(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第3698418号
(P3698418)
(45)発行日　平成17年9月21日(2005.9.21)　(24)登録日　平成17年7月15日(2005.7.15)

(51)Int.Cl.[7]　F I
G10L 19/02　G10L 19/02　143

請求項の数 2　(全 17 頁)

(21)出願番号　特願2001-157549(P2001-157549)
(22)出願日　平成13年5月25日(2001.5.25)
(62)分割の表示　特願平8-171296の分割
原出願日　平成8年7月1日(1996.7.1)
(65)公開番号　特開2002-23797(P2002-23797A)
(43)公開日　平成14年1月25日(2002.1.25)
審査請求日　平成15年4月25日(2003.4.25)

(73)特許権者　000005821
松下電器産業株式会社
大阪府門真市大字門真１００６番地
(74)代理人　100081813
弁理士　早瀬　憲一
(72)発明者　則松　武志
大阪府門真市大字門真１００６番地　松下電器産業株式会社内
(72)発明者　宮阪　修二
大阪府門真市大字門真１００６番地　松下電器産業株式会社内
(72)発明者　中藤　良久
大阪府門真市大字門真１００６番地　松下電器産業株式会社内

最終頁に続く

(54)【発明の名称】オーディオ信号圧縮方法，およびオーディオ信号圧縮装置

(57)【特許請求の範囲】

【請求項１】

入力オーディオ信号をフレーム毎に平滑化した周波数特性信号系列にＭＤＣＴ（ＭＤＣＴ：Modified Discrete Cosine Transform）変換し、変換した周波数特性信号系列の少なくとも一部の系列を量子化することにより情報量を圧縮するオーディオ信号圧縮装置のオーディオ信号圧縮方法であって、

前記オーディオ信号圧縮装置は、フレーム毎に平滑化した周波数特性信号系列の少なくとも一部の系列に対して量子化を行い第１のコードを出力する第１の量子化器と、前記第１の量子化器での量子化誤差成分に対して量子化を行い第２のコードを出力する第２の量子化器とを有する多段量子化手段を備え、

前記オーディオ信号圧縮方法は、

入力オーディオ信号のスペクトル包絡と人間の聴覚的な性質である聴覚感度特性とに基づいて、前記第１の量子化器での量子化誤差成分の周波数ブロックのうち量子化する重要度の高い周波数ブロックを選択し、前記第２の量子化器で、前記選択された周波数ブロックについて前記第１の量子化器の量子化誤差成分の量子化を行なうことを特徴とするオーディオ信号圧縮方法。

【請求項２】

入力されたオーディオ信号に対してＭＤＣＴ変換を施して周波数領域信号に変換する時間周波数変換部と、

前記入力されたオーディオ信号のスペクトル包絡を算出するスペクトル包絡算出部と、

前記時間周波数変換部で得られた周波数領域信号を前記スペクトル包絡算出部で得られたスペクトル包絡で正規化し残差信号を得る正規化部と、
前記残差信号をパワーにより正規化するパワー正規化部と、
前記パワー正規化部で正規化された残差信号に対して量子化を行い第１のコードを出力する第１の量子化器と、
入力オーディオ信号のスペクトル包絡と人間の聴覚的な性質である聴覚感度特性とに基づき、前記第１の量子化器での量子化誤差成分の周波数ブロックのうち量子化する重要度の高い周波数ブロックを選択する聴覚的選択手段と、
前記聴覚的選択手段で選択された周波数ブロックについて前記第１の量子化器の量子化誤差成分に対して量子化を行い第２のコードを出力する第２の量子化器とを備えたことを特徴とするオーディオ信号圧縮装置。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】

この発明は、オーディオ信号（音楽，音声）の、少ない容量の伝送路での情報伝送，記録メディアへの効率的な蓄積を可能とするために、特にベクトル量子化手法を応用してオーディオ信号を圧縮する場合に、従来よりも効率よく、高音質を保ったまま圧縮することのできるオーディオ信号圧縮方式，およびオーディオ信号圧縮装置に関するものである。

【０００２】

【従来の技術】

図１３は従来のオーディオ信号圧縮装置の構成を示すブロック図であり、図において、１０１は、例えば、ＭＤＣＴ（modified discrete cosine transform：変形離散コサイン変換）、あるいはＦＦＴ（高速フーリエ変換）等により入力されたディジタルオーディオ信号の時系列を、一定周期の長さ（フレーム）毎に周波数特性信号系列に変換する時間周波数変換部である。また、１０２は入力オーディオ信号をフレーム毎に線形予測分析（ＬＰＣ分析）することにより、ＬＳＰ係数（line spectrum pair coefficeent），あるいは偏自己相関係数等を抽出するとともに、これらの係数からＬＰＣスペクトル包絡を求めるスペクトル包絡算出部である。１０３は時間周波数変換部１０１で算出された周波数特性信号系列をスペクトル包絡算出部１０２で求めたＬＰＣスペクトル包絡で割り算して正規化することにより、周波数特性を平坦化し、さらにパワーの最大値，あるいは平均値等に基づいてパワーの正規化を行なう正規化部である。以降の説明では、このパワーの正規化が行なわれた時点の出力係数を残差信号とも呼ぶ。１０５は正規化部１０３で平坦化された残差信号をスペクトル包絡を重み付けとしてベクトル量子化する量子化部である。この量子化部１０５は、正規化された残差信号を量子化する第１段の量子化部１５１と、第１段の量子化部１５１での量子化誤差成分を量子化する第２段の量子化部１５２を含む多段構成を有する。

【０００３】

次に動作について説明する。入力されたオーディオ信号の時系列は、時間周波数変換部１０１において、ＭＤＣＴ，あるいはＦＦＴ等により、一定周期の長さ（フレーム）毎に周波数特性信号系列に変換される。また、入力されたオーディオ信号は、スペクトル包絡算出部１０２において、ＬＰＣスペクトル包絡が求められる。正規化部１０３では、時間周波数変換部１０１の出力をＬＰＣスペクトル包絡で割り算することにより周波数特性信号系列の周波数特性を平坦化した残差信号を求め、さらにこの残差信号について、パワーの最大値，あるいは平均値等に基づいてパワーの正規化が行なわれる。正規化された残差信号は、量子化部10５の第１段の量子化部１５１においてスペクトル包絡を重み付けとして全周波数帯域にわたってベクトル量子化され、さらに、第１段の量子化部１５１での量子化誤差信号が量子化部１０５の第２段の量子化部１５２においてスペクトル包絡を重み付けとして全周波数帯域にわたってベクトル量子化される。このように、本従来例では、量子化による誤差を少なくするために、ベクトル量子化を行なう量子化部を多段構成とし、初段のベクトル量子化器での誤差信号成分を次段のベクトル量子化器でベクトル量子化

しているものである。
【０００４】
図１４は従来の他のオーディオ信号圧縮装置の構成を示すブロック図であり、図において、２０１は入力されたオーディオ信号を複数の帯域毎の信号に分割する帯域分割部、２０２は帯域分割された帯域毎の信号のそれぞれについて、最小可聴限特性やノイズマスキング特性等の聴覚感度特性に基づいて、これらの量子化の際の重み付けを計算する聴覚重み付け計算部、２０３は帯域分割された帯域毎の信号をそれぞれ聴覚的重み付け計算部２０２で算出された重み付けを用いてスカラ量子化を行なう量子化部である。
【０００５】
次に動作について説明する。
入力されたオーディオ信号は帯域分割部２０１において周波数帯域別に複数の信号に分割され、聴覚重み付け計算部２０２に入力される。聴覚重み付け計算部２０２では、入力された信号について、最小可聴限特性やノイズマスキング特性等の聴覚感度特性に基づいて、各周波数帯域毎に量子化の際の重み付けを求める。量子化部２０３では、分割された信号のそれぞれを、聴覚重み付け計算部２０２で求められた各周波数帯域毎の重み付け係数を重み付けとしてスカラ量子化する。
【０００６】
【発明が解決しようとする課題】
図１３に示す従来例では、第２段の量子化部で、第１段の量子化部での量子化誤差信号をその全周波数帯域にわたって、スペクトル包絡を重み付けとして量子化しており、この第２段の量子化部での量子化では第１段の量子化部での量子化が良好で量子化誤差が小さい周波数帯域についても量子化のためのビットが割り当てられるため、この量子化誤差が小さい周波数帯域を量子化することによって、逆に誤差を大きくしてしまうことがあるという問題があった。また、量子化の際の重み付けをスペクトル包絡のみに基づいて行なっているため、人間の聴覚的な性質を利用して効率よく量子化することができないという問題があった。
【０００７】
また、図１４に示す従来例は、聴覚感度特性に基づいて重み付けをして量子化しているが、スカラ量子化を行なうものであるため、ベクトル量子化を行なうものに比して量子化効率が良くないという問題があった。
【０００８】
この発明は、上記のような問題を解消するためになされたものであり、ベクトル量子化を用いてオーディオ信号を圧縮する場合に、従来よりも効率よく、高音質を保ったまま圧縮することのできるオーディオ信号圧縮方法，およびオーディオ信号圧縮装置を提供することを目的とする。
【０００９】
【課題を解決するための手段】
上記課題を解決するために、本発明（請求項１）に係るオーディオ信号圧縮方法は、入力オーディオ信号をフレーム毎に平滑化した周波数特性信号系列にＭＤＣＴ変換し、変換した周波数特性信号系列の少なくとも一部の系列を量子化することにより情報量を圧縮するオーディオ信号圧縮装置のオーディオ信号圧縮方法であって、前記オーディオ信号圧縮装置は、フレーム毎に平滑化した周波数特性信号系列の少なくとも一部の系列に対して量子化を行い第１のコードを出力する第１の量子化器と、前記第１の量子化器での量子化誤差成分に対して量子化を行い第２のコードを出力する第２の量子化器とを有する多段量子化手段を備え、前記オーディオ信号圧縮方法は、入力オーディオ信号のスペクトル包絡と人間の聴覚的な性質である聴覚感度特性とに基づいて、前記第１の量子化器での量子化誤差成分の周波数ブロックのうち量子化する重要度の高い周波数ブロックを選択し、前記第２の量子化器で、前記選択された周波数ブロックについて前記第１の量子化器の量子化誤差成分の量子化を行なうようにしたものである。
【００１０】

また、本発明（請求項２）に係るオーディオ信号圧縮装置は、入力されたオーディオ信号に対してＭＤＣＴ変換を施して周波数領域信号に変換する時間周波数変換部と、前記入力されたオーディオ信号のスペクトル包絡を算出するスペクトル包絡算出部と、前記時間周波数変換部で得られた周波数領域信号を前記スペクトル包絡算出部で得られたスペクトル包絡で正規化し残差信号を得る正規化部と、前記残差信号をパワーにより正規化するパワー正規化部と、前記パワー正規化部で正規化された残差信号に対して量子化を行い第１のコードを出力する第１の量子化器と、入力オーディオ信号のスペクトル包絡と人間の聴覚的な性質である聴覚感度特性とに基づき、前記第１の量子化器での量子化誤差成分の周波数ブロックのうち量子化する重要度の高い周波数ブロックを選択する聴覚的選択手段と、前記聴覚的選択手段で選択された周波数ブロックについて前記第１の量子化器の量子化誤差成分に対して量子化を行い第２のコードを出力する第２の量子化器とを備えたことを特徴とするものである。

【００１２】

【発明の実施の形態】

実施の形態１．

図１は本発明の実施の形態１によるオーディオ信号圧縮装置の構成を示すブロック図であり、図において、１は、例えば、ＭＤＣＴ、あるいはＦＦＴ等により入力されたディジタルオーディオ信号の時系列を、一定周期の長さ（フレーム）毎に周波数特性信号系列に変換する時間周波数変換部である。また、２は入力オーディオ信号をフレーム毎に線形予測分析（ＬＰＣ分析）することにより、ＬＳＰ係数，あるいは偏自己相関係数等を抽出するとともに、これらの係数からＬＰＣスペクトル包絡を求めるスペクトル包絡算出部である。３は時間周波数変換部１で算出された周波数特性信号系列をスペクトル包絡算出部２で求めたＬＰＣスペクトル包絡で割り算して正規化することにより、周波数特性を平坦化する正規化部、４は正規化部３で平坦化された周波数特性信号系列をパワーの最大値，あるいは平均値等に基づいてパワーの正規化を行なうパワー正規化部である。５は正規化部３，パワー正規化部４で平坦化された周波数特性信号系列をベクトル量子化する多段量子化部である。多段量子化部５は縦列接続された第１段の量子化器５１，第２段の量子化器５２，…，第Ｎ段の量子化器５３を含む。６は時間周波数変換部１から出力された周波数特性信号系列とスペクトル包絡算出部２で求めたＬＰＣスペクトル包絡を入力とし、聴覚感度特性に基づいて、量子化部５での量子化の際に用いる重み付け係数を求める聴覚重み付け計算部である。

【００１３】

次に動作について説明する。入力されたディジタルオーディオ信号（以下入力信号とも記す）の時系列は、一定周期の長さ（フレーム）毎に時間周波数変換部１でＭＤＣＴ，ＦＦＴ等により周波数特性信号系列に変換される。また入力信号は、フレーム毎にスペクトル包絡算出部２で、線形予測分析（ＬＰＣ分析）によりＬＳＰ係数，あるいは偏自己相関係数等が抽出され、さらにこれらの係数からＬＰＣスペクトル包絡が求められる。正規化部３では、上記で算出された周波数特性信号系列をＬＰＣスペクトル包絡で割り算し正規化することにより、周波数特性信号系列を平坦化する。正規化部３で平坦化された周波数特性信号系列は、パワー正規化部４において、さらにパワーの最大値，あるいは平均値等に基づいてパワーの正規化が行われる。

【００１４】

一方、聴覚重み付け計算部６は、時間周波数変換部１から出力された周波数特性信号系列とスペクトル包絡算出部２で求めたＬＰＣスペクトル包絡が入力され、時間周波数変換部１から出力された周波数特性信号系列のスペクトルについて、最小可聴限特性や聴覚マスキング特性等の人間の聴覚的な性質である聴覚感度特性に基づいて、この聴覚感度特性を考慮した特性信号を算出し、さらにこの特性信号とＬＰＣスペクトル包絡に基づいて量子化に用いる重み付け係数を求める。

【００１５】

パワー正規化部４から出力された残差信号は、多段量子化部５の第１段の量子化部５１で

聴覚重み付け計算部６によって求められた重み付け係数を用いて量子化され、第１段の量子化部５１での量子化による量子化誤差成分が、多段量子化部５の第２段の量子化部５２で聴覚重み付け計算部６によって求められた重み付け係数を用いて量子化され、以下同様にして複数段の量子化部のそれぞれにおいて、前段の量子化部での量子化による量子化誤差成分の量子化が行なわれる。そして第Ｎ－１段の量子化部での量子化による量子化誤差成分を第Ｎ段の量子化部５３で聴覚重み付け計算部６によって求められた重み付け係数を用いて量子化が行なわれることによりオーディオ信号の圧縮符号化が完了する。

【００１６】

このように、本実施の形態１によるオーディオ信号圧縮方法，およびオーディオ信号圧縮装置によれば、多段量子化手段５の複数段のベクトル量子化部５１ないし５３で、聴覚重み付け計算部６において入力オーディオ信号のスペクトル，人間の聴覚的な性質である聴覚感度特性，及びＬＰＣスペクトル包絡に基づいて算出された周波数上の重み付け係数を量子化の際の重み付けとして用いてベクトル量子化を行なう構成としたので、人間の聴覚的な性質を利用して効率の良い量子化を行なうことができる。

【００１７】

なお、図１のオーディオ信号圧縮装置では、聴覚重み付け計算部６が重み付け係数の算出にＬＰＣスペクトル包絡も用いる構成としているが、入力オーディオ信号のスペクトルと人間の聴覚的な性質である聴覚感度特性のみを用いて重み付け係数を算出するようにしてもよい。

【００１８】

また、図１のオーディオ信号圧縮装置では、多段量子化手段５の複数段のベクトル量子化部の全てが聴覚重み付け計算部６において求められた聴覚感度特性に基づく重み付け係数を用いて量子化するようにしているが、多段量子化手段５の複数段のベクトル量子化器のいずれか１つが聴覚感度特性に基づく重み付け係数を用いて量子化を行なうものであれば、このような聴覚感度特性に基づく重み付け係数を用いない場合に比して、効率のよい量子化を行なうことができる。

【００１９】

実施の形態２．

図２は本発明の実施の形態２によるオーディオ信号圧縮装置の構成を示すブロック図であり、図において、図１と同一符号は同一または相当部分である。６１は入力オーディオ信号のスペクトル，人間の聴覚的な性質である聴覚感度特性，及びＬＰＣスペクトル包絡に基づいて、多段量子化手段５の第１段の量子化部５１が用いる重み付け係数を求める第１の聴覚重み付け計算部、６２は同じく入力オーディオ信号のスペクトル，人間の聴覚的な性質である聴覚感度特性，及びＬＰＣスペクトル包絡に基づいて、多段量子化手段５の第２段の量子化部５２が用いる重み付け係数を求める第２の聴覚重み付け計算部、６３は同じく入力オーディオ信号のスペクトル，人間の聴覚的な性質である聴覚感度特性，及びＬＰＣスペクトル包絡に基づいて、多段量子化手段５の第Ｎ段の量子化部５３が用いる重み付け係数を求める第３の聴覚重み付け計算部である。

【００２０】

上記実施の形態１によるオーディオ信号圧縮装置では、多段量子化手段５の複数段のベクトル量子化部の全てが聴覚重み付け計算部６において求められた同一の重み付け係数を用いて量子化するようにしているのに対し、本実施の形態２によるオーディオ信号圧縮装置では、多段量子化手段５の複数段のベクトル量子化部が、それぞれ第１ないし第３の聴覚重み付け計算部で求められた個別の重み付け係数を用いて量子化する構成としている。このような本実施の形態２によるオーディオ信号圧縮装置では、多段量子化手段５の各段で量子化による誤差が最小となるように、聴覚重み付け部６１ないし６３で求められた聴覚的な性質に基づいた周波数重み付け特性に従った重み付けによる量子化を行なうことができる。

【００２１】

たとえば、第１の聴覚重み付け部６１ではスペクトル包絡を主体として重み付け係数を算

出し、第２の聴覚重み付け部６２では最小可聴限特性を主体として重み付け係数を算出し、第３の聴覚重み付け部６３では聴覚マスキング特性を主体として重み付け係数を算出するものである。

【００２２】

このように、本実施の形態２によるオーディオ信号圧縮方法，およびオーディオ信号圧縮装置によれば、多段量子化部５の上記複数段の量子化部の複数の量子化部５１ないし５３がそれぞれ聴覚重み付け計算部６１ないし６３で得られたそれぞれ個別の重み付け係数を用いて量子化を行なう構成としたので、人間の聴覚的な性質をより有効に利用して効率の良い量子化を行なうことができる。

【００２３】

実施の形態３．

図３は本発明の実施の形態３によるオーディオ信号圧縮装置の構成を示すブロック図であり、図において、図２と同一符号は同一または相当部分である。２１はパワー正規化された残差信号をベクトル量子化する第１段の量子化部、２３は第１段の量子化部２１における量子化による量子化誤差信号を量子化する第２段の量子化部、２２は第１段の量子化部２１における量子化による量子化誤差のうち、第２段の量子化部２３で量子化すべき重要度の高い周波数領域を聴覚感度特性を考慮した基準により選択する聴覚的選択手段である。

【００２４】

次に動作について説明する。ディジタル入力信号は時間周波数変換部１において、フレーム長Ｎ（Ｎは１フレーム内のサンプル数）ごとに例えばＭＤＣＴ変換により、周波数領域係数に変換される。また入力信号はスペクトル包絡算出部（ＬＰＣ分析部）２で線形予測分析され、例えばＬＳＰ係数が求められ、それらが量子化されコードとして出力されるとともに、さらに線形予測係数を高速フーリエ変換することによりＬＰＣスペクトルの包絡が求められる。

【００２５】

ＭＤＣＴ変換により得られた周波数領域係数系列は、正規化部３においてＬＰＣスペクトル包絡で除算（正規化）され、スペクトルの平坦化された係数が得られる。スペクトルの平坦化された係数は、さらにパワー正規化部４で、例えばフレーム内のパワーの最大値で除算することにより正規化が行われる。パワー正規化部４からは正規化のために用いたパワーを示すコードが出力される。

【００２６】

パワー正規化された残差信号は、第１段の量子化部２１でベクトル量子化される。聴覚的選択手段２２では、ベクトル量子化での誤差信号が大きい部分を聴覚的な尺度に基づき判別し、そのブロックを抽出する。第２段の量子化部２３では選択されたブロック部分の誤差信号に対してベクトル量子化を施す。そして各量子化部の結果がコードとして出力される。

【００２７】

図４は図３に示すオーディオ信号圧縮装置の各量子化部，及び聴覚的選択部の詳細な構成を示すブロック図である。図において３１は残差信号をベクトル量子化する第１のベクトル量子化器、３２は第１の量子化器３１の量子化結果を逆量子化する逆量子化器であり、この逆量子化器３２の出力と残差信号ｓｉの差分をとることにより第１の量子化器３１による量子化の量子化誤差信号ｚｉが得られる。３３は人間の聴覚的な性質を示す聴覚感度特性ｈｉであり、ここでは最小可聴限特性を用いる。３５は第１の量子化器３１による量子化の量子化誤差信号ｚｉのうち、第２のベクトル量子化器３６で量子化すべき周波数領域を選択する選択器である。３４は、誤差信号ｚｉ，ＬＣＰスペクトル包絡ｌｉ，及び聴覚感度特性ｈｉに基づいて選択器３５の選択動作における選択尺度を計算する選択尺度計算部である。

【００２８】

次に、聴覚的選択部による選択動作について詳細に説明する。

第１のベクトル量子化器３１では、まずＮ個の要素により構成される１フレーム内の残差信号を、図９(a) に示す，第１のベクトル量子化器３１内のベクトル分割部で複数個のサブベクトルに分割し、第１のベクトル量子化器３１内のＮ個の量子化器１～Ｎでサブベクトルのそれぞれをベクトル量子化する。ベクトル分割，及び量子化の方法は、例えば図９(b) に示すように、周波数の低い方から順に並んでいるＮ個の要素をＮＳ個のサブブロックに等間隔に分割し、それぞれのサブブロックの第１要素のみを集めたサブベクトル、第２要素を集めたサブベクトルというようにＮ／ＮＳ個の要素で構成されるＮＳ個のサブベクトルを作成し、それぞれのサブベクトル毎にベクトル量子化を実行する。分割数等は要求される圧縮率に基づき決定される。

【００２９】

ベクトル量子化後、逆量子化器３２により量子化コードを逆量子化して入力信号との差をとることにより、図８(a) に示すような、第１のベクトル量子化器３１での誤差信号ｚｉを得る。

【００３０】

次に、選択器３５では誤差信号Ｚｉのうち第２の量子化器３６によりさらに精密に量子化すべき周波数ブロックを、選択尺度計算部３４で選択された結果に基づき選択する。

【００３１】

選択尺度計算部３４では、誤算信号Ｚｉ，ＬＰＣ分析部で得られた、図８(b) に示すようなＬＰＣスペクトル包絡ｌｉ，及び聴覚感度特性ｈｉを用い、周波数軸上でＮ個の要素に分割されたフレームの各要素毎に、

ｇ＝（ｚｉ＊ｌｉ）／ｈｉ

を計算する。

【００３２】

聴覚感度特性ｈｉとしては、例えば図８(c) に示す最小可聴限界特性といわれるものを使う。これは実験的に求められた本来人間には聞こえない領域を表す特性である。従って、この聴覚感度特性ｈｉの逆数であるｌ／ｈｉは人間の聴覚的な重要性を表すといえるものである。そして、誤差信号ｚｉ，スペクトル包絡ｌｉ，及び聴覚感度特性ｈｉの逆数を乗算した値ｇはその周波数でのさらに精密に量子化する重要度を表すと言える。

【００３３】

また、図５は図３に示すオーディオ信号圧縮装置の各量子化部，及び聴覚的選択部の他の例の詳細な構成を示すブロック図である。図において、図４と同一符号は同一または相当部分である。図５に示す例では、選択尺度（重要度）ｇを誤差信号ｚｉを使用せずに、スペクトル包絡ｌｉと聴覚感度特性ｈｉを用いて、

ｇ＝ｌｉ／ｈｉ

を計算して求める。

【００３４】

また、図６は図３に示すオーディオ信号圧縮装置の各量子化部，及び聴覚的選択部のさらに他の例の詳細な構成を示すブロック図である。図において、図４と同一符号は同一または相当部分であり、４２は時間周波数変換部でＭＤＣＴ変換された入力オーディオ周波数のスペクトルから聴覚マスキング特性によりマスキングされる量を算出するマスキング量算出部である。

【００３５】

図６に示す例では、聴覚感度特性ｈｉは以下のようにしてフレームごとに逐次求める。すなわち、入力信号の周波数スペクトル分布からマスキング特性を算出し、そのマスキング特性に最小可聴限界特性を加えることにより、そのフレームの聴覚感度特性ｈｉを求めることができる。選択尺度計算部３４の動作は図５のものと同じである。

【００３６】

また、図７は図３に示すオーディオ信号圧縮装置の各量子化部，及び聴覚的選択部のさらに他の例の詳細な構成を示すブロック図である。図において、図４と同一符号は同一または相当部分であり、４３はマスキング量計算部４２で得られたマスキング特性を、スペク

トル包絡ｌｉ，残差信号ｓｉ，及び誤差信号ｚｉを用いて補正するマスキング量補正部である。

【００３７】

図７に示す例では、聴覚感度特性ｈｉは以下のようにしてフレームごとに逐次求める。まず、入力信号の周波数スペクトル分布からマスキング量計算部４２でマスキング特性を算出する。次にマスキング量補正部４３で、スペクトル包絡ｌｉ，残差信号ｓｉ，及び誤差信号ｚｉに応じて、上記算出されたマスキング特性を補正する。この補正されたマスキング特性に最小可聴限界特性を加えることにより、そのフレームの聴覚感度特性ｈｉを求めることができる。ここで、マスキング特性を補正する方法の一例を示す。

【００３８】

まず、すでに算出されているマスキング量Ｍｉの特性が極大値を示す周波数（ｆｍ）を求める。次にその周波数ｆｍの信号がどれほどの正確さで再生されるかを、入力時の周波数ｆｍのスペクトルの強さと量子化誤差スペクトルの大きさから求める。例えば、

γ＝１－（ｆｍの量子化誤差のゲイン）／（ｆｍの入力時のゲイン）

とする。

【００３９】

このγの値が１に近ければ、すでに求めているマスキング特性を変形する必要はないが、０に近ければ、小さくする方向で補正する。例えば、

【数１】

$$hi = Mi^{\gamma}$$

のように、マスキング特性にγを乗ずることで変形することにより補正することができる。

【００４０】

次に、選択器３５の動作について説明する。

選択器３５ではフレーム内の連続した要素毎に窓（長さＷとする）を掛け、その窓内での重要度ｇの値を累積した値Ｇが最大値を示す周波数ブロックを選択する。図１０は重要度が最も高い周波数ブロック（長さＷ）を選択する一例を示す図である。窓の長さは簡単のためにはＮ／ＮＳの整数倍に設定するのがよい（図１０では整数倍でないものを示している）。この窓をＮ／ＮＳ個ずつずらしながら、その窓枠内の重要度ｇの累積値Ｇを計算しその最大値を与える長さＷの周波数ブロックを選択する。

【００４１】

選択された窓枠内のブロックに対して、第２のベクトル量子化器３２でベクトル量子化が行われる。第２のベクトル量子化器３２の動作は第１のベクトル量子化器３１と同様であるが、上述のように誤差信号ｚｉのうち、選択器３５で選択された周波数ブロックのみを量子化するものであるのでベクトル量子化されるフレーム内の要素の数は少ないものである。

【００４２】

最終的にはスペクトル包絡係数のコード，各ベクトル量子化器の量子化結果であるそれぞれのコード，及び図４，図６，及び図７に示す構成で求めた選択尺度ｇを使用する場合は、選択器３５で、どの要素から始まるブロックが選択されたかの情報がコードとして出力される。

【００４３】

一方、図５に示す構成で求めた選択尺度ｇを使用する場合は、スペクトル包絡ｌｉと聴覚感度特性ｈｉのみを用いているので、逆量子化時に、どの要素から始まるブロックが選択されたかの情報はスペクトル包絡係数のコードと既知の聴覚感度特性ｈｉから求めることが可能であるので、ブロックの選択情報をコードとして出力する必要がなく、圧縮率の点で有利となる。

【００４４】

このように、本実施の形態３によるオーディオ信号圧縮方法，およびオーディオ信号圧縮装置によれば、入力オーディオ信号のスペクトルと人間の聴覚的な性質である聴覚感度特性に基づいて、第１のベクトル量子化器での量子化誤差成分の周波数ブロックのうち量子化する重要度の高い周波数ブロックを選択し、第２のベクトル量子化器で、上記選択された周波数ブロックについて上記第１の量子化器の量子化誤差成分の量子化を行なうようにしたから、人間の聴覚的な性質を利用して効率の良い量子化を行なうことができる。また、図４，図６，及び図７に示す構成では、量子化する重要度の高い周波数ブロックの選択の際に第１のベクトル量子化器での量子化誤差に基づいて重要度を算出するようにしたから、第１のベクトル量子化器での量子化が良好である部分が再度量子化されて逆に誤差が生じてしまうことを防止でき、高品質を保持した量子化を行なうことができる。また、図５に示す構成で重要度ｇを求める場合は、図４，図６，及び図７に示す構成で重要度ｇを求める場合に比して、出力すべきコードを少なくでき、圧縮率を向上できる。

【００４５】

なお、上記実施の形態３では、量子化部が第１段の量子化部２１，第２段の量子化部２３の２段構成で、この第１段の量子化部２１と第２段の量子化部２３の間に聴覚的選択手段２２を設けたものについて説明したが、量子化部を３段以上の複数段構成とし、各量子化部の間にそれぞれ聴覚的選択手段を設ける構成としてもよく、かかる構成とした場合も、上記実施の形態３と同様、人間の聴覚的な性質を利用して効率の良い量子化を行なうことができる。

【００４６】

実施の形態４．

図１１は本発明の実施の形態４によるオーディオ信号圧縮装置の構成を示すブロック図であり、図において、図１と同一符号は同一または相当部分である。１１はパワー正規化部４から出力された残差信号ｓｉをスペクトル包絡の値ｌｉを重み係数としてベクトル量子化する第１段の量子化器、１２は第１段の量子化器１１の量子化結果を逆量子化する逆量子化器であり、この逆量子化器１２の出力とパワー正規化部４から出力された残差信号の差分をとることにより第１段の量子化器１１による量子化の量子化誤差信号ｚｉが得られる。１３は第１段の量子化器１１による量子化の量子化誤差信号ｚｉを後述する重み計算部１７の計算結果を重み係数としてベクトル量子化する第２段の量子化器、１４は第２段の量子化器１３の量子化結果を逆量子化する逆量子化器であり、この逆量子化器１４の出力と第１段の量子化器１１による量子化の量子化誤差信号の差分をとることにより第２段の量子化器１３による量子化の量子化誤差信号ｚ２ｉが得られる。１５は第２段の量子化器１３による量子化の量子化誤差信号ｚ２ｉを聴覚重み計算部６の計算結果を重み係数としてベクトル量子化する第３段の量子化器である。

【００４７】

１６は第１段の量子化器１１による量子化の量子化誤差信号ｚｉとスペクトル包絡の値ｌｉの相関を計算する相関計算部、１７は相関計算部の計算結果とスペクトル包絡の値ｌｉに基づいて第２段の量子化器１３による量子化での重み付け係数を算出する重み計算部である。

【００４８】

次に動作について説明する。本実施の形態４のオーディオ信号圧縮装置は、３段の量子化器を用い、それぞれの量子化器で異なる重み付けでベクトル量子化を行なうものである。まず第１段の量子化器１３において、入力された残差信号ｓｉが、ＬＰＣ分析部２で求められたＬＰＣスペクトル包絡の値ｌｉを重み係数としてベクトル量子化が実行される。これにより、スペクトルのエネルギーの大きな（集中した）部分に重み付けを施されることになり、結果として聴覚的に重要な部分をより高精度に量子化させる効果がある。この第１段のベクトル量子化器１３は、例えば実施の形態３における第１のベクトル量子化器３１と同様のものを使用すればよい。量子化結果は逆量子化器１２で逆量子化され、これともとの入力残差信号ｓｉの差分により量子化による誤差信号ｚｉが得られる。

【００４９】

この誤差信号ｚｉは第２段の量子化器１３でさらにベクトル量子化される。ここでは、ＬＰＣスペクトル包絡ｌｉと誤差信号ｚｉとの相関をもとに、相関計算部１６，及び重み計算部１７において重み係数を算出する。
具体的には相関計算部１６で、
α＝（Σｌｉ＊ｚｉ）／（Σｌｉ＊ｌｉ）
を計算する。このαは０＜α＜１の値をとり、両者の相関度を表す。αが０に近い時は，第１段目の量子化がスペクトル包絡の重み付けに基づき精度よく行われたことを示しており、αが１に近い時はまだ精度よく量子化できていないことを示す。そこで、このαによりスペクトル包絡ｌｉの重み付け度合いを調整する係数として、
【数２】

$$l_i^{\alpha}$$

を求め、ベクトル量子化の際の重み付け係数とする。このように第１段目の量子化の精度に応じて再度スペクトルの包絡で重み付けし、量子化することで、量子化精度が向上する。
【００５０】
第２段目の量子化器１３による量子化結果も同様に逆量子化器１４で逆量子化され、誤差信号ｚ２ｉが抽出され、この誤差信号ｚ２ｉが第３段の量子化器１５でベクトル量子化される。この時の聴覚的重み係数は、聴覚重み付け計算部６の重み計算部１９で計算される。たとえば、誤差信号ｚ２ｉ，ＬＰＣスペクトル包絡ｌｉ，及び残差信号ｓｉを用いて、
Ｎ＝Σｚ２ｉ＊ｌｉ
Ｓ＝Σｓｉ＊ｌｉ
β＝１－（Ｎ／Ｓ）
を求める。
【００５１】
一方、聴覚重み付け計算部６の聴覚マスキング計算部１８では、たとえばＭＰＥＧ（エムペグ）オーディオ標準方式の中で用いられている聴覚モデルにより聴覚マスキング特性ｍｉを計算する。これに上述した最小可聴限界特性ｈｉを重ねあわせて最終的なマスキング特性Ｍｉを求める。
【００５２】
そして、この最終的なマスキング特性Ｍｉに重み計算部１９で計算された係数βを乗じた
【数３】

$$1/m_i^{\beta}$$

第３段のベクトル量子化の際の重み係数として用いる。
【００５３】
このように、本実施の形態４によるオーディオ信号圧縮方法，およびオーディオ信号圧縮装置によれば、複数の量子化器１１，１３，１５が聴覚感度特性を考慮した重み付けを含む，それぞれ異なる重み付け係数を用いて量子化を行なう構成としたので、人間の聴覚的な性質をより有効に利用して効率の良い量子化を行なうことができる。
【００５４】
実施の形態５．
図１２は本発明の実施の形態５によるオーディオ信号圧縮装置の構成を示すブロック図であり、図において、図１，図３と同一符号は同一または相当部分である。
本実施の形態５によるオーディオ信号圧縮装置は、図３に示す実施の形態３と図１に示す実施の形態１とを組み合わせたものであり、図３に示す実施の形態３によるオーディオ信号圧縮装置において、各量子化部での量子化の際に、聴覚重み付け計算部６において聴覚感度特性を用いて求めた重み付け係数を用いるようにしたものである。本実施の形態５に

よるオーディオ信号圧縮装置では、かかる構成とすることにより、上記実施の形態１と実施の形態３により得られる効果の両方を得ることができるものである。

【００５５】

また、同様にして、図３に示す実施の形態３に実施の形態２，または実施の形態４の構成を組み合わせることも可能であり、それぞれの組み合わせにより得られるオーディオ信号圧縮装置は、それぞれ実施の形態２と実施の形態３により得られる効果の両方，及び実施の形態４と実施の形態３により得られる効果の両方が得られるものである。

【００５６】

なお、上記実施の形態１ないし５では多段量子化部は量子化部の段数が２段，または３段のものを示したが、量子化部の段数を４段以上の多段構成としてもよいことは言うまでもない。

【００５７】

また、多段量子化部の各段におけるベクトル量子化の際に用いる重み係数の順序は、上記実施の形態に示したものに限られるものではなく、たとえば、第１段で聴覚感度特性を考慮した重み付けを用い、２段目以降でＬＰＣスペクトル包絡を用いるようにしてもよい。

【００５８】

【発明の効果】

以上のように、本発明（請求項１）のオーディオ信号圧縮方法によれば、入力オーディオ信号をフレーム毎に平滑化した周波数特性信号系列にＭＤＣＴ変換し、変換した周波数特性信号系列の少なくとも一部の系列を量子化することにより情報量を圧縮するオーディオ信号圧縮装置のオーディオ信号圧縮方法であって、前記オーディオ信号圧縮装置は、フレーム毎に平滑化した周波数特性信号系列の少なくとも一部の系列に対して量子化を行い第１のコードを出力する第１の量子化器と、前記第１の量子化器での量子化誤差成分に対して量子化を行い第２のコードを出力する第２の量子化器とを有する多段量子化手段を備え、前記オーディオ信号圧縮方法は、入力オーディオ信号のスペクトル包絡と人間の聴覚的な性質である聴覚感度特性とに基づいて、前記第１の量子化器での量子化誤差成分の周波数ブロックのうち量子化する重要度の高い周波数ブロックを選択し、前記第２の量子化器で、前記選択された周波数ブロックについて前記第１の量子化器の量子化誤差成分の量子化を行なうようにしたから、人間の聴覚的な性質を利用して効率の良い量子化を行なうことができる効果がある。

【００５９】

また、本発明（請求項２）のオーディオ信号圧縮装置によれば、入力されたオーディオ信号に対してＭＤＣＴ変換を施して周波数領域信号に変換する時間周波数変換部と、前記入力されたオーディオ信号のスペクトル包絡を算出するスペクトル包絡算出部と、前記時間周波数変換部で得られた周波数領域信号を前記スペクトル包絡算出部で得られたスペクトル包絡で正規化し残差信号を得る正規化部と、前記残差信号をパワーにより正規化するパワー正規化部と、前記パワー正規化部で正規化された残差信号に対して量子化を行い第１のコードを出力する第１の量子化器と、入力オーディオ信号のスペクトル包絡と人間の聴覚的な性質である聴覚感度特性とに基づき、前記第１の量子化器での量子化誤差成分の周波数ブロックのうち量子化する重要度の高い周波数ブロックを選択する聴覚的選択手段と、前記聴覚的選択手段で選択された周波数ブロックについて前記第１の量子化器の量子化誤差成分に対して量子化を行い第２のコードを出力する第２の量子化器とを備えた構成としたから、人間の聴覚的な性質を利用して効率の良い量子化を行なうことができる効果がある。

【００６１】

【発明の効果】

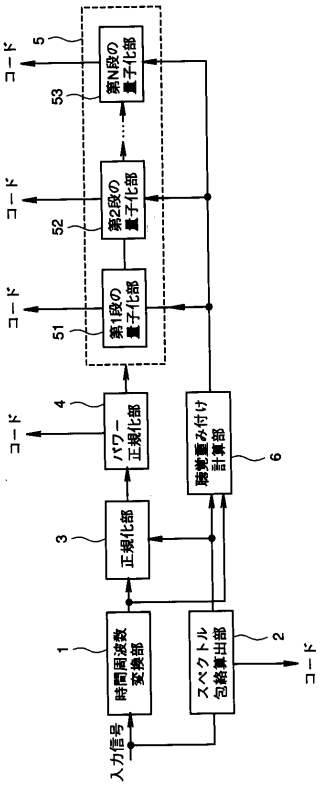
【図面の簡単な説明】

【図１】本発明の実施の形態１によるオーディオ信号圧縮装置の構成を示すブロック図である。

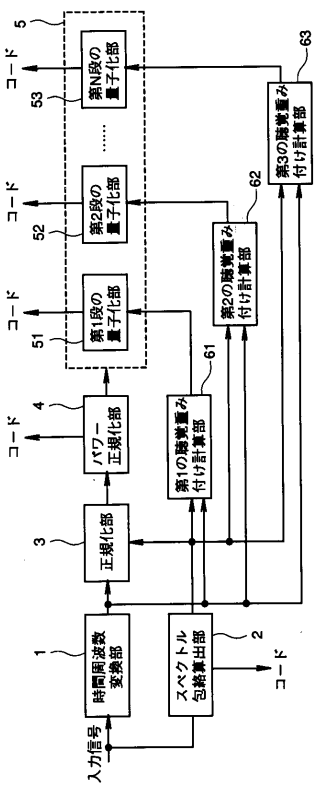
【図２】本発明の実施の形態２によるオーディオ信号圧縮装置の構成を示すブロック図で

ある。
【図３】本発明の実施の形態３によるオーディオ信号圧縮装置の構成を示すブロック図である。
【図４】本発明の実施の形態３によるオーディオ信号圧縮装置の各量子化部，及び聴覚的選択手段の詳細な構成の一例を示すブロック図である。
【図５】本発明の実施の形態３によるオーディオ信号圧縮装置の各量子化部，及び聴覚的選択手段の詳細な構成の他の例を示すブロック図である。
【図６】本発明の実施の形態３によるオーディオ信号圧縮装置の各量子化部，及び聴覚的選択手段の詳細な構成のさらに他の例を示すブロック図である。
【図７】本発明の実施の形態３によるオーディオ信号圧縮装置の各量子化部，及び聴覚的選択手段の詳細な構成のさらに他の例を示すブロック図である。
【図８】誤差信号ｚｉ，スペクトル包絡ｌｉ，及び最小可聴限特性ｈｉを示す図である。
【図９】ベクトル量子化器での量子化の動作を説明するための図である。
【図１０】選択器の動作を説明するための図である。
【図１１】本発明の実施の形態４によるオーディオ信号圧縮装置の構成を示すブロック図である。
【図１２】本発明の実施の形態５によるオーディオ信号圧縮装置の構成を示すブロック図である。
【図１３】従来のオーディオ信号圧縮装置の構成を示すブロック図である。
【図１４】従来の他のオーディオ信号圧縮装置の構成を示すブロック図である。
【符号の説明】
１　時間周波数変換部
２　スペクトル包絡算出部
３　正規化部
４　パワー正規化部
５　多段量子化部
６　聴覚重み付け計算部
１１　第１段の量子化器
１２　逆量子化器
１３　第２段の量子化器
１４　逆量子化器
１５　第３段の量子化器
１６　相関計算部
１７　重み計算部
１８　聴覚マスキング計算部
１９　重み計算部
２１　第１段の量子化部
２２　聴覚的選択手段
２３　第２段の量子化部
３１　第１のベクトル量子化器
３２　逆量子化器
３３　聴覚感度特性（最小可聴限特性）
３４　選択尺度計算部
３５　選択器
３６　第２のベクトル量子化器
４２　マスキング量算出部
４３　マスキング量補正部

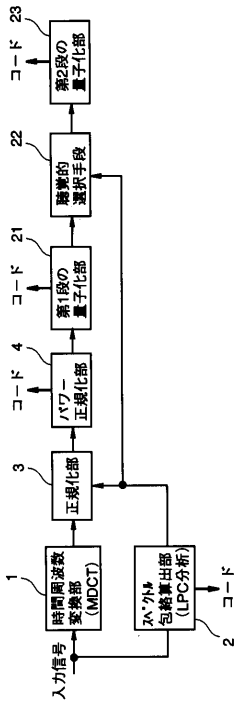
【図１】


時間周波数変換部
スペクトル包絡算出部
正規化部
パワー正規化部
聴覚重み付け計算部
第1段の量子化部
第2段の量子化部
第N段の量子化部
入力信号
コード

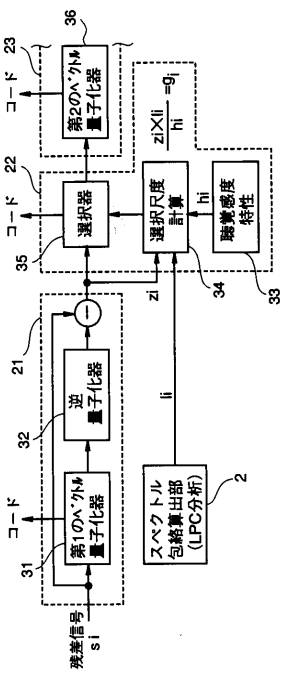
【図２】


時間周波数変換部
スペクトル包絡算出部
正規化部
パワー正規化部
第1の聴覚重み付け計算部
第2の聴覚重み付け計算部
第3の聴覚重み付け計算部
第1段の量子化部
第2段の量子化部
第N段の量子化部
入力信号
コード

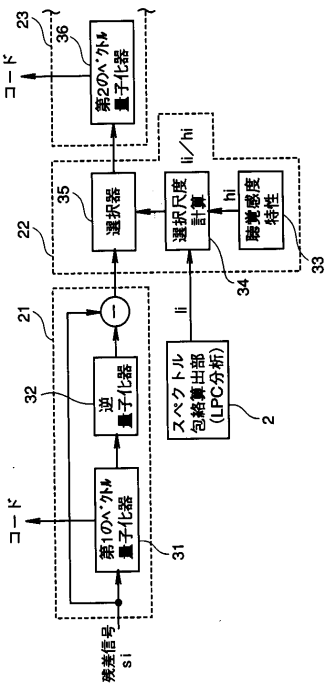
【図３】


時間周波数変換部（MDCT）
ﾍﾞｸﾄﾙ包絡算出部（LPC分析）
正規化部
パワー正規化部
第1段の量子化部
聴覚的選択手段
第2段の量子化部
入力信号
コード

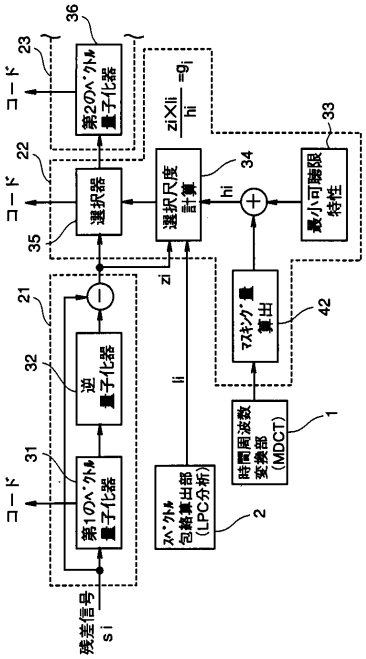
【図４】


第1のﾍﾞｸﾄﾙ量子化器
逆量子化器
第2のﾍﾞｸﾄﾙ量子化器
選択器
選択尺度計算
聴覚感度特性
スペクトル包絡算出部（LPC分析）
残差信号
コード

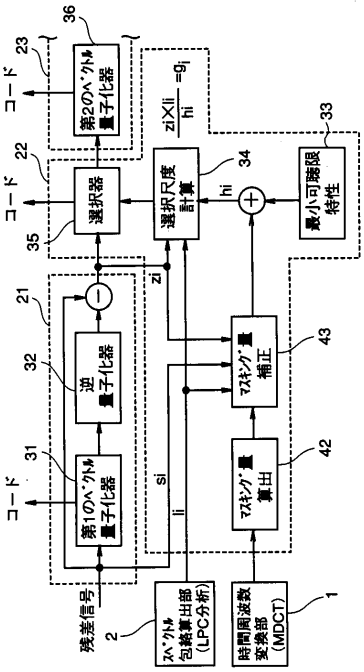
【図５】


コード
第2のベクトル量子化器
選択器
選択尺度計算
li/hi
hi
聴覚感度特性
li
スペクトル包絡算出部(LPC分析)
逆量子化器
第1のベクトル量子化器
残差信号 si
23
36
22
35
34
33
21
32
31
2

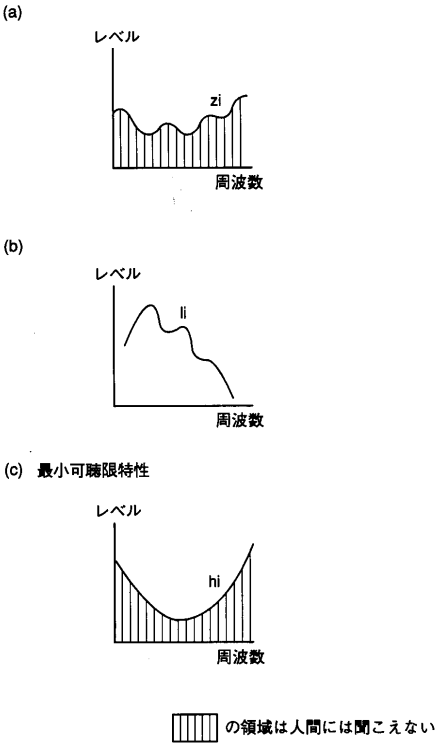
【図６】


コード
第2のベクトル量子化器
選択器
選択尺度計算
zi×li/hi =gi
hi
最小可聴限特性
zi
マスキング量算出
li
時間周波数変換部(MDCT)
スペクトル包絡算出部(LPC分析)
逆量子化器
第1のベクトル量子化器
残差信号 si
23
36
22
35
34
33
21
32
31
42
1
2

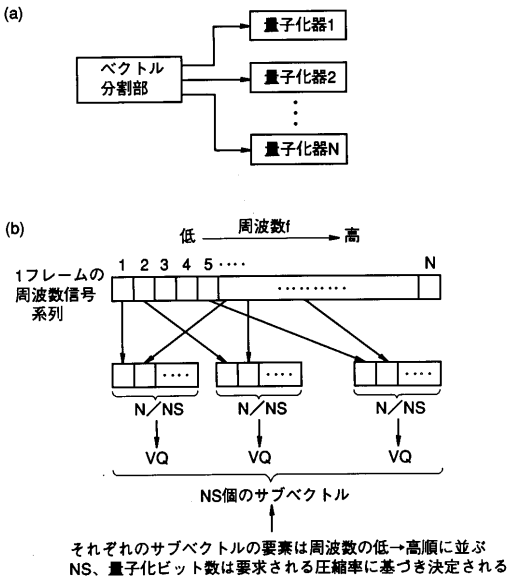
【図７】


コード
第2のベクトル量子化器
選択器
選択尺度計算
zi×li/hi =gi
hi
最小可聴限特性
zi
マスキング量補正
マスキング量算出
si
li
時間周波数変換部(MDCT)
スペクトル包絡算出部(LPC分析)
逆量子化器
第1のベクトル量子化器
残差信号
23
36
22
35
34
33
21
32
31
43
42
2
1

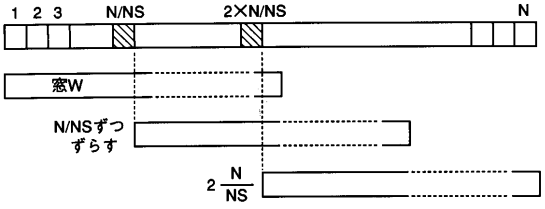
【図８】


(a)
レベル
zi
周波数
(b)
レベル
li
周波数
(c) 最小可聴限特性
レベル
hi
周波数
の領域は人間には聞こえない

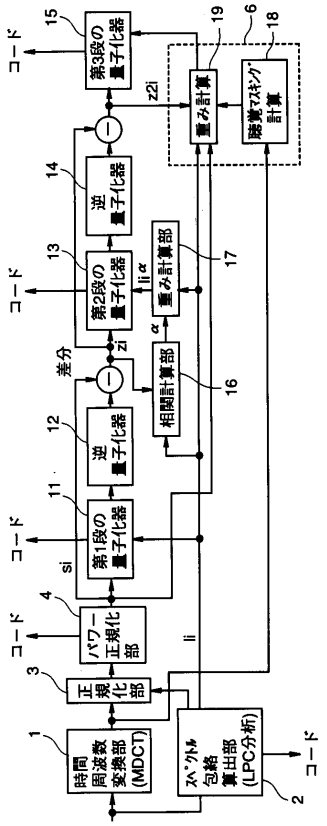
【図９】



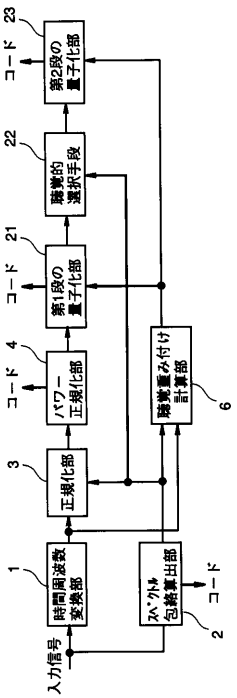
【図１０】



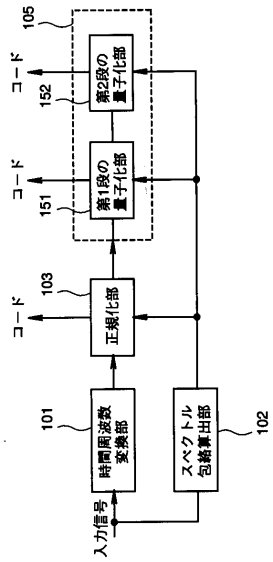
【図１１】



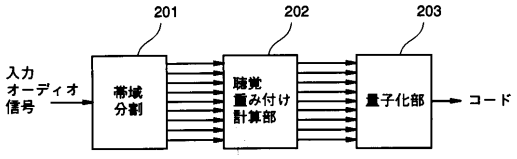
【図１２】



【図１３】


105
コード
152
第2段の量子化部
コード
151
第1段の量子化部
コード
103
正規化部
101
時間周波数変換部
スペクトル包絡算出部
102
入力信号

【図１４】


201
202
203
入力オーディオ信号
帯域分割
聴覚重み付け計算部
量子化部
コード

フロントページの続き

(72)発明者　津島　峰生
大阪府門真市大字門真１００６番地　松下電器産業株式会社内

審査官　山下　剛史

(56)参考文献　特許第３２４６７１５（ＪＰ，Ｂ２）
特開平４－３５５２７（ＪＰ，Ａ）
特開平６－１１８９９８（ＪＰ，Ａ）

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