

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2016-206244
(P2016-206244A)

(43) 公開日 平成28年12月8日(2016.12.8)

(51) Int.Cl.
G10L 19/035 (2013.01)

F I
G10L 19/035

テーマコード (参考)

E

審査請求 未請求 請求項の数 8 O L (全 31 頁)

(21) 出願番号 特願2015-83660 (P2015-83660)
(22) 出願日 平成27年4月15日 (2015.4.15)

(71) 出願人 000005223
富士通株式会社
神奈川県川崎市中原区上小田中4丁目1番1号
(74) 代理人 100074099
弁理士 大菅 義之
(74) 代理人 100133570
弁理士 ▲徳▼永 民雄
(72) 発明者 伊藤 周作
福岡県福岡市早良区百道浜2丁目2番1号
富士通九州ネットワークテクノロジーズ株式会社内
(72) 発明者 岸 洋平
神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

最終頁に続く

(54) 【発明の名称】 符号化装置、符号化方法、及びプログラム

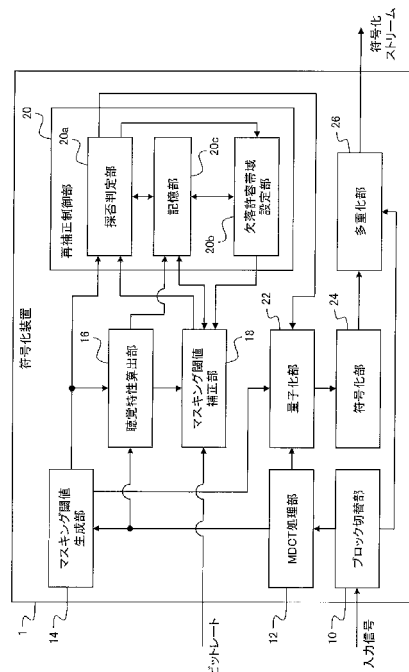
(57) 【要約】

【課題】聴覚特性に基づくマスキング閾値を用いてオーディオ信号を符号化する際の音質の劣化を抑制する。

【解決手段】オーディオ信号を周波数スペクトルに変換し、当該周波数スペクトルの量子化及び符号化を行う符号化装置1は、周波数スペクトルに基づいて量子化する際の初期マスキング閾値を生成するマスキング閾値生成部14と、周波数スペクトルの量子化に使用可能なビット量に基づいて初期マスキング閾値を補正するマスキング閾値補正部18と、補正されたマスキング閾値と初期マスキング閾値との概形の類似度が基準値以下の場合に、周波数スペクトルの帯域毎に量子化による欠落を許容するか否かを設定し、マスキング閾値補正部18に初期マスキング閾値を再度補正させる再補正制御部20と、を備える。

【選択図】 図1

本発明の第1の実施形態に係る符号化装置の構成例を示すブロック図



【特許請求の範囲】**【請求項 1】**

オーディオ信号を周波数スペクトルに変換し、当該周波数スペクトルの量子化及び符号化を行う符号化装置において、

前記周波数スペクトルに基づいて量子化する際の初期マスクング閾値を生成する閾値生成部と、

前記周波数スペクトルの量子化に使用可能なビット量に基づいて前記初期マスクング閾値を補正する閾値補正部と、

前記閾値補正部で補正されたマスクング閾値と前記初期マスクング閾値との概形の類似度が基準値以下の場合に、前記周波数スペクトルの帯域毎に量子化による欠落を許容するか否かを設定し、前記閾値補正部に前記初期マスクング閾値を再度補正させる再補正制御部と、

を備えることを特徴とする符号化装置。

【請求項 2】

前記再補正制御部は、前記初期マスクング閾値及び補正された前記マスクング閾値における欠落を許容しない帯域の概形の類似度が前記基準値以下の場合に、前記閾値補正部に前記初期マスクング閾値を再度補正させる、

ことを特徴とする請求項 1 に記載の符号化装置。

【請求項 3】

前記再補正制御部は、前記初期マスクング閾値と補正された前記マスクング閾値との概形の類似度に基づいて符号化対象の全帯域に対する欠落を許容する帯域の割合を決定し、決定した前記割合に基づいて前記周波数スペクトルの各帯域に対し欠落を許容するか否かを設定する設定部を有する、

ことを特徴とする請求項 1 に記載の符号化装置。

【請求項 4】

前記設定部は、前記マスクング閾値の概形の類似度として前記周波数スペクトル及び前記初期マスクング閾値に基づく第 1 の知覚エントロピー値と、前記周波数スペクトルの量子化に使用可能なビット数に基づく第 2 の知覚エントロピー値との差分値を用い、当該差分値に基づいて前記割合を決定する、

ことを特徴とする請求項 3 に記載の符号化装置。

【請求項 5】

前記設定部は、前記マスクング閾値の概形の類似度として符号化対象の帯域の数に対する、補正された前記マスクング閾値が前記周波数スペクトルの電力値よりも大きい帯域の数の割合を用い、当該割合に基づいて前記符号化対象の全帯域に対する欠落を許容する帯域の割合を決定する、

ことを特徴とする請求項 3 に記載の符号化装置。

【請求項 6】

前記再補正制御部は、前記周波数スペクトルの帯域のうち重要度の低い帯域から順に欠落を許容する、

ことを特徴とする請求項 1 に記載の符号化装置。

【請求項 7】

コンピュータが、

オーディオ信号から得た周波数スペクトルに基づいて生成した初期マスクング閾値が当該周波数スペクトルの量子化に用いるための条件を満たしていない場合に、

前記周波数スペクトルの各帯域に対し量子化による欠落を許容するか否かを設定し、

前記周波数スペクトルの量子化に使用可能なビット量及び各帯域の欠落を許容するか否かの設定に基づいて前記初期マスクング閾値を補正し、

補正されたマスクング閾値と前記初期マスクング閾値との概形の類似度が基準値以下の場合、1 つ以上の帯域の欠落を許容するか否かの設定を変更して前記初期マスクング閾値を再度補正する、

10

20

30

40

50

処理を実行することを特徴とする符号化方法。

【請求項 8】

オーディオ信号から得た周波数スペクトルに基づいて生成した初期マスキング閾値が当該周波数スペクトルの量子化に用いるための条件を満たしていない場合に、

前記周波数スペクトルの各帯域に対し量子化による欠落を許容するか否かを設定し、

前記周波数スペクトルの量子化に使用可能なビット量及び各帯域の欠落を許容するか否かの設定に基づいて前記初期マスキング閾値を補正し、

補正されたマスキング閾値と前記初期マスキング閾値との概形の類似度が基準値以下の場合、1つ以上の帯域の欠落を許容するか否かの設定を変更して前記初期マスキング閾値を再度補正する、

処理をコンピュータに実行させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、符号化装置、符号化方法、及びプログラムに関する。

【背景技術】

【0002】

オーディオ信号や音声信号（以下、まとめて「オーディオ信号」という）の符号化方式には、Advanced Audio Coding（AAC）方式等、人間の聴覚特性を利用して情報量を低減する方式がある。この種の符号化方式では、オーディオ信号を少ないビット数で量子化した場合に増える量子化誤差を所定のマスキング閾値以下に抑えることで、知覚できる雑音を増やすことなく符号化に必要なビット数（すなわち情報量）を低減する。

【0003】

マスキング閾値の理想的な値は、人間が知覚することのできない量子化誤差量の上限值である。そのため、マスキング閾値は、聴覚心理モデルに基づいて算出している。以下、聴覚心理モデルに基づいて算出した理想的なマスキング閾値を初期マスキング閾値という。

【0004】

ところが、低ビットレート条件（例えば64kbps以下）での符号化においては、使用可能なビット数が少ないため、量子化誤差を初期マスキング閾値以下に抑えられない場合が多い。量子化誤差を初期マスキング閾値以下に抑えられない場合、ビットレート条件に基づいて初期マスキング閾値を補正する。（例えば、非特許文献1を参照。）

【0005】

また、低ビットレート条件での符号化において限られた量のビットを効率よく利用する方法として、オーディオ信号に応じて各帯域へのビット数の割り当てを固定又は可変に適応的に切り替える方法が知られている（例えば、特許文献1を参照）。

【先行技術文献】

【特許文献】

【0006】

【特許文献1】特開2000-151413号公報

【非特許文献】

【0007】

【非特許文献1】"3GPP TS 26.403 V9.0.0", [online], 3GPP, 平成27年3月8日検索, インターネット URL: http://www.arib.or.jp/IMT-2000/V900Jul11/5_Appendix/Rel9/26/26403-900.pdf

【発明の概要】

【発明が解決しようとする課題】

【0008】

マスキング閾値の補正方法は、量子化により帯域が欠落することを許容する条件で行う方法と、帯域の欠落を許容しない条件で行う方法とに大別される。

10

20

30

40

50

【 0 0 0 9 】

帯域の欠落を許容する条件でマスキング閾値を補正した場合、補正量が多くなると、人間が知覚可能な音を含む帯域が量子化により欠落してしまうことがある。量子化（符号化）により知覚可能な音を含む帯域が欠落すると、符号化されたオーディオ信号を再生（復号化）したときに再生音を聴く人に違和感を与える。そのため、欠落する帯域が増えると音質の劣化につながる。

【 0 0 1 0 】

一方、帯域の欠落を許容しない条件でマスキング閾値を補正する場合、各帯域のマスキング閾値に上限値を設定して行う。そのため、補正量が上限値に達して更なる補正をできない帯域が生じた場合、その帯域の補正量を増やせない（言い換えると割り当てるビット数を減らせない）分、他の帯域の補正量を増やすこととなる。したがって、初期マスキング閾値と上限値との差が大きい帯域のマスキング閾値が過度に補正され、当該帯域の符号化に割り当てるビットが少なくなってしまう。初期マスキング閾値と上限値との差が大きい帯域は、音質的に重要な帯域である。すなわち、帯域の欠落を許容しない条件でマスキング閾値を補正した場合、音質的に重要な帯域に割り当てるビット数が少なくなり、音質の劣化につながる。

10

【 0 0 1 1 】

一つの側面において、本発明は、聴覚特性に基づくマスキング閾値を用いてオーディオ信号を符号化する際の音質の劣化を抑制することを目的とする。

20

【 課題を解決するための手段 】

【 0 0 1 2 】

本発明の1つの態様である符号化装置は、オーディオ信号を周波数スペクトルに変換し、当該周波数スペクトルの量子化及び符号化を行う符号化装置において、閾値生成部と、閾値補正部と、再補正制御部と、を備える。前記閾値生成部は、前記周波数スペクトルに基づいて量子化する際の初期マスキング閾値を生成する。前記閾値補正部は、前記周波数スペクトルの量子化に使用可能なビット量に基づいて前記初期マスキング閾値を補正する。前記再補正制御部は、前記閾値補正部で補正されたマスキング閾値と前記初期マスキング閾値との概形の類似度が基準値以下の場合に、前記周波数スペクトルの帯域毎に欠落を許容するか否かを設定し、前記閾値補正部に前記初期マスキング閾値を再度補正させる。

30

【 発明の効果 】

【 0 0 1 3 】

上述の態様によれば、聴覚特性に基づくマスキング閾値を用いてオーディオ信号を符号化する際の音質の劣化を抑制することができる。

【 図面の簡単な説明 】

【 0 0 1 4 】

【 図 1 】本発明の第1の実施形態に係る符号化装置の構成例を示すブロック図である。

【 図 2 A 】第1の実施形態に係る符号化装置における符号化処理を示すフローチャート（その1）である。

【 図 2 B 】第1の実施形態に係る符号化装置における符号化処理を示すフローチャート（その2）である。

40

【 図 2 C 】第1の実施形態に係る符号化装置における符号化処理を示すフローチャート（その3）である。

【 図 3 】第1の実施形態における欠落許容帯域設定処理の内容を示すフローチャートである。

【 図 4 A 】周波数スペクトル及び初期マスキング閾値の一例を示すグラフである。

【 図 4 B 】1回目の補正で得られたマスキング閾値と初期マスキング閾値との関係を説明するグラフである。

【 図 4 C 】2回目の補正で得られたマスキング閾値と1回目の補正で得られたマスキング閾値との関係を説明するグラフである。

【 図 4 D 】3回目の補正で得られたマスキング閾値と2回目の補正で得られたマスキング

50

閾値との関係を説明するグラフである。

【図5】符号化装置として用いるコンピュータのハードウェア構成例を示す模式図である。

【図6】第1の実施形態に係るマスキング閾値の補正処理の変形例を示すフローチャートである。

【図7】本発明の第2の実施形態に係るマスキング閾値の補正処理を示すフローチャートである。

【図8】第2の実施形態における欠落許容帯域設定処理の内容を示すフローチャートである。

【図9】欠落を許容する帯域の割合の決定方法を説明するグラフである。

【図10】本発明の第3の実施形態に係る欠落許容帯域設定処理の内容を示すフローチャートである。

【図11】欠落を許容する帯域の割合の決定方法を説明するグラフである。

【発明を実施するための形態】

【0015】

[第1の実施形態]

図1は、本発明の第1の実施形態に係る符号化装置の構成例を示すブロック図である。

【0016】

図1に示すように、本実施形態の符号化装置1は、ブロック切替部10、MDC T処理部12、マスキング閾値生成部14、聴覚特性算出部16、マスキング閾値補正部18、再補正制御部20、量子化部22、符号化部24、及び多重化部26を備える。

【0017】

ブロック切替部10は、入力信号（オーディオ信号）の特性に基づいて、入力信号に対しModified Discrete Cosine Transform（MDC T）処理を行う際のブロック長を切り替える。AAC方式での符号化においては、長ブロック（1024点）又は短ブロック（128点）にブロック長を切り替える。

【0018】

MDC T処理部12は、入力信号に対し長ブロック又は短ブロックに応じた窓長のMDC T処理を行い、入力信号を周波数スペクトルに変換する。AAC方式での符号化においては、ブロック長が長ブロックであれば窓長2048のMDC T処理を行い、短ブロックであれば窓長256のMDC T処理を行う。

【0019】

マスキング閾値生成部14は、入力信号に対して聴覚心理分析を行い、当該入力信号から得た周波数スペクトルの量子化において最適なマスキング閾値（初期マスキング閾値） $sfbThr_0$ を生成する。初期マスキング閾値 $sfbThr_0$ は、周波数スペクトルにおけるスケールファクタバンド sfb （以下「帯域 sfb 」ともいう）毎に生成する。また、マスキング閾値生成部14は、周波数スペクトルの各帯域 sfb における電力値（入力パワー） $mdct_pow(sfb)$ と初期マスキング閾値 $sfbThr_0(sfb)$ とに基づいて、符号化対象の帯域 sfb を決定する。更に、マスキング閾値生成部14は、符号化対象の帯域 sfb を決定した後、初期マスキング閾値を用いた量子化が可能か否か、言い換えるとマスキング閾値を補正する必要があるか否かを判断する。

【0020】

聴覚特性算出部16は、マスキング閾値の補正に必要な聴覚特性を算出する。本実施形態の聴覚特性算出部16は、聴覚特性として、各帯域 sfb の信号対マスク比（Signal Mask Ratio；SMR）を算出する。

【0021】

マスキング閾値補正部18は、聴覚特性（信号対マスク比）及び各帯域 sfb に設定されたマスキング閾値の上限値に基づいて、マスキング閾値を補正する。なお、本実施形態におけるマスキング閾値補正部18は、一組の初期マスキング閾値に対する補正が1回目の場合、補正を行う前に、全ての帯域 sfb のマスキング閾値の上限値を、量子化による欠落

10

20

30

40

50

を許容しないことを表す値に設定する。また、後述するように、一組の初期マスクング閾値に対する補正が2回目以降の場合、マスクング閾値補正部18は、いくつかの帯域の欠落を許容する条件下で補正を行う。

【0022】

再補正制御部20は、補正されたマスクング閾値の採否を判定し、採用しない場合にはマスクング閾値補正部18にマスクング閾値の補正を再度行わせる。マスクング閾値の補正を再度行わせる場合、再補正制御部20は、量子化による欠落を許容する帯域を設定する。この再補正制御部20は、採否判定部20a、欠落許容帯域設定部20b、及び記憶部20cを有する。

【0023】

採否判定部20aは、補正されたマスクング閾値の採否を判定する。本実施形態では、補正後のマスクング閾値と初期マスクング閾値との概形の類似度に基づいて、マスクング閾値の採否を判定する。概形の類似度には、補正後のマスクング閾値 $sfbThr(sfb)$ と初期マスクング閾値 $sfbThr_0(sfb)$ との相互相関値を用いる。

【0024】

欠落許容帯域設定部20bは、採否判定部20aにおいて不採用と判定した場合に、欠落を許容する帯域を設定する。本実施形態では、マスクング閾値の上限値が欠落を許容しない値に設定された帯域のうち重要度が最も低い帯域の欠落を許容する。帯域の重要度には信号対マスク比(SMR)を用い、信号対マスク比が大きいほど重要度が高い帯域とする。

【0025】

記憶部20cには、マスクング閾値を補正しなおす際に必要な情報を記憶させる。本実施形態では、符号化対象の各帯域 sfb についての初期マスクング閾値 $sfbThr_0(sfb)$ 、信号対マスク比、マスクング閾値の上限値、及び重要度を含む情報を記憶させる。

【0026】

量子化部22は、初期マスクング閾値及び補正されたマスクング閾値のいずれかを用いて、周波数スペクトルにおける符号化対象の帯域を量子化する。

【0027】

符号化部24は、周波数スペクトルを量子化して得られた値を符号化する。AAC方式での符号化の場合、符号化部24は量子化して得られた値をハフマン符号化する。

【0028】

多重化部26は、符号化されたオーディオ信号を多重化して符号化ストリームを生成する。

【0029】

次に、本実施形態に係る符号化装置1における符号化処理について説明する。

図2Aは、第1の実施形態に係る符号化装置における符号化処理を示すフローチャート(その1)である。図2Bは、第1の実施形態に係る符号化装置における符号化処理を示すフローチャート(その2)である。図2Cは、第1の実施形態に係る符号化装置における符号化処理を示すフローチャート(その3)である。

【0030】

本実施形態の符号化装置1は、入力信号(オーディオ信号)におけるフレーム等の符号化単位のデータのそれぞれに対し、図2A~図2Cに示すような符号化処理を行う。

【0031】

符号化装置1は、図2Aに示すように、まず、1フレーム分の入力信号を周波数スペクトルに変換し、各帯域 sfb の電力値 $mdct_pow(sfb)$ を算出する(ステップS10)。ステップS10の処理は、ブロック切替部10及びMDCT処理部12が行う。

【0032】

ブロック切替部10は、MDCT処理のブロック長を長ブロック及び短ブロックのいずれにするかを選択して切り替える。ブロック長は、既知の選択方法、例えば入力信号の電力変動比と予測利得変動比とに基づいて選択する。

10

20

30

40

50

【 0 0 3 3 】

また、M D C T 処理部 1 2 は、ブロック切替部 1 0 で選択したブロック長に応じた窓長の M D C T 処理を行い、入力信号を周波数スペクトルに変換する。M D C T 処理部 1 2 は、例えば、下記式 (1) により入力信号 x_{in} を周波数スペクトル $mdct(k)$ に変換する。

【 0 0 3 4 】

【 数 1 】

$$mdct(k) = 2 \cdot \sum_{n=0}^{N-1} x_{in} \cos\left(\frac{2\pi}{N}(n+n_0)\left(k+\frac{1}{2}\right)\right) \quad (1)$$

$$0 \leq k < \frac{N}{2}, \quad n_0 = \frac{1}{2}\left(\frac{N}{2}+1\right)$$

【 0 0 3 5 】

式 (1) における N は、M D C T 処理の窓長である。A A C 方式での符号化の場合、ブロック長が長ブロックであれば窓長 N を 2 0 4 8 とし、短ブロックであれば窓長 N を 2 5 6 とし、M D C T 処理を行う。

【 0 0 3 6 】

また、M D C T 処理部 1 2 は、得られた周波数スペクトルに基づき、各帯域 (スケールファクターバンド sfb) の電力値 $mdct_pow(sfb)$ を算出する。電力値 $mdct_pow(sfb)$ は、例えば、下記式 (2) により算出する。

【 0 0 3 7 】

【 数 2 】

$$mdct_pow(sfb) = \sum_{k=sboffset(x)}^{sboffset(x)+1} \sqrt{|mdct(k)|} \quad (2)$$

【 0 0 3 8 】

なお、オーディオ信号の周波数スペクトルへの変換は、式 (1) を用いた変換に限らず、既知の変換方法のいずれかを用いて行えばよい。同様に、周波数スペクトルの各帯域 sfb の電力値 $mdct_pow(sfb)$ は、式 (2) に限らず、既知の算出方法のいずれかを用いて算出すればよい。

【 0 0 3 9 】

符号化装置 1 は、次に、周波数スペクトルを量子化する際の初期マスキング閾値 $sfbThr_0(sfb)$ を生成する (ステップ S 1 2) 。ステップ S 1 2 の処理は、マスキング閾値生成部 1 4 が行う。

【 0 0 4 0 】

マスキング閾値生成部 1 4 は、入力信号に対して聴覚心理分析を行い、帯域 sfb 毎に初期マスキング閾値 $sfbThr_0(sfb)$ を求める。初期マスキング閾値 $sfbThr_0(sfb)$ は、各帯域 sfb における最小可聴レベルやマスキング効果等に基づき、既知の算出方法のいずれかを用いて算出する。

【 0 0 4 1 】

また、マスキング閾値生成部 1 4 は、初期マスキング閾値 $sfbThr_0(sfb)$ を生成すると、次に、初期マスキング閾値 $sfbThr_0(sfb)$ と周波数スペクトルの電力値 $mdct_pow(sfb)$ とに基づいて符号化対象の帯域を決定する (ステップ S 1 4) 。ステップ S 1 4 の処理において、マスキング閾値生成部 1 4 は、周波数スペクトルの全帯域のうち、 $sfbThr_0(sfb) < mdct_pow(sfb)$ である帯域のみを符号化対象とする。

【 0 0 4 2 】

符号化対象の帯域を決定した後、マスキング閾値生成部 1 4 は、マスキング閾値を補正するか否かを判定するため、初期 P E 値及び目標 P E 値を算出する (ステップ S 1 6) 。本実施形態では、初期 P E 値が目標 P E 値より大きいか否かにより、マスキング閾値を補

10

20

30

40

50

正するか否かを判定する（ステップ S 1 8）。

【 0 0 4 3 】

ここで、P E 値とは、音響パラメータの 1 つである知覚エントロピー（Perceptual Entropy）の値であり、符号化の難しさを表す。初期 P E 値は、符号化対象の帯域における電力値 $mdct_pow(sfb)$ と初期マスキング閾値 $sfbThr_0(sfb)$ とに基づいて算出される P E 値である。また、目標 P E 値は、符号化に使用可能なビット数に基づいて算出される P E 値である。初期 P E 値及び目標 P E 値は、既知の算出方法のいずれか（例えば、非特許文献 1 に記載された算出方法）を用いて算出する。

【 0 0 4 4 】

知覚エントロピーの値は、上記のように量子化に必要なビット数と関係があり、目標 P E 値に対して初期 P E 値が大きい場合、初期マスキング閾値を用いた量子化で使用するビット量がビットレートに応じて与えられるビット数を超えると判断できる。一方、初期 P E 値が目標 P E 値以下である場合、初期マスキング閾値を用いた量子化で使用するビット量がビットレートに応じて与えられるビット数内に収まると判断できる。よって、初期 P E 値と目標 P E 値との大小関係に基づいて初期マスキング閾値を用いた量子化が可能であるか否か、すなわちマスキング閾値を補正する必要があるか否かを判断できる。

10

【 0 0 4 5 】

初期 P E 値が目標 P E 値以下の場合（ステップ S 1 8 ; N o）、マスキング閾値生成部 1 4 は、マスキング閾値を補正しないと判定し、量子化部 2 2 に初期マスキング閾値 $sfbThr_0(sfb)$ を渡す。この場合、符号化装置 1 は、図 2 C に示すように、初期マスキング閾値 $sfbThr_0(sfb)$ を用いて周波数スペクトルを量子化する（ステップ S 3 2）。ステップ S 3 2 の量子化は、量子化部 2 2 が行う。量子化部 2 2 は、既知の量子化方法のいずれかを用いて周波数スペクトルを量子化する。

20

【 0 0 4 6 】

一方、初期 P E 値が目標 P E 値より大きい場合（ステップ S 1 8 ; Y e s）、マスキング閾値生成部 1 4 は、マスキング閾値を補正すると判定し、初期マスキング閾値 $sfbThr_0(sfb)$ 及び目標 P E 値を、聴覚特性算出部 1 6 及び採否判定部 2 0 a に渡す。この場合、符号化装置 1 は、図 2 B に示したステップ S 2 0 ~ S 3 0 のようなマスキング閾値の補正処理を行う。

【 0 0 4 7 】

マスキング閾値を補正する場合、符号化装置 1 は、次に、周波数スペクトル等に基づき聴覚特性を算出する（ステップ S 2 0）。ステップ S 2 0 の処理は、聴覚特性算出部 1 6 が行う。

30

【 0 0 4 8 】

聴覚特性算出部 1 6 は、各帯域 sfb における信号対マスク比（S M R）、すなわち各帯域における電力値 $mdct_pow(sfb)$ と初期マスキング閾値 $sfbThr_0(sfb)$ との差分値を算出する。信号対マスク比を算出すると、聴覚特性算出部 1 6 は、初期マスキング閾値 $sfbThr_0(sfb)$ 、目標 P E 値、及び信号対マスク比をマスキング閾値補正部 1 8 に渡すとともに、信号対マスク比を再補正制御部 2 0 の記憶部 2 0 c に記憶させる。

【 0 0 4 9 】

次に、符号化装置 1 は、欠落防止処理及びマスキング閾値の補正処理を行う。欠落防止処理及びマスキング閾値の補正処理は、マスキング閾値補正部 1 8 が行う。

40

【 0 0 5 0 】

欠落防止処理は、補正により得たマスキング閾値を用いて周波数スペクトルを量子化した際に帯域が欠落することを防止する処理である。マスキング閾値補正部 1 8 は、欠落防止処理として、各帯域 sfb における補正後のマスキング閾値 $sfbThr(sfb)$ の上限値を、欠落を許容しない値に設定する処理を行う（ステップ S 2 2）。量子化した際の帯域の欠落を防止するには、補正後のマスキング閾値 $sfbThr(sfb)$ が電力値 $mdct_pow(sfb)$ よりも大きくならないようにすればよい。よって、ステップ S 2 2 では、補正後のマスキング閾値 $sfbThr(sfb)$ の上限値を、電力値 $mdct_pow(sfb)$ よりもわずかに小さな値に設定する。

50

【 0 0 5 1 】

マスクング閾値補正部 1 8 は、次に、聴覚特性及びマスクング閾値の上限値、並びにビットレートに基づいてマスクング閾値を補正する（ステップ S 2 4）。ステップ S 2 4 では、例えば、下記式（3）を用い、電力値 $mdct_pow(sfb)$ と補正後のマスクング閾値 $sfbThr(sfb)$ とに基づいて算出される P E 値が目標 P E 値になるようマスクング閾値を補正する。

【 0 0 5 2 】

【 数 3 】

$$sfbThr(sfb) = \left\{ sfbThr_0(sfb)^{0.25} + r \right\}^4 \quad (3)$$

10

【 0 0 5 3 】

式（3）において、 r は補正パラメータである（非特許文献 1 を参照）。

【 0 0 5 4 】

ステップ S 2 4 における 1 回目のマスクング閾値の補正は、補正後のマスクング閾値を用いて周波数スペクトルを量子化した場合に欠落する帯域がない条件下で行っている。このように欠落する帯域がない条件下でマスクング閾値を補正した場合、補正後のマスクング閾値の概形が初期マスクング閾値の概形から大きくずれてしまい、音質が劣化する恐れがある。そこで、本実施形態の符号化装置 1 では、補正前後のマスクング閾値の概形の類似度と予め定めた基準値とを比較して、補正後のマスクング閾値を採用するか否かを判定する。そして、概形の類似度が基準値以下の場合、直前の補正処理により補正されたマスクング閾値を採用せず、マスクング閾値の補正を再度行う。補正後のマスクング閾値 $sfbThr(sfb)$ の採否の判定等は、再補正制御部 2 0 が行う。

20

【 0 0 5 5 】

マスクング閾値補正部 1 8 は、マスクング閾値を補正した後、得られたマスクング閾値 $sfbThr(sfb)$ を再補正制御部 2 0 の採否判定部 2 0 a に渡すとともに、マスクング閾値の上限値を記憶部 2 0 c に記憶させる。すると、採否判定部 2 0 a は、補正後のマスクング閾値 $sfbThr(sfb)$ の採否を判定するため、初期マスクング閾値 $sfbThr_0(sfb)$ と補正後のマスクング閾値 $sfbThr(sfb)$ との相互相関値 $corre$ を算出する（ステップ S 2 6）。補正後のマスクング閾値 $sfbThr(sfb)$ の採否は、算出した相互相関値 $corre$ が予め定めた基準値（相関閾値 TH_1 ）より大きいかが否かで判定する（ステップ S 2 8）。

30

【 0 0 5 6 】

閾値概形判定部 2 0 a は、例えば、下記式（4）により相互相関値 $corre$ を算出する。

【 0 0 5 7 】

【 数 4 】

$$corre = \frac{\frac{1}{maxsfb} \sum_{i=0}^{maxsfb-1} sfbThr_0(i) \cdot sfbThr(i)}{\sqrt{\frac{1}{maxsfb} \sum_{i=0}^{maxsfb-1} sfbThr_0(i)^2} \sqrt{\frac{1}{maxsfb} \sum_{i=0}^{maxsfb-1} sfbThr(i)^2}} \quad (4)$$

40

【 0 0 5 8 】

なお、相互相関値 $corre$ は、符号化対象の帯域のマスクング閾値のみを用いて算出する。

【 0 0 5 9 】

式（4）により算出される相互相関値 $corre$ は、 $0 < corre < 1$ となり、補正前後のマスクング閾値の概形の類似度が高いほど値が大きくなる。そのため、算出した相互相関値 $corre$ が相関閾値 TH_1 （例えば $TH_1 = 0.8$ ）よりも大きい場合（ステップ S 2 8；Yes）、採否判定部 2 0 a は、直前の補正処理で得た補正後のマスクング閾値 $sfbThr(sfb)$

50

)を採用すると判定する。そして、採否判定部20aは、直前の補正処理で得たマスキング閾値 $sfbThr(sfb)$ を量子化部22に渡す。この場合、符号化装置1は、図2Cに示すように、補正後のマスキング閾値 $sfbThr(sfb)$ を用いて周波数スペクトルを量子化する(ステップS34)。

【0060】

一方、相互相関値 $corre$ が相関閾値 TH_1 以下の場合(ステップS28; No)、採否判定部20aは、直前の補正処理で得た補正後のマスキング閾値 $sfbThr(sfb)$ を採用しない(すなわちマスキング閾値を補正しなおす)と判定する。この場合、採否判定部20aは、欠落許容帯域設定部20bに欠落許容帯域設定処理(ステップS30)を行わせる。

【0061】

なお、図2Bでは省略しているが、マスキング閾値を補正しなおす場合、採否判定部20aは、初期マスキング閾値 $sfbThr_0(sfb)$ を記憶部20cに記憶させる。記憶部20cに記憶させた初期マスキング閾値 $sfbThr_0(sfb)$ は、補正しなおしたマスキング閾値との相互相関値 $corre$ の算出に用いる。

【0062】

欠落許容帯域設定処理(ステップS30)では、量子化した際に欠落することを許容する帯域を設定する。本実施形態では、ステップS30の処理を行う毎に、マスキング閾値の上限値が欠落を許容しない値に設定された帯域のうち重要度が最も低い帯域を、欠落を許容する帯域に設定する。帯域の重要度は、量子化により欠落した場合の音質劣化への影響の度合いである。本実施形態では、後述するように、帯域の重要度として電力値 $mdct_power(sfb)$ と初期マスキング閾値 $sfbThr_0(sfb)$ との差分値(すなわち信号対マスク比)を用いる。また、欠落を許容する帯域は、マスキング閾値の上限値を「0」又は「-1」等の欠落を許容しない場合の値と識別可能な値に設定する。

【0063】

欠落許容帯域設定部20bは、欠落を許容する帯域の設定を終えると、記憶部20cが記憶しているマスキング閾値の上限値を更新するとともに、マスキング閾値の補正を指示する制御信号をマスキング閾値補正部18に送る。これにより、欠落許容帯域設定処理(ステップS30)が終了する。

【0064】

欠落許容帯域設定処理(ステップS30)が終了すると、符号化装置1が行う符号化処理は、マスキング閾値を補正する処理(ステップS24)に戻る。以後、符号化装置1は、相互相関値 $corre$ が相関閾値 TH_1 より大きくなるまで、ステップS24~S30の処理を繰り返す。そして、相互相関値 $corre$ が相関閾値 TH_1 より大きくなると(ステップS28; Yes)、直前の補正処理で得られたマスキング閾値 $sfbThr(sfb)$ を用いた量子化(ステップS34)を行う。

【0065】

以上の手順により初期マスキング閾値 $sfbThr_0(sfb)$ を用いた量子化(ステップS32)又は補正後のマスキング閾値 $sfbThr(sfb)$ を用いた量子化(ステップS34)を終えると、符号化装置1は、量子化された値を符号化する(ステップS36)。ステップS36は、符号化部24が行う。

【0066】

符号化部24は、固定ハフマン符号化等の既知の符号化方法を用いた符号化を行う。そして、符号化を終えると、符号化部24は、符号化したデータを多重化部26に渡す。これにより、入力信号(オーディオデータ)の1フレーム分の符号化処理が終了する。

【0067】

符号化処理を終えると、符号化装置1(多重化部26)は、符号化されたオーディオデータにヘッダ情報等を付加した符号化ストリームを生成して出力する。

【0068】

このように、本実施形態に係る符号化装置1が行う符号化処理では、初期マスキング閾値を補正する際、まず、全ての帯域 sfb に対し量子化による欠落を許容しない条件を付し

10

20

30

40

50

て補正する。そして、補正後のマスキング閾値 $\text{sfbThr}(\text{sfb})$ と初期マスキング閾値 $\text{sfbThr}_0(\text{sfb})$ との概形の類似度が基準値より低い場合、概形の類似度が基準値を超えるまで、重要度が低い帯域から順に欠落を許容しながらマスキング閾値の補正を繰り返す。これにより、帯域が欠落することによる音質の劣化を抑制しつつ、概形の類似度の低下（過度の補正）による音質劣化を抑制できる。

【0069】

次に、本実施形態における欠落許容帯域設定処理（ステップS30）の内容について説明する。

【0070】

図3は、第1の実施形態における欠落許容帯域設定処理の内容を示すフローチャートである。

10

【0071】

本実施形態では、上記のように、各帯域 sfb の重要度として電力値 $\text{mdct_pow}(\text{sfb})$ と初期マスキング閾値 $\text{sfbThr}_0(\text{sfb})$ との差分値を用いる。この差分値は、ステップS20で算出した信号対マスク比であり、1回目の欠落許容帯域設定処理（ステップS30）を開始する前に記憶部20cに記憶させている。そのため、欠落許容帯域設定部20bが欠落許容帯域設定処理を行う際には、図3に示すように、まず、記憶部20cから符号化対象の各帯域の信号対マスク比及びマスキング閾値の上限値を読み出す（ステップS3000）。

【0072】

次に、欠落許容帯域設定部20bは、マスキング閾値の上限値が欠落を許容しない値に設定された帯域のうち信号対マスク比が最も小さい帯域を、重要度が最も低い帯域に特定する（ステップS3002）。欠落を許容しないマスキング閾値の上限値は、上記のように電力値よりも小さな正の値である。

20

【0073】

次に、欠落許容帯域設定部20bは、重要度が最も低い帯域に対するマスキング閾値の上限値を、欠落を許容する値に変更して記憶部20cのデータを更新する（ステップS3004）。欠落を許容する値は、欠落を許容しないマスキング閾値の上限値との判別ができる値であればよく、例えば「0」又は「-1」にする。

【0074】

ステップS3004の処理を終えると、欠落許容帯域設定部20bは、欠落許容帯域設定処理を終了する（リターン）。

30

【0075】

初期マスキング閾値 $\text{sfbThr}_0(\text{sfb})$ は、聴覚心理モデルに基づいて算出した値であり、人間が知覚することのできない量子化誤差量の制限値である。そして、周波数スペクトルにおける電力値 $\text{mdct_pow}(\text{sfb})$ が初期マスキング閾値より小さい帯域の音は知覚できない。また、電力値が初期マスキング閾値より大きくても両者の差が非常に小さい帯域の音は知覚が困難である。よって、ステップS20で算出した信号対マスク比が非常に小さい帯域は、信号対マスク比が大きい帯域に比べ、量子化により欠落したときに音質劣化に及ぼす影響が小さい。したがって、帯域の重要度として信号対マスク比を用い、信号対マスク比が小さい帯域から順に欠落を許容することで、帯域の欠落による音質劣化を抑制することができる。

40

【0076】

更に、マスキング閾値を補正しなおす際にいくつかの帯域に対し量子化による欠落を許容した場合、欠落を許容した帯域に対するマスキング閾値は、前回の補正時の値より大きな値にすることが可能となる。欠落を許容した帯域のマスキング閾値を大きな値に補正すれば、その帯域の符号化に使用されるビット量を低減する、又は「0」にすることができ、このように、欠落を許容した帯域に使用されるビット量を減らすことができれば、減らした分のビットを他の帯域の符号化に使用することができる。欠落を許容した帯域で使用されるビット量を減らして得たビットは、例えば前回の補正において補正前後のマスキング閾値の差が大きかった帯域に充当される。そのため、補正後のマスキング閾値を初期マ

50

スキング閾値に近づけることができる。よって、量子化に用いるマスキング閾値 $\text{sfbThr}(\text{sfb})$ の概形と初期マスキング閾値 $\text{sfbThr}_0(\text{sfb})$ の概形とのずれによる音質の劣化を抑制することができる。

【0077】

上記のマスキング閾値の補正処理について、図4A～図4Dを参照しながら具体的に説明する。

【0078】

図4Aは、周波数スペクトル及び初期マスキング閾値の一例を示すグラフである。

1フレーム分の入力信号に対しステップS10～S14の処理を行うと、例えば、図4Aに示すような、周波数スペクトルの各帯域 sfb の電力値 $\text{mdct_pow}(\text{sfb})$ 及び初期マスキング閾値 $\text{sfbThr}_0(\text{sfb})$ が得られる。

10

【0079】

初期マスキング閾値 $\text{sfbThr}_0(\text{sfb})$ は、上記のように、対応する周波数スペクトルの量子化に最適なマスキング閾値である。そのため、ステップS16、S18の処理により初期マスキング閾値 $\text{sfbThr}_0(\text{sfb})$ を用いた量子化が可能であると判定した場合、符号化装置1は、初期マスキング閾値 $\text{sfbThr}_0(\text{sfb})$ を用いて周波数スペクトルを量子化する(ステップS32)。

【0080】

しかしながら、低ビットレート条件で符号化する場合、すなわち周波数スペクトルの符号化に使用可能なビット数が少ない場合、量子化誤差を初期マスキング閾値以下にできないことが多い。量子化誤差を初期マスキング閾値以下にできない場合、符号化装置1は、ビットレート条件や聴覚特性等に基づいて、音質がなるべく劣化しない範囲でマスキング閾値を大きくする(緩める)補正を行う。

20

【0081】

図4Bは、1回目の補正で得られたマスキング閾値と初期マスキング閾値との関係を説明するグラフである。

【0082】

本実施形態における初期マスキング閾値に対する1回目の補正は、全ての帯域に対し量子化による欠落を許容しない条件を付して行われる。この1回目の補正を行うと、補正後のマスキング閾値 $\text{sfbThr}(\text{sfb})$ は、例えば、図4Bに実線の折れ線で示したような概形になる。

30

【0083】

この補正後のマスキング閾値 $\text{sfbThr}(\text{sfb})$ の概形と、図4Bに点線で示した初期マスキング閾値 $\text{sfbThr}_0(\text{sfb})$ の概形とを比較すると、低周波の帯域 $\text{sfb}1\sim\text{sfb}3$ や、高周波の帯域 $\text{sfb}15\sim\text{sfb}18$ において両者の類似度が低くなっている。それでも、補正前後のマスキング閾値 $\text{sfbThr}(\text{sfb})$ 、 $\text{sfbThr}_0(\text{sfb})$ から算出した相互相関値 corre が相関閾値 TH_1 よりも大きければ、量子化による帯域の欠落を防ぎつつ、マスキング閾値の概形のずれによる音質劣化を抑制することができる。

【0084】

しかしながら、補正前後のマスキング閾値 $\text{sfbThr}(\text{sfb})$ 、 $\text{sfbThr}_0(\text{sfb})$ から算出した相互相関値 corre が相関閾値 TH_1 以下である場合、いくつかの帯域のマスキング閾値が過度に補正されており音質劣化につながる。よって、図4Bに示した補正前後のマスキング閾値 $\text{sfbThr}(\text{sfb})$ 、 $\text{sfbThr}_0(\text{sfb})$ から算出した相互相関値 corre が相関閾値 TH_1 以下である場合、重要度が最も低い帯域の欠落を許容してマスキング閾値を補正しなおす。

40

【0085】

本実施形態において重要度が最も低い帯域は、信号対マスク比が最も小さい帯域である。図4Aに示した帯域 $\text{sfb}1\sim\text{sfb}5$ 及び $\text{sfb}11\sim\text{sfb}18$ において信号対マスク比が最も小さい帯域は、帯域 $\text{sfb}5$ である。よって、符号化装置1は、帯域 $\text{sfb}5$ のマスキング閾値の上限値を、欠落を許容する値に変更して、初期マスキング閾値に対する2回目の補正を行う。

【0086】

50

帯域sfb5の欠落のみを許容する条件下でマスクング閾値を補正しなおすと、補正後のマスクング閾値sfbThr₂(sfb)は、例えば、図4Cに実線の折れ線で示したような概形になる。図4Cは、2回目の補正で得られたマスクング閾値と1回目の補正で得られたマスクング閾値との関係を説明するグラフである。

【0087】

2回目の補正で得られたマスクング閾値sfbThr₂(sfb)と、図4Cに破線で示した1回目の補正で得られたマスクング閾値sfbThr(sfb)とを比較すると、いくつかの帯域で2回目の補正で得られたマスクング閾値sfbThr₂(sfb)のほうが小さくなっている。これは、2回目の補正では帯域sfb5の欠落を許容しているためである。

【0088】

2回目の補正では、帯域sfb5に対するマスクング閾値sfbThr₂(sfb5)を1回目の補正時よりも大きくすることができる。帯域sfb5に対するマスクング閾値sfbThr₂(sfb5)が1回目の補正時よりも大きくなれば、帯域sfb5の符号化に使用されるビット量が低減し、低減した分だけ他の帯域の符号化に使用するビット量を増加させることができる。そして、1つの帯域の符号化に使用するビット量が増加すれば、その帯域のマスクング閾値は、1回目の補正時より小さくすることができる。そのため、2回目の補正で得られたマスクング閾値sfbThr₂(sfb)では、図4Cに示したように、1回目の補正において初期マスクング閾値sfbThr₀(sfb)との差が大きかった帯域sfb1, sfb2等の閾値が1回目の補正時の値よりも小さくなっている。すなわち、2回目の補正では、帯域sfb5の欠落を許容したことにより、帯域sfb1, sfb2のマスクング閾値に対する過度の補正が抑制されている。

【0089】

図4Cからもわかるように、2回目の補正で得られたマスクング閾値sfbThr₂(sfb)は、1回目の補正で得られたマスクング閾値sfbThr(sfb)に比べて、初期マスクング閾値sfbThr₀(sfb)との概形の類似度が高い。よって、2回目の補正で得られたマスクング閾値sfbThr₂(sfb)と初期マスクング閾値sfbThr₀(sfb)との相互相関値correlは、1回目の補正で得られたマスクング閾値sfbThr(sfb)と初期マスクング閾値sfbThr₀(sfb)との相互相関値より大きくなる。そして、2回目の補正で得られたマスクング閾値sfbThr₂(sfb)と初期マスクング閾値sfbThr₀(sfb)との相互相関値correlが相関閾値TH₁よりも大きい場合、符号化装置1は、2回目の補正で得られたマスクング閾値sfbThr₂(sfb)を用いて量子化する。これにより、量子化に用いたマスクング閾値と初期マスクング閾値の概形とのずれ、言い換

えるとマスクング閾値の過度の補正による音質劣化を抑制することができる。また、図4Cに示した2回目の補正で得られたマスクング閾値sfbThr₂(sfb)は、欠落を許容した帯域sfb5を含む全ての帯域でマスクング閾値が電力値mdct_pow(sfb)以下となっている。したがって、量子化により帯域が欠落することによる音質の劣化も防げる。

【0090】

しかしながら、2回目の補正で得られたマスクング閾値sfbThr₂(sfb)と初期マスクング閾値sfbThr₀(sfb)とを比較すると、高周波の帯域sfb15~sfb18の類似度が依然低い。そのため、相関閾値TH₁の値によっては、2回目の補正で得られたマスクング閾値と初期マスクング閾値との相互相関値correlが相関閾値TH₁以下になることもある。その場合、符号化装置1は、初期マスクング閾値sfbThr₀(sfb)に対し3回目の補正を行う。3回目の補正は、欠落を許容していない帯域のうち重要度が最も低い帯域、すなわち帯域sfb5の次に信号対マスク比が小さい帯域を、欠落を許容する帯域に追加設定して行う。図4A(図4C)に示した帯域sfb1~sfb5及びsfb11~sfb18において、帯域sfb5の次に信号対マスク比が最も小さい帯域は、帯域sfb15である。よって、符号化装置1は、帯域sfb5, sfb15のマスクング閾値の上限値を、欠落を許容する値に変更した状態で、初期マスクング閾値sfbThr₀(sfb)に対する3回目の補正を行う。

【0091】

帯域sfb5, sfb15の欠落を許容する条件下でマスクング閾値を補正しなおすと、補正後のマスクング閾値sfbThr₃(sfb)は、図4Dに実線の折れ線で示したような概形になる。図4Dは、3回目の補正で得られたマスクング閾値と2回目の補正で得られたマスクング閾

10

20

30

40

50

値との関係を説明するグラフである。

【0092】

3回目の補正で得られたマスクング閾値 $\text{sfbThr}_3(\text{sfb})$ と、図4Dに破線の折れ線で示した2回目の補正で得られたマスクング閾値 $\text{sfbThr}_2(\text{sfb})$ とを比較すると、いくつかの帯域で3回目の補正で得られたマスクング閾値のほうが小さくなっている。これは、欠落を許容した帯域 $\text{sfb}5$ の符号化に使用されるビット量を低減させた分、他の帯域 $\text{sfb}1$, $\text{sfb}2$, $\text{sfb}17$, $\text{sfb}18$ 等の符号化に使用されるビット量が増加したためである。

【0093】

図4D及び図4Cからわかるように、3回目の補正で得られたマスクング閾値 $\text{sfbThr}_3(\text{sfb})$ は、1回目及び2回目の補正で得られたマスクング閾値 $\text{sfbThr}(\text{sfb})$, $\text{sfbThr}_2(\text{sfb})$ に比べ、初期マスクング閾値 $\text{sfbThr}_0(\text{sfb})$ との概形の類似度が高い。よって、3回目の補正で得られたマスクング閾値 $\text{sfbThr}_3(\text{sfb})$ と初期マスクング閾値 $\text{sfbThr}_0(\text{sfb})$ との相互相関値 corre は、1回目及び2回目の補正で得られたマスクング閾値と初期マスクング閾値 $\text{sfbThr}_0(\text{sfb})$ との相互相関値より大きくなる。そして、3回目の補正で得られたマスクング閾値 $\text{sfbThr}_3(\text{sfb})$ と初期マスクング閾値 $\text{sfbThr}_0(\text{sfb})$ との相互相関値 corre が相関閾値 TH_1 よりも大きい場合、符号化装置1は、3回目の補正で得られたマスクング閾値 $\text{sfbThr}_3(\text{sfb})$ を用いて量子化する。これにより、量子化に用いたマスクング閾値と初期マスクング閾値の概形とのずれ、すなわちマスクング閾値の過度の補正による音質劣化を抑制することができる。

【0094】

また、図4Dに示した3回目の補正で得られたマスクング閾値 $\text{sfbThr}_3(\text{sfb})$ は、欠落を許容した帯域 $\text{sfb}5$ を除く全ての帯域でマスクング閾値が電力値 $\text{mdct_pow}(\text{sfb})$ 以下となっている。したがって、量子化により欠落するのは重要度が最も低い帯域 $\text{sfb}5$ のみである。よって、帯域の欠落による音質の劣化も最小限に抑えることができる。

【0095】

なお、相関閾値 TH_1 の値によっては、3回目の補正で得られたマスクング閾値 $\text{sfbThr}_3(\text{sfb})$ と初期マスクング閾値 $\text{sfbThr}_0(\text{sfb})$ との相互相関値 corre が相関閾値 TH_1 以下になることもある。その場合、符号化装置1は、初期マスクング閾値 $\text{sfbThr}_0(\text{sfb})$ に対する4回目の補正を行う。4回目の補正は、欠落を許容していない帯域のうち重要度が最も低い帯域、すなわち帯域 $\text{sfb}15$ の次に信号対マスク比が小さい帯域(例えば帯域 $\text{sfb}14$)を、欠落を許容する帯域に追加設定して行う。以後、符号化装置1は、補正後のマスクング閾値と初期マスクング閾値との相互相関値 corre が相関閾値 TH_1 より大きくなるまで、欠落を許容する帯域を追加設定しながらマスクング閾値の補正を繰り返す。これにより、帯域の欠落による音質の劣化を抑制しつつ、マスクング閾値の過度の補正による音質劣化を抑制できる。

【0096】

上記のような符号化処理を行う本実施形態の符号化装置1は、例えばコンピュータと、コンピュータに上記の符号化処理を実行させるプログラムとにより実現可能である。以下、コンピュータとプログラムにより実現される符号化装置1について、図5を参照しながら説明する。

【0097】

図5は、符号化装置として動作させるコンピュータのハードウェア構成例を示す模式図である。

【0098】

図5に示すように、符号化装置として動作させるコンピュータ5は、Central Processing Unit(CPU)50と、主記憶装置52と、補助記憶装置54と、入力装置56と、出力装置58と、を備える。また、コンピュータ5は、Digital Signal Processor(DSP)60と、記憶媒体駆動装置62と、インタフェース装置64と、を更に備える。コンピュータ5におけるこれらの要素50~64は、バス68により相互に接続されており、要素間でのデータの受け渡しが可能になっている。

【0099】

CPU 50は、オペレーティングシステムを含む各種のプログラムを実行することによりコンピュータ5の全体の動作を制御する演算処理装置である。

【0100】

主記憶装置52は、Read Only Memory (ROM) 52a及びRandom Access Memory (RAM) 52bを有する。ROM 52aには、例えばコンピュータ5の起動時にプロセッサ50が読み出す所定の基本制御プログラム等が予め記録されている。また、RAM 52bは、プロセッサ50が各種のプログラムを実行する際に、必要に応じて作業用記憶領域として使用する。本実施形態においては、例えば符号化するオーディオ信号、マスキング閾値等の一時的な記憶にRAM 52bを使用する。

10

【0101】

補助記憶装置54は、Hard Disk Drive (HDD)やSolid State Disk (SSD)等の主記憶装置52に比べて大容量の記憶装置である。補助記憶装置54には、CPU 50によって実行される各種のプログラムや各種のデータ等を記憶させる。補助記憶装置54に記憶させるプログラムとしては、例えば、オーディオ信号の符号化や再生を行うオーディオプレイヤーのプログラムが挙げられる。また、補助記憶装置54に記憶させるデータとしては、例えば、上記のプレイヤーにより符号化されたオーディオ信号のデータが挙げられる。

【0102】

入力装置56は、例えばキーボード装置やマウス装置であり、コンピュータ5のオペレータにより操作されると、その操作内容に対応付けられている入力情報をCPU 50に送信する。

20

【0103】

出力装置58は、例えば液晶ディスプレイやスピーカである。液晶ディスプレイは、CPU 50等から送信される表示データに従って各種のテキスト、画像等を表示する。また、スピーカは、CPU 50やDSP 60等から送信される音声データやオーディオデータを出力する。

【0104】

DSP 60は、CPU 50からの制御信号等に従ってオーディオ信号の符号化処理、復号化(再生)処理等を行う演算処理装置である。

30

【0105】

記憶媒体駆動装置64は、図示しない可搬型記憶媒体に記録されているプログラムやデータの読み出し、補助記憶装置54に記憶されたデータ等の可搬型記憶媒体への書き込みを行う。可搬型記憶媒体としては、例えば、USB規格のコネクタが備えられているフラッシュメモリが利用可能である。また、可搬型記憶媒体としては、Compact Disk (CD)、Digital Versatile Disc (DVD)、Blu-ray Disc (Blu-rayは登録商標)等の光ディスクも利用可能である。

【0106】

インタフェース装置64は、例えばオーディオ入出力装置や通信制御装置である。オーディオ入出力装置は、例えばコンピュータ5とマイクやオーディオ装置とを接続してオーディオ信号の入力や出力を行う。通信制御装置は、コンピュータ5とインターネット等の通信ネットワークとを通信可能に接続し、通信ネットワークを介した外部通信装置等との通信によりオーディオデータ等の送受信を行う。

40

【0107】

このコンピュータ5は、CPU 50が補助記憶装置54から上述した符号化処理を含むプログラムを読み出し、DSP 60、主記憶装置52、補助記憶装置54等と協働してオーディオ信号の符号化処理を実行する。この際、CPU 50は、符号化処理における演算処理をDSP 60に実行させる。DSP 60は、オーディオ信号を周波数スペクトルに変換し、初期マスキング閾値を生成する。オーディオ信号は、例えば音楽CD等の可搬型記憶媒体から読み出して入力してもよいし、インタフェース装置64を介した通信

50

でコンピュータ 5 に入力してもよい。また、DSP 60 は、初期 PE 値及び目標 PE 値を算出し、それらの大小関係から初期マスキング閾値を用いてオーディオ信号を量子化することができるか否かを判定する。そして、初期マスキング閾値を用いて量子化することができない場合、DSP 60 は、聴覚特性を算出し、マスキング閾値を補正する。更に、DSP 60 は、補正されたマスキング閾値の採否を判定し、不採用の場合には欠落を許容する帯域を設定してマスキング閾値の補正を再度行う。そして、補正されたマスキング閾値を採用すると判定すると、採用したマスキング閾値を用いてオーディオ信号（周波数スペクトル）を量子化、符号化する。また、DSP 60 は、上記の処理の実行中、初期マスキング閾値、目標 PE 値、補正後のマスキング閾値の上限値等を RAM 52 b や補助記憶装置 54 に記憶させる処理、及び RAM 52 b や補助記憶装置 54 から読み出す処理を行う。

10

【0108】

コンピュータ 5 で符号化したオーディオ信号のデータ（オーディオデータ）は、例えば、補助記憶装置 54 に記憶させておき、必要に応じてコンピュータ 5 で復号化（再生）する。また、インタフェース装置 64 として通信制御装置を備えたコンピュータ 5 であれば、例えば、オーディオデータを、通信ネットワークを介して他のコンピュータ等に提供（配信）することができる。

【0109】

なお、符号化装置 1 として用いるコンピュータ 5 は、図 5 に示した構成に限らず、CPU 50 においてオーディオ信号の符号化を行う構成であってもよい。また、符号化装置 1 として用いるコンピュータ 5 は、種々のプログラムを実行することにより複数の機能を実現する汎用型のものに限らず、オーディオ信号の符号化、復号化に特化されたオーディオ装置でもよい。

20

【0110】

以上説明したように、第 1 の実施形態に係る符号化装置 1 を用いた符号化方法では、聴覚心理モデルに基づいて生成した初期マスキング閾値を用いてオーディオ信号（周波数スペクトル）を量子化できない場合、マスキング閾値を補正する。その際、1 回目の補正は全ての帯域が量子化により欠落しない条件で行い、補正前後のマスキング閾値の概形の類似度が基準値を満たしているか判定する。そして、概形の類似度が基準値を満たしていない場合、補正前後のマスキング閾値の概形の類似度が基準値を満たすようになるまで、重要度が低い帯域から順に欠落を許容しながらマスキング閾値の補正を繰り返す。そのため、量子化でいくつかの帯域が欠落することによる音質劣化を抑制しつつ、量子化に用いるマスキング閾値と初期マスキング閾値との概形の類似度のずれ（すなわちマスキング閾値の過度の補正）による音質劣化を抑制することができる。

30

【0111】

なお、欠落許容帯域設定処理は、例えば、1 度の設定処理において、相互相関値 $corre$ と相関閾値 TH_1 との差の大きさに応じて重要度が最も低い帯域から順に複数の帯域の欠落を許容するようにしてもよい。

【0112】

また、補正後のマスキング閾値の採否の判定は、式（4）で表される相互相関値 $corre$ に限らず、補正前後のマスキング閾値の概形の類似度と対応する値を用いて行えばよい。

40

【0113】

また、ステップ S20 で算出する聴覚特性は、信号対マスク比に限らず、他の特性であってもよい。

【0114】

更に、符号化装置 1 は、図 1 に示したようなオーディオ信号の符号化のみを行う装置に限らず、映像信号の符号化を行う装置であってもよい。映像信号の符号化を行う装置では、図 1 に示した構成に加え、動画像の符号化を行う構成を備える。このような装置では、入力された映像信号の符号化を動画像の符号化と音声の符号化とに分けて行った後、符号化した動画像と音声とを多重化する。

50

【 0 1 1 5 】

[第 1 の実施形態の変形例]

図 2 B に示したマスキング閾値の補正処理においては、補正後のマスキング閾値 $sfbThr$ (sfb) と初期マスキング閾値 $sfbThr_0(sfb)$ との相互相関値 $corre$ が相関閾値 TH_1 より大きくなるまでマスキング閾値の補正を繰り返す。しかしながら、第 1 の実施形態に係る符号化処理においては、これに限らず、例えばマスキング閾値の補正回数に上限を設けてもよい。

【 0 1 1 6 】

図 6 は、第 1 の実施形態に係るマスキング閾値の補正処理の変形例を示すフローチャートである。なお、図 6 に示した補正処理は、図 2 A に示したステップ S 1 8 において初期 P E 値が目標 P E 値よりも大きいと判定された場合 (ステップ S 1 8 ; Y e s) に行われる。

10

【 0 1 1 7 】

初期 P E 値が目標 P E 値よりも大きい場合、符号化装置 1 は、図 6 に示すように、続けて、周波数スペクトル等に基づく聴覚特性の算出 (ステップ S 2 0)、及び補正後のマスキング閾値 $sfbThr(sfb)$ の上限値の設定 (ステップ S 2 2) を行う。その後、符号化装置 1 (マスキング閾値補正部 1 8) は、補正回数 M を $M = 1$ に初期化する (ステップ S 2 3)。

【 0 1 1 8 】

ステップ S 2 3 の後、符号化装置 1 は、マスキング閾値の補正 (ステップ S 2 4)、及び相互相関値 $corre$ の算出 (ステップ S 2 6) を行い、相互相関値 $corre$ が相関閾値 TH_1 より大きいか判定する (ステップ S 2 8)。そして、相互相関値 $corre$ が相関閾値 TH_1 より大きい場合 (ステップ S 2 8 ; Y e s)、図 2 C に示したように直前の補正で得られたマスキング閾値 $sfbThr(sfb)$ を用いて周波数スペクトルを量子化する (ステップ S 3 4)。

20

【 0 1 1 9 】

一方、相互相関値 $corre$ が相関閾値 TH_1 以下の場合 (ステップ S 2 8 ; N o)、図 6 に示したように、補正回数 M が閾値 THM 以上であるか判定する (ステップ S 2 9)。 $M \geq THM$ の場合 (ステップ S 2 9 ; Y e s)、マスキング閾値の補正処理を終了し、直前の補正で得られたマスキング閾値 $sfbThr(sfb)$ を用いて周波数スペクトルを量子化する (ステップ S 3 4)。また、 $M < THM$ の場合 (ステップ S 2 9 ; N o)、欠落許容帯域設定処理 (ステップ S 3 0) を行った後、補正回数 M を 1 だけ増加させ (ステップ S 3 1)、ステップ S 2 4 の補正処理に戻る。補正回数の閾値 THM は、適宜設定すればよく、例えば 5 ~ 1 0 程度にする。

30

【 0 1 2 0 】

図 2 B に示したマスキング閾値の補正処理では、補正回数が多くなると欠落を許容する帯域も多くなる。そのため、オーディオ信号の周波数スペクトルによっては、補正回数が多くなり重要度が高い帯域の欠落が許容され、量子化により欠落してしまう恐れがある。

【 0 1 2 1 】

これに対し、図 6 に示したように補正回数に上限を設けた場合、重要度が高い帯域の欠落を抑制することができる。また、図 6 に示したように補正回数に上限を設けた場合、マスキング閾値の補正を繰り返すことによる符号化処理の遅延等を抑制することができ、AAC-Enhanced Low Delay (AAC - E L D) 等の低遅延符号化への適用も容易になる。

40

【 0 1 2 2 】

なお、補正回数の閾値 THM は、特定の回数に固定してもよいし、例えば、重要度 (信号対マスク比) が所定の値よりも小さい帯域の数にする等、周波数スペクトルのパターンに応じて都度変えてもよい。

【 0 1 2 3 】

[第 2 の実施形態]

本実施形態は、本発明に係る符号化装置におけるマスキング閾値の補正処理を効率よく

50

行えるようにしたものである。本実施形態のマスクング閾値の補正処理は、第 1 の実施形態と同様、図 2 A に示したステップ S 1 8 の判定において初期 P E 値が目標 P E 値より大きい場合（ステップ S 1 8 ; Y e s ）に行われる。

【 0 1 2 4 】

なお、本実施形態で説明するマスクング閾値の補正処理を行う符号化装置は、図 1 に示した符号化装置 1 におけるマスクング閾値生成部 1 4、及び再補正制御部 2 0 で行う処理の一部を下記のように変更したものでよい。

【 0 1 2 5 】

本実施形態に係る符号化装置 1 におけるマスクング閾値生成部 1 4 は、初期 P E 値が目標 P E 値よりも大きい場合、初期マスクング閾値 $sfbThr_0(sfb)$ 及び目標 P E 値、並びに周波数スペクトルの電力値 $mdct_pow(sfb)$ を、採否判定部 2 0 a に渡す。また、マスクング閾値生成部 1 4 は、初期マスクング閾値 $sfbThr_0(sfb)$ 及び目標 P E 値を聴覚特性算出部 1 6 に渡す。更に、マスクング閾値生成部 1 4 は、初期 P E 値、又は目標 P E 値と初期 P E 値との差分値を再補正制御部 2 0 の記憶部 2 0 c に記憶させる。

10

【 0 1 2 6 】

一方、本実施形態に係る符号化装置 1 における再補正制御部 2 0 は、図 7 及び図 8 に示すような処理を行う。

【 0 1 2 7 】

図 7 は、本発明の第 2 の実施形態に係るマスクング閾値の補正処理を示すフローチャートである。

20

【 0 1 2 8 】

本実施形態に係るマスクング閾値の補正処理は、上記のように、初期 P E 値が目標 P E 値より大きい場合（ステップ S 1 8 ; Y e s ）に行われる。初期 P E 値が目標 P E 値より大きい場合、第 1 の実施形態で説明したとおり、初期マスクング閾値を用いて周波数スペクトルを量子化することができない。そのため、初期 P E 値が目標 P E 値より大きい場合、符号化装置 1 は、図 7 に示した手順でマスクング閾値を補正する。

【 0 1 2 9 】

本実施形態においても、マスクング閾値を補正する際には、まず聴覚特性算出部 1 6 が周波数スペクトル等に基づき聴覚特性を算出する（ステップ S 2 0 ）。聴覚特性算出部 1 6 は、聴覚特性として、周波数スペクトルの電力値 $mdct_pow(sfb)$ と初期マスクング閾値 $sfbThr_0(sfb)$ との差分値（信号対マスク比）を算出する。

30

【 0 1 3 0 】

聴覚特性（信号対マスク比）の算出を終えると、次に、マスクング閾値補正部 1 8 が、補正後のマスクング閾値 $sfbThr(sfb)$ の上限値を、欠落を許容する値に設定する（ステップ S 2 1 ）。そして、閾値の上限値の設定を終えると、マスクング閾値補正部 1 8 は、聴覚特性及び閾値の上限値に基づいてマスクング閾値を補正する（ステップ S 2 4 ）。すなわち、本実施形態では、符号化対象の全帯域が量子化による欠落を許容された条件下で、1 回目のマスクング閾値の補正を行う。なお、マスクング閾値は、第 1 の実施形態と同様、式 (3) を用い、電力値 $mdct_pow(sfb)$ と補正後のマスクング閾値 $sfbThr(sfb)$ とに基づいて算出される P E 値が目標 P E 値になるように補正する。

40

【 0 1 3 1 】

マスクング閾値の補正を終えると、次に、再補正制御部 2 0 の採否判定部 2 0 a が、補正後のマスクング閾値 $sfbThr(sfb)$ と電力値 $mdct_pow(sfb)$ とを比較する。そして、採否判定部 2 0 a は、符号化対象の全ての帯域で $sfbThr(sfb) < mdct_pow(sfb)$ であるかを判定する（ステップ S 2 7 ）。符号化対象の全ての帯域で $sfbThr(sfb) < mdct_pow(sfb)$ である場合（ステップ S 2 7 ; Y e s ）、量子化により欠落する帯域はないので、採否判定部 2 0 a は補正後のマスクング閾値 $sfbThr(sfb)$ を採用すると判定する。この場合、符号化装置 1 は、図 2 C に示したように、直前の補正により得られたマスクング閾値 $sfbThr(sfb)$ を用いて周波数スペクトルを量子化する（ステップ S 3 4 ）。

【 0 1 3 2 】

50

一方、 $sfbThr(sfb) \leq mdct_pow(sfb)$ の帯域がある場合（ステップS27；No）、その帯域は、判定に係る補正後のマスキング閾値 $sfbThr(sfb)$ を用いて周波数スペクトルを量子化すると欠落する。1回目のマスキング閾値の補正は、上記のように全ての帯域が欠落を許容された条件下で行っている。そのため、 $sfbThr(sfb) \leq mdct_pow(sfb)$ である帯域は、音質についての重要度が高い場合もありうる。重要度の高い帯域が量子化により欠落すると音質が著しく劣化する。よって、 $sfbThr(sfb) \leq mdct_pow(sfb)$ の帯域がある場合、欠落許容帯域設定処理（ステップS30）を行い、重要度が低い帯域の欠落のみを許容するようマスキング閾値の上限値の設定を変更してマスキング閾値を補正しなおす。欠落許容帯域設定処理は、第1の実施形態と同様、欠落許容帯域設定部20bが行う。

【0133】

10

図8は、第2の実施形態における欠落許容帯域設定処理の内容を示すフローチャートである。図9は、欠落を許容する帯域の割合の決定方法を説明するグラフである。

【0134】

欠落許容帯域設定部20bは、本実施形態の欠落許容帯域設定処理（ステップS30）として、図8に示すように、まず、目標PE値と初期PE値との差分値に基づいて、全帯域に対する欠落を許容する帯域の割合を決定する（ステップS3020）。ステップS3020では、例えば記憶部20cから初期PE値及び目標PE値を読み出して差分値 $diffPe$ を算出する。その後、例えば、図9に示す関数 $f(diffPe)$ に基づいて、欠落を許容する帯域の割合 $lack_ratio$ を決定する。関数 $f(diffPe)$ は、差分値 $diffPe$ をパラメータとする関数であり、下記式(5)で表される。

20

【0135】

【数5】

$$diffPe = desirePe - Pe$$

$$lack_ratio = f(diffPe)$$

$$f(diffPe) = \begin{cases} 0 & (0 \leq diffPe < TH_2, TH_3 \leq diffPe) \\ \frac{TH_4}{TH_3 - TH_2} (diffPe - TH_2) & (TH_2 \leq diffPe < TH_3) \end{cases} \quad (5)$$

30

【0136】

図9及び式(5)における TH_2 、 TH_3 、及び TH_4 は、いずれも任意の値であり、ビットレートや許容する音質劣化の度合い等に基づいて適宜設定すればよい。

【0137】

欠落を許容する帯域の割合を設定すると、欠落許容帯域設定部20bは、次に、記憶部20cから符号化対象の各帯域の信号対マスク比及びマスキング閾値の上限値を読み出す（ステップS3022）。続けて、欠落許容帯域設定部20bは、信号対マスク比（重要度）が最も低い帯域から順に、ステップS3020で決定した割合に応じた順位の帯域まで欠落を許容する（ステップS3024）。

40

【0138】

その後、欠落許容帯域設定部20bは、欠落を許容する帯域以外の帯域についてのマスキング閾値の上限値を、欠落を許容しない値に変更し、記憶部20cのデータを更新する（ステップS3026）。欠落を許容しないマスキング閾値の上限値は、第1の実施形態で説明したように、 $sfbThr(sfb) < mdct_pow(sfb)$ を満たす正の値である。

【0139】

ステップS3026の処理を終えると、欠落許容帯域設定部20bは、欠落許容帯域設定処理を終了する（リターン）。欠落許容帯域設定処理が終了すると、図7に示したステップS24に戻り、初期マスキング閾値に対する補正が再度行われる。

50

【 0 1 4 0 】

上記のように、本実施形態では、初期マスキング閾値に対する 1 回目の補正を全ての帯域の欠落を許容した条件下で行う。帯域の欠落を許容する条件下では補正後のマスキング閾値に上限値がないので、上述したような過剰な補正を防ぐことができ、補正前後のマスキング閾値の概形の類似度の低下を抑制できる。そのため、1 回目の補正で得られたマスキング閾値が全ての帯域で $\text{sfbThr}(\text{sfb}) < \text{mdct_pow}(\text{sfb})$ であれば、補正前後のマスキング閾値の概形の類似度は、音質が著しく劣化しない程度に高いといえる。したがって、1 回目の補正で得られたマスキング閾値が採用された場合、帯域が欠落することによる音質劣化がない上、マスキング閾値の概形のずれ（過度の補正）による音質劣化も抑えられる。

10

【 0 1 4 1 】

また、本実施形態では、1 回目の補正で得られたマスキング閾値に $\text{sfbThr}(\text{sfb}) > \text{mdct_pow}(\text{sfb})$ の帯域がある場合、 $\text{sfbThr}(\text{sfb}) < \text{mdct_pow}(\text{sfb})$ の帯域ではなく、音質における重要度の低い帯域の欠落を許容してマスキング閾値を再度補正する。これにより、帯域が欠落することによる音質の劣化を抑制しつつ、マスキング閾値の概形のずれ（過度の補正）による音質劣化を抑制する。

【 0 1 4 2 】

また、本実施形態では、目標 P E 値と初期 P E 値との差分値 diffPe に基づいて欠落を許容する帯域を設定してマスキング閾値を補正しなおす。

20

【 0 1 4 3 】

オーディオ信号の符号化においては、目標 P E 値と初期 P E 値との差分値 diffPe が大きくなると符号化時のビット数の不足量が増大する。そのため、差分値 diffPe が大きい場合にはマスキング閾値の補正量を多くしなければならない。ところが、帯域の欠落を許容しない条件下では、各帯域のマスキング閾値を上限値（電力値 $\text{mdct_pow}(\text{sfb})$ ）より大きくすることができない。すなわち、式（3）を用いた補正処理の過程においてマスキング閾値が上限値に達した帯域については、マスキング閾値を更に大きくして量子化に使用されるビット量を低減することができない。したがって、マスキング閾値が上限値に達した帯域があるにもかかわらずビット数が不足している場合、マスキング閾値が上限値に達していない帯域のマスキング閾値を大きくして符号化に使用するビット量を低減することとなる。よって、マスキング閾値の補正量が多くなると、欠落を許容しない条件下では補正後のマスキング閾値の概形と初期マスキング閾値との概形の類似度が低くなる。このように、帯域の欠落を許容しない条件で補正されたマスキング閾値の概形と初期マスキング閾値との概形の類似度は、目標 P E 値と初期 P E 値との差分値 diffPe の大きさから間接的に把握することができる。すなわち、帯域の欠落を許容しない条件でマスキング閾値を補正する処理、及び式（4）を用いて相互相関値 corre を算出する処理を行わなくても、帯域の欠落を許容しない条件下における補正前後のマスキング閾値の概形の類似度を把握することができる。

30

【 0 1 4 4 】

第 1 の実施形態からわかるように、マスキング閾値の概形の類似度と所定の基準値との差が大きいほど、類似度を基準値より大きくするために欠落を許容する帯域の数が多くなる。そのため、本実施形態では、図 9 に示した関数 $f(\text{diffPe})$ のように、目標 P E 値と初期 P E 値との差分値 diffPe がある閾値 TH_2 より大きくなると差分値 diffPe に比例して欠落を許容する帯域の割合（数）が増すようにしている。ただし、欠落を許容する帯域の割合（数）を大きくすると、補正後のマスキング閾値が上限値を超えてしまい量子化により欠落する帯域が多くなる恐れがある。そのため、差分値 diffPe が別の閾値 TH_3 ($> \text{TH}_2$) よりも大きい場合には、欠落を許容する帯域の割合を閾値 TH_4 に制限する。

40

【 0 1 4 5 】

このように、本実施形態では、マスキング閾値を補正しなおす際に、目標 P E 値と初期 P E 値との差分値に基づいて、補正後のマスキング閾値と初期マスキング閾値との概形の類似度が基準値を満たすよう欠落を許容する帯域の割合を設定する。そのため、帯域が欠

50

落することによる音質劣化を抑制しつつ、マスキング閾値の概形のずれ（過度の補正）による音質劣化を抑制することができる。

【0146】

また、本実施形態で欠落を許容する帯域を設定する際に用いる目標PE値と初期PE値との差分値diffPeや各帯域の重要度（信号対マスク比）等の計算量は、式（4）を用いて算出する相互相関値correの計算量よりも少ない。加えて、目標PE値、初期PE値、及び信号対マスク比は、欠落許容帯域設定処理（ステップS30）を行う前に、マスキング閾値生成部14や聴覚特性算出部16で算出している。そのため、本実施形態に係るマスキング閾値の補正処理は、マスキング閾値の相互相関値correを算出して概形の類似度を判定する第1の実施形態に比べて計算量を低減することができる。よって、本実施形態によれば、マスキング閾値の補正処理を効率よく行うことができる。

10

【0147】

更に、本実施形態に係るマスキング閾値の補正処理では、目標PE値と初期PE値との差分値diffPeが大きい場合、1度の欠落許容帯域設定処理において重要度が異なる2つ以上の帯域の欠落を許容することができる。そのため、マスキング閾値を補正しなおす際の補正処理を一層効率よく行うことができ、AAC-Enhanced Low Delay（AAC-ELD）等の低遅延符号化への適用が一層容易になる。

【0148】

なお、本実施形態では、図9及び式（5）で表される関数 $f(\text{diffPe})$ に基づいて、マスキング閾値を補正しなおす際の欠落を許容する帯域の割合lack_ratioを決定している。しかしながら、関数 $f(\text{diffPe})$ は、これに限らず、例えば任意のシグモイド関数を用いてもよい。

20

【0149】

[第3の実施形態]

本実施形態は、第2の実施形態とは異なる値に基づいて欠落を許容する帯域を設定することで、本発明に係る符号化装置におけるマスキング閾値の補正処理を効率よく行えるようにしたものである。本実施形態のマスキング閾値の補正処理は、第1の実施形態と同様、図2Aに示したステップS18の判定において初期PE値が目標PE値より大きい場合（ステップS18；Yes）に行われる。また、本実施形態に係るマスキング閾値の補正処理は、図7に示した手順で行う。すなわち、本実施形態のマスキング閾値の補正処理では、初期マスキング閾値に対する1回目の補正を、全ての帯域の欠落を許容した条件下で行う（ステップS21、S24）。そして、周波数スペクトルの電力値mdct_pow(sfb)と補正後のマスキング閾値sfbThr(sfb)との関係が $\text{sfbThr}(\text{sfb}) \leq \text{mdct_pow}(\text{sfb})$ となる帯域があった場合（ステップS27；Yes）、マスキング閾値を補正しなおす。また、マスキング閾値を補正しなおす場合、欠落許容帯域設定処理（ステップS30）を行って、量子化による欠落を許容する帯域を設定する。

30

【0150】

なお、本実施形態で説明するマスキング閾値の補正処理を行う符号化装置は、図1に示した符号化装置1におけるマスキング閾値生成部14、及び再補正制御部20で行う処理の一部を下記のように変更したものでよい。

40

【0151】

本実施形態に係る符号化装置1におけるマスキング閾値生成部14は、初期PE値が目標PE値よりも大きい場合、初期マスキング閾値 $\text{sfbThr}_0(\text{sfb})$ 及び目標PE値、並びに周波数スペクトルの電力値mdct_pow(sfb)を、採否判定部20aに渡す。また、マスキング閾値生成部14は、初期マスキング閾値 $\text{sfbThr}_0(\text{sfb})$ 及び目標PE値を聴覚特性算出部16に渡す。

【0152】

また、本実施形態に係る符号化装置1における再補正制御部20は、上記のように図7に示したマスキング閾値の補正処理を行うが、欠落許容帯域設定処理（ステップS30）として、図10に示すような処理を行う。

50

【 0 1 5 3 】

図 1 0 は、本発明の第 3 の実施形態に係る欠落許容帯域設定処理の内容を示すフローチャートである。図 1 1 は、欠落を許容する帯域の割合の決定方法を説明するグラフである。

【 0 1 5 4 】

図 1 0 に示した欠落許容帯域設定処理は、再補正制御部 2 0 の欠落許容帯域設定部 2 0 b が行う。欠落許容帯域設定部 2 0 b は、まず、符号化対象の帯域の数に対する $\text{sfbThr}(\text{sfb}) \cdot \text{mdct_pow}(\text{sfb})$ の帯域の数の割合に基づいて、欠落を許容する帯域の割合を決定する (ステップ S 3 0 2 1)。ステップ S 3 0 2 1 では、例えば、図 1 1 に示す関数 $f(\text{sat_ratio})$ に基づいて、欠落を許容する帯域の割合 lack_ratio を決定する。関数 $f(\text{sat_ratio})$ は、符号化対象の帯域の数に対する $\text{sfbThr}(\text{sfb}) \cdot \text{mdct_pow}(\text{sfb})$ の帯域の数の割合をパラメータとする関数であり、下記式 (6) で表される。

【 0 1 5 5 】

【 数 6 】

$$\left. \begin{aligned} \text{sat_ratio} &= \frac{\text{sat_sfb_num}}{\text{encode_sfb_num}} \\ \text{lack_ratio} &= f(\text{sat_ratio}) \\ f(\text{sat_ratio}) &= \begin{cases} 0 & (0 \leq \text{sat_ratio} < TH_5, TH_6 \leq \text{sat_ratio} \leq 1) \\ \frac{TH_7}{TH_6 - TH_5} (\text{sat_ratio} - TH_5) & (TH_5 \leq \text{sat_ratio} < TH_6) \end{cases} \end{aligned} \right\} (6)$$

【 0 1 5 6 】

式 (6) において、 encode_sfb_num は符号化対象の帯域の数である。また、 sat_sfb_num は、 $\text{sfbThr}(\text{sfb}) \cdot \text{mdct_pow}(\text{sfb})$ の帯域の数である。

【 0 1 5 7 】

図 1 1 及び式 (6) における TH_5 、 TH_6 、及び TH_7 は、いずれも任意の値であり、ビットレートや許容する音質劣化の度合い等に基づいて適宜設定すればよい。

【 0 1 5 8 】

欠落を許容する帯域の割合を設定すると、欠落許容帯域設定部 2 0 b は、次に、記憶部 2 0 c から符号化対象の各帯域の信号対マスク比及びマスキング閾値の上限値を読み出す (ステップ S 3 0 2 2)。続けて、欠落許容帯域設定部 2 0 b は、信号対マスク比 (重要度) が最も低い帯域から順に、ステップ S 3 0 2 1 で決定した割合に応じた順位の帯域まで欠落を許容する (ステップ S 3 0 2 4)。

【 0 1 5 9 】

その後、欠落許容帯域設定部 2 0 b は、欠落を許容する帯域以外の帯域についてのマスキング閾値の上限値を、欠落を許容しない値に変更し、記憶部 2 0 c のデータを更新する (ステップ S 3 0 2 6)。欠落を許容しないマスキング閾値の上限値は、第 1 の実施形態で説明したように、 $\text{sfbThr}(\text{sfb}) < \text{mdct_pow}(\text{sfb})$ を満たす正の値である。

【 0 1 6 0 】

ステップ S 3 0 2 6 の処理を終えると、欠落許容帯域設定部 2 0 b は、欠落許容帯域設定処理を終了する (リターン)。欠落許容帯域設定処理が終了すると、図 7 に示したステップ S 2 4 に戻り、初期マスキング閾値に対する補正が再度行われる。

【 0 1 6 1 】

上記のように、本実施形態では、初期マスキング閾値に対する 1 回目の補正を全ての帯域の欠落を許容した条件下で行う。帯域の欠落を許容する条件下でのマスキング閾値の補

正処理においては、補正前後のマスクング閾値の概形の類似度の低下が抑制される。そのため、1回目の補正で得られたマスクング閾値が全ての帯域で $\text{sfbThr}(\text{sfb}) < \text{mdct_pow}(\text{sfb})$ であれば、補正前後のマスクング閾値の概形の類似度は、音質が著しく劣化しない程度に高いといえる。したがって、1回目の補正で得られたマスクング閾値が採用された場合、帯域が欠落することによる音質劣化がない上、マスクング閾値の概形のずれ（過度の補正）による音質劣化も抑えられる。

【0162】

また、本実施形態においても、補正されたマスクング閾値に $\text{sfbThr}(\text{sfb}) < \text{mdct_pow}(\text{sfb})$ の帯域がある場合、欠落を許容する帯域を設定してマスクング閾値を補正しなおす。この際、欠落を許容する帯域は、符号化対象の帯域の数に対する $\text{sfbThr}(\text{sfb}) < \text{mdct_pow}(\text{sfb})$ の帯域の数の割合に基づいて、重要度の低い帯域から順に欠落を許容するよう設定する。これにより、帯域が欠落することによる音質の劣化を抑制しつつ、マスクング閾値の概形のずれ（過度の補正）による音質劣化を抑制する。

10

【0163】

量子化に用いるマスクング閾値に $\text{sfbThr}(\text{sfb}) < \text{mdct_pow}(\text{sfb})$ の帯域がある場合、その帯域は量子化により欠落する。そのため、量子化による帯域の欠落を抑制するには、 $\text{sfbThr}(\text{sfb}) < \text{mdct_pow}(\text{sfb})$ の帯域のマスクング閾値が $\text{sfbThr}(\text{sfb}) < \text{mdct_pow}(\text{sfb})$ になるよう補正しなおさなければならない。すなわち、 $\text{sfbThr}(\text{sfb}) < \text{mdct_pow}(\text{sfb})$ の帯域のマスクング閾値が小さくなるよう、その帯域の量子化に使用されるビット量を増加しなければならない。しかしながら、周波数スペクトルの量子化に使用可能なビット数にはビットレートに応じた上限がある。したがって、 $\text{sfbThr}(\text{sfb}) < \text{mdct_pow}(\text{sfb})$ の帯域のマスクング閾値を小さくするには、 $\text{sfbThr}(\text{sfb}) < \text{mdct_pow}(\text{sfb})$ である他の帯域のマスクング閾値を大きくしなければならない。よって、欠落を許容する条件下で補正したマスクング閾値に $\text{sfbThr}(\text{sfb}) < \text{mdct_pow}(\text{sfb})$ となる帯域の数が多い場合、欠落を許容しない条件下での補正により得られるマスクング閾値の概形と初期マスクング閾値との概形の類似度が低くなる。すなわち、帯域の欠落を許容しない条件で補正したマスクング閾値と初期マスクング閾値との概形の類似度は、欠落を許容する条件で補正したマスクング閾値において $\text{sfbThr}(\text{sfb}) < \text{mdct_pow}(\text{sfb})$ となる帯域の数から間接的に把握することができる。

20

【0164】

第1の実施形態からわかるように、マスクング閾値の概形の類似度と所定の基準値との差が大きいほど、類似度を基準値より大きくするために欠落を許容する帯域の数が増える。そのため、本実施形態では、図11に示した関数 $f(\text{sat_ratio})$ のように、 $\text{sfbThr}(\text{sfb}) < \text{mdct_pow}(\text{sfb})$ となる帯域の割合 sat_ratio が閾値 TH_5 より大きくなると割合 sat_ratio に比例して欠落を許容する帯域の割合（数）が増すようにしている。ただし、欠落を許容する帯域の割合（数）を多くすると、補正後のマスクング閾値が上限値を超えてしまい量子化により欠落する帯域が多くなる恐れがある。そのため、符号化対象の帯域の数に対する $\text{sfbThr}(\text{sfb}) < \text{mdct_pow}(\text{sfb})$ の帯域の数の割合 sat_ratio が別の閾値 TH_6 ($> \text{TH}_5$) よりも大きい場合には、欠落を許容する帯域の割合を閾値 TH_7 に制限する。

30

【0165】

このように、本実施形態では、マスクング閾値を補正しなおす際に、符号化対象の帯域の数に対する $\text{sfbThr}(\text{sfb}) < \text{mdct_pow}(\text{sfb})$ の帯域の数の割合に基づいて、重要度の低い帯域から順に欠落を許容する。そして、補正後のマスクング閾値と初期マスクング閾値との概形の類似度が基準値を満たすようになるまで、欠落を許容する帯域を追加しながらマスクング閾値の補正を繰り返す。そのため、帯域が欠落することによる音質劣化を抑制しつつ、マスクング閾値の概形のずれ（過度の補正）による音質劣化を抑制することができる。

40

【0166】

また、本実施形態で欠落を許容する帯域を設定する際に用いる値は、欠落許容帯域設定処理（ステップS30）を行う前に、マスクング閾値生成部14、聴覚特性算出部16、

50

及びマスキング閾値補正部 18 で算出している。そのため、本実施形態に係るマスキング閾値の補正処理は、マスキング閾値の相互相関値 $corre$ を算出して概形の類似度を判定する第 1 の実施形態に比べて計算量を低減することができる。よって、本実施形態によれば、マスキング閾値の補正処理を効率よく行うことができる。

【0167】

更に、本実施形態に係るマスキング閾値の補正処理では、符号化対象の帯域の数に対する $sfbThr(sfb)$ $mdct_pow(sfb)$ の帯域の数の割合が大きい場合、1 度の欠落帯域設定処理において重要度が異なる 2 つ以上の帯域の欠落を許容することができる。そのため、マスキング閾値を補正しなす際の補正処理を一層効率よく行うことができ、AAC-ELD 等の低遅延符号化への適用が一層容易になる。

10

【0168】

なお、本実施形態では、図 11 及び式 (6) で表される関数 $f(sat_ratio)$ に基づいて、マスキング閾値を補正しなす際の欠落を許容する帯域の割合 $lack_ratio$ を決定している。しかしながら、関数 $f(sat_ratio)$ は、これに限らず、例えば任意のシグモイド関数を用いてもよい。

【0169】

以上記載した各実施例を含む実施形態に関し、更に以下の付記を開示する。

(付記 1)

オーディオ信号を周波数スペクトルに変換し、当該周波数スペクトルの量子化及び符号化を行う符号化装置において、

20

前記周波数スペクトルに基づいて量子化する際の初期マスキング閾値を生成する閾値生成部と、

前記周波数スペクトルの量子化に使用可能なビット量に基づいて前記初期マスキング閾値を補正する閾値補正部と、

前記閾値補正部で補正されたマスキング閾値と前記初期マスキング閾値との概形の類似度が基準値以下の場合に、前記周波数スペクトルの帯域毎に量子化による欠落を許容するか否かを設定し、前記閾値補正部に前記初期マスキング閾値を再度補正させる再補正制御部と、

を備えることを特徴とする符号化装置。

(付記 2)

30

前記再補正制御部は、前記初期マスキング閾値及び補正された前記マスキング閾値における欠落を許容しない帯域の概形の類似度が前記基準値以下の場合に、前記閾値補正部に前記初期マスキング閾値を再度補正させる、

ことを特徴とする付記 1 に記載の符号化装置。

(付記 3)

前記再補正制御部は、前記マスキング閾値の概形の類似度として前記初期マスキング閾値と前記閾値補正部で補正されたマスキング閾値との相互相関値を算出する、

ことを特徴とする付記 2 に記載の符号化装置。

(付記 4)

40

前記再補正制御部は、前記初期マスキング閾値と補正された前記マスキング閾値との概形の類似度に基づいて符号化対象の全帯域に対する欠落を許容する帯域の割合を決定し、決定した前記割合に基づいて前記周波数スペクトルの各帯域に対し欠落を許容するか否かを設定する設定部を有する、

ことを特徴とする付記 1 に記載の符号化装置。

(付記 5)

前記設定部は、前記マスキング閾値の概形の類似度として前記周波数スペクトル及び前記初期マスキング閾値に基づく第 1 の知覚エントロピー値と、前記周波数スペクトルの量子化に使用可能なビット数に基づく第 2 の知覚エントロピー値との差分値を用い、当該差分値に基づいて前記割合を決定する、

ことを特徴とする付記 4 に記載の符号化装置。

50

(付記 6)

前記設定部は、前記マスクング閾値の概形の類似度として符号化対象の帯域の数に対する、補正された前記マスクング閾値が前記周波数スペクトルの電力値よりも大きい帯域の数の割合を用い、当該割合に基づいて前記符号化対象の全帯域に対する欠落を許容する帯域の割合を決定する、

ことを特徴とする付記 4 に記載の符号化装置。

(付記 7)

前記再補正制御部は、前記周波数スペクトルの帯域のうち重要度の低い帯域から順に欠落を許容する、

ことを特徴とする付記 1 に記載の符号化装置。

10

(付記 8)

前記再補正制御部は、前記重要度として、各帯域における電力値と初期マスクング閾値との差分値を用いる、

ことを特徴とする付記 7 に記載の符号化装置。

(付記 9)

前記再補正制御部は、前記初期マスクング閾値に対する補正回数が所定の回数に達すると、補正されたマスクング閾値と前記初期マスクング閾値との概形の類似度によらず補正されたマスクング閾値を採用する、

ことを特徴とする付記 1 に記載の符号化装置。

(付記 10)

前記閾値生成部は、聴覚心理モデルに基づいて前記初期マスクング閾値を生成する、

ことを特徴とする付記 1 に記載の符号化装置。

20

(付記 11)

コンピュータが、

オーディオ信号から得た周波数スペクトルに基づいて生成した初期マスクング閾値が当該周波数スペクトルを量子化するための条件を満たしていない場合に、

前記周波数スペクトルの各帯域に対し量子化による欠落を許容するか否かを設定し、

前記周波数スペクトルの量子化に使用可能なビット量及び各帯域の欠落を許容するか否かの設定に基づいて前記初期マスクング閾値を補正し、

補正されたマスクング閾値と前記初期マスクング閾値との概形の類似度が基準値以下の場合、1つ以上の帯域の欠落を許容するか否かの設定を変更して初期マスクング閾値を再度補正する、

処理を実行することを特徴とする符号化方法。

30

(付記 12)

前記コンピュータが、前記概形の類似度として、補正された前記マスクング閾値と前記初期マスクング閾値との相互相関値を算出する、

処理を実行することを特徴とする付記 11 に記載の符号化方法。

(付記 13)

前記コンピュータが、

前記周波数スペクトル及び前記初期マスクング閾値に基づく第 1 の知覚エントロピー値と、前記周波数スペクトルの量子化に使用可能なビット数に基づく第 2 の知覚エントロピー値との差分値を算出し、

算出した前記差分値に基づいて符号化対象の帯域に対する欠落を許容する帯域の割合を決定する、

処理を実行することを特徴とする付記 11 に記載の符号化方法。

40

(付記 14)

前記コンピュータが、

符号化対象の帯域の数に対する、補正された前記マスクング閾値が前記周波数スペクトルの電力値よりも大きい帯域の数の割合に基づいて、符号化対象の帯域に対する欠落を許容する帯域の割合を決定する、

50

処理を実行することを特徴とする付記 1 1 に記載の符号化方法。

(付記 1 5)

補正された前記マスキング閾値が所定の採用条件を満たしていない場合、前記周波数スペクトルの各帯域のうち電力値と前記初期マスキング閾値との差分値が小さい帯域から順に欠落を許容する、

ことを特徴とする付記 1 1 に記載の符号化方法。

(付記 1 6)

オーディオ信号から得た周波数スペクトルに基づいて生成した初期マスキング閾値が当該周波数スペクトルを量子化するための条件を満たしていない場合に、

前記周波数スペクトルの各帯域に対し量子化による欠落を許容するか否かを設定し、

前記周波数スペクトルの量子化に使用可能なビット量及び各帯域の欠落を許容するか否かの設定に基づいて前記初期マスキング閾値を補正し、

補正されたマスキング閾値と前記初期マスキング閾値との概形の類似度が基準値以下の場合、1つ以上の帯域欠落を許容するか否かの設定を変更して初期マスキング閾値を再度補正する、

処理をコンピュータに実行させるためのプログラム。

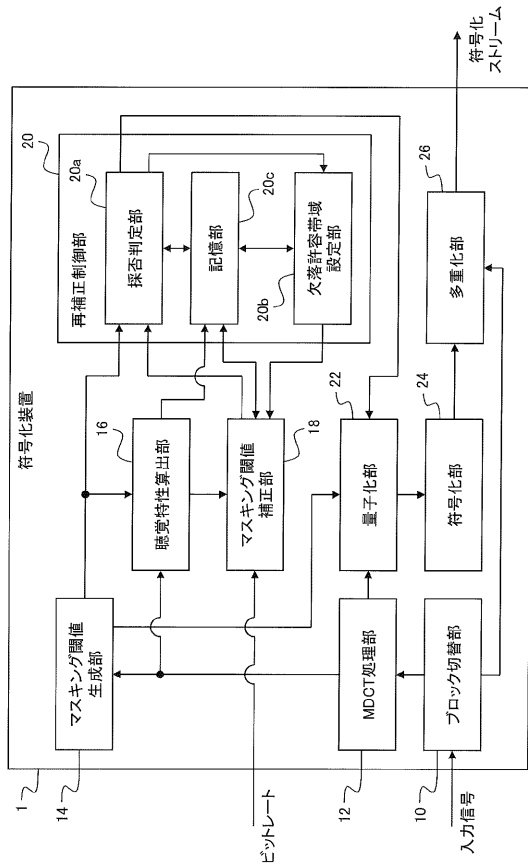
【符号の説明】

【0170】

| | | |
|-------|-------------|----|
| 1 | 符号化装置 | |
| 1 0 | ブロック切替部 | 20 |
| 1 2 | M D C T 処理部 | |
| 1 4 | マスキング閾値生成部 | |
| 1 6 | 聴覚特性算出部 | |
| 1 8 | マスキング閾値補正部 | |
| 2 0 | 再補正制御部 | |
| 2 0 a | 採否判定部 | |
| 2 0 b | 欠落許容帯域設定部 | |
| 2 0 c | 記憶部 | |
| 2 2 | 量子化部 | |
| 2 4 | 符号化部 | 30 |
| 2 6 | 多重化部 | |
| 5 | コンピュータ | |
| 5 0 | C P U | |
| 5 2 | 主記憶装置 | |
| 5 2 a | R O M | |
| 5 2 b | R A M | |
| 5 4 | 補助記憶装置 | |
| 5 6 | 入力装置 | |
| 5 8 | 出力装置 | |
| 6 0 | D S P | 40 |
| 6 2 | 記憶媒体駆動装置 | |
| 6 4 | インタフェース装置 | |

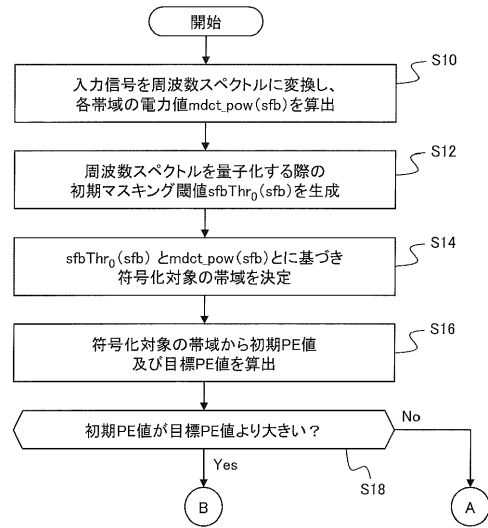
【 図 1 】

本発明の第1の実施形態に係る符号化装置の構成例を示すブロック図



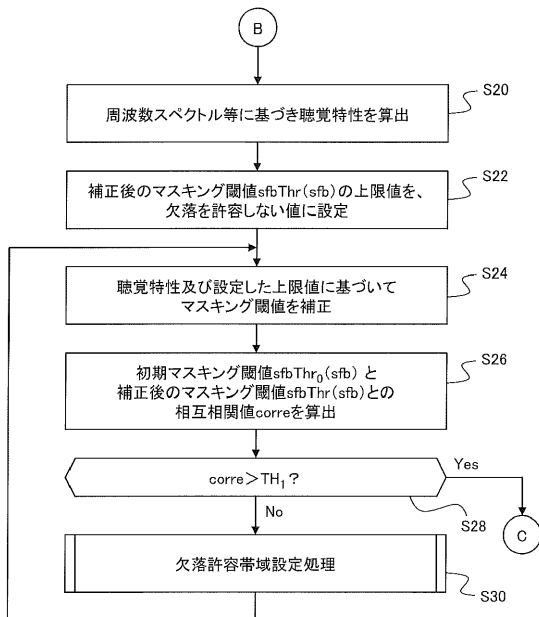
【 図 2 A 】

第1の実施形態に係る符号化装置における符号化処理を示すフローチャート(その1)



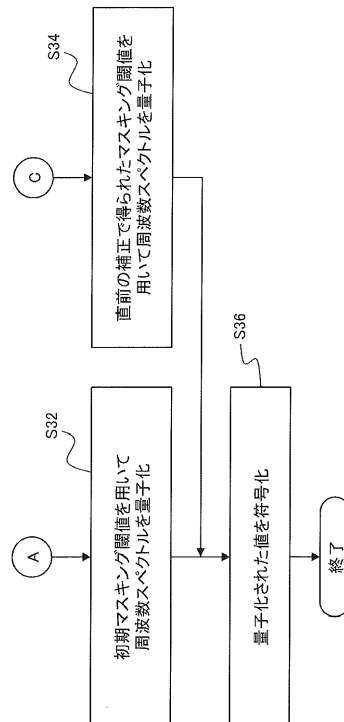
【 図 2 B 】

第1の実施形態に係る符号化装置における符号化処理を示すフローチャート(その2)



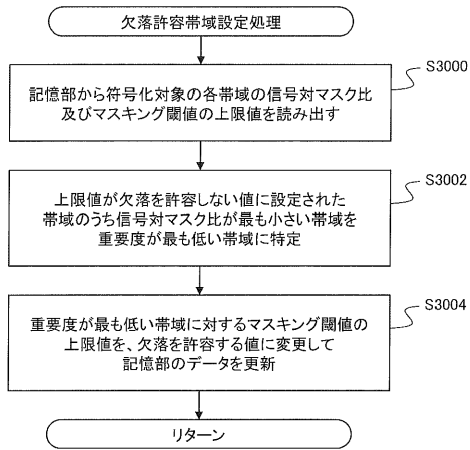
【 図 2 C 】

第1の実施形態に係る符号化装置における符号化処理を示すフローチャート(その3)



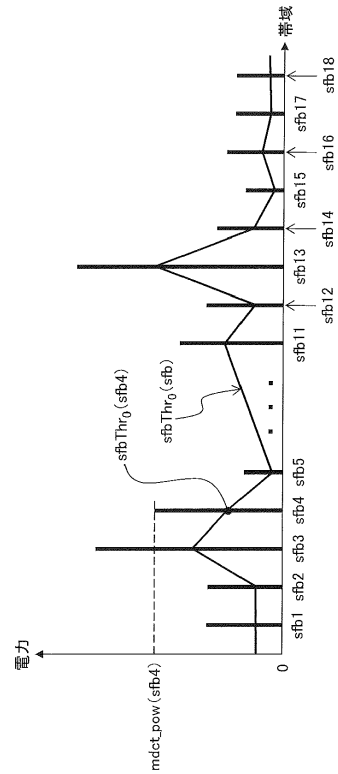
【 図 3 】

第1の実施形態における欠落許容帯域設定処理の内容を示すフローチャート



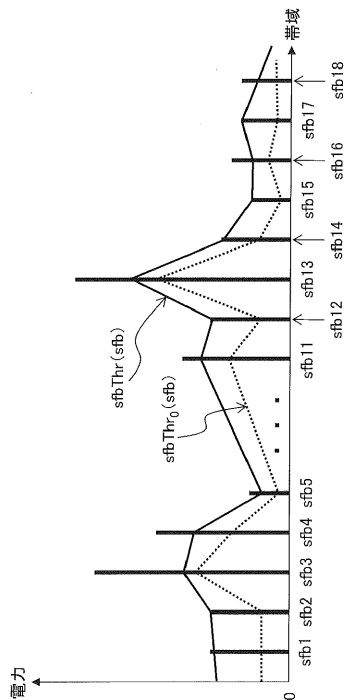
【 図 4 A 】

周波数スペクトル及び初期マスクング閾値の一例を示すグラフ



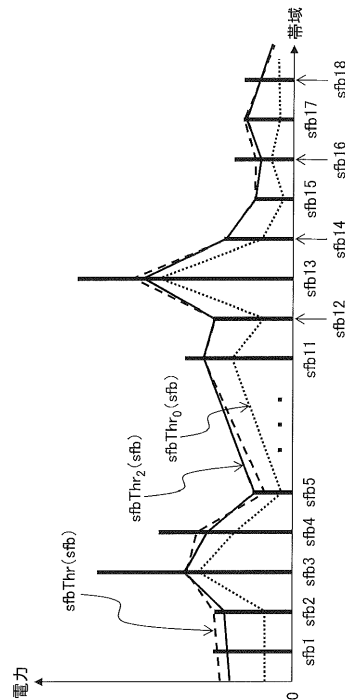
【 図 4 B 】

1回目の補正で得られたマスクング閾値と初期マスクング閾値との関係を説明するグラフ



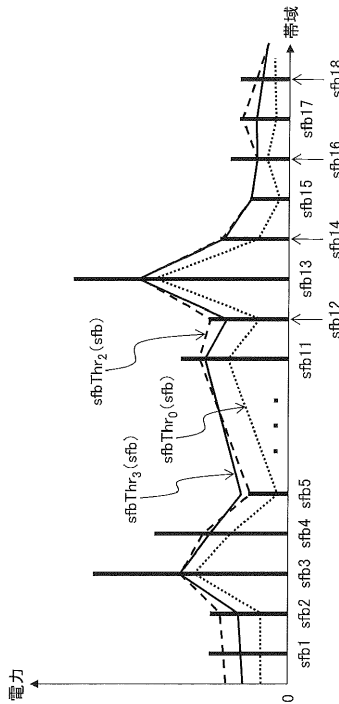
【 図 4 C 】

2回目の補正で得られたマスクング閾値と1回目の補正で得られたマスクング閾値との関係を説明するグラフ



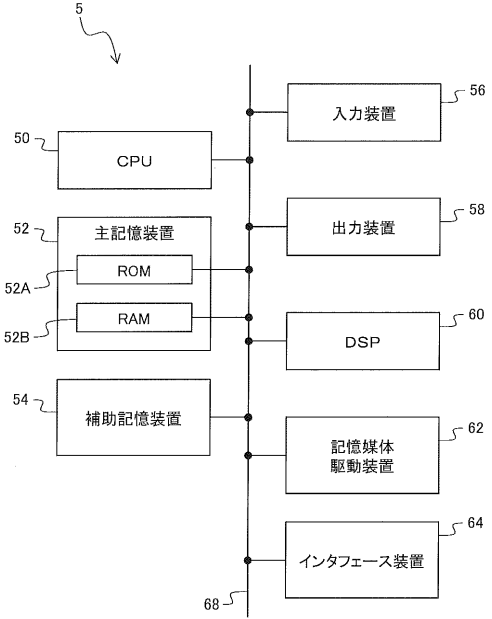
【図4D】

3回目の補正で得られたマスクング閾値と2回目の補正で得られたマスクング閾値との関係を説明するグラフ



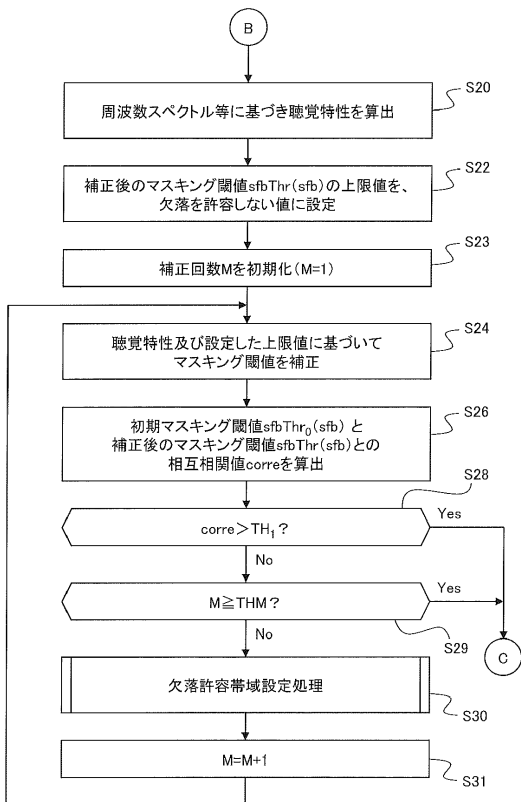
【図5】

符号化装置として動作させるコンピュータのハードウェア構成例を示す模式図



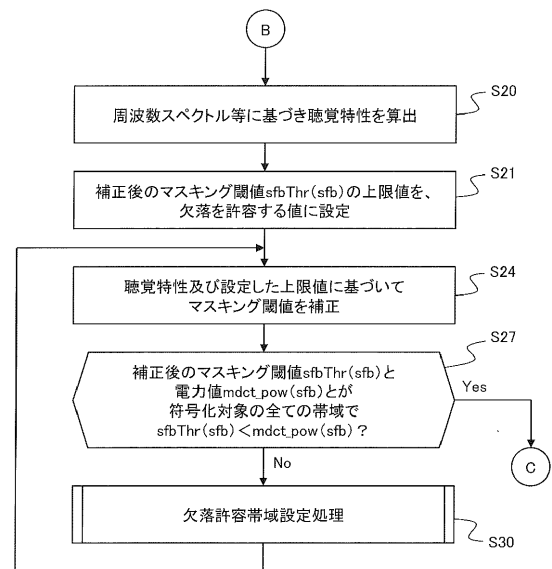
【図6】

第1の実施形態に係るマスクング閾値の補正処理の変形例を示すフローチャート



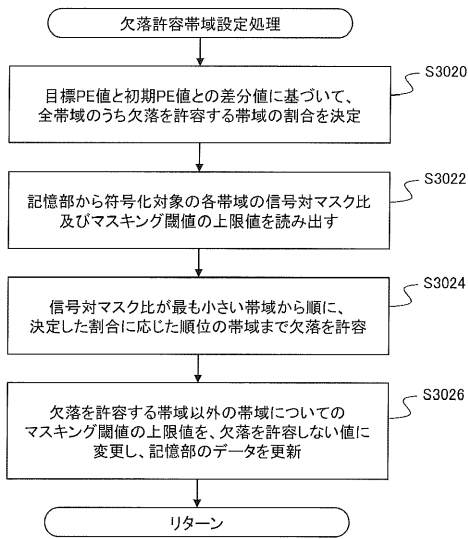
【図7】

本発明の第2の実施形態に係るマスクング閾値の補正処理を示すフローチャート



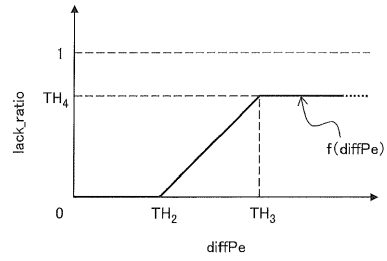
【 図 8 】

第2の実施形態における欠落許容帯域設定処理の内容を示すフローチャート



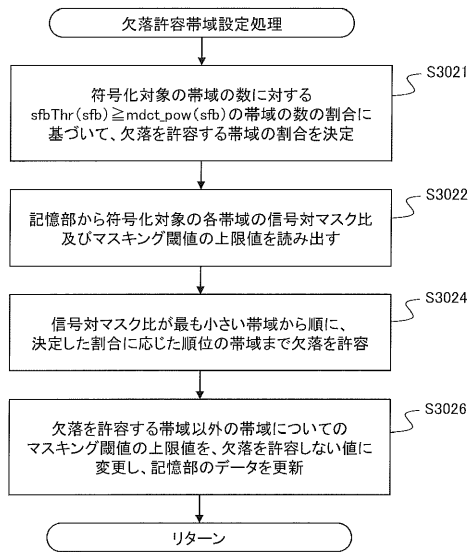
【 図 9 】

欠落を許容する帯域の割合の決定方法を説明するグラフ



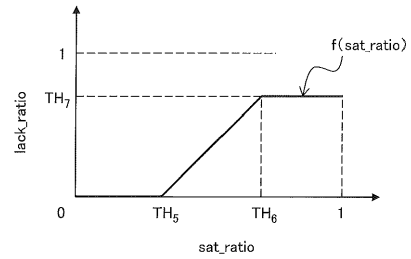
【 図 1 0 】

本発明の第3の実施形態に係る欠落許容帯域設定処理の内容を示すフローチャート



【 図 1 1 】

欠落を許容する帯域の割合の決定方法を説明するグラフ



フロントページの続き

- (72)発明者 平原 舞子
福岡県福岡市早良区百道浜2丁目2番1号 富士通九州ネットワークテクノロジーズ株式会社内
- (72)発明者 土永 義照
福岡県福岡市早良区百道浜2丁目2番1号 富士通九州ネットワークテクノロジーズ株式会社内
- (72)発明者 白川 美由紀
福岡県福岡市早良区百道浜2丁目2番1号 富士通九州ネットワークテクノロジーズ株式会社内
- (72)発明者 中村 祥吾
福岡県福岡市早良区百道浜2丁目2番1号 富士通九州ネットワークテクノロジーズ株式会社内
- (72)発明者 釜野 晃
神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
- (72)発明者 大谷 猛
神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内