

(19) 日本国特許庁 (JP)

(12) 特許公報 (B2)

(11) 特許番号

特許第3590342号

(P3590342)

(45) 発行日 平成16年11月17日 (2004.11.17)

(24) 登録日 平成16年8月27日 (2004.8.27)

(51) Int. Cl.⁷

F I

G 1 0 L 19/02

G 1 0 L 7/04

G

H 0 3 M 7/30

H 0 3 M 7/30

A

請求項の数 6

(全 15 頁)

(21) 出願番号 特願2000-318017 (P2000-318017)
 (22) 出願日 平成12年10月18日 (2000.10.18)
 (65) 公開番号 特開2002-123298 (P2002-123298A)
 (43) 公開日 平成14年4月26日 (2002.4.26)
 審査請求日 平成13年12月21日 (2001.12.21)

(73) 特許権者 000004226
 日本電信電話株式会社
 東京都千代田区大手町二丁目3番1号
 (74) 代理人 100123788
 弁理士 宮崎 昭夫
 (74) 代理人 100106297
 弁理士 伊藤 克博
 (72) 発明者 神 明夫
 東京都千代田区大手町二丁目3番1号 日
 日本電信電話株式会社内
 (72) 発明者 守谷 健弘
 東京都千代田区大手町二丁目3番1号 日
 日本電信電話株式会社内

最終頁に続く

(54) 【発明の名称】 信号符号化方法、装置及び信号符号化プログラムを記録した記録媒体

(57) 【特許請求の範囲】

【請求項 1】

入力信号に対して量子化を行う信号符号化方法であって、
 前記入力信号に対して時間軸 / 周波数軸変換を行って周波数軸上の係数列を得る工程と、
 前記係数列に基づいてスペクトル包絡を算出する工程と、
 前記スペクトル包絡における山・谷の位置を推定する工程と、
 前記スペクトル包絡において、推定された山・谷の位置へ情報量重み付けを行う工程と、
 情報量重み付けされたスペクトル包絡に基づいて、量子化用聴覚重みを算出する工程と、
 前記量子化用聴覚重みに基づいて量子化を行う工程と、
 を有し、
 前記推定する工程は、前記スペクトル包絡の1階微分値を求める工程と、前記1階微分値
の1階微分値を求めて2階微分値とする工程と、を有し、
前記情報量重み付けを行う工程は、山の付近を高く持ち上げ谷の付近を深く下げるか、ま
たは山の付近を低く下げ谷の付近を浅くなるように持ち上げるような重み関数を用いて山
・谷の位置へ重み付け操作を行う工程を有し、
 前記1階微分値が正の値から負の値に変化し、かつ、該変化の近傍で前記2階微分値が常
 に負の値ならば、その周波数を山の位置とし、前記1階微分値が負の値から正の値に変化
 し、かつ、該変化の近傍で前記2階微分値が常に正の値ならば、その周波数を谷の位置と
 する、信号符号化方法。

【請求項 2】

入力信号に対して量子化を行う信号符号化方法であって、
前記入力信号に対して時間軸 / 周波数軸変換を行って周波数軸上の係数列を得る工程と、

前記係数列に基づいてスペクトル包絡を算出する工程と、
前記スペクトル包絡における山・谷の位置を推定する工程と、
前記スペクトル包絡において、推定された山・谷の位置へ情報量重み付けを行う工程と、

情報量重み付けされたスペクトル包絡に基づいて、量子化用聴覚重みを算出する工程と、

前記量子化用聴覚重みに基づいて量子化を行う工程と、 10
を有し、

前記推定する工程は、前記スペクトル包絡の1階微分値を求める工程と、前記1階微分値の相加平均値を求める工程と、前記1階微分値の相加平均値の1階微分値を求めて2階微分値とする工程と、前記2階微分値の相加平均値を求める工程と、を有し、

前記情報量重み付けを行う工程は、山の付近を高く持ち上げ谷の付近を深く下げるか、または山の付近を低く下げ谷の付近を浅くなるように持ち上げるような重み関数を用いて山・谷の位置へ重み付け操作を行う工程を有し、

前記1階微分値の相加平均値が正の値から負の値に変化し、かつ、該変化の近傍で前記2階微分値の相加平均値が常に負の値ならば、その周波数を山の位置とし、前記1階微分値の相加平均値が負の値から正の値に変化し、かつ、該変化の近傍で前記2階微分値の相加平均値が常に正の値ならば、その周波数を谷の位置とする、信号符号化方法。 20

【請求項3】

入力信号に対して量子化を行う信号符号化装置であって、
前記入力信号に対して時間軸 / 周波数軸変換を行って周波数軸上の係数列を得る変換手段と、

前記係数列に基づいてスペクトル包絡を算出する包絡算出手段と、
前記スペクトル包絡における山・谷の位置を推定する山・谷推定手段と、
前記スペクトル包絡において、推定された山・谷の位置へ情報量重み付けを行う重み付け手段と、

情報量重み付けされたスペクトル包絡に基づいて、量子化用聴覚重みを算出する聴覚重み算出手段と、 30

前記量子化用聴覚重みに基づいて量子化を行う量子化手段と、
を有し、

前記重み付け手段は、山の付近を高く持ち上げ谷の付近を深く下げるか、または山の付近を低く下げ谷の付近を浅くなるように持ち上げるような重み関数を用いて山・谷の位置へ重み付け操作を行い、

前記山・谷推定手段は、前記スペクトル包絡の1階微分値を求め、前記1階微分値の1階微分値を求めて2階微分値とし、前記1階微分値が正の値から負の値に変化し、かつ、該変化の近傍で前記2階微分値が常に負の値ならば、その周波数を山の位置とし、前記1階微分値が負の値から正の値に変化し、かつ、該変化の近傍で前記2階微分値が常に正の値ならば、その周波数を谷の位置とする、信号符号化装置。 40

【請求項4】

入力信号に対して量子化を行う信号符号化装置であって、
前記入力信号に対して時間軸 / 周波数軸変換を行って周波数軸上の係数列を得る変換手段と、

前記係数列に基づいてスペクトル包絡を算出する包絡算出手段と、
前記スペクトル包絡における山・谷の位置を推定する山・谷推定手段と、
前記スペクトル包絡において、推定された山・谷の位置へ情報量重み付けを行う重み付け手段と、

情報量重み付けされたスペクトル包絡に基づいて、量子化用聴覚重みを算出する聴覚重み 50

算出手段と、

前記量子化用聴覚重みに基づいて量子化を行う量子化手段と、

を有し、

前記重み付け手段は、山の付近を高く持ち上げ谷の付近を深く下げるか、または山の付近を低く下げ谷の付近を浅くなるように持ち上げるような重み関数を用いて山・谷の位置へ重み付け操作を行い、

前記山・谷推定手段は、前記スペクトル包絡の1階微分値を求め、前記1階微分値の相加平均値を求め、前記1階微分値の相加平均値の1階微分値を求めて2階微分値とし、前記2階微分値の相加平均値を求め、前記1階微分値の相加平均値が正の値から負の値に変化し、かつ、該変化の近傍で前記2階微分値の相加平均値が常に負の値ならば、その周波数を山の位置とし、前記1階微分値の相加平均値が負の値から正の値に変化し、かつ、該変化の近傍で前記2階微分値の相加平均値が常に正の値ならば、その周波数を谷の位置とする、信号符号化装置。

10

【請求項5】

計算機が読取り可能な記録媒体であって、

前記計算機に、

前記入力信号に対して時間軸/周波数軸変換を行って周波数軸上の係数列を得る工程と、

前記係数列に基づいてスペクトル包絡を算出する工程と、

前記スペクトル包絡における山・谷の位置を推定する工程と、

前記スペクトル包絡において、推定された山・谷の位置へ情報量重み付けを行う工程と、

20

情報量重み付けされたスペクトル包絡に基づいて、量子化用聴覚重みを算出する工程と、

前記量子化用聴覚重みに基づいて量子化を行う工程と、

を実行させ、

前記推定する工程は、前記スペクトル包絡の1階微分値を求める工程と、前記1階微分値の1階微分値を求めて2階微分値とする工程と、を有し、

前記情報量重み付けを行う工程は、山の付近を高く持ち上げ谷の付近を深く下げるか、または山の付近を低く下げ谷の付近を浅くなるように持ち上げるような重み関数を用いて山・谷の位置へ重み付け操作を行う工程を有し、

前記1階微分値が正の値から負の値に変化し、かつ、該変化の近傍で前記2階微分値が常に負の値ならば、その周波数を山の位置とし、前記1階微分値が負の値から正の値に変化し、かつ、該変化の近傍で前記2階微分値が常に正の値ならば、その周波数を谷の位置とする、信号符号化プログラムを記録した記録媒体。

30

【請求項6】

計算機が読取り可能な記録媒体であって、

前記計算機に、

前記入力信号に対して時間軸/周波数軸変換を行って周波数軸上の係数列を得る工程と、

前記係数列に基づいてスペクトル包絡を算出する工程と、

前記スペクトル包絡における山・谷の位置を推定する工程と、

前記スペクトル包絡において、推定された山・谷の位置へ情報量重み付けを行う工程と、

40

情報量重み付けされたスペクトル包絡に基づいて、量子化用聴覚重みを算出する工程と、

前記量子化用聴覚重みに基づいて量子化を行う工程と、

を実行させ、

前記推定する工程は、前記スペクトル包絡の1階微分値を求める工程と、前記1階微分値の相加平均値を求める工程と、前記1階微分値の相加平均値の1階微分値を求めて2階微分値とする工程と、前記2階微分値の相加平均値を求める工程と、を有し、

前記情報量重み付けを行う工程は、山の付近を高く持ち上げ谷の付近を深く下げるか、または山の付近を低く下げ谷の付近を浅くなるように持ち上げるような重み関数を用いて山

50

・谷の位置へ重み付け操作を行う工程を有し、

前記1階微分値の相加平均値が正の値から負の値に変化し、かつ、該変化の近傍で前記2階微分値の相加平均値が常に負の値ならば、その周波数を山の位置とし、前記1階微分値の相加平均値が負の値から正の値に変化し、かつ、該変化の近傍で前記2階微分値の相加平均値が常に正の値ならば、その周波数を谷の位置とする、信号符号化プログラムを記録した記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、入力信号を時間軸/周波数軸変換して量子化を行う信号符号化方法及び装置に関し、特に、符号化に際して発生する量子化誤差を、人間の耳が知覚しづらいように変形するための聴覚マスキング方法と、この聴覚マスキング方法による信号符号化装置に関する。

10

【0002】

【従来の技術】

音声・楽音を符号化する従来の信号符号化方法における聴覚マスキング方法としては、入力信号を時間軸上または時間軸/周波数軸変換した上で、線形予測分析方法等によりその入力信号のスペクトル包絡曲線を推定し、その推定された曲線に妥当な変形操作を加えることによってマスキング曲線を求めて聴覚マスキングを行なうという方法があった。あるいは、入力信号を時間軸/周波数軸変換した信号から直接、スペクトル包絡曲線を求め、この曲線に妥当な変形操作を加えることによってマスキング曲線を求めて、聴覚マスキングによる量子化を行なう方法もあった。

20

【0003】

【発明が解決しようとする課題】

聴覚マスキング方法では、周波数軸上でのマスキングとして、スペクトル包絡曲線の谷付近の量子化雑音を減らし、その代わりにスペクトル包絡曲線の山付近の量子化雑音を増加させるようなノイズシェイピングを行うことによって、人間の耳には量子化雑音が聞こえにくいようにすることができる。ここで、上述したような従来法では、スペクトル包絡における山と谷の推定位置が不正確となる場合があったため、ノイズシェイピングが適切に行われずに、結果として符号化再生音の音質が悪い場合があった。

30

【0004】

そこで本発明の目的は、スペクトル包絡曲線における山と谷の位置を正確に推定することができ、これによって精度の高い聴覚マスキング方法を実行できる信号符号化方法及び装置を提供することにある。

【0005】

【課題を解決するための手段】

本発明は、聴感ベースでの歪みが最小となるように量子化できる信号符号化を実現するためのものであって、上述した課題を解決するために、スペクトル包絡曲線の山と谷の位置を正確に推定し、正確に推定した山と谷の位置から適切なノイズシェイピングを行う手法を取る。スペクトル包絡曲線の山と谷の位置推定は、時間軸/周波数軸変換した信号の正確なスペクトル包絡曲線から必要に応じて、微細な凹凸を取り除き、さらに必要に応じて1階微分、2階微分を求めて、これらの微分値または、微分値の相加平均値から、山と谷の正確な位置を決定する。こうして得られた山と谷の位置において適切な重みづけを行ない、効果的なノイズシェイピングを実現する。

40

【0006】

【発明の実施の形態】

次に、本発明の好ましい実施の形態について、図面を参照して説明する。図1は本発明の実施の一形態の信号符号化装置の構成を示すブロック図である。

【0007】

この信号符号化装置は、典型的には音声信号あるいは楽音信号である時系列の入力信号x

50

(t) に対して時間軸 / 周波数軸変換 (T / F 変換) を施して周波数軸上の信号列 $X(n)$ を得る T / F 変換部 11 と、この信号列 $X(n)$ に対してベクトル量子化 (VQ) 及びスカラー量子化 (SQ) を施して量子化インデックスを得る量子化部 12 を備えている。ここで T / F 変換部 11 は、例えば、MDC T (modified discrete cosine transform; 変形離散コサイン変換) などの変換を実行し、 $X(n)$ はこの変換によって得られた変換係数列などを指す。さらにこの信号符号化装置では、どの周波数帯域にどれだけの情報量を配分するのかを決定するための“聴覚重み”を算出し、量子化部 11 での量子化に際し、人間の耳に量子化雑音が聞こえ難いようにこの聴覚重みに基づく聴覚重み付け量子化が行われるようになっている。聴覚重みの算出のために、この信号符号化装置は、信号列 $X(n)$ に基づいてスペクトル包絡を算出する包絡算出部 13 と、算出されたスペクトル包絡に基づいてスペクトルの山と谷の位置を推定する山・谷推定部 14 と、推定されたスペクトルの山と谷の位置に基づき、情報量の配分が「山の位置で特に小さく」かつ「谷の位置で特に大きく」なるように、山の付近と谷の付近において適切な重み付けを行う重み付け部 15 と、“聴覚重み”として量子化部 12 に出力する聴覚重み算出部 16 と、を備えている。ここで“聴覚重み”の原形としては、スペクトル包絡の逆数を用いている。

【0008】

なお、山、谷については、横軸を周波数軸として信号列 $X(n)$ をプロットし、ならした (平滑化した) ときに、周囲に比べて信号列の値が大きいところを山と称し、周囲に比べて値が小さいところを谷と称している。後述するように、平滑化は、例えばある区間長 (平均区間長ともいう) での相加平均を算出する (その区間長による移動平均を算出する) ことによって行われているが、このとき、その区間長を変化させることにより、微細な山・谷、やや微細な山・谷、大まかな山・谷の位置などが推定されることになる。ここで相加平均とは、1 フレーム内のスペクトルを周波数区間で平滑化するためのものである。本発明では、平滑化の度合いが異なる山・谷の位置の推定を組み合わせることにより、より精度の高い聴覚マスキングを可能にしている。

【0009】

次に、この信号符号化装置の動作を説明する。

【0010】

時系列の信号として入力する時系列の入力信号 $x(t)$ は、T / F 変換部 11 によって周波数軸上の信号列 $X(n)$ に変換される。この信号列 $X(n)$ は、ベクトル量子化及びスカラー量子化のために量子化部 12 に供給されるとともに、そのスペクトル包絡を算出するために、包絡算出部 13 にも送られる。包絡算出部 13 は、信号列 $X(n)$ のスペクトル包絡を算出し、山・谷推定部 14 は、算出されたスペクトル包絡に基づいて、スペクトルにおける山と谷の位置を推定し、推定した位置を重み付け部 15 に出力する。重み付け部 15 は、包絡算出部 13 において得られたスペクトル包絡の逆数に基づいて、スペクトルの山と谷の位置においてそれぞれ、情報量の配分が「山の位置で特に小さく」、「谷の位置で特に大きく」なるように、山の付近と谷の付近において、適切な情報量重み付けを行う。具体的には、山の付近を高く持ち上げかつ谷の付近を深く下げるか、あるいは、山の付近を低く下げ谷の付近を浅くなるように持ち上げるような重み関数を用いて、山・谷の位置へ重み付け操作を行う。重み付け部 15 には、包絡算出部 13 からスペクトル包絡曲線が供給されており、重み付け操作が施されたスペクトル包絡曲線が重み付け部 15 から聴覚重み算出部 16 に供給される。

【0011】

聴覚重み算出部 16 は、重み付けされたスペクトル包絡曲線に基づいて量子化用聴覚重みを算出してそれを量子化部 12 に向けて出力する。その結果、量子化部 13 は、供給された量子化用聴覚重みを使用して、T / F 変換部 11 からの信号列 $X(n)$ に対するベクトル量子化及びスカラー量子化を実行する。これにより、量子化部 13 から、精度の高い聴覚マスキングがなされた量子化インデックス (出力インデックス) が出力される。

【0012】

10

20

30

40

50

以上、この実施の形態の信号符号化装置の基本的動作を説明したが、本発明では、聴覚重み付けの方法として、上述した重み付けの方法と、従来から一般的に用いられている線形予測分析法等によりスペクトル包絡を予測し包絡曲線の山と谷をべき乗演算によりなまらせ重みとする方法とを併用してもよい。

【0013】

次に、この実施の形態における重み付けの過程を詳細を説明する。

【0014】

図2は、スペクトルの山・谷へ重み付けを行う過程を示すブロック図である。ここでは、スペクトル包絡算出部13において得られたスペクトル包絡曲線から、山・谷推定部14において、スペクトルの微細な山・谷の周波数位置を推定し、次にやや微細な山・谷の周波数位置を推定し、というように、この手順を必要な回数だけ繰り返し、最後に、スペクトルの大まかな山・谷の周波数位置を推定する。重み付け部15は、これらの推定された山と谷の付近に対して、各々、妥当な重み関数によって重み付け操作を行う。

【0015】

図3は、包絡算出部13における処理の詳細を示すブロック図である。包絡算出部13は、周波数領域の信号列 $X(n)$ に対して相加平均処理を施すことにより、スペクトル包絡曲線を得るものである。図において、相加平均(1)から相加平均(k)までは、それぞれ、区間長が異なる移動平均区間における相加平均である。ここでは、信号列 $X(n)$ に対し、まず、第1の相加平均(1)が適用され、その結果 $Y_1(n)$ に対して第2の相加平均(2)が適用され、さらにその結果 $Y_2(n)$ に対して第3の相加平均(3)が適用され、というようにして、k回の相加平均を順次行うようにしている。ここでkは1以上の整数の定数である。このようにして得られた各相加平均の結果 $Y_1(n)$ 、 $Y_2(n)$ 、 \dots 、 $Y_k(n)$ は、それぞれ山・谷推定部14に送られる。各回の相加平均での区間長は、各々の用途に応じて決定されるものであるが、主として、相加平均(1)では平均区間長を短くして微細な山と谷の位置を検出し、相加平均(2)では相加平均(1)よりも平均区間長を長くして大まかな山と谷の位置を検出する。以下、相加平均(k)まで同様の操作とし、各回の相加平均での平均区間長を徐々に長くして行くとよい。

また、前述した“相加平均(k)”の演算は、必要に応じて、平均区間長を変えて複数回実施してもよい。

【0017】

次に、山・谷推定部14での処理を説明する。図4は、山・谷推定部14での処理を説明するブロック図である。

【0018】

山・谷推定部14は、包絡算出部13からの各回の相加平均によるスペクトル包絡を表す係数列 $Y_1(n)$ 、 $Y_2(n)$ 、 \dots 、 $Y_k(n)$ を入力として、係数列ごとに、以下のようにして山と谷の位置を推定する。すなわち、入力した係数列 $Y_j(n)$ ($1 \leq j \leq k$)をまずnで微分して系列 $Y'_j(n)$ を求め、この系列 $Y'_j(n)$ に対して適切な区間で相加平均をとり、微細な変動成分を取り除いた系列

【0019】

【外1】

$$\overline{Y'_j}(n)$$

【0020】

を求める。さらにこれをnで再び微分して系列 $Y''_j(n)$ を求め、この系列 $Y''_j(n)$ の微細な変動成分を取り除いた系列

【0021】

【外2】

$$\overline{Y''_j}(n)$$

【0022】

を求める。そして、図4中に式で示したように、これらの値の正負からスペクトル包絡曲線の山と谷の位置を推定する。また、前述した、微細な変動成分を取り除くための“相加平均”の演算は、必要に応じて、平均区間長を変えて複数回実施してもよいし、これを実施しなくてもよい。

【0023】

図5は、以上のようにして係数列 $X(n)$ からスペクトル包絡の山と谷が検出された様子を例示する図である。ここでは、 $k=2$ 、すなわち包絡算出部13において2段階に相加平均を求める場合を示している。この図において、平均を取る前の係数列 $X(n)$ の絶対値 $|X(n)|$ を1、相加平均(1)による係数列 $Y_1(n)$ における絶対値 $|Y_1(n)|$ を2、相加平均(2)による係数列 $Y_2(n)$ における絶対値 $|Y_2(n)|$ を3とする。相加平均(1)から推定した山の位置を m_1, m_2, \dots, m_{12} 、谷の位置を v_1, v_2, \dots, v_{11} で表し、相加平均(2)から推定した山の位置を M_1, M_2, M_3 、谷の位置を V_1, V_2 で表している。ここでは、相加平均(1)での区間長よりも相加平均(2)での区間長を長くしており、2が微細な山・谷の周波数位置に相当し、3が大まかな山・谷の周波数位置に相当する。

【0024】

次に、このようにして、複数種類の山・谷の周波数位置が求められたとして、どのように情報量の重み付けを行うかを説明する。図6は、スペクトル包絡曲線の山・谷付近に情報量の重み付けを行った例を示す図である。ここでは、説明を分かりやすくするために、おおまかな波形を使って説明を行う。

【0025】

図6において、あらかじめ推定されたスペクトル包絡曲線1 ($|Y_2(n)|$)の逆数2 ($1/|Y_2(n)|$)を聴覚重みの原形とし、これの山と谷の推定位置付近において、重み関数を使って重み付けを行う。この図の例では、重み付け関数4を2に乗算することによって、山と谷の位置で情報量を補正した聴覚重み3 (W_L)を作成している。重み付け関数4及び5としては、種々の形のものが可能であるが、ここでは、一例として、重み付けを行う区間長が2t、山の中心で0.5倍、山の端で1.0倍、谷の中心で2.0倍、谷の端で1.0倍となるような直線関数による重み付けを行った結果を3として示している。図6から分かるように、山と谷の正確な位置を推定し、谷の付近に情報量を多くし、山の付近に情報量を少なく割り当てる重みを作成することができる。

【0026】

ここでtの値は、例えば、ピッチ周波数を表す山・谷の構造に重み付けしたい場合には100~200Hz、ホルマント周波数を表す山・谷の構造に重み付けしたい場合には300~600Hz程度とすることが好ましい。

【0027】

実際には、スペクトル包絡の“微細な曲線”と“おおまかな曲線”の各々の山・谷の付近において、前述した方法により重み付けを行う。例えば、図5に示すようにスペクトル包絡の“微細な曲線”と“おおまかな曲線”の各々について山と谷の位置が推定されている場合には、微細構造を表すスペクトル包絡2の逆数 $1/|Y_1(n)|$ を聴覚重みの原形とし、この包絡曲線の山と谷の位置 $m_1, v_1, m_2, v_2, \dots$ の付近において、図6と同様にして聴覚重みの原形である $1/|Y_1(n)|$ に対して適切な重み付けを行い、さらに、おおまかなスペクトル構造を表す曲線3の山と谷の位置 $M_1, V_1, M_2, V_2, \dots$ の付近において、同様に聴覚重みの原形である $1/|Y_1(n)|$ に対して適切な重み付けを行う。

【0028】

山に対する重み付け関数及び谷に対する重み付け関数としては、各種のものが考えられる。図7は、そうした重み付け関数を例示するものである。

【0029】

10

20

30

40

50

図7中、(a)、(b)はいずれも山に対する重み付け関数の例を示しており、(a)は直線により構成されたもの、(b)は放物線により構成されたものである。いずれも山の中心 $n = M$ の両側に t ずつ、合計 $2t$ の区間を重み付け区間としている。重み付け関数の値は、重み付け区間の両端($M \pm t$)においては 1.0 であるものとする。また、山の中心 $n = M$ における重みの値は、通常、 $0 < < 1.0$ における妥当な定数とすればよい。同様に図7中、(c)、(d)は、谷に対する重み付け関数の例を示しており、(c)は直線により構成されたもの、(d)は放物線により構成されたものである。山の場合と同様に、谷に対する重み付け関数も、その値は、重み付け区間の両端($V \pm t$)においては 1.0 である。また、谷の中心 $n = V$ における重みの値は、通常、 > 1.0 における妥当な定数を使用する。しかしながら、場合によっては、 > 1.0 、 $0 < < 1.0$ とすると効果的なこともある。

10

【0030】

このようにして聴覚重み付けを行った場合に、量子化雑音は図8に示すように変形される。すなわち、聴覚重み付けを行わない場合には、量子化ノイズは周波数によらずに一定であると考えられるが(図中2)、入力信号のスペクトル包絡が図中1に示すようなものであるとすると、上述した聴覚重み付けを行うことにより、ノイズは、図中3に示すようにその周波数特性が変形され、入力信号のスペクトル特性である1に隠されて、聴感的に聞こえ難くなる。

【0031】

したがって、従来法よりも精度の高い聴覚マスキングが行なえ、高品質な符号化を行なうことが可能となる。

20

【0032】

次に、上述した本発明の信号符号化方法を一般的な変換符号化方式の聴覚重み付けに適用した例を説明する。図9はそのような聴覚重み付けを行う信号符号化装置の構成を示している。

【0033】

図9に示す信号符号化装置は、入力信号に対してMDC Tを施すMDC T変換部31と、MDC T後の信号のスペクトルを平坦化するスペクトル平坦化部32と、平坦化後のスペクトルに基づいてフレームゲインを正規化し量子化した後、ゲインインデックスを出力するフレームゲイン正規化部33と、正規化されたフレームゲインに基づいて残差成分を量子化(ベクトル量子化あるいはスカラー量子化)し、量子化インデックスを出力する残差成分量子化部34と、MDC T後の信号のスペクトルからスペクトル包絡を推定するスペクトル包絡推定部35と、残差成分量子化部34での量子化に際して情報量重み付けを行うために、推定されたスペクトル包絡から聴覚重みを計算する聴覚重み計算部36と、推定されたスペクトル包絡に基づいてスペクトル情報を量子化しスペクトルインデックスを出力するスペクトル情報量子化部37とを備えている。この信号符号化装置では、MDC T変換部31が図1に示した信号符号化装置のT/F変換部11に相当し、また、スペクトル包絡推定部35は、図1に示す装置の包絡算出部13及び山・谷推定部14で構成され、聴覚重み計算部36は、図1に示す装置の重み付け部15及び聴覚重み算出部16で構成される。

30

40

【0034】

本発明の信号符号化方法により、分析フレーム内におけるスペクトルの山と谷を正確かつ細かに分析し、その形に合わせて量子化の際に精度の高い聴覚マスキングを行うことができる。この聴覚マスキングは、ベクトル量子化や、サブバンドスカラー量子化に対して適用できる。

【0035】

さらに図10は、特開平8-44399号公報に開示される符号器及び復号器に本発明の聴覚重み付けを適用した例を示している。図10に示されるものにおいて、符号器110は、入力端子111に与えられた入力信号をフレームに分割するフレーム分割部114と、フレームに時間窓を描ける時間窓掛部115と、時間窓が掛けられたフレームにN次の

50

MDC Tを施すMDC T部 1 1 6 と、時間窓が掛けられたフレームに対して線形予測分析を行い予測係数を出力する線形予測分析部 1 1 7 と、予測係数を量子化してインデックス I_p を得る量子化部 1 1 8 と、予測係数のスペクトラム概形を求めるスペクトラム概形計算部 1 2 1 と、MDC T部 1 1 6 からのスペクトラム振幅をスペクトラム概形により正規化し残差係数 $R(F)$ を得る正規化部 1 2 2 と、残差係数概形 $E_R(F)$ を計算する残差概形計算部 1 2 3 と、残差係数概形及びスペクトラム概形に基づいて重み付け係数(ベクトル W) を計算する重み計算部 1 2 4 と、重み付け係数に基づいて量子化しインデックス I_m と量子化小系列ベクトル $C(m)$ を出力する量子化部 1 2 5 と、残差係数 $R(F)$ を残差係数概形 $E_R(F)$ で正規化して微細構造係数を得る残差係数正規化部 1 2 6 と、現フレームの微細構造係数を正規化し正規化微細構造係数 $X(F)$ として量子化部 1 2 5 に与えるとともにインデックス I_G を出力するパワー正規化部 1 2 7 と、量子化小系列ベクトル $C(m)$ を逆正規化し量子化残差係数 $R_Q(F)$ を残差概形計算部 1 2 3 に出力する逆正規化部 1 3 1 とを備えている。

10

【0036】

符号器 1 1 0 において本発明に基づく聴覚重み付けを行うためには、スペクトラム概形計算部 1 2 1 において、従来法に加えてさらに図 1 に示した信号符号化装置の包絡算出部 1 3 及び山・谷推定部 1 4 での処理と同様の処理を行わせ、その結果に基づいて、重み計算部 1 2 4 においては、従来法に加えてさらに図 1 に示した装置の重み付け部 1 5 及び聴覚重み算出部 1 6 での処理と同様の処理を行い、得られた量子化用聴覚重みを量子化部 1 2 5 に供給するようにすればよい。

20

【0037】

これに対して復号器 1 5 0 は、インデックス I_m から正規化微細構造係数を再生する再生部 1 5 1 と、インデックス I_G から正規化ゲインを再生する正規化ゲイン再生部 1 5 2 と、正規化微細構造係数を正規化ゲインにより逆正規化して微細構造係数を得るパワー逆正規化部 1 5 3 と、微細構造係数を残差概形 E_R で逆正規化して残差係数 $R(F)$ を再生する残差逆正規化部 1 5 4 と、残差概形 E_R を計算する残差概形計算部 1 5 5 と、インデックス I_p から線形予測係数を再生しスペクトラム概形を計算する再生・スペクトラム概形計算部 1 5 6 と、スペクトラム概形を残差係数 $R(F)$ で逆正規化し周波数領域係数を再生する逆正規化部 1 5 7 と、周波数領域係数にフレームごとに逆MDC Tを施し時間領域信号を得る逆MDC T部 1 5 8 と、時間領域信号にフレームごとに時間窓を掛ける窓掛部 1 5 9 と、窓掛け出力に対してフレーム重ね合わせを行い再生音響信号を得てこれを出力端子 1 9 1 に出力するフレーム重ね合わせ部 1 6 1 と、を備えている。

30

【0038】

なお、図 1 0 に示す符号器 1 1 0 においては、逆正規化部 1 3 1 を設けることなく、正規化部 1 2 2 の出力のみに基づいて残差概形計算部 1 2 3 が残差係数概形 $E_R(F)$ とインデックス I_Q を算出するようにすることが可能であり、この場合、復号器 1 5 0 において残差概形計算部 1 5 5 はインデックス I_Q に基づいて残差概形 E_R を計算する。

【0039】

次に、時間領域の符号化方式である CELP (Code - Excited Linear Prediction) 符号化の聴覚マスクングに本発明を適用した例を説明する。CELP 符号化では、時間領域で聴覚マスクングが行われるため、本発明に基づく聴覚重み付けを周波数領域で適用し、得られた聴覚重みを時間領域に戻してから量子化に適用する。図 1 1 はそのような符号化を行う信号符号化装置の構成を示すブロック図である。

40

【0040】

図 1 1 に示す装置は、入力信号に対して FFT (高速フーリエ変換) を施す FFT 部 3 8 と、FFT 部の出力(周波数領域の信号列)に基づき、スペクトル包絡を推定するスペクトル包絡推定部 3 5 と、推定されたスペクトル包絡から聴覚重みを計算する聴覚重み計算部 3 6 と、聴覚重みを時間領域に戻すための逆 FFT 部 3 9 と、時間領域の聴覚重みに基づいて入力信号の CELP 符号化を行い、インデックスを出力する CELP 符号化部 4 0 とを備えている。この信号符号化装置においては、FFT 部 3 8 が図 1 に示した信号符号

50

化装置のT/F変換部11に相当し、また、スペクトル包絡推定部35は、図1に示す装置の包絡算出部13及び山・谷推定部14で構成され、聴覚重み計算部36は、図1に示す装置の重み付け部15及び聴覚重み算出部16で構成される。

【0041】

さらに図12は、特開平6-282298号公報の図1に開示される音声符号化装置に本発明の聴覚重み付けを適用した例を示している。図12に示される音声符号化装置は、入力端子201を介して入力した音声信号をフレームに分割して線形予測分析を行い、予測係数を決定する予測係数決定部202と、合成フィルタ203と、予測係数を量子化して合成フィルタ203に予測係数を設定する予測係数量子化部204と、複数のピッチ周期ベクトルを記憶する適応符号帳217と、複数の雑音波形ベクトルを記憶する雑音符号帳218と、適応符号帳217から選択されたピッチ周期ベクトルに利得を加える利得部219a及び雑音符号帳218から選択された雑音波形ベクトルに利得を加える利得部219bとを有する利得符号帳219と、利得部219bの過去の出力パワーに基づいて次の雑音波形ベクトルの予測利得を得る予測利得決定部215と、利得部219bの入力側に設けられ選択された雑音波形ベクトルにこの予測利得を加える予測利得部216と、利得部219a、219bからの出力ベクトルを加算して駆動ベクトルとして合成フィルタ203に供給する加算器209と、入力音声ベクトル(入力信号)から合成フィルタ203の出力(合成音声ベクトル)を減算して歪データとして出力する減算器211と、歪データに対して聴覚重み付けを行う聴覚重み付けフィルタ220と、聴覚重み付け後の歪データに基づいて歪パワーを計算し、歪パワーが最小になるように各符号帳217~219での選択を行う歪パワー計算部212と、符号を出力する符号出力部213と、を備えている。

【0042】

この音声符号化装置において本発明に基づく聴覚重み付けを行う場合には、上述の図11に示した信号符号化装置をここでの聴覚重み付けフィルタ220として、または聴覚重み付けフィルタ220と併用して用いればよい。これにより、歪データに対して、本発明に基づく聴覚重み付けがなされることになる。さらに、ここでは図面を用いては説明しないが、特開平6-282298号公報の図2に開示される音声符号化装置においても、その聴覚重み付けフィルタとして、図11に示した信号符号化装置を上述のように変形したものを使用することができる。

【0043】

以上説明した本発明に基づく信号符号化方法及び装置は、それを実現するための計算機プログラムを、計算機(コンピュータ)に読み込ませ、そのプログラムを実行させることによっても実現できる。信号符号化を行うためのプログラムは、磁気テープやCD-ROMなどの記録媒体によって、あるいは、ネットワークを介して、計算機に読み込まれる。図13は、上述の信号符号化方法を実行する計算機の構成を示すブロック図である。

【0044】

この計算機は、中央処理装置(CPU)21と、プログラムやデータを格納するためのハードディスク装置22と、主メモリ23と、キーボードやマウス、マイクロホンなどの入力装置24と、CRTやスピーカなどの表示装置25と、磁気テープやCD-ROM等の記録媒体27を読み取る読み取り装置26と、ネットワークに接続した通信インタフェース28とから構成されている。ハードディスク装置22、主メモリ23、入力装置24、表示装置25、読み取り装置26及び通信インタフェース28は、いずれも中央処理装置21に接続している。ハードディスク装置22の代わりに、フラッシュROMなどの不揮発性半導体記憶装置を用いてもよい。この計算機は、信号符号化を行うためのプログラムを格納した記録媒体27を読み取り装置26に装着し、記録媒体27からプログラムを読み出してハードディスク装置22に格納し、ハードディスク装置22に格納されたプログラムを中央処理装置21が実行することにより、信号符号化装置として機能するようになる。もちろん、ネットワークを介して、信号符号化を行うためのプログラムをこの計算機にダウンロードするようにしてもよい。

【 0 0 4 5 】

【 発明の効果 】

以上説明したように、本発明によれば、音声・楽音信号を符号化する際に、従来法よりも精度の高い聴覚マスキングが行なえ、高品質な符号化を行なうことが可能となる。具体的には、例えばMDC T変換等によって時系列信号を周波数領域の係数列に変換して量子化する際に、本発明を用いれば、人間の聴覚マスキング特性を利用して、量子化誤差を知覚し難いように、周波数軸上で従来法よりも高精度で配分することが可能となる。

【 図面の簡単な説明 】

【 図 1 】 本発明の実施の一形態の信号符号化装置の構成を示すブロック図である。

【 図 2 】 スペクトルの山・谷へ重み付けを行う過程を示すブロック図である。

10

【 図 3 】 包絡算出部における処理の詳細を示すブロック図である。

【 図 4 】 山・谷推定部における処理の詳細を示すブロック図である。

【 図 5 】 山・谷推定部により検出された、スペクトラム包絡における山及び谷の様子の一例を示す図である。

【 図 6 】 スペクトル包絡の山・谷付近に重み付けを行った例を示す図である。

【 図 7 】 (a) ~ (d) は、山・谷付近への重み付け関数の例を示す図である。

【 図 8 】 聴覚重み付け処理によって量子化雑音がスペクトル包絡にマスキングされる様子を示した図である。

【 図 9 】 本発明に基づく信号符号化装置の構成の一例を示すブロック図である。

【 図 1 0 】 本発明に基づく聴覚重み付けが適用される符号器及び復号器の構成の一例を示すブロック図である。

20

【 図 1 1 】 信号符号化装置の構成の一例を示すブロック図である。

【 図 1 2 】 信号符号化装置の構成の一例を示すブロック図である。

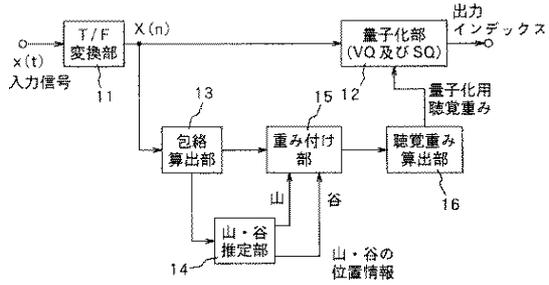
【 図 1 3 】 信号符号化装置を構成するために使用される計算機システムの一例を示すブロック図である。

【 符号の説明 】

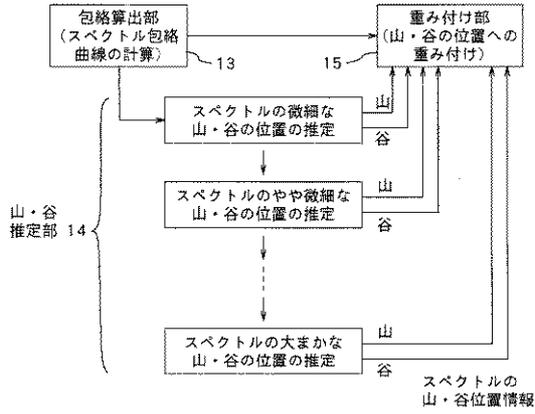
- 1 1 T / F 変換部
- 1 2 量子化部
- 1 3 包絡算出部
- 1 4 山・谷推定部
- 1 5 重み付け部
- 1 6 聴覚重み算出部

30

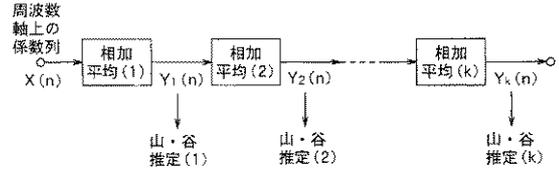
【図1】



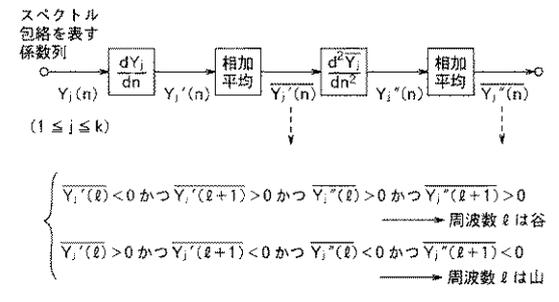
【図2】



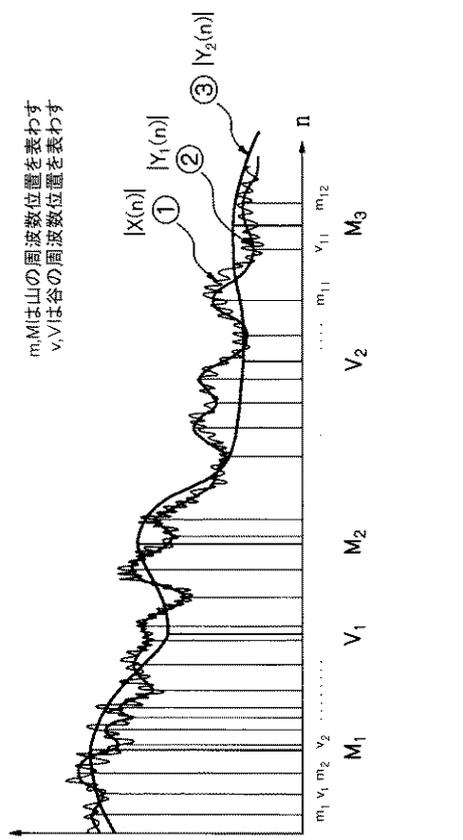
【図3】



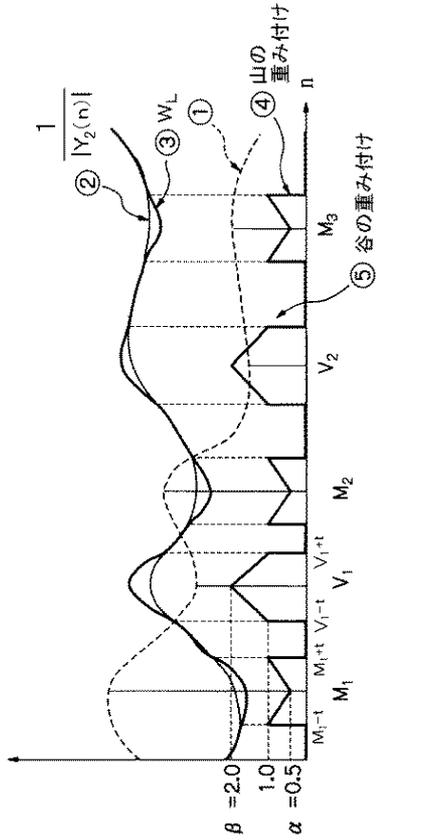
【図4】



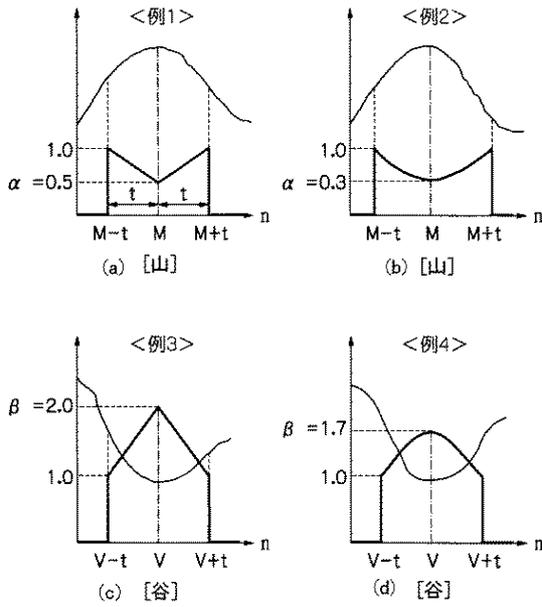
【図5】



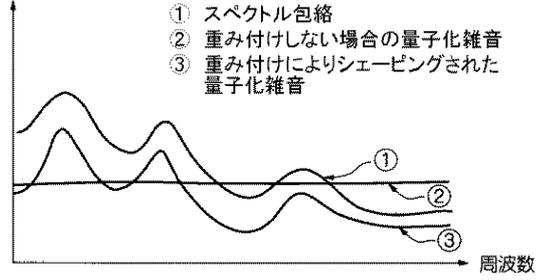
【図6】



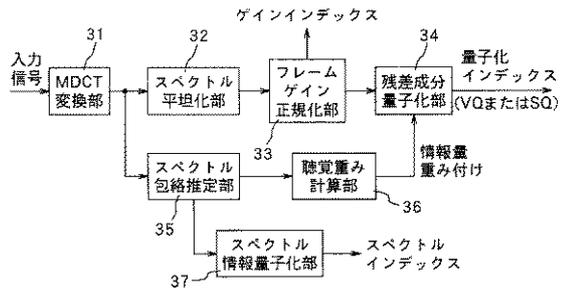
【図7】



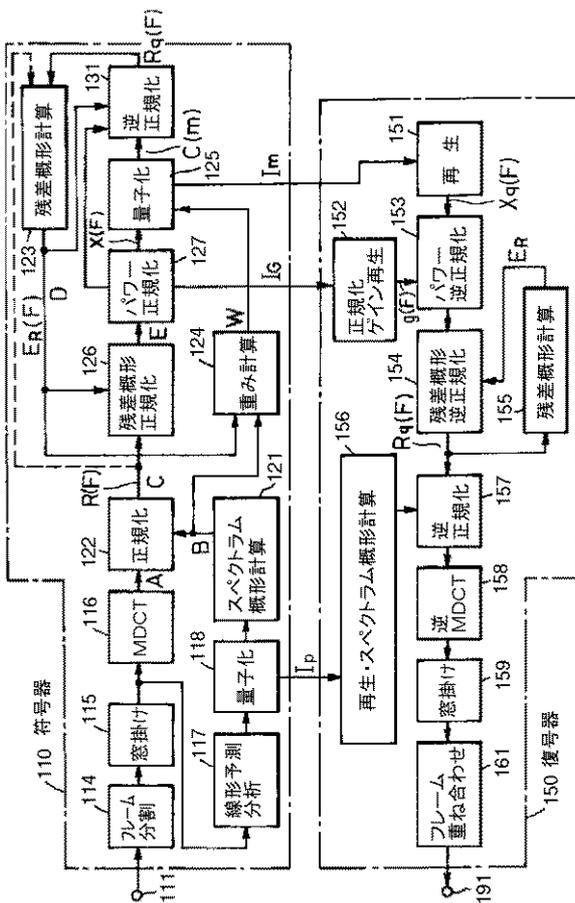
【図8】



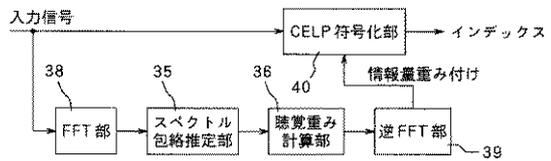
【図9】



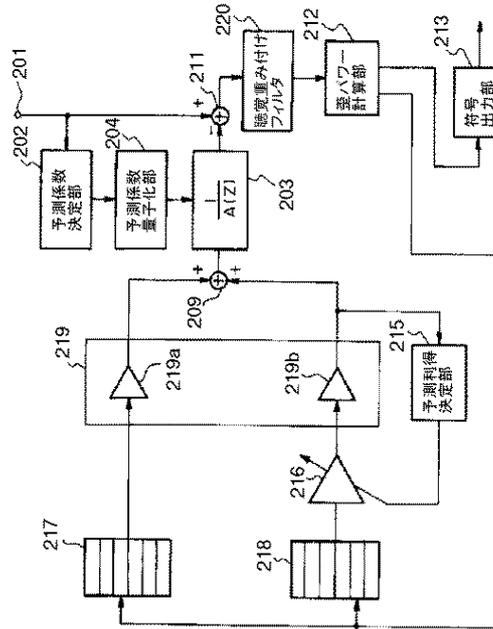
【図10】



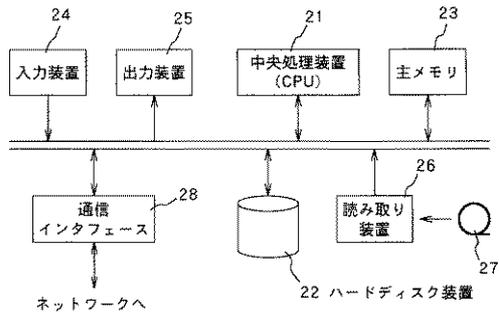
【図11】



【図12】



【図13】



フロントページの続き

- (72)発明者 岩上 直樹
東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内
- (72)発明者 森 岳至
東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内
- (72)発明者 千喜良 和明
東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内

審査官 山下 剛史

- (56)参考文献 特開平07-248794(JP,A)
特開平05-073093(JP,A)
特開平10-207489(JP,A)
特開平10-070467(JP,A)
特開平10-340100(JP,A)
特開昭62-296200(JP,A)
特開平07-325583(JP,A)
特開平08-328600(JP,A)
特開平09-230898(JP,A)
特開平11-194799(JP,A)
特開昭53-135561(JP,A)
特開2001-117573(JP,A)
特表2002-536679(JP,A)

(58)調査した分野(Int.Cl.⁷, DB名)

G10L 19/00
G10L 19/02
H03M 7/30
JICSTファイル(JOIS)