

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関  
国際事務局

(43) 国際公開日  
2014年1月16日(16.01.2014)



(10) 国際公開番号  
WO 2014/010175 A1

- (51) 国際特許分類:  
G10L 19/22 (2013.01) G10L 25/81 (2013.01)  
G10L 19/022 (2013.01)
- (21) 国際出願番号: PCT/JP2013/003814
- (22) 国際出願日: 2013年6月19日(19.06.2013)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:  
特願 2012-153563 2012年7月9日(09.07.2012) JP
- (71) 出願人: パナソニック株式会社 (PANASONIC CORPORATION) [JP/JP]; 〒5718501 大阪府門真市大字門真1006番地 Osaka (JP).
- (72) 発明者: 江原 宏幸 (EHARA, Hiroyuki). 堀 貴子 (HORI, Takako). 押切 正浩 (OSHIKIRI, Masahiro).
- (74) 代理人: 鷺田 公一 (WASHIDA, Kimihito); 〒1600023 東京都新宿区西新宿1-23-7 新宿ファーストウェスト8階 Tokyo (JP).
- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA,

BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

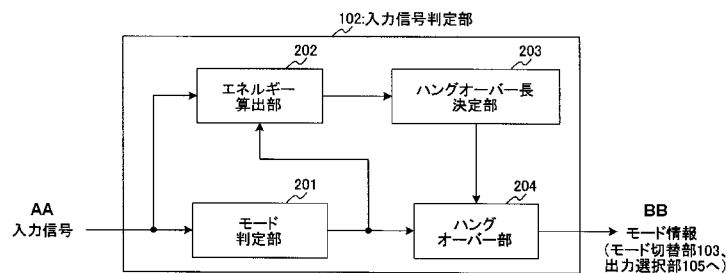
- (84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

- 国際調査報告 (条約第 21 条(3))

(54) Title: ENCODING DEVICE AND ENCODING METHOD

(54) 発明の名称: 符号化装置及び符号化方法



- 102 INPUT SIGNAL DETERMINATION UNIT
- 201 MODE DETERMINATION UNIT
- 202 ENERGY CALCULATION UNIT
- 203 HANGOVER LENGTH DETERMINATION UNIT
- 204 HANGOVER UNIT
- AA INPUT SIGNAL
- BB MODE INFORMATION (TO MODE SWITCHING UNIT (103), OUTPUT SELECTION UNIT (105))

(57) Abstract: Provided are an encoding device and encoding method capable of improving the accuracy of determining whether a BGM signal is in a voice signal mode or a music signal mode and minimizing degradation in sound quality. A mode determination unit (201) determines whether an input signal is in a voice signal mode or music signal mode, and an energy calculation unit (202) calculates the average energy of the input signal included in a frame determined to be a music signal. A hangover length determination unit (203) increases the hangover length when the amount of calculated energy information is large, and decreases the hangover length when the amount of calculated energy information is small. A hangover unit (204) uses the mode information of the previous frame, the mode information of the present frame, and the determined hangover length and corrects the mode information for the present frame when a prescribed condition is satisfied.

(57) 要約:

[続葉有]



WO 2014/010175 A1

---

BGM信号に対する音声／音楽判定の精度を向上させ、音質劣化を抑制する符号化装置及び符号化方法を提供する。モード判定部(201)は、入力信号が音声信号又は音楽信号のいずれのモードかを判定し、エネルギー算出部(202)は、音楽信号と判定されたフレームに含まれる入力信号の平均エネルギーを算出する。ハングオーバー長決定部(203)は、算出されたエネルギー情報が大きい場合には、ハングオーバー長を長くし、エネルギー情報が小さい場合には、ハングオーバー長を短くする。ハングオーバー部(204)は、前フレームのモード情報と、現フレームのモード情報と、決定されたハングオーバー長とを用いて、所定の条件を満足する場合に、現フレームのモード情報を修正する。

## 明 細 書

発明の名称：符号化装置及び符号化方法

### 技術分野

[0001] 本発明は、音声信号及び音楽信号を符号化する符号化装置及び符号化方法に関する。

### 背景技術

[0002] 現在、3 G P P (3rd Generation Partnership Project) において、E P S (Evolved Packet System) に適した音声コーデックであるE V S (Enhanced Voice Service) の標準化が行われている。非特許文献1には、E V S に対する要求条件が、近年の携帯端末を用いた通話サービスを意識して決められている。例えば、携帯端末の保留音が音楽付き音声であること、または、コールセンターからのB G M付き音声案内を携帯端末で聞きながら処理することを想定しているため、音楽も良い品質で再生することが求められている。

[0003] 音楽を良い品質で再生する技術として、特許文献1には、V S E L P (Vector Sum Excited Linear Prediction) を用いた携帯端末において、保留が指示された際にはV O X (Voice Operated Transmission) 制御をオフにする方法が開示されている。また、特許文献2には、携帯端末へ音楽を提供する場合に、同時に送出する周波数の数である和音数の異なる複数の音源ファイルを携帯端末が予め蓄積しておき、使用するコーデックに応じた音源ファイルを選択する方法が開示されている。

[0004] また、音声信号の符号化に適するよう設計された音声符号化部と、音楽信号の符号化に適するよう設計された音楽符号化部とを、入力信号に応じてフレーム単位で切り替える技術は、非特許文献2に開示の音声／音楽汎用符号化方式 (USAC: Unified Speech and Audio Coding) に代表されるように、音声信号及び音楽信号を高音質に符号化できる有望な技術として知られている。このとき、フレーム単位で符号化方式を切り替えるため、切り替え方法の

性能が音質に与える影響が大きい。なお、フレーム長は、符号化方式によって異なるが、多くの場合20 msecが用いられている。

- [0005] 一方で、音声／音楽の判定をフレーム単位で行う方式がG. 720. 1（または、GSAD: Generic Sound Activity Detector）としてITU-Tにて標準化されている（非特許文献3参照）。GSADでは、フレーム単位に特徴パラメータを使って音声／音楽の判定を行うが、判定結果が不安定となり、音声／音楽の頻繁な切り替えが生じてしまう場合がある。このような頻繁な切り替えを避けるため、ハングオーバーと呼ばれる技術が適用される。ハングオーバーとは、前フレームで選択された判定結果を規定回数だけ強制的に使用する技術であり、これにより頻繁な切り替えを避けることが可能となる。

## 先行技術文献

### 特許文献

- [0006] 特許文献1：特許第2983829号  
特許文献2：特許第4507822号

### 非特許文献

- [0007] 非特許文献1：3GPP TR 22.813 V10.0.0, “Study of Use Cases and Requirements for Enhanced Voice Codecs for the Evolved Packet System (EPS)”  
非特許文献2：T. Moriya, “Technologies for Speech and Audio Coding”,  
proc. ISCE2009, pp.148-149.  
非特許文献3：Recommendation ITU-T G.720.1

## 発明の概要

### 発明が解決しようとする課題

- [0008] ハングオーバーは、音声／音楽の頻繁な切り替えを避けるために有効な技術であるが、音楽が背景で流れている環境で音声通話を行う場合に問題がある。以後、このような信号をBGM信号と呼ぶ。BGM信号は、非音声区間では音楽が流れており、音声区間では音声に音楽が重畳している。

[0009] 図1にBGM信号、各フレームの信号成分（音声または音楽）、ハングオーバー値が2（すなわち、2フレームは前回の判定結果を強制的に使用）の場合における判定結果を示す。図1に示すように、信号成分が切り替わるときに、ハングオーバーがあるために、すぐには判定結果が切り替わらない。そのため誤判定となってしまう、音質が劣化するという問題がある。

[0010] この現象は、例えば、背景が無音の場合（クリーン信号）、及び背景雑音になっている場合（背景雑音信号）には、通常、音声と判定されるため、問題とはならない。このように、音楽が背景で流れている場合に音楽と判定してしまうため、上述したハングオーバーの問題はBGM信号特有のものである。

[0011] 本発明の目的は、BGM信号に対する音声／音楽判定の精度を向上させ、音質劣化を抑制する符号化装置及び符号化方法を提供することである。

#### 課題を解決するための手段

[0012] 本発明の符号化装置は、入力信号の符号化処理に許容される遅延時間を判定する処理遅延判定手段と、前記入力信号が音声信号又は音楽信号のいずれのモードかを所定の区間毎に判定するモード判定手段と、前記遅延時間に応じて、ハングオーバー長を決定するハングオーバー長決定手段と、前記モード判定手段によって判定された前区間のモードと現区間のモードとを比較し、前記比較の結果及び前記ハングオーバー長を用いて、前記現区間のモードを確定するハングオーバー手段と、確定されたモードに応じた符号化方式によって前記入力信号を符号化する符号化手段と、を具備する構成を採る。

[0013] 本発明の符号化方法は、入力信号の符号化処理に許容される遅延時間を判定する処理遅延判定工程と、前記入力信号が音声信号又は音楽信号のいずれのモードかを所定の区間毎に判定するモード判定工程と、前記遅延時間に応じて、ハングオーバー長を決定するハングオーバー長決定工程と、前記モード判定工程において判定された前区間のモードと現区間のモードとを比較し、前記比較の結果及び前記ハングオーバー長を用いて、前記現区間のモードを確定するハングオーバー工程と、確定されたモードに応じた符号化方式に

よって前記入力信号を符号化する符号化工程と、を具備するようにした。

### 発明の効果

- [0014] 本発明によれば、BGM信号に対する音声／音楽判定の精度を向上させ、音質劣化を抑制することができる。

### 図面の簡単な説明

- [0015] [図1]ハングオーバーに起因する音声／音楽判定の誤判定の様子を示す図  
[図2]本発明の実施の形態1に係る符号化装置の構成を示すブロック図  
[図3]図2に示した入力信号判定部の内部構成を示すブロック図  
[図4]エネルギー情報とハングオーバー長との対応関係を示す図  
[図5]図3に示したハングオーバー部の内部構成を示すブロック図  
[図6]本発明の実施の形態1に係る符号化装置の効果を示す図  
[図7]本発明の実施の形態2に係る符号化装置の構成を示すブロック図  
[図8]図7に示した処理遅延判定部の処理手順を示すフロー図  
[図9]図7に示した符号化部の内部構成を示すブロック図  
[図10]処理遅延とハングオーバー長との対応関係を示す図

### 発明を実施するための形態

- [0016] 以下、本発明の実施の形態について、図面を参照して詳細に説明する。ただし、実施の形態において、同一機能を有する構成には、同一符号を付し、重複する説明は省略する。

- [0017] (実施の形態1)

図2は、本発明の実施の形態1に係る符号化装置100の構成を示すブロック図である。以下、符号化装置100の構成について図2を用いて説明する。

- [0018] 入力バッファ101は、入力信号を入力信号判定部102に出力すると共に、入力信号を一時記憶して、モード切替部103に出力する。

- [0019] 入力信号判定部102は、入力バッファ101から出力された入力信号が音声信号か音楽信号かを判定し、判定結果をモード情報としてモード切替部103及び出力選択部105に出力する。なお、入力信号判定部102の詳細

細については後述する。

[0020] モード切替部 103 は、入力信号判定部 102 から出力されたモード情報に基づいて、切替スイッチを符号化部コア 104 の符号化モード 1 又は符号化モード 2 に接続し、入力バッファ 101 から出力された入力信号を接続した符号化モード 1 または符号化モード 2 に出力する。具体的には、モード切替部 103 は、モード情報が音声信号を示す場合には、切替スイッチを符号化モード 1 に接続し、モード情報が音楽信号を示す場合には、切替スイッチを符号化モード 2 に接続する。

[0021] 符号化部コア 104 は、符号化モード 1 と符号化モード 2 の符号化方式を備え、符号化モード 1 は、例えば、G. 729 などの音声信号に適した符号化方式であり、符号化モード 2 は、例えば、MP3 (MPEG Audio Layer-3) または AAC (Advanced Audio Coding) などの音楽信号に適した符号化方式である。符号化部コア 104 は、入力バッファ 101 から出力された入力信号を符号化モード 1 または符号化モード 2 によって符号化し、符号化情報を出力選択部 105 に出力する。なお、符号化モード 1 と符号化モード 2 のいずれを用いて符号化したかを示す情報も符号化情報の一部として出力してもよい。この場合、符号化情報のみから復号処理が可能である。

[0022] 出力選択部 105 は、入力信号判定部 102 から出力されたモード情報に基づいて、符号化部コア 104 の符号化モード 1 又は符号化モード 2 に接続し、符号化モード 1 又は符号化モード 2 から出力された符号化情報を符号化装置 100 の出力とする。

[0023] 図 3 は、図 2 に示した入力信号判定部 102 の内部構成を示すブロック図である。以下、入力信号判定部 102 の内部構成について図 3 を用いて説明する。図 3 において、予め定められた時間長によって区切られた入力信号（フレーム化された入力信号）がモード判定部 201 及びエネルギー算出部 202 に入力される。

[0024] モード判定部 201 は、既存の手法により入力信号の分析を行って特徴パラメータを算出し、特徴パラメータを用いて入力信号が音声信号又は音楽信

号のいずれのモードかを判定する。モード判定部201は、判定結果（モード情報）をエネルギー算出部202及びハングオーバー部204に出力する。

[0025] エネルギー算出部202は、モード判定部201において音楽信号と判定されたフレームに含まれる入力信号の平均エネルギー（またはエネルギー情報）を算出し、エネルギー情報をハングオーバー長決定部203に出力する。エネルギー情報の算出方法としては、例えば、次のような処理を行う。

[0026] エネルギー算出部202は、音楽信号と判定された過去のフレームの平均エネルギーを記憶するバッファを有し、現フレームが音楽信号と判定されたときに、現フレームのエネルギーを反映するように、バッファに記憶されている値を更新する。更新は次の式（1）に従って行われる。

[数1]

$$E_{avg} = \alpha \cdot E_n + (1 - \alpha) \cdot E_{avg} \quad \dots (1)$$

[0027] ここで、 $E_{avg}$ は、バッファに記憶されている音楽信号と判定された過去のフレームの平均エネルギーを表す。また、 $E_n$ は、現フレームが音楽信号と判定されたときの現フレームに含まれる信号のエネルギーを表す。さらに、 $\alpha$ は、更新速度を制御する0以上1未満の係数を表し、例えば $\alpha = 0.95$ などの数値を用いる。現フレームが音声信号と判定された場合には、このバッファの更新は行われない。これにより、音楽信号の平均的なエネルギーを算出することができる。

[0028] エネルギー算出部202は、このように算出した平均エネルギー $E_{avg}$ をエネルギー情報として、ハングオーバー長決定部203に出力する。

[0029] ハングオーバー長決定部203は、エネルギー算出部202から出力されたエネルギー情報と所定の閾値とを比較し、エネルギー情報が閾値より大きい場合には、ハングオーバー長を長くしてハングオーバー部204に出力する。一方、エネルギー情報が閾値より小さい場合には、ハングオーバー長決定部203は、ハングオーバー長を短くしてハングオーバー部204に出力

する。ハングオーバー長の具体的な値としては、例えば、図4に示すように、エネルギー情報が大きい場合には、ハングオーバー長を2フレームとし、エネルギー情報が小さい場合には、ハングオーバー長を1フレームなどとする。

[0030] このように、BGM信号における非音声区間の音楽信号のエネルギー情報が小さい場合には、音声／音楽判定の誤判定が起きにくく、判定の精度が向上するため、ハングオーバー長を短くしても、ハングオーバーに起因する音声／音楽判定の誤判定を低減することができる。一方、BGM信号における音声区間（音声信号）のエネルギー情報が大きい場合には、音声／音楽判定の誤判定が起りやすく、判定の精度が劣化するため、ハングオーバー長を長くすることにより、音声／音楽判定の誤判定を低減することができる。

[0031] ハングオーバー部204は、前フレームで決定したモード情報を記憶しており、前フレームのモード情報と、モード判定部201から出力された現フレームのモード情報と、ハングオーバー長決定部203から出力されたハングオーバー長とを用いて、モード判定部201から出力された現フレームのモード情報を確定して出力する。

[0032] 図5は、図3に示したハングオーバー部204の内部構成を示すブロック図である。以下、ハングオーバー部204の内部構成について図5を用いて説明する。記憶部301には、前フレームにてハングオーバー部204から出力されたモード情報が記憶されており、前フレームのモード情報が判定部302に出力される。

[0033] 判定部302は、記憶部301から出力された前フレームのモード情報と、モード判定部201から出力された現フレームのモード情報とを比較する。前フレームのモード情報と現フレームのモード情報とが一致する場合、判定部302に内蔵されているカウンタをゼロにリセットし、パス(B)が有効となるようにスイッチ303、304を切り替える。パス(B)は、モード判定部201から出力されたモード情報をそのまま出力する経路である。このため、モード判定部201から出力されたモード情報は、何ら処理され

ることなくハングオーバー部204から出力される。

[0034] 一方、前フレームのモード情報と現フレームのモード情報とが一致しない場合、判定部302に内蔵されているカウンタをインクリメントする。判定部302は、カウンタ値とハングオーバー長決定部203から出力されたハングオーバー長とを比較する。カウンタ値がハングオーバー長以下の場合、判定部302は、ハングオーバー処理が有効、すなわち、パス(A)が有効となるようにスイッチ303、304を切り替える。パス(A)は、モード判定部201から出力されたモード情報をモード情報修正部305において修正して、ハングオーバー部204から出力する経路である。また、カウンタ値がハングオーバー長を超えている場合、判定部302は、カウンタをゼロにリセットすると共に、パス(B)が有効となるようにスイッチ303、304を切り替える。

[0035] モード情報修正部305は、パス(A)が有効なときにのみ動作し、モード判定部201から出力されたモード情報を、記憶部301に記憶されている前フレームのモード情報に置き換え、出力する。

[0036] パス(A)またはパス(B)を通過してハングオーバー部204から出力される現フレームのモード情報は、記憶部301に記憶されることにより、それまで記憶されていたモード情報と置き換えられ、次のフレームの処理に備える。

[0037] 図6は、上述した符号化装置100の効果を示す図である。図6では、図1と同一のBGM信号が入力されており、図1ではハングオーバー長が2と固定されているのに対し、図6ではエネルギー算出部202で算出されたエネルギー情報が閾値より小さく、そのため、ハングオーバー長を1としたときの様子を示している。

[0038] 図6から分かるように、ハングオーバーに起因する誤判定の長さが、図1に比べて半分になっている。このため、上述した符号化装置100では音質を改善することが可能となる。

[0039] このように、実施の形態1によれば、入力信号のエネルギーを算出し、算

出したエネルギーが小さい場合に、ハングオーバー長を短くすることにより、音声／音楽判定を短い区間で行うことができ、BGM信号に対する音声／音楽判定の精度を向上させ、音質劣化を抑制することができる。

[0040] (実施の形態2)

音声信号は、通常、相手との会話、すなわち双方向通信にて使用される。そのため、遅延が長くなると会話が成り立たなくなるので、音声信号を短い遅延（以下、「低遅延」という）で符号化する必要がある。また、音声信号は、無音区間、無声区間及び有声区間というように、比較的短時間に信号の特性が大きく変わるという特徴がある。このため、符号化用のバッファに長い時間信号を格納（すなわち、長い遅延（以下、「高遅延」という））して分析しても符号化効率は高くなりにくい。このような理由により、音声信号の符号化には低遅延が向いている。

[0041] 一方、音楽信号は、音声信号のように信号特性が短時間で大きく変わることは稀である。そのため、符号化用の分析バッファに長い時間信号を格納して分析することにより、符号化効率が大きく向上する。また、音楽信号は、通常、サーバーから端末に片方向にデータを伝送するストリーミングが主要なアプリケーションとなるため、片方向通信は双方向通信より遅延の要求が厳しくない。このような理由により、音楽信号の符号化には高遅延が適しているといえる。

[0042] 以下、上述した音声信号と音楽信号の特徴を踏まえ、本発明の実施の形態2について説明する。図7は、本発明の実施の形態2に係る符号化装置400の構成を示すブロック図である。以下、符号化装置400の構成について図7を用いて説明する。

[0043] ユーザインタフェース401は、例えば、キーボード、タッチパネル等であり、入力源、すなわち、マイクロフォン402及びデータ蓄積部403のON動作を切り替える入力源切替信号をマイクロフォン402、データ蓄積部403及び入力源特定部404に出力する。

[0044] マイクロフォン402は、ユーザインタフェース401から出力された入

力源切替信号に従って、音声を入力し、入力した音声を変換して、符号化部406に出力する。また、データ蓄積部403は、保留音又はメッセージ等のデータを蓄積し、ユーザインタフェース401から出力された入力源切替信号に従って、蓄積したデータを符号化部406に出力する。

[0045] 入力源特定部404は、ユーザインタフェース401から出力された入力源切替信号によって入力源を特定し、特定した入力源を示す入力源情報を処理遅延判定部405に出力する。

[0046] 処理遅延判定部405は、入力源特定部404から出力された入力源情報に応じて、符号化処理に許容される遅延時間を判定する。具体的には、図8に示すように、入力源がマイクロフォンであった場合（ST501：YES）、処理遅延判定部405は、符号化処理の対象となるデータが、音声通話等の双方向のリアルタイム処理が必要なデータであると判断し、処理遅延を低遅延と判定する（ST502）。一方、入力源がデータ蓄積部403であった場合（ST501：NO）、処理遅延判定部405は、符号化処理の対象となるデータが、保留音またはメッセージ等の双方向リアルタイム処理の必要のないデータであると判断し、処理遅延を高遅延と判定する（ST503）。処理遅延判定部405は、判定結果（遅延情報）を符号化部406に出力する。

[0047] 符号化部406は、処理遅延判定部405から出力された遅延情報に基づいて、マイクロフォン402から出力された音声信号又はデータ蓄積部403から出力されたデータをそれぞれの特性に適した符号化方式によって符号化し、符号化情報を出力する。

[0048] 図9は、図7に示した符号化部406の内部構成を示すブロック図である。以下、符号化部406の内部構成について図9を用いて説明する。ただし、図9の符号化部406は、図2の符号化装置100に対して、入力信号判定部102を削除し、モード切替部103をモード切替部601に変更し、出力選択部105を出力選択部602に変更した点が異なる。

[0049] モード切替部601は、処理遅延判定部405から出力された遅延情報に

基づいて、切替スイッチを符号化部コア104の符号化モード1または符号化モード2に接続し、入力バッファ101から出力された入力信号を接続した符号化モード1または符号化モード2に出力する。具体的には、モード切替部601は、遅延情報が低遅延を示す場合には、符号化モード1に接続し、遅延情報が高遅延を示す場合には、符号化モード2に接続する。なお、符号化モード1は、例えば、G. 729などの音声信号に適した符号化方式であり、符号化モード2は、例えば、MP3又はAACなどの音楽信号に適した符号化方式である。

[0050] 出力選択部602は、処理遅延判定部405から出力された遅延情報に基づいて、符号化部コア104の符号化モード1または符号化モード2に接続し、符号化モード1または符号化モード2から出力された符号化情報を符号化装置400の出力とする。

[0051] このように、実施の形態2によれば、入力源に応じて、符号化処理に許容される処理遅延を判定し、判定結果を示す遅延情報に応じて、符号化方式を切り替えて入力信号を符号化することにより、精度良く符号化することができ、音質劣化を抑制することができる。

[0052] なお、実施の形態1では、入力信号のエネルギー情報に応じて、ハングオーバー長を制御する場合について説明したが、符号化処理に許容される処理遅延に応じて、ハングオーバー長を制御してもよい。

[0053] 高遅延では、入力信号を長い時間バッファに格納できるので、未来のデータまで参照可能である。このため、音声／音楽判定の性能自体が高くなる。この場合、長いハングオーバーは不要であり、ハングオーバー長を短くする。これにより、ハングオーバーに起因する音質劣化を回避することができ、全体的な音質を改善することができる。

[0054] 一方、低遅延では、入力信号を短い時間しかバッファに格納できないので、未来のデータを参照することができない。このため、現在のデータしか分析に使用することができない。この場合、音声／音楽判定の判定性能が低下してしまうので、ハングオーバー長を長くし、性能の低い音声／音楽判定に

よる頻繁な切り替えの発生を防ぎ、音質劣化を回避する。図10は、低遅延の場合には、ハングオーバー長を長く（ハングオーバー長=2）、高遅延の場合には、ハングオーバー長を短く（ハングオーバー長=1）したときの例を示している。

[0055] 図4及び図10では、ハングオーバー長が1~2フレームの場合について説明しているが、本発明はこれに限定されるものではない。符号化方式のフレーム長が短い場合（例えば、10 msec以下）、または入力信号に雑音を重ねているため音声/音楽判定性能が悪い場合に、ハングオーバー長を長くしてもよい。また、符号化方式のフレーム長が長い場合（例えば、40 msec以上）には、ハングオーバー長を短くしてもよい。

[0056] なお、上記各実施の形態では、本発明をハードウェアで構成する場合を例にとって説明したが、本発明はハードウェアとの連携においてソフトウェアによって実現することも可能である。

[0057] また、上記各実施の形態の説明に用いた各機能ブロックは、典型的には集積回路であるLSIとして実現される。これらは個別に1チップ化されてもよいし、一部又は全てを含むように1チップ化されてもよい。ここでは、LSIとしたが、集積度の違いにより、IC、システムLSI、スーパーLSI、ウルトラLSIと呼称されることもある。

[0058] また、集積回路化の手法はLSIに限るものではなく、専用回路又は汎用プロセッサで実現してもよい。LSI製造後に、プログラムすることが可能なFPGA（Field Programmable Gate Array）、またはLSI内部の回路セルの接続及び設定を再構成可能なリプログラマブル・プロセッサを利用してもよい。

[0059] さらには、半導体技術の進歩又は派生する別技術によりLSIに置き換わる集積回路化の技術が登場すれば、当然、その技術を用いて機能ブロックの集積化を行ってもよい。バイオ技術の適用等が可能性としてありえる。

[0060] 2012年7月9日出願の特願2012-153563の日本出願に含まれる明細書、図面及び要約書の開示内容は、すべて本願に援用される。

## 産業上の利用可能性

[0061] 本発明にかかる符号化装置及び符号化方法は、例えば、通話機能を備えた携帯電話等の通信端末に適用できる。

## 符号の説明

- [0062]
- 101 入力バッファ
  - 102 入力信号判定部
  - 103、601 モード切替部
  - 104 符号化部コア
  - 105、602 出力選択部
  - 201 モード判定部
  - 202 エネルギー算出部
  - 203 ハングオーバー長決定部
  - 204 ハングオーバー部
  - 301 記憶部
  - 302 判定部
  - 303、304 スイッチ
  - 305 モード情報修正部
  - 401 ユーザインタフェース
  - 402 マイクロフォン
  - 403 データ蓄積部
  - 404 入力源特定部
  - 405 処理遅延判定部
  - 406 符号化部

## 請求の範囲

- [請求項1] 入力信号の符号化処理に許容される遅延時間を判定する処理遅延判定手段と、
- 前記入力信号が音声信号又は音楽信号のいずれのモードかを所定の区間毎に判定するモード判定手段と、
- 前記遅延時間に応じて、ハングオーバー長を決定するハングオーバー長決定手段と、
- 前記モード判定手段によって判定された前区間のモードと現区間のモードとを比較し、前記比較の結果及び前記ハングオーバー長を用いて、前記現区間のモードを確定するハングオーバー手段と、
- 確定されたモードに応じた符号化方式によって前記入力信号を符号化する符号化手段と、
- を具備する符号化装置。
- [請求項2] 前記ハングオーバー長決定手段は、前記遅延時間が短い低遅延の場合、前記ハングオーバー長を長くし、前記遅延時間が長い高遅延の場合、前記ハングオーバー長を短くする、
- 請求項1に記載の符号化装置。
- [請求項3] 音楽信号と判定された区間に含まれる前記入力信号のエネルギーを算出するエネルギー算出手段を具備し、
- 前記ハングオーバー長決定手段は、前記エネルギーに応じて、ハングオーバー長を決定する、
- 請求項1に記載の符号化装置。
- [請求項4] 前記ハングオーバー長決定手段は、前記エネルギーが大きい場合、前記ハングオーバー長を長くし、前記エネルギーが小さい場合、前記ハングオーバー長を短くする、
- 請求項3に記載の符号化装置。
- [請求項5] 前記入力信号の入力源を特定する入力源特定手段を具備し、
- 前記処理遅延判定手段は、前記入力源に応じて、前記遅延時間を判

定する、

請求項 1 に記載の符号化装置。

[請求項6] 前記入力源は、マイクロフォン及び蓄積されたデータである、  
請求項 5 に記載の符号化装置。

[請求項7] 前記処理遅延判定手段は、前記入力源がマイクロフォンである場合、前記遅延時間が低遅延であると判定し、前記入力源が蓄積されたデータである場合、前記遅延時間が高遅延であると判定する、  
請求項 6 に記載の符号化装置。

[請求項8] 入力信号の符号化処理に許容される遅延時間を判定する処理遅延判定工程と、

前記入力信号が音声信号又は音楽信号のいずれのモードかを所定の区間毎に判定するモード判定工程と、

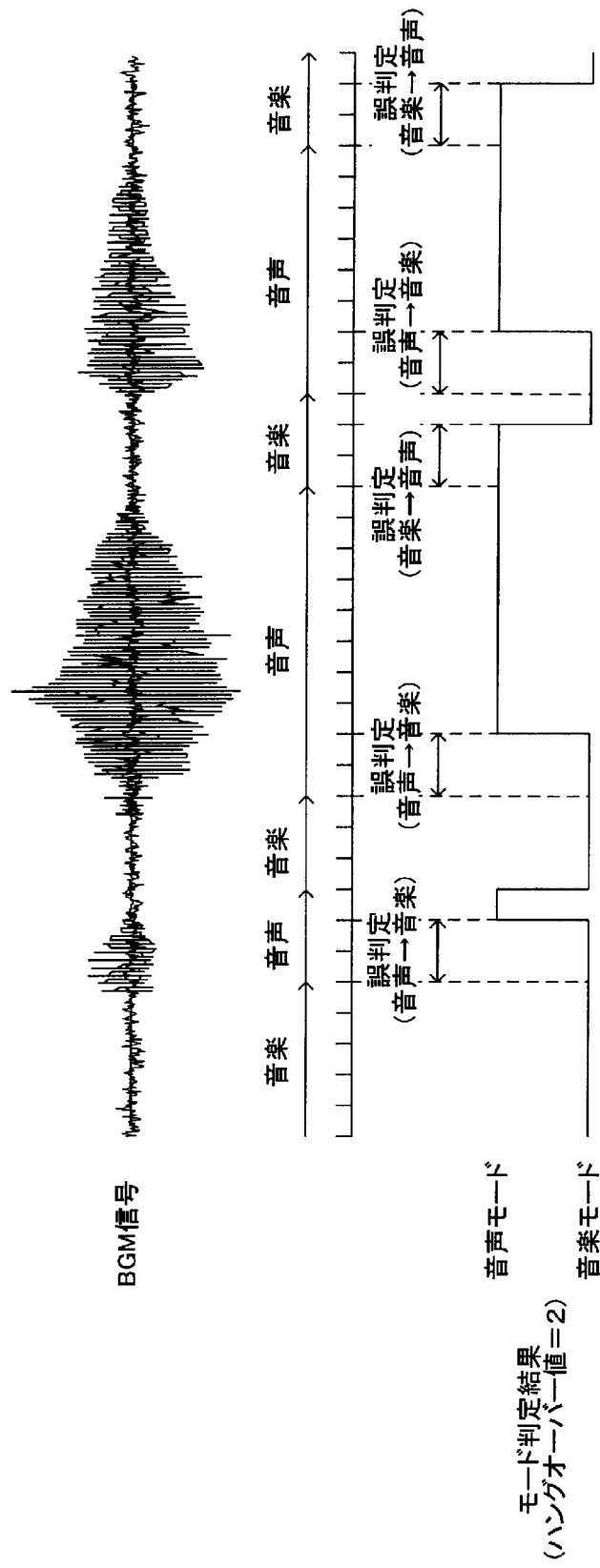
前記遅延時間に応じて、ハングオーバー長を決定するハングオーバー長決定工程と、

前記モード判定工程において判定された前区間のモードと現区間のモードとを比較し、前記比較の結果及び前記ハングオーバー長を用いて、前記現区間のモードを確定するハングオーバー工程と、

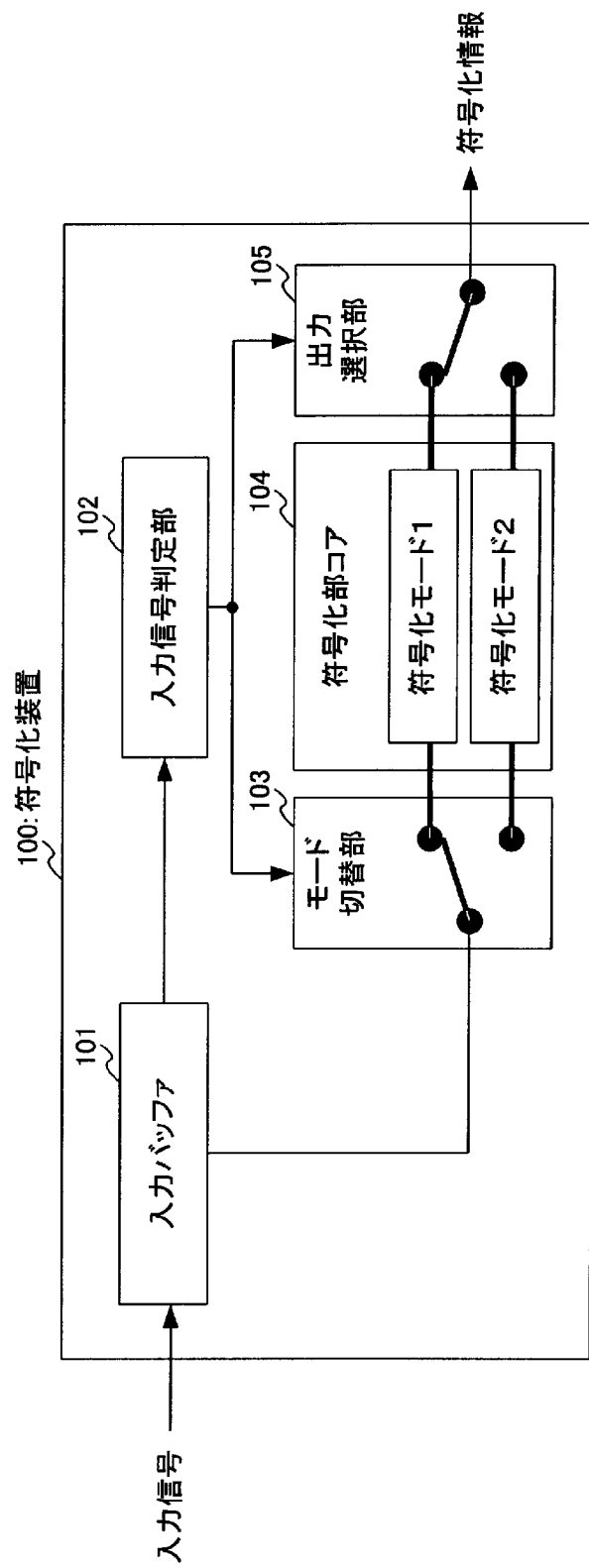
確定されたモードに応じた符号化方式によって前記入力信号を符号化する符号化工程と、

を具備する符号化方法。

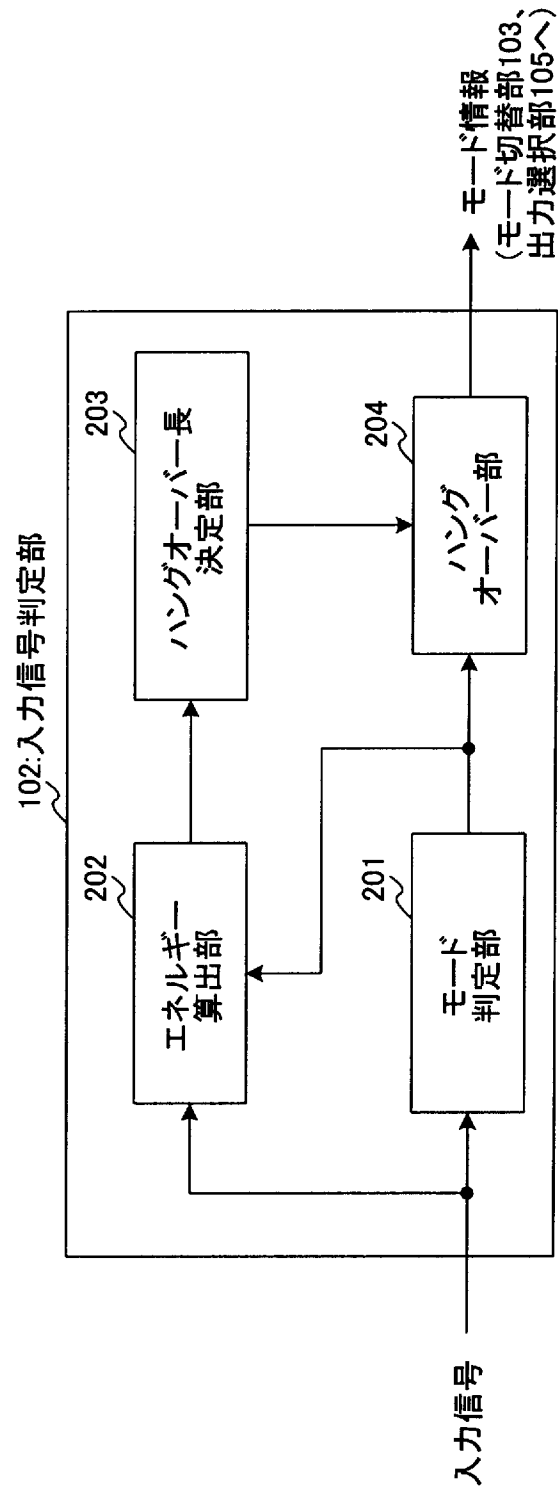
[図1]



[図2]



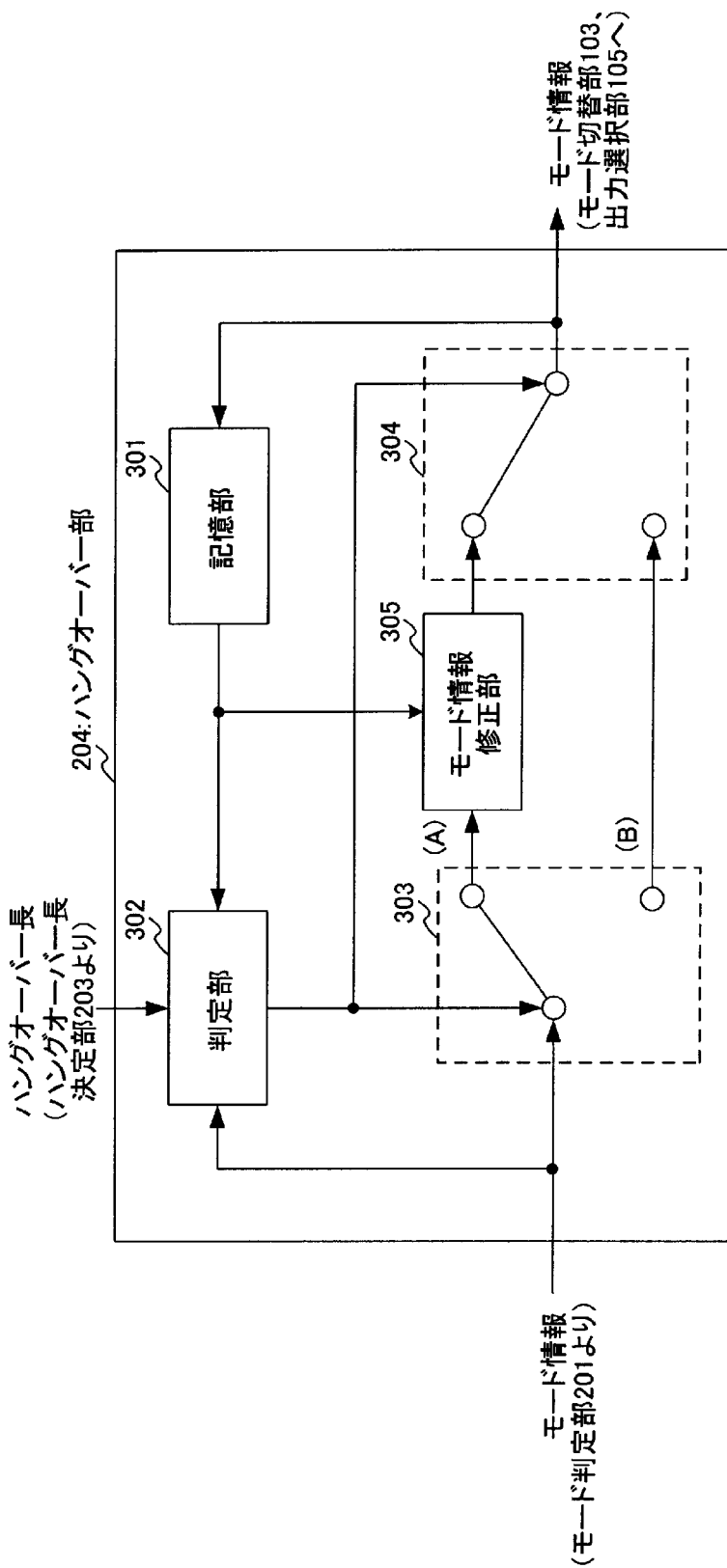
[図3]



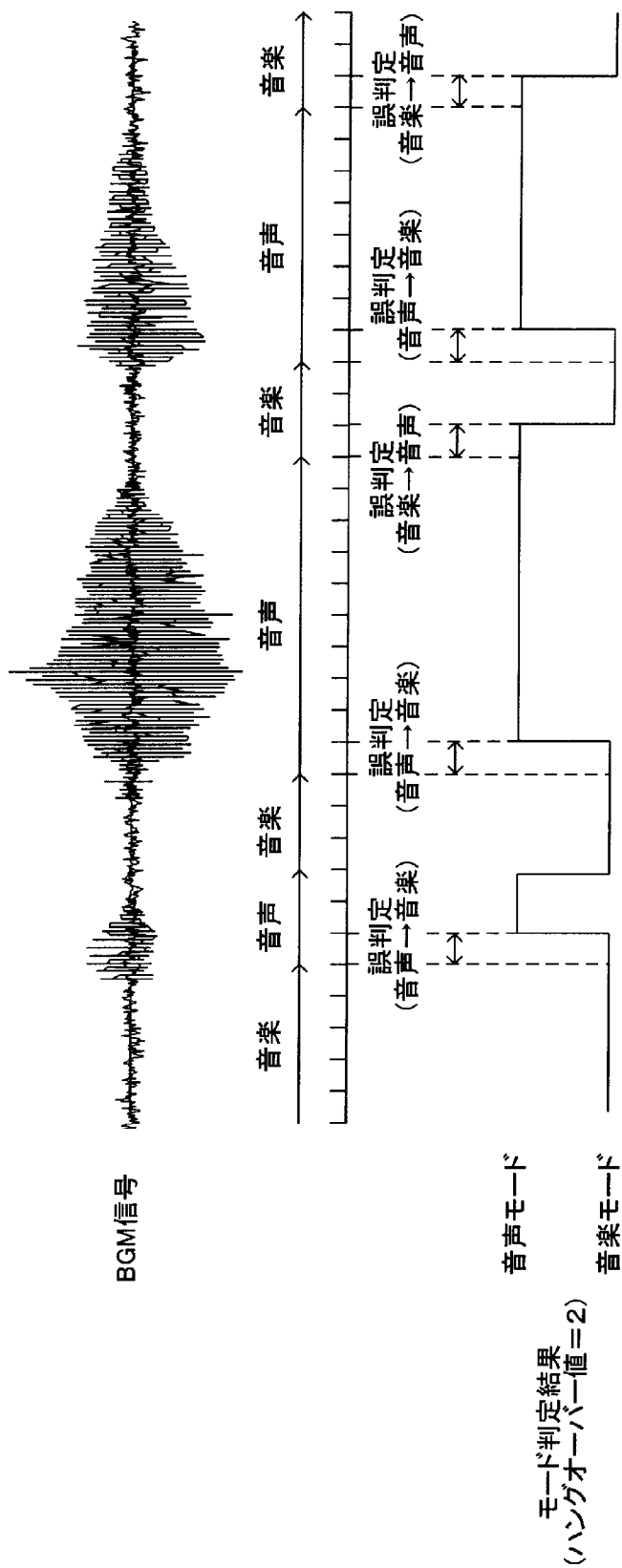
[図4]

エネルギー情報:小	エネルギー情報:大
ハンダグオーバー長= 1フレーム	ハンダグオーバー長= 2フレーム

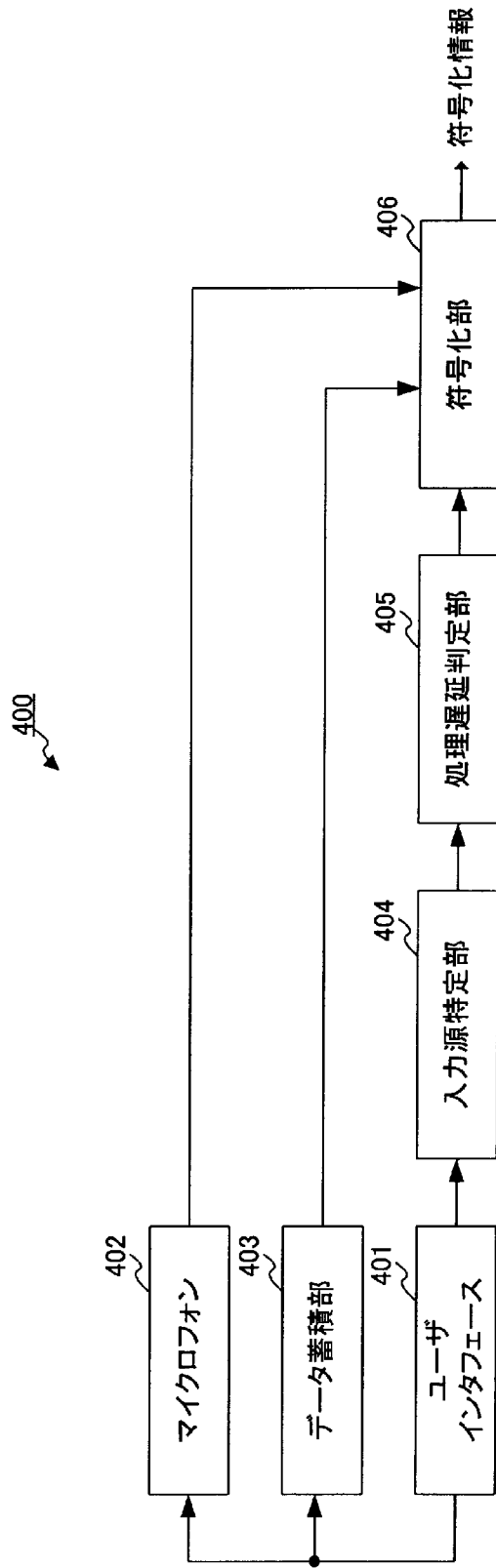
[図5]



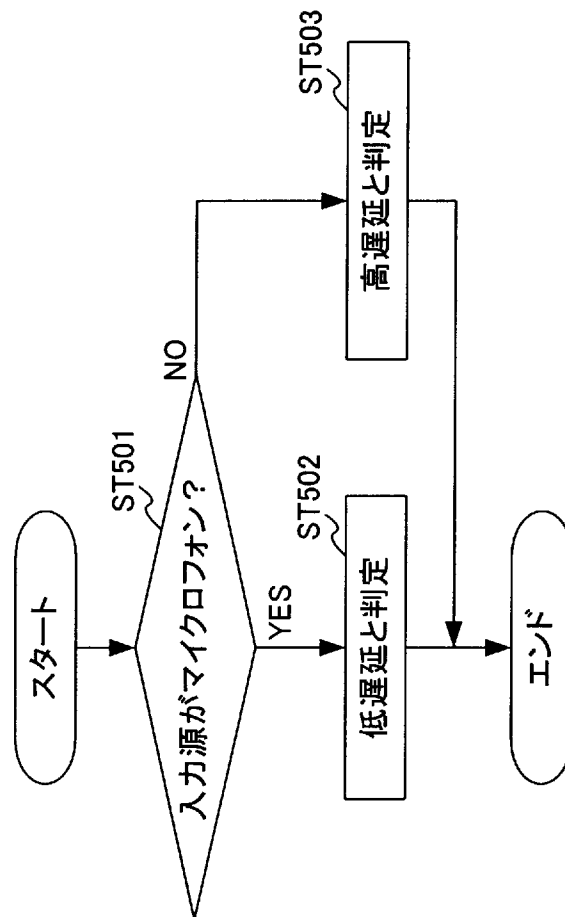
[図6]



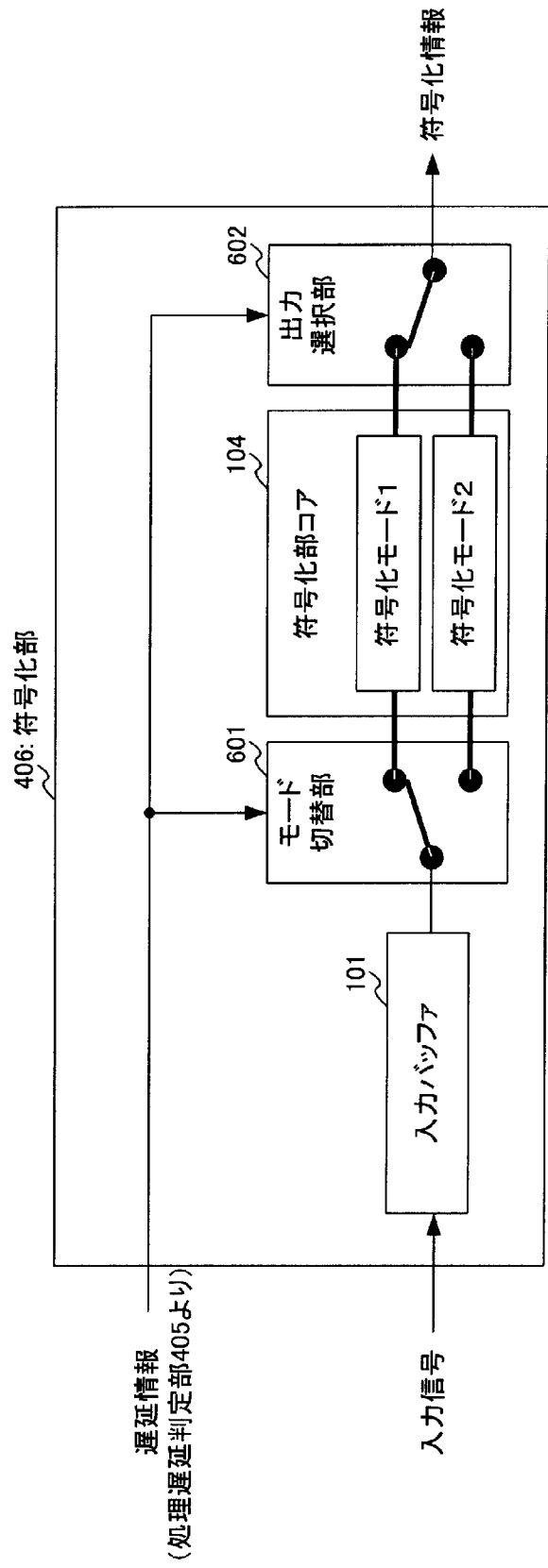
[図7]



[図8]



[図9]



[図10]

高遅延	ハングオーバー長= 1フレーム
低遅延	ハングオーバー長= 2フレーム

**INTERNATIONAL SEARCH REPORT**

International application No. PCT/JP2013/003814
--

**A. CLASSIFICATION OF SUBJECT MATTER**  
*G10L19/22(2013.01) i, G10L19/022(2013.01) i, G10L25/81(2013.01) i*

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)  
 G10L19/22, G10L19/022, G10L25/81

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2013
Kokai Jitsuyo Shinan Koho	1971-2013	Toroku Jitsuyo Shinan Koho	1994-2013

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2004-301907 A (Matsushita Electric Industrial Co., Ltd.), 28 October 2004 (28.10.2004), paragraph [0036] (Family: none)	1-8
Y	US 6202046 B1 (Kabushiki Kaisha Toshiba), 13 March 2001 (13.03.2001), column 16, lines 8 to 37; column 19, lines 11 to 28 & JP 10-207491 A	1-8

Further documents are listed in the continuation of Box C.       See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 17 July, 2013 (17.07.13)	Date of mailing of the international search report 30 July, 2013 (30.07.13)
---	--

Name and mailing address of the ISA/ Japanese Patent Office	Authorized officer
Facsimile No.	Telephone No.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2013/003814

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2008/121035 A1 (TELEFONAKTIEBOLAGET LM ERICSSON (publ)), 09 October 2008 (09.10.2008), page 4, line 17 to page 11, line 11 & JP 2010-525376 A & US 2010/0106490 A1 & EP 2143103 A & KR 10-2009-0122976 A	1-8
A	US 5410632 A (Motorola Inc.), 25 April 1995 (25.04.1995), column 2, lines 10 to 68; column 3, lines 36 to 45; column 5, lines 25 to 29; column 6, lines 21 to 33 & WO 1993/013516 A1	1-8

A. 発明の属する分野の分類 (国際特許分類 (IPC))  
 Int.Cl. G10L19/22(2013.01)i, G10L19/022(2013.01)i, G10L25/81(2013.01)i

B. 調査を行った分野  
 調査を行った最小限資料 (国際特許分類 (IPC))  
 Int.Cl. G10L19/22, G10L19/022, G10L25/81

最小限資料以外の資料で調査を行った分野に含まれるもの  
 日本国実用新案公報 1922-1996年  
 日本国公開実用新案公報 1971-2013年  
 日本国実用新案登録公報 1996-2013年  
 日本国登録実用新案公報 1994-2013年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
Y	JP 2004-301907 A (松下電器産業株式会社) 2004.10.28, 段落【036】 (ファミリーなし)	1-8
Y	US 6202046 B1 (Kabushiki Kaisha Toshiba) 2001.03.13, 第16欄第8-37行, 第19欄第11-28行 & JP 10-207491 A	1-8
A	WO 2008/121035 A1 (TELEFONAKTIEBOLAGET LM ERICSSON (publ)) 2008.10.09, 第4頁第17行-第11頁第11行 & JP 2010-525376 A & US 2010/0106490 A1 & EP 2143103 A & KR 10-2009-0122976 A	1-8

C欄の続きにも文献が列挙されている。  パテントファミリーに関する別紙を参照。

\* 引用文献のカテゴリー  
 「A」 特に関連のある文献ではなく、一般的技術水準を示すもの  
 「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの  
 「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)  
 「O」 口頭による開示、使用、展示等に言及する文献  
 「P」 国際出願日前で、かつ優先権の主張の基礎となる出願日の後に公表された文献  
 「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの  
 「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの  
 「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの  
 「&」 同一パテントファミリー文献

国際調査を完了した日 17.07.2013	国際調査報告の発送日 30.07.2013		
国際調査機関の名称及びあて先 日本国特許庁 (ISA/J P) 郵便番号100-8915 東京都千代田区霞が関三丁目4番3号	特許庁審査官 (権限のある職員) 浜岸 広明	5 Z	9 8 4 5
電話番号 03-3581-1101 内線 3591			

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	US 5410632 A (Motorola Inc.) 1995.04.25, 第2欄第10-68行, 第3欄第36-45行, 第5欄第25-29行, 第6欄第21-33行 & W0 1993/013516 A1	1-8