

(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
2006年5月4日 (04.05.2006)

PCT

(10) 国際公開番号
WO 2006/046546 A1

(51) 国際特許分類:

G10L 19/02 (2006.01) H03M 7/30 (2006.01)

(72) 発明者; および

(75) 発明者/出願人(米国についてのみ): 押切 正浩 (OSHIKIRI, Masahiro).

(21) 国際出願番号:

PCT/JP2005/019578

(22) 国際出願日:

2005年10月25日 (25.10.2005)

(74) 代理人: 鷺田 公一 (WASHIDA, Kimihito); 〒2060034 東京都多摩市鶴牧1丁目24-1 新都市センタービル5階 Tokyo (JP).

(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 2004-311143

2004年10月26日 (26.10.2004) JP

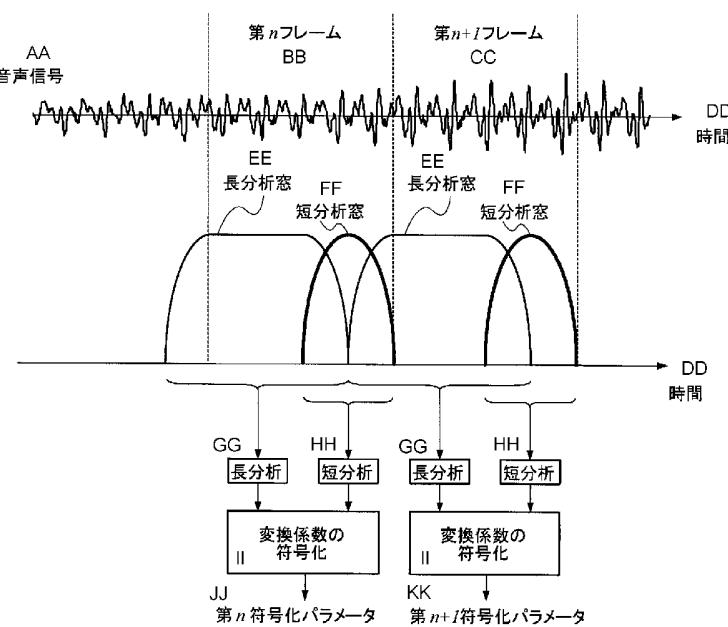
(71) 出願人(米国を除く全ての指定国について): 松下電器産業株式会社 (MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD.) [JP/JP]; 〒5718501 大阪府門真市大字門真1006番地 Osaka (JP).

(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, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU,

[続葉有]

(54) Title: SOUND ENCODING DEVICE AND SOUND ENCODING METHOD

(54) 発明の名称: 音声符号化装置および音声符号化方法



AA... SOUND SIGNAL

BB... n-TH FRAME

CC... (n+1)-TH FRAME

DD... TIME

EE... LONG ANALYSIS WINDOW

FF... SHORT ANALYSIS WINDOW

GG... LONG ANALYSIS

HH... SHORT ANALYSIS

II... ENCODING OF TRANSFORM COEFFICIENT

JJ... n-TH ENCODED PARAMETER

KK... (n+1)-TH ENCODED PARAMETER

対して分析窓を乗じ、この分析窓を乗じた信号をMDCT部(212)に出力され、MDCT部(212)が、入力

(57) Abstract: A sound encoding device enabling the amount of delay to be kept small and the connection distortion to be mitigated. In the sound encoding device, a window multiplication part (211) of a long analysis section (21) multiplies a long analysis frame signal of analysis length M1 by an analysis window, the resultant signal multiplied by the analysis window is outputted to an MDCT section (212), and the MDCT section (212) performs MDCT of the input signal to obtain the transform coefficient of the long analysis frame and outputs it to a transform coefficient encoding section (30). The window multiplication part (221) of a short analysis section (22) multiplies a short analysis frame signal of analysis length M2 ($M2 < M1$) by an analysis window and the resultant signal multiplied by the analysis window is outputted to the MDCT section (222). The MDCT section (222) performs MDCT of the input signal to obtain the transform coefficient of the short analysis frame and outputs it to the transform coefficient encoding section (30). A transform coefficient encoding section (30) encodes these transform coefficients and outputs them.

(57) 要約: 遅延量を低遅延に抑え、かつ、接続歪を緩和することができる音声符号化装置。この装置では、長分析部(21)において、窓掛け部(211)が、分析長M1の長分析フレーム信号に

[続葉有]

WO 2006/046546 A1



SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

- (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, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR),

添付公開書類:
— 国際調査報告書

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

信号にMDCTを施して長分析フレームの変換係数を得て変換係数符号化部（30）に出力する。一方、短分析部（22）では、窓掛け部（221）が、分析長M2（M2 < M1）の短分析フレーム信号に対して分析窓を乗じ、この分析窓を乗じた信号をMDCT部（222）に出力し、MDCT部（222）が、入力された信号にMDCTを施して短分析フレームの変換係数を得て変換係数符号化部（30）に出力する。変換係数符号化部（30）は、これらの変換係数を符号化して出力する。

明細書

音声符号化装置および音声符号化方法

技術分野

[0001] 本発明は、音声符号化装置および音声符号化方法に関する。

背景技術

[0002] 音声符号化において、時間信号を周波数領域に変換し変換係数を符号化する変換符号化は、時間信号に含まれる冗長性を効率的に除去できる。また、変換符号化では、周波数領域で表される人間の聴覚特性を利用することにより、低ビットレートにしても量子化歪が知覚され難い符号化を実現することが可能である。

[0003] 近年の変換符号化には、重ね合わせ直交変換(LOT; Lapped Orthogonal Transform)と呼ばれる変換技術が利用されることが多い。これは、ブロック内だけの直交性だけでなく、隣り合うブロックとの直交性を考慮した直交基底に基づき変換を行うものであり、その代表的な手法としてMDCT(Modified Discrete Cosine Transform)が挙げられる。MDCTでは現分析フレームとその前後の分析フレームとを重ね合わせるようにして分析フレームを配置し分析を行う。この際、変換後の係数のうち、分析長の半分の係数のみを符号化の対象とすればよいため、MDCTを用いることにより効率的な符号化が行える。また、合成時には現フレームと隣接するフレームとを重ね合わせ加算するため、フレーム毎に異なる量子化歪が発生する状況においてもフレーム境界での不連続が発生しにくいという特徴がある。

[0004] 通常、MDCTによる分析／合成を行う際には窓関数である分析窓および合成窓をそれぞれ対象信号に乘じる。このとき用いる分析窓／合成窓は、隣接フレームとの重ね合わせに位置する部分が傾斜(スロープ)している。この重ね合わせ区間の長さ(つまり、スロープの長さ)と入力フレームのバッファリングに必要な遅延が、MDCT分析／合成により発生する遅延の長さに相当する。双方向通信においてこの遅延が大きくなると、通信相手端末からのレスポンスが自端末に届くまでに時間がかかるてしまい、円滑な会話が成り立たなくなる。そのため、遅延はできるだけ短いことが望ましい。

[0005] 以下、従来のMDCTについて説明する。

[0006] MDCTに使用される分析窓／合成窓は、式(1)で表される条件を満たす場合に完全再構成(量子化歪がない前提で変換による歪が0になる)を実現する。

[数1]

$$w_{in}(i) \cdot w_{out}(i) + w_{in}(i + N/2) \cdot w_{out}(i + N/2) = 1 \quad \dots (1)$$

$$(0 \leq i < N)$$

[0007] 式(1)の条件を満たす代表的な窓として、非特許文献1では、式(2)で表されるsin窓が提案されている。また、このsin窓を図示すると図1のようになる。このようなsin窓を用いる場合、sin窓のスペクトル特性はサイドローブが十分に減衰されたため、精度よくスペクトルを分析することが可能である。

[数2]

$$w(i) = \sin\left(\frac{i\pi}{N}\right) \quad \dots (2)$$

$$(0 \leq i < N)$$

[0008] また、非特許文献2では、式(1)の条件を満たす窓として式(3)で表される窓を用いてMDCT分析／合成を行う方法が提案されている。ここで、Nは分析窓の長さ、Lは重ね合わせ区間の長さを示している。また、式(3)で表される窓を図示すると図2のようになる。このような窓を用いた場合、重ね合わせ区間がLとなるため、この窓による遅延はLと表される。よって、重ね合わせ区間Lを短く設定することにより遅延の発生を抑制することが可能になる。

[数3]

$$w(i) = \begin{cases} 0 & 0 \leq i < \frac{1}{4}N - \frac{1}{2}L \\ \cos\left(\frac{\pi \cdot (i - N/4 - L/2)}{2L}\right) & \frac{1}{4}N - \frac{1}{2}L \leq i < \frac{1}{4}N + \frac{1}{2}L \\ 1 & \frac{1}{4}N + \frac{1}{2}L \leq i < \frac{3}{4}N - \frac{1}{2}L \\ \cos\left(\frac{\pi \cdot (i - 3N/4 + L/2)}{2L}\right) & \frac{3}{4}N - \frac{1}{2}L \leq i < \frac{3}{4}N + \frac{1}{2}L \\ 0 & \frac{3}{4}N + \frac{1}{2}L \leq i < N \end{cases} \dots (3)$$

非特許文献1:守谷健弘著、「音声符号化」、(社)電子情報通信学会、1998年10月20日、pp.36-38

非特許文献2:M. Iwadare, et al., "A 128 kb/s Hi-Fi Audio CODEC Based on Adaptive Transform Coding with Adaptive Block Size MDCT," IEEE Journal on Selected Areas in Communications, Vol. 10, No.1, pp.138-144, Jan. 1992.

発明の開示

発明が解決しようとする課題

- [0009] 式(2)で表されるsin窓を用いると、図1に示すように、隣接する分析フレームの重ね合わせ区間が分析フレームの1/2の長さになる。この例では、分析フレーム長がNであるので、重ね合わせ区間はN/2となる。よって、合成側では、N/2~N-1に位置する信号を合成するには次の分析フレームの情報が得られなければ信号を合成することができない。つまり、(3N/2)-1に位置するサンプル値が得られるまで次の分析フレームでMDCT分析を行うことができず、(3N/2)-1の位置のサンプルが得られて初めて次の分析フレームでMDCT分析が行われ、その分析フレームの変換係数を用いてN/2~N-1の信号を合成することが可能になる。従って、sin窓を用いる場合には、N/2の長さの遅延が発生することになる。
- [0010] 一方、式(3)で表される窓を用いると、重ね合わせ区間Lが短いためにフレーム間の不連続が生じやすい。現分析フレームと次分析フレーム各々でMDCT分析を行い、その変換係数を量子化する場合、各々独立に量子化が行われるため、現分析フ

レームと次分析フレームとではそれぞれ異なる量子化歪が発生する。量子化歪が付加された変換係数を時間領域に逆変換すると、その時間信号には合成フレーム全体にわたって量子化歪が付加されることになる。つまり、現合成フレームの量子化歪と次の合成フレームの量子化歪は無相関に発生する。そのため、重ね合わせ区間が短い場合、合成フレーム間の隣接部分において、量子化歪に起因する復号信号の不連続性が十分に吸収できず、フレーム間の接続歪が知覚されてしまう。この傾向は、重ね合わせ区間Lを短くするほどに顕著に発生する。

- [0011] 本発明の目的は、遅延量を低遅延に抑え、かつ、接続歪を緩和することができる音声符号化装置および音声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] 本発明の音声符号化装置は、時間領域の音声信号の1フレームを長い分析長と短い分析長の双方でMDCT分析して、周波数領域の2種の変換係数を得る第1の分析手段と、前記2種の変換係数を符号化する符号化手段と、を具備する構成を探る。

発明の効果

- [0013] 本発明によれば、遅延量を低遅延に抑え、かつ、接続歪を緩和することができる。

図面の簡単な説明

- [0014] [図1]従来の分析窓を示す図

- [図2]従来の分析窓を示す図

- [図3]本発明の実施の形態1に係る音声符号化装置および音声復号化装置の構成を示すブロック図

- [図4]本発明の実施の形態1に係る音声符号化装置の構成を示すブロック図

- [図5]本発明の実施の形態1に係る音声符号化装置の信号状態遷移図

- [図6]本発明の実施の形態1に係る分析窓を示す図

- [図7]本発明の実施の形態1に係る音声復号化装置の構成を示すブロック図

- [図8]本発明の実施の形態1に係る音声復号化装置の信号状態遷移図

- [図9]本発明の実施の形態1に係る音声符号化装置の動作説明図

- [図10]本発明の実施の形態1に係る分析窓を示す図

- [図11]本発明の実施の形態1に係る分析窓を示す図

[図12]本発明の実施の形態2に係る分析窓を示す図

[図13]本発明の実施の形態2に係る音声符号化装置の構成を示すブロック図

[図14]本発明の実施の形態2に係る音声復号化装置の構成を示すブロック図

発明を実施するための最良の形態

[0015] 以下、本発明の実施の形態について、添付図面を参照して詳細に説明する。

[0016] (実施の形態1)

本発明の実施の形態1に係る音声符号化装置および音声復号化装置の構成は図3に示すようになる。この図に示すように、音声符号化装置は、フレーム構成部10、分析部20、変換係数符号化部30を備える。また、音声復号化装置は、変換係数復号化部50、合成部60、フレーム接続部70を備える。

[0017] 音声符号化装置において、フレーム構成部10は、入力される時間領域の音声信号をフレーム化する。分析部20は、フレーム化された時間領域の音声信号をMDCT分析により周波数領域信号に変換する。変換係数符号化部30は、分析部20により得られた変換係数を符号化し符号化パラメータを出力する。この符号化パラメータは、伝送路を介して音声復号化装置へ送られる。

[0018] 音声復号化装置において、変換係数復号化部50は、伝送路を介して送られてきた符号化パラメータを復号する。合成部60は、復号された変換係数からMDCT合成により時間領域信号を生成する。フレーム接続部70は、隣接するフレーム間で不連続が生じないように時間領域信号を接続して復号音声信号を出力する。

[0019] 次いで、音声符号化装置についてより詳細に説明する。音声符号化装置のより詳細な構成は図4に示すようになり、信号の状態遷移は図5に示すようになる。なお、図4に示す信号A～Gは、図5に示す信号A～Gにそれぞれ対応する。

[0020] フレーム構成部10に音声信号Aが入力されると、フレーム構成部10にて長分析用の分析フレーム区間(長分析フレーム)および短分析用の分析フレーム区間(短分析フレーム)が定められる。そして、フレーム構成部10は、長分析フレーム信号Bを長分析部21の窓掛け部211に出力し、短分析フレーム信号Cを短分析部22の窓掛け部221に出力する。なお、長分析フレーム長(長分析窓長)と短分析フレーム長(短分析窓長)はあらかじめ定められており、ここでは長分析フレーム長をM1、短分析フレ

ーム長をM2(M1>M2)として説明する。よって、生じる遅延は、M2／2となる。

- [0021] 長分析部21では、窓掛け部211が、分析長(分析窓長)M1の長分析フレーム信号Bに対して分析窓を乗じ、この分析窓を乗じた信号DをMDCT部212に出力する。分析窓には、図6に示す長分析窓を用いる。この長分析窓は、式(3)を基に、分析長M1、重ね合わせ区間M2／2として設計したものである。
- [0022] MDCT部212は、式(4)に従って信号Dに対してMDCTを施す。そして、MDCT部212は、このMDCTにより得られる変換係数Fを変換係数符号化部30に出力する。なお、式(4)において、{s1(i); 0≤i< M1}は長分析フレームに含まれる時間信号、{X1(k); 0≤k< M1／2}は長分析により得られる変換係数Fを表す。

[数4]

$$X1(k) = \sqrt{\frac{2}{M1}} \sum_{i=0}^{M1-1} s1(i) \cos\left(\frac{(2i+1+M1/2)(2k+1)\pi}{2 \cdot M1}\right) \quad \dots (4)$$

- [0023] 一方、短分析部22では、窓掛け部221が、分析長(分析窓長)M2の短分析フレーム信号Cに対して分析窓を乗じ、この分析窓を乗じた信号EをMDCT部222に出力する。分析窓には、図6に示す短分析窓を用いる。この短分析窓は、式(2)を基に、分析長M2(M2< M1)として設計したものである。

- [0024] MDCT部222は、式(5)に従って信号Eに対してMDCTを施す。そして、MDCT部222は、このMDCTにより得られる変換係数Gを変換係数符号化部30に出力する。なお、式(5)において、{s2(i); 0≤i< M2}は短分析フレームに含まれる時間信号、{X2(k); 0≤k< M2／2}は短分析により得られる変換係数Gを表す。

[数5]

$$X2(k) = \sqrt{\frac{2}{M2}} \sum_{i=0}^{M2-1} s2(i) \cos\left(\frac{(2i+1+M2/2)(2k+1)\pi}{2 \cdot M2}\right) \quad \dots (5)$$

- [0025] 変換係数符号化部30は、変換係数F:{X1(k)}および変換係数G:{X2(k)}を符号化し、それぞれの符号化パラメータを時間多重して出力する。この際、変換係数符

号化部30は、変換係数 $\{X_2(k)\}$ に対し、変換係数 $\{X_1(k)\}$ の符号化に比べより精度の高い(より量子化誤差の小さい)符号化を行う。例えば、変換係数符号化部30は、変換係数 $\{X_2(k)\}$ の変換係数当りの符号化ビット数を変換係数 $\{X_1(k)\}$ の変換係数当りの符号化ビット数より高く設定して、変換係数 $\{X_1(k)\}$ および変換係数 $\{X_2(k)\}$ を符号化する。すなわち、変換係数符号化部30では、変換係数 $\{X_1(k)\}$ に比べ変換係数 $\{X_2(k)\}$ の量子化歪みが小さくなる符号化を行う。なお、変換係数符号化部30での符号化方法は、例えば特開2003-323199号公報に記載された符号化方法等を用いることができる。

- [0026] 次いで、音声復号化装置についてより詳細に説明する。音声復号化装置のより詳細な構成は図7に示すようになり、信号の状態遷移は図8に示すようになる。なお、図7に示す信号A～Iは、図8に示す信号A～Iにそれぞれ対応する。
- [0027] 変換係数復号化部50に符号化パラメータが入力されると、変換係数復号化部50にて復号変換係数(長分析) $\{X_{1q}(k); 0 \leq k < M_1/2\}$:Aと復号変換係数(短分析) $\{X_{2q}(k); 0 \leq k < M_2/2\}$:Bとが復号される。そして、変換係数復号化部50は、復号変換係数 $\{X_{1q}(k)\}$:Aを長合成部61のIMDCT部611に出力し、復号変換係数 $\{X_{2q}(k)\}$:Bを短合成部62のIMDCT部621に出力する。
- [0028] 長合成部61では、IMDCT部611が、復号変換係数 $\{X_{1q}(k)\}$ に対してIMDCT(MDCT部212で行われたMDCTの逆変換)を行って長合成信号Cを生成し、この長合成信号Cを窓掛け部612に出力する。
- [0029] 窓掛け部612は、長合成信号Cに対して合成窓を乗じ、この合成窓を乗じた信号Eをフレーム内接続部71に出力する。合成窓には、音声符号化装置の窓掛け部211同様、図6に示す長分析窓を用いる。
- [0030] 一方、短合成部62では、IMDCT部621が、復号変換係数 $\{X_{2q}(k)\}$ に対してIMDCT(MDCT部222で行われたMDCTの逆変換)を行って短合成信号Dを生成し、この短合成信号Dを窓掛け部622に出力する。
- [0031] 窓掛け部622は、短合成信号Dに対して合成窓を乗じ、この合成窓を乗じた信号Fをフレーム内接続部71に出力する。合成窓には、音声符号化装置の窓掛け部221同様、図6に示す短分析窓を用いる。

- [0032] そして、フレーム内接続部71にて第nフレームの復号信号Gを生成し、次に、フレーム間接続部73にて第nフレームの復号信号Gと第n-1フレームの復号信号Hの対応する区間を重ね合わせ加算して復号音声信号を生成する。すなわち、フレーム内接続部71では、信号Eと信号Fの対応する区間を重ね合わせ加算して第nフレームの復号信号{sq(i); 0≤i< M1}:Gを生成する。そして、フレーム間接続部73では、第nフレームの復号信号Gとバッファ72にバッファリングされている第n-1フレームの復号信号Hの対応する区間を重ね合わせ加算して復号音声信号Iを生成する。その後、第nフレームの復号信号Gは、次フレーム(第n+1フレーム)の処理のためにバッファ72に格納される。
- [0033] 次いで、図9に、音声信号の各フレームと分析部20での分析位置の対応関係を示す。図9に示すように、本実施の形態では、1フレーム区間(符号化パラメータを生成する単位)の音声信号の分析は、いずれのフレームも必ず長分析と短分析の組み合わせにより行われる。
- [0034] このように、本実施の形態では、MDCT分析を長い分析長(長分析)と短い分析長(短分析)の組み合わせで行い、かつ、短分析により得られる変換係数の量子化誤差が小さくなるように符号化処理を行うため、遅延が短い状態で分析長を長くとて冗長性を効率的に取り除くことができるとともに、短分析の変換係数の量子化歪を小さくすることができる。よって、遅延量をM2/2という低遅延に抑え、かつ、接続歪を緩和することができる。
- [0035] なお、1フレーム区間における長分析窓と短分析窓の配置について、図6では短分析窓を長分析窓より時間的に後方に配置するようにしているが、例えば図10に示すように、長分析窓を短分析窓より時間的に後方に配置するようにしてもよい。図10に示すような配置にしても、図6に示す配置と同様に、遅延量を低遅延に抑え、かつ、接続歪を緩和することができる。
- [0036] また、本実施の形態では、短分析窓を式(2)を基に設計したが、短分析窓の分析長M2と長分析窓の分析長M1との関係がM2< M1となっていれば、式(3)で表される窓を短分析窓として用いてもよい。つまり、式(3)を基に、分析長M2として設計した窓を短分析窓として用いてもよい。この一例を図11に示す。このような分析窓の

構成においても、遅延量を低遅延に抑え、かつ、接続歪を緩和することができる。

[0037] (実施の形態2)

音声符号化装置に入力される音声信号が語頭部や、特性が急激に変化する遷移部である場合、周波数分解能よりもむしろ時間分解能が必要とされる。そのような音声信号に対しては、分析フレームが全て短分析フレームにより分析される方が音声品質が向上する。

[0038] そこで、本実施の形態では、各フレームに対して行うMDCT分析を、入力音声信号の特性に応じて、(1)長分析と短分析の組み合わせで行うモード(長短混合分析モード)と(2)短分析を繰り返し複数回行うモード(全短分析モード)とで切り替えて行う。全短分析モードのときに各フレームに対して使用する分析／合成窓の例を図12に示す。なお、長短混合分析モードについては、実施の形態1で説明したとおりである。

[0039] 本発明の実施の形態2に係る音声符号化装置の構成を図13に示す。この図に示すように、本実施の形態に係る音声符号化装置は、実施の形態1の構成(図4)にさらに判定部15、多重化部35、SW(スイッチ)11、SW12を備える。図13において図4と同一の構成には同一符号を付して説明を省略する。なお、フレーム構成部10から分析部20への出力および分析部20から変換係数符号化部30への出力は実際に図4に示すように並列的に行われるが、ここでは図示の便宜上1本の信号線で示すものとする。

[0040] 判定部15は、入力音声信号を分析して信号の特性を判定する。特性判定では、音声信号の特徴量の時間的な変化を監視し、その変化量が所定量未満の場合には定常部、所定量以上の場合には非定常部であると判定する。音声信号の特徴量としては、例えば、短区間パワー、短区間スペクトルなどが挙げられる。

[0041] そして、判定部15は、判定結果に応じてMDCT分析の分析モードを長短混合分析モードと全短分析モードとで切り替える。すなわち、判定部15は、入力された音声信号が定常部である場合は、SW11およびSW12を分析部20側に接続して、分析部20を使用して長短混合分析モードでMDCT分析を行う。一方、入力された音声信号が非定常部である場合は、SW11およびSW12を全短分析部25側に接続して

、全短分析部25を使用して全短分析モードでMDCT分析を行う。この切替えにより、音声信号のフレームは、定常部である場合には実施の形態1のように長分析と短分析の組み合わせで分析され、非定常部である場合には短分析を繰り返し複数回行われる。

- [0042] 判定部15により全短分析モードが選択された場合には、全短分析部25は、分析窓長をM2としたときの式(2)で表される分析窓を用い、式(5)に示されるMDCTにより分析を行う。
- [0043] また、判定部15は、定常部であるか非定常部であるかを示す判定情報を符号化して多重化部35に出力する。この判定情報は、多重化部35により、変換係数符号化部30から出力される符号化パラメータと多重されて出力される。
- [0044] 一方、本発明の実施の形態2に係る音声復号化装置の構成を図14に示す。この図に示すように、本実施の形態に係る音声復号化装置は、実施の形態1の構成(図7)にさらに分離部45、判定情報復号化部55、全短合成部65、SW21、SW22を備える。図14において図7と同一の構成には同一符号を付して説明を省略する。なお、変換係数復号化部50から合成部60への出力および合成部60からフレーム内接続部71への出力は実際には図7に示すように並列的に行われるが、ここでは図示の便宜上1本の信号線で示すものとする。
- [0045] 分離部45は、入力される符号化パラメータを判定情報を表す符号化パラメータと変換係数を表す符号化パラメータに分離し、それぞれ判定情報復号化部55と変換係数復号化部50に出力する。
- [0046] 判定情報復号化部55は、入力された判定情報を復号する。そして、判定情報復号化部55は、判定情報が定常部を示す場合は、SW21およびSW22を合成部60側に接続して、合成部60を使用して合成信号を生成する。なお、合成部60を使用した合成信号の生成については、実施の形態1で説明したとおりである。一方、判定情報が非定常部を示す場合は、SW21およびSW22を全短合成部65側に接続して、全短合成部65を使用して合成信号を生成する。全短合成部65は、1フレームにおける複数の復号変換係数(短分析)それぞれにIMDCT処理を施して合成信号を生成する。
- [0047] このように、本実施の形態では、1フレームにおいて、入力音声信号が定常部で安

定している場合はそのフレームの音声信号を長分析および短分析の組合せで分析し、入力音声信号が非定常部である場合(入力音声信号が急激に変化する場合)はそのフレームの音声信号を全て短分析で分析して時間分解能を上げるため、入力音声信号の特性に応じて最適なMDCT分析を行うことができ、入力音声信号の特性が変化する場合でも音声品質を良好に保つことができる。

- [0048] また、本実施の形態では、長短混合分析モードの重ね合わせ区間と全短分析モードの重ね合わせ区間を同一にする。よって、例えばISO/IEC IS 13818-7 Information technology -- Generic coding of moving pictures and associated audio information -- Part 7: Advanced Audio Coding (AAC)に記載されているLONG_START_WINDOW またはLONG_STOP_WINDOWのような遷移用の分析フレームを用いる必要がない。
- [0049] また、長短混合分析モードか全短分析モードかを判定する別の方法として、次フレームとの接続部に位置する信号の原信号に対するSNRの大きさによって判定する方法が挙げられる。この判定方法を用いることにより、接続部のSNRによって次フレームの分析モードを決定することができるため、分析モードの誤判定を減らすことができる。
- [0050] なお、上記各実施形態は、レイヤ数が2以上の階層符号化の拡張レイヤに対して適用可能である。
- [0051] また、上記各実施の形態に係る音声符号化装置、音声復号化装置を、移動体通信システムにおいて使用される無線通信移動局装置や無線通信基地局装置等の無線通信装置に搭載することも可能である。
- [0052] また、上記実施の形態では、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、本発明はソフトウェアで実現することも可能である。
- [0053] また、上記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。
- [0054] 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。
- [0055] 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)や、LSI内部の回路セルの接続や設定を再構成可能なリコンフィギュラブル・プロセッサーを利用してもよい。

[0056] 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。バイオ技術の適応等が可能性としてありえる。

[0057] 本明細書は、2004年10月26日出願の特願2004-311143に基づくものである。この内容はすべてここに含めておく。

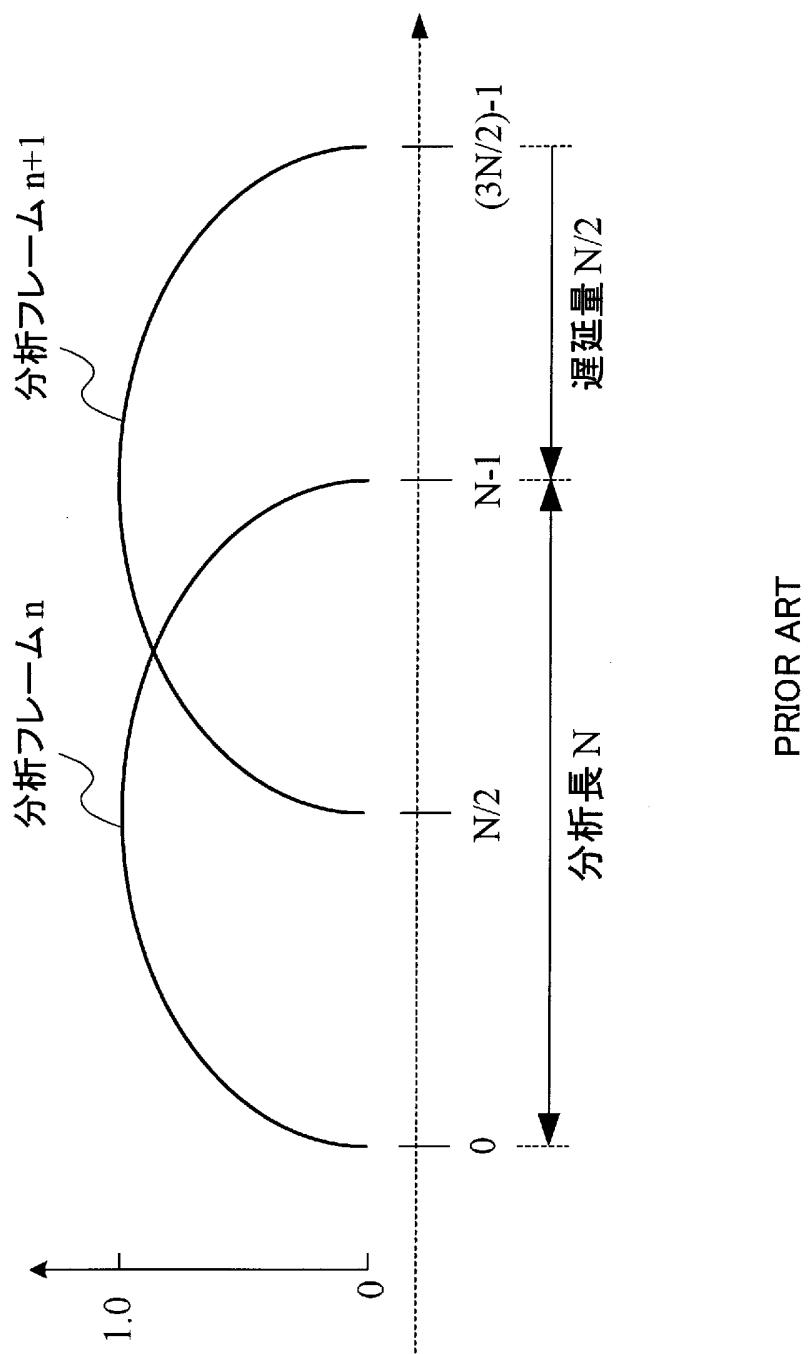
産業上の利用可能性

[0058] 本発明は、移動体通信システムやインターネットプロトコルを用いたパケット通信システム等における通信装置の用途に適用できる。

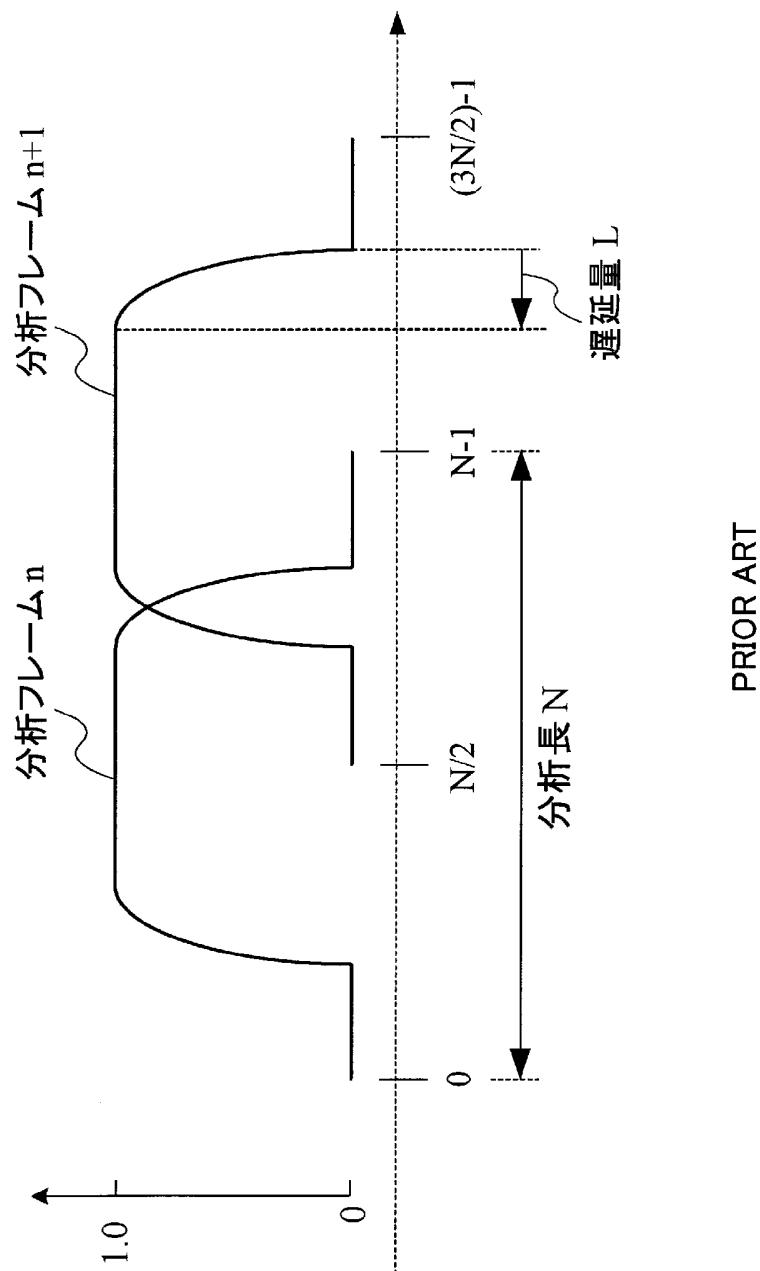
請求の範囲

- [1] 時間領域の音声信号の1フレームを長い分析長と短い分析長の双方でMDCT分析して、周波数領域の2種の変換係数を得る第1の分析手段と、
前記2種の変換係数を符号化する符号化手段と、
を具備する音声符号化装置。
- [2] 前記符号化手段は、前記長い分析長により得られる第1の変換係数に対する符号化に比べより精度の高い符号化を前記短い分析長により得られる第2の変換係数に対して行う、
請求項1記載の音声符号化装置。
- [3] 前記音声信号が定常部または非定常部のいずれであるかを判定する判定手段と、
非定常部である場合に、前記1フレームを前記短い分析長で繰り返し複数回MDCT分析する第2の分析手段と、
をさらに具備する請求項1記載の音声符号化装置。
- [4] 請求項1記載の音声符号化装置を具備する無線通信移動局装置。
- [5] 請求項1記載の音声符号化装置を具備する無線通信基地局装置。
- [6] 時間領域の音声信号の1フレームを長い分析長と短い分析長の双方でMDCT分析して、周波数領域の2種の変換係数を得る分析工程と、
前記2種の変換係数を符号化する符号化工程と、
を具備する音声符号化方法。

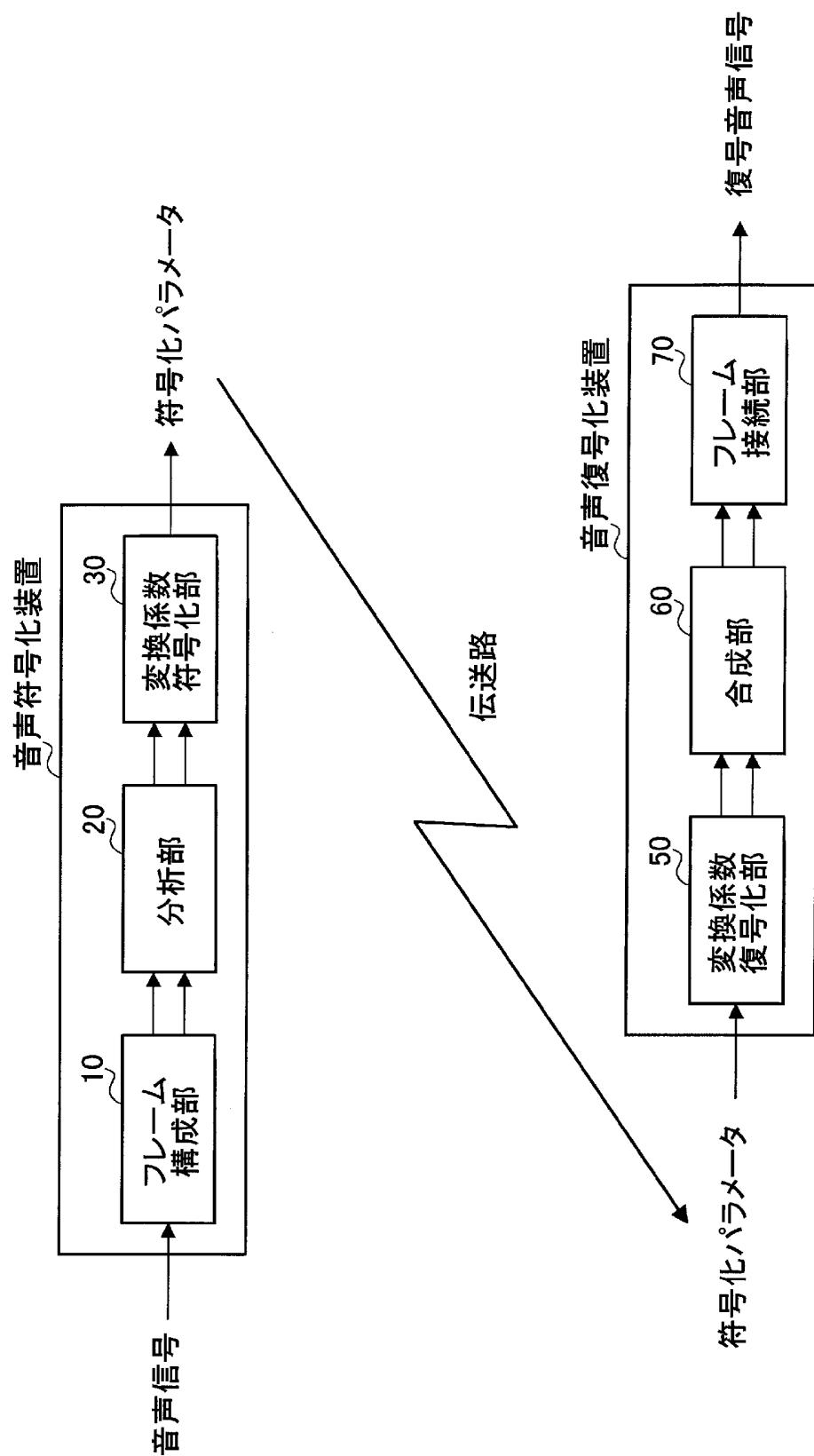
[図1]



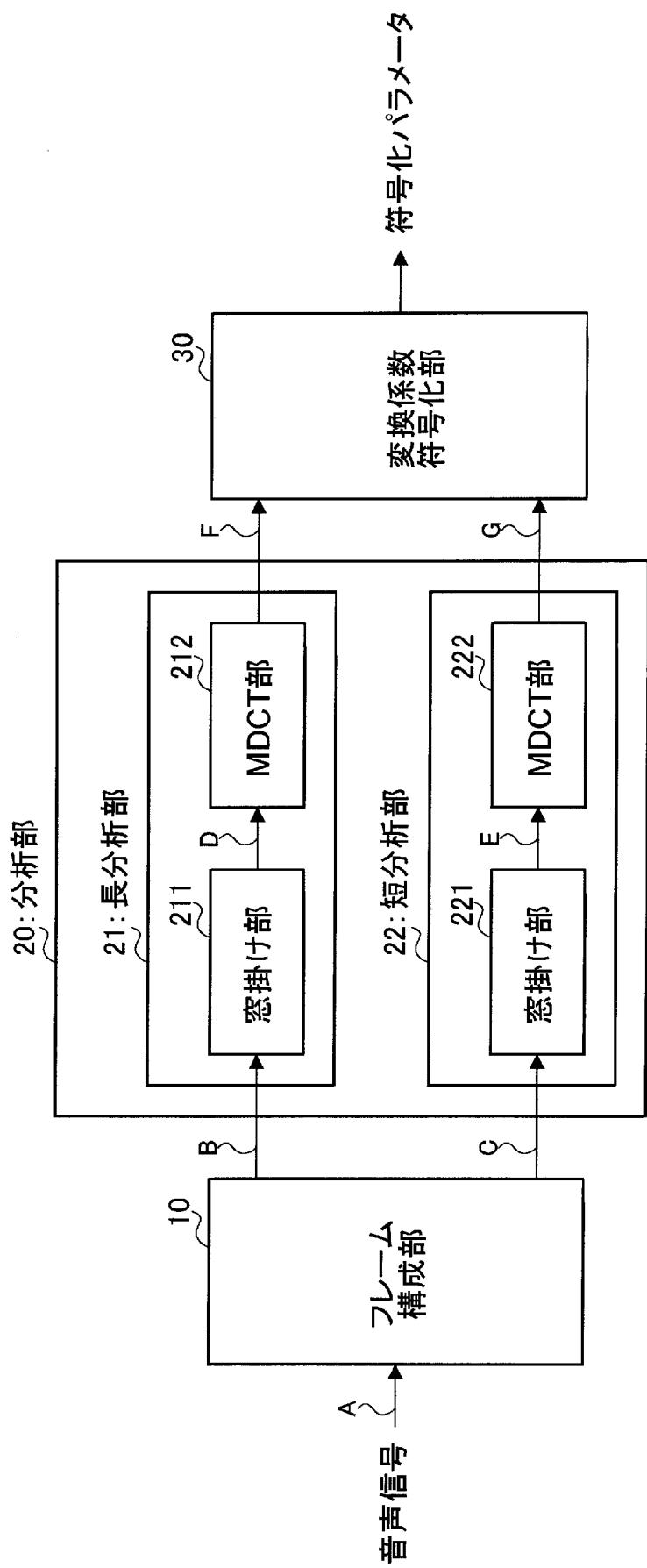
[図2]



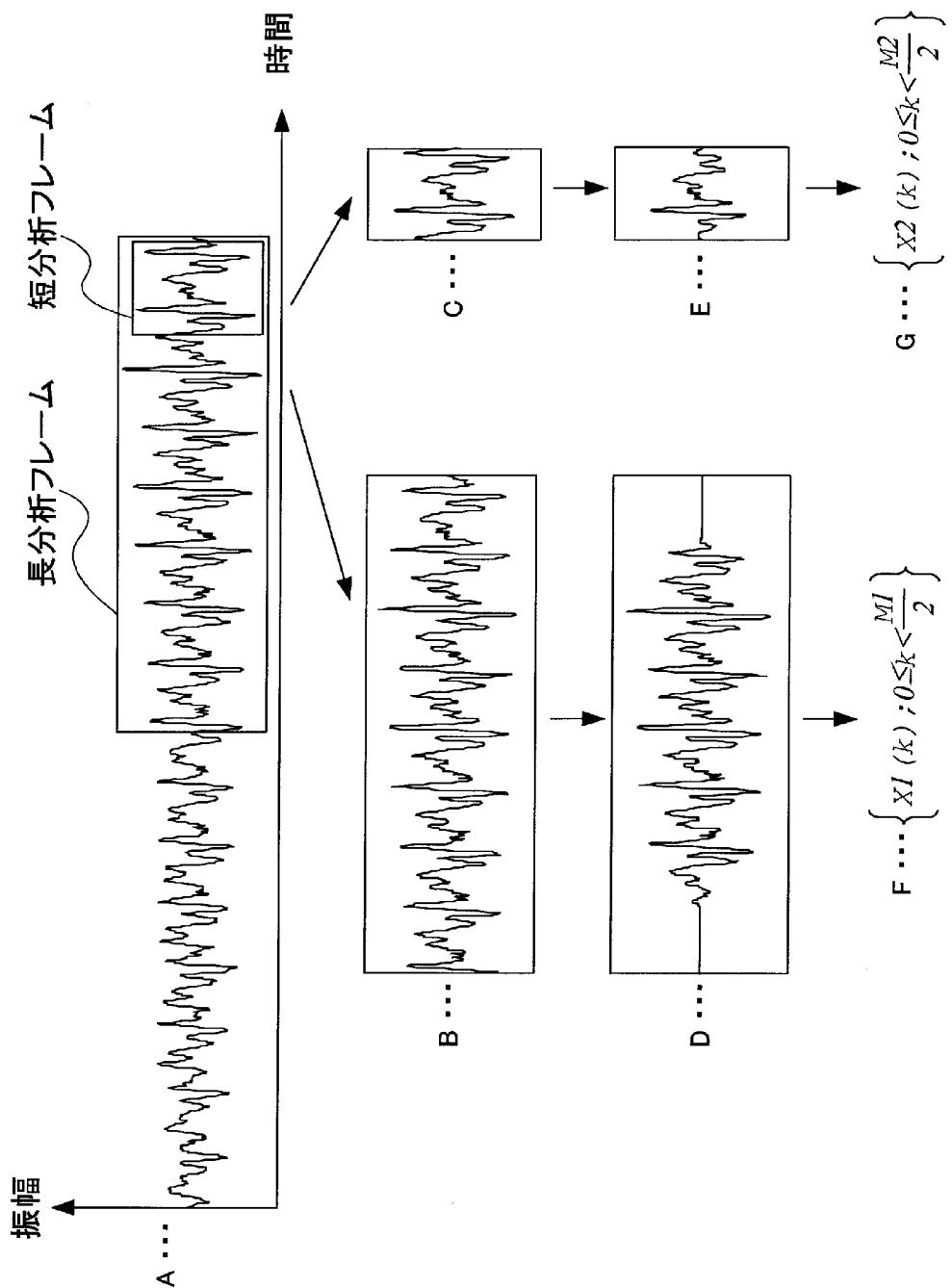
[図3]



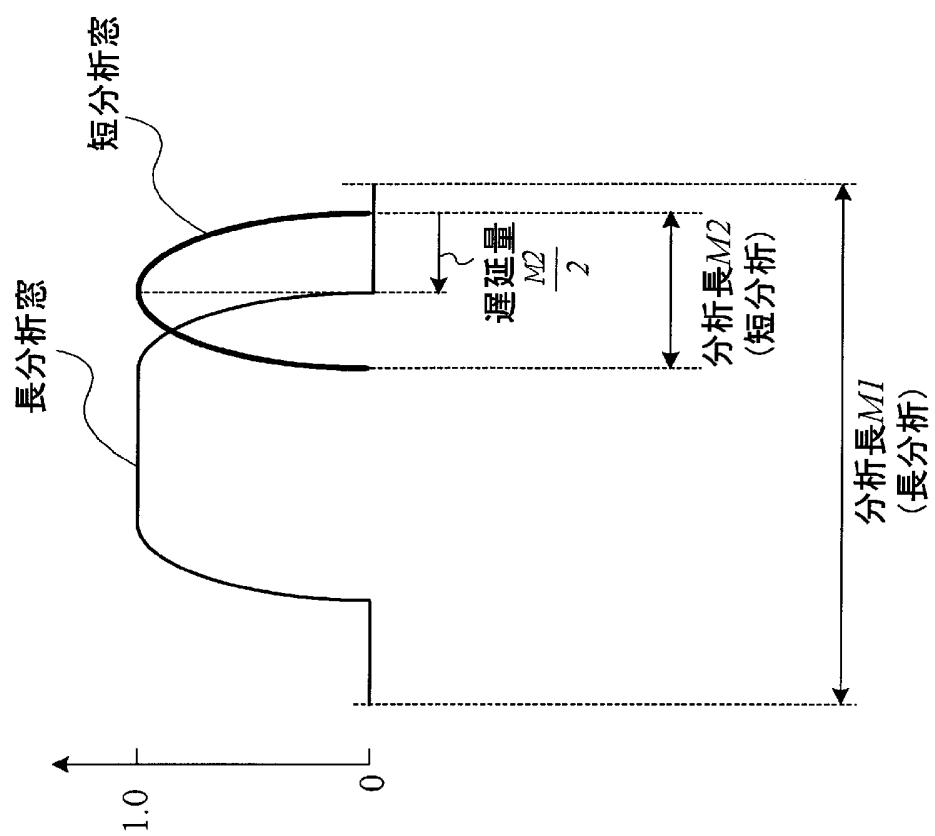
[図4]



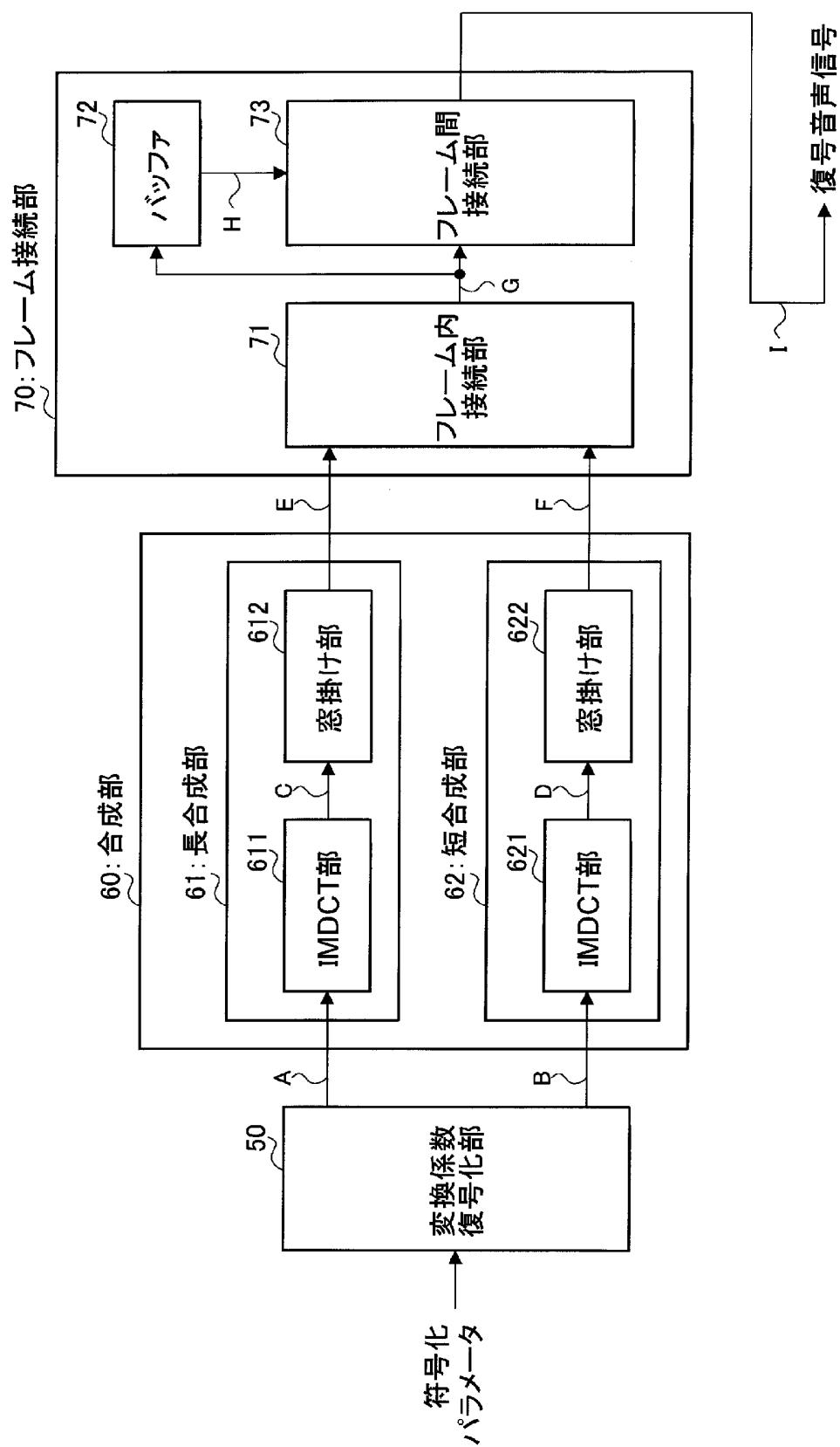
[図5]



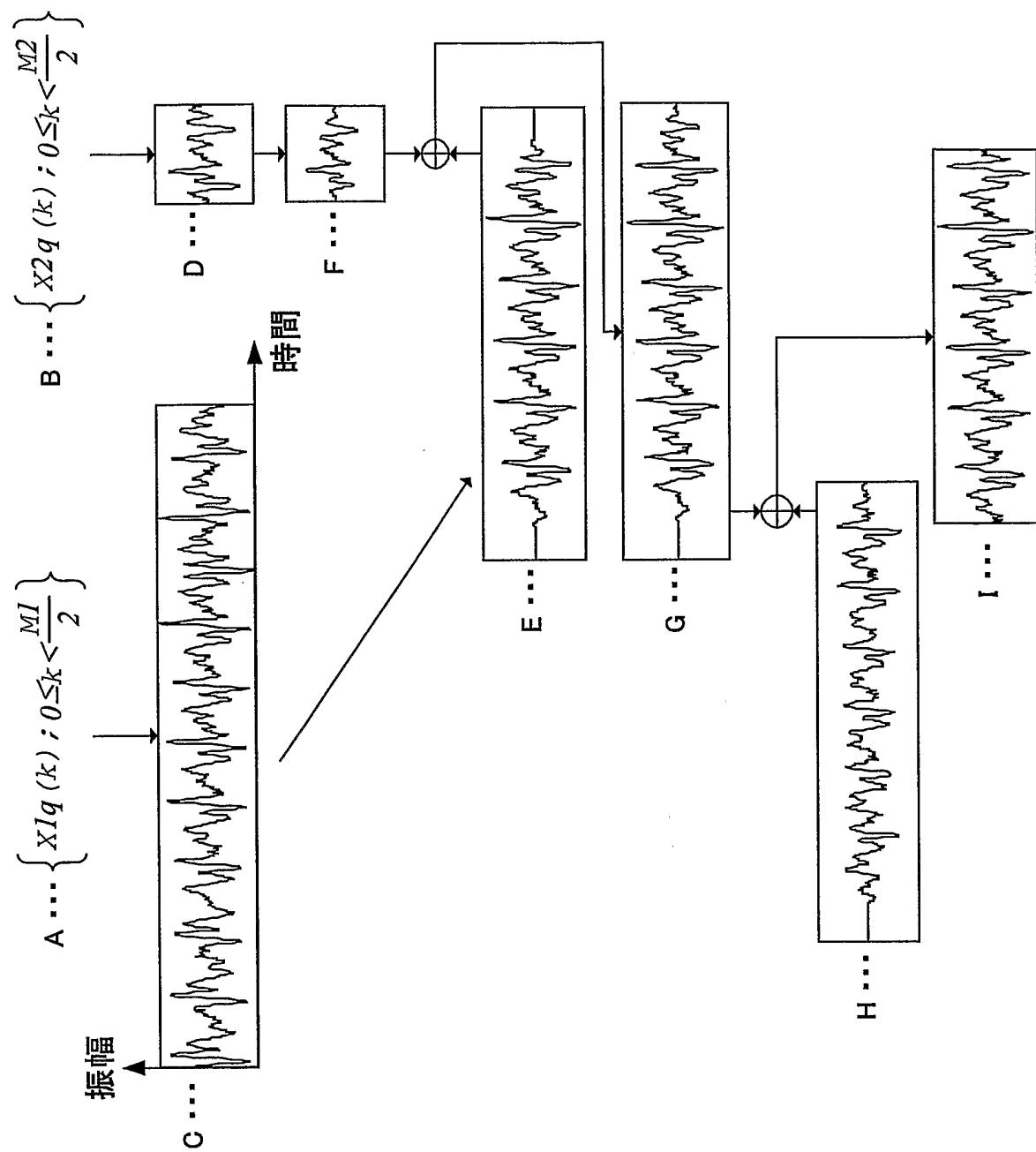
[図6]



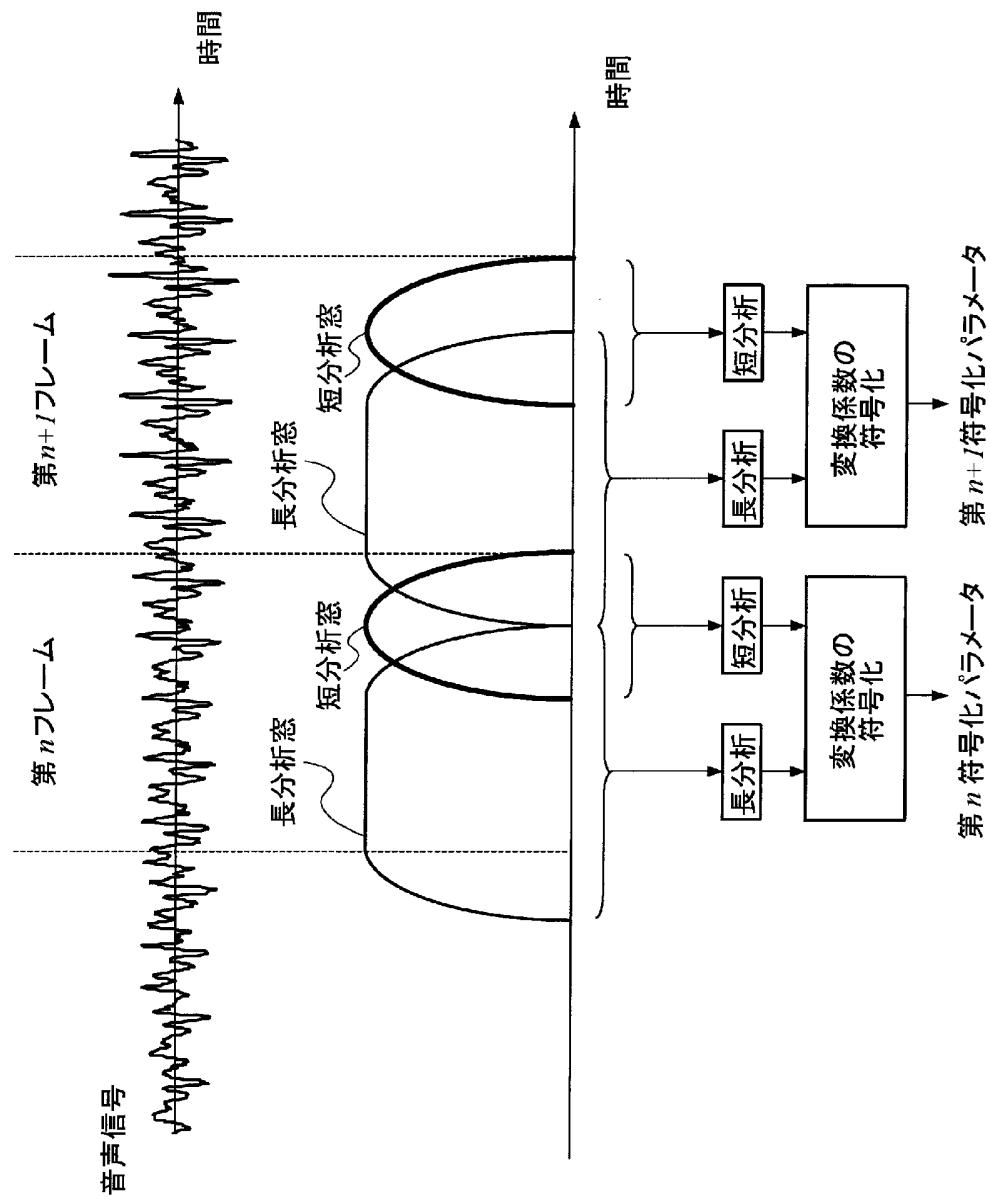
[図7]



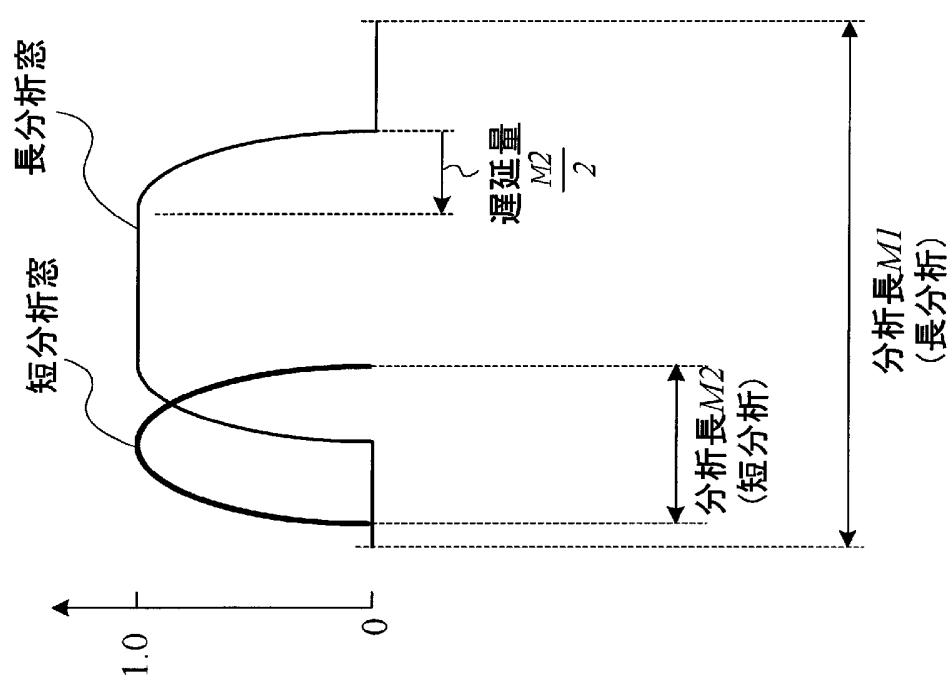
[図8]



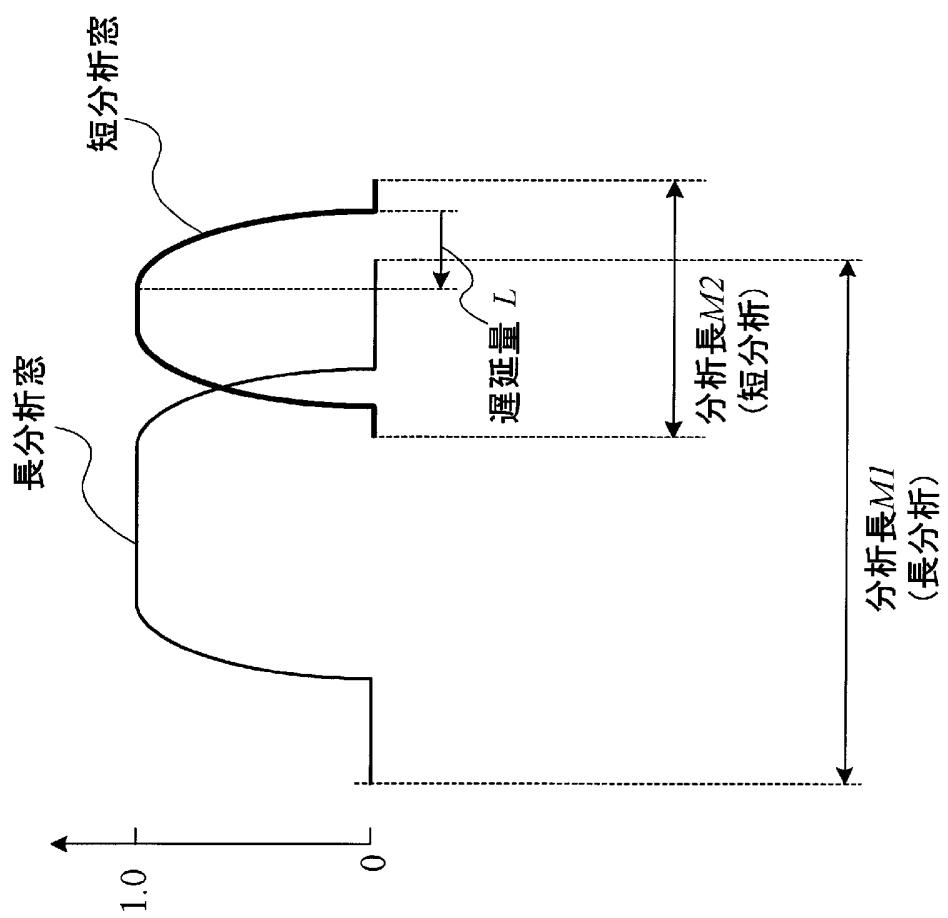
[図9]



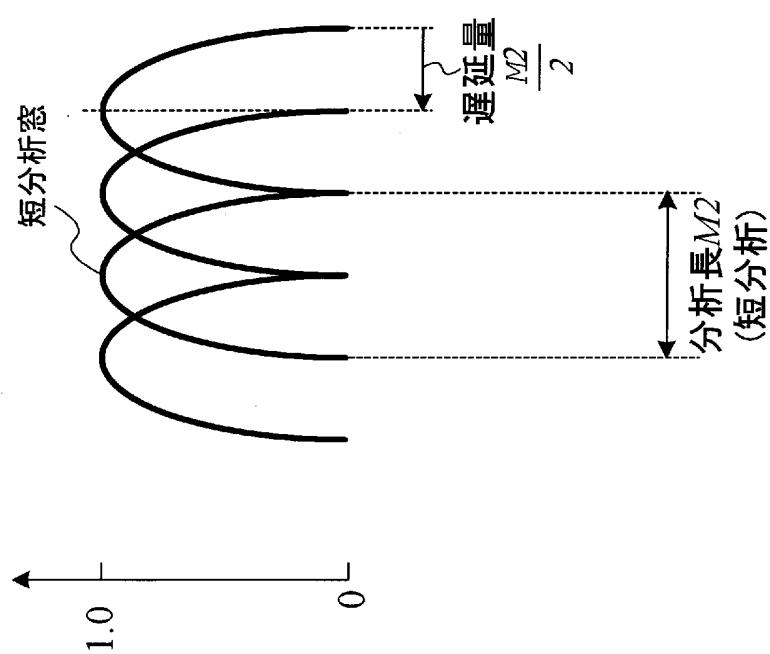
[図10]



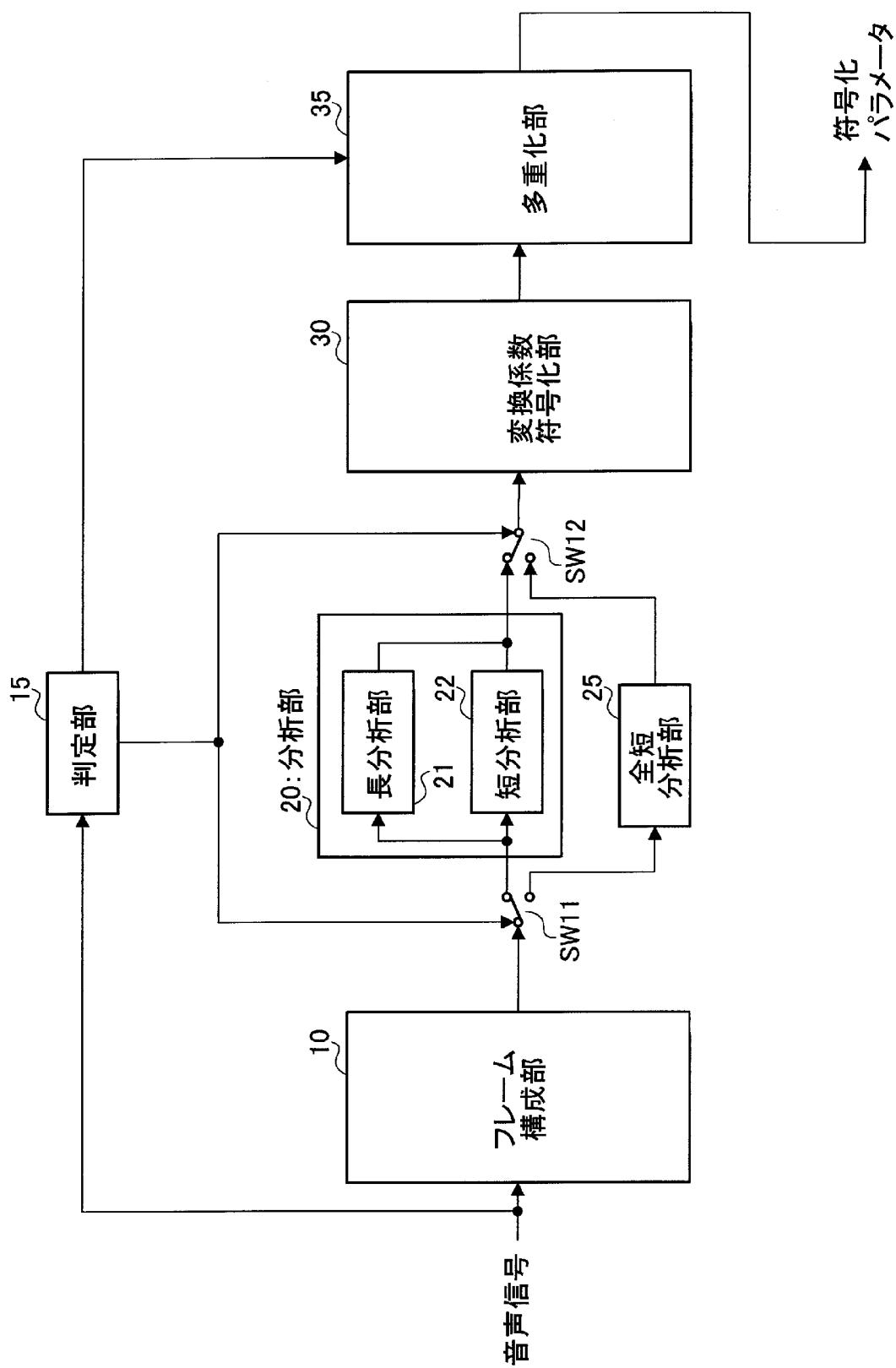
[図11]



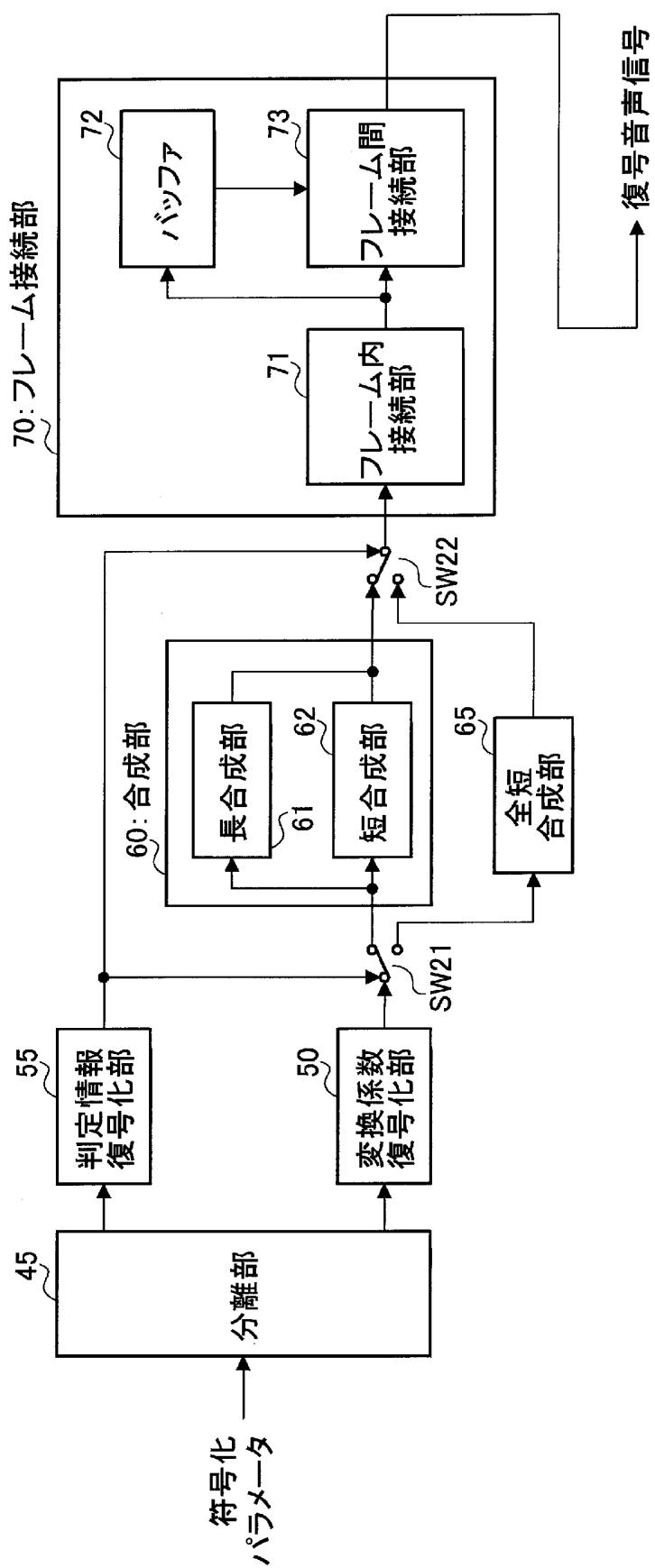
[図12]



[図13]



[図14]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2005/019578

A. CLASSIFICATION OF SUBJECT MATTER
G10L19/02 (2006.01), H03M7/30 (2006.01)

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)

G10L19/02 (2006.01), H03M7/30 (2006.01)

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	1996-2005
Kokai Jitsuyo Shinan Koho	1971-2005	Toroku Jitsuyo Shinan Koho	1994-2005

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	JP 6-268608 A (Sony Corp.), 22 September, 1994 (22.09.94), Par. No. [0037]; Fig. 3 & US 5761642 A1 & EP 0620653 A2 & AU 5635994 A1	1, 3-6 2
Y	JP 2003-66998 A (Mitsubishi Electric Corp.), 05 March, 2003 (05.03.03), Par. Nos. [0045], [0046] (Family: none)	2
X	JP 2000-500247 A (Fraunhofer-Gesellschaft zur Forderung der Angewandten Forschung E.V.), 11 January, 2000 (11.01.00), Full text & WO 1998/002971 A1 & US 5848391 A & EP 0910900 A1	1, 3-6

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"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)

"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

"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

"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

"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

"&" document member of the same patent family

Date of the actual completion of the international search
12 December, 2005 (12.12.05)

Date of mailing of the international search report
27 December, 2005 (27.12.05)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

A. 発明の属する分野の分類（国際特許分類（IPC））

Int.Cl. G10L19/02 (2006.01), H03M7/30 (2006.01)

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. G10L19/02 (2006.01), H03M7/30 (2006.01)

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	1971-2005年
日本国実用新案登録公報	1996-2005年
日本国登録実用新案公報	1994-2005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）

JSTPlus(JOIS)

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 請求の範囲の番号
X	J P 6-268608 A (ソニー株式会社) 1994.09.22, 段落【0037】，第3図	1, 3-6
Y	& U S 5761642 A1 & E P 0620653 A2 & A U 5635994 A1	2
Y	J P 2003-66998 A (三菱電機株式会社) 2003.03.05, 段落【0045】，【0046】(ファミリー なし)	2

 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

の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

12.12.2005

国際調査報告の発送日

27.12.2005

国際調査機関の名称及びあて先

日本国特許庁 (ISA/JP)

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）

山下 剛史

5Z 3454

電話番号 03-3581-1101 内線 3541

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 請求の範囲の番号
X	JP 2000-500247 A (フラオホッフェルーゲゼルシヤフト ツル フェルデルング デル アンゲヴァンドテン フォルシュング エー. ヴェー.) 2000. 01. 11, 全文 & WO 1998/002971 A1 & US 5848391 A & EP 0910900 A1	1, 3-6