が出力される P E から、 P E 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。 M P E G - 2 A A C の場合、 I S O 規格書に記載されている P E ビットの計算式は次のようになる。

【 手続補正 1 2 】

【 補正対象書類名 】 明細書

【 補正対象項目名 】 0 1 7 7

【補正方法】変更

【補正の内容】

【0177】

ステップS311では、ステップS309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S310で求めた量子化スペクトル予測情報量を引き、ステップS312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。

【手続補正13】

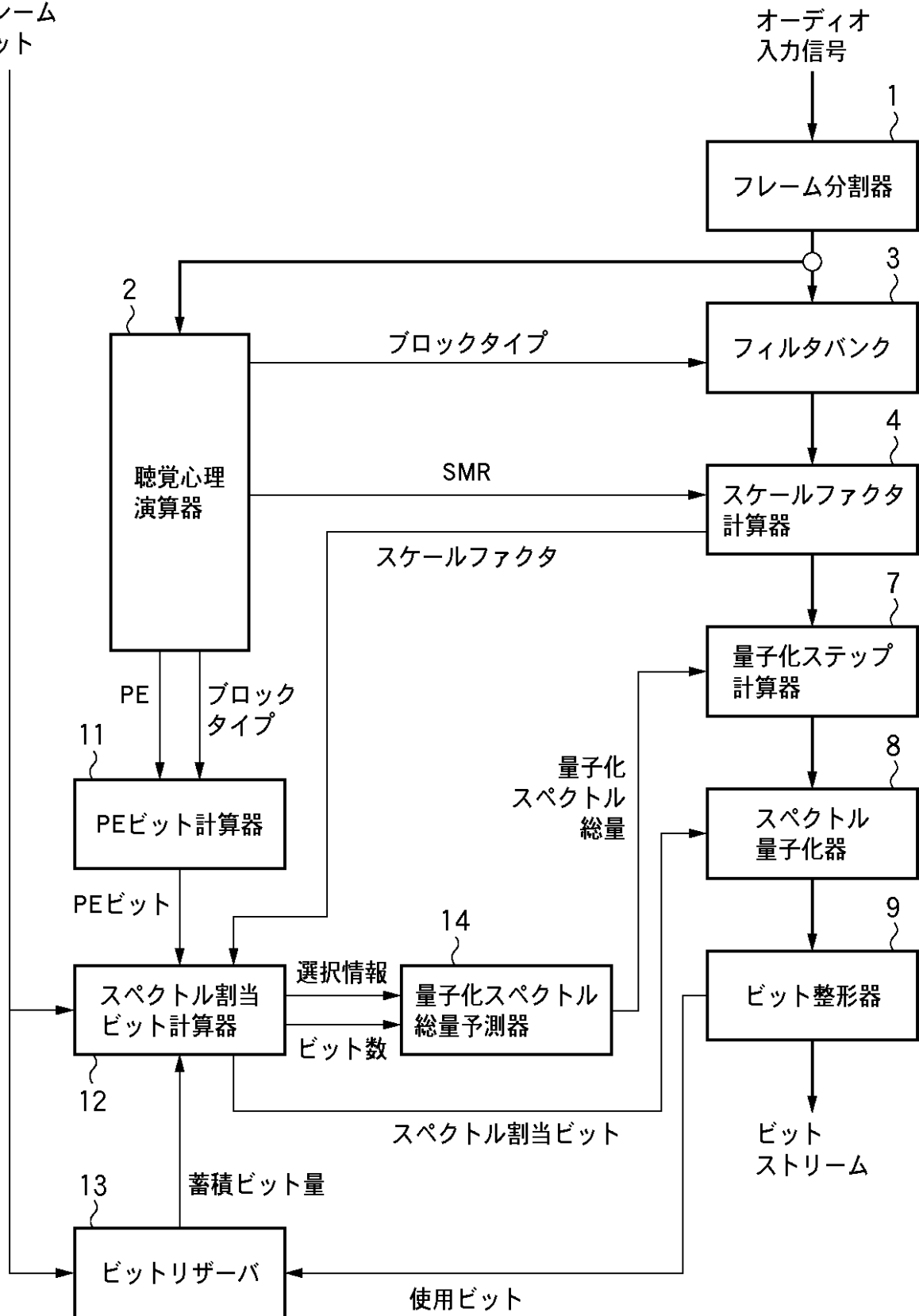
【補正対象書類名】図面

【補正対象項目名】図10

【補正方法】変更

【補正の内容】

【図 10】

フレーム
ビット

【手続補正 14】

【補正対象書類名】図面

【補正対象項目名】図 11

【補正方法】変更

【補正の内容】

【図 11】

