

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号
特許第4218271号
(P4218271)

(45) 発行日 平成21年2月4日(2009.2.4)

(24) 登録日 平成20年11月21日(2008.11.21)

(51) Int.Cl.
G 1 O L 19/02 (2006.01)

F I
G 1 O L 19/02 1 7 O Z

請求項の数 13 (全 62 頁)

(21) 出願番号	特願2002-210899 (P2002-210899)	(73) 特許権者	000002185
(22) 出願日	平成14年7月19日 (2002.7.19)		ソニー株式会社
(65) 公開番号	特開2004-53891 (P2004-53891A)		東京都港区港南1丁目7番1号
(43) 公開日	平成16年2月19日 (2004.2.19)	(74) 代理人	100082131
審査請求日	平成17年6月30日 (2005.6.30)		弁理士 稲本 義雄
		(72) 発明者	近藤 哲二郎
			東京都品川区北品川6丁目7番35号 ソ
			ニー株式会社内
		(72) 発明者	渡辺 勉
			東京都品川区北品川6丁目7番35号 ソ
			ニー株式会社内
		審査官	菊池 智紀

最終頁に続く

(54) 【発明の名称】 データ処理装置およびデータ処理方法、並びにプログラムおよび記録媒体

(57) 【特許請求の範囲】

【請求項 1】

時間領域のオーディオデータを周波数領域のオーディオデータに変換することにより得られる第1のデータを、前記第1のデータよりも高品質のデータである第2のデータに変換するデータ処理装置であって、

前記第2のデータのうちの、注目しているデータである注目データをクラス分けするクラス分類に用いる、前記注目データに近接する複数のデータからなるクラスタップを、前記第1のデータから抽出するクラスタップ抽出手段と、

前記クラスタップを構成する複数のデータの値から、前記クラスに対応するクラスコードを算出することにより、そのクラスコードに対応するクラスに、前記注目データをクラス分類するクラス分類手段と、

前記注目データを求める積和演算に用いる、前記注目データに近接する複数のデータからなる予測タップを、前記第1のデータから抽出する予測タップ抽出手段と、

学習の教師となる、前記第2のデータに対応する教師データと、前記学習の生徒となる、前記第1のデータに対応する生徒データを用い、注目している前記教師データの予測値を、その教師データに近接する複数の前記生徒データとタップ係数との積和演算により求め、注目している前記教師データと、求められた前記予測値との差分値である予測誤差を最小にする学習を、前記1以上のクラスごとに行うことにより求められたタップ係数から、前記注目データのクラスのタップ係数を取得する取得手段と、

前記注目データのクラスのタップ係数と、前記予測タップとを用いた積和演算を行うこ

10

20

とにより前記注目データを求めることで、知覚符号化技術または聴覚心理符号化技術により一部の周波数帯域に対するビット割り当てが0とされることがある、複数の周波数帯域それぞれごとのオーディオデータである前記第1のデータを、前記第1のデータにおいてビット割り当てが0とされた前記第1のデータにおいて抜けている周波数帯域のオーディオデータである前記第2のデータに変換する予測演算手段と

を備えることを特徴とするデータ処理装置。

【請求項2】

前記予測演算手段は、前記第1のデータを、複数の周波数帯域それぞれごとのオーディオデータである前記第2のデータに変換する

ことを特徴とする請求項1に記載のデータ処理装置。

10

【請求項3】

前記第2のデータである前記複数の周波数帯域それぞれごとのオーディオデータを合成して、時間領域のオーディオデータを求める合成手段をさらに備える

ことを特徴とする請求項2に記載のデータ処理装置。

【請求項4】

前記第2のデータは、時間領域または周波数領域のオーディオデータであり、

前記第2のデータが周波数領域のオーディオデータである場合には、その周波数領域のオーディオデータを、時間領域のオーディオデータに変換する変換手段をさらに備える

ことを特徴とする請求項1に記載のデータ処理装置。

【請求項5】

20

前記第2のデータである時間領域のオーディオデータ、または前記第2のデータである周波数領域のオーディオデータを変換して得られる時間領域のオーディオデータを補正する補正手段をさらに備える

ことを特徴とする請求項4に記載のデータ処理装置。

【請求項6】

時間領域のオーディオデータを周波数領域のオーディオデータに変換することにより得られる第1のデータを、前記第1のデータよりも高品質のデータである第2のデータに変換するデータ処理方法であって、

前記第2のデータのうちの、注目しているデータである注目データをクラス分けするクラス分類に用いる、前記注目データに近接する複数のデータからなるクラスタップを、前記第1のデータから抽出するクラスタップ抽出ステップと、

30

前記クラスタップを構成する複数のデータの値から、前記クラスに対応するクラスコードを算出することにより、そのクラスコードに対応するクラスに、前記注目データをクラス分類するクラス分類ステップと、

前記注目データを求める積和演算に用いる、前記注目データに近接する複数のデータからなる予測タップを、前記第1のデータから抽出する予測タップ抽出ステップと、

学習の教師となる、前記第2のデータに対応する教師データと、前記学習の生徒となる、前記第1のデータに対応する生徒データを用い、注目している前記教師データの予測値を、その教師データに近接する複数の前記生徒データとタップ係数との積和演算により求め、注目している前記教師データと、求められた前記予測値との差分値である予測誤差を

40

最小にする学習を、前記1以上のクラスごとに行うことにより求められたタップ係数から、前記注目データのクラスのタップ係数を取得する取得ステップと、
前記注目データのクラスのタップ係数と、前記予測タップとを用いた積和演算を行うことにより前記注目データを求めることで、知覚符号化技術または聴覚心理符号化技術により一部の周波数帯域に対するビット割り当てが0とされることがある、複数の周波数帯域それぞれごとのオーディオデータである前記第1のデータを、前記第1のデータにおいてビット割り当てが0とされた前記第1のデータにおいて抜けている周波数帯域のオーディオデータである前記第2のデータに変換する予測演算ステップと

を備えることを特徴とするデータ処理方法。

【請求項7】

50

時間領域のオーディオデータを周波数領域のオーディオデータに変換することにより得られる第1のデータを、前記第1のデータよりも高品質のデータである第2のデータに変換するデータ処理を、コンピュータに行わせるプログラムであって、

前記第2のデータのうちの、注目しているデータである注目データをクラス分けするクラス分類に用いる、前記注目データに近接する複数のデータからなるクラスタップを、前記第1のデータから抽出するクラスタップ抽出ステップと、

前記クラスタップを構成する複数のデータの値から、前記クラスに対応するクラスコードを算出することにより、そのクラスコードに対応するクラスに、前記注目データをクラス分類するクラス分類ステップと、

前記注目データを求める積和演算に用いる、前記注目データに近接する複数のデータからなる予測タップを、前記第1のデータから抽出する予測タップ抽出ステップと、

学習の教師となる、前記第2のデータに対応する教師データと、前記学習の生徒となる、前記第1のデータに対応する生徒データを用い、注目している前記教師データの予測値を、その教師データに近接する複数の前記生徒データとタップ係数との積和演算により求め、注目している前記教師データと、求められた前記予測値との差分値である予測誤差を最小にする学習を、前記1以上のクラスごとに行うことにより求められたタップ係数から、前記注目データのクラスのタップ係数を取得する取得ステップと、

前記注目データのクラスのタップ係数と、前記予測タップとを用いた積和演算を行うことにより前記注目データを求めることで、知覚符号化技術または聴覚心理符号化技術により一部の周波数帯域に対するビット割り当てが0とされることがある、複数の周波数帯域それぞれごとのオーディオデータである前記第1のデータを、前記第1のデータにおいてビット割り当てが0とされた前記第1のデータにおいて抜けている周波数帯域のオーディオデータである前記第2のデータに変換する予測演算ステップと

をコンピュータに実行させるためのプログラム。

【請求項8】

時間領域のオーディオデータを周波数領域のオーディオデータに変換することにより得られる第1のデータを、前記第1のデータよりも高品質のデータである第2のデータに変換するデータ処理を、コンピュータに行わせるプログラムが記録されている記録媒体であって、

前記第2のデータのうちの、注目しているデータである注目データをクラス分けするクラス分類に用いる、前記注目データに近接する複数のデータからなるクラスタップを、前記第1のデータから抽出するクラスタップ抽出ステップと、

前記クラスタップを構成する複数のデータの値から、前記クラスに対応するクラスコードを算出することにより、そのクラスコードに対応するクラスに、前記注目データをクラス分類するクラス分類ステップと、

前記注目データを求める積和演算に用いる、前記注目データに近接する複数のデータからなる予測タップを、前記第1のデータから抽出する予測タップ抽出ステップと、

学習の教師となる、前記第2のデータに対応する教師データと、前記学習の生徒となる、前記第1のデータに対応する生徒データを用い、注目している前記教師データの予測値を、その教師データに近接する複数の前記生徒データとタップ係数との積和演算により求め、注目している前記教師データと、求められた前記予測値との差分値である予測誤差を最小にする学習を、前記1以上のクラスごとに行うことにより求められたタップ係数から、前記注目データのクラスのタップ係数を取得する取得ステップと、

前記注目データのクラスのタップ係数と、前記予測タップとを用いた積和演算を行うことにより前記注目データを求めることで、知覚符号化技術または聴覚心理符号化技術により一部の周波数帯域に対するビット割り当てが0とされることがある、複数の周波数帯域それぞれごとのオーディオデータである前記第1のデータを、前記第1のデータにおいてビット割り当てが0とされた前記第1のデータにおいて抜けている周波数帯域のオーディオデータである前記第2のデータに変換する予測演算ステップと

をコンピュータに実行させるためのプログラムが記録されている

ことを特徴とする記録媒体。

【請求項 9】

時間領域のオーディオデータを周波数領域のオーディオデータに変換することにより得られる第 1 のデータを、前記第 1 のデータよりも高品質のデータである第 2 のデータに変換するのに用いられるタップ係数を学習するデータ処理装置であって、

前記第 2 のデータのうちの、注目しているデータである注目データをクラス分けするクラス分類に用いる、前記注目データに近接する複数のデータからなるクラスタップを、前記第 1 のデータから抽出するクラスタップ抽出手段と、

前記クラスタップを構成する複数のデータの値から、前記クラスに対応するクラスコードを算出することにより、そのクラスコードに対応するクラスに、前記注目データをクラス分類するクラス分類手段と、

前記タップ係数の学習の教師となる、前記第 2 のデータに対応する教師データのうちの注目している注目データを求めるのに用いる、前記注目データに近接する複数のデータからなる予測タップを、前記学習の生徒となる、前記第 1 のデータに対応する生徒データから抽出する予測タップ抽出手段と、

前記注目データと予測タップを用い、前記注目データの予測値を、その注目データに近接する複数の前記生徒データとタップ係数との積和演算により求め、前記注目データと、求められた前記予測値との差分値である予測誤差を最小にする学習を、前記 1 以上のクラスごとに行うことにより、前記タップ係数を求める学習手段と

を備え、

前記第 1 のデータは、知覚符号化技術または聴覚心理符号化技術により一部の周波数帯域に対するビット割り当てが 0 とされることがある、複数の周波数帯域それぞれごとのオーディオデータであり、

前記学習手段は、前記第 1 のデータを、前記第 1 のデータにおいてビット割り当てが 0 とされた前記第 1 のデータにおいて抜けている周波数帯域のオーディオデータである前記第 2 のデータに変換するのに用いられる前記タップ係数を求める

ことを特徴とするデータ処理装置。

【請求項 10】

前記学習手段は、前記第 1 のデータを、複数の周波数帯域それぞれごとのオーディオデータである前記第 2 のデータに変換するのに用いられる前記タップ係数を求める

ことを特徴とする請求項 9 に記載のデータ処理装置。

【請求項 11】

時間領域のオーディオデータを周波数領域のオーディオデータに変換することにより得られる第 1 のデータを、前記第 1 のデータよりも高品質のデータである第 2 のデータに変換するのに用いられるタップ係数を学習するデータ処理方法であって、

前記第 2 のデータのうちの、注目しているデータである注目データをクラス分けするクラス分類に用いる、前記注目データに近接する複数のデータからなるクラスタップを、前記第 1 のデータから抽出するクラスタップ抽出ステップと、

前記クラスタップを構成する複数のデータの値から、前記クラスに対応するクラスコードを算出することにより、そのクラスコードに対応するクラスに、前記注目データをクラス分類するクラス分類ステップと、

前記タップ係数の学習の教師となる、前記第 2 のデータに対応する教師データのうちの注目している注目データを求めるのに用いる、前記注目データに近接する複数のデータからなる予測タップを、前記学習の生徒となる、前記第 1 のデータに対応する生徒データから抽出する予測タップ抽出ステップと、

前記注目データと予測タップを用い、前記注目データの予測値を、その注目データに近接する複数の前記生徒データとタップ係数との積和演算により求め、前記注目データと、求められた前記予測値との差分値である予測誤差を最小にする学習を、前記 1 以上のクラスごとに行うことにより、前記タップ係数を求める学習ステップと

を備え、

前記第 1 のデータは、知覚符号化技術または聴覚心理符号化技術により一部の周波数帯域に対するビット割り当てが 0 とされることがある、複数の周波数帯域それぞれごとのオーディオデータであり、

前記学習ステップにおいて、前記第 1 のデータにおいてビット割り当てが 0 とされた前記第 1 のデータにおいて抜けている周波数帯域のオーディオデータである前記第 2 のデータに変換するのに用いられる前記タップ係数が求められる

ことを特徴とするデータ処理方法。

【請求項 1 2】

時間領域のオーディオデータを周波数領域のオーディオデータに変換することにより得られる第 1 のデータを、前記第 1 のデータよりも高品質のデータである第 2 のデータに変換するのに用いられるタップ係数を学習するデータ処理を、コンピュータに行わせるプログラムであって、

10

前記第 2 のデータのうちの、注目しているデータである注目データをクラス分けするクラス分類に用いる、前記注目データに近接する複数のデータからなるクラスタップを、前記第 1 のデータから抽出するクラスタップ抽出ステップと、

前記クラスタップを構成する複数のデータの値から、前記クラスに対応するクラスコードを算出することにより、そのクラスコードに対応するクラスに、前記注目データをクラス分類するクラス分類ステップと、

前記タップ係数の学習の教師となる、前記第 2 のデータに対応する教師データのうちの注目している注目データを求めるのに用いる、前記注目データに近接する複数のデータからなる予測タップを、前記学習の生徒となる、前記第 1 のデータに対応する生徒データから抽出する予測タップ抽出ステップと、

20

前記注目データと予測タップを用い、前記注目データの予測値を、その注目データに近接する複数の前記生徒データとタップ係数との積和演算により求め、前記注目データと、求められた前記予測値との差分値である予測誤差を最小にする学習を、前記 1 以上のクラスごとに行うことにより、前記タップ係数を求める学習ステップと

をコンピュータに実行させ、

前記第 1 のデータは、知覚符号化技術または聴覚心理符号化技術により一部の周波数帯域に対するビット割り当てが 0 とされることがある、複数の周波数帯域それぞれごとのオーディオデータであり、

30

前記学習ステップにおいて、前記第 1 のデータを、前記第 1 のデータにおいてビット割り当てが 0 とされた前記第 1 のデータにおいて抜けている周波数帯域のオーディオデータである前記第 2 のデータに変換するのに用いられる前記タップ係数が求められる

ことを特徴とするプログラム。

【請求項 1 3】

時間領域のオーディオデータを周波数領域のオーディオデータに変換することにより得られる第 1 のデータを、前記第 1 のデータよりも高品質のデータである第 2 のデータに変換するのに用いられるタップ係数を学習するデータ処理を、コンピュータに行わせるプログラムが記録されている記録媒体であって、

前記第 2 のデータのうちの、注目しているデータである注目データをクラス分けするクラス分類に用いる、前記注目データに近接する複数のデータからなるクラスタップを、前記第 1 のデータから抽出するクラスタップ抽出ステップと、

40

前記クラスタップを構成する複数のデータの値から、前記クラスに対応するクラスコードを算出することにより、そのクラスコードに対応するクラスに、前記注目データをクラス分類するクラス分類ステップと、

前記タップ係数の学習の教師となる、前記第 2 のデータに対応する教師データのうちの注目している注目データを求めるのに用いる、前記注目データに近接する複数のデータからなる予測タップを、前記学習の生徒となる、前記第 1 のデータに対応する生徒データから抽出する予測タップ抽出ステップと、

前記注目データと予測タップを用い、前記注目データの予測値を、その注目データに近

50

接する複数の前記生徒データとタップ係数との積和演算により求め、前記注目データと、求められた前記予測値との差分値である予測誤差を最小にする学習を、前記1以上のクラスごとに行うことにより、前記タップ係数を求める学習ステップと

をコンピュータに実行させ、

前記第1のデータは、知覚符号化技術または聴覚心理符号化技術により一部の周波数帯域に対するビット割り当てが0とされることがある、複数の周波数帯域それぞれごとのオーディオデータであり、

前記学習ステップにおいて、前記第1のデータを、前記第1のデータにおいてビット割り当てが0とされた前記第1のデータにおいて抜けている周波数帯域のオーディオデータである前記第2のデータに変換するのに用いられる前記タップ係数が求められる

10

プログラムが記録されている

ことを特徴とする記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、データ処理装置およびデータ処理方法、並びにプログラムおよび記録媒体に関し、特に、音質の良いオーディオデータを提供することができるようにするデータ処理装置およびデータ処理方法、並びにプログラムおよび記録媒体に関する。

【0002】

【従来の技術】

20

オーディオデータの符号化方法としては、知覚符号化技術または聴覚心理符号化技術により、データ量を削減するものが知られている。知覚符号化技術または聴覚心理符号化技術による符号化方法は、簡単には、人間が知覚することのできる音圧と時間との関係、または周波数と音量との関係を利用して、（ほとんど）知覚することができないオーディオデータは、符号化のためのビット割り当てを0としてマスクする、つまり符号化の対象としないもので、そのような符号化方法としては、例えば、MPEG(Moving Picture Experts Group)オーディオがある。

【0003】

ここで、オーディオデータの符号化方法としてのMPEGオーディオ(MPEG Audio)には、MPEG1オーディオレイヤ1, 2, 3(MPEG1 Audio Layer1, Layer2, Layer3)や、MPEG2オーディオAAC(MPEG2 Audio Advanced Audio Coding)などがある。

30

【0004】

なお、MPEGオーディオについては、例えば、「最新MPEG教科書」、株式会社アスキー発行などに、その詳細が記載されている。また、MPEG1オーディオレイヤ3については、例えば、特開2000-323993号公報、特開2001-102932号公報、特許2923406号公報などに、MPEG2オーディオAACについては、例えば、特開平11-259096号公報などにも記載されている。

【0005】

図1は、MPEG1オーディオレイヤ1方式により、オーディオデータを符号化するエンコーダの構成例を示している。

40

【0006】

エンコーダには、16ビット直線量子化されたPCM(Pulse Code Modulation)信号としての時間領域のオーディオデータが入力される。エンコーダに入力されたオーディオデータは、サブバンド分割部1と聴覚心理分析部2に供給される。

【0007】

サブバンド分割部1は、例えば、ポリフェーズフィルタバンクで構成され、そのポリフェーズフィルタバンクでオーディオデータをフィルタリングすることにより、オーディオデータをサブバンド符号化(サブバンド合成)する。即ち、サブバンド分割部1は、そこに供給されるオーディオデータの384サンプルを1フレームとしてサブバンド符号化することにより、その1フレームの時間領域のオーディオデータを、32の同一周波数帯域幅

50

のバンド（サブバンド）の周波数領域のオーディオデータとする。

【 0 0 0 8 】

ここで、サブバンド分割部 1 では、384 サンプルのオーディオデータが、32 のサブバンドのオーディオデータとされるので、1 サブバンドのサンプル数は、12 (= 384 / 32) サンプルとなる。

【 0 0 0 9 】

なお、MPEG1オーディオレイヤ2方式でも、MPEG1オーディオレイヤ1方式と同様に、オーディオデータは、ポリフェーズフィルタバンクによるフィルタリングによってサブバンド符号化されることにより、32 のサブバンドのオーディオデータとされる。また、MPEG1オーディオレイヤ3方式では、オーディオデータに対しては、ポリフェーズフィルタバンクによるフィルタリングが行われる他、MDCT(Modified Discrete Cosine Transform)演算、および折り返し歪み削減バタフライ演算が施され、やはり、32 バンドの周波数領域のオーディオデータとされる。

10

【 0 0 1 0 】

サブバンド分割部 1 で得られた32 のサブバンドの周波数領域のオーディオデータ（以下、適宜、サブバンドデータという）は、聴覚心理分析部 2 と量子化 / 符号化部 4 に供給される。

【 0 0 1 1 】

聴覚心理分析部 2 は、聴覚心理モデル計算部 1 1 とスケーリングファクタ計算部 1 2 などで構成され、信号対マスク比(Signal to Mask Ratio)やスケーリングファクタなどを求める。

20

【 0 0 1 2 】

即ち、聴覚心理モデル計算部 1 1 は、所定の聴覚心理モデルを用い、オーディオデータをFFT(Fast Fourier Transform)分析したり、各サブバンドの音圧を計算することなどによって、各サブバンドについて、そのサブバンドにおけるオーディオデータの最大振幅レベルと、人間の聴覚または知覚特性に基づいてマスキングされる最大振幅レベル（マスキング閾値）との比である信号対マスク比を計算し、ビット割り当て部 3 に供給する。また、スケーリングファクタ計算部 1 2 は、各サブバンドのサブバンドデータなどに基づき、各サブバンドのスケーリングファクタを計算し、ビット割り当て部 3 に供給する。

【 0 0 1 3 】

30

ビット割り当て部 3 は、聴覚心理分析部 2（の聴覚心理モデル計算部 1 1）から供給される信号対マスク比などに基づいて、サブバンド分割部 1 が出力する各サブバンドのサブバンドデータに対するビット割り当てを決定し、そのビット割り当てを表す情報（ビット割り当て情報）を、量子化 / 符号化部 4 とビットストリーム生成部 5 に供給する。さらに、ビット割り当て部 3 は、聴覚心理分析部 2（のスケーリングファクタ計算部 1 2）から供給されるスケーリングファクタを、量子化 / 符号化部 4 とビットストリーム生成部 5 に供給する。

【 0 0 1 4 】

量子化 / 符号化部 4 は、サブバンド分割部 1 から供給される各サブバンドのサブバンドデータを、ビット割り当て部 3 から供給されるスケーリングファクタで正規化する。さらに、量子化 / 符号化部 4 は、その正規化後の各サブバンドのサブバンドデータを、ビット割り当て部 3 から供給されるビット割り当て情報に対応する量子化ステップで量子化し、ビットストリーム生成部 5 に供給する。

40

【 0 0 1 5 】

ここで、MPEG1オーディオレイヤ2方式でも、MPEG1オーディオレイヤ1方式と同様に、量子化 / 符号化部 4 において、各サブバンドのサブバンドデータが量子化される。また、MPEG1オーディオレイヤ3方式では、量子化 / 符号化部 4 において、各サブバンドのサブバンドデータが量子化された後、ハフマン符号化される。

【 0 0 1 6 】

ビットストリーム生成部 5 は、量子化 / 符号化部 4 から供給される、量子化されたサブバ

50

ンドデータ（以下、適宜、量子化サブバンドデータという）、並びにビット割り当て部 3 から供給されるビット割り当て情報およびスケールファクタを多重化し、さらに、ヘッダを付加することにより、オーディオデータを符号化した符号化ビットストリームを生成して出力する。

【 0 0 1 7 】

次に、図 2 は、オーディオデータを MPEG1 オーディオレイヤ 1 方式で符号化した符号化ビットストリーム、即ち、図 1 のエンコーダが出力する符号化ビットストリームを復号するデコーダの構成例を示している。

【 0 0 1 8 】

符号化ビットストリームは、ビットストリーム分解部 2 1 に供給される。ビットストリーム分解部 2 1 は、符号化ビットストリームを、ヘッダ、量子化サブバンドデータ、ビット割り当て情報、およびスケールファクタに分解し、再構成部 2 2 に供給する。

10

【 0 0 1 9 】

再構成部 2 2 は、ビット割り当て復号部 3 1 および復号 / 逆量子化部 3 2 で構成され、ビットストリーム分解部 2 1 の出力から、各サブバンドのサブバンドデータを再構成（復号）する。

【 0 0 2 0 】

即ち、ビット割り当て復号部 3 1 は、ビットストリーム分解部 2 1 から、ビット割り当て情報とスケールファクタを受信し、ビット割り当て情報が表す量子化ステップとスケールファクタを、復号 / 逆量子化部 3 2 に供給する。復号 / 逆量子化部 3 2 は、ビットストリーム分解部 2 1 から、各サブバンドの量子化サブバンドデータを受信し、その量子化サブバンドデータを、ビット割り当て部 3 1 から供給される量子化ステップで逆量子化する。さらに、復号 / 逆量子化部 3 2 は、量子化サブバンドデータの逆量子化結果に、スケールファクタを乗算すること等により、3 2 のサブバンドのサブバンドデータを復号し、サブバンド合成部 2 3 に供給する。

20

【 0 0 2 1 】

サブバンド合成部 2 3 は、再構成部 2 2（の復号 / 逆量子化部 3 2）から供給される 3 2 のサブバンドのサブバンドデータを合成し、これにより、1 フレームの時間領域のオーディオデータを復号して出力する。

【 0 0 2 2 】

次に、図 3 は、MPEG2 オーディオ AAC 方式により、オーディオデータを符号化するエンコーダの構成例を示している。

30

【 0 0 2 3 】

オーディオデータは、聴覚心理分析部 4 1 とフィルタバンク 4 2 に供給される。

【 0 0 2 4 】

聴覚心理分析部 4 1 は、そこに供給されるオーディオデータを FFT することにより、パワースペクトルを求め、そのパワースペクトルに基づき、複数のバンド（周波数帯域）それぞれごとに許容される量子化雑音を求める。さらに、聴覚心理分析部 4 1 は、その量子化雑音に基づいて、人間の聴覚または知覚特性によって雑音がマスキングされるように量子化を行うために 1 フレームのオーディオデータに割り当てべき割り当てビット数を表すビット割り当て情報を求め、フィルタバンク 4 2 に供給する。

40

【 0 0 2 5 】

フィルタバンク 4 2 は、MDCT を行うための複数のフィルタで構成され、聴覚心理分析部 4 1 から供給されるビット割り当て情報に基づいて、1 フレームのオーディオデータを MDCT 処理する際のブロック長を決定し、そのブロック長ごとに、オーディオデータをフィルタリングすることで、そのオーディオデータを MDCT 処理（適応ブロック長 MDCT 処理）する。フィルタバンク 4 2 においてオーディオデータを MDCT 処理することにより得られる、複数の周波数帯域それぞれごとのオーディオデータとしての MDCT 係数（周波数領域のオーディオデータ）は、演算部 4 3 に供給される。

【 0 0 2 6 】

50

演算部 4 3 は、予測部 5 0 より供給される MDCT 係数の予測値について、フィルタバンク 4 2 から供給される MDCT 係数の真値から、MDCT 係数の予測値を減算した予測誤差を求め、その MDCT 係数の予測誤差を、量子化部 4 4 に供給する。量子化部 4 4 は、演算部 4 3 からの MDCT 係数の予測誤差を量子化し、その結果得られる量子化予測誤差を、符号化部 4 5 と逆量子化部 4 7 に供給する。

【 0 0 2 7 】

符号化部 4 5 は、量子化部 4 4 からの量子化予測誤差をハフマン符号化し、その結果得られるハフマン符号化データを、ビットストリーム生成部 4 6 に供給する。ビットストリーム生成部 4 6 は、符号化部 4 5 から供給されるハフマン符号化データに、ヘッダその他の情報を多重化し、これにより、オーディオデータを符号化した符号化ビットストリームを生成して出力する。

10

【 0 0 2 8 】

一方、逆量子化部 4 7 は、量子化部 4 4 の出力を逆量子化することにより、MDCT 係数の予測誤差を求め、演算部 4 8 に供給する。演算部 4 8 は、逆量子化部 4 7 から供給される MDCT 係数の予測誤差と、予測部 5 0 が出力する MDCT 係数の予測値とを加算することにより、MDCT 係数をローカルデコードし、遅延部 4 9 に供給する。遅延部 4 9 は、演算部 4 8 から供給される MDCT 係数を所定の時間だけ遅延し、予測部 5 0 に供給する。

【 0 0 2 9 】

予測部 5 0 は、遅延部 4 9 から供給される MDCT 係数を用いて線形予測を行うことにより、フィルタバンク 4 2 が出力する MDCT 係数の予測値を求め、演算部 4 3 に供給する。

20

【 0 0 3 0 】

演算部 4 3 では、このようにして予測部 5 0 から供給される MDCT 係数の予測値について、上述したように予測誤差が求められる。

【 0 0 3 1 】

次に、図 4 は、オーディオデータを MPEG2 オーディオ AAC 方式で符号化した符号化ビットストリーム、即ち、図 3 のエンコーダが出力する符号化ビットストリームを復号するデコーダの構成例を示している。

【 0 0 3 2 】

符号化ビットストリームは、ビットストリーム分解部 6 1 に供給される。ビットストリーム分解部 6 1 は、符号化ビットストリームを、ヘッダその他の情報と、ハフマン符号化データに分解し、復号部 6 2 に供給する。

30

【 0 0 3 3 】

復号部 6 2 は、ビットストリーム分解部 6 1 から供給されるハフマン符号化データを量子化予測誤差に復号し、逆量子化部 6 3 に供給する。逆量子化部 6 3 は、復号部 6 2 から供給される量子化予測誤差を逆量子化し、これにより、MDCT 係数の予測誤差を得て、演算部 6 4 に供給する。

【 0 0 3 4 】

演算部 6 4 は、図 3 の演算部 4 8 と同様に、逆量子化部 6 3 から供給される MDCT 係数の予測誤差と、予測部 6 7 が出力する MDCT 係数の予測値とを加算することにより、MDCT 係数を復号し、フィルタバンク 6 5 と遅延部 6 6 に供給する。

40

【 0 0 3 5 】

フィルタバンク 6 5 は、逆 MDCT を行うための複数のフィルタで構成され、演算部 6 4 から供給される MDCT 係数を逆 MDCT 処理することにより、複数の周波数帯域それぞれごとのオーディオデータを合成した時間領域のオーディオデータを求めて出力する。

【 0 0 3 6 】

一方、遅延部 6 6 は、図 3 の遅延部 4 9 と同様に、演算部 6 4 から供給される MDCT 係数を所定の時間だけ遅延し、予測部 6 7 に供給する。

【 0 0 3 7 】

予測部 6 7 は、図 3 の予測部 5 0 と同様に、遅延部 6 6 から供給される MDCT 係数を用いて線形予測を行うことにより、逆量子化部 6 3 が予測誤差を出力する MDCT 係数の予測値を求

50

め、演算部 6 4 に供給する。

【 0 0 3 8 】

演算部 6 4 では、このようにして予測部 6 7 から供給される MDCT 係数の予測値が、上述したように、逆量子化部 6 3 から供給される MDCT 係数の予測誤差と加算されることにより、MDCT 係数が復号される。

【 0 0 3 9 】

【発明が解決しようとする課題】

以上のように、知覚符号化技術または聴覚心理符号化技術により、データ量を削減する MP EG オーディオ方式などでは、信号対マスク比に基づくビット数や、人間の聴覚または知覚特性によって雑音がマスキングされるように量子化を行うために必要なビット数が、オーディオデータに割り当てられる。従って、周波数帯域によっては、サブバンドデータまたは MDCT 係数としての周波数領域のオーディオデータに対するビット割り当てが 0 とされることがある。

10

【 0 0 4 0 】

このように、ビット割り当てが 0 とされた周波数帯域のオーディオデータは、図 1 または図 3 のエンコーダが出力する符号化ビットストリームに含まれないから、図 2 または図 4 のデコーダでは、その周波数帯域の周波数成分が存在しないオーディオデータが復号される。従って、復号されたオーディオデータの音質が、少なからず劣化する課題があった。

【 0 0 4 1 】

かかるオーディオデータの音質の劣化を低減（防止）する方法としては、各周波数帯域のオーディオデータに対して、必ず、1 ビット以上を割り当てる方法があるが、この方法では、符号化ビットストリームのデータ量が増加することとなる。

20

【 0 0 4 2 】

本発明は、このような状況に鑑みてなされたものであり、符号化ビットストリームのデータ量を増加させずに、高音質のオーディオデータを復号することができるようにするものである。

【 0 0 4 3 】

【課題を解決するための手段】

本発明の第 1 のデータ処理装置は、第 2 のデータのうちの、注目しているデータである注目データをクラス分けするクラス分類に用いる、注目データに近接する複数のデータからなるクラスタップを、第 1 のデータから抽出するクラスタップ抽出手段と、クラスタップを構成する複数のデータの値から、クラスに対応するクラスコードを算出することにより、そのクラスコードに対応するクラスに、注目データをクラス分類するクラス分類手段と、注目データを求める積和演算に用いる、注目データに近接する複数のデータからなる予測タップを、第 1 のデータから抽出する予測タップ抽出手段と、学習の教師となる、第 2 のデータに対応する教師データと、学習の生徒となる、第 1 のデータに対応する生徒データを用い、注目している教師データの予測値を、その教師データに近接する複数の生徒データとタップ係数との積和演算により求め、注目している教師データと、求められた予測値との差分値である予測誤差を最小にする学習を、1 以上のクラスごとに行うことにより求められたタップ係数から、注目データのクラスのタップ係数を取得する取得手段と、
注目データのクラスのタップ係数と、予測タップとを用いた積和演算を行うことにより注目データを求めることで、知覚符号化技術または聴覚心理符号化技術により一部の周波数帯域に対するビット割り当てが 0 とされることがある、複数の周波数帯域それぞれごとのオーディオデータである第 1 のデータを、第 1 のデータにおいてビット割り当てが 0 とされた第 1 のデータにおいて抜けている周波数帯域のオーディオデータである第 2 のデータに変換する予測演算手段とを備えることを特徴とする。

30

40

【 0 0 4 4 】

本発明の第 1 のデータ処理方法は、第 2 のデータのうちの、注目しているデータである注目データをクラス分けするクラス分類に用いる、注目データに近接する複数のデータからなるクラスタップを、第 1 のデータから抽出するクラスタップ抽出ステップと、クラス

50

タップを構成する複数のデータの値から、クラスに対応するクラスコードを算出することにより、そのクラスコードに対応するクラスに、注目データをクラス分類するクラス分類ステップと、注目データを求める積和演算に用いる、注目データに近接する複数のデータからなる予測タップを、第1のデータから抽出する予測タップ抽出ステップと、学習の教師となる、第2のデータに対応する教師データと、学習の生徒となる、第1のデータに対応する生徒データを用い、注目している教師データの予測値を、その教師データに近接する複数の生徒データとタップ係数との積和演算により求め、注目している教師データと、求められた予測値との差分値である予測誤差を最小にする学習を、1以上のクラスごとに行うことにより求められたタップ係数から、注目データのクラスのタップ係数を取得する取得ステップと、注目データのクラスのタップ係数と、予測タップとを用いた積和演算を行うことにより注目データを求めることで、知覚符号化技術または聴覚心理符号化技術により一部の周波数帯域に対するビット割り当てが0とされることがある、複数の周波数帯域それぞれごとのオーディオデータである第1のデータを、第1のデータにおいてビット割り当てが0とされた第1のデータにおいて抜けている周波数帯域のオーディオデータである第2のデータに変換する予測演算ステップとを備えることを特徴とする。

10

【0045】

本発明の第1のプログラムは、第2のデータのうちの、注目しているデータである注目データをクラス分けするクラス第2のデータのうちの、注目しているデータである注目データをクラス分けするクラス分類に用いる、注目データに近接する複数のデータからなるクラスタップを、第1のデータから抽出するクラスタップ抽出ステップと、クラスタップを構成する複数のデータの値から、クラスに対応するクラスコードを算出することにより、そのクラスコードに対応するクラスに、注目データをクラス分類するクラス分類ステップと、注目データを求める積和演算に用いる、注目データに近接する複数のデータからなる予測タップを、第1のデータから抽出する予測タップ抽出ステップと、学習の教師となる、第2のデータに対応する教師データと、学習の生徒となる、第1のデータに対応する生徒データを用い、注目している教師データの予測値を、その教師データに近接する複数の生徒データとタップ係数との積和演算により求め、注目している教師データと、求められた予測値との差分値である予測誤差を最小にする学習を、1以上のクラスごとに行うことにより求められたタップ係数から、注目データのクラスのタップ係数を取得する取得ステップと、注目データのクラスのタップ係数と、予測タップとを用いた積和演算を行うことにより注目データを求めることで、知覚符号化技術または聴覚心理符号化技術により一部の周波数帯域に対するビット割り当てが0とされることがある、複数の周波数帯域それぞれごとのオーディオデータである第1のデータを、第1のデータにおいてビット割り当てが0とされた第1のデータにおいて抜けている周波数帯域のオーディオデータである第2のデータに変換する予測演算ステップとを備えることを特徴とする。

20

30

【0046】

本発明の第1の記録媒体は、第2のデータのうちの、注目しているデータである注目データをクラス分けするクラス分類に用いる、注目データに近接する複数のデータからなるクラスタップを、第1のデータから抽出するクラスタップ抽出ステップと、クラスタップを構成する複数のデータの値から、クラスに対応するクラスコードを算出することにより、そのクラスコードに対応するクラスに、注目データをクラス分類するクラス分類ステップと、注目データを求める積和演算に用いる、注目データに近接する複数のデータからなる予測タップを、第1のデータから抽出する予測タップ抽出ステップと、学習の教師となる、第2のデータに対応する教師データと、学習の生徒となる、第1のデータに対応する生徒データを用い、注目している教師データの予測値を、その教師データに近接する複数の生徒データとタップ係数との積和演算により求め、注目している教師データと、求められた予測値との差分値である予測誤差を最小にする学習を、1以上のクラスごとに行うことにより求められたタップ係数から、注目データのクラスのタップ係数を取得する取得ステップと、注目データのクラスのタップ係数と、予測タップとを用いた積和演算を行うことにより注目データを求めることで、知覚符号化技術または聴覚心理符号化技術により一

40

50

部の周波数帯域に対するビット割り当てが0とされることがある、複数の周波数帯域それぞれごとのオーディオデータである第1のデータを、第1のデータにおいてビット割り当てが0とされた第1のデータにおいて抜けている周波数帯域のオーディオデータである第2のデータに変換する予測演算ステップとを備えるプログラムが記録されていることを特徴とする。

【0047】

本発明の第2のデータ処理装置は、第2のデータのうちの、注目しているデータである注目データをクラス分けするクラス分類に用いる、注目データに近接する複数のデータからなるクラスタップを、第1のデータから抽出するクラスタップ抽出手段と、クラスタップを構成する複数のデータの値から、クラスに対応するクラスコードを算出することにより、そのクラスコードに対応するクラスに、注目データをクラス分類するクラス分類手段と、タップ係数の学習の教師となる、第2のデータに対応する教師データのうちの注目している注目データを求めるのに用いる、注目データに近接する複数のデータからなる予測タップを、学習の生徒となる、第1のデータに対応する生徒データから抽出する予測タップ抽出手段と、注目データと予測タップを用い、注目データの予測値を、その注目データに近接する複数の生徒データとタップ係数との積和演算により求め、注目データと、求められた予測値との差分値である予測誤差を最小にする学習を、1以上のクラスごとに行うことにより、タップ係数を求める学習手段とを備え、第1のデータは、知覚符号化技術または聴覚心理符号化技術により一部の周波数帯域に対するビット割り当てが0とされることがある、複数の周波数帯域それぞれごとのオーディオデータであり、学習手段は、第1

10

20

【0048】

本発明の第2のデータ処理方法は、時間領域のオーディオデータを周波数領域のオーディオデータに変換することにより得られる第1のデータを、第1のデータよりも高品質のデータである第2のデータに変換するのに用いられるタップ係数を学習するデータ処理方法であって、第2のデータのうちの、注目しているデータである注目データをクラス分けするクラス分類に用いる、注目データに近接する複数のデータからなるクラスタップを、第1のデータから抽出するクラスタップ抽出ステップと、クラスタップを構成する複数のデータの値から、クラスに対応するクラスコードを算出することにより、そのクラスコードに対応するクラスに、注目データをクラス分類するクラス分類ステップと、タップ係数の学習の教師となる、第2のデータに対応する教師データのうちの注目している注目データを求めるのに用いる、注目データに近接する複数のデータからなる予測タップを、学習の生徒となる、第1のデータに対応する生徒データから抽出する予測タップ抽出ステップと、注目データと予測タップを用い、注目データの予測値を、その注目データに近接する複数の生徒データとタップ係数との積和演算により求め、注目データと、求められた予測値との差分値である予測誤差を最小にする学習を、1以上のクラスごとに行うことにより、タップ係数を求める学習ステップとを備え、第1のデータは、知覚符号化技術または聴覚心理符号化技術により一部の周波数帯域に対するビット割り当てが0とされることがある、複数の周波数帯域それぞれごとのオーディオデータであり、学習ステップにおいて、第1のデータを、第1のデータにおいてビット割り当てが0とされた第1のデータにおいて抜けている周波数帯域のオーディオデータである第2のデータに変換するのに用いられるタップ係数が求められることを特徴とする。

30

40

【0049】

本発明の第2のプログラムは、第2のデータのうちの、注目しているデータである注目データをクラス分けするクラス分類に用いる、注目データに近接する複数のデータからなるクラスタップを、第1のデータから抽出するクラスタップ抽出ステップと、クラスタップを構成する複数のデータの値から、クラスに対応するクラスコードを算出することにより、そのクラスコードに対応するクラスに、注目データをクラス分類するクラス分類ステ

50

ップと、タップ係数の学習の教師となる、第2のデータに対応する教師データのうちの注目している注目データを求めるのに用いる、注目データに近接する複数のデータからなる予測タップを、学習の生徒となる、第1のデータに対応する生徒データから抽出する予測タップ抽出ステップと、注目データと予測タップを用い、注目データの予測値を、その注目データに近接する複数の生徒データとタップ係数との積和演算により求め、注目データと、求められた予測値との差分値である予測誤差を最小にする学習を、1以上のクラスごとに行うことにより、タップ係数を求める学習ステップとを備え、第1のデータは、知覚符号化技術または聴覚心理符号化技術により一部の周波数帯域に対するビット割り当てが0とされることがある、複数の周波数帯域それぞれごとのオーディオデータであり、学習ステップにおいて、第1のデータを、第1のデータにおいてビット割り当てが0とされた第1のデータにおいて抜けている周波数帯域のオーディオデータである第2のデータに変換するのに用いられるタップ係数が求められることを特徴とする。

10

【0050】

本発明の第2の記録媒体は、第2のデータのうちの、注目しているデータである注目データをクラス分けするクラス分類に用いる、注目データに近接する複数のデータからなるクラスタップを、第1のデータから抽出するクラスタップ抽出ステップと、クラスタップを構成する複数のデータの値から、クラスに対応するクラスコードを算出することにより、そのクラスコードに対応するクラスに、注目データをクラス分類するクラス分類ステップと、タップ係数の学習の教師となる、第2のデータに対応する教師データのうちの注目している注目データを求めるのに用いる、注目データに近接する複数のデータからなる予測タップを、学習の生徒となる、第1のデータに対応する生徒データから抽出する予測タップ抽出ステップと、注目データと予測タップを用い、注目データの予測値を、その注目データに近接する複数の生徒データとタップ係数との積和演算により求め、注目データと、求められた予測値との差分値である予測誤差を最小にする学習を、1以上のクラスごとに行うことにより、タップ係数を求める学習ステップとを備え、第1のデータは、知覚符号化技術または聴覚心理符号化技術により一部の周波数帯域に対するビット割り当てが0とされることがある、複数の周波数帯域それぞれごとのオーディオデータであり、学習ステップにおいて、第1のデータを、第1のデータにおいてビット割り当てが0とされた第1のデータにおいて抜けている周波数帯域のオーディオデータである第2のデータに変換するのに用いられるタップ係数が求められるプログラムが記録されていることを特徴とする。

20

30

【0051】

本発明の第1のデータ処理装置およびデータ処理方法、並びにプログラムおよび記録媒体においては、第2のデータのうちの、注目しているデータである注目データをクラス分けするクラス分類に用いる、注目データに近接する複数のデータからなるクラスタップが、第1のデータから抽出され、クラスタップを構成する複数のデータの値から、クラスに対応するクラスコードを算出することにより、そのクラスコードに対応するクラスに、注目データがクラス分類され、注目データを求める積和演算に用いる、注目データに近接する複数のデータからなる予測タップが、第1のデータから抽出される。また、学習の教師となる、第2のデータに対応する教師データと、学習の生徒となる、第1のデータに対応する生徒データを用い、注目している教師データの予測値を、その教師データに近接する複数の生徒データとタップ係数との積和演算により求め、注目している教師データと、求められた予測値との差分値である予測誤差を最小にする学習を、1以上のクラスごとに行うことにより求められたタップ係数から、注目データのクラスのタップ係数が取得される。そして、その注目データのクラスのタップ係数と、予測タップとを用いた積和演算を行うことにより注目データを求めることで、知覚符号化技術または聴覚心理符号化技術により一部の周波数帯域に対するビット割り当てが0とされることがある、複数の周波数帯域それぞれごとのオーディオデータである第1のデータが、第1のデータにおいてビット割り当てが0とされた第1のデータにおいて抜けている周波数帯域のオーディオデータである第2のデータに変換される。

40

50

【 0 0 5 2 】

本発明の第2のデータ処理装置およびデータ処理方法、並びにプログラムおよび記録媒体においては、第2のデータのうちの、注目しているデータである注目データをクラス分けするクラス分類に用いる、注目データに近接する複数のデータからなるクラスタップが、第1のデータから抽出され、クラスタップを構成する複数のデータの値から、クラスに対応するクラスコードを算出することにより、そのクラスコードに対応するクラスに、注目データがクラス分類される。また、タップ係数の学習の教師となる、第2のデータに対応する教師データのうちの注目している注目データを求めるのに用いる、注目データに近接する複数のデータからなる予測タップが、学習の生徒となる、第1のデータに対応する生徒データから抽出される。そして、注目データと予測タップを用い、注目データの予測値を、その注目データに近接する複数の生徒データとタップ係数との積和演算により求め、注目データと、求められた予測値との差分値である予測誤差を最小にする学習を、1以上のクラスごとに行うことにより、タップ係数が求められる。また、第1のデータは、知覚符号化技術または聴覚心理符号化技術により一部の周波数帯域に対するビット割り当てが0とされることがある、複数の周波数帯域それぞれごとのオーディオデータであり、学習において、第1のデータを、第1のデータにおいてビット割り当てが0とされた第1のデータにおいて抜けている周波数帯域のオーディオデータである第2のデータに変換するのに用いられるタップ係数が求められる。

10

【 0 0 5 3 】

【発明の実施の形態】

20

図5は、本発明を適用したデータ処理装置の一実施の形態の構成例を示している。

【 0 0 5 4 】

このデータ処理装置においては、ディジタルオーディオデータを符号化し、その結果得られる符号化ビットストリームが、ディジタルオーディオデータに復号されるようになっている。

【 0 0 5 5 】

即ち、符号化対象のオーディオデータは、符号化装置101に供給される。符号化装置101は、知覚符号化技術または聴覚心理符号化技術により、データ量を削減するもので、そこに供給されるオーディオデータを所定の符号化方式で符号化し、符号化ビットストリームを出力する。

30

【 0 0 5 6 】

ここで、符号化装置101における符号化方式としては、例えば、前述したMPEG1オーディオレイヤ1, 2, 3やMPEG2オーディオAACなどを採用することができる。なお、ここでは、符号化装置101の符号化方式として、例えば、MPEG1オーディオレイヤ1を採用することとする。従って、符号化装置101は、前述の図1のエンコーダと同様に構成される。

【 0 0 5 7 】

符号化装置101が出力する符号化ビットストリームは、例えば、地上波や、衛星回線、CATV(Cable Television)、インターネット、電話回線、その他の有線または無線の伝送媒体103を介して伝送され、あるいは、例えば、半導体メモリや、光磁気ディスク、磁気ディスク、相変化ディスクなどの記録媒体104に記録される。

40

【 0 0 5 8 】

復号装置102には、伝送媒体103を介して伝送されてくる符号化ビットストリーム、あるいは、記録媒体104から再生された符号化ビットストリームが供給される。

【 0 0 5 9 】

復号装置102は、そこに供給される符号化ビットストリームをオーディオデータに復号して出力する。

【 0 0 6 0 】

なお、符号化装置101と復号装置102は、別々の筐体でなる、独立した装置として構成することもできるし、1つの筐体でなる1つの装置として構成することも可能である。

50

【 0 0 6 1 】

次に、図 6 は、図 5 の復号装置 1 0 2 の第 1 の構成例を示している。

【 0 0 6 2 】

復号装置 1 0 2 に供給される符号化ビットストリームは、デコーダ 1 1 1 で受信される。

【 0 0 6 3 】

デコーダ 1 1 1 は、符号化装置 1 0 1 (図 5) が出力する符号化ビットストリームを復号するデコーダと同様に構成されている。従って、本実施の形態では、符号化装置 1 0 1 は、MPEG1 オーディオレイヤ 1 方式でオーディオデータを符号化した符号化ビットストリームを出力するので、デコーダ 1 1 1 は、例えば、図 2 のデコーダと同様に構成されている。但し、デコーダ 1 1 1 は、図 2 のサブバンド合成部 2 3 で得られる、符号化ビットストリームを MPEG1 オーディオレイヤ 1 で復号した時間領域のオーディオデータではなく、再構成部 2 2 で得られる周波数領域のオーディオデータとしての各サブバンドごとのサブバンドデータを出力する。さらに、デコーダ 1 1 1 は、ビットストリーム分解部 2 1 において、符号化ビットストリームから得られた、量子化サブバンドデータ以外のデータ、即ち、ヘッダや、スケーリングファクタ、ビット割り当て情報を、サイドインフォメーションとして出力する。

10

【 0 0 6 4 】

なお、デコーダ 1 1 1 は、フレームごとに、各サブバンドのサブバンドデータとサイドインフォメーションを出力するが、サブバンドデータは、クラス分類適応処理部 1 1 3 とサブバンド合成部 1 1 4 に供給され、サイドインフォメーションは、コントローラ 1 1 2 と

20

【 0 0 6 5 】

コントローラ 1 1 2 は、クラス分類適応処理部 1 1 3 において求めようとするサブバンドデータを注目データとして選択する。さらに、コントローラ 1 1 2 は、注目データのサブバンドを表すサブバンドナンバと、そのサブバンドにおける注目データのデータ位置などを、注目データを特定するための特定情報として、クラス分類適応処理部 1 1 3 およびサブバンド合成部 1 1 4 に供給する。また、コントローラ 1 1 2 は、デコーダ 1 1 1 から供給されるサイドインフォメーションに基づき、各フレームにおいてサブバンドデータが抜けているサブバンド、即ち、符号化装置 1 0 1 (図 5) においてビット割り当てが 0 ビットとされたサブバンドがあるかどうかを判定し、さらに、ビット割り当てが 0 ビットとされたサブバンドがある場合には、そのサブバンドを表すサブバンドナンバを認識する。

30

【 0 0 6 6 】

ここで、本実施の形態では、符号化装置 1 0 1 (図 5) において、オーディオデータは、図 1 で説明したように、32 のサブバンドのサブバンドデータに符号化される。いま、この 32 のサブバンドを、低周波数帯域から、サブバンド sb0, sb1, . . . , sb31 というように、sb に、0 からのシーケンシャルな数字を付して表すこととすると、あるサブバンド sb #n のサブバンドナンバとは、例えば、sb に付されている数字 #n を表す。また、1 つのサブバンドには、図 1 で説明したように、12 のサブバンドデータが存在するが、そのサブバンドデータ (のサンプル) を、d0, d1, . . . , d12 というように、d に、0 からの数字を付して表すこととすると、あるサブバンドデータ d #i のデータ位置とは、例えば、d に付されている数字 #i を表す。

40

【 0 0 6 7 】

クラス分類適応処理部 1 1 3 は、デコーダ 1 1 1 から供給されるサブバンドデータ、およびコントローラ 1 1 2 から供給される特定情報としての注目データ (サブバンドデータ) のサブバンドナンバとデータ位置などを用いて、例えば、後述するクラス分類適応処理を行うことにより、注目データとしてのサブバンドデータを求め (予測し) 、サブバンド合成部 1 1 4 に供給する。

【 0 0 6 8 】

サブバンド合成部 1 1 4 は、デコーダ 1 1 1 から供給されるサブバンドのサブバンドデータと、クラス分類適応処理部 1 1 3 から供給されるサブバンドのサブバンドデータとを、

50

図 2 のサブバンド合成部 2 3 における場合と同様にして合成し、これにより、時間領域のオーディオデータを求めて出力する。

【 0 0 6 9 】

次に、図 6 のクラス分類適応処理部 1 1 3 において行われるクラス分類適応処理について説明する。

【 0 0 7 0 】

クラス分類適応処理は、クラス分類処理と適応処理とからなり、クラス分類処理によって、データを、その性質に基づいてクラス分けし、各クラスごとに適応処理を施すものであり、適応処理とは、以下のような手法の処理である。

【 0 0 7 1 】

即ち、適応処理では、第 1 のデータが、所定のタップ係数を用いてマッピング（写像）されることにより、第 2 のデータに変換される。

【 0 0 7 2 】

いま、このタップ係数を用いてのマッピング方法として、例えば、線形 1 次結合モデルを採用するとともに、第 1 のデータとして、デコーダ 1 1 1 が出力するサブバンドデータ、即ち、オーディオデータを MPEG1 オーディオレイヤ 1 方式により符号化し、さらに、MPEG1 オーディオレイヤ 1 方式により復号することにより得られるサブバンドデータを用い、第 2 のデータとして、そのサブバンドデータの真値、即ち、符号化装置 1 0 1 でオーディオデータをサブバンド符号化することにより得られるサブバンドデータを用いることとして、適応処理について説明する。

【 0 0 7 3 】

なお、オーディオデータを MPEG1 オーディオレイヤ 1 方式により符号化し、さらに、MPEG1 オーディオレイヤ 1 方式により復号することにより得られるサブバンドデータから、時間領域のオーディオデータを求めた場合と、符号化装置 1 0 1 でオーディオデータをサブバンド符号化することにより得られるサブバンドデータから時間領域のオーディオデータを求めた場合とでは、量子化の有無によって、音質に差が生じる。即ち、前者では後者より低音質のオーディオデータが得られ、後者では前者より高音質のオーディオデータを得ることができる。そこで、以下、適宜、オーディオデータを MPEG1 オーディオレイヤ 1 方式により符号化し、さらに、MPEG1 オーディオレイヤ 1 方式により復号することにより得られるサブバンドデータを、低音質サブバンドデータというとともに、オーディオデータをサブバンド符号化することにより得られるサブバンドデータを、高音質サブバンドデータという。

【 0 0 7 4 】

上述の条件下において、高音質サブバンドデータ y は、例えば、低音質サブバンドデータから、高音質サブバンドデータを予測するための予測タップとして抽出される複数の低音質サブバンドデータと、タップ係数とを用いて、次の線形 1 次式（線形結合）によって求められる。

【 0 0 7 5 】

【 数 1 】

$$y = \sum_{n=1}^N w_n x_n$$

・・・ (1)

【 0 0 7 6 】

但し、式 (1) において、 x_n は、高音質サブバンドデータ y についての予測タップを構成する、 n 番目の低音質サブバンドデータ（のサンプル値）を表し、 w_n は、 n 番目の低音質サブバンドデータと乗算される n 番目のタップ係数を表す。なお、式 (1) では、予測タップが、 N 個の低音質サブバンドデータ x_1, x_2, \dots, x_N で構成されるものとしてある。

【 0 0 7 7 】

ここで、高音質サブバンドデータ y は、式 (1) に示した線形 1 次式ではなく、2 次以上の高次の式によって求めるようにすることも可能である。

【 0 0 7 8 】

いま、第 k サンプルの高音質サブバンドデータの真値を y_k と表すとともに、式 (1) によって得られるその真値 y_k の予測値を y_k' と表すと、その予測誤差 e_k は、例えば、次式で表される。

【 0 0 7 9 】

【 数 2 】

$$e_k = y_k - y_k'$$

10

... (2)

【 0 0 8 0 】

式 (2) の予測値 y_k' は、式 (1) にしたがって求められるため、式 (2) の y_k' を、式 (1) にしたがって置き換えると、次式が得られる。

【 0 0 8 1 】

【 数 3 】

$$e_k = y_k - \left(\sum_{n=1}^N w_n x_{n,k} \right)$$

20

... (3)

【 0 0 8 2 】

但し、式 (3) において、 $x_{n,k}$ は、第 k サンプルの高音質サブバンドデータについての予測タップを構成する n 番目の低音質サブバンドデータを表す。

【 0 0 8 3 】

式 (3) の予測誤差 e_k を 0 とするタップ係数 w_n が、高音質サブバンドデータを予測するのに最適なものとなるが、すべての高音質サブバンドデータについて、そのようなタップ係数 w_n を求めることは、一般には困難である。

30

【 0 0 8 4 】

そこで、タップ係数 w_n が最適なものであることを表す規範として、例えば、最小自乗法を採用することとすると、最適なタップ係数 w_n は、統計的な誤差としての、例えば、次式で表される自乗誤差の総和 E を最小にすることで求めることができる。

【 0 0 8 5 】

【 数 4 】

$$E = \sum_{k=1}^K e_k^2$$

40

... (4)

【 0 0 8 6 】

但し、式 (4) において、 K は、高音質サブバンドデータ y_k と、その高音質サブバンドデータ y_k についての予測タップを構成する低音質サブバンドデータ $x_{1,k}, x_{2,k}, \dots, x_{N,k}$ とのセットのサンプル数を表す。

【 0 0 8 7 】

式 (4) の自乗誤差の総和 E を最小 (極小) にするタップ係数 w_n は、その総和 E をタップ係数 w_n で偏微分したものを 0 とするものであり、従って、次式を満たす必要がある。

【 0 0 8 8 】

【 数 5 】

50

$$\frac{\partial E}{\partial w_n} = e_1 \frac{\partial e_1}{\partial w_n} + e_2 \frac{\partial e_2}{\partial w_n} + \dots + e_k \frac{\partial e_n}{\partial w_n} = 0 \quad (n=1, 2, \dots, N)$$

... (5)

【 0 0 8 9 】

そこで、上述の式 (3) をタップ係数 w_n で偏微分すると、次式が得られる。

【 0 0 9 0 】

【 数 6 】

$$\frac{\partial e_k}{\partial w_1} = -x_{1,k}, \quad \frac{\partial e_k}{\partial w_2} = -x_{2,k}, \quad \dots, \quad \frac{\partial e_k}{\partial w_N} = -x_{N,k}, \quad (k=1, 2, \dots, K)$$

... (6)

【 0 0 9 1 】

式 (5) と (6) から、次式が得られる。

【 0 0 9 2 】

【 数 7 】

$$\sum_{k=1}^K e_k x_{1,k} = 0, \quad \sum_{k=1}^K e_k x_{2,k} = 0, \quad \dots, \quad \sum_{k=1}^K e_k x_{N,k} = 0$$

... (7)

【 0 0 9 3 】

式 (7) の e_k に、式 (3) を代入することにより、式 (7) は、式 (8) に示す正規方程式で表すことができる。

【 0 0 9 4 】

【 数 8 】

$$\begin{bmatrix} \left(\sum_{k=1}^K x_{1,k} x_{1,k} \right) & \left(\sum_{k=1}^K x_{1,k} x_{2,k} \right) & \dots & \left(\sum_{k=1}^K x_{1,k} x_{N,k} \right) \\ \left(\sum_{k=1}^K x_{2,k} x_{1,k} \right) & \left(\sum_{k=1}^K x_{2,k} x_{2,k} \right) & \dots & \left(\sum_{k=1}^K x_{2,k} x_{N,k} \right) \\ \vdots & \vdots & \ddots & \vdots \\ \left(\sum_{k=1}^K x_{N,k} x_{1,k} \right) & \left(\sum_{k=1}^K x_{N,k} x_{2,k} \right) & \dots & \left(\sum_{k=1}^K x_{N,k} x_{N,k} \right) \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix} = \begin{bmatrix} \left(\sum_{k=1}^K x_{1,k} y_k \right) \\ \left(\sum_{k=1}^K x_{2,k} y_k \right) \\ \vdots \\ \left(\sum_{k=1}^K x_{N,k} y_k \right) \end{bmatrix}$$

... (8)

【 0 0 9 5 】

式 (8) の正規方程式は、高音質サブバンドデータ y_k と低音質サブバンドデータ $x_{n,k}$ のセットを、ある程度の数だけ用意することで、求めるべきタップ係数 w_n の数と同じ数だけたてることができ、従って、式 (8) を解くことで (但し、式 (8) を解くには、式 (8) において、タップ係数 w_n にかかる左辺の行列が正則である必要がある)、最適なタップ係数 w_n を求めることができる。なお、式 (8) を解くにあたっては、例えば、掃き出し法 (Gauss-Jordan の消去法) などを採用することが可能である。

【 0 0 9 6 】

10

20

30

40

50

以上のように、多数の高音質サブバンドデータ y_1, y_2, \dots, y_K を、タップ係数の学習の教師となる教師データとするとともに、各高音質サブバンドデータ y_k についての予測タップを構成する低音質サブバンドデータ $x_{1,k}, x_{2,k}, \dots, x_{N,k}$ を、タップ係数の学習の生徒となる生徒データとして、式(8)を解くことにより、最適なタップ係数 w_n を求める学習を行っておき、さらに、そのタップ係数 w_n を用い、式(1)により、低音質サブバンドデータを、高音質サブバンドデータにマッピング(変換)するのが適応処理である。

【0097】

なお、適応処理は、低音質サブバンドデータには含まれていないが、高音質サブバンドデータに含まれる成分が再現される点で、例えば、単なる補間処理等とは異なる。即ち、適応処理では、式(1)だけを見る限りは、いわゆる補間フィルタを用いての補間処理と同一であるが、その補間フィルタのタップ係数に相当するタップ係数 w_n が、教師データとしての高音質サブバンドデータと生徒データとしての低音質サブバンドデータとを用いての学習により求められるため、高音質サブバンドデータに含まれる成分を再現することができる。

10

【0098】

ここで、タップ係数 w_n の学習では、教師データ y と生徒データ x との組み合わせとして、どのようなものを採用するかによって、各種の変換を行うタップ係数 w_n を求めることができる。

【0099】

20

即ち、例えば、上述のように、教師データ y として、高音質サブバンドデータを採用するとともに、生徒データ x として、低音質サブバンドデータを採用した場合には、低音質サブバンドデータを高音質サブバンドデータに変換するタップ係数 w_n を得ることができる。また、例えば、教師データ y として、時間領域の高音質のオーディオデータを採用するとともに、生徒データ x として、その高音質のオーディオデータを処理して得られる低音質サブバンドデータを採用した場合には、周波数領域のデータとしての低音質サブバンドデータを、時間領域のデータとしての高音質のオーディオデータに変換するタップ係数 w_n を得ることができる。さらに、例えば、教師データ y として、時間領域の高音質のオーディオデータを採用するとともに、生徒データ x として、その高音質のオーディオデータをMPEG2オーディオACC方式により符号化して復号することにより得られるMDCT係数を採用した場合には、周波数領域のデータとしてのMDCT係数を、時間領域のデータとしての高音質のオーディオデータに変換するタップ係数 w_n を得ることができる。

30

【0100】

なお、上述の場合には、予測タップを、低音質サブバンドデータだけから構成することとしたが、予測タップは、低音質サブバンドデータ以外の情報、即ち、例えば、サイドインフォメーションなどを含めて構成することが可能である。

【0101】

図6のクラス分類適応処理部113は、タップ生成部121および122、クラス分類部123、タップ係数記憶部124、並びに予測演算部125で構成され、上述のようなクラス分類適応処理を行うことにより、高音質サブバンドデータを求めるようになっている。

40

【0102】

即ち、タップ生成部121は、コントローラ112から供給される特定情報としてのサブバンドナンバとデータ位置によって特定される高音質サブバンドデータを注目データとして選択し、その注目データ(の予測値)を求めるのに用いる予測タップを、デコーダ111から供給される、MPEG1オーディオレイヤ1方式により復号されたサブバンドデータ、即ち、低音質サブバンドデータ、さらには、やはり、デコーダ111から供給されるサイドインフォメーションから生成し、予測演算部125に供給する。

【0103】

具体的には、タップ生成部121は、例えば、注目データのフレームの低音質サブバンド

50

データとサイドインフォメーションの幾つかを抽出し、それらを所定の順番で並べたものを、予測タップとする。あるいは、また、タップ生成部 1 2 1 は、例えば、注目データのフレームの他、そのフレームに近接するフレームからも、低音質サブバンドデータとサイドインフォメーションの幾つかを抽出し、それらを所定の順番で並べたものを、予測タップとする。

【 0 1 0 4 】

タップ生成部 1 2 2 は、タップ生成部 1 2 1 と同様に、コントローラ 1 1 2 から供給される特定情報から注目データを選択し、その注目データのクラス分類に用いるクラスタップを、デコーダ 1 1 1 から供給される低音質サブバンドデータとサイドインフォメーションから生成し、クラス分類部 1 2 3 に出力する。

10

【 0 1 0 5 】

ここで、上述したように、予測タップやクラスタップは、デコーダ 1 1 1 から供給される低音質サブバンドデータの他、サイドインフォメーションをも用いて生成することが可能であるが、デコーダ 1 1 1 から供給される低音質サブバンドデータのみから生成することも可能である。

【 0 1 0 6 】

そこで、以下では、説明を簡単にするために、予測タップとクラスタップは、デコーダ 1 1 1 から供給される低音質サブバンドデータのみから生成されるものとする。

【 0 1 0 7 】

クラス分類部 1 2 3 は、タップ生成部 1 2 2 からのクラスタップ、さらには、例えば、コントローラ 1 1 2 から供給される特定情報としてのサブバンドナンバとデータ位置に基づいて、注目データをクラス分類し、その結果得られる注目データのクラスに対応するクラスコードを、タップ係数記憶部 1 2 4 に供給する。

20

【 0 1 0 8 】

ここで、クラスタップによるクラス分類を行うにあたっては、そのクラスタップを構成する低音質サブバンドデータの各サンプル値を表すビット列をそのまま所定の順番で並べて得られるビット列を、クラスコードとすることが可能であるが、この場合、クラス数（クラスの総数）が膨大な数になることがある。そこで、クラス分類には、例えば、KビットADRC(Adaptive Dynamic Range Coding)処理などの圧縮処理を採用することができる。

【 0 1 0 9 】

KビットADRC処理においては、クラスタップを構成する低音質サブバンドデータのサンプル値の最大値MAXと最小値MINが検出され、 $DR=MAX-MIN$ を、局所的なダイナミックレンジとし、このダイナミックレンジDRに基づいて、クラスタップを構成する低音質サブバンドデータがKビットに再量子化される。即ち、クラスタップを構成する各低音質サブバンドデータのサンプル値から、最小値MINが減算され、その減算値が $DR/2^K$ で除算（量子化）される。従って、クラスタップが、例えば、1ビットADRC処理される場合には、そのクラスタップを構成する各低音質サブバンドデータのサンプル値は1ビットとされることになる。そして、この場合、以上のようにして得られる、クラスタップを構成する各低音質サブバンドデータについての1ビットのサンプル値を、所定の順番で並べたビット列が、ADRCコードとして出力され、このADRCコードがクラスコードとされる。

30

40

【 0 1 1 0 】

なお、クラス分類は、その他、例えば、クラスタップを構成する低音質サブバンドデータ（のサンプル値）を、ベクトルのコンポーネントとみなし、そのベクトルをベクトル量子化すること等によって行うことも可能である。また、クラス分類としては、1クラスのクラス分類を行うことも可能である。この場合、クラス分類部 1 2 3 は、どのようなクラスタップおよび特定情報が供給されても、固定のクラスコードを出力するものとなる。但し、1クラスのクラス分類を行うということは、クラス分類部 1 2 3 を設けないことと等価である。

【 0 1 1 1 】

図 6 の実施の形態では、クラス分類部 1 2 3 は、クラスタップの他、注目データを特定情

50

報としてのサブバンドナンバとデータ位置にも基づいて、クラス分類を行う。ここで、サブバンドナンバとデータ位置に基づくクラス分類も、上述のKビットADRC処理やベクトル量子化などを採用することができるが、サブバンドナンバやデータ位置の数は、それほど多くないので、サブバンドナンバとデータ位置に基づくクラス分類では、そのサブバンドナンバとデータ位置それぞれを表すビット列をそのまま並べて得られるビット列を、クラスコードとすることが可能である。

【0112】

即ち、本実施の形態では、上述したように、サブバンドナンバは、0乃至31の5ビットで表される値であり、データ位置は、0乃至11の4ビットで表される値である。従って、サブバンドナンバとデータ位置に基づくクラス分類において、そのサブバンドナンバとデータ位置それぞれを表すビット列をそのまま並べて得られるビット列を、クラスコードとして採用しても、クラスコードは、9ビット(=5+4ビット)で表されるものとなり、クラス数は、それほど多くならない。

10

【0113】

いま、クラスタップに基づくクラス分類の結果得られるクラスコードを、タップコードというとともに、特定情報としてのサブバンドナンバとデータ位置に基づくクラス分類の結果得られるクラスコードを、情報コードというものとする。クラス分類部123は、タップコードと情報コードを求めた後、そのタップコードと情報コードそれぞれを表すビット列を並べて得られるビット列を、注目データについての最終的なクラスコードとして、タップ係数記憶部124に供給する。

20

【0114】

タップ係数記憶部124は、学習の教師となる高音質サブバンドデータである教師データと、学習の生徒となる低音質サブバンドデータである生徒データとの関係を、1以上のクラスごとに学習することにより得られたタップ係数を記憶している。そして、タップ係数記憶部124は、クラス分類部123から注目データのクラスコードが供給されると、そのクラスコードに対応するクラスのタップ係数を読み出すことにより、注目データのクラスのタップ係数を取得し、予測演算部125に供給する。なお、タップ係数記憶部124に記憶されるタップ係数の学習方法についての詳細は、後述する。

【0115】

予測演算部125は、タップ係数記憶部124から供給される、注目データのクラスのタップ係数 w_1, w_2, \dots と、タップ生成部121からの注目データの予測タップ(を構成する各低音質サブバンドデータのサンプル値) x_1, x_2, \dots を用いて、式(1)に示した積和演算を行うことにより、注目データ y (の予測値)を求め、これを、注目データとしての高音質サブバンドデータのサンプル値とする。予測演算部125は、以上のようにして、1サブバンドを構成する高音質サブバンドデータのサンプル値すべてが得られると、その1サブバンドの高音質サブバンドデータを、サブバンド合成部114に供給する。

30

【0116】

次に、図7乃至図9を参照して、図6のタップ生成部121と122でそれぞれ生成される予測タップとクラスタップについて説明する。

40

【0117】

なお、以下、適宜、予測タップとクラスタップを、特に区別する必要がない場合には、単に、タップという。また、ここでは、説明を簡単にするために、予測タップとクラスタップのタップの構造(タップ構造)は、同一であるとする。但し、予測タップとクラスタップとは、異なるタップ構造のものとすることが可能である。

【0118】

いま、注目データのサブバンドを注目サブバンドというとともに、注目データのフレームを注目フレームというものとする。

【0119】

注目フレームは、図7Aに示すように、32のサブバンドsb0乃至sb31で構成されるが、

50

いま、そのうちのサブバンドsb#nが注目サブバンドであるとする、注目データのタップは、例えば、注目フレームを構成する32のサブバンドsb0乃至sb31それぞれの低音質サブバンドデータすべてで構成することが可能である。

【0120】

しかしながら、この場合、注目データのタップを構成する低音質サブバンドデータのサンプル数（タップ数）が多くなる。そして、その結果、クラス分類部123でのクラス分類により得られるクラスの総数も多くなり、また、予測演算部125での演算量も多くなる。

【0121】

一方、クラス分類適応処理によれば、基本的には、タップ数が多いほど、注目データを精度良く求めることができる（高い予測制度で予測することができる）。しかしながら、タップ数が多くても、その中に、注目データとまったく相関のない、またはほとんど相関のないものが含まれている場合には、逆に、注目データの予測精度が劣化することがある。

【0122】

そこで、タップ生成部121および122は、例えば、注目フレームを構成する32のサブバンドsb0乃至sb31のうちの、幾つかのサブバンドを、注目データのタップを構成するのに用いるタップ用サブバンドとして選択し、そのタップ用サブバンドの低音質サブバンドデータを、注目データのタップとして抽出する。

【0123】

タップ用サブバンドの選択方法としては、例えば、注目サブバンドsb#nだけを選択する方法がある。しかしながら、オーディオデータがMPEG1オーディオレイヤ1で符号化された場合には、上述したように、ビット割り当てが0とされ、サブバンドデータが抜けているサブバンド（以下、適宜、ヌルサブバンドという）が存在することがある。

【0124】

従って、タップ用サブバンドを、注目サブバンドsb#nだけとした場合に、その注目サブバンドsb#nがヌルサブバンドであるときには、注目データのタップは、例えば、すべて0となる。予測タップがすべて0の場合、式(1)を演算しても、その演算結果は、必ず0となるから、予測タップがすべて0となるのは好ましくない。

【0125】

そこで、タップ用サブバンドとしては、複数のサブバンドを選択するのが望ましい。タップ用サブバンドとして複数のサブバンドを選択する場合、その選択方法としては、例えば、注目サブバンドsb#nと、その注目サブバンドsb#nに近接するサブバンド（例えば、隣接するサブバンドsb#n-1およびsb#n+1など）を選択する方法がある。

【0126】

ところで、クラス分類適応処理では、タップに、注目データとの相関が高いデータが含まれる方が、相関が低いデータが含まれる場合よりも、注目データを精度良く求めることができる。

【0127】

一方、楽曲などの音の音色は、その音に含まれる倍音などにより決まり、楽曲には、多くの倍音が含まれる。従って、オーディオデータが、楽曲などのデータである場合には、ある周波数帯域の成分は、その 2^{-N} 倍や 2^N 倍の周波数帯域の成分と大きな相関を有することが多い（Nは正の整数値）。

【0128】

そこで、本実施の形態では、例えば、図7Bに示すように、注目サブバンドsb#nと、その注目サブバンドsb#nの1/2倍の周波数帯域のサブバンドsb#nL、および2倍の周波数帯域のサブバンドsb#nHが、タップ用サブバンドとして選択される。

【0129】

この場合、タップが、注目サブバンドsb#nと相関の大きいサブバンドのサブバンドデータで構成されることとなるので、注目データとしての高音質データを高精度で求めることが可能となる。

10

20

30

40

50

【0130】

なお、注目サブバンドsb#nの1/2倍または2倍の周波数帯域のサブバンドsb#nLまたはsb#nHのサブバンドナンバ#nLまたは#nHは、例えば、 $\text{INT}[\#n/2]$ または $\text{INT}[\#n \times 2]$ や、あるいは、 $\text{INT}[\#n/2]+1$ または $\text{INT}[\#n \times 2]+1$ などで、それぞれ表される。但し、 $\text{INT}[]$ は、 $[]$ 内の値以下の最大の整数値を表す。

【0131】

また、タップ用サブバンドとしては、注目サブバンドsb#nと、その注目サブバンドsb#nの1/2倍の周波数帯域のサブバンド（以下、適宜、ローサブバンドという）sb#nL、および2倍の周波数帯域のサブバンド（以下、適宜、ハイサブバンドという）sb#nHの他、例えば、注目サブバンドsb#nの1/4倍や4倍、1/8倍、8倍などの周波数帯域のサブバンドを選択することも可能である。

10

【0132】

ここで、本実施の形態では、サブバンドは、32のサブバンドsb0乃至sb31しか存在しないから、注目サブバンドsb#nによっては、その1/2倍の周波数帯域のローサブバンドsb#nL、あるいは2倍の周波数帯域のハイサブバンドsb#nHが存在しないことがある。ローサブバンドsb#nLまたはハイサブバンドsb#nHが存在しない場合、タップ生成部121および122は、その存在しないサブバンドのサブバンドデータが、例えば0であるとして、タップを構成する。

【0133】

次に、タップ生成部121および122は、上述したように、タップ用サブバンドである注目サブバンドsb#n、ローサブバンドsb#nL、およびハイサブバンドsb#nHのサブバンドデータから、タップを生成する。この場合、タップ生成部121および122では、例えば、注目サブバンドsb#n、ローサブバンドsb#nL、およびハイサブバンドsb#nHのサブバンドデータのすべてを、タップとすることができる他、そのサブバンドデータの一部だけを、タップとすることも可能である。

20

【0134】

注目サブバンドsb#n、ローサブバンドsb#nL、およびハイサブバンドsb#nHのサブバンドデータの一部をタップとする場合には、例えば、図8に示すように、注目サブバンドsb#n、ローサブバンドsb#nL、ハイサブバンドsb#nHそれぞれから、注目データのデータ位置に近いデータ位置のサブバンドデータを、タップとすることができる。

30

【0135】

即ち、図8では、注目サブバンドsb#nの先頭から6番目のデータ位置にある高音質サブバンドデータd5が注目データとされている。この場合、例えば、注目サブバンドsb#n、ローサブバンドsb#nL、ハイサブバンドsb#nHそれぞれにおいて、注目データd5のデータ位置に近い低音質サブバンドデータd3乃至d7の、合計で25のサブバンドデータを、注目データのタップとすることができる。

【0136】

なお、上述の場合には、注目フレームのサブバンドのサブバンドデータのみを用いて、注目データのタップを構成するようにしたが、注目データのタップは、その他、注目フレームに近接するフレームのサブバンドのサブバンドデータを用いて構成することが可能である。

40

【0137】

即ち、注目データのタップは、例えば、図9Aに示すように、注目フレームである第fフレームの注目サブバンドsb#nと、その1フレーム前または後の第f-1フレームまたは第f+1フレームそれぞれの、注目サブバンドsb#nと同一のサブバンドナンバ#nのサブバンドsb#nを、タップ用サブバンドとして、そのタップ用サブバンドのサブバンドデータを用いて構成することができる。注目サブバンドsb#nの1フレーム前または後の第f-1フレームまたは第f+1フレームそれぞれの、注目サブバンドsb#nと同一のサブバンドナンバ#nのサブバンドsb#nは、注目フレームである第fフレームの注目サブバンドsb#nとの相関が大きいので、注目データとしての高音質データを高精度で求めることが可能となる。

50

【 0 1 3 8 】

さらに、注目データのタップは、例えば、図 9 B に示すように、第 $f - 1$ 乃至第 $f + 1$ フレームそれぞれのサブバンド $sb\#n$ 、第 $f - 1$ 乃至第 $f + 1$ フレームそれぞれのローサブバンド $sb\#nL$ 、および第 $f - 1$ 乃至第 $f + 1$ フレームそれぞれのハイサブバンド $sb\#nH$ を、タップ用サブバンドとして、そのタップ用サブバンドのサブバンドデータを用いて構成することができる。

【 0 1 3 9 】

次に、図 10 のフローチャートを参照して、図 6 の復号装置 102 の処理（復号処理）について説明する。

【 0 1 4 0 】

デコーダ 111 は、あるフレームの符号化ビットストリームを受信すると、そのフレームを、注目フレームとし、ステップ S 1 において、その注目フレームについて、通常デコード処理を行う。

【 0 1 4 1 】

即ち、本実施の形態では、符号化ビットストリームは、オーディオデータを MPEG1 オーディオレイヤ 1 方式で符号化したものであり、ステップ S 1 では、注目フレームの符号化ビットストリームが、MPEG1 オーディオレイヤ 1 方式で復号される。

【 0 1 4 2 】

具体的には、ステップ S 1 の通常デコード処理では、まず最初に、ステップ S 2 1 において、デコーダ 111 は、注目フレームの符号化ビットストリームを、量子化サブバンドデータと、その他のデータであるサイドインフォメーションに分解し、注目フレームのサイドインフォメーションを、コントローラ 112 とクラス分類適応処理部 113 に供給する。そして、ステップ S 2 2 に進み、デコーダ 111 は、注目フレームの量子化サブバンドデータを逆量子化し、これにより、MPEG1 オーディオレイヤ 1 方式で符号化されて復号されたサブバンドデータ、即ち、低音質サブバンドデータを得る。この注目フレームの低音質サブバンドデータは、クラス分類適応処理部 113 とサブバンド合成部 114 に供給される。

【 0 1 4 3 】

以上のステップ S 2 1 および 2 2 の処理によって、ステップ S 1 の通常デコード処理が終了し、その後は、ステップ S 2 に進む。ステップ S 2 では、コントローラ 112 は、注目フレームのサブバンドのサブバンドナンバを表す変数 n に、初期値としての、例えば 0 をセットし、ステップ S 3 に進む。ステップ S 3 では、コントローラ 112 は、注目フレームのサイドインフォメーションを参照することにより、サブバンドナンバ $\#n$ で特定される注目フレームのサブバンド $sb\#n$ に、サブバンドデータが存在するかどうか、即ち、注目フレームのサブバンド $sb\#n$ に対するビット割り当てが 0 かどうかを判定する。

【 0 1 4 4 】

ステップ S 3 において、注目フレームのサブバンド $sb\#n$ に、サブバンドデータが存在すると判定された場合、ステップ S 4 乃至 S 9 をスキップして、ステップ S 10 に進む。

【 0 1 4 5 】

また、ステップ S 3 において、注目フレームのサブバンド $sb\#n$ に、サブバンドデータが存在しないと判定された場合、即ち、注目フレームのサブバンド $sb\#n$ に対するビット割り当てが 0 である場合、ステップ S 4 に進み、コントローラ 112 は、その注目フレームのサブバンド $sb\#n$ を、注目サブバンドとして選択し、ステップ S 5 に進む。

【 0 1 4 6 】

ステップ S 5 では、コントローラ 112 は、サブバンドの高音質サブバンドデータのデータ位置を表す変数 i に、初期値としての、例えば 0 をセットし、ステップ S 6 に進む。

【 0 1 4 7 】

ステップ S 6 では、コントローラ 112 は、注目サブバンド $sb\#n$ の、データ位置 $\#i$ で特定される高音質サブバンドデータ $d\#i$ を、注目データとして選択する。さらに、ステップ S 6 では、コントローラ 112 は、注目データ $d\#i$ を特定する特定情報として、注目データ d

10

20

30

40

50

#iのデータ位置#iと、注目サブバンドsb#nのサブバンドナンバ#nを、クラス分類適応処理部113とサブバンド合成部114に供給し、ステップS7に進む。

【0148】

ステップS7では、クラス分類適応処理部113が、注目データd#iを求めるクラス分類適応処理を行う。

【0149】

即ち、ステップS7のクラス分類適応処理では、まず最初に、ステップS31において、タップ生成部121と122が、コントローラ112からの特定情報としてのデータ位置#iとサブバンドナンバ#nで特定される高音質サブバンドデータd#iを、注目データd#iとして認識し、その注目データd#iについて、デコーダ111から供給される低音質サブバン
ドデータを用い、例えば、図7乃至図9で説明した予測タップとクラスタップをそれぞれ生成する。

10

【0150】

ここで、本実施の形態では、説明を簡単にするために、予測タップとクラスタップを、低音質サブバンドデータのみを用いて生成することとしているが、予測タップとクラスタップは、上述したように、デコーダ111から供給されるサイドインフォメーションなどをも用いて生成することが可能である。

【0151】

ステップS31で得られた予測タップは、タップ生成部121から予測演算部125に供給され、また、クラスタップは、タップ生成部122からクラス分類部123に供給され、その後、ステップS31からステップS32に進む。

20

【0152】

ステップS32では、クラス分類部123が、タップ生成部122から供給されるクラスタップの他、コントローラ112から供給される注目サブバンドsb#nのサブバンドナンバ#n、および注目データd#iのデータ位置#iに基づいて、注目データをクラス分類し、その結果得られる注目データのクラスを表すクラスコードを、タップ係数記憶部124に供給して、ステップS33に進む。

【0153】

ここで、クラス分類部123でのクラス分類は、クラスタップ、サブバンドナンバ#n、およびデータ位置#iに基づいて行う他、例えば、クラスタップのみに基づいて行うようにすることも可能である。

30

【0154】

ステップS33では、タップ係数記憶部124が、そこに記憶されているクラスごとのタップ係数から、クラス分類部123から供給されるクラスコードが表す注目データd#iのクラスのタップ係数を読み出すことにより取得し、その注目データd#iのクラスのタップ係数を、予測演算部125に供給して、ステップS34に進む。

【0155】

ステップS34では、予測演算部125が、タップ生成部121から供給される注目データd#iについての予測タップと、タップ係数記憶部124から供給される注目データd#iのクラスのタップ係数とを用いて、式(1)の積和演算を行うことにより、注目データd#i
を求める。

40

【0156】

以上のステップS31乃至S34の処理によって、ステップS7のクラス分類適応処理が終了し、ステップS8に進む。ステップS8では、コントローラ112が、変数iが、データ位置の最大値である11に等しいかどうかを判定し、11に等しくないと判定した場合、ステップS9に進む。ステップS9では、コントローラ112は、変数iを1だけインクリメントして、ステップS6に戻り、以下、同様の処理が繰り返される。

【0157】

また、ステップS8において、変数iが、データ位置の最大値である11に等しいと判定された場合、即ち、注目サブバンドsb#nのすべての高音質サブバンドデータを注目データ

50

として、そのサンプル値が、クラス分類適応処理によって求められた場合、クラス分類適応処理部 1 1 3 の予測演算部 1 2 5 は、その注目サブバンドsb#nの高音質サブバンドデータを、サブバンド合成部 1 1 4 に供給して、ステップ S 1 0 に進む。

【 0 1 5 8 】

ステップ S 1 0 では、コントローラ 1 1 2 は、変数 n が、1 フレームのサブバンドのサブバンドナンバの最大値である 3 1 に等しいかどうかを判定する。ステップ S 1 0 において、変数 n が、サブバンドナンバの最大値である 3 1 に等しくないとは判定された場合、ステップ S 1 1 に進み、コントローラ 1 1 2 は、変数 n を 1 だけインクリメントして、ステップ S 3 に戻り、以下、同様の処理が繰り返される。

【 0 1 5 9 】

また、ステップ S 1 0 において、変数 n が、サブバンドナンバの最大値である 3 1 に等しいとは判定された場合、即ち、注目フレームのサブバンドすべてについて、そのサブバンドデータがあるかどうかを判定し、サブバンドデータがないサブバンドについては、クラス分類適応処理によって、高音質サブバンドデータを求めた場合、ステップ S 1 2 に進み、サブバンド合成部 1 1 4 は、デコーダ 1 1 1 から供給される、サブバンドデータが存在するサブバンドと、クラス分類適応処理部 1 1 3 から供給される、クラス分類適応処理によってサブバンドデータが生成されたサブバンドとを合成することにより、注目フレームの時間領域のオーディオデータを復号して出力する。

【 0 1 6 0 】

従って、この場合、サブバンドデータがないサブバンド、即ち、ビット割り当てが 0 のサブバンドについては、クラス分類適応処理によってサブバンドデータが生成され、そのサブバンドデータを用いて、時間領域のオーディオデータが復号される。その結果、ビット割り当てが 0 とされたサブバンドに対応する周波数成分を有するオーディオデータ、即ち、高音質のオーディオデータを、符号化ビットストリーム of データ量を増加させずに復号することができる。

【 0 1 6 1 】

なお、サブバンド合成部 1 1 4 は、クラス分類適応処理部 1 1 3 でサブバンドデータが求められるサブバンドを、コントローラ 1 1 2 から供給されるサブバンドナンバから認識する。

【 0 1 6 2 】

ステップ S 1 2 の処理後は、ステップ S 1 3 に進み、デコーダ 1 1 1 は、注目フレームの次のフレームの符号化ビットストリームが供給されたかどうかを判定する。ステップ S 1 3 において、注目フレームの次のフレームの符号化ビットストリームが供給されたと判定された場合、その符号化ビットストリームが供給されたフレームを、新たに注目フレームとして、ステップ S 1 に戻り、以下、同様の処理が繰り返される。

【 0 1 6 3 】

また、ステップ S 1 3 において、注目フレームの次のフレームの符号化ビットストリームが供給されていないとは判定された場合、処理を終了する。

【 0 1 6 4 】

次に、図 1 1 は、図 6 の復号装置 1 0 2 のタップ係数記憶部 1 2 4 に記憶させるクラスごとのタップ係数を学習する学習装置の一実施の形態の構成例を示している。

【 0 1 6 5 】

教師データ生成部 1 3 1 は、タップ係数の学習の教師となる教師データを生成し、学習部 1 3 4 に供給する。

【 0 1 6 6 】

即ち、図 6 の復号装置 1 0 2 のクラス分類適応処理部 1 1 3 では、オーディオデータを MP EG1 オーディオレイヤ 1 方式で符号化して復号することにより得られるサブバンドデータ（低音質サブバンドデータ）から予測タップを生成し、その予測タップを用いて、式（1）の積和演算を行うことにより、オーディオデータを単にサブバンド符号化して得られるサブバンドデータ（高音質サブバンドデータ）が求められる。

10

20

30

40

50

【0167】

従って、式(1)の積和演算に用いるタップ係数としては、オーディオデータを単にサブバンド符号化して得られるサブバンドデータを、教師データとするとともに、オーディオデータをMPEG1オーディオレイヤ1方式で符号化して復号することにより得られるサブバンドデータを生徒データとして、その生徒データと教師データとを、式(1)によって関係付けるものを求めれば良い。

【0168】

そこで、教師データ生成部131は、学習用データベース141とエンコーダ142で構成されており、オーディオデータを単にサブバンド符号化して得られる高音質サブバンドデータを、教師データとして生成する。

10

【0169】

即ち、学習用データベース141は、タップ係数の学習に用いる学習用データとして、時間領域のデジタルオーディオデータを記憶している。なお、ここでは、学習用データベース141に記憶させるオーディオデータは、例えば、図5の符号化装置101で符号化対象となるオーディオデータと同一の音質のものとする。但し、学習用データベース141に記憶させるオーディオデータは、例えば、図5の符号化装置101で符号化対象となるオーディオデータよりも高音質のものとするのが可能である。

【0170】

エンコーダ142は、図5の符号化装置101と同様に構成されており、学習用データベース141に記憶されたオーディオデータを読み出し、MPEG1オーディオレイヤ1方式で符号化し、その結果得られる符号化ビットストリームを、生徒データ生成部132に供給する。さらに、エンコーダ142は、オーディオデータを、MPEG1オーディオレイヤ1方式で符号化する過程で得られる高音質サブバンドデータ、即ち、オーディオデータをサブバンド符号化して得られるサブバンドデータを、教師データとして、学習部134の足し込み部154に供給する。

20

【0171】

生徒データ生成部132は、デコーダ143で構成されており、オーディオデータをMPEG1オーディオレイヤ1方式で符号化して復号することにより得られるサブバンドデータを、生徒データとして生成する。

【0172】

即ち、デコーダ143は、図6のデコーダ111と同様に構成されており、教師データ生成部131のエンコーダ142が出力する符号化ビットストリームをMPEG1オーディオレイヤ1方式で復号し、その復号の過程で得られるサブバンドデータ(低音質サブバンドデータ)とサイドインフォメーションを、生徒データとして、学習部134に供給する。さらに、デコーダ143は、符号化ビットストリームをMPEG1オーディオレイヤ1方式で復号する過程で得られるサイドインフォメーションを、コントローラ133に供給する。

30

【0173】

コントローラ133は、デコーダ143から供給されるサイドインフォメーションのフレームを注目フレームとして、その注目フレームにおいて、ビット割り当てが0ビットとされたサブバンドがあるかどうかを、デコーダ143から供給されるサイドインフォメーションに基づいて判定し、さらに、ビット割り当てが0ビットとされたサブバンド(ヌルサブバンド)がある場合には、そのヌルサブバンドを表すサブバンドナンバを認識する。また、コントローラ133は、ヌルサブバンドの高音質サブバンドデータとしての教師データを、順次、注目データとして選択し、その注目データのデータ位置と、その注目データのサブバンド(ヌルサブバンド)を表すサブバンドナンバを、注目データを特定するための特定情報として、学習部134に供給する。

40

【0174】

学習部134は、教師データ生成部131から供給される教師データとしての高音質サブバンドデータと、生徒データ生成部132から供給される生徒データとしての低音質サブバンドデータおよびサイドインフォメーションとの関係を学習することにより、式(1)

50

により、教師データと生徒データとを関係付けるタップ係数を、クラスごとに求める。

【0175】

即ち、学習部134は、タップ生成部151および152、クラス分類部153、足し込み部154、タップ係数演算部155、並びにタップ係数記憶部156で構成されている。

【0176】

タップ生成部151と152は、図6のタップ生成部121と122における場合と同様に、コントローラ133が出力する注目データの特定情報としてのデータ位置とサブバンドナンバから、注目データとしての教師データを認識し、その注目データについて、生徒データ生成部143から供給される生徒データから、予測タップとクラスタップを、それぞれ生成する。

10

【0177】

即ち、タップ生成部151は、注目データについて、生徒データ生成部143から供給される生徒データを用い、図6のタップ生成部121が生成するのと同じのタップ構造の予測タップを構成し、足し込み部154に供給する。タップ生成部152も、注目データについて、生徒データ生成部143から供給される生徒データを用い、図6のタップ生成部122が生成するのと同じのタップ構造のクラスタップを構成し、クラス分類部153に供給する。

【0178】

クラス分類部153は、タップ生成部152から供給されるクラスタップと、コントローラ133から供給される注目データの特定情報としてのデータ位置およびサブバンドナンバに基づき、注目データのクラス分類を、図6のクラス分類部123と同様に行う。クラス分類部153でクラス分類が行われることにより得られる注目データのクラスを表すクラスコードは、クラス分類部153から足し込み部154に供給される。

20

【0179】

足し込み部154およびタップ係数演算部155は、教師データ生成部131から供給される教師データのうちの注目データと、タップ生成部151から供給される予測タップとを用い、教師データと生徒データとの関係を、クラス分類部153から供給されるクラスごとに学習することにより、クラスごとのタップ係数を求める。

【0180】

即ち、足し込み部154は、クラス分類部153が出力するクラスコードごとに、タップ生成部151から供給される予測タップと、教師データ生成部131から供給される教師データのうちの注目データとを対象とした、式(8)の足し込みを行う。

30

【0181】

具体的には、足し込み部154は、クラス分類部153から供給されるクラスコードに対応するクラスごとに、予測タップを構成する生徒データとしての低音質サブバンドデータ(あるいは、サイドインフォメーション) $x_{n,k}$ を用い、式(8)の左辺の行列における低音質サブバンドデータどうしの乗算($x_{n,k} x_{n',k}$)と、サメーション()に相当する演算を行う。

【0182】

さらに、足し込み部154は、やはり、クラス分類部153から供給されるクラスコードに対応するクラスごとに、予測タップを構成する生徒データとしての低音質サブバンドデータ $x_{n,k}$ と注目データとなっている教師データである高音質サブバンドデータ y_k を用い、式(8)の右辺のベクトルにおける低音質サブバンドデータ $x_{n,k}$ および高音質サブバンドデータ y_k の乗算($x_{n,k} y_k$)と、サメーション()に相当する演算を行う。

40

【0183】

足し込み部154は、教師データ生成部131から供給される教師データとしての高音質サブバンドデータすべてを注目データとして、上述の足し込みを行うことにより、各クラスについて、式(8)に示した正規方程式をたてると、その正規方程式を、タップ係数演算部155に供給する。

50

【 0 1 8 4 】

タップ係数演算部 1 5 5 は、足し込み部 1 5 4 から、各クラスについての式 (8) の正規方程式を受信し、その正規方程式を解くことにより、クラスごとのタップ係数を求めて出力する。

【 0 1 8 5 】

タップ係数記憶部 1 5 6 は、タップ係数演算部 1 5 5 が出力するクラスごとのタップ係数を記憶する。

【 0 1 8 6 】

次に、図 1 2 のフローチャートを参照して、図 1 1 の学習装置において行われる、クラスごとのタップ係数を求める学習処理について説明する。

10

【 0 1 8 7 】

まず最初に、教師データ生成部 1 3 1 は、学習用データベース 1 4 1 に記憶されたオーディオデータのフレームのうちの 1 フレームを注目フレームとし、ステップ S 4 1 において、その注目フレームの教師データを生成する。即ち、教師データ生成部 1 3 1 では、エンコーダ 1 4 2 が、学習用データベース 1 4 1 から、注目フレームのオーディオデータを読み出し、MPEG1オーディオレイヤ 1 方式で符号化する。そして、エンコーダ 1 4 2 は、注目フレームのオーディオデータの符号化の過程で得られる注目フレームの高音質サブバンドデータを、教師データとして、学習部 1 3 4 の足し込み部 1 5 4 に供給するとともに、その符号化の結果得られる注目フレームの符号化ビットストリームを生徒データ生成部 1 3 2 に供給する。

20

【 0 1 8 8 】

その後、ステップ S 4 2 に進み、生徒データ生成部 1 3 2 は、教師データ生成部 1 3 1 から供給される注目フレームの符号化ビットストリームから生徒データを生成する。即ち、ステップ S 4 2 では、生徒データ生成部 1 3 2 のデコーダ 1 4 3 が、注目フレームの符号化ビットストリームを、MPEG1オーディオレイヤ 1 方式で復号する。そして、デコーダ 1 4 3 は、注目フレームの符号化ビットストリームの復号過程で得られる低音質サブバンドデータとサイドインフォメーションを、生徒データとして、学習部 1 3 4 に供給する。さらに、デコーダ 1 4 3 は、サイドインフォメーションを、コントローラ 1 3 3 にも供給する。

【 0 1 8 9 】

30

そして、ステップ S 4 3 に進み、コントローラ 1 3 3 は、注目フレームのサブバンドのサブバンドナンバを表す変数 n に、初期値としての、例えば 0 をセットし、ステップ S 4 4 に進む。ステップ S 4 4 では、コントローラ 1 3 3 は、注目フレームのサイドインフォメーションを参照することにより、サブバンドナンバ $\#n$ で特定される注目フレームのサブバンド $sb\#n$ に、生徒データとしてのサブバンドデータが存在するかどうか、即ち、注目フレームのサブバンド $sb\#n$ に対するビット割り当てが 0 かどうかを判定する。

【 0 1 9 0 】

ステップ S 4 4 において、注目フレームのサブバンド $sb\#n$ に、サブバンドデータが存在すると判定された場合、ステップ S 4 5 乃至 S 5 2 をスキップして、ステップ S 5 3 に進む。

40

【 0 1 9 1 】

また、ステップ S 4 4 において、注目フレームのサブバンド $sb\#n$ に、生徒データとしてのサブバンドデータが存在しないと判定された場合、即ち、注目フレームのサブバンド $sb\#n$ に対するビット割り当てが 0 である場合、ステップ S 4 5 に進み、コントローラ 1 3 3 は、その注目フレームのサブバンド $sb\#n$ を、注目サブバンドとして選択し、ステップ S 4 6 に進む。

【 0 1 9 2 】

ステップ S 4 6 では、コントローラ 1 3 3 は、サブバンドの高音質サブバンドデータのデータ位置を表す変数 i に、初期値としての、例えば 0 をセットし、ステップ S 4 7 に進む。

50

【0193】

ステップS47では、コントローラ133は、注目サブバンドsb#nの、データ位置#iで特定される高音質サブバンドデータとしての教師データd#iを、注目データとして選択する。さらに、ステップS47では、コントローラ133は、注目データd#iを特定する特定情報として、注目データd#iのデータ位置#iと、注目サブバンドsb#nのサブバンドナンバ#nを、学習部134に供給し、ステップS48に進む。

【0194】

ステップS48では、タップ生成部151と152が、コントローラ133からの特定情報としてのデータ位置#iとサブバンドナンバ#nで特定される高音質サブバンドデータとしての教師データd#iを、注目データd#iとして認識し、その注目データd#iについて、生徒データ生成部132から供給される生徒データとしての低音質サブバンドデータとサイドインフォメーションを用い、図6のタップ生成部121と122における場合と同一のタップ構造の予測タップとクラスタップを、それぞれ生成する。

10

【0195】

ステップS48で得られた予測タップは、タップ生成部151から足し込み部154に供給され、また、クラスタップは、タップ生成部152からクラス分類部153に供給され、その後、ステップS48からステップS49に進む。

【0196】

ステップS49では、クラス分類部153が、タップ生成部152から供給されるクラスタップの他、コントローラ133から供給される注目サブバンドsb#nのサブバンドナンバ#n、および注目データd#iのデータ位置#iに基づき、注目データについて、図6のクラス分類部123と同一のクラス分類を行い、その結果得られる注目データのクラスを表すクラスコードを、足し込み部154に供給して、ステップS50に進む。

20

【0197】

ステップS50では、足し込み部154が、教師データ生成部131から供給される注目フレームの教師データのうちの注目データとなっているものを選択する。さらに、ステップS50では、足し込み部154が、クラス分類部153から供給されるクラスコードが表す注目データのクラスごとに、タップ生成部151から供給される予測タップ、および注目データを対象とした、上述した式(8)における足し込みを行い、ステップS51に進む。

30

【0198】

ステップS51では、コントローラ133が、変数iが、データ位置の最大値である11に等しいかどうかを判定し、11に等しくないと判定した場合、ステップS52に進む。ステップS52では、コントローラ133は、変数iを1だけインクリメントして、ステップS47に戻り、以下、同様の処理が繰り返される。

【0199】

また、ステップS51において、変数iが、データ位置の最大値である11に等しいと判定された場合、即ち、注目サブバンドsb#nのすべての高音質サブバンドデータを注目データとして、ステップS50の足し込みが行われた場合、ステップS53に進み、コントローラ133は、変数nが、1フレームのサブバンドのサブバンドナンバの最大値である31に等しいかどうかを判定する。ステップS53において、変数nが、サブバンドナンバの最大値である31に等しくないと判定された場合、ステップS54に進み、コントローラ133は、変数nを1だけインクリメントして、ステップS44に戻り、以下、同様の処理が繰り返される。

40

【0200】

また、ステップS53において、変数nが、サブバンドナンバの最大値である31に等しいと判定された場合、即ち、注目フレームのサブバンドすべてについて、そのサブバンドデータがあるかどうかを判定し、サブバンドデータがないサブバンドを注目サブバンドとして、ステップS50の足し込みを行った場合、ステップS55に進み、教師データ生成部131は、注目フレームの次のフレームのオーディオデータが、学習用データベース1

50

4 1 に記憶されているかどうかを判定する。ステップ S 5 5 において、注目フレームの次のフレームのオーディオデータが、学習用データベース 1 4 1 に記憶されていると判定された場合、その注目フレームの次のフレームを、新たに注目フレームとして、ステップ S 4 1 に戻り、以下、同様の処理が繰り返される。

【0201】

また、ステップ S 5 5 において、注目フレームの次のフレームのオーディオデータが、学習用データベース 1 4 1 に記憶されていないと判定された場合、即ち、学習用データベース 1 4 1 に記憶されたオーディオデータのフレームすべてを注目フレームとして、ステップ S 5 0 の足し込みが行われた場合、足し込み部 1 5 4 は、いままでのステップ S 5 0 における足し込みによって、クラスごとに得られた式 (8) の正規方程式を、タップ係数演算部 1 5 5 に供給して、ステップ S 5 6 に進む。

10

【0202】

ステップ S 5 6 では、タップ係数演算部 1 5 5 は、足し込み部 1 5 4 から供給される、各クラスごとの式 (8) の正規方程式を解くことにより、クラスごとのタップ係数を求め、タップ係数記憶部 1 5 6 に供給して記憶させ、処理を終了する。

【0203】

以上のようにして、タップ係数記憶部 1 5 6 に記憶されたクラスごとのタップ係数が、図 6 の復号装置 1 0 2 におけるタップ係数記憶部 1 2 4 に記憶されている。

【0204】

なお、以上のようなタップ係数の学習処理において、用意する学習用データ等によっては、タップ係数を求めるのに必要な数の正規方程式が得られないクラスが生じる場合があり得るが、そのようなクラスについては、例えば、タップ係数演算部 1 5 5 において、デフォルトのタップ係数を出力するようにすること等が可能である。あるいは、タップ係数を求めるのに必要な数の正規方程式が得られないクラスが生じた場合には、新たに学習用データを用意して学習用データベース 1 4 1 に記憶させ、再度、タップ係数の学習を行うようにしても良い。

20

【0205】

以上のように、図 1 1 の学習装置では、オーディオデータをサブバンド符号化して得られるサブバンドデータを、教師データとするとともに、そのオーディオデータを MPEG1 オーディオレイヤ 1 方式で符号化して復号することにより得られるサブバンドデータ (さらには、サイドインフォメーション) を生徒データとして、その教師データと生徒データとの関係を学習することにより、式 (1) によって、教師データと生徒データとを関係付けるタップ係数を、クラスごとに求めるようにしたので、図 6 の復号装置 1 0 2 において、そのようなタップ係数を用いて、ビット割り当てが 0 とされたサブバンドのサブバンドデータを求めることにより、高音質のオーディオデータを復号することが可能となる。

30

【0206】

次に、図 1 3 は、図 5 の復号装置 1 0 2 の第 2 の構成例を示している。なお、図中、図 6 における場合と対応する部分については、同一の符号を付してあり、以下では、その説明は、適宜省略する。即ち、図 1 3 の復号装置 1 0 2 は、基本的に、図 6 における場合と同様に構成されている。但し、サブバンド合成部 1 1 4 には、デコーダ 1 1 1 が出力する低音質サブバンドデータが供給されないようになっている。さらに、クラス分類適応処理部 1 1 3 は、ビット割り当てが 0 のサブバンドの高音質サブバンドデータの他、ビット割り当てが 1 ビット以上のサブバンドの高音質サブバンドデータも求めて、サブバンド合成部 1 1 4 に供給するようになっている。即ち、クラス分類適応処理部 1 1 3 は、すべてのサブバンドについて、高音質サブバンドデータを求めて、サブバンド合成部 1 1 4 に供給するようになっている。

40

【0207】

なお、図 1 3 の実施の形態では、クラス分類適応処理部 1 1 3 からサブバンド合成部 1 1 4 に対して、すべてのサブバンドの高音質サブバンドデータが供給されることから、図 6 の実施の形態における場合のように、サブバンド合成部 1 1 4 において、クラス分類適

50

処理部 1 1 3 から供給されるサブバンドデータのサブバンドが、いずれのサブバンドナンバのものであるかを、特に認識する必要がないので、サブバンド合成部 1 1 4 には、コントローラ 1 1 2 が出力する特定情報としてのサブバンドナンバとデータ位置が供給されないようになっている（供給する必要がない）。

【 0 2 0 8 】

次に、図 1 4 のフローチャートを参照して、図 1 3 の復号装置 1 0 2 による復号処理について説明する。

【 0 2 0 9 】

図 1 3 の復号装置 1 0 2 では、上述したように、クラス分類適応処理部 1 1 3 は、すべてのサブバンドについて、高音質サブバンドデータを求めて、サブバンド合成部 1 1 4 に供給するようになっている。

10

【 0 2 1 0 】

即ち、デコーダ 1 1 1 は、あるフレームの符号化ビットストリームを受信すると、そのフレームを、注目フレームとし、ステップ S 6 1 において、その注目フレームについて、図 1 0 のステップ S 1 における場合と同様の通常デコード処理を行う。

【 0 2 1 1 】

ステップ S 6 1 において、デコーダ 1 1 1 が通常デコード処理を行うことにより得られる注目フレームのサイドインフォメーションは、コントローラ 1 1 2 とクラス分類適応処理部 1 1 3 に供給され、さらに、注目フレームの低音質サブバンドデータ（MPEG1 オーディオレイヤ 1 方式で符号化されて復号されたサブバンドデータ）は、クラス分類適応処理部 1 1 3 に供給される。

20

【 0 2 1 2 】

そして、ステップ S 6 2 に進み、コントローラ 1 1 2 は、注目フレームのサブバンドのサブバンドナンバを表す変数 n に、初期値としての、例えば 0 をセットし、ステップ S 6 3 に進む。ステップ S 6 3 では、コントローラ 1 1 2 は、注目フレームのサブバンド $sb\#n$ を、注目サブバンドとして選択し、ステップ S 6 4 に進む。

【 0 2 1 3 】

ステップ S 6 4 では、コントローラ 1 1 2 は、サブバンドの高音質サブバンドデータのデータ位置を表す変数 i に、初期値としての、例えば 0 をセットし、ステップ S 6 5 に進む。

30

【 0 2 1 4 】

ステップ S 6 5 では、コントローラ 1 1 2 は、注目サブバンド $sb\#n$ の、データ位置 $\#i$ で特定される高音質サブバンドデータ $d\#i$ を、注目データとして選択する。さらに、ステップ S 6 5 では、コントローラ 1 1 2 は、注目データ $d\#i$ を特定する特定情報として、注目データ $d\#i$ のデータ位置 $\#i$ と、注目サブバンド $sb\#n$ のサブバンドナンバ $\#n$ を、クラス分類適応処理部 1 1 3 に供給し、ステップ S 6 6 に進む。

【 0 2 1 5 】

ステップ S 6 6 では、クラス分類適応処理部 1 1 3 が、注目データ $d\#i$ を求めるクラス分類適応処理を行う。即ち、ステップ S 6 6 では、クラス分類適応処理部 1 1 3 が、図 1 0 のステップ S 7 における場合と同様のクラス分類適応処理を行うことにより、注目データ $d\#i$ を求め、ステップ S 6 7 に進む。

40

【 0 2 1 6 】

ステップ S 6 7 では、コントローラ 1 1 2 が、変数 i が、データ位置の最大値である 1 1 に等しいかどうかを判定し、1 1 に等しくないとは判定した場合、ステップ S 6 8 に進む。ステップ S 6 8 では、コントローラ 1 1 2 は、変数 i を 1 だけインクリメントして、ステップ S 6 5 に戻り、以下、同様の処理が繰り返される。

【 0 2 1 7 】

また、ステップ S 6 7 において、変数 i が、データ位置の最大値である 1 1 に等しいとは判定された場合、即ち、注目サブバンド $sb\#n$ のすべての高音質サブバンドデータを注目データとして、そのサンプル値が、クラス分類適応処理によって求められた場合、クラス分類

50

適応処理部 113 の予測演算部 125 は、その注目サブバンド sb#n の高音質サブバンドデータを、サブバンド合成部 114 に供給して、ステップ S69 に進む。

【0218】

ステップ S69 では、コントローラ 112 は、変数 n が、1 フレームのサブバンドのサブバンドナンバの最大値である 31 に等しいかどうかを判定する。ステップ S69 において、変数 n が、サブバンドナンバの最大値である 31 に等しくないとは判定された場合、ステップ S70 に進み、コントローラ 112 は、変数 n を 1 だけインクリメントして、ステップ S63 に戻り、以下、同様の処理が繰り返される。

【0219】

また、ステップ S69 において、変数 n が、サブバンドナンバの最大値である 31 に等しいとは判定された場合、即ち、注目フレームのサブバンドすべてについて、クラス分類適応処理によって、高音質サブバンドデータを求めた場合、ステップ S71 に進み、サブバンド合成部 114 は、クラス分類適応処理部 113 から供給された、クラス分類適応処理によってサブバンドデータが生成された注目フレームのサブバンドを合成することにより、注目フレームの時間領域のオーディオデータを復号して出力する。

10

【0220】

従って、この場合、ビット割り当てが 0 とされたサブバンド（ヌルサブバンド）についてだけでなく、すべてのサブバンドについて、クラス分類適応処理によって高音質サブバンドデータが生成され、その高音質サブバンドデータを用いて、時間領域のオーディオデータが復号される。即ち、この場合、符号化装置 101（図 5）においてビット割り当てが 0 とされたサブバンドについては、そのサブバンドのサブバンドデータが、いわば創造され、サブバンド合成に用いられる。また、符号化装置 101 においてビット割り当てが 1 以上とされたサブバンドについては、符号化装置 101 における量子化などに起因して生じる歪みを除去（軽減）したサブバンドデータが創造され、サブバンド合成に用いられる。その結果、より高音質のオーディオデータを、符号化ビットストリームのデータ量を増加させずに復号することができる。

20

【0221】

ステップ S71 の処理後は、ステップ S72 に進み、デコーダ 111 は、注目フレームの次のフレームの符号化ビットストリームが供給されたかどうかを判定する。ステップ S72 において、注目フレームの次のフレームの符号化ビットストリームが供給されたと判定された場合、その符号化ビットストリームが供給されたフレームを、新たに注目フレームとして、ステップ S61 に戻り、以下、同様の処理が繰り返される。

30

【0222】

また、ステップ S72 において、注目フレームの次のフレームの符号化ビットストリームが供給されていないとは判定された場合、処理を終了する。

【0223】

次に、図 13 の復号装置 102 のクラス分類適応処理部 113 では、ヌルサブバンドの高音質サブバンドデータだけでなく、すべてのサブバンドの高音質サブバンドデータが求められる。このため、図 13 のタップ係数記憶部 124 は、ヌルサブバンドの高音質サブバンドデータだけでなく、すべてのサブバンドの高音質サブバンドデータを求めるためのタップ係数を記憶している。

40

【0224】

ここで、図 13 の復号装置 102 のクラス分類適応処理部 113 では、オーディオデータを MPEG1 オーディオレイヤ 1 方式で符号化して復号することにより得られるサブバンドデータ（低音質サブバンドデータ）から予測タップを生成し、その予測タップを用いて、式（1）の積和演算を行うことにより、オーディオデータを単にサブバンド符号化して得られるサブバンドデータ（高音質サブバンドデータ）が求められる。

【0225】

従って、式（1）の積和演算に用いるタップ係数としては、オーディオデータを単にサブバンド符号化して得られるサブバンドデータを、教師データとするとともに、オーディオ

50

データをMPEG1オーディオレイヤ1方式で符号化して復号することにより得られるサブバンドデータを生徒データとして、その生徒データと教師データとを、式(1)によって関係付けるものを求めれば良い。

【0226】

このような教師データと生徒データとを関係付けるタップ係数は、上述した図11の学習装置において求めることができるが、図11の学習装置で行われる図12の学習処理では、ヌルサブバンドだけが注目サブバンドとされ、ヌルサブバンドについて得られる教師データと生徒データを用いて、タップ係数が求められる。

【0227】

一方、図13の復号装置102では、ヌルサブバンドの高音質サブバンドデータだけでなく、すべてのサブバンドの高音質サブバンドデータが求められるため、その高音質サブバンドデータを求めるのに用いるタップ係数は、ヌルサブバンドだけでなく、すべてのサブバンドについて得られる教師データと生徒データを用いて求める(学習する)必要がある。

【0228】

そこで、図13の復号装置102のタップ係数記憶部124に記憶させるタップ係数は、図11の学習装置に、図12の学習処理に代えて、図15のフローチャートにしたがった学習処理を行わせることで求めることができる。

【0229】

即ち、図15の学習処理では、教師データ生成部131は、学習用データベース141に記憶されたオーディオデータのフレームのうちの1フレームを注目フレームとし、ステップS81において、図12のステップS41における場合と同様に、注目フレームのオーディオデータをMPEG1オーディオレイヤ1方式で符号化し、その符号化の過程で得られる注目フレームの高音質サブバンドデータを、教師データとして、学習部134の足し込み部154に供給するとともに、その符号化の結果得られる注目フレームの符号化ビットストリームを生徒データ生成部132に供給する。

【0230】

その後、ステップS82に進み、生徒データ生成部132は、図12のステップS42における場合と同様に、教師データ生成部131から供給される注目フレームの符号化ビットストリームを、MPEG1オーディオレイヤ1方式で復号し、その復号過程で得られる低音質サブバンドデータとサイドインフォメーションを、生徒データとして、学習部134に供給する。さらに、ステップS82では、生徒データ生成部132は、サイドインフォメーションを、コントローラ133にも供給する。

【0231】

そして、ステップS83に進み、コントローラ133は、注目フレームのサブバンドのサブバンドナンバを表す変数nに、初期値としての、例えば0をセットし、ステップS84に進む。ステップS84では、コントローラ133は、注目フレームのサブバンドsb#nを、注目サブバンドとして選択し、ステップS85に進む。

【0232】

ステップS85では、コントローラ133は、サブバンドの高音質サブバンドデータのデータ位置を表す変数iに、初期値としての、例えば0をセットし、ステップS86に進む。

【0233】

ステップS86では、コントローラ133は、注目サブバンドsb#nの、データ位置#iで特定される高音質サブバンドデータとしての教師データd#iを、注目データとして選択する。そして、ステップS87に進み、以下、ステップS87乃至95において、図12のステップS48乃至S56における場合とそれぞれ同様の処理が行われ、処理を終了する。

【0234】

図15の学習処理では、図12の学習処理における場合のように、サブバンドsb#nがヌルサブバンドであるかどうかを判定せず(図12のS45)、すべてのサブバンドが注目サ

10

20

30

40

50

ブバンドとして処理されるので、すべてのサブバンドについて得られる教師データと生徒データを用いて、タップ係数が求められる。

【0235】

従って、図13の復号装置102では、図15の学習処理により得られるタップ係数を用いて、すべてのサブバンドについて、教師データと同音質の高音質サブバンドデータが求められるので、より高音質のオーディオデータを復号することができる。

【0236】

次に、図16は、図5の復号装置102の第3の構成例を示している。なお、図中、図6または図13における場合と対応する部分については、同一の符号を付してあり、以下では、その説明は、適宜省略する。即ち、図16の復号装置102は、サブバンド合成部114が設けられていない他は、図13における場合と、基本的に同様に構成されている。

10

【0237】

但し、図6および図13の復号装置102は、周波数領域のオーディオデータである低音質サブバンドデータから高音質サブバンドデータを求め、その高音質サブバンドデータを用いて、サブバンド合成部114においてサブバンド合成を行うことにより、高音質の時間領域のオーディオデータを復号するようになっていたが、図16の復号装置102は、周波数領域のオーディオデータである低音質サブバンドデータから、いわば直接、高音質の時間領域のオーディオデータを復号するようになっている。

【0238】

即ち、図16の実施の形態では、クラス分類適応処理113は、クラス分類適応処理によって、周波数領域のオーディオデータである低音質サブバンドデータを、高音質の時間領域のオーディオデータ（以下、適宜、高音質オーディオデータという）に変換する。

20

【0239】

そのため、図16の復号装置102では、コントローラ112は、クラス分類適応処理部113において求めようとする高音質の時間領域のオーディオデータを、順次、注目データとして選択する。さらに、コントローラ112は、注目データのフレームにおける、その注目データのデータ位置などを、注目データを特定するための特定情報として、クラス分類適応処理部113に供給する。また、コントローラ112は、デコーダ111から供給されるサイドインフォメーションに基づき、注目データのフレームである注目フレームにおいてサブバンドデータが抜けているサブバンド、即ち、符号化装置101（図5）においてビット割り当てが0ビットとされたサブバンド（ヌルサブバンド）があるかどうかを判定し、さらに、ヌルサブバンドがある場合には、そのヌルサブバンドを表すサブバンドナンバを認識し、クラス分類適応処理部113に供給する。

30

【0240】

クラス分類適応処理部113は、デコーダ111から供給されるサブバンドデータ（低音質サブバンドデータ）、およびコントローラ112から供給される注目データのデータ位置と注目フレームにおけるヌルサブバンドのサブバンドナンバなどを用いて、クラス分類適応処理を行うことにより、注目データとしての高音質オーディオデータを求めて（予測して）出力する。

【0241】

40

即ち、図16のクラス分類適応処理部113において、タップ生成部121は、コントローラ112から供給されるデータ位置によって特定される高音質オーディオデータを注目データとして選択し、その注目データ（の予測値）を求めるのに用いる予測タップを、デコーダ111から供給される、MPEG1オーディオレイヤ1方式により復号されたサブバンドデータである低音質サブバンドデータ、さらには、やはり、デコーダ111から供給されるサイドインフォメーションから生成し、予測演算部125に供給する。

【0242】

具体的には、タップ生成部121は、例えば、注目データのフレームである注目フレームの低音質サブバンドデータの幾つか、さらには注目フレームのサイドインフォメーションの幾つかを抽出し、それらを所定の順番で並べたものを、予測タップとする。

50

【 0 2 4 3 】

ここで、本実施の形態においては、少なくとも、注目フレームの低音質サブバンドデータのすべてを含む予測タップを構成するのが望ましい。

【 0 2 4 4 】

即ち、本実施の形態において、符号化装置 1 0 1 (図 5) では、上述したように、オーディオデータがMPEG1オーディオレイヤ 1 方式で符号化される。MPEG1オーディオレイヤ 1 方式では、図 1 7 に示すように、3 8 4 サンプルのオーディオデータ s_0, s_1, \dots, s_{383} を 1 フレームとして、その 1 フレームのオーディオデータがサブバンド符号化 (分割) されることにより、3 2 のサブバンド $sb_0, sb_1, \dots, sb_{31}$ のサブバンドデータとされる。従って、3 2 のサブバンド sb_0 乃至 sb_{31} それぞれには、そのフレームの時間領域のオーディオデータすべての周波数成分が反映されるから、注目データとされた時間領域のオーディオデータを予測するのに、3 2 のサブバンド sb_0 乃至 sb_{31} のうちの一部のサブバンドデータしか用いないのは、注目データの予測精度が劣化することがあり得る。このため、図 1 6 の復号装置 1 0 2 では、予測タップは、注目フレームの低音質サブバンドデータのすべてを含むものとするのが望ましい。

10

【 0 2 4 5 】

なお、図 1 7 に示すように、1 つのサブバンドは、1 2 サンプルのサブバンドデータから構成され、さらに、1 フレームは、3 2 のサブバンドから構成されるから、予測タップを、注目フレームの低音質サブバンドデータのすべてを含めて構成した場合、予測タップのタップ数は、少なくとも、3 8 4 (= 1 2 × 3 2) タップとなる。

20

【 0 2 4 6 】

また、予測タップは、注目フレームのサブバンドデータやサイドインフォメーションの他、例えば、注目フレームに近接するフレームのサブバンドデータやサイドインフォメーションなどを含めて構成することが可能である。

【 0 2 4 7 】

図 1 6 に戻り、タップ生成部 1 2 2 は、タップ生成部 1 2 1 と同様に、コントローラ 1 1 2 から供給されるデータ位置の高音質オーディオデータを、注目データとして選択し、その注目データのクラス分類に用いるクラスタップを、デコーダ 1 1 1 から供給される低音質サブバンドデータとサイドインフォメーションから生成し、クラス分類部 1 2 3 に出力する。

30

【 0 2 4 8 】

クラス分類部 1 2 3 は、タップ生成部 1 2 2 からのクラスタップ、さらには、例えば、コントローラ 1 1 2 から供給される注目フレームにおけるヌルサブバンドのサブバンドナンバーや、注目データのデータ位置に基づいて、注目データをクラス分類し、その結果得られる注目データのクラスに対応するクラスコードを、タップ係数記憶部 1 2 4 に供給する。

【 0 2 4 9 】

タップ係数記憶部 1 2 4 は、後述する図 1 9 の学習装置によって、学習の教師となる高音質オーディオデータである教師データと、学習の生徒となる低音質サブバンドデータである生徒データとの関係を、1 以上のクラスごとに学習することにより得られたタップ係数を記憶している。そして、タップ係数記憶部 1 2 4 は、クラス分類部 1 2 3 から注目データのクラスコードが供給されると、そのクラスコードに対応するクラスのタップ係数を読み出すことにより、注目データのクラスのタップ係数を取得し、予測演算部 1 2 5 に供給する。

40

【 0 2 5 0 】

予測演算部 1 2 5 は、タップ係数記憶部 1 2 4 から供給される、注目データのクラスのタップ係数 w_1, w_2, \dots と、タップ生成部 1 2 1 からの注目データの予測タップ (を構成する各低音質サブバンドデータのサンプル値) x_1, x_2, \dots とを用いて、式 (1) に示した積和演算を行うことにより、注目データ y (の予測値) として的高音質オーディオデータを求める。予測演算部 1 2 5 は、以上のようにして、1 フレームを構成する高音質オーディオデータのサンプル値すべてが得られると、その 1 フレームの高音質オーディ

50

オデータを出力する。

【0251】

次に、図18のフローチャートを参照して、図16の復号装置102の処理（復号処理）について説明する。

【0252】

デコーダ111は、あるフレームの符号化ビットストリームを受信すると、そのフレームを、注目フレームとし、ステップS101において、その注目フレームについて、図10のステップS1における場合と同様の通常デコード処理を行う。

【0253】

ステップS101の通常デコード処理において得られる、MPEG1オーディオレイヤ1方式で符号化されて復号された注目フレームのサブバンドデータ、即ち、低音質サブバンドデータは、クラス分類適応処理部113に供給され、また、注目フレームのサイドインフォメーションは、コントローラ112とクラス分類適応処理部113に供給される。

10

【0254】

その後は、ステップS102に進み、コントローラ112は、注目フレームの高音質オーディオデータのデータ位置を表す変数mに、初期値としての、例えば0をセットし、ステップS103に進む。

【0255】

ステップS103では、コントローラ112は、注目フレームの、データ位置#mで特定される高音質オーディオデータs#mを、注目データとして選択する。さらに、ステップS103では、コントローラ112は、注目データs#mを特定するデータ位置#mと、注目フレームにおけるヌルサブバンドのサブバンドナンバを、クラス分類適応処理部113に供給し、ステップS104に進む。

20

【0256】

ステップS104では、クラス分類適応処理部113が、注目データs#mを求めるクラス分類適応処理を行う。

【0257】

即ち、ステップS104のクラス分類適応処理では、まず最初に、ステップS121において、タップ生成部121と122が、コントローラ112からのデータ位置#mで特定される高音質オーディオデータs#mを、注目データs#mとして認識し、その注目データs#mについて、デコーダ111から供給される低音質サブバンドデータ、さらには、必要に応じてサイドインフォメーションを用い、予測タップとクラスタップを生成する。

30

【0258】

ステップS121で得られた予測タップは、タップ生成部121から予測演算部125に供給され、また、クラスタップは、タップ生成部121からクラス分類部123に供給され、その後、ステップS121からステップS122に進む。

【0259】

ステップS122では、クラス分類部123が、タップ生成部122から供給されるクラスタップの他、コントローラ112から供給される注目フレームにおけるヌルサブバンドのサブバンドナンバ、および注目データs#mのデータ位置#mに基づいて、注目データをクラス分類し、その結果得られる注目データのクラスを表すクラスコードを、タップ係数記憶部124に供給して、ステップS123に進む。

40

【0260】

ステップS123では、タップ係数記憶部124が、そこに記憶されているクラスごとのタップ係数から、クラス分類部123から供給されるクラスコードが表す注目データs#mのクラスのタップ係数を読み出すことにより取得し、その注目データs#mのクラスのタップ係数を、予測演算部125に供給して、ステップS124に進む。

【0261】

ステップS124では、予測演算部125が、タップ生成部121から供給される注目データs#mについての予測タップと、タップ係数記憶部124から供給される注目データs#m

50

のクラスのタップ係数とを用いて、式(1)の積和演算を行うことにより、注目データ s_m を求める。

【0262】

以上のステップS121乃至S124の処理によって、ステップS104のクラス分類適応処理が終了し、ステップS105に進む。ステップS105では、コントローラ112が、変数 m が、時間領域の1フレームのオーディオデータのデータ位置の最大値である383(図17)に等しいかどうかを判定し、383に等しくないと判定した場合、ステップS106に進む。ステップS106では、コントローラ112は、変数 m を1だけインクリメントして、ステップS103に戻り、以下、同様の処理が繰り返される。

【0263】

また、ステップS105において、変数 m が、データ位置の最大値である383に等しいと判定された場合、即ち、注目フレームのすべての高音質オーディオデータを注目データとして、そのサンプル値が、クラス分類適応処理によって求められた場合、クラス分類適応処理部113の予測演算部125は、その注目フレームの高音質オーディオデータを出力する。

【0264】

従って、この場合、符号化装置101(図5)が出力する符号化ビットストリームのデータ量を増加させずに、高音質のオーディオデータを復号することができる。

【0265】

その後、ステップS107に進み、デコーダ111は、注目フレームの次のフレームの符号化ビットストリームが供給されたかどうかを判定する。ステップS107において、注目フレームの次のフレームの符号化ビットストリームが供給されたと判定された場合、その符号化ビットストリームが供給されたフレームを、新たに注目フレームとして、ステップS101に戻り、以下、同様の処理が繰り返される。

【0266】

また、ステップS107において、注目フレームの次のフレームの符号化ビットストリームが供給されていないと判定された場合、処理を終了する。

【0267】

次に、図19は、図16の復号装置102のタップ係数記憶部124に記憶させるクラスごとのタップ係数を学習する学習装置の一実施の形態の構成例を示している。なお、図中、図11における場合と対応する部分については、同一の符号を付してあり、以下では、その説明は、適宜省略する。即ち、図19の学習装置は、基本的に、図11の学習装置と同様に構成されている。

【0268】

但し、図11の学習装置では、教師データ生成部131が学習用データベース141とエンコーダ142で構成されているのに対して、図19の学習装置では、教師データ生成部131が学習用データベース141で構成されている。また、図11の学習装置では、生徒データ生成部132がデコーダ143で構成されているのに対して、図19の学習装置では、生徒データ生成部132がエンコーダ142とデコーダ143で構成されている。

【0269】

ここで、図16の復号装置102におけるクラス分類適応処理部113では、オーディオデータをMPEG1オーディオレイヤ1方式で符号化して復号することにより得られるサブバンドデータ(低音質サブバンドデータ)から予測タップを生成し、その予測タップを用いて、式(1)の積和演算を行うことにより、高音質オーディオデータが求められる。

【0270】

従って、式(1)の積和演算に用いるタップ係数としては、高音質オーディオデータを、教師データとするとともに、その高音質オーディオデータをMPEG1オーディオレイヤ1方式で符号化して復号することにより得られる低音質サブバンドデータを生徒データとして、その生徒データと教師データとを、式(1)によって関係付けるものを求めれば良い。

【0271】

10

20

30

40

50

そこで、図 19 の学習装置では、教師データ生成部 131 は、学習用データベース 141 で構成されており、その学習用データベース 141 に記憶されているオーディオデータを読み出し、教師データとして、生徒データ生成部 132 と、学習部 134 の足し込み部 154 に供給する。

【0272】

生徒データ生成部 132 は、エンコーダ 142 とデコーダ 143 で構成されている。エンコーダ 142 は、教師データ生成部 131 から供給される教師データとしての高音質オーディオデータを MPEG1 オーディオレイヤ 1 方式で符号化し、その符号化によって得られる符号化ビットストリームを、デコーダ 143 に供給する。デコーダ 143 は、エンコーダ 142 が出力する符号化ビットストリームを MPEG1 オーディオレイヤ 1 方式で復号し、その復号の過程で得られるサブバンドデータ（低音質サブバンドデータ）とサイドインフォメーションを、生徒データとして、学習部 134 に供給する。さらに、デコーダ 143 は、符号化ビットストリームを MPEG1 オーディオレイヤ 1 方式で復号する過程で得られるサイドインフォメーションを、コントローラ 133 に供給する。

10

【0273】

コントローラ 133 は、デコーダ 143 から供給されるサイドインフォメーションのフレームを注目フレームとして、その注目フレームにおいて、ビット割り当てが 0 ビットとされたサブバンド（ヌルサブバンド）がある場合には、そのヌルサブバンドを表すサブバンドナンバを、学習部 134 に供給する。また、コントローラ 133 は、注目フレームの高音質データとしての教師データを、順次、注目データとして選択し、その注目データのデータ位置を、学習部 134 に供給する。

20

【0274】

学習部 134 は、教師データ生成部 131 から供給される教師データとしての高音質オーディオデータと、生徒データ生成部 132 から供給される生徒データとしての低音質サブバンドデータ（およびサイドインフォメーション）との関係を学習することにより、式（1）により、教師データと生徒データとを関係付けるタップ係数を、クラスごとに求める。

【0275】

即ち、タップ生成部 151 と 152 は、コントローラ 133 が出力するデータ位置から、注目フレームにおいて注目データとなっている教師データを認識し、その注目データについて、生徒データ生成部 143 から供給される生徒データから、図 16 のタップ生成部 121 と 122 における場合と同一のタップ構造の予測タップとクラスタップを、それぞれ生成する。そして、タップ生成部 151 は、注目データの予測タップを、足し込み部 154 に供給し、タップ生成部 152 は、注目データのクラスタップを、クラス分類部 153 に供給する。

30

【0276】

クラス分類部 153 は、タップ生成部 152 から供給されるクラスタップと、コントローラ 133 から供給される注目データのデータ位置と、注目フレームにおけるヌルサブバンドのサブバンドナンバに基づき、図 16 のクラス分類部 123 における場合と同一のクラス分類を行う。クラス分類部 153 でクラス分類が行われることにより得られる注目データのクラスを表すクラスコードは、クラス分類部 153 から足し込み部 154 に供給される。

40

【0277】

足し込み部 154 およびタップ係数演算部 155 は、図 11 における場合と同様に、教師データ生成部 131 から供給される教師データのうちの注目データと、タップ生成部 151 から供給される予測タップとを用い、教師データと生徒データとの関係を、クラス分類部 153 から供給されるクラスごとに学習することにより、クラスごとのタップ係数を求める。タップ係数記憶部 156 は、図 11 における場合と同様に、タップ係数演算部 155 で求められるクラスごとのタップ係数を記憶する。

【0278】

50

次に、図20フローチャートを参照して、図19の学習装置において行われる、クラスごとのタップ係数を求める学習処理について説明する。

【0279】

まず最初に、教師データ生成部131は、学習用データベース141に記憶された高音質オーディオデータのフレームのうちの1フレームを注目フレームとし、ステップS131において、その注目フレームの教師データを生成する。即ち、教師データ生成部131は、学習用データベース141から、注目フレームの高音質オーディオデータを読み出し、教師データとして、生徒データ生成部132と、学習部134の足し込み部154に供給する。

【0280】

その後、ステップS132に進み、生徒データ生成部132は、教師データ生成部131から供給される注目フレームの高音質オーディオデータから生徒データを生成する。即ち、ステップS132では、生徒データ生成部132のデコーダ143が、注目フレームの高音質オーディオデータをMPEG1オーディオレイヤ1方式で符号化し、その符号化により得られる符号化ビットストリームを、デコーダ143に供給する。デコーダ143は、その符号化ビットストリームをMPEG1オーディオレイヤ1方式で復号し、その復号過程で得られる低音質サブバンドデータとサイドインフォメーションを、生徒データとして、学習部134に供給する。さらに、デコーダ143は、サイドインフォメーションを、コントローラ133にも供給する。

【0281】

そして、ステップS133に進み、コントローラ133は、注目フレームの高音質オーディオデータのデータ位置を表す変数*i*に、初期値としての、例えば0をセットし、ステップS134に進む。

【0282】

ステップS134では、コントローラ133は、注目フレームの、データ位置#*m*で特定される高音質オーディオデータとしての教師データ*s*#*m*を、注目データとして選択する。さらに、ステップS134では、コントローラ133は、注目データ*s*#*m*のデータ位置#*m*と、注目フレームにおけるヌルサブバンドのサブバンドナンバ#*n*を、学習部134に供給し、ステップS135に進む。

【0283】

ステップS135では、タップ生成部151と152が、コントローラ133からのデータ位置#*m*で特定される高音質オーディオデータとしての教師データ*s*#*m*を、注目データ*s*#*m*として認識し、その注目データ*s*#*m*について、生徒データ生成部132から供給される生徒データとしての低音質サブバンドデータ（とサイドインフォメーション）を用い、図16のタップ生成部121と122における場合と同一のタップ構造の予測タップとクラスタップを、それぞれ生成する。

【0284】

ステップS135で得られた予測タップは、タップ生成部151から足し込み部154に供給され、また、クラスタップは、タップ生成部151からクラス分類部153に供給され、その後、ステップS135からステップS136に進む。

【0285】

ステップS136では、クラス分類部153が、タップ生成部152から供給されるクラスタップの他、コントローラ133から供給される注目フレームにおけるヌルサブバンドのサブバンドナンバ、および注目データ*s*#*m*のデータ位置#*m*に基づき、注目データについて、図16のクラス分類部123と同一のクラス分類を行い、その結果得られる注目データのクラスを表すクラスコードを、足し込み部154に供給して、ステップS137に進む。

【0286】

ステップS137では、足し込み部154が、教師データ生成部131から供給される注目フレームの教師データのうちの注目データとなっているものを選択する。さらに、ステ

10

20

30

40

50

ップS 37では、足し込み部154が、クラス分類部153から供給されるクラスコードが表す注目データのクラスごとに、タップ生成部151から供給される予測タップ、および注目データを対象とした、上述した式(8)における足し込みを行い、ステップS 138に進む。

【0287】

ステップS 138では、コントローラ133が、変数mが、1フレームにおける高音質オーディオデータのデータ位置の最大値である383に等しいかどうかを判定し、383に等しくないと判定した場合、ステップS 139に進む。ステップS 139では、コントローラ133は、変数mを1だけインクリメントして、ステップS 134に戻り、以下、同様の処理が繰り返される。

10

【0288】

また、ステップS 138において、変数mが、データ位置の最大値である383に等しいと判定された場合、即ち、注目フレームの高音質オーディオデータとしての教師データすべてを注目データとして、ステップS 137の足し込みが行われた場合、ステップS 140に進み、教師データ生成部131は、注目フレームの次のフレームの高音質オーディオデータが、学習用データベース141に記憶されているかどうかを判定する。ステップS 140において、注目フレームの次のフレームの高音質オーディオデータが、学習用データベース141に記憶されていると判定された場合、その注目データの次のフレームを、新たに注目フレームとして、ステップS 131に戻り、以下、同様の処理が繰り返される。

20

【0289】

また、ステップS 140において、注目フレームの次のフレームの高音質オーディオデータが、学習用データベース141に記憶されていないと判定された場合、即ち、学習用データベース141に記憶された高音質オーディオデータのフレームすべてを注目フレームとして、ステップS 137の足し込みが行われた場合、足し込み部154は、いままでのステップS 137における足し込みによって、クラスごとに得られた式(8)の正規方程式を、タップ係数演算部155に供給して、ステップS 141に進む。

【0290】

ステップS 141では、タップ係数演算部155は、足し込み部154から供給される、各クラスごとの式(8)の正規方程式を解くことにより、クラスごとのタップ係数を求め、タップ係数記憶部156に供給して記憶させ、処理を終了する。

30

【0291】

以上のようにして、タップ係数記憶部156に記憶されたクラスごとのタップ係数が、図16の復号装置102におけるタップ係数記憶部124に記憶されている。

【0292】

以上のように、図19の学習装置では、高音質オーディオデータを、教師データとするとともに、その高音質オーディオデータをMPEG1オーディオレイヤ1方式で符号化して復号することにより得られるサブバンドデータ(さらには、サイドインフォメーション)を生徒データとして、その教師データと生徒データとの関係を学習することにより、式(1)によって、教師データと生徒データとを関係付けるタップ係数を、クラスごとに求めるようにしたので、図16の復号装置102において、そのようなタップ係数を用いてクラス分類適応処理を行うことにより、教師データに相当する高音質のオーディオデータを復号することが可能となる。

40

【0293】

次に、図21は、図5の復号装置102の第4の構成例を示している。なお、図中、図16における場合と対応する部分については、同一の符号を付してあり、以下では、その説明は、適宜省略する。即ち、図21の復号装置102は、基本的に、図16における場合と同様に構成されている。

【0294】

但し、図16の実施の形態では、デコーダ111が、符号化ビットストリームをMPEG1オ

50

ーディオレイヤ1方式で復号する過程で得られる低音質サブバンドデータを出力するのに対して、図21の実施の形態では、デコーダ111が、その低音質サブバンドデータに代えて、符号化ビットストリームをMPEG1オーディオレイヤ1方式で復号することにより得られる時間領域のオーディオデータ、即ち、低音質のオーディオデータ（以下、適宜、低音質オーディオデータという）を出力するようになっている。

【0295】

そして、図21の復号装置102では、クラス分類適応処理部113が、デコーダ111が出力する時間領域の低音質オーディオデータを、時間領域の高音質オーディオデータに変換する。

【0296】

そのため、図21の復号装置102では、クラス分類適応処理部113のタップ生成部121と122が、低音質サブバンドデータに代えて、デコーダ111が出力する低音質オーディオデータを用いて、予測タップとクラスタップを、それぞれ生成するようになっている。

【0297】

次に、図22のフローチャートを参照して、図21の復号装置102の処理（復号処理）について説明する。

【0298】

デコーダ111は、あるフレームの符号化ビットストリームを受信すると、そのフレームを、注目フレームとし、ステップS151において、その注目フレームについて、通常デコード処理を行う。

【0299】

即ち、本実施の形態では、符号化ビットストリームは、オーディオデータをMPEG1オーディオレイヤ1方式で符号化したものであり、ステップS1では、注目フレームの符号化ビットストリームが、MPEG1オーディオレイヤ1方式で復号される。

【0300】

具体的には、ステップS151の通常デコード処理では、まず最初に、ステップS161において、デコーダ111は、注目フレームの符号化ビットストリームを、量子化サブバンドデータと、その他のデータであるサイドインフォメーションに分解し、注目フレームのサイドインフォメーションを、コントローラ112とクラス分類適応処理部113に供給する。そして、デコーダ111は、ステップS162において、注目フレームの量子化サブバンドデータを逆量子化し、さらに、ステップS163に進み、その逆量子化の結果得られるサブバンドデータを用いてサブバンド合成を行う。これにより、デコーダ111は、MPEG1オーディオレイヤ1方式で符号化されて復号されたオーディオデータ、即ち、低音質オーディオデータを得て、クラス分類適応処理部113に供給する。

【0301】

以上のステップS161乃至S163の処理によって、ステップS151の通常デコード処理が終了し、その後は、ステップS152に進む。

【0302】

ステップS152に進み、コントローラ112は、注目フレームの高音質オーディオデータのデータ位置を表す変数mに、初期値としての、例えば0をセットし、ステップS153に進む。

【0303】

ステップS153では、コントローラ112は、注目フレームの、データ位置#mで特定される高音質オーディオデータs#mを、注目データとして選択する。さらに、ステップS153では、コントローラ112は、注目データs#mを特定するデータ位置#mと、注目フレームにおけるヌルサブバンドのサブバンドナンバを、クラス分類適応処理部113に供給し、ステップS154に進む。

【0304】

ステップS154では、クラス分類適応処理部113が、注目データs#mを求めるクラス

10

20

30

40

50

分類適応処理を行う。

【0305】

即ち、ステップS154のクラス分類適応処理では、まず最初に、ステップS171において、タップ生成部121と122が、コントローラ112からのデータ位置#mで特定される高音質オーディオデータs#mを、注目データs#mとして認識し、その注目データs#mについて、デコーダ111から供給される低音質オーディオデータ、さらには、必要に応じてサイドインフォメーションを用い、予測タップとクラスタップを生成する。

【0306】

ステップS171で得られた予測タップは、タップ生成部121から予測演算部125に供給され、また、クラスタップは、タップ生成部121からクラス分類部123に供給され、その後、ステップS171からステップS172に進む。

10

【0307】

ステップS172では、クラス分類部123が、タップ生成部122から供給されるクラスタップの他、コントローラ112から供給される注目フレームにおけるヌルサブバンドのサブバンドナンバ、および注目データs#mのデータ位置#mに基づいて、注目データをクラス分類し、その結果得られる注目データのクラスを表すクラスコードを、タップ係数記憶部124に供給して、ステップS173に進む。

【0308】

ステップS173では、タップ係数記憶部124が、そこに記憶されているクラスごとのタップ係数から、クラス分類部123から供給されるクラスコードが表す注目データs#mのクラスのタップ係数を読み出すことにより取得し、その注目データs#mのクラスのタップ係数を、予測演算部125に供給して、ステップS174に進む。

20

【0309】

ステップS174では、予測演算部125が、タップ生成部121から供給される注目データs#mについての予測タップと、タップ係数記憶部124から供給される注目データs#mのクラスのタップ係数とを用いて、式(1)の積和演算を行うことにより、注目データs#mを求める。

【0310】

以上のステップS171乃至S174の処理によって、ステップS154のクラス分類適応処理が終了し、ステップS155に進む。ステップS155では、コントローラ112が、変数mが、時間領域の1フレームのオーディオデータのデータ位置の最大値である383(図17)に等しいかどうかを判定し、383に等しくないと判定した場合、ステップS156に進む。ステップS156では、コントローラ112は、変数mを1だけインクリメントして、ステップS153に戻り、以下、同様の処理が繰り返される。

30

【0311】

また、ステップS155において、変数mが、データ位置の最大値である383に等しいと判定された場合、即ち、注目フレームのすべての高音質オーディオデータを注目データとして、そのサンプル値が、クラス分類適応処理によって求められた場合、クラス分類適応処理部113の予測演算部125は、その注目フレームの高音質オーディオデータを出力する。

40

【0312】

従って、この場合も、図18における場合と同様に、符号化装置101(図5)が出力する符号化ビットストリームのデータ量を増加させずに、高音質のオーディオデータを復号することができる。

【0313】

その後、ステップS157に進み、デコーダ111は、注目フレームの次のフレームの符号化ビットストリームが供給されたかどうかを判定する。ステップS157において、注目フレームの次のフレームの符号化ビットストリームが供給されたと判定された場合、その符号化ビットストリームが供給されたフレームを、新たに注目フレームとして、ステップS151に戻り、以下、同様の処理が繰り返される。

50

【0314】

また、ステップS157において、注目フレームの次のフレームの符号化ビットストリームが供給されていないと判定された場合、処理を終了する。

【0315】

次に、図23は、図21の復号装置102のタップ係数記憶部124に記憶させるクラスごとのタップ係数を学習する学習装置の一実施の形態の構成例を示している。なお、図中、図19における場合と対応する部分については、同一の符号を付してあり、以下では、その説明は、適宜省略する。即ち、図23の学習装置は、基本的に、図19の学習装置と同様に構成されている。

【0316】

但し、図19の学習装置では、生徒データ生成部132において、デコーダ143が、エンコーダ142が出力する符号化ビットストリームをMPEG1オーディオレイヤ1方式により復号し、その復号の結果得られる低音質オーディオデータを、低音質サブバンドデータに代えて、生徒データとして出力するようになっている。

【0317】

即ち、図21の復号装置102におけるクラス分類適応処理部113では、オーディオデータをMPEG1オーディオレイヤ1方式で符号化して復号することにより得られるオーディオデータ（低音質オーディオデータ）から予測タップを生成し、その予測タップを用いて、式(1)の積和演算を行うことにより、高音質オーディオデータが求められる。

【0318】

従って、式(1)の積和演算に用いるタップ係数としては、高音質オーディオデータを、教師データとするとともに、その高音質オーディオデータをMPEG1オーディオレイヤ1方式で符号化して復号することにより得られる低音質オーディオデータを生徒データとして、その生徒データと教師データとを、式(1)によって関係付けるものを求めれば良い。

【0319】

そこで、図23の学習装置では、上述したように、生徒データ生成部132が、低音質サブバンドデータに代えて、低音質オーディオデータを、生徒データとして出力するようになっている。

【0320】

次に、図24フローチャートを参照して、図23の学習装置において行われる、クラスごとのタップ係数を求める学習処理について説明する。

【0321】

まず最初に、教師データ生成部131は、学習用データベース141に記憶された高音質オーディオデータのフレームのうちの1フレームを注目フレームとし、ステップS181において、その注目フレームの教師データを生成する。即ち、教師データ生成部131は、学習用データベース141から、注目フレームの高音質オーディオデータを読み出し、教師データとして、生徒データ生成部132と、学習部134の足し込み部154に供給する。

【0322】

その後、ステップS182に進み、生徒データ生成部132は、教師データ生成部131から供給される注目フレームの高音質オーディオデータから生徒データを生成する。即ち、ステップS182では、生徒データ生成部132のデコーダ143が、注目フレームの高音質オーディオデータをMPEG1オーディオレイヤ1方式で符号化し、その符号化により得られる符号化ビットストリームを、デコーダ143に供給する。デコーダ143は、その符号化ビットストリームをMPEG1オーディオレイヤ1方式で復号し、その復号過程で得られる注目フレームのサイドインフォメーションと、その復号の結果得られる注目データの低音質オーディオデータを、生徒データとして、学習部134に供給する。さらに、デコーダ143は、サイドインフォメーションを、コントローラ133にも供給する。

【0323】

そして、ステップS183に進み、コントローラ133は、注目フレームの高音質オーデ

10

20

30

40

50

ィオデータのデータ位置を表す変数 i に、初期値としての、例えば 0 をセットし、ステップ S 1 8 4 に進む。ステップ S 1 8 4 では、コントローラ 1 3 3 は、注目フレームの、データ位置 $\#m$ で特定される高音質オーディオデータとしての教師データ $s\#m$ を、注目データとして選択する。さらに、ステップ S 1 8 4 では、コントローラ 1 3 3 は、注目データ $s\#m$ のデータ位置 $\#m$ と、注目フレームにおけるヌルサブバンドのサブバンドナンバ $\#n$ を、学習部 1 3 4 に供給し、ステップ S 1 8 5 に進む。

【 0 3 2 4 】

ステップ S 1 8 5 では、タップ生成部 1 5 1 と 1 5 2 が、コントローラ 1 3 3 からのデータ位置 $\#m$ で特定される高音質オーディオデータとしての教師データ $s\#m$ を、注目データ $s\#m$ として認識し、その注目データ $s\#m$ について、生徒データ生成部 1 3 2 から供給される生徒データとしての低音質オーディオデータとサイドインフォメーションを用い、図 2 1 のタップ生成部 1 2 1 と 1 2 2 における場合と同一のタップ構造の予測タップとクラスタップを、それぞれ生成する。

10

【 0 3 2 5 】

ステップ S 1 8 5 で得られた予測タップは、タップ生成部 1 5 1 から足し込み部 1 5 4 に供給され、また、クラスタップは、タップ生成部 1 5 1 からクラス分類部 1 5 3 に供給され、その後、ステップ S 1 8 5 からステップ S 1 8 6 に進む。

【 0 3 2 6 】

ステップ S 1 8 6 では、クラス分類部 1 5 3 が、タップ生成部 1 5 2 から供給されるクラスタップの他、コントローラ 1 3 3 から供給される注目フレームにおけるヌルサブバンドのサブバンドナンバ、および注目データ $s\#m$ のデータ位置 $\#m$ に基づき、注目データについて、図 2 1 のクラス分類部 1 2 3 と同一のクラス分類を行い、その結果得られる注目データのクラスを表すクラスコードを、足し込み部 1 5 4 に供給して、ステップ S 1 8 7 に進む。

20

【 0 3 2 7 】

ステップ S 1 8 7 では、足し込み部 1 5 4 が、教師データ生成部 1 3 1 から供給される注目フレームの教師データのうちの注目データとなっているものを選択する。さらに、ステップ S 1 8 7 では、足し込み部 1 5 4 が、クラス分類部 1 5 3 から供給されるクラスコードが表す注目データのクラスごとに、タップ生成部 1 5 1 から供給される予測タップ、および注目データを対象とした、上述した式 (8) における足し込みを行い、ステップ S 1 8 8 に進む。

30

【 0 3 2 8 】

ステップ S 1 8 8 では、コントローラ 1 3 3 が、変数 m が、1 フレームにおける高音質オーディオデータのデータ位置の最大値である 3 8 3 に等しいかどうかを判定し、3 8 3 に等しくないと判定した場合、ステップ S 1 8 9 に進む。ステップ S 1 8 9 では、コントローラ 1 3 3 は、変数 m を 1 だけインクリメントして、ステップ S 1 8 4 に戻り、以下、同様の処理が繰り返される。

【 0 3 2 9 】

また、ステップ S 1 8 8 において、変数 m が、データ位置の最大値である 3 8 3 に等しいと判定された場合、即ち、注目フレームの高音質オーディオデータとしての教師データすべてを注目データとして、ステップ S 1 8 7 の足し込みが行われた場合、ステップ S 1 9 0 に進み、教師データ生成部 1 3 1 は、注目フレームの次のフレームの高音質オーディオデータが、学習用データベース 1 4 1 に記憶されているかどうかを判定する。ステップ S 1 9 0 において、注目フレームの次のフレームの高音質オーディオデータが、学習用データベース 1 4 1 に記憶されていると判定された場合、その注目データの次のフレームを、新たに注目フレームとして、ステップ S 1 8 1 に戻り、以下、同様の処理が繰り返される。

40

【 0 3 3 0 】

また、ステップ S 1 9 0 において、注目フレームの次のフレームの高音質オーディオデータが、学習用データベース 1 4 1 に記憶されていないと判定された場合、即ち、学習用デ

50

ータベース 141 に記憶された高音質オーディオデータのフレームすべてを注目フレームとして、ステップ S187 の足し込みが行われた場合、足し込み部 154 は、いままでのステップ S187 における足し込みによって、クラスごとに得られた式 (8) の正規方程式を、タップ係数演算部 155 に供給して、ステップ S191 に進む。

【0331】

ステップ S191 では、タップ係数演算部 155 は、足し込み部 154 から供給される、各クラスごとの式 (8) の正規方程式を解くことにより、クラスごとのタップ係数を求め、タップ係数記憶部 156 に供給して記憶させ、処理を終了する。

【0332】

以上のようにして、タップ係数記憶部 156 に記憶されたクラスごとのタップ係数が、図 21 の復号装置 102 におけるタップ係数記憶部 124 に記憶されている。

10

【0333】

以上のように、図 23 の学習装置では、高音質オーディオデータを、教師データとするとともに、その高音質オーディオデータを MPEG1 オーディオレイヤ 1 方式で符号化して復号することにより得られる低音質オーディオデータ (さらには、サイドインフォメーション) を生徒データとして、その教師データと生徒データとの関係を学習することにより、式 (1) によって、教師データと生徒データとを関係付けるタップ係数を、クラスごとに求めるようにしたので、図 21 の復号装置 102 において、そのようなタップ係数を用いてクラス分類適応処理を行うことにより、教師データに相当する高音質のオーディオデータを復号することが可能となる。

20

【0334】

次に、図 25 は、図 5 の復号装置 102 の第 5 の構成例を示している。

【0335】

図 25 の実施の形態においては、復号装置 102 は、復号ブロック 161 とレベル補正部 162 で構成されている。

【0336】

復号ブロック 161 には、図 5 の符号化装置 101 が出力する符号化ビットストリームが供給されるようになっている。復号ブロック 161 は、例えば、図 6、図 13、図 16、または図 21 に示した復号装置 102 と同様に構成されており、そこに供給される符号化ビットストリームのフレームを、順次、注目フレームとし、その注目フレームの符号化ビットストリームを、時間領域のデータであるオーディオデータに復号して、レベル補正部 162 に供給する。さらに、復号ブロック 161 は、注目フレームの符号化ビットストリームの復号の過程で認識することができる注目フレームにおけるヌルサブバンドのサブバンドナンバも、レベル補正部 162 に供給する。

30

【0337】

ここで、注目フレームにおいては、必ずしも、ヌルサブバンドが存在するわけではなく、存在しない場合もある。このため、注目フレームにヌルバンドが存在しない場合、即ち、注目フレームのすべてのサブバンドに、1 ビット以上が割り当てられている場合、復号ブロック 161 は、サブバンドナンバとして取り得ない値を、ヌルバンドが存在しない旨の情報として、レベル補正部 162 に供給する。なお、このことは、上述した図 16 等の復号装置 102 においても同様である。

40

【0338】

レベル補正部 162 は、サブバンドナンバによって特定されるサブバンドの周波数成分に基づいて、復号ブロック 161 から供給される注目ブロックのオーディオデータを補正して出力する。

【0339】

即ち、図 26 および図 27 は、オーディオデータのパワースペクトルと静寂時の最小可聴限界を示している。

【0340】

ここで、図 26 および図 27 において、細い実線がオーディオデータのパワースペクトル

50

を示し、太い実線が最小可聴限界を示している。

【 0 3 4 1 】

MPEG1オーディオレイヤ 1 方式では、オーディオデータのパワースペクトルが最小可聴限界以上の各サブバンドについて、そのオーディオデータを人間が感知することができるレベルとしてのマスキングスレッショルドが求められる。ここで、図 2 6 および図 2 7 では、点線がマスキングスレッショルドを示している。

【 0 3 4 2 】

さらに、MPEG1オーディオレイヤ 1 方式では、各サブバンドに、サブバンドデータを量子化することにより生じる量子化雑音が、マスキングスレッショルドより小さくなるように、ビット割り当てが行われる。

10

【 0 3 4 3 】

従って、MPEG1オーディオレイヤ 1 方式では、最小可聴限界以下のレベルのサブバンドデータしかないサブバンドについては、ビット割り当てが 0 とされる。

【 0 3 4 4 】

このことは、MPEG1オーディオレイヤ 1 方式において、ビット割り当てが 0 とされたサブバンド、即ち、ヌルサブバンドのパワーは、最小可聴限界以下であることを意味する。

【 0 3 4 5 】

従って、例えば、図 2 5 の復号ブロック 1 6 1 が出力する注目フレームのオーディオデータにおけるサブバンドsb#nがヌルサブバンドであるとした場合、そのヌルサブバンドsb#nに対応する周波数帯域成分は、図 2 6 に示すように、最小可聴限界以下でなければ、復号ブロック 1 6 1 が出力する注目ブロックのオーディオデータは、図 5 の符号化装置 1 0 1 で符号化されたオーディオデータを精度良く再現したものとは言えない。

20

【 0 3 4 6 】

即ち、ヌルサブバンドsb#nに対応する周波数帯域成分が、図 2 7 に示すように、最小可聴限界以上である（より大きい）場合には、復号ブロック 1 6 1 が出力する注目ブロックのオーディオデータは、図 5 の符号化装置 1 0 1 で符号化されたオーディオデータを精度良く再現したものとは言えない。

【 0 3 4 7 】

そこで、図 2 5 のレベル補正部 1 6 2 は、復号ブロック 1 6 1 が出力するオーディオデータにおけるヌルサブバンドsb#nに対応する周波数帯域成分が、最小可聴限界以上である場合に、そのオーディオデータのレベルを補正し、これにより、図 5 の符号化装置 1 0 1 で符号化されたオーディオデータを精度良く再現したオーディオデータを求めて出力するようになっている。

30

【 0 3 4 8 】

即ち、レベル補正部 1 6 2 は、ラッチ部 1 7 1、FFT部 1 7 2、レベル調整部 1 7 3、逆FFT部 1 7 4、選択部 1 7 5、および判定部 1 7 6 から構成されており、ラッチ部 1 7 1 には、復号ブロック 1 6 1 が出力する注目フレームのオーディオデータが供給され、判定部 1 7 6 には、復号ブロック 1 6 1 が出力する注目ブロックにおけるヌルサブバンドを表すサブバンドナンバが供給されるようになっている。

【 0 3 4 9 】

ラッチ部 1 7 1 は、復号ブロック 1 6 1 が出力する注目フレームのオーディオデータをラッチし、FFT部 1 7 2 と選択部 1 7 5 に供給する。

40

【 0 3 5 0 】

FFT部 1 7 2 は、ラッチ部 1 7 1 から供給される注目フレームのオーディオデータをFFT処理し、これにより、パワースペクトルに変換して、判定部 1 7 6 とレベル調整部 1 7 3 に供給する。

【 0 3 5 1 】

レベル調整部 1 7 3 は、判定部 1 7 6 の制御にしたがい、パワースペクトルのレベルを調整し、そのレベル調整が行われた後のパワースペクトル（以下、適宜、正規化パワースペクトルという）を、逆FFT部 1 7 4 に供給する。

50

【 0 3 5 2 】

逆 F F T 部 1 7 4 は、レベル調整部 1 7 3 から供給される正規化パワースペクトルを逆 F F T 処理し、これにより、時間領域のオーディオデータに変換する。ここで、正規化パワースペクトルを逆 F F T することにより得られる時間領域のオーディオデータを、以下、適宜、正規化オーディオデータという。

【 0 3 5 3 】

選択部 1 7 5 には、上述したラッチ部 1 7 1 から、復号ブロック 1 6 1 が出力する注目フレームのオーディオデータが供給される他、逆 F F T 部 1 7 4 から、注目フレームの正規化オーディオデータが供給されるようになっている。選択部 1 7 5 は、判定部 1 7 6 の制御にしたがい、ラッチ部 1 7 1 からの注目フレームのオーディオデータか、または逆 F F T 部 1 7 4 からの注目フレームの正規化オーディオデータを選択し、その選択したオーディオデータを、注目フレームの最終的な復号結果として出力する。

10

【 0 3 5 4 】

判定部 1 7 6 は、復号ブロック 1 6 1 から供給されるサブバンドナンバが表すサブバンド、即ち、ヌルサブバンドの最小可聴限界に応じた所定の閾値と、F F T 部 1 7 2 から供給される注目フレームのパワースペクトルのうちの、復号ブロック 1 6 1 から供給されるサブバンドナンバが表すヌルサブバンドのレベルとを比較し、その比較結果に基づいて、レベル調整部 1 7 3 を制御する。また、判定部 1 7 6 は、選択部 1 7 5 も制御する。

【 0 3 5 5 】

次に、図 2 8 のフローチャートを参照して、図 2 5 のレベル補正部 1 6 2 の処理（レベル補正処理）について説明する。

20

【 0 3 5 6 】

復号ブロック 1 6 1 が注目フレームのオーディオデータと、その注目フレームにおけるヌルサブバンドのサブバンドナンバを出力すると、ステップ S 2 0 1 において、ラッチ部 1 7 1 は、復号ブロック 1 6 1 が出力する注目フレームのオーディオデータをラッチし、ステップ S 2 0 2 に進む。

【 0 3 5 7 】

ステップ S 2 0 2 では、判定部 1 7 6 が、復号ブロック 1 6 1 から供給されるサブバンドナンバに基づいて、注目フレームにヌルサブバンドが存在するかどうかを判定する。

【 0 3 5 8 】

ステップ S 2 0 2 において、注目フレームにヌルサブバンドが存在しないと判定された場合、即ち、復号ブロック 1 6 1 が出力したサブバンドナンバが、1 フレームのサブバンドナンバとして取り得る値ではない場合、ステップ S 2 0 3 に進み、判定部 1 7 6 は、ラッチ部 1 7 1 の出力を選択するように、選択部 1 7 5 を制御する。これにより、さらに、ステップ S 2 0 3 では、選択部 1 7 5 が、ラッチ部 1 7 1 でラッチされている注目フレームのオーディオデータを選択し、注目フレームの最終的な復号結果として出力して、ステップ S 2 1 0 に進む。

30

【 0 3 5 9 】

従って、注目フレームにヌルバンドが存在しない場合は、復号ブロック 1 6 1 が出力する注目フレームのオーディオデータが、そのまま、注目フレームの最終的な復号結果として出力される。

40

【 0 3 6 0 】

一方、ステップ S 2 0 2 において、注目フレームにヌルサブバンドが存在すると判定された場合、即ち、復号ブロック 1 6 1 が出力したサブバンドナンバが、1 フレームのサブバンドナンバとして取り得る値である場合、ステップ S 2 0 4 に進み、F F T 部 1 7 2 は、ラッチ部 1 7 1 でラッチされている注目フレームのオーディオデータを F F T 処理し、これにより、注目フレームの時間領域のオーディオデータを、パワースペクトルに変換して、判定部 1 7 6 とレベル調整部 1 7 3 に供給し、ステップ S 2 0 5 に進む。

【 0 3 6 1 】

ステップ S 2 0 5 では、判定部 1 7 6 が、復号ブロック 1 7 6 から供給されるサブバンド

50

ナンバが表すヌルサブバンドの最小可聴限界と、FFT部172から供給される注目フレームのパワースペクトルのうちの、復号ブロック176から供給されるサブバンドナンバが表すヌルサブバンドの最大レベル（以下、適宜、ヌルサブバンドレベルという）とを比較し、ヌルサブバンドレベルが、そのヌルサブバンドの最小可聴限界に、所定のオフセットを加えた所定の閾値よりも大きいヌルサブバンドが存在するかどうかを判定する。

【0362】

即ち、1フレームには、1つのヌルサブバンドだけでなく、複数のヌルサブバンドが存在する場合がある。ステップS205では、注目フレームにおける1以上のヌルサブバンドの中に、ヌルサブバンドレベルが、そのヌルサブバンドの最小可聴限界に所定のオフセットを加えた所定の閾値よりも大きいヌルサブバンドがあるかどうか判定される。

10

【0363】

ステップS205において、注目フレームに、ヌルサブバンドレベルが所定の閾値よりも大きいヌルサブバンドがないと判定された場合、即ち、注目フレームにヌルサブバンドは存在するが、そのヌルサブバンドの最小可聴限界に所定のオフセットを加えた所定の閾値よりも大きいヌルサブバンドレベルのサブバンドは存在しない場合、注目ブロックのオーディオデータの補正は必要ないとして、ステップS203に進み、上述したように、選択部203は、ラッチ部171でラッチされている注目フレームのオーディオデータを選択し、注目フレームの最終的な復号結果として出力して、ステップS210に進む。

【0364】

また、ステップS205において、注目フレームに、ヌルサブバンドレベルが所定の閾値よりも大きいヌルサブバンドがあると判定された場合、即ち、注目フレームに、そのヌルサブバンドの最小可聴限界を大きく超えるヌルサブバンドレベルのサブバンドが存在する場合、ステップS206に進み、判定部176は、ヌルサブバンドレベルと最小可聴限界との差分（[ヌルサブバンドレベル] - [最小可聴限界]）が最大のヌルサブバンド（以下、適宜、最大ヌルサブバンドという）を検出する。

20

【0365】

即ち、注目フレームに、ヌルサブバンドレベルが所定の閾値よりも大きいヌルサブバンドが1つだけある場合、ステップS205では、そのヌルサブバンドが、最大ヌルサブバンドとして検出される。また、注目フレームに、ヌルサブバンドレベルが所定の閾値よりも大きいヌルサブバンドが複数ある場合、ステップS205では、その複数のヌルサブバンドの中から、ヌルサブバンドレベルと最小可聴限界との差分が最大のヌルサブバンドが、最大ヌルサブバンドとして検出される。

30

【0366】

その後、ステップS207に進み、判定部207は、レベル調整部173を制御することにより、FFT部172が出力する注目フレームのパワースペクトルのレベルを、最大ヌルサブバンドのヌルサブバンドレベルに基づいて調整させる。即ち、これにより、レベル調整部173は、例えば、FFT部172が出力する注目フレームのパワースペクトルの各サブバンドのパワーそれぞれを、最大ヌルサブバンドのヌルサブバンドレベルで除算することにより、あるいは、FFT部172が出力する注目フレームのパワースペクトルの各サブバンドのパワーそれぞれから、最大ヌルサブバンドのヌルサブバンドレベルを減算することにより、注目フレームのパワースペクトルのレベルを調整した正規化パワースペクトルを求め、逆FFT部174に供給する。

40

【0367】

そして、ステップS208に進み、逆FFT部174は、レベル調整部173から供給される正規化パワースペクトルを逆FFT処理し、これにより、時間領域のオーディオデータである正規化オーディオデータを求めて、選択部175に出力し、ステップS209に進む。

【0368】

ステップS209では、判定部176は、逆FFT部174の出力を選択するように、選択部175を制御する。これにより、さらに、ステップS209では、選択部175が、

50

逆FFT部174が出力する注目フレームの正規化オーディオデータを選択し、注目フレームの最終的な復号結果として出力して、ステップS210に進む。

【0369】

従って、注目フレームに、そのヌルサブバンドの最小可聴限界を大きく超えるヌルサブバンドレベルのサブバンドが存在する場合には、復号ブロック161が出力する注目フレームのオーディオデータが補正され、これにより、ヌルサブバンドレベルが最小可聴限界を超えない正規化オーディオデータとされる。そして、この正規化オーディオデータが、注目フレームの最終的な復号結果として出力される。

【0370】

以上のように、ヌルサブバンドの最小可聴限界を大きく超えるヌルサブバンドレベルのサブバンドが存在する場合に、復号ブロック161が出力するオーディオデータを補正するようにしたので、符号化装置101(図5)で符号化されたオーディオデータにより近い復号結果を得ることが可能となる。

10

【0371】

ところで、図5の符号化装置101では、オーディオデータが、MPEG1オーディオレイヤ1方式で符号化されるため、時間領域のオーディオデータは、サブバンド分割によって、同一周波数帯域幅のサブバンドに分割される。従って、各サブバンドの周波数帯域幅は、同一である。

【0372】

サブバンド分割では、このような周波数帯域幅が同一のサブバンドへの分割の他、例えば、いわゆるログスケール(log scale)等で周波数帯域幅が増加するサブバンドへの分割を行うことも可能である。

20

【0373】

そこで、例えば、2の累乗で周波数帯域幅が増加するサブバンド分割を考えると、即ち、図29に示すように、サブバンドsb0の周波数帯域幅が $B \times 2^0$ 、サブバンドsb1の周波数帯域幅が $B \times 2^1$ 、・・・、サブバンドsb#nの周波数帯域幅が $B \times 2^n$ 、・・・となるサブバンド分割を考えると、各サブバンドのサブバンドデータのサンプル数は、そのサブバンドの周波数帯域幅に比例した数になる。

【0374】

即ち、同一周波数帯域幅のサブバンドへのサブバンド分割では、各サブバンドの周波数帯域幅は同一幅なので、サブバンドデータのサンプル数も、各サブバンドで同一になるが、2の累乗で周波数帯域幅が増加するサブバンド分割では、各サブバンドの周波数帯域幅が異なるので、サブバンドデータのサンプル数も、各サブバンドで異なる数になる。

30

【0375】

この場合、例えば、図7で説明したように、注目サブバンドsb#nと、その注目サブバンドsb#nの1/2倍の周波数帯域のローサブバンドsb#nL、および2倍の周波数帯域のハイサブバンドsb#nHを、タップ用サブバンドとして選択し、そのタップ用サブバンドのすべてのサブバンドデータを、タップとすることとすると、注目サブバンドがいずれのサブバンドであるかによって、タップ数が増減してしまい、クラス分類適応処理を行うにあたり、あまり望ましくない。

40

【0376】

そこで、符号化装置101(図5)において、2の累乗で周波数帯域幅が増加するサブバンド分割が行われる場合には、上述のようなタップ数の変化を防止するために、タップ用サブバンドとされたサブバンドsb#nについては、例えば、図29に示すように、そのサブバンドデータのうちの 2^n サンプルごとのサブバンドデータをタップとするようにすることができる。この場合、例えば、サブバンドsb1がタップ用サブバンドとして選択されたときには、そのサブバンドデータのうちの2サンプルごとのサブバンドが、タップとされることになる。

【0377】

以上のように、タップ用サブバンドとされたサブバンドsb#nについて、そのサブバンドデ

50

ータのうちの 2^n サンプルごとのサブバンドデータをタップとすることにより、タップ用サブバンドとされたサブバンドそれぞれからは、同一の数のサブバンドデータがタップとされることとなり、その結果、注目サブバンドがいずれのサブバンドであるかによるタップ数の変化を防止することが可能となる。

【0378】

ここで、上述の場合には、タップ用サブバンドとされたサブバンドsb#nについて、そのサブバンドデータのうちの 2^n サンプルごとのサブバンドデータをタップとすることにしたが、その他、例えば、 2^n サンプルごとのサブバンドデータの平均値などをタップとすることによっても可能である。

【0379】

なお、周波数帯域幅が同一のサブバンドへのサブバンド分割では、上述したように、注目サブバンドsb#nのサブバンドナンバ#nから、注目サブバンドsb#nの $1/2$ 倍の周波数帯域のローサブバンドsb#nLと、注目サブバンドsb#nの2倍の周波数帯域のハイサブバンドsb#nHを特定することができるが、2の累乗で周波数帯域幅が増加するサブバンド分割が行われる場合には、注目サブバンドsb#nの $1/2$ 倍の周波数帯域のローサブバンドsb#nLと、注目サブバンドsb#nの2倍の周波数帯域のハイサブバンドsb#nHは、例えば、次のようにして特定される。

【0380】

即ち、注目サブバンドsb#nの中心周波数（周波数帯域幅の最小周波数と最大周波数の平均値）が求められ、例えば、その中心周波数の $1/2$ 倍または2倍の周波数に最も近い中心周波数のサブバンドが、それぞれ、ローサブバンドsb#nLまたはハイサブバンドsb#nHとされる。

【0381】

次に、上述した一連の処理は、ハードウェアにより行うこともできるし、ソフトウェアにより行うこともできる。一連の処理をソフトウェアによって行う場合には、そのソフトウェアを構成するプログラムが、汎用のコンピュータ等にインストールされる。

【0382】

そこで、図30は、上述した一連の処理を実行するプログラムがインストールされるコンピュータの一実施の形態の構成例を示している。

【0383】

プログラムは、コンピュータに内蔵されている記録媒体としてのハードディスク205やROM203に予め記録しておくことができる。

【0384】

あるいはまた、プログラムは、フレキシブルディスク、CD-ROM(Compact Disc Read Only Memory)、MO(Magneto Optical)ディスク、DVD(Digital Versatile Disc)、磁気ディスク、半導体メモリなどのリムーバブル記録媒体211に、一時的あるいは永続的に格納（記録）しておくことができる。このようなリムーバブル記録媒体211は、いわゆるパッケージソフトウェアとして提供することができる。

【0385】

なお、プログラムは、上述したようなリムーバブル記録媒体211からコンピュータにインストールする他、ダウンロードサイトから、デジタル衛星放送用の人工衛星を介して、コンピュータに無線で転送したり、LAN(Local Area Network)、インターネットといったネットワークを介して、コンピュータに有線で転送し、コンピュータでは、そのようにして転送されてくるプログラムを、通信部208で受信し、内蔵するハードディスク205にインストールすることができる。

【0386】

コンピュータは、CPU(Central Processing Unit)202を内蔵している。CPU202には、バス201を介して、入出力インタフェース210が接続されており、CPU202は、入出力インタフェース210を介して、ユーザによって、キーボードや、マウス、マイク等で構成される入力部207が操作等されることにより指令が入力されると、それにした

10

20

30

40

50

がって、ROM(Read Only Memory) 2 0 3 に格納されているプログラムを実行する。あるいは、また、CPU 2 0 2 は、ハードディスク 2 0 5 に格納されているプログラム、衛星若しくはネットワークから転送され、通信部 2 0 8 で受信されてハードディスク 2 0 5 にインストールされたプログラム、またはドライブ 2 0 9 に装着されたリムーバブル記録媒体 2 1 1 から読み出されてハードディスク 2 0 5 にインストールされたプログラムを、RAM(Random Access Memory) 2 0 4 にロードして実行する。これにより、CPU 2 0 2 は、上述したフローチャートにしたがった処理、あるいは上述したブロック図の構成により行われる処理を行う。そして、CPU 2 0 2 は、その処理結果を、必要に応じて、例えば、入出力インタフェース 2 1 0 を介して、LCD(Liquid Crystal Display)やスピーカ等で構成される出力部 2 0 6 から出力、あるいは、通信部 2 0 8 から送信、さらには、ハードディスク 2 0 5 に記録等させる。

10

【 0 3 8 7 】

ここで、本明細書において、コンピュータに各種の処理を行わせるためのプログラムを記述する処理ステップは、必ずしもフローチャートとして記載された順序に沿って時系列に処理する必要はなく、並列的あるいは個別に実行される処理（例えば、並列処理あるいはオブジェクトによる処理）も含むものである。

【 0 3 8 8 】

また、プログラムは、1のコンピュータにより処理されるものであっても良いし、複数のコンピュータによって分散処理されるものであっても良い。さらに、プログラムは、遠方のコンピュータに転送されて実行されるものであっても良い。

20

【 0 3 8 9 】

なお、本実施の形態では、注目サブバンドsb#nと、その注目サブバンドsb#nの1/2倍の周波数帯域のローサブバンドsb#nLと2倍の周波数帯域のハイサブバンドsb#nHを、タップ用サブバンドとして選択するようにしたが、ローサブバンドsb#nLとハイサブバンドsb#nHとしては、その他、例えば、1/3倍と3倍や、1/9倍と9倍、1/5倍と5倍の周波数帯域などの1/PとPのN乗倍のサブバンドを採用することが可能である（P、Nは、正の整数）。

【 0 3 9 0 】

また、本実施の形態では、符号化装置 1 0 1（図 5）において、オーディオデータをMPEG1オーディオレイヤ1方式で符号化するようにしたが、オーディオデータの符号化方式としては、MPEG1オーディオレイヤ1方式以外に、例えば、前述したMPEG1オーディオレイヤ2、3や、MPEG2オーディオAAC、その他、時間領域のオーディオデータを周波数領域のオーディオデータに変換する周波数変換を利用するものを採用することが可能である。

30

【 0 3 9 1 】

ここで、符号化装置 1 0 1 において、周波数変換によって、時間領域のオーディオデータを、例えば、MDCT係数に変換する場合には、復号装置 1 0 2 のクラス分類適応処理部 1 1 3 では、例えば、MDCTの基底のMDCT係数すべてによって、予測タップを構成するとともに、MDCT係数の変化パターンやレベルなどに基づいてクラス分類を行うようにすることができる。また、符号化装置 1 0 1 において、周波数変換としてFFTを採用し、時間領域のオーディオデータを、周波数領域のオーディオデータとしての周波数スペクトルに変換する場合には、復号装置 1 0 2 のクラス分類適応処理部 1 1 3 では、例えば、FFTによって得られる周波数スペクトルを表す複素数を構成する実数部や虚数部から予測タップを構成するとともに、周波数スペクトルの変化パターンや各周波数帯域ごとのパワー（周波数スペクトルを表す複素数の平方和の平方根）などに基づいてクラス分類を行うようにすることができる。

40

【 0 3 9 2 】

【発明の効果】

以上の如く、本発明によれば、符号化ビットストリームのデータ量を増加させずに、高音質のオーディオデータを復号することが可能となる。

【図面の簡単な説明】

50

【図 1】 MPEG1オーディオレイヤ 1 方式のエンコーダの構成例を示すブロック図である。

【図 2】 MPEG1オーディオレイヤ 1 方式のデコーダの構成例を示すブロック図である。

【図 3】 MPEG2オーディオAAC方式のエンコーダの構成例を示すブロック図である。

【図 4】 MPEG2オーディオAAC方式のデコーダの構成例を示すブロック図である。

【図 5】 本発明を適用したデータ処理装置の一実施の形態の構成例を示すブロック図である。

【図 6】 復号装置 102 の第 1 の構成例を示すブロック図である。

【図 7】 タップ構造の例を説明する図である。

【図 8】 タップ構造の例を説明する図である。

10

【図 9】 タップ構造の例を説明する図である。

【図 10】 復号装置 102 の処理を説明するフローチャートである。

【図 11】 タップ係数を学習する学習装置の第 1 の構成例を示すブロック図である。

【図 12】 学習装置の処理を説明するフローチャートである。

【図 13】 復号装置 102 の第 2 の構成例を示すブロック図である。

【図 14】 復号装置 102 の処理を説明するフローチャートである。

【図 15】 学習装置の処理を説明するフローチャートである。

【図 16】 復号装置 102 の第 3 の構成例を示すブロック図である。

【図 17】 タップ構造の例を説明する図である。

【図 18】 復号装置 102 の処理を説明するフローチャートである。

20

【図 19】 学習装置の第 2 の構成例を示すブロック図である。

【図 20】 学習装置の処理を説明するフローチャートである。

【図 21】 復号装置 102 の第 4 の構成例を示すブロック図である。

【図 22】 復号装置 102 の処理を説明するフローチャートである。

【図 23】 学習装置の第 4 の構成例を示すブロック図である。

【図 24】 学習装置の処理を説明するフローチャートである。

【図 25】 復号装置 102 の第 5 の構成例を示すブロック図である。

【図 26】 オーディオデータのパワースペクトルと静寂時の最小可聴限界を示す図である。

。

【図 27】 オーディオデータのパワースペクトルと静寂時の最小可聴限界を示す図である

30

。

【図 28】 レベル補正部 162 の処理を説明するフローチャートである。

【図 29】 ログスケールでのサブバンド分割を示す図である。

【図 30】 本発明を適用したコンピュータの一実施の形態の構成例を示すブロック図である。

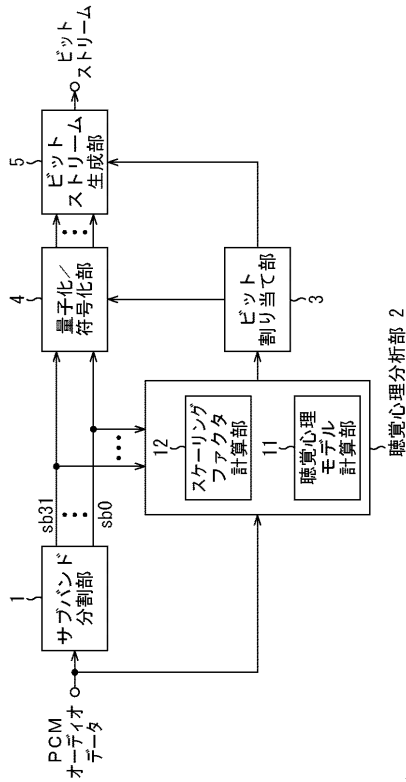
【符号の説明】

101 符号化装置, 102 復号装置, 103 伝送媒体, 104 記録媒体,
111 デコーダ, 112 コントローラ, 113 クラス分類適応処理部, 1
14 サブバンド合成部, 121, 122 タップ生成部, 123 クラス分類部,
124 タップ係数記憶部, 125 予測演算部, 131 教師データ生成部,
132 生徒データ生成部, 133 コントローラ, 134 学習部, 141 学
習用データベース, 142 エンコーダ, 143 デコーダ, 151, 152 タ
ップ生成部, 153 クラス分類部, 154 足し込み部, 155 タップ係数演
算部, 156 タップ係数記憶部, 161 復号ブロック, 162 レベル補正部,
171 ラッチ部, 172 FFT部, 173 レベル調整部, 174 逆FFT
部, 175 選択部, 176 判定部, 201 バス, 202 CPU, 20
3 ROM, 204 RAM, 205 ハードディスク, 206 出力部, 207 入
力部, 208 通信部, 209 ドライブ, 210 入出力インタフェース, 21
1 リムーバブル記録媒体

40

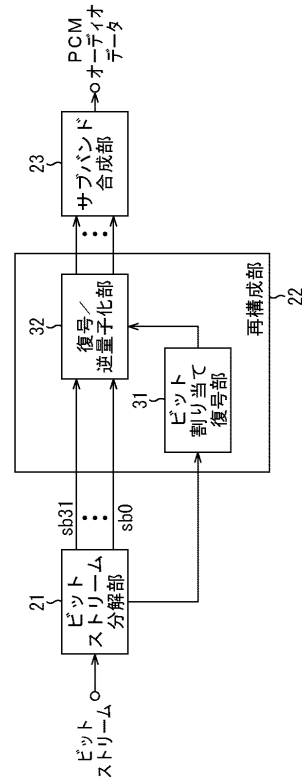
【図 1】

図1



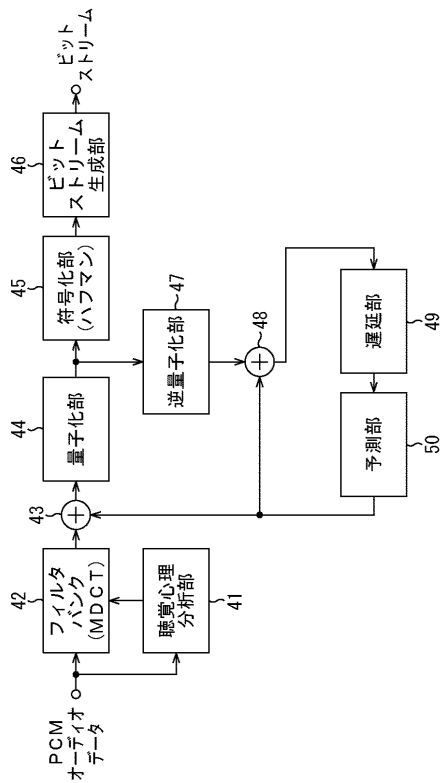
【図 2】

図2



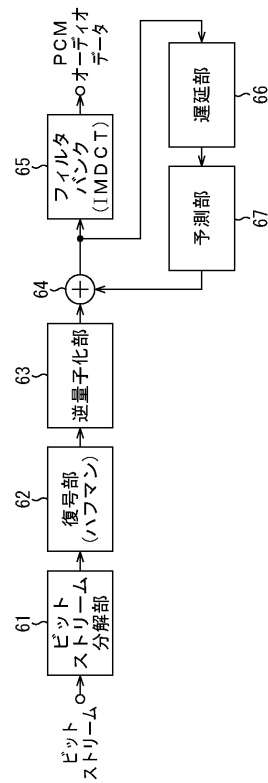
【図 3】

図3



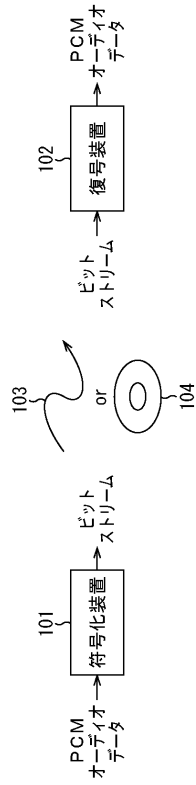
【図 4】

図4



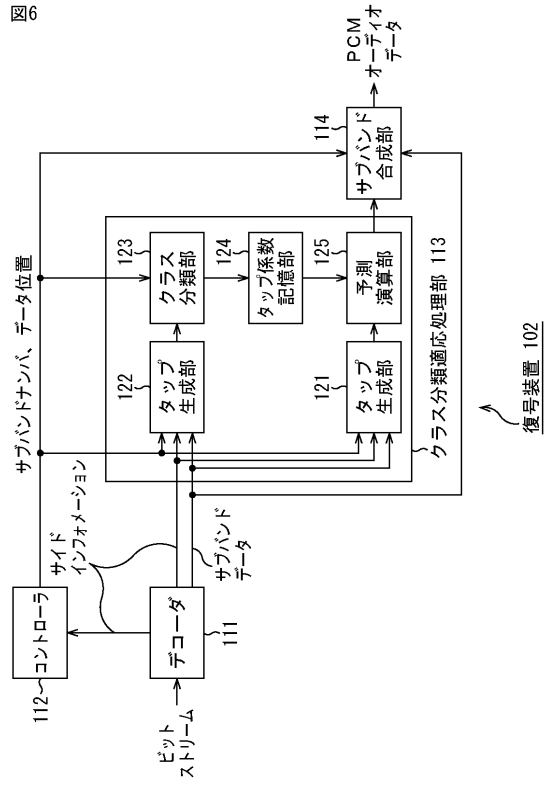
【図 5】

図5



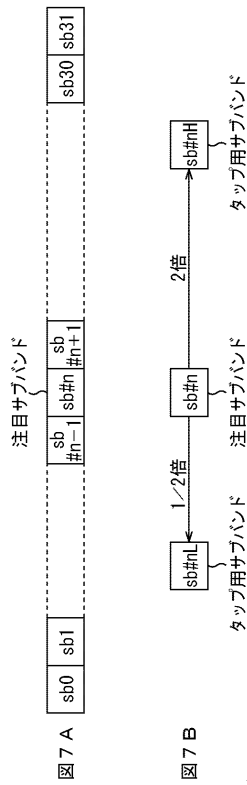
【図 6】

図6



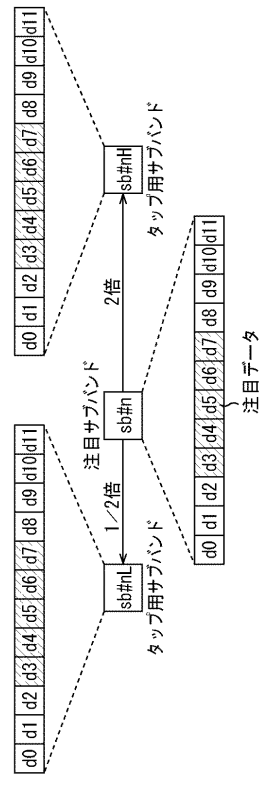
【図 7】

図7

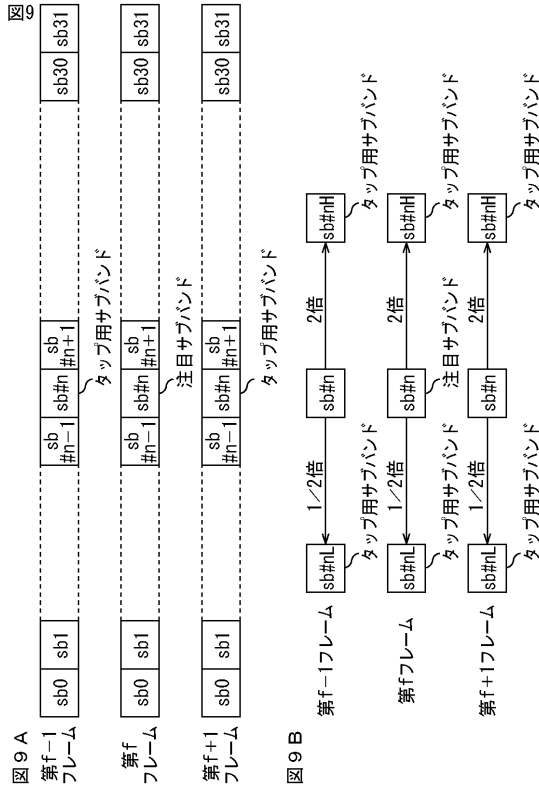


【図 8】

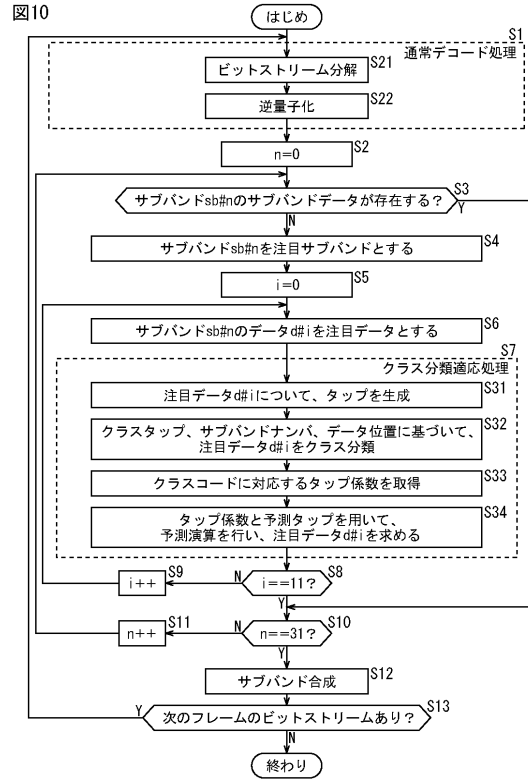
図8



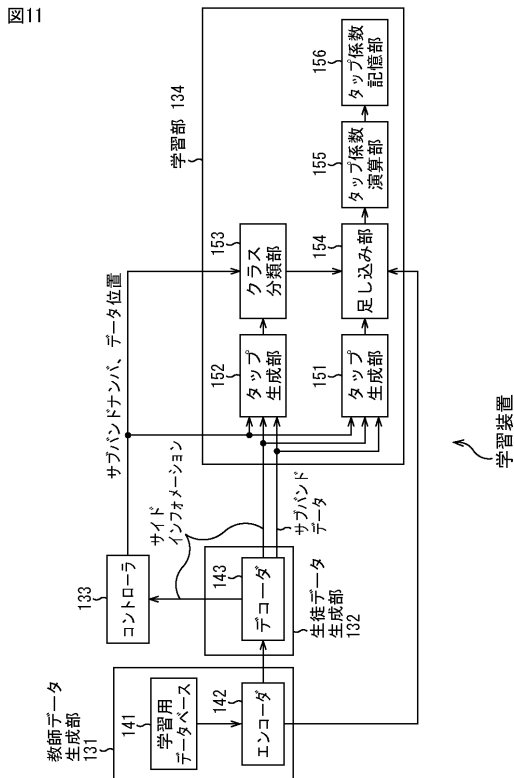
【図 9】



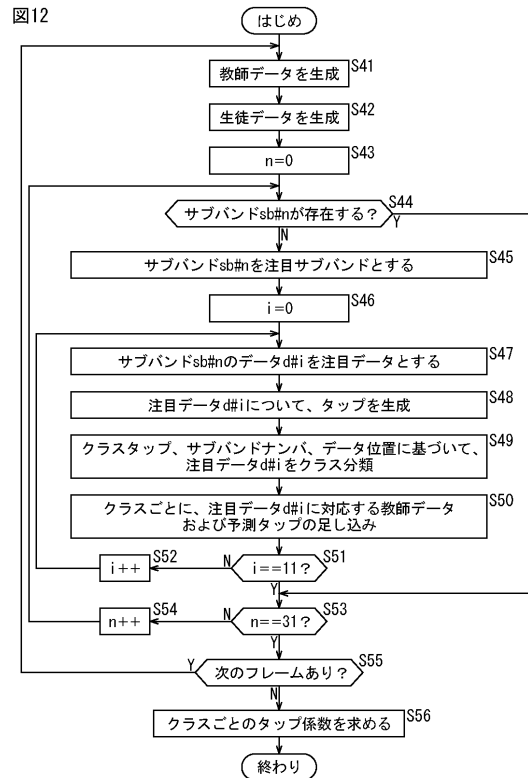
【図 10】



【図 11】

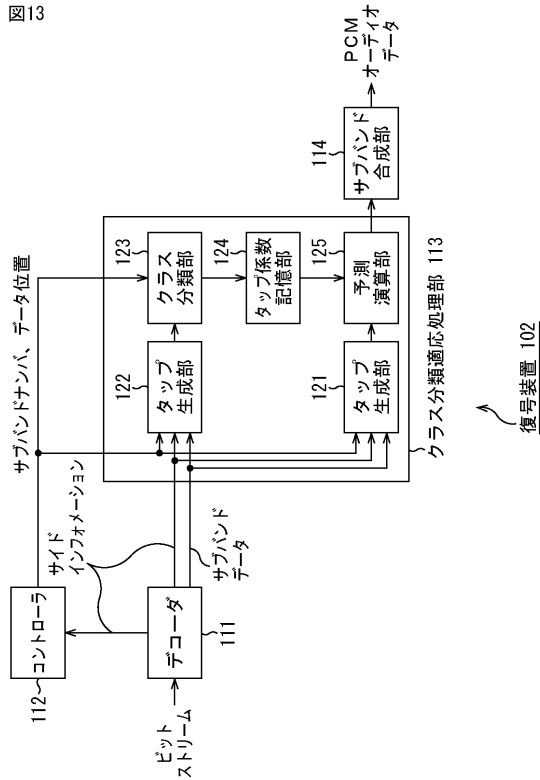


【図 12】



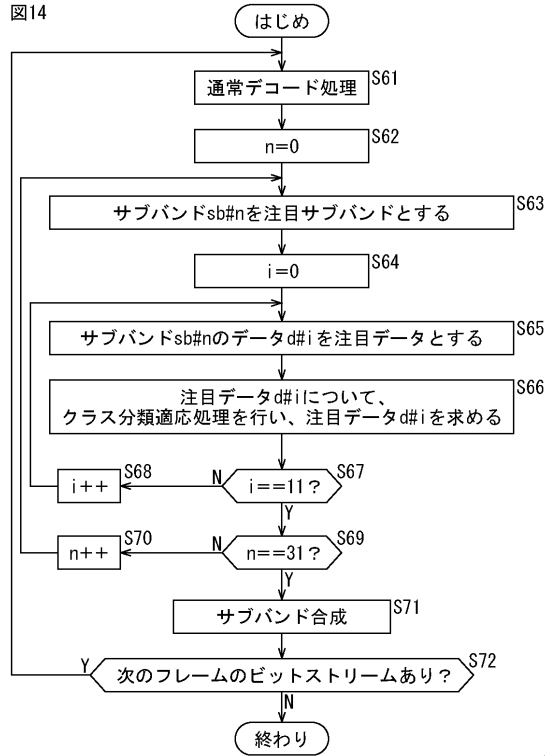
【図 13】

図13



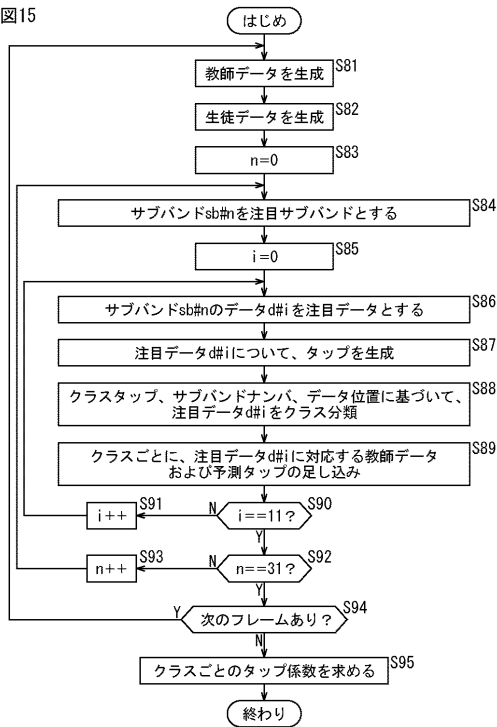
【図 14】

図14



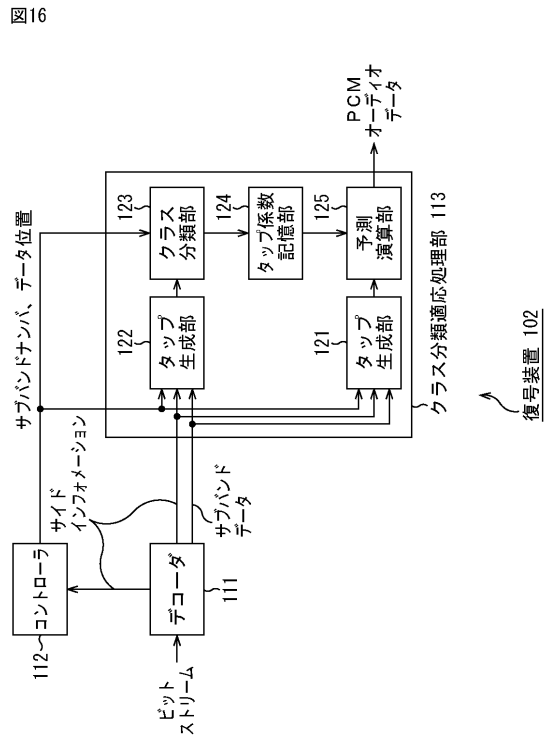
【図 15】

図15



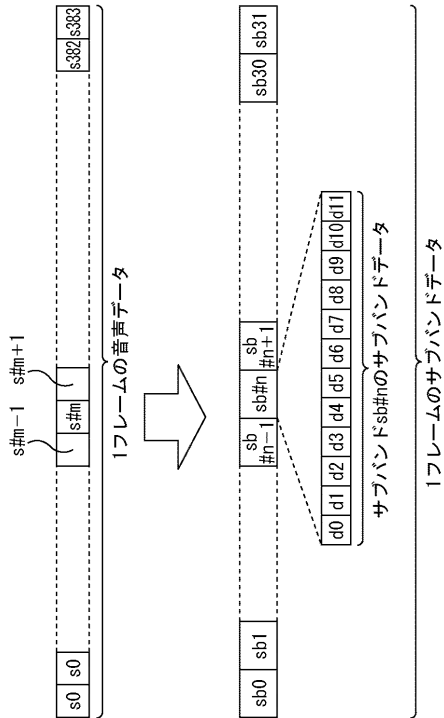
【図 16】

図16



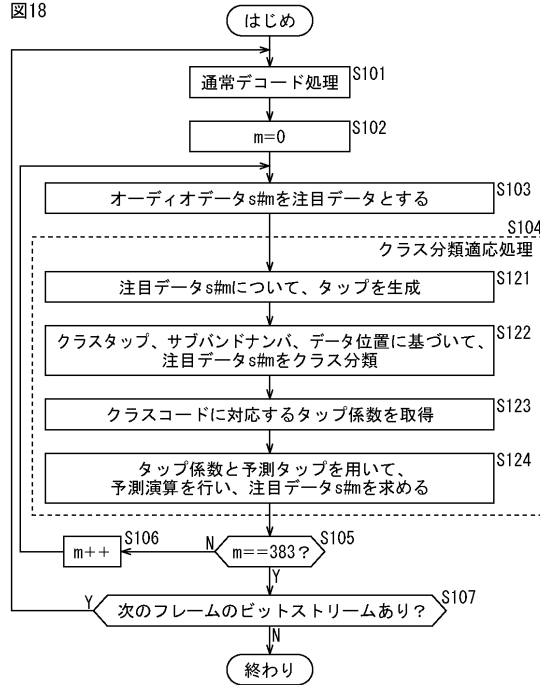
【図 17】

図17



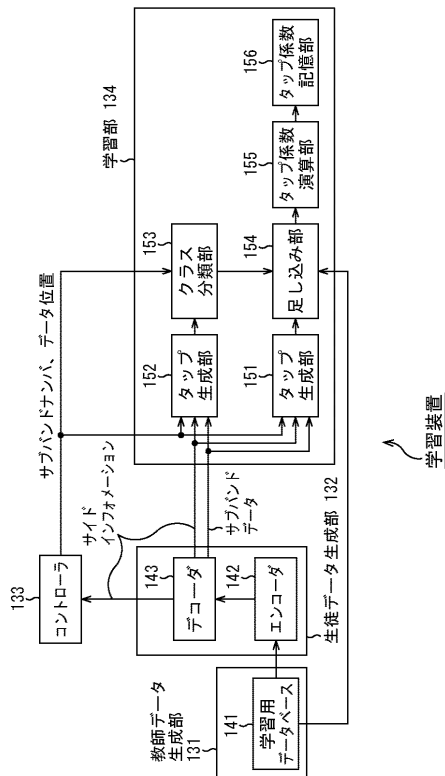
【図 18】

図18



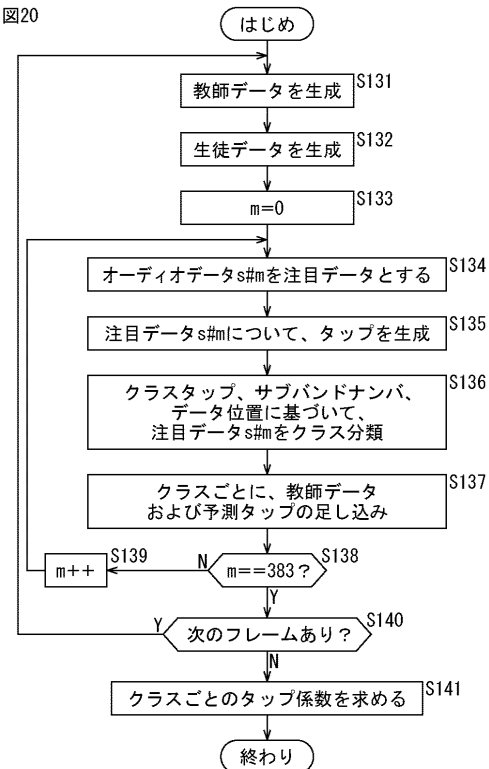
【図 19】

図19



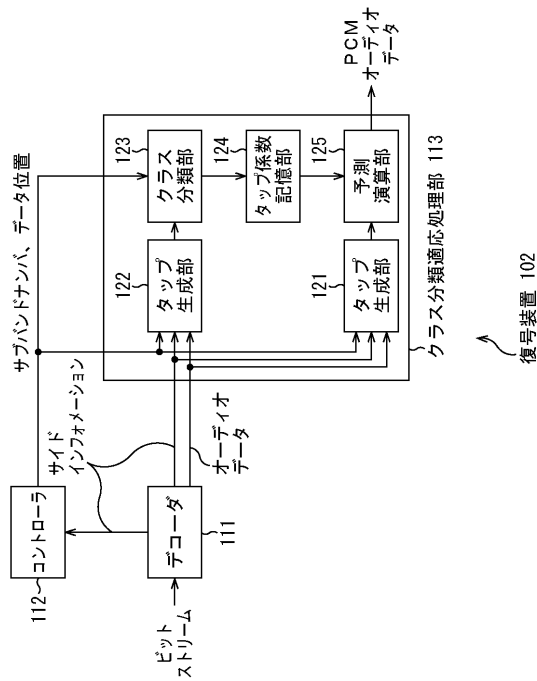
【図 20】

図20



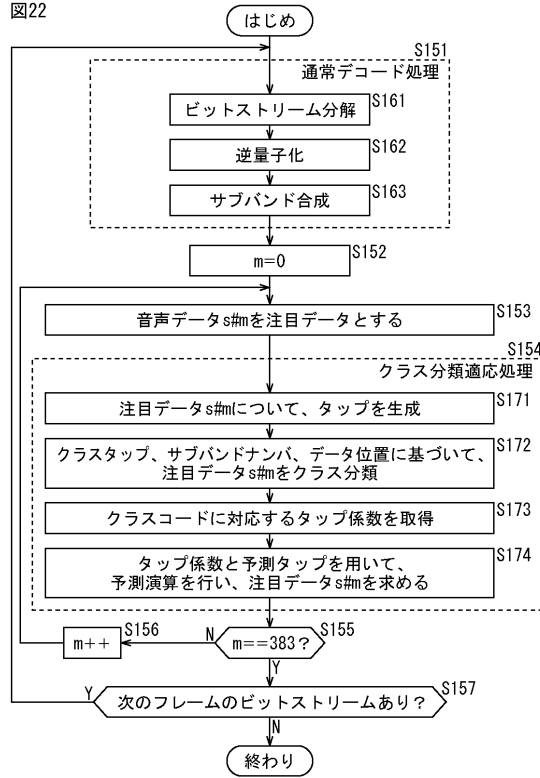
【図 2 1】

図21



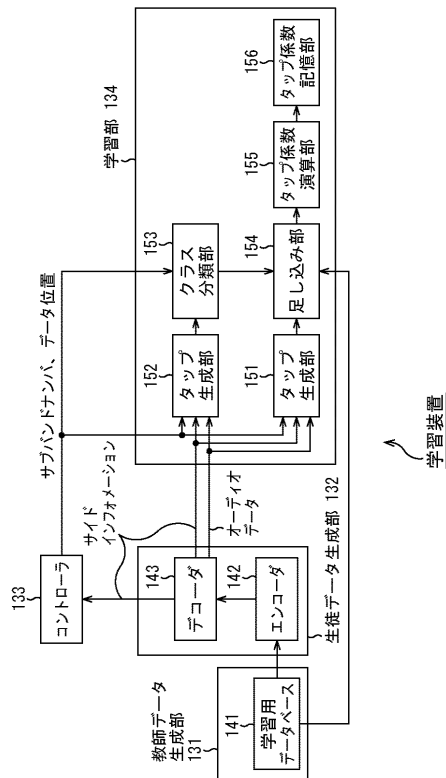
【図 2 2】

図22



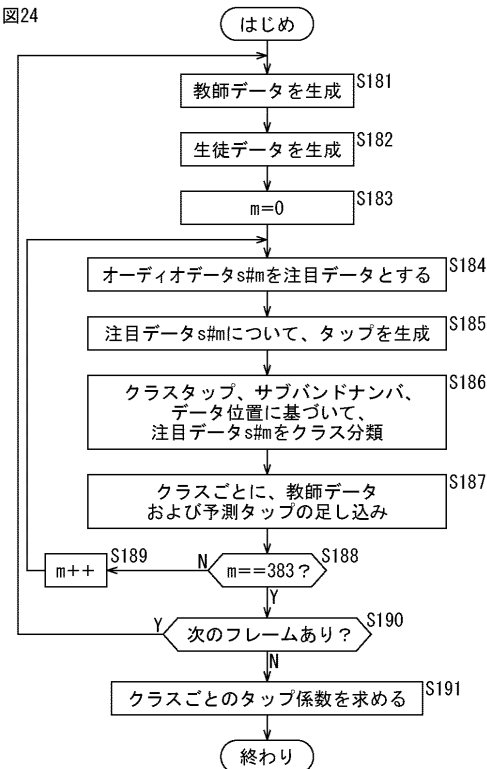
【図 2 3】

図23



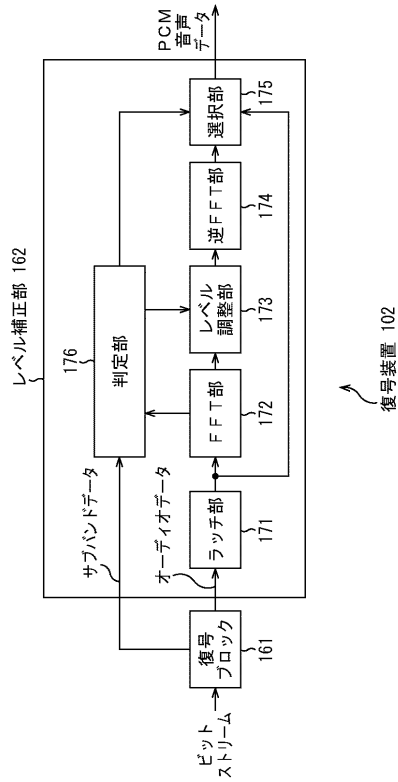
【図 2 4】

図24



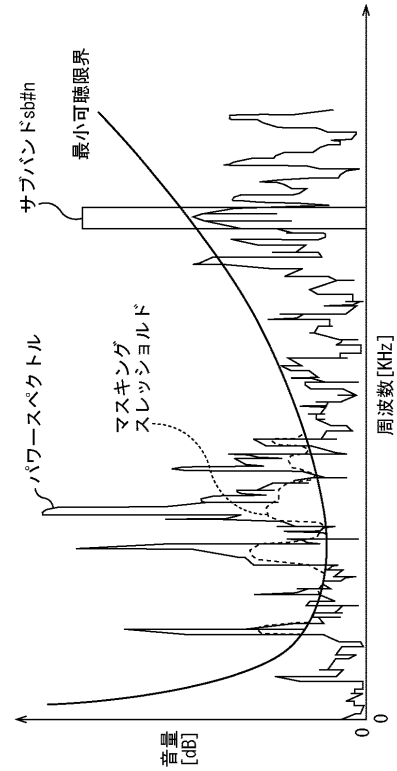
【図 25】

図25



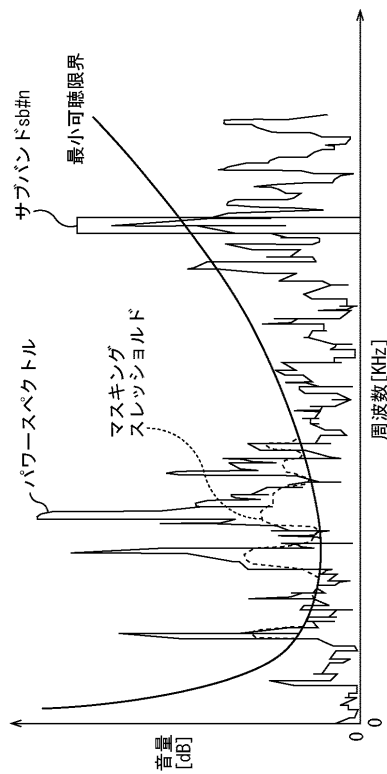
【図 26】

図26



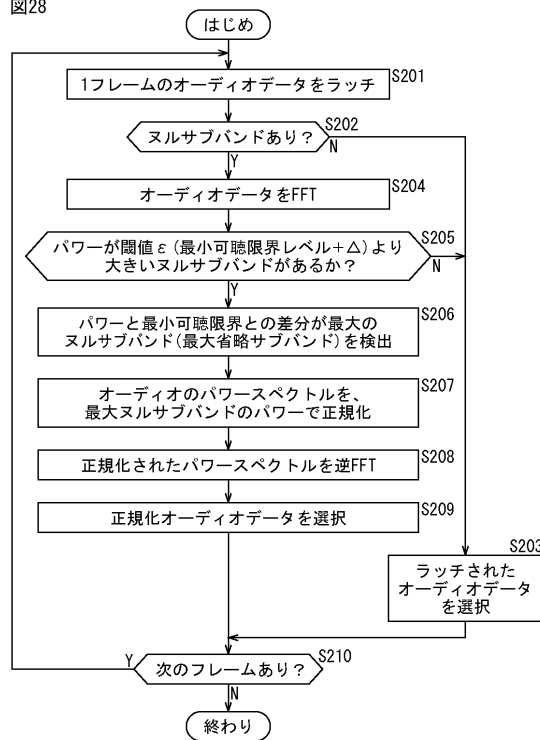
【図 27】

図27

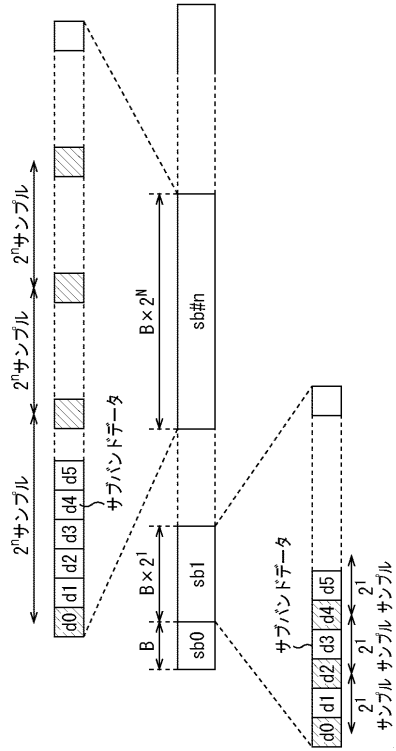


【図 28】

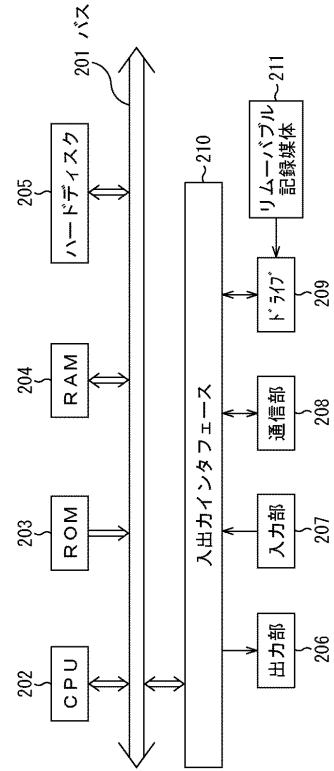
図28



【図 29】
図29



【図 30】
図30



コンピュータ

フロントページの続き

(56)参考文献 特開 2 0 0 1 - 3 4 6 2 1 0 (J P , A)
特開 2 0 0 2 - 0 4 9 3 9 8 (J P , A)
特開 2 0 0 2 - 0 4 9 3 9 5 (J P , A)

(58)調査した分野(Int.Cl. , D B 名)
G10L 19/00-21/06