

(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
2013年5月2日(02.05.2013)



(10) 国際公開番号
WO 2013/061584 A1

- (51) 国際特許分類:
G10L 19/14 (2006.01) G10L 19/02 (2013.01)
- (21) 国際出願番号: PCT/JP2012/006802
- (22) 国際出願日: 2012年10月24日(24.10.2012)
- (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 2011-236912 2011年10月28日(28.10.2011) 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: 石川 智一 (ISHIKAWA, Tomokazu). 則松武志 (NORIMATSU, Takeshi). チョン コック セン (CHONG, Kok Seng). ザオ ダン (ZHAO, Dan).
- (74) 代理人: 新居 広守 (NIJ, Hiromori); 〒5320011 大阪府大阪市淀川区西中島5丁目3番10号タナカ・イトーピア新大阪ビル6階新居国際特許事務所内 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, 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, KM, 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,

[続葉有]

(54) Title: HYBRID SOUND-SIGNAL DECODER, HYBRID SOUND-SIGNAL ENCODER, SOUND-SIGNAL DECODING METHOD, AND SOUND-SIGNAL ENCODING METHOD

(54) 発明の名称: 音信号ハイブリッドデコーダ、音信号ハイブリッドエンコーダ、音信号復号方法、及び音信号符号化方法

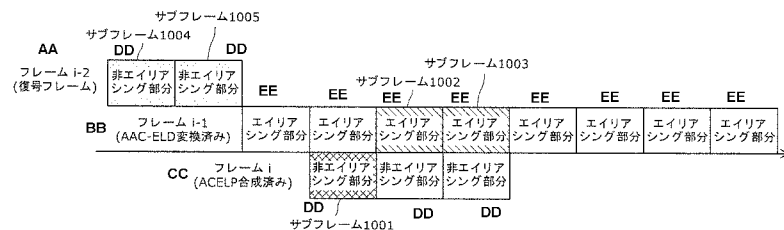
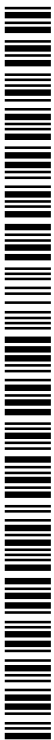


FIG. 11
 1001, 1002, 1003, 1004, 1005 Subframe
 AA Frame i-2 (decoded frame)
 BB Frame i-1 (AAC-ELD conversion complete)
 CC Frame i (ACELP synthesis complete)
 DD Aliasing region
 EE Non-aliasing region

(57) Abstract: A hybrid sound-signal decoder that decodes a bit stream containing the following: audio frames encoded by an audio encoding process using a low-delay filter bank; and voice frames encoded by a voice encoding process using linear prediction coefficients. If the frame being decoded (the *i*th frame) is the first voice frame after a switch from audio frames to voice frames, subframes (1101 and 1102) that are signals corresponding to a pre-encoding (*i*-1)th frame are generated using a subframe (1001) that is a signal generated using a pre-encoding (*i*-1)th-frame signal, obtained by decoding the *i*th frame.

(57) 要約: 低遅延フィルタバンクを用いた音響符号化処理で符号化された音響フレームと、線形予測係数を用いた音声符号化処理で符号化された音声フレームとが含まれるビットストリームを復号する音信号ハイブリッドデコーダであって、復号対象フレームが、音響フレームから音声フレームに切り替えられた最初の音声フレームである第 *i* フレームであるとき、第 *i* フレームを復号することで得られる、符号化前の第 *i*-1 フレームの信号を用いて生成された信号であるサブフレーム (1001) を用いて符号化前の第 *i*-1 フレームに対応する信号であるサブフレーム (1101) 及びサブフレーム (1102) を生成する。



WO 2013/061584 A1

NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI 添付公開書類:
(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, — 国際調査報告 (条約第 21 条(3))
NE, SN, TD, TG).

明 細 書

発明の名称：

音信号ハイブリッドデコーダ、音信号ハイブリッドエンコーダ、音信号復号方法、及び音信号符号化方法

技術分野

[0001] 本発明は、音声コーデックと音響コーデックとを切替可能な音信号ハイブリッドデコーダ及び音信号ハイブリッドエンコーダに関する。

背景技術

[0002] ハイブリッドコーデック（例えば、特許文献1参照）は、音響コーデックと音声コーデック（例えば、非特許文献1参照）の利点を組み合わせたコーデックである。ハイブリッドコーデックによれば、音声信号主体のコンテンツと音響信号主体のコンテンツが混合した音信号を、音響コーデックと音声コーデックとを切替えることによりそれぞれに適した符号化方法により符号化することができる。したがって、ハイブリッドコーデックによれば、低ビットレートで安定した音信号の圧縮符号化が実現される。

先行技術文献

特許文献

[0003] 特許文献1：Fuchs, Guillaume 「Apparatus and method for encoding/decoding an audio signal using an aliasing switch scheme」、国際公開第2010/003532 A1号

非特許文献

[0004] 非特許文献1：Milan Jelinek, 「Wideband Speech Coding Advances in VMR-WB Standard」、IEEE Transactions on Audio, Speech and Language Processing, 15 (4), 1167-1179 (2007)

非特許文献2: Chi-Min Liu and Wen-Chieh Lee, 「A unified fast algorithm for cosine modulated filterbanks in current audio standards」, J. Audio Engineering 47 (12), 1061-1075 (1999)

発明の概要

発明が解決しようとする課題

[0005] ハイブリッドコーデックの音質を向上させるために、音響コーデックとして、例えば、AAC-ELD (Advanced Audio Coding - Enhanced Low Delay) モードを用いることで音質向上が期待できる。

[0006] しかしながら、AAC-ELDモードのような符号化方式では、先行フレームとオーバーラップしたサンプルを用いて符号化を行うために、対象フレーム内のサンプルだけで符号化が完結する音声コーデックに切り替わるときにエイリアシングが生じ、不自然な音が発生する。特許文献1では、このように符号化モードが切り替わる箇所における信号処理について開示されているが、このような処理は、AAC-ELDモードのような複数の先行フレームによるオーバーラップ処理が必要な符号化方式に対応しておらず、特許文献1の方法では、上記エイリアシングを低減することができない。

[0007] 本発明の目的は、音響コーデックとしてAAC-ELDモードのような複数の先行フレームによるオーバーラップ処理が必要な符号化方式を用いた場合に、音声コーデックと音響コーデックとの切替部分において発生するエイリアシングを低減するハイブリッドコーデック（音信号ハイブリッドデコーダ及び音信号ハイブリッドエンコーダ）を提供することである。

課題を解決するための手段

[0008] 本発明の一態様に係る音信号ハイブリッドデコーダは、低遅延フィルタバンクを用いた音響符号化処理で符号化された音響フレームと、線形予測係数を用いた音声符号化処理で符号化された音声フレームとが含まれるビットス

トリームを復号する音信号ハイブリッドデコーダであって、前記音響フレームを低遅延逆フィルタバンク処理を用いて復号する低遅延変換デコーダと、前記音声フレームを復号する音声信号デコーダと、前記ビットストリームのうちの復号対象フレームが前記音響フレームである場合、当該復号対象フレームを前記低遅延変換デコーダによって復号し、前記復号対象フレームが前記音声フレームである場合、当該復号対象フレームを前記音声信号デコーダによって復号する制御を行うブロック切替部とを備え、前記復号対象フレームが、前記音響フレームから前記音声フレームに切り替わった最初の前記音声フレームである第 i フレームであるとき、前記第 i フレームには、前記第 i フレームよりも1フレーム先行するフレームである第 $i-1$ フレームの符号化前の信号を用いて生成された第1信号が符号化された状態で含まれ、前記ブロック切替部は、(1)前記第 i フレームよりも2フレーム先行するフレームである第 $i-2$ フレームを前記低遅延変換デコーダによって復号することで得られる、前記第 i フレームよりも3フレーム先行するフレームである第 $i-3$ フレームの再構成された信号を窓処理した信号である第2信号のフレームの前半部分に相当する信号に、前記第2信号のフレームの後半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第 i フレームを前記音声信号デコーダによって復号することで得られる、前記第1信号に窓処理を行った信号と、前記第 $i-1$ フレームを前記低遅延逆フィルタバンク処理及び窓処理した信号の前記第 $i-3$ フレームに対応する部分である第3信号のフレームの前半部分の信号と、を加算する処理を行って符号化前の前記第 $i-1$ フレームの前半部分に対応する信号を生成し、前記第2信号のフレームの後半部分に相当する信号に、前記第2信号のフレームの前半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第1信号に畳み込み処理及び窓処理を行った信号と、前記第3信号のフレームの後半部分に相当する信号と、を加算する処理を行って符号化前の前記第 $i-1$ フレームの後半部分に対応する信号を生成する、または(2)前記第2信号のフレームの前半部分に相当する信号に、前記

第2信号のフレームの後半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第1信号に畳み込み処理及び窓処理を行った信号と、前記第3信号のフレームの前半部分に相当する信号と、を加算する処理を行って符号化前の前記第*i*-1フレームの前半部分に対応する信号を生成し、前記第2信号のフレームの後半部分に相当する信号に、前記第2信号のフレームの前半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第1信号に窓処理を行った信号と、前記第3信号のフレームの後半部分に相当する信号と、を加算する処理を行って符号化前の前記第*i*-1フレームの後半部分に対応する信号を生成することを特徴とする。

[0009] なお、これらの全般的または具体的な態様は、システム、方法、集積回路、コンピュータプログラムまたはコンピュータ読み取り可能なCD-ROMなどの記録媒体で実現されてもよく、システム、方法、集積回路、コンピュータプログラム及び記録媒体の任意な組み合わせで実現されてもよい。

発明の効果

[0010] 本発明によれば、AAC-ELDモードのように複数の先行フレームによるオーバーラップ処理が必要な符号化方式の音響コーデックを用いたハイブリッドコーデック（音信号ハイブリッドデコーダ及び音信号ハイブリッドエンコーダ）において、音声コーデックと音響コーデックとの切替部分において発生するエイリアシングを低減することができる。

図面の簡単な説明

[0011] [図1]図1は、AAC-ELDのエンコーダにおける分析窓を示す図である。
[図2]図2は、AAC-ELDのデコーダにおける復号処理を示す図である。
[図3]図3は、AAC-ELDのデコーダにおける合成窓を示す図である。
[図4]図4は、AAC-ELDの符号化・復号処理の遅延量を示す図である。
[図5]図5は、遷移フレームを説明するための図である。
[図6]図6は、実施の形態1に係る音信号ハイブリッドエンコーダの構成を示すブロック図である。

[図7]図7は、符号化モードがFD符号化モードからACELP符号化モードに切り替えられる場合の符号化されたフレームを示す図である。

[図8A]図8Aは、成分Xの生成方法の一例を示す図である。

[図8B]図8Bは、成分Xの生成方法のフローチャートである。

[図9]図9は、TCXエンコーダを備える音信号ハイブリッドエンコーダの構成を示すブロック図である。

[図10]図10は、実施の形態1に係る音信号ハイブリッドデコーダの構成を示すブロック図である。

[図11]図11は、FD符号化モードで符号化された信号がACELP符号化モードで符号化された信号に切り替わるときの、ブロック切替部の切替制御を示す模式図である。

[図12A]図12Aは、フレーム*i*-1の信号を再構成する方法を示す図である。

[図12B]図12Bは、フレーム*i*-1の信号を再構成する方法のフローチャートである。

[図13]図13は、実施の形態1に係る符号化・復号処理の遅延量を示す図である。

[図14]図14は、TCXデコーダを備える音信号ハイブリッドデコーダの構成を示すブロック図である。

[図15]図15は、合成エラー補償装置を用いてフレーム*i*-1の信号を再構成する方法を示す図である。

[図16]図16は、合成エラー情報の復号処理を示す図である。

[図17]図17は、符号化モードがACELP符号化モードからFD符号化モードに切り替えられる場合の符号化されたフレームを示す図である。

[図18]図18は、ACELP符号化モードで符号化された信号がFD符号化モードで符号化された信号に切り替わるときの、ブロック切替部の切替制御を示す模式図である。

[図19]図19は、実施の形態2に係るフレーム*i*-1の信号を再構成する方

法のフローチャートである。

[図20A]図20Aは、実施の形態2に係るフレーム $i-1$ の信号を再構成する方法の一例を示す図である。

[図20B]図20Bは、実施の形態2に係るフレーム $i-1$ の信号を再構成する方法の一例を示すもう一つの図である。

[図21]図21は、実施の形態2に係るフレーム i の信号を再構成する方法の一例を示す図である。

[図22]図22は、実施の形態2に係るフレーム $i+1$ の信号を再構成する方法の一例を示す図である。

[図23]図23は、実施の形態2に係る符号化・復号処理の遅延量を示す図である。

[図24]図24は、SEC装置を用いてフレーム $i-1$ の信号を再構成する方法を示す図である。

[図25]図25は、SEC装置を用いてフレーム i の信号を再構成する方法を示す図である。

[図26]図26は、SEC装置を用いてフレーム $i-1$ の信号を再構成する方法を示す図である。

[図27]図27は、符号化モードがFD符号化モードがTCX符号化モードに切り替えられる場合の符号化されたフレームを示す図である。

[図28]図28は、FD符号化モードで符号化された信号がTCX符号モードで符号化された信号に切り替わるときの、ブロック切替部の切替制御を示す模式図である。

[図29]図29は、実施の形態3に係る符号化・復号処理の遅延量を示す図である。

[図30]図30は、符号化モードがTCX符号化モードからFD符号化モードに切り替えられる場合の符号化されたフレームを示す図である。

[図31]図31は、符号化モードがTCX符号化モードからFD符号化モードに切り替えられる場合の符号化されたフレームを示す図である。

[図32]図32は、実施の形態4に係るフレーム $i-1$ の信号を再構成する方法の一例を示す図である。

[図33]図33は、実施の形態4に係る符号化・復号処理の遅延量を示す図である。

発明を実施するための形態

[0012] (発明の基礎となった知見)

音声コーデックは、特に、音声信号の特徴に応じて音声信号を符号化するためのコーデックである（非特許文献1参照）。音声コーデックは、音声信号を低ビットレートで符号化した場合、低遅延で良好な音質が実現される。しかしながら、音声コーデックは、音響信号の符号化には適していない。したがって、音声コーデックによって音響信号を符号化した場合、例えば、AACなどの音響コーデックで符号化した場合に比べて音質は低下する。

[0013] 現在では、ACELP符号化モード (Algebraic Code Excited Linear Prediction)、またはTCX符号化モード (Transform Coded Excitation) などの一般的な音声コーデックは、線形予測領域符号化に基づいている（特許文献1参照）。ACELP符号化モードにおいては、線形予測分析の後、代数コードブックが励振信号の符号化に適用される。TCX符号化モードにおいては、線形予測分析の後、変換符号化が励振信号に対して用いられる。

[0014] 一方、音響コーデックは、音響信号の符号化に適したコーデックである。しかしながら、音響コーデックが音声信号に用いられた場合、音声コーデックのように安定した音質を実現するためには、通常、高いビットレートを要する。

[0015] ハイブリッドコーデックは、音響コーデックと音声コーデックの利点を組み合わせるものである。ハイブリッドコーデックでは、符号化モードは、2つの系統に分かれている。1つは、上記音響コーデックに相当する、AACなどの周波数領域 (FD: Frequency Domain) 符号化モードである。もう1つは、上記音声コーデックに相当する、線形予測領域 (L

PD: Linear Prediction Domain) 符号化モードである。

[0016] FD符号化モードとして、一般的には、AAC-LD符号化モードやAAC符号化モードなどの直行変換符号化が用いられる。また、LPD符号化モードとしては、一般的にLPC (Linear Prediction Coefficient) 残差の周波数領域表示であるTCX符号化モードと、LPC残差の時間領域表示であるACELP符号化モードとが用いられる。

[0017] ハイブリッドコーデックでは、符号化対象の信号が音声信号であるか音響信号であるかに応じて、符号化モードを切替える (特許文献1参照)。なお、ACELP符号化モードと、TCX符号化モードとのどちらを選択するかは、例えば、閉ループ分析合成技術に基づき選択される。

[0018] ここで、VoIP (Voice over Internet Protocol) やビデオ会議などのリアルタイム通信を行う場合、低遅延のハイブリッドコーデックがより望ましい。ここで低遅延を実現するために、上記FD符号化モードとして、AAC及びAAC-LDを拡張したAAC-ELD符号化方式 (以下、単にAAC-ELDとも記載する) が用いられる。AAC-ELD符号化方式は、十分な低遅延を実現するために、下記の特徴を有する。

[0019] 1. AAC-ELDの1つのフレームにおけるサンプル数 (フレームサイズ N 、以下本明細書中においても同様である) が、512時間領域サンプル及び480時間領域サンプルと少ない。

[0020] 2. 先読み処理及びブロック切り替え処理が無効とされている。

[0021] 3. 低遅延フィルタバンクを採用するために、分析及び合成フィルタバンクが変更される。具体的には、長さ $4N$ のロングウィンドウが、過去とのオーバーラップを多く、未来とのオーバーラップを少なくして用いられる (値 $N/4$ は、実際はゼロである)。

[0022] 4. ビットリザーバが最小化される、またはビットリザーバが一切使用さ

れない。

[0023] 5. 時間領域ノイズ整形及び長期予測機能が低遅延のフレームサイズに応じて適応される。

[0024] ここで、AAC-ELD低遅延フィルタバンクの変換及び逆変換について説明する。なお、以下で説明される背景知識は、後続の記載においてもそのまま用いられる。

[0025] 既に述べたように、AAC-ELDにおいては、低遅延分析及び合成フィルタバンクが用いられる。低遅延フィルタバンクは、以下のように定義される。

[0026] [数1]

$$X_k = -2 \sum_{n=-2N}^{2N-1} x_n \cos\left[\frac{\pi}{N}\left(n + \frac{1}{2} - \frac{N}{2}\right)\left(k + \frac{1}{2}\right)\right], 0 \leq k < N \cdots \text{式 (1)}$$

[0027] ここで、 x_n は、窓処理された入力信号（符号化対象）である。一方、AAC-ELDの低遅延逆フィルタバンクは、以下のように定義される。

[0028] [数2]

$$y_n = -\frac{1}{N} \sum_{k=0}^{N-1} X_k \cos\left[\frac{\pi}{N}\left(n + \frac{1}{2} - \frac{N}{2}\right)\left(k + \frac{1}{2}\right)\right], 0 \leq n < 4N$$

[0029] ここで、 X_k は、復号された変換係数である。

[0030] まず、AAC-ELDのエンコーダにおける変換処理（AAC-ELDの符号化処理）を説明する。

[0031] AAC-ELDにおいては、1つのフレームに対応して4フレームが符号化される。具体的には、フレーム*i*-1が符号化される場合、当該フレーム*i*-1に先行する3つのフレーム*i*-4、*i*-3、*i*-2が連結された長さ4*N*の拡張フレームが形成され、この拡張フレームが符号化される。1つのフレームサイズが*N*である場合、符号化されるフレームサイズは、4*N*である。

[0032] 図1は、AAC-ELDのエンコーダにおける分析窓（エンコーダ窓）を

示し、これは w_{enc} と示される。なお、上述のように分析窓の長さは、 $4N$ である。

[0033] なお、便宜上、1つのフレームは、2つのサブフレームに分割される。例えば、フレーム $i-1$ は分割され、 $[a_{i-1}, b_{i-1}]$ のようにベクトルの形式で表される。 a_{i-1} 、 b_{i-1} の長さは、それぞれ $N/2$ サンプルである。これに対応して、長さが $4N$ のエンコーダ窓は8分割され、図1に示されるように、これらは $[w_1, w_2, w_3, w_4, w_5, w_6, w_7, w_8]$ である。一方、拡張フレームは、 $[a_{i-4}, b_{i-4}, a_{i-3}, b_{i-3}, a_{i-2}, b_{i-2}, a_{i-1}, b_{i-1}]$ と示される。エンコーダ窓が拡張フレームに適用されて、窓処理された信号である $x_n = [a_{i-4}w_1, b_{i-4}w_2, a_{i-3}w_3, b_{i-3}w_4, a_{i-2}w_5, b_{i-2}w_6, a_{i-1}w_7, b_{i-1}w_8]$ が得られる。

[0034] ここで、式(1)で定義された低遅延フィルタバンクが、窓処理された信号 x_n を変換するために用いられる。上記低遅延フィルタバンクにしたがって、フレームサイズ $4N$ の窓処理された信号 x_n から、フレームサイズ N の変換スペクトル係数が生成される。

[0035] なお、低遅延フィルタバンクの基本的なアルゴリズムは、MDCT (Modified Discrete Cosine Transform) と同様である。ここで、MDCTは、DCT-IVに基づくフーリエ変換の類似形であるので、低遅延フィルタバンクとDCT-IVとの間には、基本的に等価な関係性が存在することとなる(非特許文献2参照)。DCT-IVは、以下のように定義される。

[0036] [数3]

$$X_k = DCT-IV(x_n) = \sum_{n=0}^{N-1} x_n \cos\left[\frac{\pi}{N}\left(n + \frac{1}{2}\right)\left(k + \frac{1}{2}\right)\right]$$

[0037] DCT-IVは、以下のような偶数/奇数交互の境界条件を有する。

[0038]

[数4]

$$\cos\left[\frac{\pi}{N}\left(-n-1+\frac{1}{2}\right)\left(k+\frac{1}{2}\right)\right] = \cos\left[\frac{\pi}{N}\left(n+\frac{1}{2}\right)\left(k+\frac{1}{2}\right)\right]$$

$$\cos\left[\frac{\pi}{N}\left(2N-n-1+\frac{1}{2}\right)\left(k+\frac{1}{2}\right)\right] = -\cos\left[\frac{\pi}{N}\left(n+\frac{1}{2}\right)\left(k+\frac{1}{2}\right)\right]$$

[0039] これらの境界条件を用いて、低遅延フィルタバンクによって変換されたフレーム $i-1$ の信号は、DCT-IVでは以下のように表される。

[0040] [数5]

$$\left[DCT-IV\left(-\left(a_{i-4}w_1\right)_R - b_{i-4}w_2 + \left(a_{i-2}w_5\right)_R + b_{i-2}w_6\right), \right. \\ \left. DCT-IV\left(-a_{i-3}w_3 + \left(b_{i-3}w_4\right)_R + a_{i-1}w_7 - \left(b_{i-1}w_8\right)_R\right) \right]$$

[0041] 式中の、 $\left(a_{i-4}w_1\right)_R$ 、 $\left(a_{i-2}w_5\right)_R$ 、 $\left(b_{i-3}w_4\right)_R$ 、 $\left(b_{i-1}w_8\right)_R$ は、それぞれ、ベクトル $a_{i-4}w_1$ 、 $a_{i-2}w_5$ 、 $b_{i-3}w_4$ 、 $b_{i-1}w_8$ の逆順である。

[0042] 次に、AAC-ELDのデコーダにおける逆変換処理（AAC-ELDの復号処理）について説明する。

[0043] 図2は、AAC-ELDのデコーダにおける復号処理を示す図である。復号後の出力信号の長さ（フレームサイズ）は、 $4N$ である。同様に、逆MDCTとDCT-IVとの関係性が等価であることを考慮すると（非特許文献2参照）、フレーム $i-1$ に対する逆変換信号は以下の通りである。

[0044]

[数6]

$$\begin{aligned}
y_{i-1} = & [-a_{i-4}w_1 - (b_{i-4}w_2)_R + a_{i-2}w_5 + (b_{i-2}w_6)_R, \\
& -(a_{i-4}w_1)_R - b_{i-4}w_2 + (a_{i-2}w_5)_R + b_{i-2}w_6, \\
& -a_{i-3}w_3 + (b_{i-3}w_4)_R + a_{i-1}w_7 - (b_{i-1}w_8)_R, \\
& (a_{i-3}w_3)_R - b_{i-3}w_4 - (a_{i-1}w_7)_R + b_{i-1}w_8, \\
& a_{i-4}w_1 + (b_{i-4}w_2)_R - a_{i-2}w_5 - (b_{i-2}w_6)_R, \\
& (a_{i-4}w_1)_R + b_{i-4}w_2 - (a_{i-2}w_5)_R - b_{i-2}w_6, \\
& a_{i-3}w_3 - (b_{i-3}w_4)_R - a_{i-1}w_7 + (b_{i-1}w_8)_R, \\
& -(a_{i-3}w_3)_R + b_{i-3}w_4 + (a_{i-1}w_7)_R - b_{i-1}w_8]
\end{aligned}$$

[0045] y_{i-1} にAAC-ELDのデコーダにおける合成窓が適用されることで、

[0046] [数7]

$$\bar{y}_{i-1}$$

が得られる。図3は、AAC-ELDのデコーダにおける合成窓を示し、これは w_{dec} と示される。合成窓は、AAC-ELDのエンコーダにおける分析窓をそのまま逆順にしたものである。また、AAC-ELDのエンコーダにおける分析窓と同様に、便宜上、図3に示されるように合成窓は8分割される。合成窓は、以下のようにベクトルの形式で表される。

[0047] [数8]

$$[w_{R,8}, w_{R,7}, w_{R,6}, w_{R,5}, w_{R,4}, w_{R,3}, w_{R,2}, w_{R,1}]$$

[0048] したがって、窓処理された逆変換信号である

[0049] [数9]

$$\bar{y}_{i-1}$$

は、以下の通りである。

[0050] [数10]

$$\begin{aligned}
\bar{y}_{i-1} = & [(-a_{i-4}w_1 - (b_{i-4}w_2)_R + a_{i-2}w_5 + (b_{i-2}w_6)_R)w_{R,8}, \\
& (-(a_{i-4}w_1)_R - b_{i-4}w_2 + (a_{i-2}w_5)_R + b_{i-2}w_6)w_{R,7}, \\
& (-a_{i-3}w_3 + (b_{i-3}w_4)_R + a_{i-1}w_7 - (b_{i-1}w_8)_R)w_{R,6}, \\
& ((a_{i-3}w_3)_R - b_{i-3}w_4 - (a_{i-1}w_7)_R + b_{i-1}w_8)w_{R,5}, \\
& (a_{i-4}w_1 + (b_{i-4}w_2)_R - a_{i-2}w_5 - (b_{i-2}w_6)_R)w_{R,4}, \\
& ((a_{i-4}w_1)_R + b_{i-4}w_2 - (a_{i-2}w_5)_R - b_{i-2}w_6)w_{R,3}, \\
& (a_{i-3}w_3 - (b_{i-3}w_4)_R - a_{i-1}w_7 + (b_{i-1}w_8)_R)w_{R,2}, \\
& (-(a_{i-3}w_3)_R + b_{i-3}w_4 + (a_{i-1}w_7)_R - b_{i-1}w_8)w_{R,1}] \\
= & [c_{-4}, d_{-4}, c_{-3}, d_{-3}, c_{-2}, d_{-2}, c_{-1}, d_{-1}]_{i-1}
\end{aligned}$$

[0051] AAC-ELDの復号処理では、フレーム*i*-1の信号 $[a_{i-1}, b_{i-1}]$ を再構成するために復号対象フレーム*i*を復号処理する。すなわち、フレーム*i*とこれに先行する3つのフレームの窓処理された逆変換信号をそれぞれ用いて重複加算処理を行う。よって、図2に示す重複加算処理は、以下の式で表される。

[0052] [数11]

$$\begin{aligned}
out_{i,n} = & \bar{y}_{i,n} + \bar{y}_{i-1,n+N} + \bar{y}_{i-2,n+2N} + \bar{y}_{i-3,n+3N}, 0 \leq n < N \\
= & [c_{-4}, d_{-4}]_i + [c_{-3}, d_{-3}]_{i-1} + [c_{-2}, d_{-2}]_{i-2} + [c_{-1}, d_{-1}]_{i-3}
\end{aligned}$$

[0053] 再構成された信号の長さはNである。

[0054] エイリアシングの低減は、上記重複加算式に基づき導き出される。

[0055] [数12]

$$0 \leq n < \frac{N}{2}$$

については、以下の通りである。

[0056] [数13]

$$\begin{aligned}
 out_{i,n} &= [c_{-4}]_i + [c_{-3}]_{i-1} + [c_{-2}]_{i-2} + [c_{-1}]_{i-3} \\
 &= (-a_{i-3}w_1 - (b_{i-3}w_2)_R + a_{i-1}w_5 + (b_{i-1}w_6)_R)w_{R,8} + (-a_{i-3}w_3 + (b_{i-3}w_4)_R + a_{i-1}w_7 - (b_{i-1}w_8)_R)w_{R,6} \\
 &\quad + (a_{i-5}w_1 + (b_{i-5}w_2)_R - a_{i-3}w_5 - (b_{i-3}w_6)_R)w_{R,4} + (a_{i-5}w_3 - (b_{i-5}w_4)_R - a_{i-3}w_7 + (b_{i-3}w_8)_R)w_{R,2} \\
 &= a_{i-5}(w_3w_{R,2} + w_1w_{R,4}) + a_{i-3}(-w_7w_{R,2} - w_5w_{R,4} - w_1w_{R,8} - w_3w_{R,6}) + a_{i-1}(w_7w_{R,6} + w_5w_{R,8})
 \end{aligned}$$

また、

[0057] [数14]

$$\frac{N}{2} \leq n < N$$

については、以下の通りである。

[0058] [数15]

$$\begin{aligned}
 out_{i,n} &= [d_{-4}]_i + [d_{-3}]_{i-1} + [d_{-2}]_{i-2} + [d_{-1}]_{i-3} \\
 &= (-a_{i-3}w_1)_R - b_{i-3}w_2 + (a_{i-1}w_5)_R + b_{i-1}w_6)w_{R,7} + ((a_{i-3}w_3)_R - b_{i-3}w_4 - (a_{i-1}w_7)_R + b_{i-1}w_8)w_{R,5} \\
 &\quad + ((a_{i-5}w_1)_R + b_{i-5}w_2 - (a_{i-3}w_5)_R - b_{i-3}w_6)w_{R,3} + (-a_{i-5}w_3)_R + b_{i-5}w_4 + (a_{i-3}w_7)_R - b_{i-3}w_8)w_{R,1} \\
 &= b_{i-5}(w_4w_{R,1} + w_2w_{R,3}) + b_{i-3}(-w_8w_{R,1} - w_6w_{R,3} - w_4w_{R,5} - w_2w_{R,7}) + b_{i-1}(w_8w_{R,5} + w_6w_{R,7})
 \end{aligned}$$

[0059] さらに、以下の窓の特性から、フレーム $i-1$ の信号 $[a_{i-1}, b_{i-1}]$ が、重複加算処理により再構成される。

[0060] [数16]

$$\begin{aligned}
 w_3w_{R,2} + w_1w_{R,4} &\approx 0 \\
 -w_7w_{R,2} - w_5w_{R,4} - w_1w_{R,8} - w_3w_{R,6} &\approx 0 \\
 w_7w_{R,6} + w_5w_{R,8} &\approx 1 \\
 w_4w_{R,1} + w_2w_{R,3} &\approx 0 \\
 -w_8w_{R,1} - w_6w_{R,3} - w_4w_{R,5} - w_2w_{R,7} &\approx 0 \\
 w_8w_{R,5} + w_6w_{R,7} &\approx 1
 \end{aligned}$$

[0061] ここで、AAC-ELDの符号化・復号処理の遅延量について説明する。

[0062] 図4は、AAC-ELDの符号化・復号処理の遅延量を示す図である。図4において、フレーム*i* - 1に対する符号化処理は、時間*t*において開始されるものとする。

[0063] 図1に示されるように、AAC-ELDのエコーダにおける分析窓の $w_{s,8}$ の後半 $N/4$ 個のサンプルに相当する部分は、ゼロである。したがって、図4に示すように、時間 $t + 3 * N/4$ サンプルの時には、 x_{i-1} は、MDCT変換できる状態でありIMDCT変換された信号 y_{i-1} が得られる。

[0064] 同様に、図4に示すように、時間 $t + 7 * N/4$ サンプルの時には、IMDCT変換された信号 y_i が得られる。

[0065] 続いて、窓処理及び重複加算処理が y_{i-1} 、 y_i に適用されて、 $out_{i,n}$ が得られる。ここでも、図3に示されるように、AAC-ELDのデコーダにおける合成窓 $w_{R,8}$ の前半 $N/4$ 個のサンプルに相当する部分は、ゼロであるので、

[0066] [数17]

$$\bar{y}_i$$

が利用可能になる $N/4$ サンプル前に音の出力を開始することができる。つまり、音の出力は $(t + 7 * N/4) - N/4 = t + 3 * N/2$ サンプルにおいて開始される。すなわち、AAC-ELD符号化・復号処理の遅延量は、 $3 * N/2$ サンプルであり、低遅延である。

[0067] 以上説明したように、AAC-ELDでは、MDCTが4つの連続したフレームに対して行われ、4つのフレームは、図2に示されるように重複加算処理される。このようなAAC-ELDをハイブリッドコーデックに用いることで、音質を向上させ、遅延量をさらに低減させることができる。なお、MDCT変換は、TCX符号化モードにおいても用いられるが、TCX符号化モードでは、MDCT変換は、1つのフレーム内に1つ以上の複数のブロックが存在し、その連続するブロックに対して行われ、1つのブロックの後半が、次のブロックの前半と一致するように後続のブロックが重複される。

- [0068] AAC-ELDでは、上述のように重複加算処理により先行フレーム及び後続フレームを用いて復号を行うために、符号化モードをLPD符号化モードからAAC-ELD、またはAAC-ELDからLPD符号化モードに切替えた最初のフレームである遷移フレームを復号する場合にエイリアシングが発生する。
- [0069] 図5は、遷移フレームを説明するための図である。図5中のフレーム*i*が遷移フレームである。例えば、モード1がAAC-ELDであり、モード2がLPD符号化モードであるとき、フレーム*i*を復号する場合にエイリアシングが発生する。同様に、モード1がLPD符号化モードであり、モード2がAAC-ELDであるとき、フレーム*i*を復号する場合にエイリアシングが発生する。
- [0070] 遷移フレームにおいて発生したエイリアシングによって、通常、可聴のアーチファクトが発生する。しかしながら、特許文献1に記載のような方法は、AAC-ELDのような複数の先行フレームによるオーバーラップ処理が必要な符号化方式に対応していないため、発生したエイリアシングを低減することができない。
- [0071] このような課題を解決するために、本発明の一態様に係る音信号ハイブリッドデコーダは、低遅延フィルタバンクを用いた音響符号化処理で符号化された音響フレームと、線形予測係数を用いた音声符号化処理で符号化された音声フレームとが含まれるビットストリームを復号する音信号ハイブリッドデコーダであって、前記音響フレームを低遅延逆フィルタバンク処理を用いて復号する低遅延変換デコーダと、前記音声フレームを復号する音声信号デコーダと、前記ビットストリームのうちの復号対象フレームが前記音響フレームである場合、当該復号対象フレームを前記低遅延変換デコーダによって復号し、前記復号対象フレームが前記音声フレームである場合、当該復号対象フレームを前記音声信号デコーダによって復号する制御を行うブロック切替部とを備え、前記復号対象フレームが、前記音響フレームから前記音声フレームに切り替わった最初の前記音声フレームである第*i*フレームであると

き、前記第 i フレームには、前記第 i フレームよりも 1 フレーム先行するフレームである第 $i - 1$ フレームの符号化前の信号を用いて生成された第 1 信号が符号化された状態で含まれ、前記ブロック切替部は、(1) 前記第 i フレームよりも 2 フレーム先行するフレームである第 $i - 2$ フレームを前記低遅延変換デコーダによって復号することで得られる、前記第 i フレームよりも 3 フレーム先行するフレームである第 $i - 3$ フレームの再構成された信号を窓処理した信号である第 2 信号のフレームの前半部分に相当する信号に、前記第 2 信号のフレームの後半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第 i フレームを前記音声信号デコーダによって復号することで得られる、前記第 1 信号に窓処理を行った信号と、前記第 $i - 1$ フレームを前記低遅延逆フィルタバンク処理及び窓処理した信号の前記第 $i - 3$ フレームに対応する部分である第 3 信号のフレームの前半部分の信号と、を加算する処理を行って符号化前の前記第 $i - 1$ フレームの前半部分に対応する信号を生成し、前記第 2 信号のフレームの後半部分に相当する信号に、前記第 2 信号のフレームの前半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第 1 信号に畳み込み処理及び窓処理を行った信号と、前記第 3 信号のフレームの後半部分に相当する信号と、を加算する処理を行って符号化前の前記第 $i - 1$ フレームの後半部分に対応する信号を生成する、または (2) 前記第 2 信号のフレームの前半部分に相当する信号に、前記第 2 信号のフレームの後半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第 1 信号に畳み込み処理及び窓処理を行った信号と、前記第 3 信号のフレームの前半部分に相当する信号と、を加算する処理を行って符号化前の前記第 $i - 1$ フレームの前半部分に対応する信号を生成し、前記第 2 信号のフレームの後半部分に相当する信号に、前記第 2 信号のフレームの前半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第 1 信号に窓処理を行った信号と、前記第 3 信号のフレームの後半部分に相当する信号と、を加算する処理を行って符号化前の前記第 $i - 1$ フレームの後半部

分に対応する信号を生成することを特徴とする。

[0072] つまり、ブロック切替部は、図12Aに示される処理を行う。これにより、符号化モードがFD符号化モードからLPD符号化モードに切り替えられた最初のフレームを復号する場合に発生するエイリアシングを低減することができる。したがって、FD復号技術とLPD復号技術とのシームレスな切替が実現される。

[0073] また、本発明の一態様において、低遅延フィルタバンクを用いた音響符号化処理で符号化された音響フレームと、線形予測係数を用いた音声符号化処理で符号化された音声フレームとが含まれるビットストリームを復号する音信号ハイブリッドデコーダであって、前記音響フレームを低遅延逆フィルタバンク処理によって復号する低遅延変換デコーダと、前記音声フレームを復号する音声信号デコーダと、前記ビットストリームのうちの復号対象フレームが前記音響フレームである場合、当該復号対象フレームを前記低遅延変換デコーダによって復号し、前記復号対象フレームが前記音声フレームである場合、当該復号対象フレームを前記音声信号デコーダによって復号する制御を行うブロック切替部とを備え、前記ブロック切替部は、前記復号対象フレームが、前記音声フレームから前記音響フレームに切り替わった最初の音響フレームである第 i フレームであるとき、前記第 i フレームよりも1フレーム先行するフレームである第 $i-1$ フレームを前記音声信号デコーダによって復号することで得られる信号を窓処理した第4信号に、当該第4信号を畳み込み処理した信号を加算し、窓処理を行った第5信号と、前記第 i フレームよりも3フレーム先行するフレームである第 $i-3$ フレームを前記音声信号デコーダによって復号することで得られる信号を窓処理した第6信号に、当該第6信号を畳み込み処理した信号を加算し、窓処理を行った第7信号と、前記第 i フレームを前記低遅延逆フィルタバンク処理及び窓処理した信号の前記第 $i-3$ フレームに対応する部分である第8信号と、を加算する処理を行って符号化前の前記第 $i-1$ フレームに対応する信号である再構成信号を生成してもよい。

[0074] つまり、ブロック切替部は、図20A及び図20Bに示される処理を行う。これにより、符号化モードがLPD符号化モードからFD符号化モードに切り替えられた最初のフレームを復号する場合に発生するエイリアシングを低減することができる。したがって、FD復号技術とLPD復号技術とのシームレスな切替が実現される。

[0075] また、本発明の一態様において、前記ブロック切替部は、前記復号対象フレームが、前記第*i*フレームの1フレーム後のフレームである第*i*+1フレームであるとき、前記第*i*+1フレームを前記低遅延逆フィルタバンク処理及び窓処理した信号のうちの、前記第*i*フレームよりも2フレーム先行するフレームである第*i*-2フレームに対応する部分である第9信号と、前記第*i*フレームを前記低遅延逆フィルタバンク処理及び窓処理した信号の前記第*i*-2フレームに対応する部分である第10信号と、前記第*i*-2フレームを前記音声信号デコーダによって復号することで得られる第11信号に第1の窓処理を行なった信号のフレームの前半部分に相当する信号に、前記第11信号に前記第1の窓処理を行った信号のフレームの後半部分に相当する信号に畳み込み処理した信号を加算することで得られる第12信号に、当該第12信号を畳み込み処理した信号を連結し、窓処理を行った第13信号と、前記第11信号に前記第1の窓処理とは異なる第2の窓処理を行った信号のフレームの前半部分に相当する信号に、前記第11信号に前記第2の窓処理を行った信号のフレームの後半部分に相当する信号に畳み込み処理した信号を加算することで得られる第14信号に、当該第14信号を畳み込み処理して符号を反転させた信号を連結し、窓処理を行った第15信号と、を加算する処理を行って、符号化前の前記第*i*フレームに対応する信号を生成してもよい。

[0076] つまり、ブロック切替部は、図21に示される処理を行う。これにより、符号化モードがLPD符号化モードからFD符号化モードに切り替えられた最初のフレームの1フレーム後のフレームを復号する場合に発生するエイリアシングを低減することができる。

[0077] また、本発明の一態様において、前記ブロック切替部は、前記復号対象フレームが、前記第 i フレームの 2 フレーム後のフレームである第 $i + 2$ フレームであるとき、前記 $i + 2$ フレームを前記低遅延逆フィルタバンク処理及び窓処理した信号の前記第 $i - 1$ フレームに対応する部分である第 16 信号と、前記第 $i + 1$ フレームを前記低遅延逆フィルタバンク処理及び窓処理した信号の前記第 $i - 1$ フレームに対応する部分である第 17 信号と、前記第 i フレームを前記低遅延逆フィルタバンク処理及び窓処理した信号の前記第 $i - 1$ フレームに対応する部分である第 18 信号と、前記第 $i - 3$ フレームを前記音声信号デコーダによって復号することで得られる第 19 信号に窓処理を行なった信号のフレームの前半部分に相当する信号に、前記第 19 信号に窓処理を行った信号のフレームの後半部分に相当する信号に畳み込み処理した信号を加算することで得られる第 20 信号に、当該第 20 信号を畳み込み処理した信号を連結し、窓処理を行った第 21 信号と、前記再構成信号に窓処理を行った信号のフレームの前半部分に相当する信号に、前記再構成信号に窓処理を行った信号のフレームの後半部分に相当する信号に畳み込み処理した信号を加算することで得られる第 22 信号に、当該第 22 信号を畳み込み処理して符号を反転させた信号を連結し、窓処理を行った第 23 信号と、を加算する処理を行って、符号化前の前記第 $i + 1$ フレームに対応する信号を生成してもよい。

[0078] つまり、ブロック切替部は、図 22 に示される処理を行う。これにより、符号化モードが LPD 符号化モードから FD 符号化モードに切り替えられた最初のフレームの 2 フレーム後のフレームを復号する場合に発生するエイリアシングを低減することができる。

[0079] また、本発明の一態様において、低遅延フィルタバンクを用いた音響符号化処理で符号化された音響フレームと、線形予測係数を用いた音声符号化処理で符号化された音声フレームとが含まれるビットストリームを復号する音信号ハイブリッドデコーダであって、前記音響フレームを低遅延逆フィルタバンク処理を用いて復号する低遅延変換デコーダと、TCX (T r a n s f

ormCodedExcitation)方式によって符号化された前記音声フレームを復号するTCXデコーダと、前記ビットストリームのうちの復号対象フレームが前記音響フレームである場合、当該復号対象フレームを前記低遅延変換デコーダによって復号し、前記復号対象フレームが前記音声フレームである場合、当該復号対象フレームを前記音声信号デコーダによって復号する制御を行うブロック切替部とを備え、前記復号対象フレームが、前記音響フレームから前記音声フレームに切り替わった最初の前記音声フレームであって、過渡信号が符号化されたフレームである第 i フレームであるとき、前記第 i フレームには、前記第 i フレームよりも1フレーム先行するフレームである第 $i-1$ フレームの符号化前の信号を用いて生成された第1信号が符号化された状態で含まれ、前記ブロック切替部は、(1)前記第 i フレームよりも2フレーム先行するフレームである第 $i-2$ フレームを前記低遅延変換デコーダによって復号することで得られる、前記第 i フレームよりも3フレーム先行するフレームである第 $i-3$ フレームの再構成された信号を窓処理した信号である第2信号のフレームの前半部分に相当する信号に、前記第2信号のフレームの後半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第 i フレームを前記音声信号デコーダによって復号することで得られる、前記第1信号に窓処理を行った信号と、前記第 $i-1$ フレームを前記低遅延逆フィルタバンク処理及び窓処理した信号の前記第 $i-3$ フレームに対応する部分である第3信号のフレームの前半部分の信号と、を加算する処理を行って符号化前の前記第 $i-1$ フレームの前半部分に対応する信号を生成し、前記第2信号のフレームの後半部分に相当する信号に、前記第2信号のフレームの前半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第1信号に畳み込み処理及び窓処理を行った信号と、前記第3信号のフレームの後半部分に相当する信号と、を加算する処理を行って符号化前の前記第 $i-1$ フレームの後半部分に対応する信号を生成する、または(2)前記第2信号のフレームの前半部分に相当する信号に、前記第2信号のフレームの後半部分に相当す

る信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第1信号に畳み込み処理及び窓処理を行った信号と、前記第3信号のフレームの前半部分に相当する信号と、を加算する処理を行って符号化前の前記第*i* - 1フレームの前半部分に対応する信号を生成し、前記第2信号のフレームの後半部分に相当する信号に、前記第2信号のフレームの前半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第1信号に窓処理を行った信号と、前記第3信号のフレームの後半部分に相当する信号と、を加算する処理を行って符号化前の前記第*i* - 1フレームの後半部分に対応する信号を生成してもよい。

[0080] つまり、ブロック切替部は、FD符号化モードにおいて過渡信号（過渡フレーム）が発生した場合の符号化信号の復号において、図12Aに示される処理を行う。これにより、過渡フレームを復号した場合の音の音質を向上させることができる。

[0081] また、本発明の一態様において、前記低遅延変換デコーダは、前記音響フレーム及び当該音響フレームに時間的に連続して先行する3つのフレームのそれぞれについて低遅延逆フィルタバンク処理及び窓処理を行った信号のそれぞれを重複加算処理することによって、当該音響フレームを復号するAAC-ELD (Advanced Audio Coding-Enhanced Low Delay) デコーダであってもよい。

[0082] また、本発明の一態様において、前記音声信号デコーダは、ACELP (Algebraic Code Excited Linear Prediction) 係数を用いて符号化された前記音声フレームを復号するACELPデコーダであってもよい。

[0083] また、本発明の一態様において、前記音声信号デコーダは、TCX方式によって符号化された前記音声フレームを復号するTCXデコーダであってもよい。

[0084] また、本発明の一態様において、さらに、前記復号対象フレームとともに符号化された合成エラー情報を復号する合成エラー補償装置を備え、前記合

成エラー情報は、前記ビットストリームが符号化される前の信号と、前記ビットストリームを復号した信号との差分を表す情報であり、前記合成エラー補償装置は、前記ブロック切替部が生成した前記符号化前の前記第 $i - 1$ フレームの信号、前記ブロック切替部が生成した前記符号化前の前記第 i フレームの信号、または前記ブロック切替部が生成した前記符号化前の前記第 $i + 1$ フレームの信号を、復号した前記合成エラー情報を用いて修正してもよい。

[0085] これにより、符号化モードが切り替えられることによって、音信号ハイブリッドデコーダにおいて発生する合成エラーが低減され、音質を向上することができる。

[0086] また、本発明の一態様に係る音信号ハイブリッドエンコーダは、音信号の音響特性を分析し、前記音信号に含まれるフレームが音響信号であるか音声信号であるかを判断する信号分類部と、低遅延フィルタバンクを用いて前記フレームを符号化する低遅延変換エンコーダと、前記フレームの線形予測係数を算出することによって当該フレームを符号化する音声信号エンコーダと、前記信号分類部が前記音響信号であると判断した符号化対象フレームを前記低遅延変換エンコーダによって符号化し、前記信号分類部が前記音声信号であると判断した前記符号化対象フレームを前記音声信号エンコーダによって符号化する制御を行うブロック切替部とを備え、前記ブロック切替部は、
(1) 前記符号化対象フレームが、前記信号分類部が前記音声信号であると判断したフレームである第 $i - 1$ フレームの 1 フレーム後のフレームであって、前記信号分類部が前記音響信号であると判断したフレームである第 i フレームであるとき、前記第 $i - 1$ フレームの前半部分に相当する信号を窓処理した信号に前記第 $i - 1$ フレームの後半部分に相当する信号を窓処理して畳み込み処理した信号を加算した信号と、前記第 i フレームとを前記音声信号エンコーダによって符号化する、または (2) 前記第 $i - 1$ フレームの後半部分に相当する信号を窓処理した信号に前記第 $i - 1$ フレームの前半部分に相当する信号を窓処理して畳み込み処理した信号を加算した信号と、前記

第 i フレームとを前記音声信号エンコーダによって符号化することを特徴とする。

[0087] つまり、ブロック切替部は、図 7 及び図 8 A に示される処理を行う。これにより、符号化モードが FD 符号化モードから LPD 符号化モードに切り替えられた最初のフレームを復号する場合に発生するエイリアシングを低減することができる。したがって、FD 復号技術と LPD 復号技術とのシームレスな切替が実現される。

[0088] また、本発明の一態様において、音信号の音響特性を分析し、前記音信号に含まれるフレームが音響信号であるか音声信号であるかを判断する信号分類部と、低遅延フィルタバンクを用いて前記フレームを符号化する低遅延変換エンコーダと、前記フレームの線形予測係数の残差を MDCT (Modified Discrete Cosine Transform) 処理した TCX 方式によって前記フレームを符号化する TCX エンコーダと、前記信号分類部が前記音響信号であると判断した符号化対象フレームを前記低遅延変換エンコーダによって符号化し、前記信号分類部が前記音声信号であると判断した前記符号化対象フレームを前記音声信号エンコーダによって符号化する制御を行うブロック切替部とを備え、前記ブロック切替部は、前記符号化対象フレームである第 i フレームが、前記信号分類部が前記音響信号であり、なおかつエネルギーが急激に変化する過渡信号であると判断したフレームであるとき、(1) 前記第 i フレームの 1 フレーム前のフレームである第 $i-1$ フレームの前半部分に相当する信号を窓処理した信号に前記第 $i-1$ フレームの後半部分に相当する信号を窓処理して畳み込み処理した信号を加算した信号と、前記第 i フレームとを前記音声信号エンコーダによって符号化する、または (2) 前記第 $i-1$ フレームの後半部分に相当する信号を窓処理した信号に前記第 $i-1$ フレームの前半部分に相当する信号を窓処理して畳み込み処理した信号を加算した信号と、前記第 i フレームとを前記音声信号エンコーダによって符号化してもよい。

[0089] つまり、ブロック切替部は、FD 符号化モードにおいて過渡信号 (過渡フ

レーム)が発生した場合の符号化において、図7及び図8Aに示される処理を行う。これにより、過渡フレームを復号した場合の音の音質を向上させることができる。

[0090] また、本発明の一態様において、前記低遅延変換エンコーダは、前記フレームと、当該フレームに時間的に連続して先行する3つのフレームとを連結した拡張フレームについて窓処理及び低遅延フィルタバンク処理をすることによって、前記フレームを符号化するAAC-ELDエンコーダであってもよい。

[0091] また、本発明の一態様において、前記音声信号エンコーダは、ACELP係数を生成することによって前記フレームを符号化するACELPエンコーダであってもよい。

[0092] また、本発明の一態様において、前記音声信号エンコーダは、前記線形予測係数の残差をMDCT処理して前記フレームを符号化するTCXエンコーダであってもよい。

[0093] また、本発明の一態様において、さらに、符号化した前記音信号を復号するローカルデコーダと、前記音信号と、前記ローカルデコーダが復号した前記音信号との差分である合成エラー情報を符号化するローカルエンコーダとを備えてもよい。

[0094] なお、これらの全般的または具体的な態様は、システム、方法、集積回路、コンピュータプログラムまたはコンピュータ読み取り可能なCD-ROMなどの記録媒体で実現されてもよく、システム、方法、集積回路、コンピュータプログラム及び記録媒体の任意な組み合わせで実現されてもよい。

[0095] 以下、実施の形態について、図面を参照しながら具体的に説明する。

[0096] 以下の各実施の形態では、以下の5通りの符号化モードの遷移において、エイリアシングの影響を低減し、シームレスな符号化モードの切替を実現する音信号ハイブリッドエンコーダ及び音信号ハイブリッドデコーダについて説明する。

[0097] ・FD符号化モードからACELP符号化モードへの遷移（実施の形態

1)

- ・ A C E L P 符号化モードから F D 符号化モードへの遷移（実施の形態

2)

- ・ F D 符号化モードから T C X 符号化モードへの遷移（実施の形態 3）
- ・ T C X 符号化モードから F D 符号化モードへの遷移（実施の形態 4）
- ・ F D 符号化モードから過渡信号符号化モードへの遷移（実施の形態 5

)

[0098] なお、以下で説明する実施の形態は、いずれも包括的または具体的な例を示すものである。以下の実施の形態で示される数値、形状、材料、構成要素、構成要素の配置位置及び接続形態、ステップ、ステップの順序などは、一例であり、本発明を限定する主旨ではない。また、以下の実施の形態における構成要素のうち、最上位概念を示す独立請求項に記載されていない構成要素については、任意の構成要素として説明される。

[0099] （実施の形態 1）

実施の形態 1 では、符号化モードを F D 符号化モードから A C E L P 符号化モードに切り替える場合の音信号ハイブリッドエンコーダの符号化方法及び音信号ハイブリッドデコーダの復号方法について説明する。なお、以下の実施の形態の説明においては、特に断りのない限り、F D 符号化モードとは A A C - E L D を意味するものとする。

[0100] [1 - 1 . 符号化方法]

図 6 は、実施の形態 1 に係る音信号ハイブリッドエンコーダの構成を示すブロック図である。

[0101] 音信号ハイブリッドエンコーダ 5 0 0 は、高周波エンコーダ 5 0 1 と、ブロック切替部 5 0 2 と、信号分類部 5 0 3 と、A C E L P エンコーダ 5 0 4 と、F D エンコーダ 5 0 5 と、ビットマルチプレクサ 5 0 6 とを備える。

[0102] 入力信号は、高周波エンコーダ 5 0 1 と信号分類部 5 0 3 とにそれぞれ送信される。

[0103] 高周波エンコーダ 5 0 1 は、入力信号のうちの高周波帯域を抽出して符号

化した信号である高周波パラメータと、入力信号のうちの低周波帯域を抽出した信号である低周波信号とを生成する。高周波パラメータは、ビットマルチプレクサ506に送信される。低周波信号は、ブロック切替部502に送信される。

[0104] 信号分類部503は、上記低周波信号の音響特性を分析し、低周波信号をサンプル数Nごと（フレームごと）に当該フレームが音響信号であるか音声信号であるか判断する。具体的には、信号分類部503は、当該フレームの3kHz以上の帯域のスペクトル強度と、当該フレームの3kHz以下の帯域のスペクトル強度とを算出する。3kHz以下のスペクトル強度がそれ以外の帯域のスペクトル強度よりも大きい場合、信号分類部503は、当該フレームが音声信号主体の信号である、すなわち音声信号であると判断し、判断結果を表すモード指標をブロック切替部502及びビットマルチプレクサ506に送信する。同様に、3kHz以下のスペクトル強度がそれ以外の帯域のスペクトル強度よりも小さい場合、信号分類部503は、当該フレームが音響信号主体の信号である、すなわち音響信号であると判断し、モード指標をブロック切替部502及びビットマルチプレクサ506に送信する。

[0105] ブロック切替部502は、モード指標が音響信号であることを表すフレームをFDエンコーダ505によって符号化し、モード指標が音声信号であることを表すフレームをACELPエンコーダ504によって符号化する切替制御を行う。すなわち、ブロック切替部502は、高周波エンコーダから受信した低周波信号をモード指標に従ってFDエンコーダ505とACELPエンコーダ504とにフレームごとに送信する。

[0106] FDエンコーダ505は、ブロック切替部502の制御に基づき、フレームをAAC-ELD符号化モードで符号化し、当該符号化によって生成されたFD変換係数をビットマルチプレクサ506に送信する。

[0107] ACELPエンコーダ504は、ブロック切替部502の制御に基づき、フレームをACELP符号化モードによって符号化し、当該符号化によって生成されたACELP係数をビットマルチプレクサ506に送信する。

- [0108] ビットマルチプレクサ506は、符号化モード指標と、高帯域パラメータと、FD変換係数と、ACELP係数とを合成したビットストリームを生成する。
- [0109] なお、図示しないが、音信号ハイブリッドエンコーダ500は、フレーム(信号)を一時的に記憶する記憶部を備えてもよい。
- [0110] 次に、符号化モードがFD符号化モードからACELP符号化モードに切り替えられる場合のブロック切替部502の制御について説明する。
- [0111] 図7は、符号化モードがFD符号化モードがACELP符号化モードに切り替えられる場合の符号化されたフレームを示す図である。
- [0112] この場合、フレーム*i*を符号化するとき、先行するフレーム*i-1*の信号 $[a_{i-1}, b_{i-1}]$ から生成される成分Xを加えた信号が符号化される。具体的には、ブロック切替部502は、成分Xと、フレーム*i*の信号 $[a_i, b_i]$ とを合わせた拡張フレームを生成する。拡張フレームは、 $(N+N/2)$ の長さである。拡張フレームは、ブロック切替部502によりACELPエンコーダ504に送信され、ACELP符号化モードで符号化される。
- [0113] 成分Xは、具体的には、次のように生成される。
- [0114] 図8Aは、成分Xの生成方法の一例を示す図である。図8Bは、成分Xの生成方法のフローチャートである。
- [0115] まず、フレーム*i-1*の信号の前半部分である入力部分 a_{i-1} に窓 w_5 を適用することで、成分 $a_{i-1}w_5$ が得られる(図8BのS101)。同様に、フレーム*i-1*の信号の後半部分である入力部分 b_{i-1} に窓 w_6 を適用することで、 $b_{i-1}w_6$ が得られる(図8BのS102)。次に、 $b_{i-1}w_6$ に、さらに畳み込み処理(折り畳み処理)を適用する(図8BのS103)。
- [0116] なお、本明細書中では、「信号に畳み込み処理をする」とは、当該信号ベクトル毎に信号ベクトルを構成するサンプルを時間的に逆順に並べ変えることを意味する。
- [0117] これにより、 $b_{i-1}w_6$ の逆順 $(b_{i-1}w_6)_R$ が得られる。最後に、 $a_{i-1}w_5$ と $(b_{i-1}w_6)_R$ とが加算された成分Xが得られる(図8BのS104)

- 。
- [0118] 得られた成分 X は、デコーダにおいて、複数の先行フレームとともに復号に用いられる。これにより、フレーム $i-1$ の信号 $[a_{i-1}, b_{i-1}]$ が適切に再構成される。
- [0119] なお、以上の説明では、 $b_{i-1}w_6$ に、さらに畳み込み処理を適用したが、 $a_{i-1}w_5$ にさらに畳み込み処理を適用してもよい。すなわち、成分 X は、 $(a_{i-1}w_5)_R + b_{i-1}w_6$ であってもよい。
- [0120] なお、図9に示されるように、音信号ハイブリッドエンコーダ500は、さらに、TCXエンコーダ507を備えてもよい。
- [0121] TCXエンコーダ507は、ブロック切替部502の制御に基づき、フレームをTCX符号化モードによって符号化し、当該符号化によって生成されたTCX係数をビットマルチプレクサ506に送信する。
- [0122] [1-2. 復号方法]
- 以下、音信号ハイブリッドエンコーダ500によって図8Aに示されるように符号化された符号化信号を復号する音信号ハイブリッドデコーダについて説明する。
- [0123] 図10は、実施の形態1に係る音信号ハイブリッドデコーダの構成を示すブロック図である。
- [0124] 音信号ハイブリッドデコーダ900は、デマルチプレクサ901と、FDデコーダ902と、ACELPデコーダ903と、ブロック切替部904と、高周波デコーダ905とを備える。
- [0125] デマルチプレクサ901は、ビットストリームを逆多重化する。具体的には、デマルチプレクサ901は、ビットストリームをモード指標と、高帯域パラメータと、符号化信号とに分割する。モード指標は、ブロック切替部904に送信され、高周波パラメータは、高周波デコーダ905に送信され、符号化信号（FD変換係数、及びACELP係数）は、フレームごとに対応するFDデコーダ902とACELPデコーダ903とに送信される。
- [0126] FDデコーダ902は、図2を用いて説明したAAC-ELD復号処理に

- よってFD変換係数からFD逆変換信号を生成する。すなわち、FDデコーダ902は、FD符号化モードによって符号化されたフレームを復号する。
- [0127] ACELPデコーダ903は、ACELP復号処理によってACELP係数からACELP合成信号を生成する。すなわち、ACELPデコーダ903は、ACELP符号化モードによって符号化されたフレームを復号する。
- [0128] FD逆変換信号とACELP合成信号とは、ブロック切替部904に送信される。
- [0129] ブロック切替部904は、モード指標が音響信号であることを表すフレームをFDデコーダ902によって復号してFD逆変換信号を受信し、モード指標が音声信号であることを表すフレームをACELPデコーダ903によって復号してACELP合成信号を受信する。
- [0130] 高周波デコーダ905は、デマルチプレクサから送信される高周波パラメータとブロック切替部904から送信される低周波帯域の時間領域信号とを用いて入力信号を再構成する。
- [0131] なお、図示しないが、音信号ハイブリッドデコーダ900は、フレーム（信号）を一時的に記憶する記憶部を備えてもよい。
- [0132] 次に、FD符号化モードで符号化された信号がACELP符号化モードで符号化された信号に切り替わるときの、ブロック切替部904の切替制御（復号方法）について説明する。
- [0133] 図11は、FD符号化モードで符号化された信号がACELP符号モードで符号化された信号に切り替わるときの、ブロック切替部904の切替制御（復号方法）を示す模式図である。図11に示されるようにフレーム*i*-1はFD符号化モードによって符号化されたフレームであり、復号対象フレームであるフレーム*i*は、ACELP符号化モードで符号化されたフレームである。
- [0134] 上述のように、FD符号化モードで符号化された信号が連続する場合は、復号対象フレーム*i*を復号してフレーム*i*-1の信号を再構成することができる。つまり、図11に示される場合、フレーム*i*-2の信号までは通常の

FD復号処理によって再構成できる。しかしながら、フレーム $i-1$ の信号は、復号対象フレーム i が ACELP 符号化モードで符号化されているため、通常の方法で再構成すると、エイリアシング成分による不自然な音が発生してしまう。すなわち、フレーム $i-1$ の信号は、図 11 に示されるようにエイリアシング部分となる。

[0135] エイリアシング成分を低減するために、ブロック切替部 904 は、次の 3 つの信号を用いて復号処理を行う。

[0136] 第 1 に、復号対象フレーム i を ACELP 復号処理することで得られる ACELP 合成信号の成分 X の信号（第 1 信号）がエイリアシング成分を低減したフレーム $i-1$ の信号を再構成するために用いられる。この信号は、図 11 においてサブフレーム 1001 と示される信号であり、図 8A を用いて説明した成分 X である。

[0137] 復号対象フレーム i は、ACELP 符号化モードで符号化された長さ $3N/2$ のフレームである。すなわち、フレーム i を ACELP 復号処理することで得られる ACELP 合成信号は、 $y_{i,n}^{acelp}$ と示され、

[0138] [数18]

$$0 \leq n < \frac{3}{2}N$$

である。したがって、成分 X に相当する拡張部分は、以下のようになる。

[0139] [数19]

$$X_{i,n} = y_{i,n}^{acelp}, \quad 0 \leq n < \frac{N}{2}$$

[0140] 図 8A を用いて説明したように、成分 X は、具体的には、 $a_{i-1}w_5 + (b_{i-1}w_6)_R$ である。

[0141] 第 2 に、復号対象フレーム $i-1$ を AAC-ELD 低遅延フィルタバンクによって逆変換した後、窓処理した信号のうちのフレーム $i-3$ に対応する部分の信号（第 3 信号）がエイリアシング成分を低減したフレーム $i-1$ の

信号を再構成するために用いられる。この信号は、図11においてサブフレーム1002及びサブフレーム1003と示される。

[0142] より具体的には、この信号は、フレーム*i*−1をAAC−ELD低遅延フィルタバンクによって、通常フレームとして長さ4*N*で逆変換し、さらに窓処理することによって得られる。逆変換信号は、

[0143] [数20]

$$\bar{y}_{i-1} = [c_{-4}, d_{-4}, c_{-3}, d_{-3}, c_{-2}, d_{-2}, c_{-1}, d_{-1}]_{i-1}$$

と示される。このうち、フレーム*i*−3に対応する部分の信号（図11においてサブフレーム1002及びサブフレーム1003と示される2つのエイリアシング部分）は、上記逆変換信号から以下のように抽出される。すなわち、

[0144] [数21]

$$[c_{-3}]_{i-1} = -a_{i-3}w_3w_{R,6} + (b_{i-3}w_4)_R w_{R,6} + a_{i-1}w_7w_{R,6} - (b_{i-1}w_8)_R w_{R,6}$$

及び

[0145] [数22]

$$[d_{-3}]_{i-1} = (a_{i-3}w_3)_R w_{R,5} - b_{i-3}w_4w_{R,5} - (a_{i-1}w_7)_R w_{R,5} + b_{i-1}w_8w_{R,5}$$

がサブフレーム1002とサブフレーム1003にそれぞれ対応する信号である。

[0146] 第3に、復号対象フレーム*i*−2をFD復号処理することによって得られるフレーム*i*−3の信号 $[a_{i-3}, b_{i-3}]$ （第2信号）がエイリアシング成分を低減したフレーム*i*−1の信号を再構成するために用いられる。フレーム*i*−3の信号は、図11においてサブフレーム1004及びサブフレーム1005と示される。

[0147] 以上、説明したように、図11においてサブフレーム1001と示される

信号 $a_{i-1}w_5 + (b_{i-1}w_6)_R$ と、サブフレーム 1002 と示される信号 $[c_{-3}]_{i-1}$ と、サブフレーム 1003 と示される信号 $[d_{-3}]_{i-1}$ と、サブフレーム 1004、1005 と示される信号 $[a_{i-3}, b_{i-3}]$ とが、エイリアシング成分を低減したフレーム $i-1$ の信号を再構成するために用いられる。

[0148] 上記の信号を用いたエイリアシング成分を低減したフレーム $i-1$ の信号を再構成する方法について具体的に説明する。

[0149] 図 12A の (a) は、フレーム $i-1$ の信号の前半のサンプル部分である a_{i-1} を再構成する方法を示す図である。図 12B は、フレーム $i-1$ の信号の前半のサンプル部分である a_{i-1} を再構成する方法のフローチャートである。

[0150] まず、サブフレーム 1004 (第 2 信号のフレーム前半部分) である a_{i-3} に窓 w_3 を適用することで、 $a_{i-3}w_3$ が得られる (図 12B の S201)。次に、サブフレーム 1005 (第 2 信号のフレーム後半部分) である b_{i-3} に窓 w_4 を適用することで、 $b_{i-3}w_4$ が得られ、さらに、畳み込み処理を適用することで、 $b_{i-3}w_4$ の逆順である $(b_{i-3}w_4)_R$ が得られる (図 12B の S202)。

[0151] 次に、 $a_{i-3}w_3$ と及び $(b_{i-3}w_4)_R$ とが加算処理された信号に、窓処理が適用されることで、 $a_{i-3}w_3w_{R,6} - (b_{i-3}w_4)_Rw_{R,6}$ が得られる (図 12B の S203)。

[0152] サブフレーム 1001 (成分 X、第 1 信号) である $a_{i-1}w_5 + (b_{i-1}w_6)_R$ には、合成窓 $w_{R,8}$ が適用されて、 $a_{i-1}w_5w_{R,8} + (b_{i-1}w_6)_Rw_{R,8}$ が得られる (図 12B の S204)。

[0153] これに加えて、逆変換信号であるサブフレーム 1002 (第 3 信号のフレーム前半部分) は、

[0154] [数 23]

$$-a_{i-3}w_3w_{R,6} + (b_{i-3}w_4)_Rw_{R,6} + a_{i-1}w_7w_{R,6} - (b_{i-1}w_8)_Rw_{R,6}$$

となる。上記それぞれの信号は、加算され、 $a_{i-1} (w_5 w_{R,8} + w_7 w_{R,6})$ が得られる (図 12B の S205)。

[0155] 上述の窓の特性を考慮すると、

[0156] [数24]

$$w_5 w_{R,8} + w_7 w_{R,6} \approx 1$$

であることから、エイリアシング成分を低減したフレーム $i-1$ の信号の前半部分であるサブフレーム 1101 が得られる。

[0157] 同様に、図 12A の (b) は、フレーム $i-1$ の信号の後半のサンプル部分である b_{i-1} を再構成する方法を示す図である。図 12A の (a) とは、サブフレーム 1001 に畳み込み処理を行う点で異なるが、その他の処理は、同様である。これにより、エイリアシング成分を低減したフレーム $i-1$ の信号の後半部分であるサブフレーム 1102 が得られる。

[0158] したがって、復号対象フレーム i を復号することで、サブフレーム 1101 とサブフレーム 1102 とを連結した信号フレーム $i-1$ の信号 $[a_{i-1}, b_{i-1}]$ が得られる。

[0159] なお、以上の説明では、図 12A の (a) に示されるサブフレーム 1001 に窓処理を適用した、図 12A の (b) に示されるサブフレーム 1001 には、畳み込み処理と窓処理とを適用した。これは、上述のように成分 X が $a_{i-1} w_5 + (b_{i-1} w_6)_R$ と表される場合の処理である。成分 X が、 $(a_{i-1} w_5)_R + b_{i-1} w_6$ である場合、図 12A の (a) に示されるサブフレーム 1001 には畳み込み処理と窓処理とが適用され、図 12A の (b) に示されるサブフレーム 1001 には、窓処理が適用される。

[0160] [1-3. 遅延量]

次に、以上説明した実施の形態 1 に係る符号化・復号処理の遅延量について説明する。

[0161] 図 13 は、実施の形態 1 に係る符号化・復号処理の遅延量を示す図である。なお、図 13 において、フレーム $i-1$ に対する符号化処理は、時間 t に

において開始されるものとする。

[0162] 既に述べたように、AAC-ELDにおける低遅延フィルタバンクの窓の特徴により、フレーム $i-1$ のIMDCT変換された出力

[0163] [数25]

$$\bar{y}_{i-1}$$

は、時間 $t + 3 * N / 4$ サンプルにおいて得られる。すなわち、サブフレーム1002、及び1003は、時間 $t + 3 * N / 4$ サンプルにおいて得られる。

[0164] サブフレーム1004及びサブフレーム1005は、先行フレームを復号して再構成された信号であるため、既に取得されている。

[0165] また、時間 $t + 2 N$ サンプルにおいて、フレーム i のACELP合成信号が得られる。すなわち、サブフレーム1001（成分X）は、時間 $t + 2 N$ サンプルにおいて得られる。しかしながら、サブフレーム1001には、前半 $N / 4$ 個のサンプルに相当する部分がゼロである合成窓 $w_{R, s}$ が適用されることから、サブフレーム1001を完全に取得する $N / 4$ サンプル前に音の出力を開始することができる。

[0166] このため、上述のようにサブフレーム1001～1005を用いた信号 $[a_{i-1}, b_{i-1}]$ が再構成されて出力される場合の遅延量は、 $2 N - N / 4 = 7 * N / 4$ サンプルである。

[0167] [1-4. まとめ]

以上、説明したように、音信号ハイブリッドエンコーダ500及び音信号ハイブリッドデコーダ900によれば、符号化モードがFD符号化モードからACELP符号化モードに切り替えられた最初のフレームである遷移フレームを復号する場合に発生するエイリアシングを低減することができ、FD復号技術とACELP復号技術とのシームレスな切替が実現される。

[0168] なお、図14に示されるように、音信号ハイブリッドデコーダ900は、さらに、TCXデコーダ906を備えてもよい。

- [0169] 図14に示されるTCXデコーダ906は、TCX復号処理によってTCX係数からTCX合成信号を生成する。すなわち、TCXデコーダ906は、TCX符号化モードによって符号化されたフレームを復号する。
- [0170] なお、さらなる高音質を実現するために、音信号ハイブリッドデコーダ900は、さらに合成エラー補償（SEC: Synthesis Error Compensation）装置を備えてもよい。
- [0171] SEC処理は、最終的な合成信号を生成するために復号対象フレーム*i*を復号する時点において行われる。SEC装置を追加する目的は、音質を向上させるために、音信号ハイブリッドデコーダ900において符号化モードが切り替えられることによって発生する合成エラーを低減（除去）することである。
- [0172] 図15は、合成エラー補償装置を用いてフレーム*i-1*の信号を再構成する方法を示す図である。ここでは、時間領域のエイリアシングの影響を効率的に補償するために、再構成後の信号 $[a_{i-1}, b_{i-1}]$ にSEC処理が行われる。
- [0173] SEC装置は、復号対象フレームにおいて、符号化処理時にDCT-IV、AVQ方式などで変換して算出した合成エラー情報を復号する。復号された合成エラー情報は、SEC処理により再構成後の信号 $[a_{i-1}, b_{i-1}]$ に加算され、再構成後の信号は修正される。具体的には、図15の(a)に示されるようにサブフレーム1101は、サブフレーム2901に修正され、図15の(b)に示されるようにサブフレーム1102は、サブフレーム2902に修正される。
- [0174] 音信号ハイブリッドデコーダ900側でSEC処理を行うためには、上記合成エラー情報を音信号ハイブリッドエンコーダ500側で符号化しておくことが必要である。
- [0175] 図16は、合成エラー情報の符号化、復号方法を示す図である。
- [0176] 図16に示されるように、合成エラー情報を符号化する場合、音信号ハイブリッドエンコーダ500は、ローカルデコーダ508と、ローカルエンコ

ーダとを備える。

[0177] ローカルデコーダ508は、エンコーダ（ACELPエンコーダ504、FDエンコーダ505、またはTCXエンコーダ507）で符号化された元の信号（符号化前の信号）を復号する。再構成された信号（復号された元の信号）と、元の信号との差分が合成エラー情報である。

[0178] ローカルエンコーダ509は、DCT-IV、AVQ（Adaptive Vector Quantization）などを用いて合成エラー情報を符号化（変換）する。符号化された合成エラー情報は、音信号ハイブリッドデコーダ900が備えるSEC装置907によって復号（逆変換）され、図15を用いて説明したようなSEC処理による再構成後の信号の修正に用いられる。

[0179] （実施の形態2）

実施の形態2では、符号化モードをACELP符号化モードからFD符号化モードに切り替える場合の音信号ハイブリッドエンコーダ500の符号化方法及び音信号ハイブリッドデコーダ900の復号方法について説明する。なお、音信号ハイブリッドエンコーダ500及び音信号ハイブリッドデコーダ900の構成は、実施の形態1と同じである。

[0180] [2-1. 符号化方法]

図17は、符号化モードがACELP符号化モードからFD符号化モードに切り替えられる場合の符号化されたフレームを示す図である。

[0181] フレーム*i*-1は、ACELP符号化モードによって符号化される。フレーム*i*は、FD符号化モードによって先行する3つのフレーム*i*-3、*i*-2、*i*-1と連結されて符号化される。

[0182] [2-2. 復号方法]

以下、音信号ハイブリッドエンコーダ500によって図17に示されるように符号化された符号化信号を復号する音信号ハイブリッドデコーダ900の復号方法について説明する。

[0183] 通常、復号対象フレーム*i*を復号する場合、上述のように先行する3つの

フレーム $i-3$ 、 $i-2$ 、 $i-1$ と重複加算処理を行うことでフレーム $i-1$ の信号が得られる。

[0184] しかしながら、上記重複加算処理は、連続するフレームが全てFD符号化モードによって符号化されていることが前提の処理である。ここで、フレーム i が、符号化モードがACELP符号化モードからFD符号化モードに切り替えられた場合の遷移フレームである場合、先行する3つのフレームである、フレーム $i-3$ 、 $i-2$ 、 $i-1$ はACELP符号化モードで符号化されている。このため、復号対象フレーム i を通常のFD復号処理した場合、エイリアシングが発生することとなる。同様に、フレーム $i+1$ 、フレーム $i+2$ においても先行する3つのフレームにACELP符号化モードで符号化されたフレームが含まれるため、エイリアシングが発生することとなる。

[0185] [2-2-1. 復号対象フレーム i の復号方法]

図18は、ACELP符号化モードで符号化された信号がFD符号化モードで符号化された信号に切り替わるときの、ブロック切替部904の切替制御（復号方法）を示す模式図である。

[0186] 復号対象フレーム i を復号してフレーム $i-1$ の信号 $[a_{i-1}, b_{i-1}]$ を再構成する場合、エイリアシング成分を低減するために、ブロック切替部904は、次の3つの信号を用いて復号処理を行う。

[0187] 第1に、復号対象フレーム i をAAC-ELD低遅延フィルタバンクによって逆変換した後、窓処理した信号のうちのフレーム $i-3$ に対応する部分の信号が用いられる。この信号は、図18においてサブフレーム1401及びサブフレーム1402と示される。

[0188] 第2に、復号対象フレーム $i-1$ をACELP復号処理することで得られるACELP合成信号 $[a_{i-1}, b_{i-1}]$ が用いられる。この信号は、図18においてサブフレーム1403及び1404と示される信号である。

[0189] 第3に、復号対象フレーム $i-3$ をACELP復号処理することによって得られるフレーム $i-3$ の信号 $[a_{i-3}, b_{i-3}]$ が用いられる。フレーム $i-3$ の信号は、図18においてサブフレーム1407及びサブフレーム14

08と示される。

[0190] 次に、上記3つの信号を用いた復号処理についてより詳細に説明する。

[0191] 図19は、フレーム*i*-1の信号 $[a_{i-1}, b_{i-1}]$ を再構成する方法のフローチャートである。

[0192] 復号対象フレーム*i*をAAC-ELD低遅延フィルタバンクによって逆変換した後、窓処理した信号（第8信号）を生成する（図19のS301）。第8信号は、以下の式で示される。

[0193] [数26]

$$\begin{aligned} \bar{y}_i = & [(-a_{i-3}w_1 - (b_{i-3}w_2)_R + a_{i-1}w_5 + (b_{i-1}w_6)_R)w_{R,8}, \\ & (-a_{i-3}w_1)_R - b_{i-3}w_2 + (a_{i-1}w_5)_R + b_{i-1}w_6)w_{R,7}, \\ & (-a_{i-2}w_3 + (b_{i-2}w_4)_R + a_iw_7 - (b_iw_8)_R)w_{R,6}, \\ & ((a_{i-2}w_3)_R - b_{i-2}w_4 - (a_iw_7)_R + b_iw_8)w_{R,5}, \\ & (a_{i-3}w_1 + (b_{i-3}w_2)_R - a_{i-1}w_5 - (b_{i-1}w_6)_R)w_{R,4}, \\ & ((a_{i-3}w_1)_R + b_{i-3}w_2 - (a_{i-1}w_5)_R - b_{i-1}w_6)w_{R,3}, \\ & (a_{i-2}w_3 - (b_{i-2}w_4)_R - a_iw_7 + (b_iw_8)_R)w_{R,2}, \\ & (-(a_{i-2}w_3)_R + b_{i-2}w_4 + (a_iw_7)_R - b_iw_8)w_{R,1}] \end{aligned}$$

[0194] このうち、フレーム*i*-3に対応する部分の信号（図18においてサブフレーム1401及びサブフレーム1402と示される信号）は、それぞれ以下の式で示される。

[0195] [数27]

$$[c_{-4}]_i = (-a_{i-3}w_1 - (b_{i-3}w_2)_R + a_{i-1}w_5 + (b_{i-1}w_6)_R)w_{R,8}$$

[0196] [数28]

$$[d_{-4}]_i = (-a_{i-3}w_1)_R - b_{i-3}w_2 + (a_{i-1}w_5)_R + b_{i-1}w_6)w_{R,7}$$

[0197] 図20Aは、フレーム*i*−1の信号 $[a_{i-1}, b_{i-1}]$ を再構成する方法の一例を示す図である。第*i*−1フレームをACELP復号処理によって復号することで得られる信号を窓処理した窓処理した信号（第4信号）に、当該第4信号を畳み込み処理した信号を加算した信号は、

[0198] [数29]

$$[a_{i-1}w_7 - (b_{i-1}w_8)_R, -(a_{i-1}w_7)_R + b_{i-1}w_8]$$

のように示される。窓 $[w_{R,6}, w_{R,5}]$ を

[0199] [数30]

$$[a_{i-1}w_7 - (b_{i-1}w_8)_R, -(a_{i-1}w_7)_R + b_{i-1}w_8]$$

に適用し、信号

[0200] [数31]

$$[a_{i-1}w_7w_{R,6} - (b_{i-1}w_8)_Rw_{R,6}, -(a_{i-1}w_7)_Rw_{R,5} + b_{i-1}w_8w_{R,5}]$$

（第5信号）が生成される（図19のS302）。第5信号は、図20Aにおいてサブフレーム1501及びサブフレーム1502と示される。

[0201] 図20Bは、フレーム*i*−1の信号 $[a_{i-1}, b_{i-1}]$ を再構成する方法の一例を示すもう一つの図である。第*i*−3フレームをACELP復号処理によって復号することで得られる信号を窓処理した第6信号に、当該第6信号を畳み込み処理した信号を加算した信号は、

[0202] [数32]

$$[a_{i-3}w_1 + (b_{i-3}w_2)_R, (a_{i-3}w_1)_R + b_{i-3}w_2]$$

のように示される。この信号に窓 $[w_{R,8}, w_{R,7}]$ を適用することで、

[0203]

[数33]

$$[a_{i-3}w_1w_{R,8} + (b_{i-3}w_2)_R w_{R,8}, (a_{i-3}w_1)_R w_{R,7} + b_{i-3}w_2w_{R,7}]$$

(第7信号)が得られる(図19のS303)。

[0204] 図20Bに示されるように、この第7信号と、第6信号(サブフレーム1501及びサブフレーム1502)、ならびにフレーム*i*から拡張されたエイリアシング成分である第8信号(サブフレーム1401及びサブフレーム1402)とが加算されることにより、フレーム*i-1*の再構成信号 $[a_{i-1}, b_{i-1}]$ が生成される(図19のS304)。

[0205] [2-2-2. 復号対象フレーム*i+1*の復号方法]

復号対象フレーム*i+1*を復号してフレーム*i*の信号 $[a_i, b_i]$ を再構成する場合、エイリアシング成分を低減するために、ブロック切替部904は、次の3つの信号を用いて復号処理を行う。

[0206] 第1に、復号対象フレーム*i+1*をAAC-ELD低遅延フィルタバンクによって逆変換した後、窓処理した信号のうちのフレーム*i-2*に対応する部分の信号(第9信号)が用いられる。復号対象フレーム*i+1*をAAC-ELD低遅延フィルタバンクによって逆変換し、窓処理した信号は、

[0207] [数34]

$$\bar{y}_{i+1}$$

と示される。

[0208] [数35]

$$\bar{y}_{i+1}$$

から抽出される、フレーム*i-2*に対応する部分(エイリアシング部分)は、以下の通りである。

[0209]

[数36]

$$[c_{-4}, d_{-4}]_{i+1} = [(-a_{i-2}w_1 - (b_{i-2}w_2)_R + a_iw_5 + (b_iw_6)_R)w_{R,8}, \\ (-a_{i-2}w_1)_R - b_{i-2}w_2 + (a_iw_5)_R + b_iw_6]w_{R,7}]$$

[0210] 第2に、復号対象フレーム i を AAC-ELD 低遅延フィルタバンクによって逆変換した後、窓処理した信号のうちのフレーム $i-2$ に対応する部分の信号（第10信号）が用いられる。復号対象フレーム i を AAC-ELD 低遅延フィルタバンクによって逆変換し、窓処理した信号は、

[0211] [数37]

$$\bar{y}_i$$

と示され、この式から抽出される、フレーム $i-2$ に対応する部分は、以下の通りである。

[0212] [数38]

$$[c_{-3}, d_{-3}]_i = [(-a_{i-2}w_3 + (b_{i-2}w_4)_R + a_iw_7 - (b_iw_8)_R)w_{R,6}, \\ ((a_{i-2}w_3)_R - b_{i-2}w_4 - (a_iw_7)_R + b_iw_8)w_{R,5}]$$

[0213] 第3に、

[0214] [数39]

$$\bar{y}_i$$

から抽出される上記フレーム $i-2$ に対応する部分と、

[0215] [数40]

$$\bar{y}_{i+1}$$

から抽出されるフレーム $i-2$ に対応する部分に加えて、復号対象フレーム $i-2$ を ACELP 復号処理することによって得られるフレーム $i-2$ の信

号 $[a_{i-2}, b_{i-2}]$ が用いられる。この信号は、図 18 において、サブフレーム 1405 及びサブフレーム 1406 と示される。

[0216] 図 21 は、フレーム i の信号を再構成する方法の一例を示す図である。

[0217] フレーム $i-2$ の信号 $[a_{i-2}, b_{i-2}]$ (第 11 信号) に窓処理 $[w_1, w_2]$ (第 1 の窓処理) をした信号のうちフレーム前半に相当する信号は、 $a_{i-2}W_1$ と示される。この信号にフレーム $i-2$ の信号に窓処理をした信号のうちフレーム後半に相当する信号である $b_{i-2}W_2$ を畳み込み処理した信号 ($b_{i-2}W_2$)_R が加算されることにより第 12 信号が生成される。

[0218] さらに、この第 12 信号に、当該第 12 信号を畳み込み処理した信号を組み合わせる (連結する) ことによって、信号

[0219] [数41]

$$[(a_{i-2}w_1 + (b_{i-2}w_2)_R), (a_{i-2}w_1)_R + b_{i-2}w_2]$$

が得られる。ここで、窓 $[w_{R,8}, w_{R,7}]$ が

[0220] [数42]

$$[(a_{i-2}w_1 + (b_{i-2}w_2)_R), (a_{i-2}w_1)_R + b_{i-2}w_2]$$

に適用されて、第 13 信号 (エイリアシング成分)

[0221] [数43]

$$[(a_{i-2}w_1 + (b_{i-2}w_2)_R)w_{R,8}, ((a_{i-2}w_1)_R + b_{i-2}w_2)w_{R,7}]$$

が得られる。

[0222] 一方、フレーム $i-2$ の信号に窓処理 $[w_3, w_4]$ (第 2 の窓処理) をした信号のうちフレーム前半に相当する信号は、 $a_{i-2}W_3$ と示される。この信号にフレーム $i-2$ の信号に窓処理をした信号のうちフレーム後半に相当する信号である $b_{i-2}W_4$ を畳み込み処理した信号 ($b_{i-2}W_4$)_R が加算されることにより第 14 信号が生成される。

[0223] さらに、この第14信号に、当該15信号を畳み込み処理し、なおかつ符号を反転させた（-1を掛け合わせた）信号を組み合わせる（連結する）ことによって、信号

[0224] [数44]

$$[(a_{i-2}w_3 - (b_{i-2}w_4)_R, -(a_{i-2}w_3)_R + b_{i-2}w_4]$$

が得られる。ここで、窓 $[w_{R,6}, w_{R,5}]$ が

[0225] [数45]

$$[(a_{i-2}w_3 - (b_{i-2}w_4)_R, -(a_{i-2}w_3)_R + b_{i-2}w_4]$$

に適用されて、第15信号（エイリアシング成分）

[0226] [数46]

$$[(a_{i-2}w_3 - (b_{i-2}w_4)_R)w_{R,6}, (-(a_{i-2}w_3)_R + b_{i-2}w_4)w_{R,5}]$$

が得られる。

[0227] 最後に、図21に示されるように、エイリアシングが低減されたフレーム i の信号 $[a_i, b_i]$ を取得するために、

[0228] [数47]

$$\bar{y}_{i+1}$$

及び

[0229] [数48]

$$\bar{y}_i$$

から抽出された第9信号及び第10信号に第15信号が加算される。

[0230]

[数49]

$$\begin{aligned}
& [(-a_{i-2}w_1 - (b_{i-2}w_2)_R + a_iw_5 + (b_iw_6)_R)w_{R,8}, \\
& \quad (-(a_{i-2}w_1)_R - b_{i-2}w_2 + (a_iw_5)_R + b_iw_6)w_{R,7}] \\
& + [(-a_{i-2}w_3 + (b_{i-2}w_4)_R + a_iw_7 - (b_iw_8)_R)w_{R,6}, \\
& \quad ((a_{i-2}w_3)_R - b_{i-2}w_4 - (a_iw_7)_R + b_iw_8)w_{R,5}] \\
& + [(a_{i-2}w_1 + (b_{i-2}w_2)_R)w_{R,8}, ((a_{i-2}w_1)_R + b_{i-2}w_2)w_{R,7}] \\
& + [(a_{i-2}w_3 - (b_{i-2}w_4)_R)w_{R,6}, (-(a_{i-2}w_3)_R + b_{i-2}w_4)w_{R,5}] \\
& = [a_i(w_5w_{R,8} + w_7w_{R,6}), b_i(w_6w_{R,7} + w_8w_{R,5})]
\end{aligned}$$

[0231] ここで、上述の窓の特性を考慮すると、復号対象フレーム $i+1$ からフレーム i の信号 $[a_i, b_i]$ (サブフレーム 1701 及び 1702) が再構成される。

[0232] [2-2-3. 復号対象フレーム $i+2$ の復号方法]

復号対象フレーム $i+2$ を復号してフレーム $i+1$ の信号 $[a_{i+1}, b_{i+1}]$ を再構成する場合、エイリアシング成分を低減するために、ブロック切替部 904 は、次の 5 つの信号を用いて復号処理を行う。

[0233] 第 1 に、フレーム $i+2$ を AAC-ELD 低遅延フィルタバンクによって逆変換した後、窓処理した信号のうちのフレーム $i-1$ に対応する部分 (エイリアシング部分) の信号 (第 16 信号) が用いられる。フレーム $i+2$ を AAC-ELD 低遅延フィルタバンクによって逆変換し、窓処理した信号は、

[0234] [数50]

$$\bar{y}_{i+2}$$

と示される。

[0235] [数51]

$$\bar{y}_{i+2}$$

から抽出される、フレーム $i-1$ に対応する部分（エイリアシング部分）は、以下の通りである。

[0236] [数52]

$$[c_{-4}, d_{-4}]_{i+2} = [(-a_{i-1}w_1 - (b_{i-1}w_2)_R + a_{i+1}w_5 + (b_{i+1}w_6)_R)w_{R,8}, \\ -(a_{i-1}w_1)_R - b_{i-1}w_2 + (a_{i+1}w_5)_R + b_{i+1}w_6]w_{R,7}]$$

[0237] 第2に、フレーム i を AAC-ELD 低遅延フィルタバンクによって逆変換した後、窓処理した信号のうちのフレーム $i-1$ に対応する部分（エイリアシング部分）の信号（第18信号）が用いられる。フレーム i を AAC-ELD 低遅延フィルタバンクによって逆変換し、窓処理した信号は、

[0238] [数53]

$$\bar{y}_i$$

と示される。

[0239] 第3に、フレーム $i+1$ を AAC-ELD 低遅延フィルタバンクによって逆変換した後、窓処理した信号のうちのフレーム $i-1$ に対応する部分（エイリアシング部分）の信号（第17信号）が用いられる。フレーム $i+1$ を AAC-ELD 低遅延フィルタバンクによって逆変換し、窓処理した信号は、

[0240] [数54]

$$\bar{y}_{i+1}$$

と示される。上記第18信号は、以下の通りである。

[0241]

[数55]

$$[c_{-3}, d_{-3}]_{i+1} = [(-a_{i-1}w_3 + (b_{i-1}w_4)_R + a_{i+1}w_7 - (b_{i+1}w_8)_R)w_{R,6}, \\ ((a_{i-1}w_3)_R - b_{i-1}w_4 - (a_{i+1}w_7)_R + b_{i+1}w_8)w_{R,5}]$$

また、上記第17信号は、以下の通りである。

[0242] [数56]

$$[c_{-2}, d_{-2}]_i = [(a_{i-3}w_1 + (b_{i-3}w_2)_R - a_{i-1}w_5 - (b_{i-1}w_6)_R)w_{R,4}, \\ ((a_{i-3}w_1)_R + b_{i-3}w_2 - (a_{i-1}w_5)_R - b_{i-1}w_6)w_{R,3}]$$

[0243] 第4に、

[0244] [数57]

$$\bar{y}_i$$

から抽出される上記第18信号と、

[0245] [数58]

$$\bar{y}_{i+1}$$

から抽出される上記第17信号と、

[0246] [数59]

$$\bar{y}_{i+2}$$

から抽出される上記第16信号に加えて、図18においてサブフレーム1407及びサブフレーム1408と示される信号（第19信号）が用いられる。サブフレーム1407及びサブフレーム1408は、フレーム*i*-3をACELP復号処理によって復号した信号 $[a_{i-3}, b_{i-3}]$ である。

[0247] 第5に、図20Bにおいてサブフレーム1601及びサブフレーム1602と示されるフレーム*i*-1の再構成信号 $[a_{i-1}, b_{i-1}]$ が用いられる。

[0248] 図22は、フレーム*i*+1の信号を再構成する方法の一例を示す図である。

[0249] フレーム*i*-3の信号 $[a_{i-3}, b_{i-3}]$ （第19信号）に窓処理 $[w_1, w_2]$ をした信号のうちフレーム前半に相当する信号は、 $a_{i-3}W_1$ と示される。この信号にフレーム*i*-3の信号に窓処理をした信号のうちフレーム後半に相当する信号である $b_{i-3}W_2$ を畳み込み処理した信号 $(b_{i-3}W_2)_R$ が加算されることにより第20信号が生成される。

[0250] さらに、この第20信号に、当該第20信号を畳み込み処理した信号を組み合わせる（連結する）ことによって、信号

[0251] [数60]

$$-[a_{i-3}w_1 + (b_{i-3}w_2)_R, (a_{i-3}w_1)_R + b_{i-3}w_2]$$

が得られる。ここで、窓 $[w_{R,4}, w_{R,3}]$ が

[0252] [数61]

$$-[a_{i-3}w_1 + (b_{i-3}w_2)_R, (a_{i-3}w_1)_R + b_{i-3}w_2]$$

に適用されて、第21信号（エイリアシング成分）

[0253] [数62]

$$-[(a_{i-3}w_1 + (b_{i-3}w_2)_R)w_{R,4}, ((a_{i-3}w_1)_R + b_{i-3}w_2)w_{R,3}]$$

が得られる。

[0254] 一方、フレーム*i*-1の再構成信号 $[a_{i-1}, b_{i-1}]$ に窓処理 $[w_7, w_8]$ をした信号のうちフレーム前半に相当する信号は、 $a_{i-1}W_7$ と示される。この信号にフレーム*i*-1の再構成信号に窓処理をした信号のうちフレーム後半に相当する信号である $b_{i-1}W_8$ を畳み込み処理した信号 $(b_{i-1}W_8)_R$ が加算されることにより第22信号が生成される。

[0255] さらに、この第22信号に、当該第22信号を畳み込み処理し、なおかつ符号を反転させた（-1を掛け合わせた）信号を組み合わせる（連結する）

ことによって、信号

[0256] [数63]

$$[-a_{i-1}w_7 + (b_{i-1}w_8)_R, (a_{i-1}w_7)_R - b_{i-1}w_8]$$

が得られる。ここで、窓 $[w_{R,2}, w_{R,1}]$ が

[0257] [数64]

$$[-a_{i-1}w_7 + (b_{i-1}w_8)_R, (a_{i-1}w_7)_R - b_{i-1}w_8]$$

に適用されて、第23信号（エイリアシング成分）

[0258] [数65]

$$[(-a_{i-1}w_7 + (b_{i-1}w_8)_R)w_{R,2}, ((a_{i-1}w_7)_R - b_{i-1}w_8)w_{R,1}]$$

が得られる。

[0259] 最後に、図22に示されるように、エイリアシングが低減されたフレーム $i+1$ の信号 $[a_i, b_i]$ を取得するために、

[0260] [数66]

$$\bar{y}_i$$

[0261] [数67]

$$\bar{y}_{i+1}$$

及び

[0262] [数68]

$$\bar{y}_{i+2}$$

から抽出された第16信号、第17信号、及び第18信号と、上記第21信号と、上記第23信号とが加算される。

[0263] [数69]

$$\begin{aligned}
 & [(-a_{i-1}w_1 - (b_{i-1}w_2)_R + a_{i+1}w_5 + (b_{i+1}w_6)_R)w_{R,8}, \\
 & \quad (-a_{i-1}w_1)_R - b_{i-1}w_2 + (a_{i+1}w_5)_R + b_{i+1}w_6)w_{R,7}] \\
 & + [(-a_{i-1}w_3 + (b_{i-1}w_4)_R + a_{i+1}w_7 - (b_{i+1}w_8)_R)w_{R,6}, \\
 & \quad ((a_{i-1}w_3)_R - b_{i-1}w_4 - (a_{i+1}w_7)_R + b_{i+1}w_8)w_{R,5}] \\
 & + [(a_{i-3}w_1 + (b_{i-3}w_2)_R - a_{i-1}w_5 - (b_{i-1}w_6)_R)w_{R,4}, \\
 & \quad ((a_{i-3}w_1)_R + b_{i-3}w_2 - (a_{i-1}w_5)_R - b_{i-1}w_6)w_{R,3}] \\
 & - [(a_{i-3}w_1 + (b_{i-3}w_2)_R)w_{R,4}, ((a_{i-3}w_1)_R + b_{i-3}w_2)w_{R,3}] \\
 & + [(-a_{i-1}w_7 + (b_{i-1}w_8)_R)w_{R,2}, ((a_{i-1}w_7)_R - b_{i-1}w_8)w_{R,1}] \\
 \\
 & = [-a_{i-1}(w_1w_{R,8} + w_3w_{R,6} + w_5w_{R,4} + w_7w_{R,2}) + a_{i+1}(w_5w_{R,8} + w_7w_{R,6}), \\
 & \quad -b_{i-1}(w_2w_{R,7} + w_4w_{R,5} + w_6w_{R,3} + w_8w_{R,1}) + b_{i+1}(w_6w_{R,7} + w_8w_{R,5})]
 \end{aligned}$$

[0264] ここで、上述の窓の特性を考慮すると、復号対象フレーム $i+2$ からフレーム $i+1$ の信号 $[a_{i+1}, b_{i+1}]$ (サブフレーム 1801 及び 1802) が再構成される。

[0265] [2-3. 遅延量]

次に、以上説明した実施の形態 2 に係る符号化・復号処理の遅延量について説明する。

[0266] 図 23 は、実施の形態 2 に係る符号化・復号処理の遅延量を示す図である。なお、図 23 において、フレーム $i-1$ に対する符号化処理は、時間 t において開始されるものとする。

[0267] フレーム $i-1$ の ACELP 合成信号は、時間 $t+N$ サンプルにおいて得られる。すなわち、サブフレーム 1501、及び 1502 (サブフレーム 1403、及び 1404) は、時間 $t+N$ サンプルにおいて得られる。

[0268] サブフレーム 1407 及びサブフレーム 1408 は、先行フレームを復号して再構成された信号であるため、既に取り得られている。

[0269] また、既に述べたように、AAC-ELD における低遅延フィルタバンクの窓の特徴により、フレーム i の IMDCT 変換された出力は、時間 $t+7 * N / 4$ サンプルにおいて得られる。すなわち、サブフレーム 1401、及び 1402 は、時間 $t+7 * N / 4$ サンプルにおいて得られる。しかしなが

ら、サブフレーム1401には、前半 $N/4$ 個のサンプルに相当する部分がゼロである合成窓 $w_{R,8}$ が適用されることから、サブフレーム1401を完全に取得する $N/4$ サンプル前に音の出力を開始することができる。

[0270] このため、上述のように再構成された信号 $[a_{i-1}, b_{i-1}]$ は、時間 $t + 3 * N / 2$ サンプルにおいて出力が開始され、遅延量は、 $(t + 3 * N / 2) - t = 3 * N / 2$ サンプルである。

[0271] [2-4. まとめ]

以上、実施の形態2において説明したように、音信号ハイブリッドエンコーダ500及び音信号ハイブリッドデコーダ900によれば、符号化モードがACELP符号化モードからFD符号化モードに切り替えられた最初のフレームである遷移フレームを復号する場合に発生するエイリアシングを低減することができ、ACELP復号処理とFD復号処理とのシームレスな切替が実現される。

[0272] なお、実施の形態1と同様に、実施の形態2に係る音信号ハイブリッドデコーダ900は、さらに、図14に示されるようなTCXデコーダ906を備えてもよい。

[0273] なお、実施の形態1と同様に、さらなる高音質を実現するために、実施の形態2に係る音信号ハイブリッドデコーダ900は、さらに合成エラー補償(SEC)装置を備えてもよい。

[0274] 図24は、SEC装置を用いてフレーム $i-1$ の信号 $[a_{i-1}, b_{i-1}]$ を再構成する方法を示す図である。図24に示される構成は、図20Bに示される構成にSEC装置を追加したものである。図24に示されるようにサブフレーム1601及び1602は、SEC処理によってそれぞれサブフレーム3101及び3102に修正される。

[0275] 図25は、SEC装置を用いてフレーム i の信号 $[a_i, b_i]$ を再構成する方法を示す図である。図25に示される構成は、図21に示される構成にSEC装置を追加したものである。図25に示されるようにサブフレーム1701及び1702は、SEC処理によってそれぞれサブフレーム3201

及び3202に修正される。

[0276] 図26は、SEC装置を用いてフレーム $i-1$ の信号 $[a_{i+1}, b_{i+1}]$ を再構成する方法を示す図である。図26に示される構成は、図22に示される構成にSEC装置を追加したものである。図26に示されるようにサブフレーム1801及び1802は、SEC処理によってそれぞれサブフレーム3301及び3302に修正される。

[0277] このように、再構成された信号含まれる合成エラーをデコーダに設けられたSEC装置により補償することで、さらに音質を向上させることができる。

[0278] (実施の形態3)

実施の形態3では、符号化モードをFD符号化モードからTCX符号化モードに切り替える場合の音信号ハイブリッドエンコーダ500の符号化方法及び音信号ハイブリッドデコーダ900の復号方法について説明する。

[0279] 音信号ハイブリッドエンコーダ500の構成は、図9に示される構成と同様であるが、図9中のACELPエンコーダ504は、省略可能である。また、音信号ハイブリッドデコーダ900の構成は、図14に示される構成と同様であるが、図14中のACELPデコーダ903は、省略可能である。

[0280] [3-1. 符号化方法]

まず、符号化モードがFD符号化モードからTCX符号化モードに切り替えられる場合のブロック切替部502の制御について説明する。

[0281] 図27は、符号化モードがFD符号化モードがTCX符号化モードに切り替えられる場合の符号化されたフレームを示す図である。

[0282] この場合、フレーム i を符号化するとき、先行するフレーム $i-1$ の信号 $[a_{i-1}, b_{i-1}]$ から生成される成分 X を加えた信号が符号化される。具体的には、ブロック切替部502は、成分 X と、フレーム i の信号 $[a_i, b_i]$ とを合わせた拡張フレームを生成する。拡張フレームは、 $(N+N/2)$ の長さである。拡張フレームは、ブロック切替部502によりTCXエンコーダ507に送信され、TCX符号化モードで符号化される。また、成分 X

は、図 8 A 及び図 8 B を用いて説明した方法と同じ方法で生成される。

[0283] [3-2. 復号方法]

次に、FD 符号化モードで符号化された信号が TCX 符号化モードで符号化された信号に切り替わるときの、ブロック切替部 904 の切替制御（復号方法）について説明する。

[0284] 図 28 は、FD 符号化モードで符号化された信号が TCX 符号モードで符号化された信号に切り替わるときの、ブロック切替部 904 の切替制御（復号方法）を示す模式図である。図 28 に示されるようにフレーム $i-1$ は FD 符号化モードによって符号化されたフレームであり、復号対象フレームであるフレーム i は、TCX 符号化モードで符号化されたフレームである。

[0285] 上述のように、FD 符号化モードで符号化された信号が連続する場合は、復号対象フレーム i を復号してフレーム $i-1$ の信号を再構成することができる。つまり、図 11 に示される場合、フレーム $i-2$ の信号までは通常の FD 復号処理によって再構成できる。しかしながら、フレーム $i-1$ の信号は、復号対象フレーム i が ACELP 符号化モードで符号化されているため、通常の方法で再構成すると、エイリアシング成分による不自然な音が発生してしまう。すなわち、フレーム $i-1$ の信号は、図 11 に示されるようにエイリアシング部分となる。

[0286] エイリアシング成分を低減するために、ブロック切替部 904 は、次の 3 つの信号を用いて復号処理を行う。

[0287] 第 1 に、復号対象フレーム i を TCX 復号処理することで得られる TCX 合成信号の成分 X の信号がエイリアシング成分を低減したフレーム $i-1$ の信号を再構成するために用いられる。この信号は、図 11 においてサブフレーム 2001 と示される信号であり、図 8 A を用いて説明した成分 X である。

[0288] 図 8 A を用いて説明したように、成分 X は、具体的には、 $a_{i-1}w_5 + (b_{i-1}w_6)_R$ である。

[0289] 第 2 に、復号対象フレーム $i-1$ を AAC-ELD 低遅延フィルタバンク

によって逆変換した後、窓処理した信号のうちのフレーム $i-3$ に対応する部分の信号がエイリアシング成分を低減したフレーム $i-1$ の信号を再構成するために用いられる。この信号は、図 28 においてサブフレーム 2002 及びサブフレーム 2003 と示される。

[0290] より具体的には、この信号は、フレーム $i-1$ を AAC-ELD 低遅延フィルタバンクによって、通常フレームとして長さ $4N$ で逆変換し、さらに窓処理することによって得られる。逆変換信号は、

[0291] [数70]

$$\bar{y}_{i-1}$$

と示される。このうち、フレーム $i-3$ に対応する部分の信号（図 28 においてサブフレーム 2002 及びサブフレーム 2003 と示されるエイリアシング部分）は、上記逆変換信号から以下のように抽出される。すなわち、

[0292] [数71]

$$[c_{-3}]_{i-1} = -a_{i-3} w_3 w_{R,6} + (b_{i-3} w_4)_R w_{R,6} + a_{i-1} w_7 w_{R,6} - (b_{i-1} w_8)_R w_{R,6}$$

及び

[0293] [数72]

$$[d_{-3}]_{i-1} = (a_{i-3} w_3)_R w_{R,5} - b_{i-3} w_4 w_{R,5} - (a_{i-1} w_7)_R w_{R,5} + b_{i-1} w_8 w_{R,5}$$

がサブフレーム 2002 とサブフレーム 2003 にそれぞれ対応する信号である。

[0294] 第3に、復号対象フレーム $i-2$ を FD 復号処理することによって得られるフレーム $i-3$ の信号 $[a_{i-3}, b_{i-3}]$ がエイリアシング成分を低減したフレーム $i-1$ の信号を再構成するために用いられる。フレーム $i-3$ の信号は、図 28 においてサブフレーム 2004 及びサブフレーム 2005 と示される。

[0295] 上記の信号を用いたエイリアシング成分を低減したフレーム $i-1$ の信号

を再構成する方法は、図12A及び図12Bを用いて説明した方法と同じである。具体的には、図12Aにおけるサブフレーム1001、1002、1003、1004、1005が、図28のサブフレーム2001、2002、2003、2004、2005にそれぞれ置き換えたと考えればよい。これにより、フレーム*i*の信号 $[a_{i-1}, b_{i-1}]$ が再構成される。

[0296] [3-3. 遅延量]

次に、以上説明した実施の形態1に係る符号化・復号処理の遅延量について説明する。

[0297] 図29は、実施の形態3に係る符号化・復号処理の遅延量を示す図である。なお、図29において、フレーム*i-1*に対する符号化処理は、時間*t*において開始されるものとする。

[0298] 既に述べたように、AAC-ELDにおける低遅延フィルタバンクの窓の特徴により、フレーム*i-1*のIMDCT変換された出力

[0299] [数73]

$$\bar{y}_{i-1}$$

は、時間*t* + 3 * *N* / 4 サンプルにおいて得られる。すなわち、サブフレーム2002、及び2003は、時間*t* + 3 * *N* / 4 サンプルにおいて得られる。

[0300] サブフレーム2004及びサブフレーム2005は、先行フレームを復号して再構成された信号であるため、既に取得されている。

[0301] また、時間*t* + 2 *N* サンプルにおいて、フレーム*i*のTCX合成信号が得られる。すなわち、サブフレーム2001（成分X）は、時間*t* + 2 *N* サンプルにおいて得られる。しかしながら、サブフレーム2001には、前半*N* / 4個のサンプルに相当する部分がゼロである合成窓 $w_{R, s}$ が適用されることから、サブフレーム2001を完全に取得する*N* / 4 サンプル前に音の出力を開始することができる。

[0302] このため、上述のようにサブフレーム2001～2005を用いて信号 [

a_{i-1} , b_{i-1}] が再構成されて出力される場合の遅延量は、 $2N/4 - N/4 = 7 * N/4$ サンプルである。

[0303] [3-4. まとめ]

以上、説明したように、音信号ハイブリッドエンコーダ500及び音信号ハイブリッドデコーダ900によれば、符号化モードがFD符号化モードからTCX符号化モードに切り替えられた最初のフレームである遷移フレームを復号する場合に発生するエイリアシングを低減することができ、FD復号技術とTCX復号技術とのシームレスな切替が実現される。

[0304] なお、さらなる高音質を実現するために、音信号ハイブリッドデコーダ900は、さらに合成エラー補償(SEC)装置を備えてもよい。この場合の信号の再構成方法は、図15に示されるものと同様である。

[0305] (実施の形態4)

実施の形態4では、符号化モードをTCX符号化モードからFD符号化モードに切り替える場合の音信号ハイブリッドエンコーダ500符号化方法及び音信号ハイブリッドデコーダ900の復号方法について説明する。

[0306] 音信号ハイブリッドエンコーダ500の構成は、図9に示される構成と同様であるが、図9中のACELPエンコーダ504は、省略可能である。また、音信号ハイブリッドデコーダ900の構成は、図14に示される構成と同様であるが、図14中のACELPデコーダ903は、省略可能である。

[0307] [4-1. 符号化方法]

図30は、符号化モードがTCX符号化モードからFD符号化モードに切り替えられる場合の符号化されたフレームを示す図である。

[0308] フレーム*i*-1は、TCX符号化モードによって符号化される。フレーム*i*は、FD符号化モードによって先行する3つのフレーム*i*-3、*i*-2、*i*-1と連結されて符号化される。

[0309] [4-2. 復号方法]

以下、音信号ハイブリッドエンコーダ500によって図31に示されるように符号化された符号化信号を復号する音信号ハイブリッドデコーダ900

の復号方法について説明する。

[0310] [4-2-1. 復号対象フレーム i の復号方法]

復号対象フレーム i を復号する場合、エイリアシング成分を低減するために、ブロック切替部 904 は、次の 3 つの信号を用いて復号処理を行う。

[0311] 第 1 に、復号対象フレーム i を AAC-ELD 低遅延フィルタバンクによって逆変換した後、窓処理した信号のうちのフレーム $i-3$ に対応する部分の信号が用いられる。この信号は、図 31 において、サブフレーム 2301 及びサブフレーム 2302 と示される。

[0312] 第 2 に、復号対象フレーム $i-1$ を TCX 復号処理することで得られる TCX 合成信号 $[a_{i-1}, b_{i-1}]$ が用いられる。この信号は、図 31 においてサブフレーム 2303 及び 2304 と示される信号である。

[0313] 第 3 に、復号対象フレーム $i-3$ を TCX 復号処理することによって得られるフレーム $i-3$ の信号 $[a_{i-3}, b_{i-3}]$ が用いられる。フレーム $i-3$ の信号は、図 31 においてサブフレーム 2307 及びサブフレーム 2308 と示される。

[0314] 復号対象フレーム i を AAC-ELD 低遅延フィルタバンクによって逆変換した後、窓処理した信号（第 8 信号）のフレーム $i-3$ に対応する部分の信号（図 31 においてサブフレーム 2301 及びサブフレーム 2302 と示される信号）は、それぞれ以下の式で示される。

[0315] [数 74]

$$[c_{-4}]_i = (-a_{i-3}w_1 - (b_{i-3}w_2)_R + a_{i-1}w_5 + (b_{i-1}w_6)_R)w_{R,8}$$

[0316] [数 75]

$$[d_{-4}]_i = (-(a_{i-3}w_1)_R - b_{i-3}w_2 + (a_{i-1}w_5)_R + b_{i-1}w_6)w_{R,7}$$

[0317] ここで、復号対象フレーム $i-1$ を TCX 復号処理することで得られる TCX 合成信号 $[a_{i-1}, b_{i-1}]$ は、説明の便宜上、

[0318]

[数76]

$$\left[\underbrace{a_{i-1}}_{N/2}, \underbrace{b_{i-1,1}}_{N/4}, \underbrace{b_{i-1,2}}_{N/4} \right]$$

のように分割される。これに対応して、窓 $[w_7, w_8]$ は、

[0319] [数77]

$$\left[\underbrace{w_7}_{N/2}, \underbrace{w_{8,1}}_{N/4}, \underbrace{w_{8,2}}_{N/4} \right]$$

に分割される。サブフレーム 2303 及び 2304 と示される TCX 合成信号は、後続するフレームが TCX 符号化モードで符号化されていないため、エイリアシング成分を含み、

[0320] [数78]

$$\left[\underbrace{a_{i-1}}_{N/2}, \underbrace{b_{i-1,1}}_{N/4}, \underbrace{b_{i-1,2} + \text{aliasing}}_{N/4} \right]$$

と示される。ここで、分析窓 w_8 の特性、すなわち $w_{8,2} = 0$ を考慮して窓 $[w_7, w_8]$ を TCX 合成信号

[0321] [数79]

$$\left[\underbrace{a_{i-1}}_{N/2}, \underbrace{b_{i-1,1}}_{N/4}, \underbrace{b_{i-1,2} + \text{aliasing}}_{N/4} \right]$$

に適用すると、

[0322] [数80]

$$\left[\underbrace{a_{i-1} w_7}_{N/2}, \underbrace{b_{i-1,1} w_{8,1}}_{N/4}, 0 \right]$$

が得られる。これは、図 32 に示される

[0323]

[数81]

$$\left[\underbrace{a_{i-1}}_{N/2} w_7, \underbrace{b_{i-1}}_{N/2} w_8 \right]$$

と実際には等価である。

[0324] したがって、図32に示されるサブフレーム2401、2402を生成する方法は、図20Aに示される方法と同じである。

[0325] つまり、以降の処理は、図20Bを用いて説明した方法と同様である。具体的には、図20Bにおいて、サブフレーム1401、1402、1407、1408、1501、及び1502が、それぞれ、サブフレーム2301、2302、2307、2308、2401、及び2402に置き換えられたと考えればよい。

[0326] [4-2-2. 復号対象フレーム*i*+1の復号方法]

復号対象フレーム*i*+1を復号する場合、エイリアシング成分を低減するために、ブロック切替部904は、次の3つの信号を用いて復号処理を行う。

[0327] 第1に、復号対象フレーム*i*+1をAAC-ELD低遅延フィルタバンクによって逆変換した後、窓処理した信号のうちのフレーム*i*-2に対応する部分の信号（第9信号）が用いられる。

[0328] 第2に、復号対象フレーム*i*をAAC-ELD低遅延フィルタバンクによって逆変換した後、窓処理した信号のうちのフレーム*i*-2に対応する部分の信号（第10信号）が用いられる。

[0329] 以上の第9信号及び第10信号については、図21を用いて説明したものと同様である。

[0330] 第3に、復号対象フレーム*i*-2をTCX復号処理することによって得られるフレーム*i*-2の信号 $[a_{i-2}, b_{i-2}]$ が用いられる。この信号は、図31において、サブフレーム2305及びサブフレーム2306と示される。

- [0331] 上記の3つの信号を用いた、復号対象フレーム $i + 1$ の復号方法は、図21を用いて説明した方法と同様である。具体的には、図21において、サブフレーム1405、1406が、それぞれ、サブフレーム2305、及び2306に置き換えられたと考えればよい。
- [0332] [4-2-3. 復号対象フレーム $i + 2$ の復号方法]
復号対象フレーム $i + 2$ を復号する場合、エイリアシング成分を低減するために、ブロック切替部904は、次の5つの信号を用いて復号処理を行う。
- [0333] 第1に、フレーム $i + 2$ をAAC-ELD低遅延フィルタバンクによって逆変換した後、窓処理した信号のうちのフレーム $i - 1$ に対応する部分（エイリアシング部分）の信号（第16信号）が用いられる。
- [0334] 第2に、フレーム i をAAC-ELD低遅延フィルタバンクによって逆変換した後、窓処理した信号のうちのフレーム $i - 1$ に対応する部分（エイリアシング部分）の信号（第18信号）が用いられる。
- [0335] 第3に、フレーム $i + 1$ をAAC-ELD低遅延フィルタバンクによって逆変換した後、窓処理した信号のうちのフレーム $i - 1$ に対応する部分（エイリアシング部分）の信号（第17信号）が用いられる。
- [0336] これら第16信号、第17信号、及び第18信号の3つの信号は、図22を用いて説明したものと同様である。
- [0337] 第4に、フレーム $i - 3$ をTCX復号処理によって復号した信号 $[a_{i-3}, b_{i-3}]$ が用いられる。
- [0338] 第5に、フレーム $i - 1$ をTCX復号処理によって復号した信号 $[a_{i-1}, b_{i-1}]$ が用いられる。
- [0339] 上記の5つの信号を用いた、復号対象フレーム $i + 2$ の復号方法は、図22を用いて説明した方法と同様である。具体的には、図22において、サブフレーム1407、1408が、それぞれ、サブフレーム2307、及び2308に置き換えられたと考えればよい。また、図22に示されるサブフレーム1601、及び1602は、復号対象フレーム i の復号方法で説明した

方法（図20BにおいてフレームをTCX符号化モードのフレームに置き換える方法）で生成されたフレームに置き換えられたと考えればよい。

[0340] [4-3. 遅延量]

次に、以上説明した実施の形態4に係る符号化・復号処理の遅延量について説明する。

[0341] 図33は、実施の形態4に係る符号化・復号処理の遅延量を示す図である。なお、図33において、フレーム $i-1$ に対する符号化処理は、時間 t において開始されるものとする。

[0342] フレーム $i-1$ のTCX合成信号は、時間 $t+N$ サンプルにおいて得られる。すなわち、サブフレーム2401、及び2402（サブフレーム2303、及び2304）は、時間 $t+N$ サンプルにおいて得られる。

[0343] サブフレーム2307及びサブフレーム2308は、先行フレームを復号して再構成された信号であるため、既に取り得られている。

[0344] また、既に述べたように、AAC-ELDにおける低遅延フィルタバンクの窓の特徴により、フレーム i のIMDCT変換された出力は、時間 $t+7*N/4$ サンプルにおいて得られる。すなわち、サブフレーム2301、及びサブフレーム2302は、時間 $t+7*N/4$ サンプルにおいて得られる。しかしながら、サブフレーム2301には、前半 $N/4$ 個のサンプルに相当する部分がゼロである合成窓 $w_{R,8}$ が適用されることから、サブフレーム2301を完全に取得する $N/4$ サンプル前に音の出力を開始することができる。

[0345] このため、上述のように再構成された信号 $[a_{i-1}, b_{i-1}]$ は、時間 $t+3*N/2$ サンプルにおいて出力が開始され、遅延量は、 $(t+3*N/2) - t = 3*N/2$ サンプルである。

[0346] [4-4. まとめ]

以上、説明したように、音信号ハイブリッドエンコーダ500及び音信号ハイブリッドデコーダ900によれば、符号化モードがTCX符号化モードからFD符号化モードに切り替えられた最初のフレームである遷移フレーム

を復号する場合に発生するエイリアシングを低減することができ、TCX復号技術とFD復号技術とのシームレスな切替が実現される。

[0347] なお、さらなる高音質を実現するために、音信号ハイブリッドデコーダ900は、さらに合成エラー補償(SEC)装置を備えてもよい。この場合の信号の再構成方法は、図24～図26に示されるものと同様である。

[0348] (実施の形態5)

実施の形態5では、過渡信号を符号化する場合の音信号ハイブリッドエンコーダの符号化方法、及び過渡信号を復号する場合の音信号ハイブリッドデコーダの復号方法について説明する。実施の形態5において、音信号ハイブリッドエンコーダ500の構成は、図9に示される構成と同様であるが、図9中のACELPエンコーダ504は、省略可能である。また、音信号ハイブリッドデコーダ900の構成は、図14に示される構成と同様であるが、図14中のACELPデコーダ903は、省略可能である。

[0349] FD符号化モードでは、ロングウィンドウが用いられる(時間幅が大きい窓が用いられる)ため、エネルギー(=信号電力、すなわち符号化フレームにおける音信号の振幅の自乗和に比例する値)が急激に変化する過渡信号の符号化には適していない。すなわち、過渡信号を処理するとき、ショートウィンドウ(時間幅が小さい窓)を用いてもよい。

[0350] [5-1. 符号化方法]

まず、符号化対象フレーム*i*が過渡信号(過渡フレーム)である場合、符号化対象フレーム*i*を符号化するときには、先行するフレーム*i-1*の信号 $[a_{i-1}, b_{i-1}]$ から生成される成分Xを加えた信号が符号化される。具体的には、ブロック切替部502は、成分Xと、フレーム*i*の信号 $[a_i, b_i]$ とを合わせた拡張フレームを生成する。拡張フレームは、 $(N+N/2)$ の長さである。拡張フレームは、ブロック切替部502によりTCXエンコーダ507に送信され、TCX符号化モードで符号化される。なお、このとき、TCXエンコーダ507は、MDCTフィルタバンクのショートウィンドウモードを用いたTCX符号化を行う。このとき、符号化されたフレーム

は、図 27 を用いて説明したものと同様である。また、成分 X は、図 8 A 及び図 8 B を用いて説明した方法と同じ方法で生成される。

[0351] なお、符号化対象フレーム i が過渡信号であるか否かの判断は、例えば、符号化対象フレームにおけるエネルギーが所定の閾値を越えるか否かによってなされるが、このような方法に限定されるものではない。

[0352] [5-2. 復号方法]

上記のように符号化された過渡フレームの復号方法は、FD 符号化モードで符号化された信号が TCX 符号化モードで符号化された信号に切り替わる際の復号方法と同様である。すなわち、図 12 A または図 28 を用いて説明した方法と同様である。

[0353] なお、実施の形態 5 の符号化・復号処理の遅延量は、実施の形態 1 及び 3 と同じであり、 $7 * N / 4$ サンプルである。

[0354] [5-3. まとめ]

以上、説明したように、音信号ハイブリッドデコーダ 900 によれば、FD 符号化モードで符号化を行っているときの過渡フレームにおいて、TCX 符号化モードで符号化し、復号することによって、より音質を向上させることができる。

[0355] なお、さらなる高音質を実現するために、音信号ハイブリッドデコーダ 900 は、さらに合成エラー補償 (SEC) 装置を備えてもよい。この場合の信号の再構成方法は、図 15 に示されるものと同様である。

[0356] (変形例)

以上、本発明を上記実施の形態に基づいて説明してきたが、本発明は、上記の実施の形態に限定されないのはもちろんである。

[0357] 例えば、例えば、LPD 符号化モードとして VSELP (Vector Sum Excited Linear Prediction) 符号化モード等、ACELP 以外の CELP 方式を用いてもよい。復号処理についても同様に、ACELP 以外の CELP 方式を用いてもよい。

[0358] また、本実施の形態では、FD 符号化モードの一例として AAC-ELD

モードについて主に説明したが、本発明は、A A C - E L Dモードのみならず、複数の先行フレームによるオーバーラップ処理が必要な符号化方式に適用可能である。

[0359] また、以下のような場合も本発明に含まれる。

[0360] (1) 上記の各装置は、具体的には、マイクロプロセッサ、ROM、RAM、ハードディスクユニット、ディスプレイユニット、キーボード、マウスなどから構成されるコンピュータシステムで実現され得る。RAMまたはハードディスクユニットには、コンピュータプログラムが記憶されている。マイクロプロセッサが、コンピュータプログラムにしたがって動作することにより、各装置は、その機能を達成する。ここでコンピュータプログラムは、所定の機能を達成するために、コンピュータに対する指令を示す命令コードが複数個組み合わされて構成されたものである。

[0361] (2) 上記の各装置を構成する構成要素の一部または全部は、1個のシステムLSI (Large Scale Integration: 大規模集積回路) から構成されているとしてもよい。システムLSIは、複数の構成部を1個のチップ上に集積して製造された超多機能LSIであり、具体的には、マイクロプロセッサ、ROM、RAMなどを含んで構成されるコンピュータシステムである。ROMには、コンピュータプログラムが記憶されている。マイクロプロセッサが、ROMからRAMにコンピュータプログラムをロードし、ロードしたコンピュータプログラムにしたがって演算等の動作することにより、システムLSIは、その機能を達成する。

[0362] (3) 上記の各装置を構成する構成要素の一部または全部は、各装置に脱着可能なICカードまたは単体のモジュールから構成されてもよい。ICカードまたはモジュールは、マイクロプロセッサ、ROM、RAMなどから構成されるコンピュータシステムである。ICカードまたはモジュールには、上記の超多機能LSIが含まれてもよい。マイクロプロセッサが、コンピュータプログラムにしたがって動作することにより、ICカードまたはモジュールは、その機能を達成する。このICカードまたはこのモジュールは、耐

タンパ性を有してもよい。

[0363] (4) 本発明は、上記に示す方法で実現されてもよい。また、これらの方法をコンピュータにより実現するコンピュータプログラムで実現してもよいし、コンピュータプログラムからなるデジタル信号で実現してもよい。

[0364] また、本発明は、コンピュータプログラムまたはデジタル信号をコンピュータ読み取り可能な記録媒体、例えば、フレキシブルディスク、ハードディスク、CD-ROM、MO、DVD、DVD-ROM、DVD-RAM、BD (Blu-ray Disc)、半導体メモリなどに記録したもので実現してもよい。また、これらの記録媒体に記録されているデジタル信号で実現してもよい。

[0365] また、本発明は、コンピュータプログラムまたはデジタル信号を、電気通信回線、無線または有線通信回線、インターネットを代表とするネットワーク、データ放送等を経由して伝送してもよい。

[0366] また、本発明は、マイクロプロセッサとメモリを備えたコンピュータシステムであって、メモリは、コンピュータプログラムを記憶しており、マイクロプロセッサは、コンピュータプログラムにしたがって動作してもよい。

[0367] また、プログラムまたはデジタル信号を記録媒体に記録して移送することにより、またはプログラムまたはデジタル信号をネットワーク等を経由して移送することにより、独立した他のコンピュータシステムにより実施するとしてもよい。

[0368] (5) 上記実施の形態及び上記変形例をそれぞれ組み合わせるとしてもよい。

[0369] なお、本発明は、これらの実施の形態またはその変形例に限定されるものではない。本発明の趣旨を逸脱しない限り、当業者が思いつく各種変形を本実施の形態またはその変形例に施したものの、あるいは異なる実施の形態またはその変形例における構成要素を組み合わせる構築される形態も、本発明の範囲内に含まれる。

産業上の利用可能性

[0370] 本発明の音信号ハイブリッドデコーダ及び音信号ハイブリッドエンコーダは、高音質で低遅延な音信号の符号化及び復号が可能であり、放送システム、携帯用テレビ、携帯電話通信、テレビ会議などに用いることができる。

符号の説明

- [0371] 500 音信号ハイブリッドエンコーダ
501 高周波エンコーダ
502 ブロック切替部
503 信号分類部
504 ACELPエンコーダ
505 FDエンコーダ
506 ビットマルチプレクサ
507 TCXエンコーダ
508 ローカルデコーダ
509 ローカルエンコーダ
900 音信号ハイブリッドデコーダ
901 デマルチプレクサ
902 FDデコーダ
903 ACELPデコーダ
904 ブロック切替部
905 高周波デコーダ
906 TCXデコーダ
907 SEC装置
1001~1005、1101、1102 サブフレーム
1401~1408、1501、1502、1601、1602 サブフレーム
1701、1702、1801、1802 サブフレーム
2001~2005、2301~2308、2401、2402 サブフレーム

2901、2902、3101、3102、3201、3202 サブフ
レーム

3301、3302 サブフレーム

請求の範囲

[請求項1]

低遅延フィルタバンクを用いた音響符号化処理で符号化された音響フレームと、線形予測係数を用いた音声符号化処理で符号化された音声フレームとが含まれるビットストリームを復号する音信号ハイブリッドデコーダであって、

前記音響フレームを低遅延逆フィルタバンク処理を用いて復号する低遅延変換デコーダと、

前記音声フレームを復号する音声信号デコーダと、

前記ビットストリームのうちの復号対象フレームが前記音響フレームである場合、当該復号対象フレームを前記低遅延変換デコーダによって復号し、前記復号対象フレームが前記音声フレームである場合、当該復号対象フレームを前記音声信号デコーダによって復号する制御を行うブロック切替部とを備え、

前記復号対象フレームが、前記音響フレームから前記音声フレームに切り替わった最初の前記音声フレームである第 i フレームであるとき、

前記第 i フレームには、前記第 i フレームよりも1フレーム先行するフレームである第 $i - 1$ フレームの符号化前の信号を用いて生成された第1信号が符号化された状態で含まれ、

前記ブロック切替部は、

(1)

前記第 i フレームよりも2フレーム先行するフレームである第 $i - 2$ フレームを前記低遅延変換デコーダによって復号することで得られる、前記第 i フレームよりも3フレーム先行するフレームである第 $i - 3$ フレームの再構成された信号を窓処理した信号である第2信号のフレームの前半部分に相当する信号に、前記第2信号のフレームの後半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第 i フレームを前記音声信号デコーダによって復号

することで得られる、前記第1信号に窓処理を行った信号と、前記第 $i - 1$ フレームを前記低遅延逆フィルタバンク処理及び窓処理した信号の前記第 $i - 3$ フレームに対応する部分である第3信号のフレームの前半部分の信号と、を加算する処理を行って符号化前の前記第 $i - 1$ フレームの前半部分に対応する信号を生成し、

前記第2信号のフレームの後半部分に相当する信号に、前記第2信号のフレームの前半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第1信号に畳み込み処理及び窓処理を行った信号と、前記第3信号のフレームの後半部分に相当する信号と、を加算する処理を行って符号化前の前記第 $i - 1$ フレームの後半部分に対応する信号を生成する、または

(2)

前記第2信号のフレームの前半部分に相当する信号に、前記第2信号のフレームの後半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第1信号に畳み込み処理及び窓処理を行った信号と、前記第3信号のフレームの前半部分に相当する信号と、を加算する処理を行って符号化前の前記第 $i - 1$ フレームの前半部分に対応する信号を生成し、

前記第2信号のフレームの後半部分に相当する信号に、前記第2信号のフレームの前半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第1信号に窓処理を行った信号と、前記第3信号のフレームの後半部分に相当する信号と、を加算する処理を行って符号化前の前記第 $i - 1$ フレームの後半部分に対応する信号を生成する

音信号ハイブリッドデコーダ。

[請求項2]

低遅延フィルタバンクを用いた音響符号化処理で符号化された音響フレームと、線形予測係数を用いた音声符号化処理で符号化された音声フレームとが含まれるビットストリームを復号する音信号ハイブリ

ッドデコーダであって、

前記音響フレームを低遅延逆フィルタバンク処理によって復号する低遅延変換デコーダと、

前記音声フレームを復号する音声信号デコーダと、

前記ビットストリームのうちの復号対象フレームが前記音響フレームである場合、当該復号対象フレームを前記低遅延変換デコーダによって復号し、前記復号対象フレームが前記音声フレームである場合、当該復号対象フレームを前記音声信号デコーダによって復号する制御を行うブロック切替部とを備え、

前記ブロック切替部は、

前記復号対象フレームが、前記音声フレームから前記音響フレームに切り替わった最初の音響フレームである第 i フレームであるとき、

前記第 i フレームよりも 1 フレーム先行するフレームである第 $i - 1$ フレームを前記音声信号デコーダによって復号することで得られる信号を窓処理した第 4 信号に、当該第 4 信号を畳み込み処理した信号を加算し、窓処理を行った第 5 信号と、前記第 i フレームよりも 3 フレーム先行するフレームである第 $i - 3$ フレームを前記音声信号デコーダによって復号することで得られる信号を窓処理した第 6 信号に、当該第 6 信号を畳み込み処理した信号を加算し、窓処理を行った第 7 信号と、前記第 i フレームを前記低遅延逆フィルタバンク処理及び窓処理した信号の前記第 $i - 3$ フレームに対応する部分である第 8 信号と、を加算する処理を行って符号化前の前記第 $i - 1$ フレームに対応する信号である再構成信号を生成する

音信号ハイブリッドデコーダ。

[請求項3]

前記ブロック切替部は、

前記復号対象フレームが、前記第 i フレームの 1 フレーム後のフレームである第 $i + 1$ フレームであるとき、

前記第 $i + 1$ フレームを前記低遅延逆フィルタバンク処理及び窓処

理した信号のうちの、前記第 i フレームよりも 2 フレーム先行するフレームである第 $i - 2$ フレームに対応する部分である第 9 信号と、前記第 i フレームを前記低遅延逆フィルタバンク処理及び窓処理した信号の前記第 $i - 2$ フレームに対応する部分である第 10 信号と、前記第 $i - 2$ フレームを前記音声信号デコーダによって復号することで得られる第 11 信号に第 1 の窓処理を行なった信号のフレームの前半部分に相当する信号に、前記第 11 信号に前記第 1 の窓処理を行った信号のフレームの後半部分に相当する信号に畳み込み処理した信号を加算することで得られる第 12 信号に、当該第 12 信号を畳み込み処理した信号を連結し、窓処理を行った第 13 信号と、前記第 11 信号に前記第 1 の窓処理とは異なる第 2 の窓処理を行った信号のフレームの前半部分に相当する信号に、前記第 11 信号に前記第 2 の窓処理を行った信号のフレームの後半部分に相当する信号に畳み込み処理した信号を加算することで得られる第 14 信号に、当該第 14 信号を畳み込み処理して符号を反転させた信号を連結し、窓処理を行った第 15 信号と、を加算する処理を行って、符号化前の前記第 i フレームに対応する信号を生成する

請求項 2 に記載の音信号ハイブリッドデコーダ。

[請求項 4]

前記ブロック切替部は、

前記復号対象フレームが、前記第 i フレームの 2 フレーム後のフレームである第 $i + 2$ フレームであるとき、

前記 $i + 2$ フレームを前記低遅延逆フィルタバンク処理及び窓処理した信号の前記第 $i - 1$ フレームに対応する部分である第 16 信号と、前記第 $i + 1$ フレームを前記低遅延逆フィルタバンク処理及び窓処理した信号の前記第 $i - 1$ フレームに対応する部分である第 17 信号と、前記第 i フレームを前記低遅延逆フィルタバンク処理及び窓処理した信号の前記第 $i - 1$ フレームに対応する部分である第 18 信号と、前記第 $i - 3$ フレームを前記音声信号デコーダによって復号するこ

とで得られる第19信号に窓処理を行なった信号のフレームの前半部分に相当する信号に、前記第19信号に窓処理を行った信号のフレームの後半部分に相当する信号に畳み込み処理した信号を加算することで得られる第20信号に、当該第20信号を畳み込み処理した信号を連結し、窓処理を行った第21信号と、前記再構成信号に窓処理を行った信号のフレームの前半部分に相当する信号に、前記再構成信号に窓処理を行った信号のフレームの後半部分に相当する信号に畳み込み処理した信号を加算することで得られる第22信号に、当該第22信号を畳み込み処理して符号を反転させた信号を連結し、窓処理を行った第23信号と、を加算する処理を行って、符号化前の前記第*i*+1フレームに対応する信号を生成する

請求項3に記載の音信号ハイブリッドデコーダ。

[請求項5]

低遅延フィルタバンクを用いた音響符号化処理で符号化された音響フレームと、線形予測係数を用いた音声符号化処理で符号化された音声フレームとが含まれるビットストリームを復号する音信号ハイブリッドデコーダであって、

前記音響フレームを低遅延逆フィルタバンク処理を用いて復号する低遅延変換デコーダと、

TCX (Transform Coded Excitation) 方式によって符号化された前記音声フレームを復号するTCXデコーダと、

前記ビットストリームのうちの復号対象フレームが前記音響フレームである場合、当該復号対象フレームを前記低遅延変換デコーダによって復号し、前記復号対象フレームが前記音声フレームである場合、当該復号対象フレームを前記音声信号デコーダによって復号する制御を行うブロック切替部とを備え、

前記復号対象フレームが、前記音響フレームから前記音声フレームに切り替わった最初の前記音声フレームであって、過渡信号が符号化

されたフレームである第 i フレームであるとき、

前記第 i フレームには、前記第 i フレームよりも 1 フレーム先行するフレームである第 $i - 1$ フレームの符号化前の信号を用いて生成された第 1 信号が符号化された状態で含まれ、

前記ブロック切替部は、

(1)

前記第 i フレームよりも 2 フレーム先行するフレームである第 $i - 2$ フレームを前記低遅延変換デコーダによって復号することで得られる、前記第 i フレームよりも 3 フレーム先行するフレームである第 $i - 3$ フレームの再構成された信号を窓処理した信号である第 2 信号のフレームの前半部分に相当する信号に、前記第 2 信号のフレームの後半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第 i フレームを前記音声信号デコーダによって復号することで得られる、前記第 1 信号に窓処理を行った信号と、前記第 $i - 1$ フレームを前記低遅延逆フィルタバンク処理及び窓処理した信号の前記第 $i - 3$ フレームに対応する部分である第 3 信号のフレームの前半部分の信号と、を加算する処理を行って符号化前の前記第 $i - 1$ フレームの前半部分に対応する信号を生成し、

前記第 2 信号のフレームの後半部分に相当する信号に、前記第 2 信号のフレームの前半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第 1 信号に畳み込み処理及び窓処理を行った信号と、前記第 3 信号のフレームの後半部分に相当する信号と、を加算する処理を行って符号化前の前記第 $i - 1$ フレームの後半部分に対応する信号を生成する、または

(2)

前記第 2 信号のフレームの前半部分に相当する信号に、前記第 2 信号のフレームの後半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第 1 信号に畳み込み処理及び窓処

理を行った信号と、前記第3信号のフレームの前半部分に相当する信号と、を加算する処理を行って符号化前の前記第 $i - 1$ フレームの前半部分に対応する信号を生成し、

前記第2信号のフレームの後半部分に相当する信号に、前記第2信号のフレームの前半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第1信号に窓処理を行った信号と、前記第3信号のフレームの後半部分に相当する信号と、を加算する処理を行って符号化前の前記第 $i - 1$ フレームの後半部分に対応する信号を生成する

音信号ハイブリッドデコーダ。

[請求項6]

前記低遅延変換デコーダは、前記音響フレーム及び当該音響フレームに時間的に連続して先行する3つのフレームのそれぞれについて低遅延逆フィルタバンク処理及び窓処理を行った信号のそれぞれを重複加算処理することによって、当該音響フレームを復号するAAC-ELD (Advanced Audio Coding - Enhanced Low Delay) デコーダである

請求項1～5のいずれか1項に記載の音信号ハイブリッドデコーダ。

[請求項7]

前記音声信号デコーダは、ACELP (Algebraic Code Excited Linear Prediction) 係数を用いて符号化された前記音声フレームを復号するACELPデコーダである

請求項1～4のいずれか1項に記載の音信号ハイブリッドデコーダ。

[請求項8]

前記音声信号デコーダは、TCX方式によって符号化された前記音声フレームを復号するTCXデコーダである

請求項1～4のいずれか1項に記載の音信号ハイブリッドデコーダ。

[請求項9] さらに、前記復号対象フレームとともに符号化された合成エラー情報を復号する合成エラー補償装置を備え、

前記合成エラー情報は、前記ビットストリームが符号化される前の信号と、前記ビットストリームを復号した信号との差分を表す情報であり、

前記合成エラー補償装置は、前記ブロック切替部が生成した前記符号化前の前記第 $i - 1$ フレームの信号、前記ブロック切替部が生成した前記符号化前の前記第 i フレームの信号、または前記ブロック切替部が生成した前記符号化前の前記第 $i + 1$ フレームの信号を、復号した前記合成エラー情報を用いて修正する

請求項 1 ~ 8 のいずれか 1 項に記載の音信号ハイブリッドデコーダ。

[請求項10] 音信号の音響特性を分析し、前記音信号に含まれるフレームが音響信号であるか音声信号であるかを判断する信号分類部と、

低遅延フィルタバンクを用いて前記フレームを符号化する低遅延変換エンコーダと、

前記フレームの線形予測係数を算出することによって当該フレームを符号化する音声信号エンコーダと、

前記信号分類部が前記音響信号であると判断した符号化対象フレームを前記低遅延変換エンコーダによって符号化し、前記信号分類部が前記音声信号であると判断した前記符号化対象フレームを前記音声信号エンコーダによって符号化する制御を行うブロック切替部とを備え、

前記ブロック切替部は、

(1) 前記符号化対象フレームが、前記信号分類部が前記音声信号であると判断したフレームである第 $i - 1$ フレームの 1 フレーム後のフレームであって、前記信号分類部が前記音響信号であると判断したフレームである第 i フレームであるとき、

前記第 $i - 1$ フレームの前半部分に相当する信号を窓処理した信号に前記第 $i - 1$ フレームの後半部分に相当する信号を窓処理して畳み込み処理した信号を加算した信号と、前記第 i フレームとを前記音声信号エンコーダによって符号化する、または

(2) 前記第 $i - 1$ フレームの後半部分に相当する信号を窓処理した信号に前記第 $i - 1$ フレームの前半部分に相当する信号を窓処理して畳み込み処理した信号を加算した信号と、前記第 i フレームとを前記音声信号エンコーダによって符号化する

音信号ハイブリッドエンコーダ。

[請求項11]

音信号の音響特性を分析し、前記音信号に含まれるフレームが音響信号であるか音声信号であるかを判断する信号分類部と、

低遅延フィルタバンクを用いて前記フレームを符号化する低遅延変換エンコーダと、

前記フレームの線形予測係数の残差をMDC T (Modified Discrete Cosine Transform) 処理したTCX方式によって前記フレームを符号化するTCXエンコーダと、

前記信号分類部が前記音響信号であると判断した符号化対象フレームを前記低遅延変換エンコーダによって符号化し、前記信号分類部が前記音声信号であると判断した前記符号化対象フレームを前記音声信号エンコーダによって符号化する制御を行うブロック切替部とを備え、

前記ブロック切替部は、

前記符号化対象フレームである第 i フレームが、前記信号分類部が前記音響信号であり、なおかつエネルギーが急激に変化する過渡信号であると判断したフレームであるとき、

(1) 前記第 i フレームの1フレーム前のフレームである第 $i - 1$ フレームの前半部分に相当する信号を窓処理した信号に前記第 $i - 1$ フレームの後半部分に相当する信号を窓処理して畳み込み処理した信

号を加算した信号と、前記第 i フレームとを前記音声信号エンコーダによって符号化する、または

(2) 前記第 $i - 1$ フレームの後半部分に相当する信号を窓処理した信号に前記第 $i - 1$ フレームの前半部分に相当する信号を窓処理して畳み込み処理した信号を加算した信号と、前記第 i フレームとを前記音声信号エンコーダによって符号化する

音信号ハイブリッドエンコーダ。

[請求項12] 前記低遅延変換エンコーダは、前記フレームと、当該フレームに時間的に連続して先行する3つのフレームとを連結した拡張フレームについて窓処理及び低遅延フィルタバンク処理をすることによって、前記フレームを符号化するAAC-ELDエンコーダである

請求項10または11に記載の音信号ハイブリッドエンコーダ。

[請求項13] 前記音声信号エンコーダは、ACELP係数を生成することによって前記フレームを符号化するACELPエンコーダである

請求項10～12のいずれか1項に記載の音信号ハイブリッドエンコーダ。

[請求項14] 前記音声信号エンコーダは、前記線形予測係数の残差をMDCT処理して前記フレームを符号化するTCXエンコーダである

請求項10～12のいずれか1項に記載の音信号ハイブリッドエンコーダ。

[請求項15] さらに、

符号化した前記音信号を復号するローカルデコーダと、

前記音信号と、前記ローカルデコーダが復号した前記音信号との差分である合成エラー情報を符号化するローカルエンコーダとを備える

請求項10～14のいずれか1項に記載の音信号ハイブリッドエンコーダ。

[請求項16] 低遅延フィルタバンクを用いた音響符号化処理で符号化された音響フレームと、線形予測係数を用いた音声符号化処理で符号化された音

声フレームとが含まれるビットストリームを復号する音信号復号方法であって、

前記音響フレームを低遅延逆フィルタバンク処理を用いて復号する低遅延変換デコードステップと、

前記音声フレームを復号する音声信号デコードステップと、

前記ビットストリームのうちの復号対象フレームが前記音響フレームである場合、当該復号対象フレームを前記低遅延変換デコードステップによって復号し、前記復号対象フレームが前記音声フレームである場合、当該復号対象フレームを前記音声信号デコードステップによって復号する制御を行う制御ステップとを含み、

前記復号対象フレームが、前記音響フレームから前記音声フレームに切り替わった最初の前記音声フレームである第 i フレームであるとき、

前記第 i フレームには、前記第 i フレームよりも 1 フレーム先行するフレームである第 $i - 1$ フレームの符号化前の信号を用いて生成された第 1 信号が符号化された状態で含まれ、

前記制御ステップでは、

(1)

前記第 i フレームよりも 2 フレーム先行するフレームである第 $i - 2$ フレームを前記低遅延変換デコードステップによって復号することで得られる、前記第 i フレームよりも 3 フレーム先行するフレームである第 $i - 3$ フレームの再構成された信号を窓処理した信号である第 2 信号のフレームの前半部分に相当する信号に、前記第 2 信号のフレームの後半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第 i フレームを前記音声信号デコードステップによって復号することで得られる、前記第 1 信号に窓処理を行った信号と、前記第 $i - 1$ フレームを前記低遅延逆フィルタバンク処理及び窓処理した信号の前記第 $i - 3$ フレームに対応する部分である第

3 信号のフレームの前半部分の信号と、を加算する処理を行って符号化前の前記第 $i - 1$ フレームの前半部分に対応する信号を生成し、

前記第 2 信号のフレームの後半部分に相当する信号に、前記第 2 信号のフレームの前半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第 1 信号に畳み込み処理及び窓処理を行った信号と、前記第 3 信号のフレームの後半部分に相当する信号と、を加算する処理を行って符号化前の前記第 $i - 1$ フレームの後半部分に対応する信号を生成する、または

(2)

前記第 2 信号のフレームの前半部分に相当する信号に、前記第 2 信号のフレームの後半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第 1 信号に畳み込み処理及び窓処理を行った信号と、前記第 3 信号のフレームの前半部分に相当する信号と、を加算する処理を行って符号化前の前記第 $i - 1$ フレームの前半部分に対応する信号を生成し、

前記第 2 信号のフレームの後半部分に相当する信号に、前記第 2 信号のフレームの前半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第 1 信号に窓処理を行った信号と、前記第 3 信号のフレームの後半部分に相当する信号と、を加算する処理を行って符号化前の前記第 $i - 1$ フレームの後半部分に対応する信号を生成する

音信号復号方法。

[請求項17]

低遅延フィルタバンクを用いた音響符号化処理で符号化された音響フレームと、線形予測係数を用いた音声符号化処理で符号化された音声フレームとが含まれるビットストリームを復号する音信号復号方法であって、

前記音響フレームを低遅延逆フィルタバンク処理によって復号する低遅延変換デコードステップと、

前記音声フレームを復号する音声信号デコードステップと、
前記ビットストリームのうちの復号対象フレームが前記音響フレームである場合、当該復号対象フレームを前記低遅延変換デコードステップによって復号し、前記復号対象フレームが前記音声フレームである場合、当該復号対象フレームを前記音声信号デコードステップによって復号する制御を行う制御ステップとを含み、

前記制御ステップは、

前記復号対象フレームが、前記音声フレームから前記音響フレームに切り替わった最初の音響フレームである第 i フレームであるとき、

前記第 i フレームよりも 1 フレーム先行するフレームである第 $i-1$ フレームを前記音声信号デコードステップによって復号することで得られる信号を窓処理した第 4 信号に、当該第 4 信号を畳み込み処理した信号を加算し、窓処理を行った第 5 信号と、前記第 i フレームよりも 3 フレーム先行するフレームである第 $i-3$ フレームを前記音声信号デコードステップによって復号することで得られる信号を窓処理した第 6 信号に、当該第 6 信号を畳み込み処理した信号を加算し、窓処理を行った第 7 信号と、前記第 i フレームを前記低遅延逆フィルタバンク処理及び窓処理した信号の前記第 $i-3$ フレームに対応する部分である第 8 信号と、を加算する処理を行って符号化前の前記第 $i-1$ フレームに対応する信号である再構成信号を生成する

音信号復号方法。

[請求項18]

低遅延フィルタバンクを用いた音響符号化処理で符号化された音響フレームと、線形予測係数を用いた音声符号化処理で符号化された音声フレームとが含まれるビットストリームを復号する音信号復号方法であって、

前記音響フレームを低遅延逆フィルタバンク処理を用いて復号する低遅延変換デコードステップと、

TCX方式によって符号化された前記音声フレームを復号するTC

Xデコードステップと、

前記ビットストリームのうちの復号対象フレームが前記音響フレームである場合、当該復号対象フレームを前記低遅延変換デコードステップによって復号し、前記復号対象フレームが前記音声フレームである場合、当該復号対象フレームを前記音声信号デコードステップによって復号する制御を行う制御ステップとを含み、

前記復号対象フレームが、前記音響フレームから前記音声フレームに切り替わった最初の前記音声フレームであって、エネルギーが急激に変化する過渡信号が符号化されたフレームである第 i フレームであるとき、

前記第 i フレームには、前記第 i フレームよりも1フレーム先行するフレームである第 $i - 1$ フレームの符号化前の信号を用いて生成された第1信号が符号化された状態で含まれ、

前記制御ステップでは、

(1)

前記第 i フレームよりも2フレーム先行するフレームである第 $i - 2$ フレームを前記低遅延変換デコードステップによって復号することで得られる、前記第 i フレームよりも3フレーム先行するフレームである第 $i - 3$ フレームの再構成された信号を窓処理した信号である第2信号のフレームの前半部分に相当する信号に、前記第2信号のフレームの後半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第 i フレームを前記音声信号デコードステップによって復号することで得られる、前記第1信号に窓処理を行った信号と、前記第 $i - 1$ フレームを前記低遅延逆フィルタバンク処理及び窓処理した信号の前記第 $i - 3$ フレームに対応する部分である第3信号のフレームの前半部分の信号と、を加算する処理を行って符号化前の前記第 $i - 1$ フレームの前半部分に対応する信号を生成し、

前記第2信号のフレームの後半部分に相当する信号に、前記第2信

号のフレームの前半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第 1 信号に畳み込み処理及び窓処理を行った信号と、前記第 3 信号のフレームの後半部分に相当する信号と、を加算する処理を行って符号化前の前記第 $i - 1$ フレームの後半部分に対応する信号を生成する、または

(2)

前記第 2 信号のフレームの前半部分に相当する信号に、前記第 2 信号のフレームの後半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第 1 信号に畳み込み処理及び窓処理を行った信号と、前記第 3 信号のフレームの前半部分に相当する信号と、を加算する処理を行って符号化前の前記第 $i - 1$ フレームの前半部分に対応する信号を生成し、

前記第 2 信号のフレームの後半部分に相当する信号に、前記第 2 信号のフレームの前半部分に相当する信号を畳み込み処理した信号を加算して窓処理を行った信号と、前記第 1 信号に窓処理を行った信号と、前記第 3 信号のフレームの後半部分に相当する信号と、を加算する処理を行って符号化前の前記第 $i - 1$ フレームの後半部分に対応する信号を生成する

音信号復号方法。

[請求項19]

音信号の音響特性を分析し、前記音信号に含まれるフレームが音響信号であるか音声信号であるかを判断する判断ステップと、

低遅延フィルタバンクを用いて前記フレームを符号化する低遅延変換エンコードステップと、

前記フレームの線形予測係数を算出することによって当該フレームを符号化する音声信号エンコードステップと、

前記判断ステップにおいて前記音響信号であると判断した符号化対象フレームを前記低遅延変換エンコードステップによって符号化し、前記判断ステップにおいて前記音声信号であると判断した前記符号化

対象フレームを前記音声信号エンコードステップによって符号化する制御を行う制御ステップとを含み、

前記制御ステップでは、

(1) 前記符号化対象フレームが、前記判断ステップにおいて前記音声信号であると判断したフレームである第 $i-1$ フレームの 1 フレーム後のフレームであって、前記判断ステップにおいて前記音響信号であると判断したフレームである第 i フレームであるとき、

前記第 $i-1$ フレームの前半部分に相当する信号を窓処理した信号に前記第 $i-1$ フレームの後半部分に相当する信号を窓処理して畳み込み処理した信号を加算した信号と、前記第 i フレームとを前記音声信号エンコードステップによって符号化する、または

(2) 前記第 $i-1$ フレームの後半部分に相当する信号を窓処理した信号に前記第 $i-1$ フレームの前半部分に相当する信号を窓処理して畳み込み処理した信号を加算した信号と、前記第 i フレームとを前記音声信号エンコードステップによって符号化する

音信号符号化方法。

[請求項20]

音信号の音響特性を分析し、前記音信号に含まれるフレームが音響信号であるか音声信号であるかを判断する判断ステップと、

低遅延フィルタバンクを用いて前記フレームを符号化する低遅延変換エンコードステップと、

前記フレームの線形予測係数の残差をMDC T処理したTCX方式によって前記フレームを符号化するTCXエンコードステップと、

前記判断ステップにおいて前記音響信号であると判断した符号化対象フレームを前記低遅延変換エンコードステップによって符号化し、前記判断ステップにおいて前記音声信号であると判断した前記符号化対象フレームを前記音声信号エンコードステップによって符号化する制御を行う制御ステップとを含み、

前記制御ステップでは、

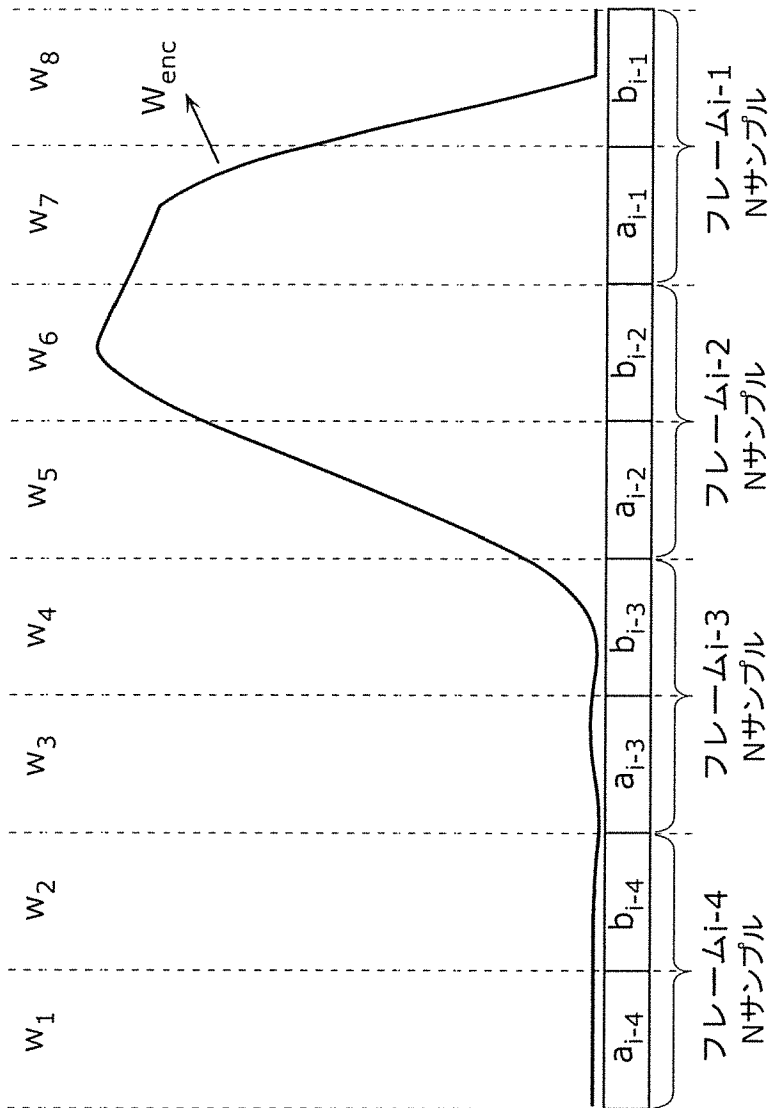
前記符号化対象フレームである第 i フレームが、前記判断ステップにおいて前記音響信号であり、なおかつエネルギーが急激に変化する過渡信号であると判断したフレームであるとき、

(1) 前記第 i フレームの 1 フレーム前のフレームである第 $i - 1$ フレームの前半部分に相当する信号を窓処理した信号に前記第 $i - 1$ フレームの後半部分に相当する信号を窓処理して畳み込み処理した信号を加算した信号と、前記第 i フレームとを前記音声信号エンコードステップによって符号化する、または

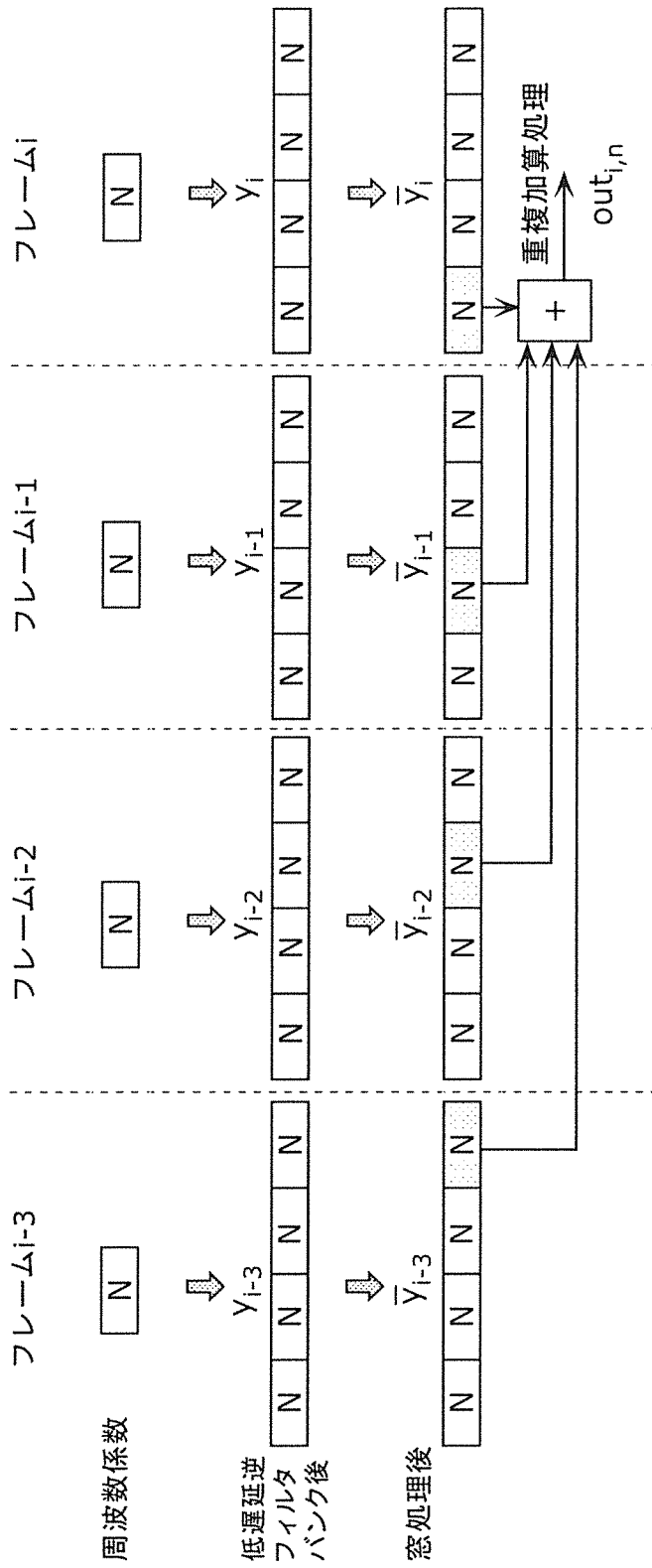
(2) 前記第 $i - 1$ フレームの後半部分に相当する信号を窓処理した信号に前記第 $i - 1$ フレームの前半部分に相当する信号を窓処理して畳み込み処理した信号を加算した信号と、前記第 i フレームとを前記音声信号エンコードステップによって符号化する

音信号符号化方法。

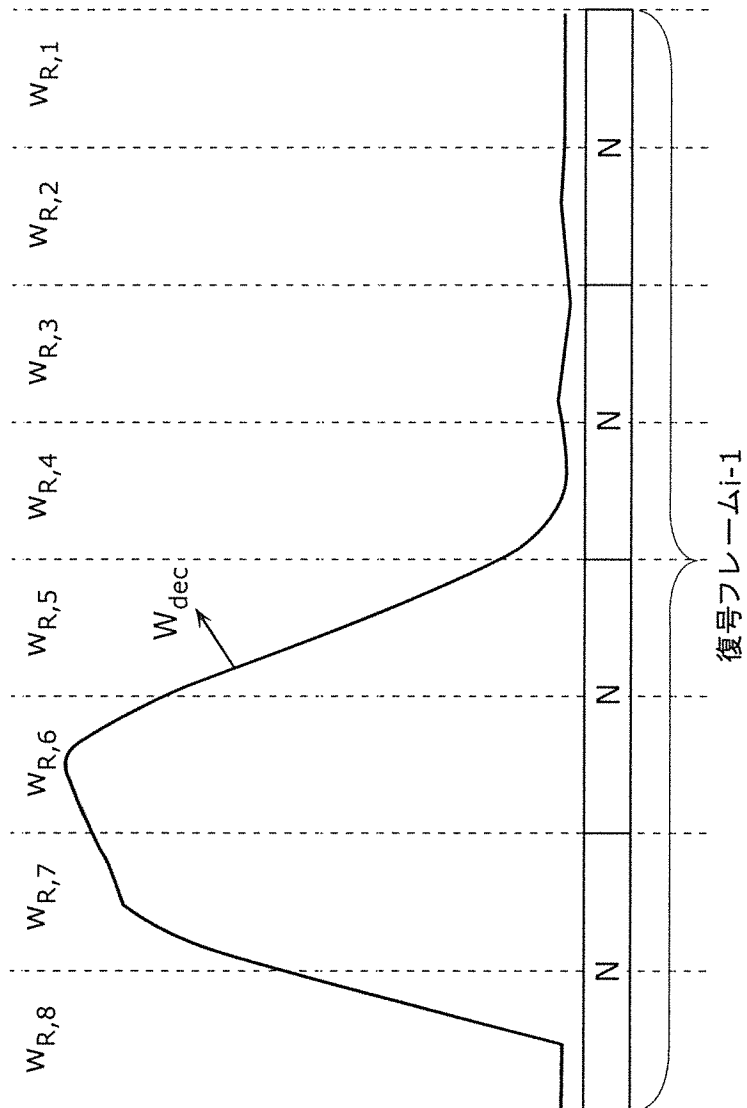
[図1]



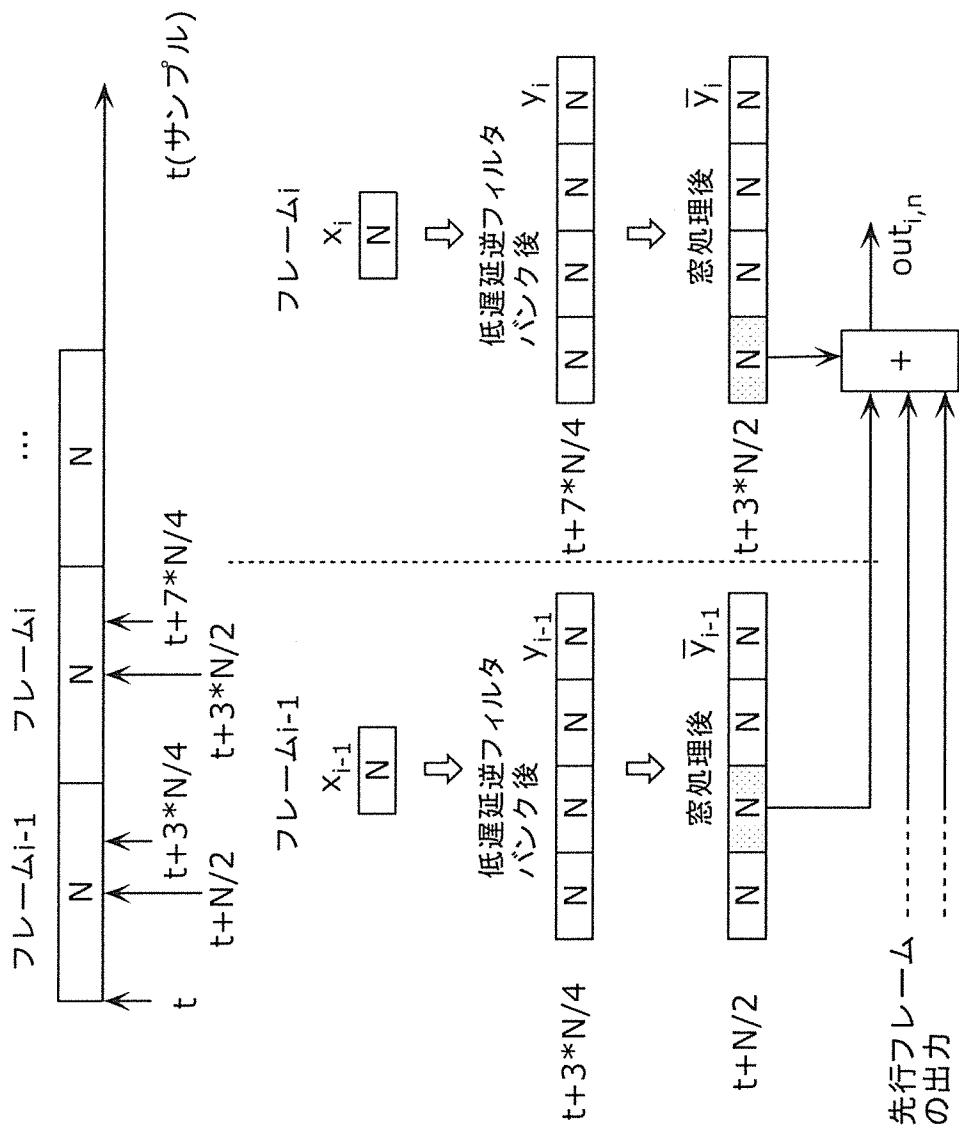
[図2]



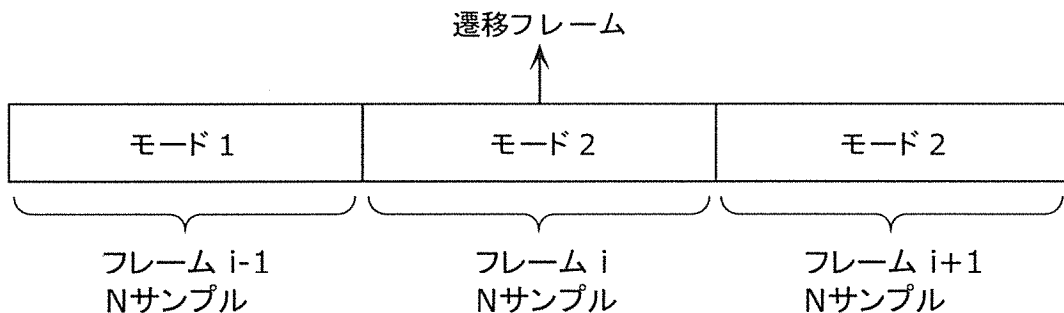
[図3]



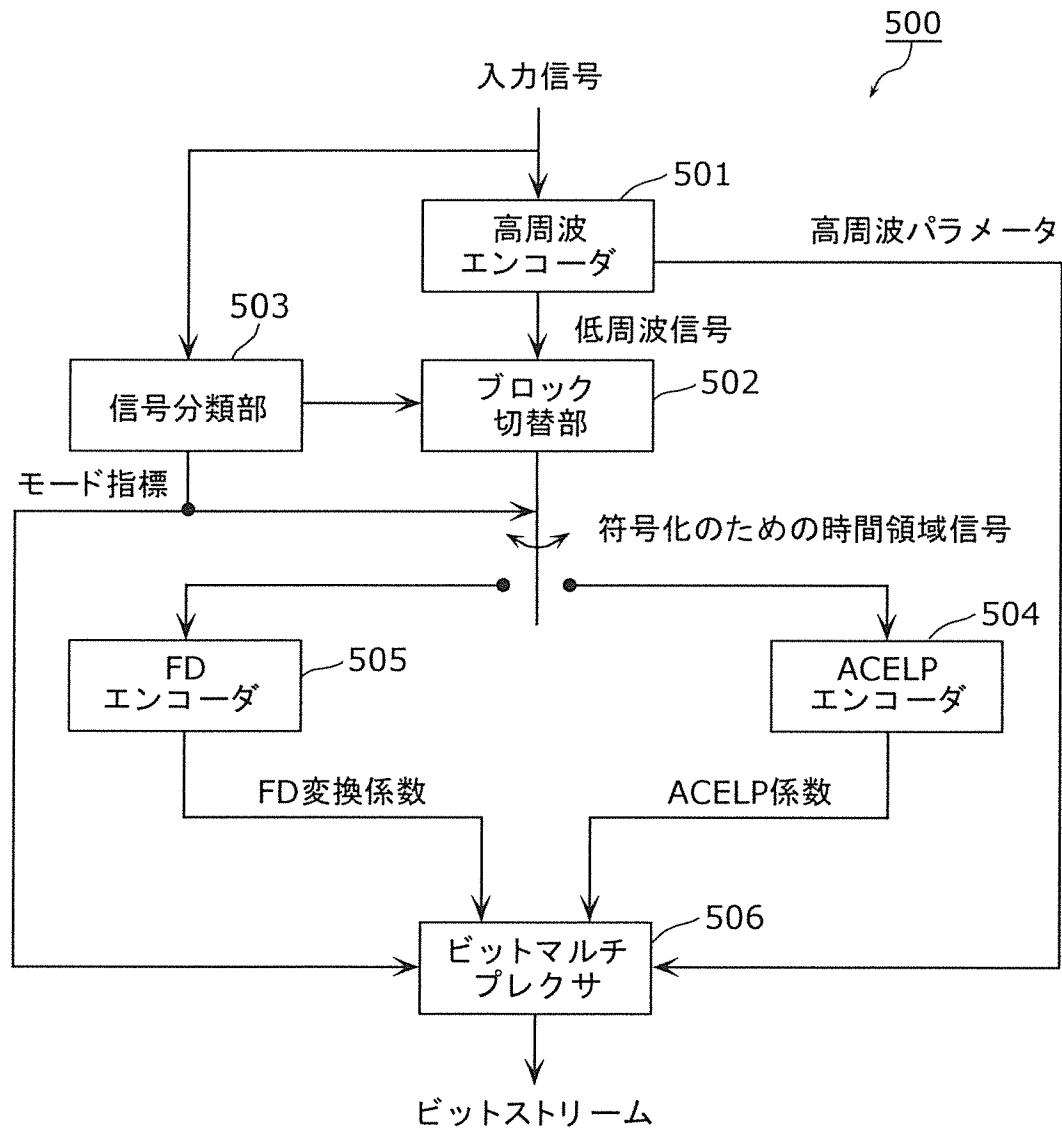
[図4]



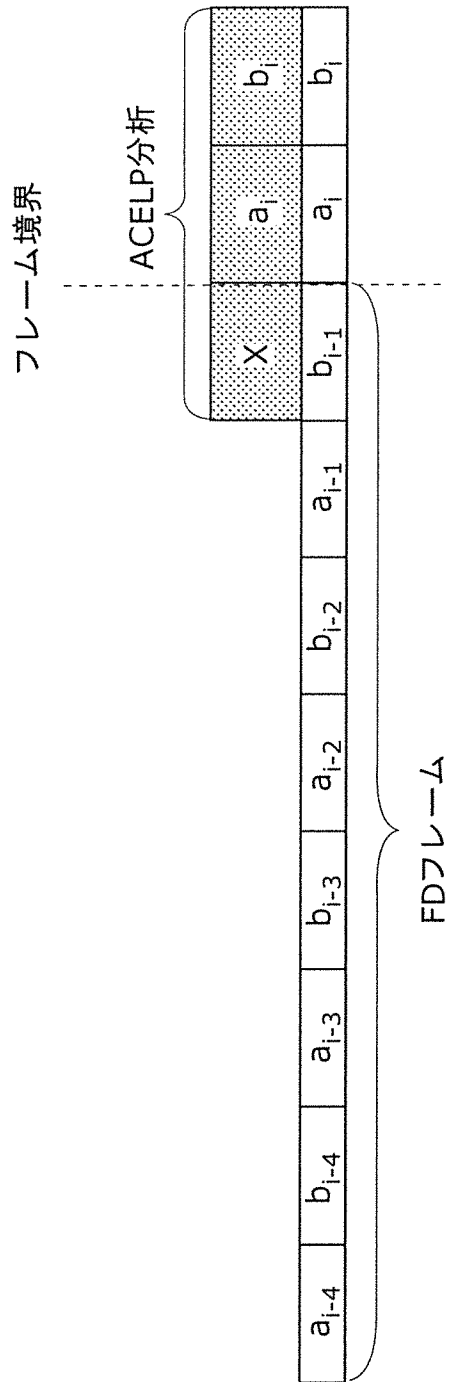
[図5]



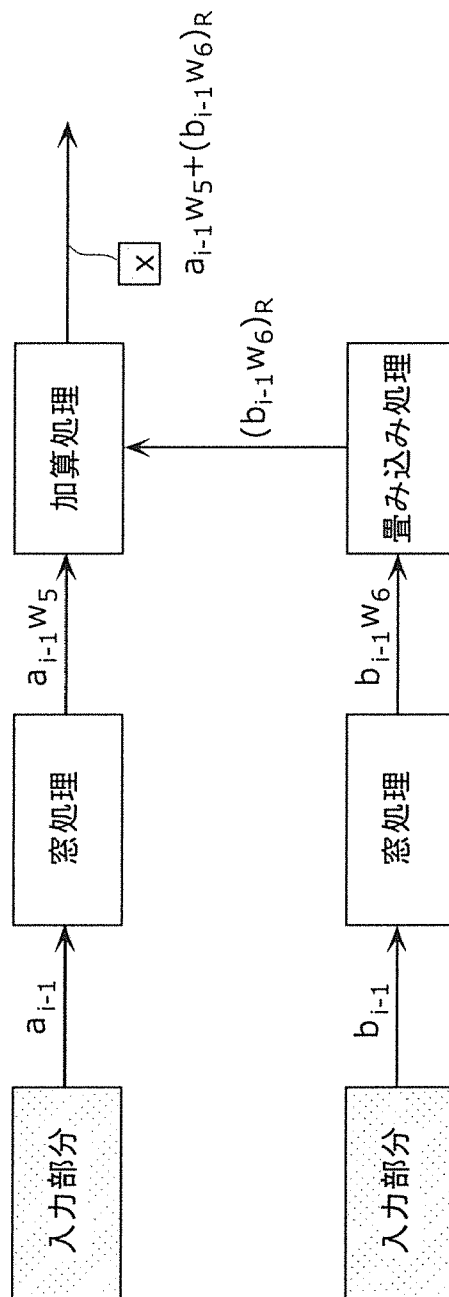
[図6]



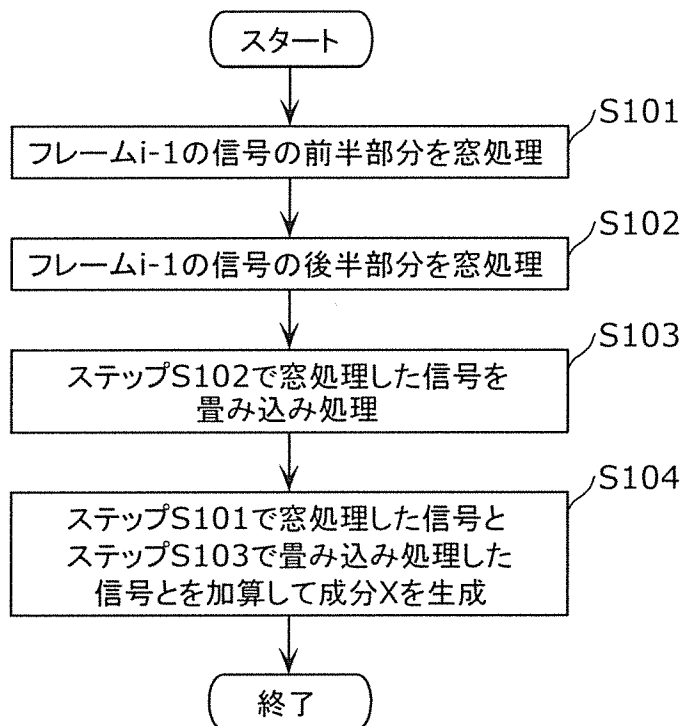
[図7]



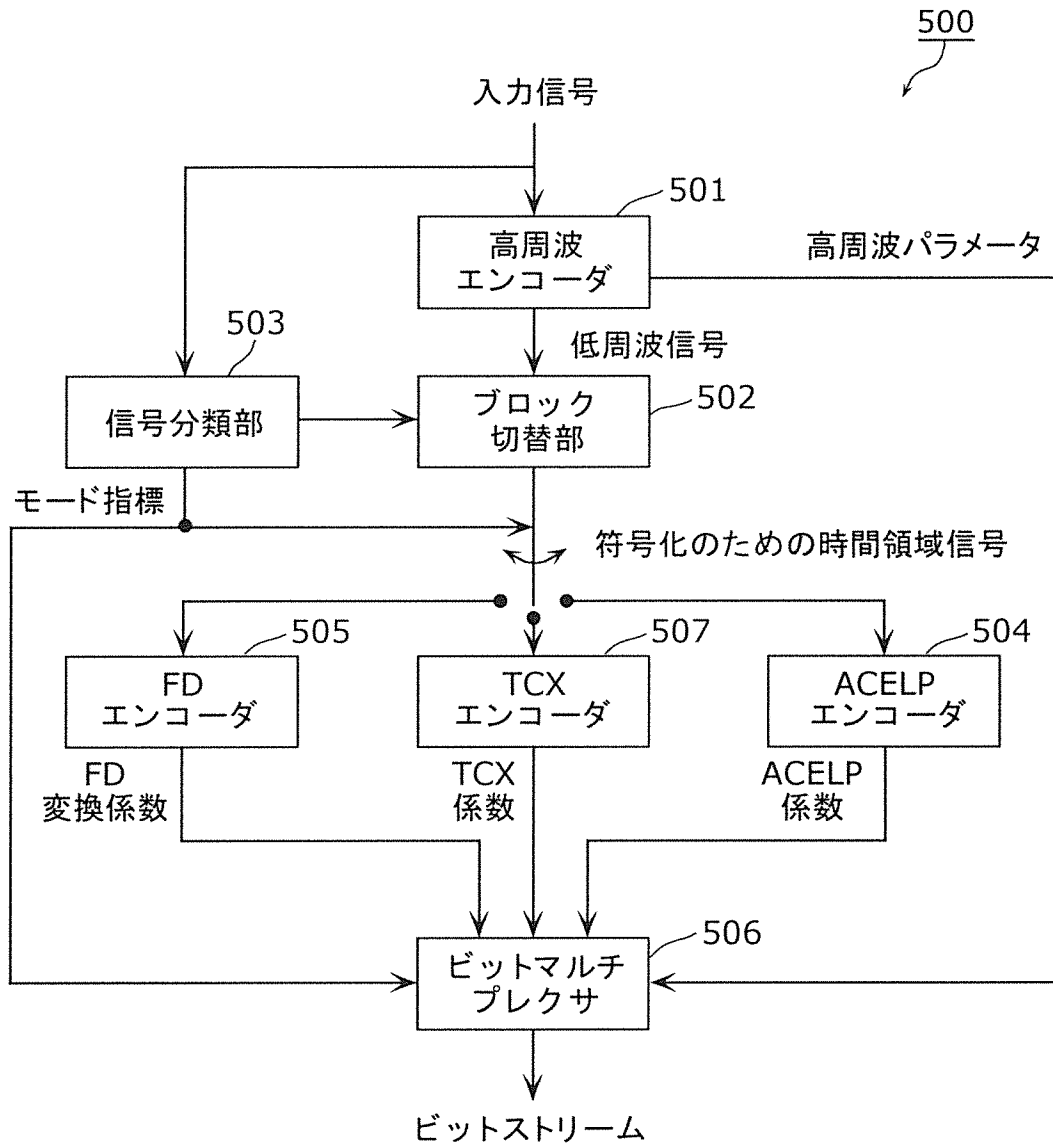
[図8A]



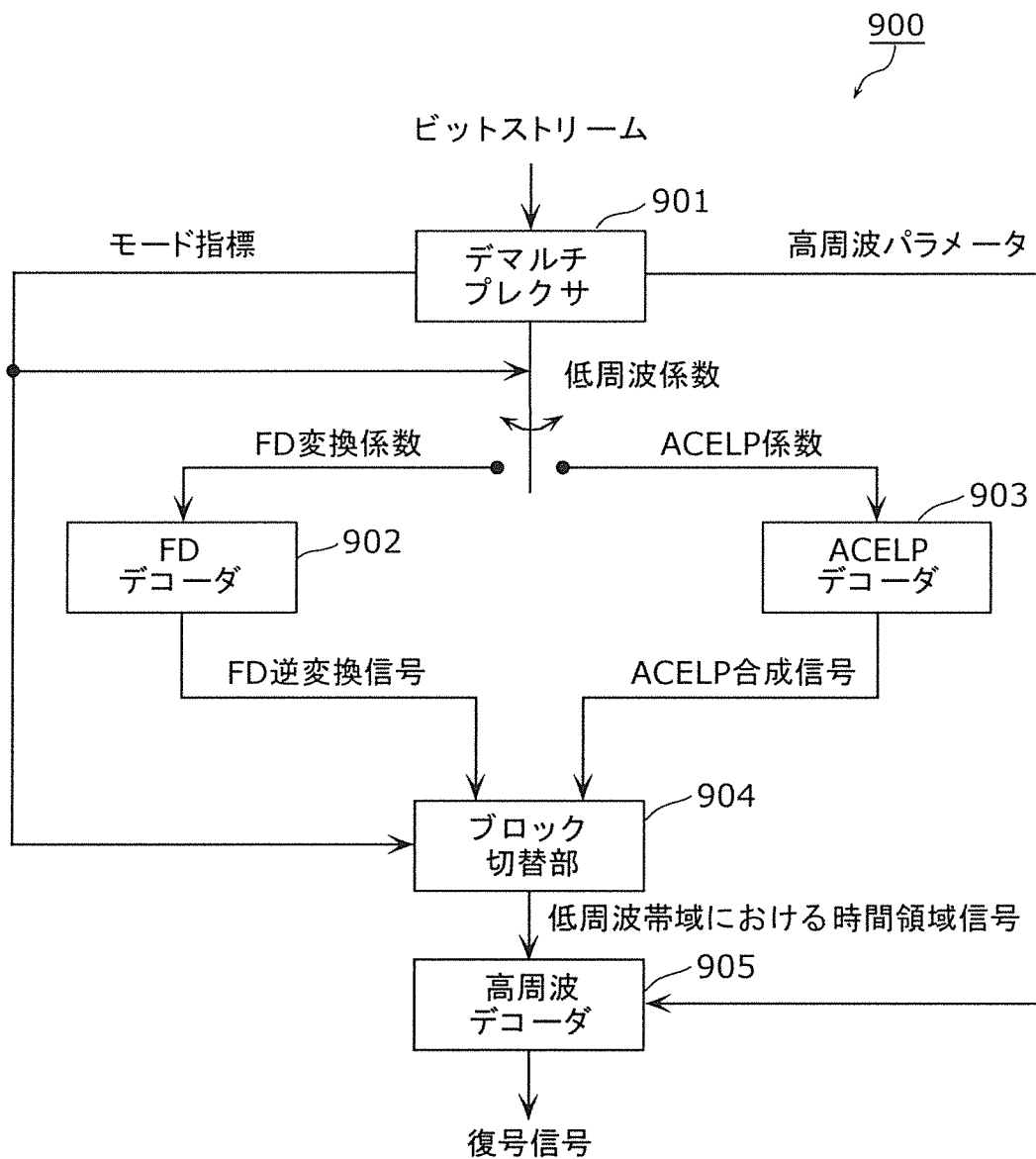
[図8B]



[図9]

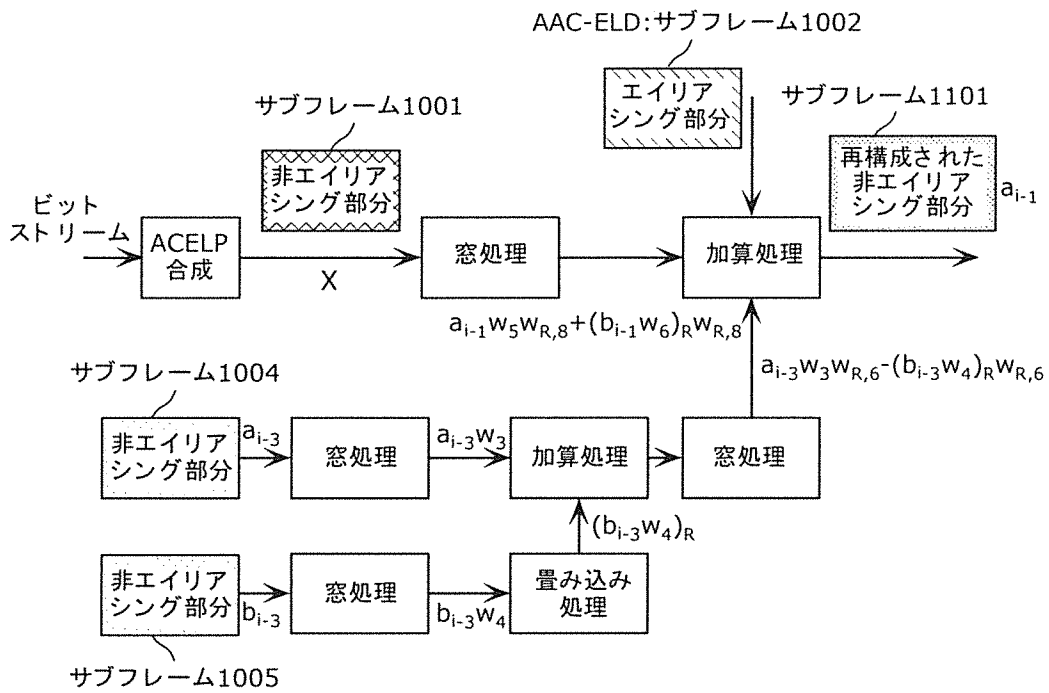


[図10]

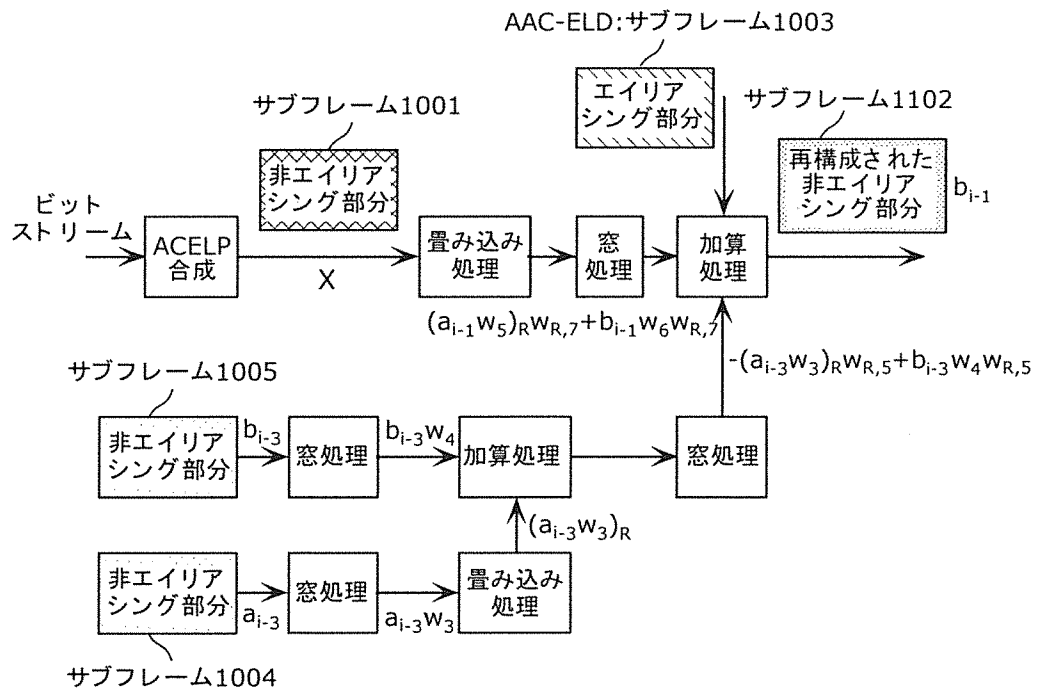


[図12A]

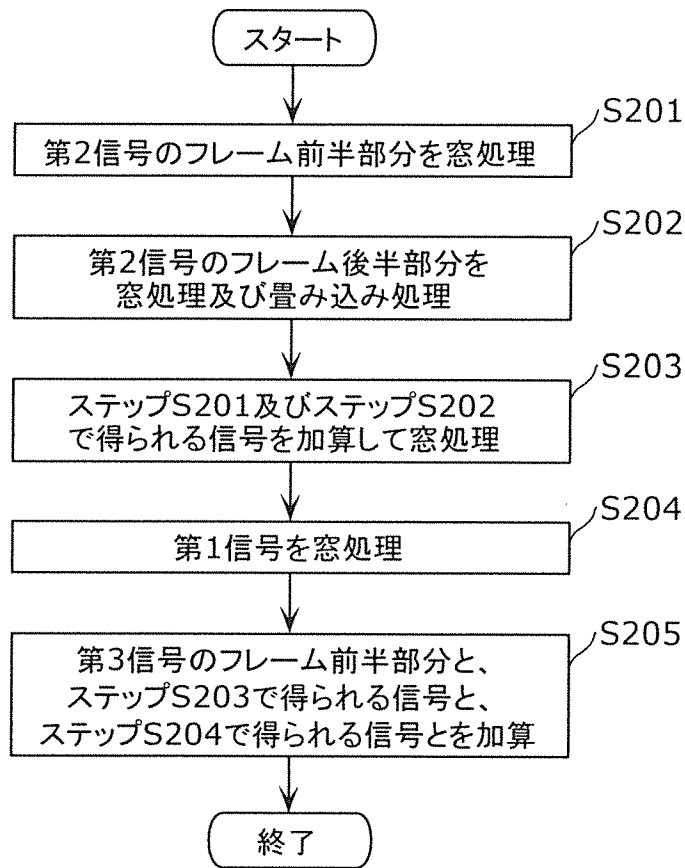
(a)



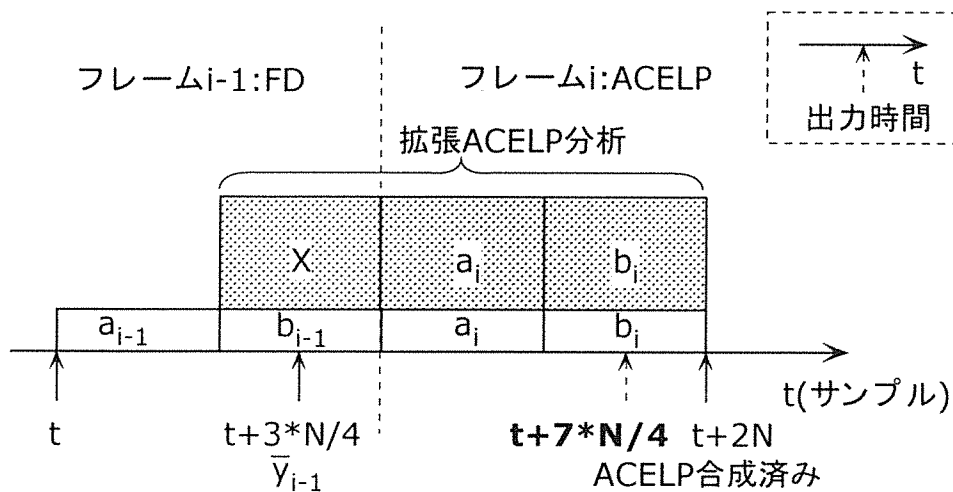
(b)



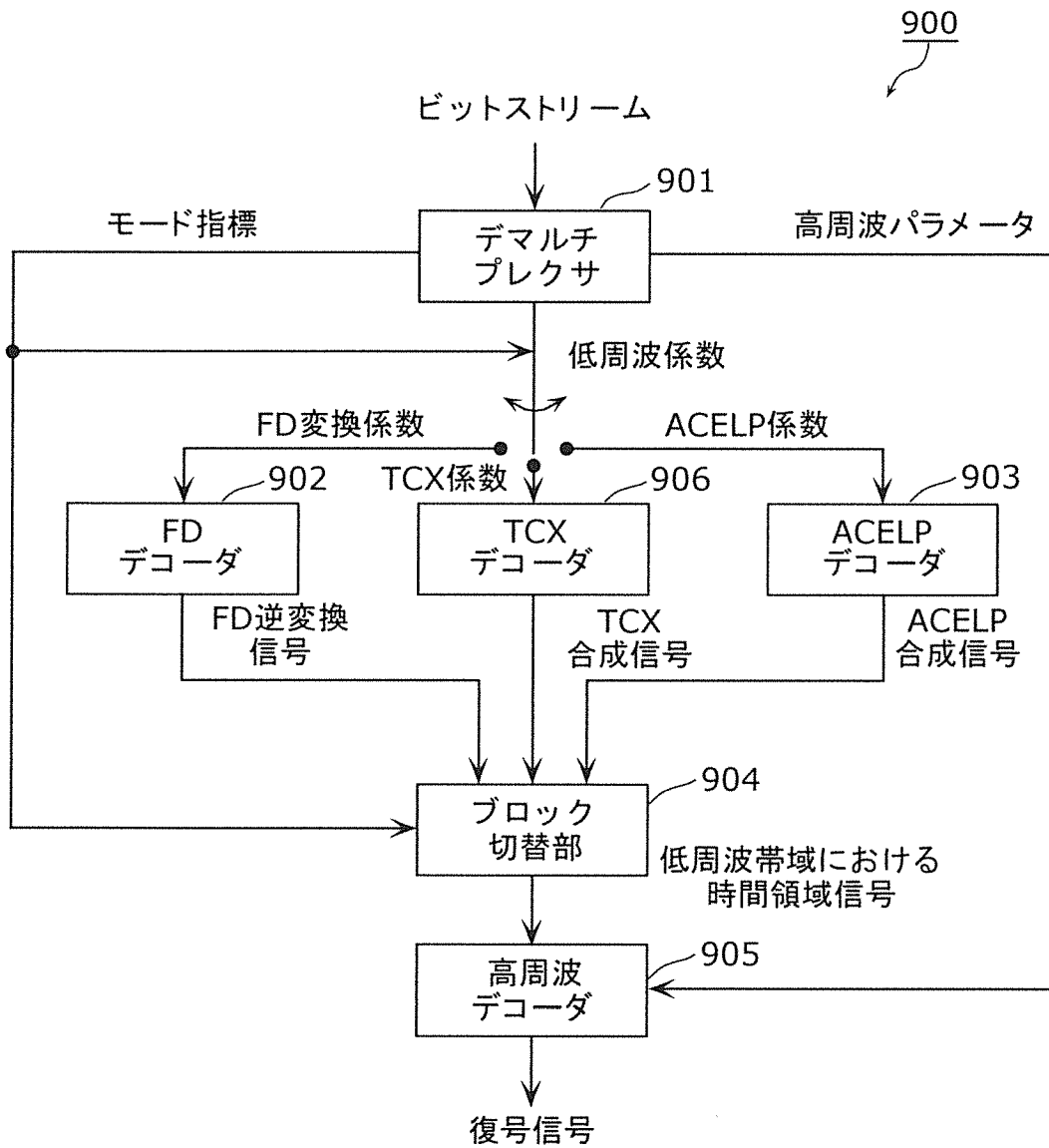
[図12B]



[図13]

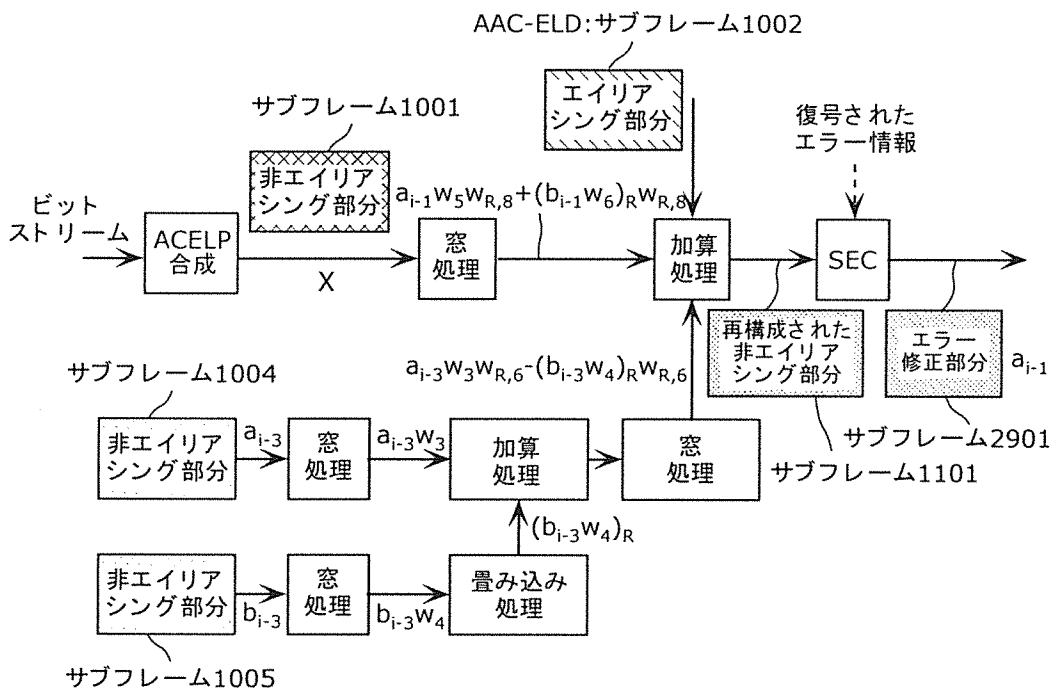


[図14]

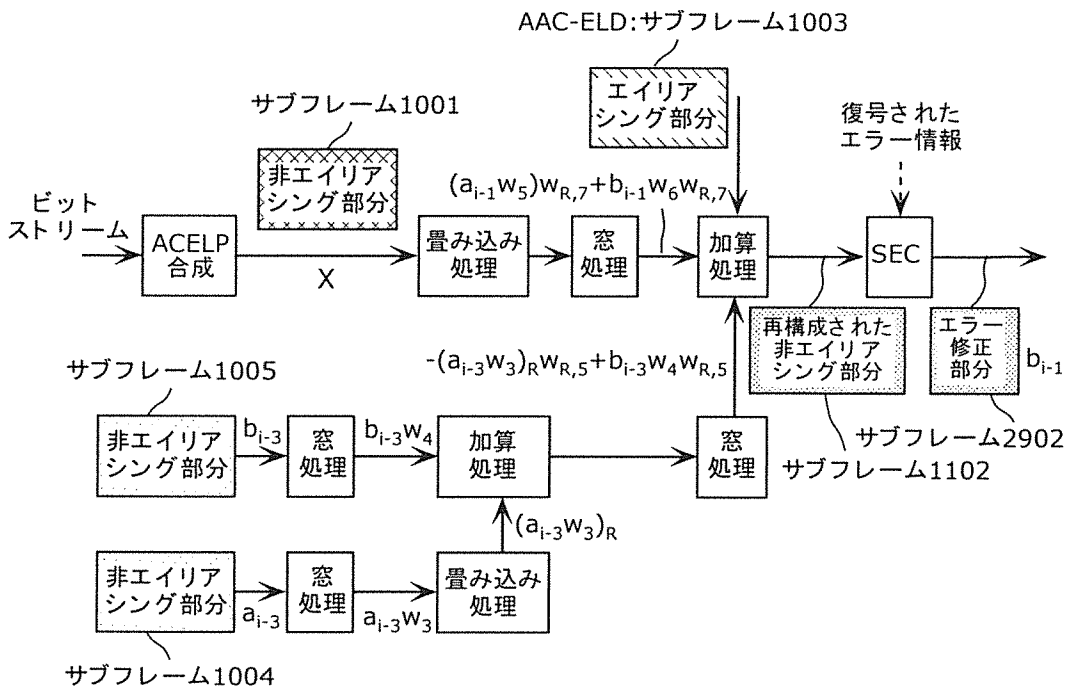


[図15]

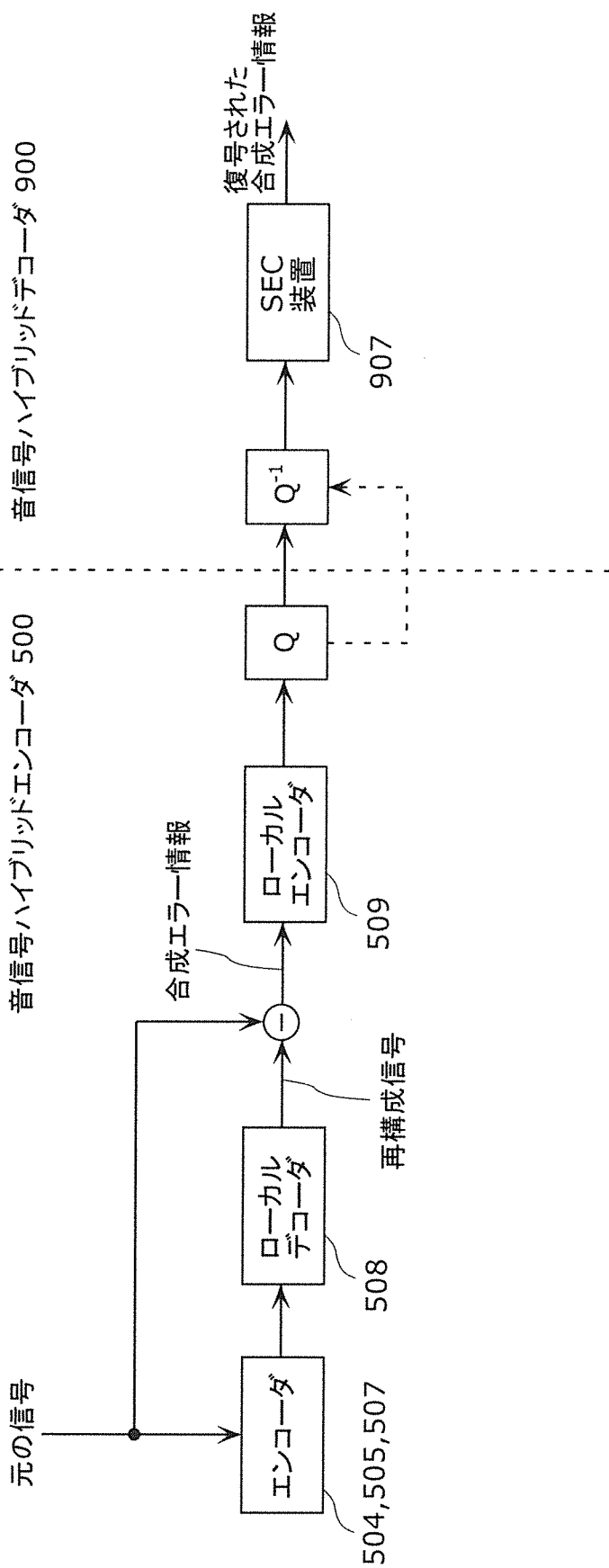
(a)



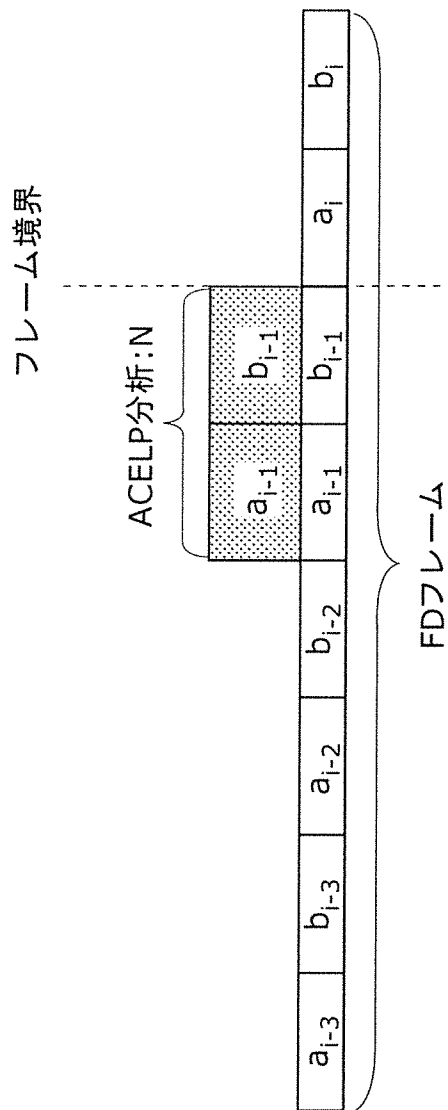
(b)



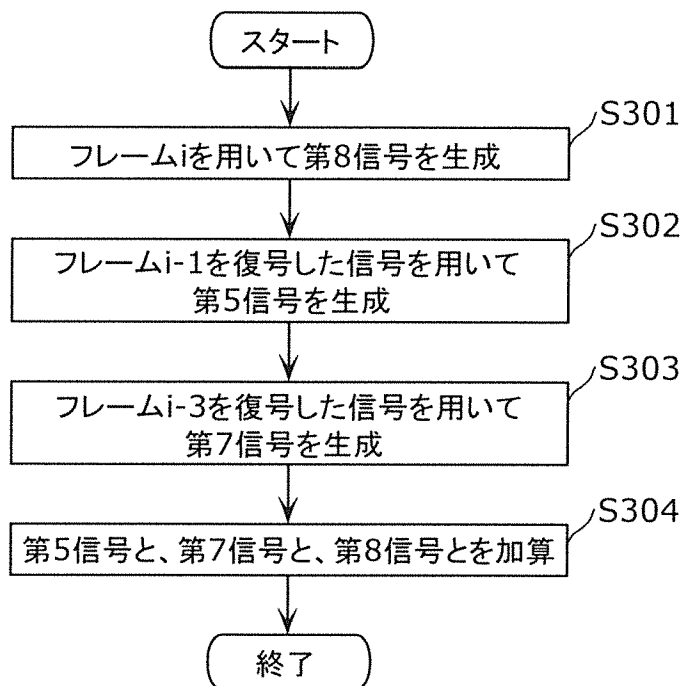
[図16]



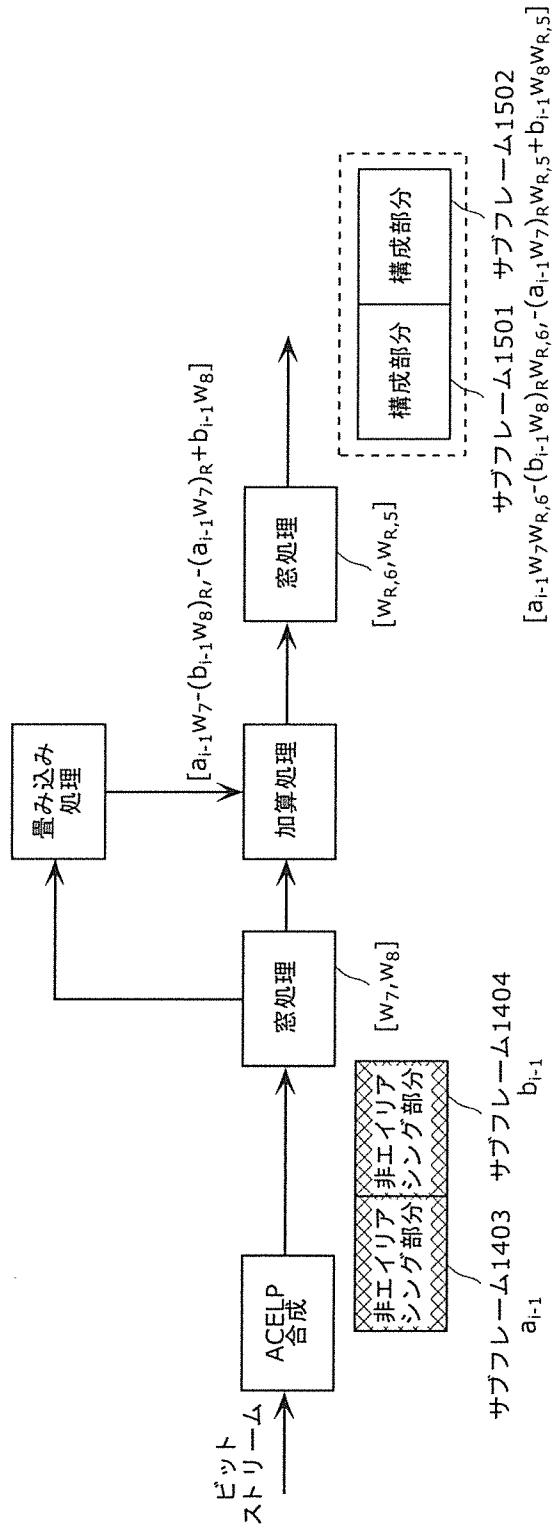
[図17]



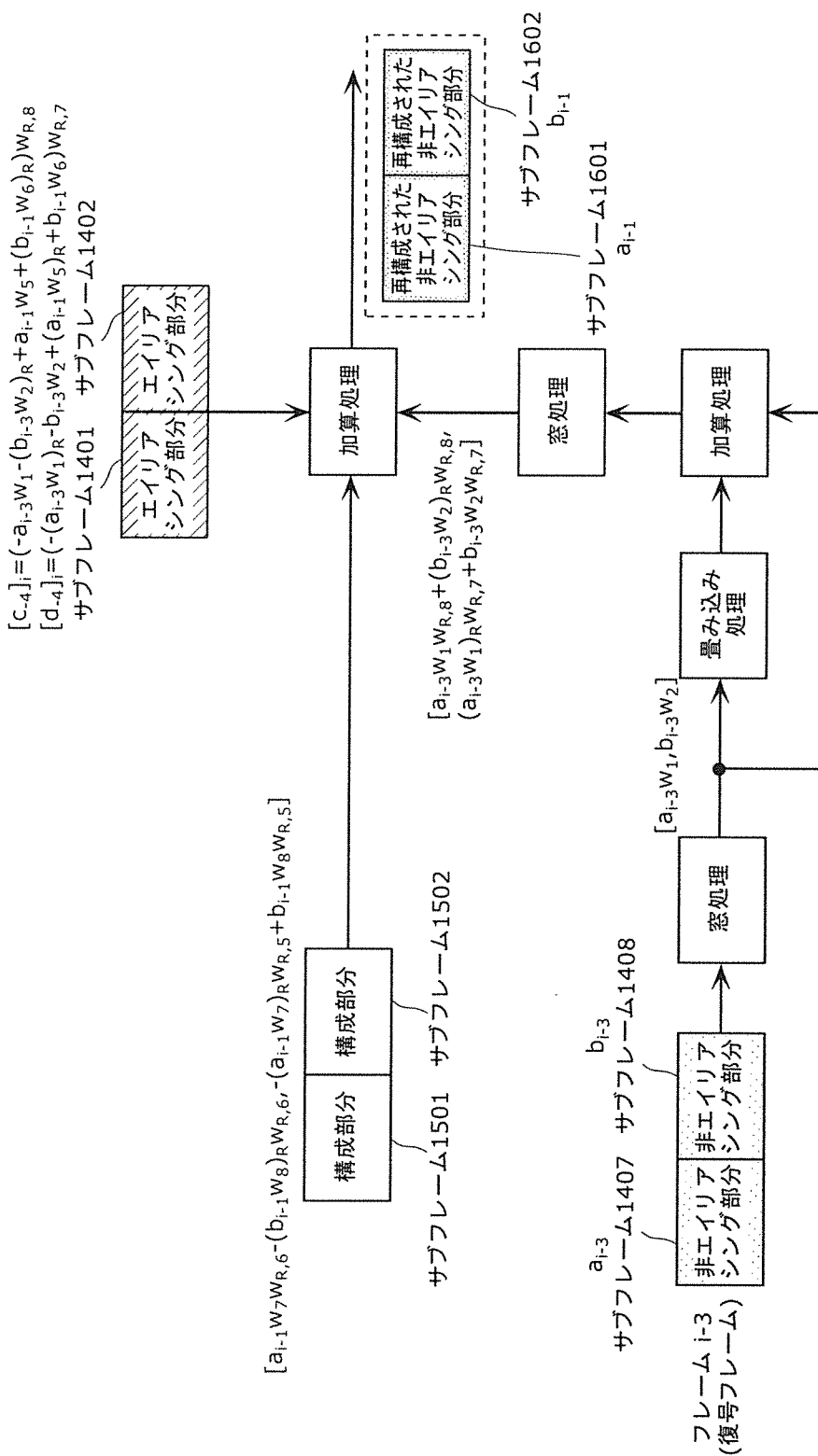
[図19]



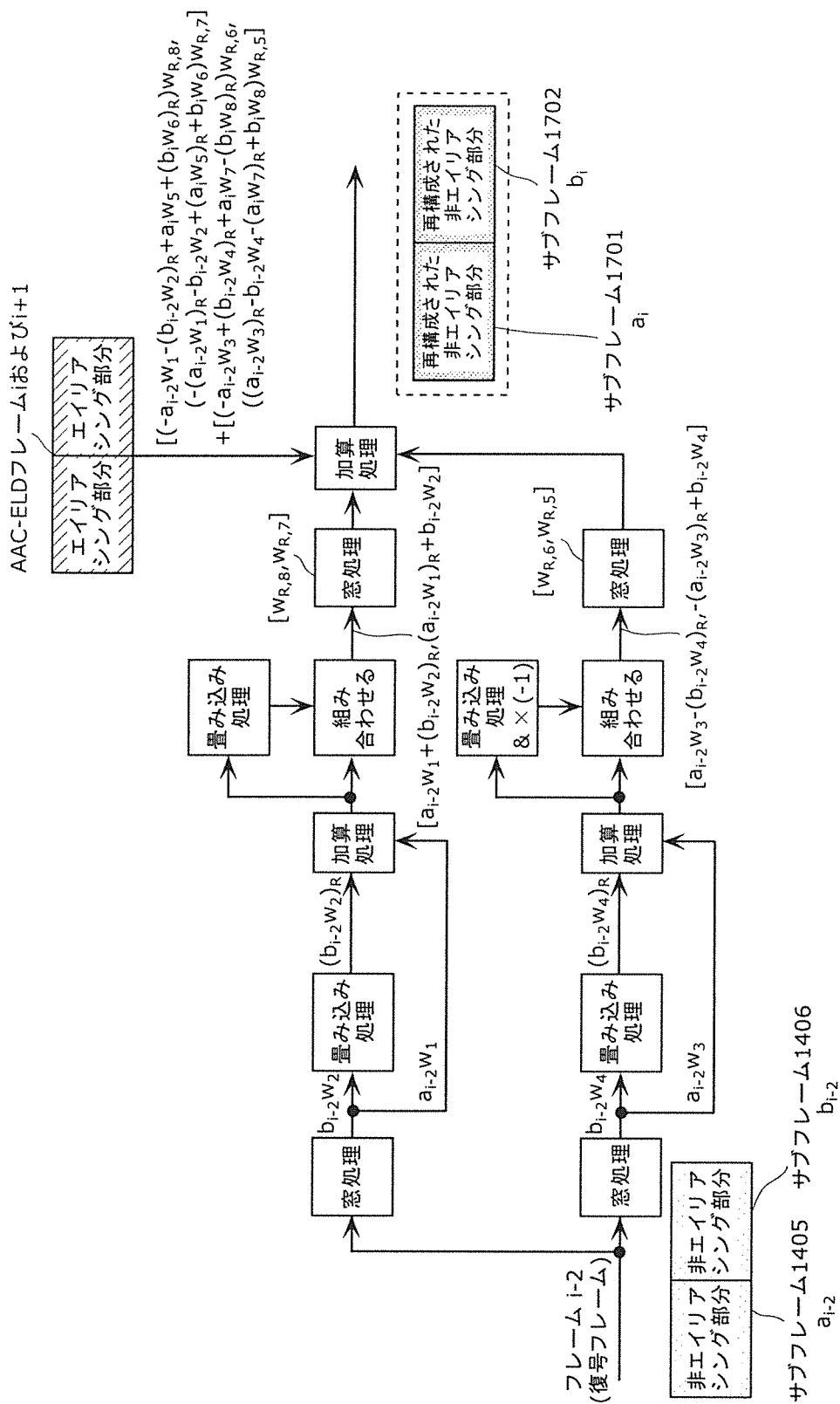
[図20A]



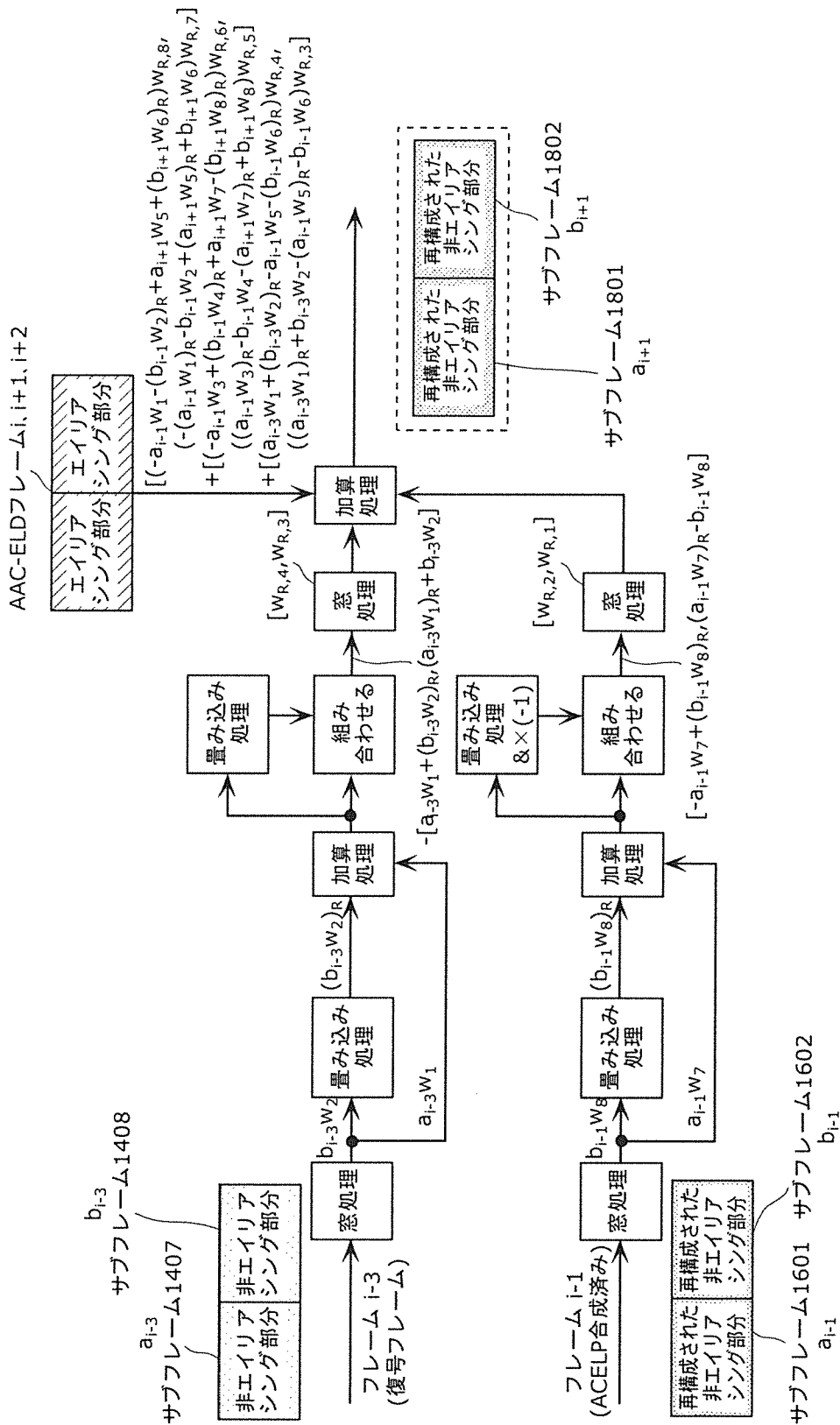
[図20B]



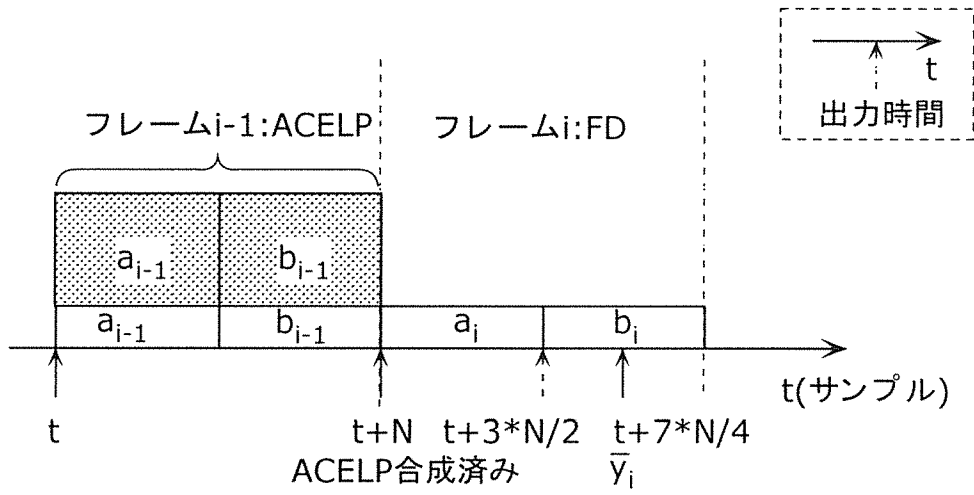
[図21]



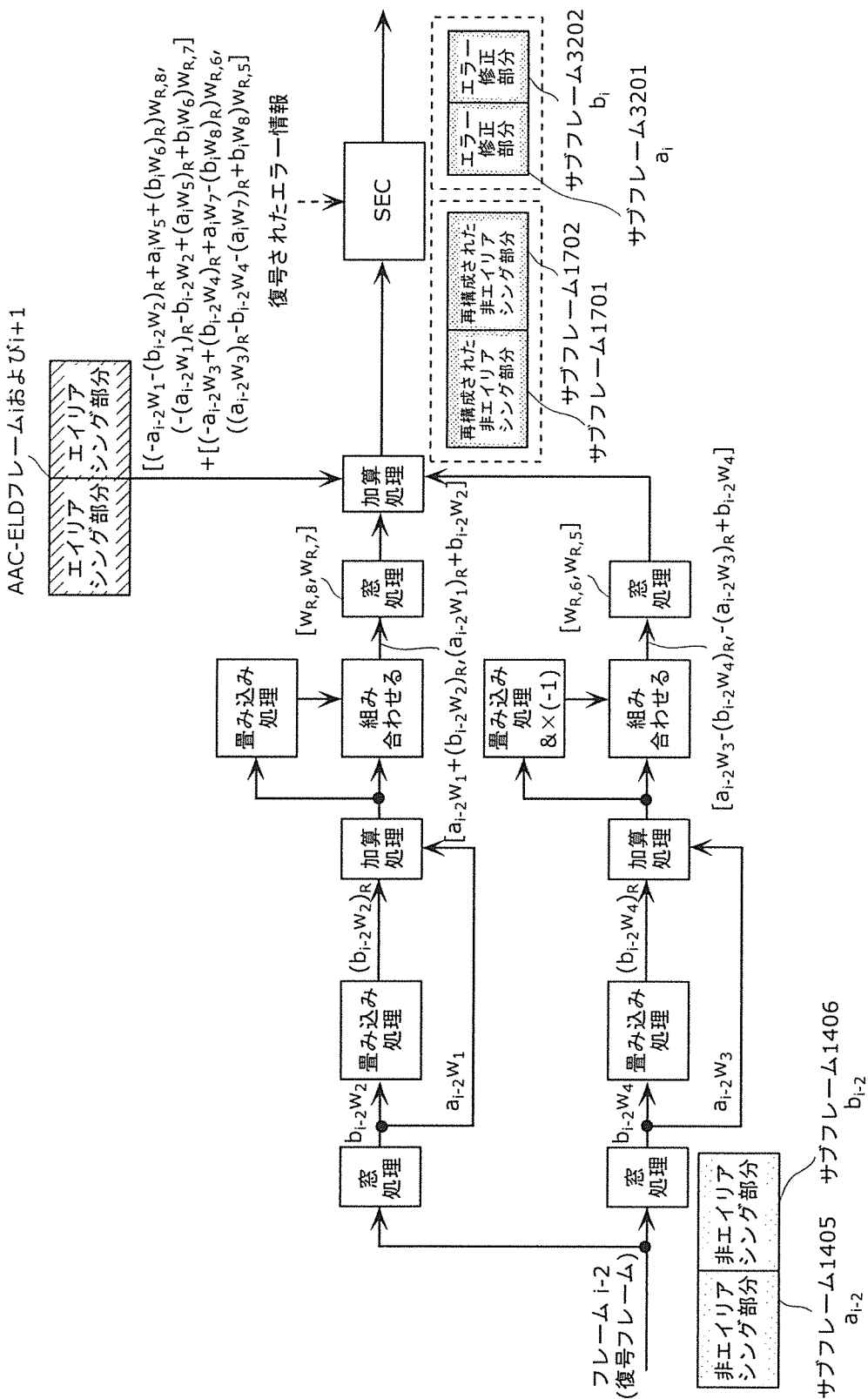
[図22]



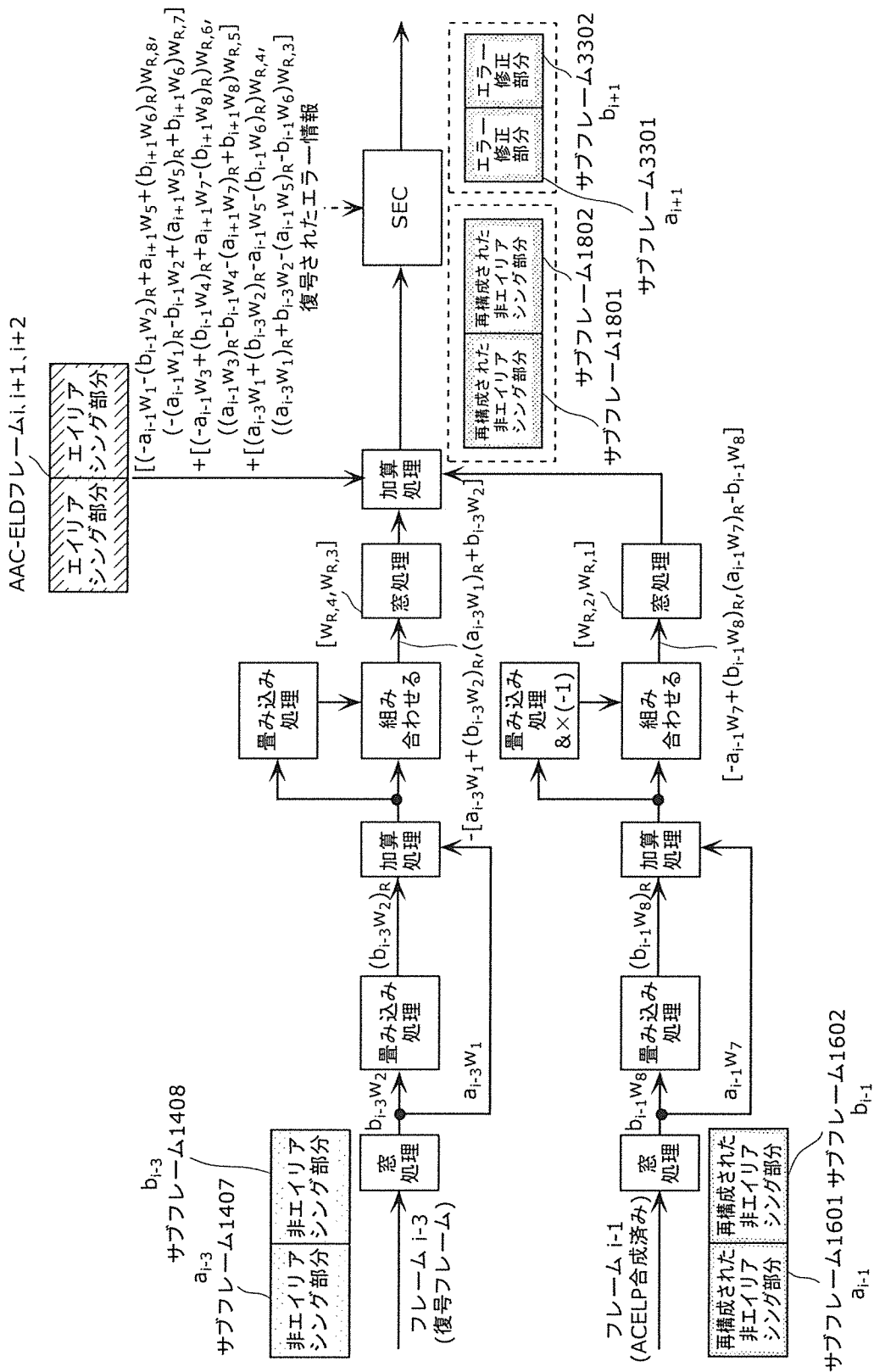
[図23]



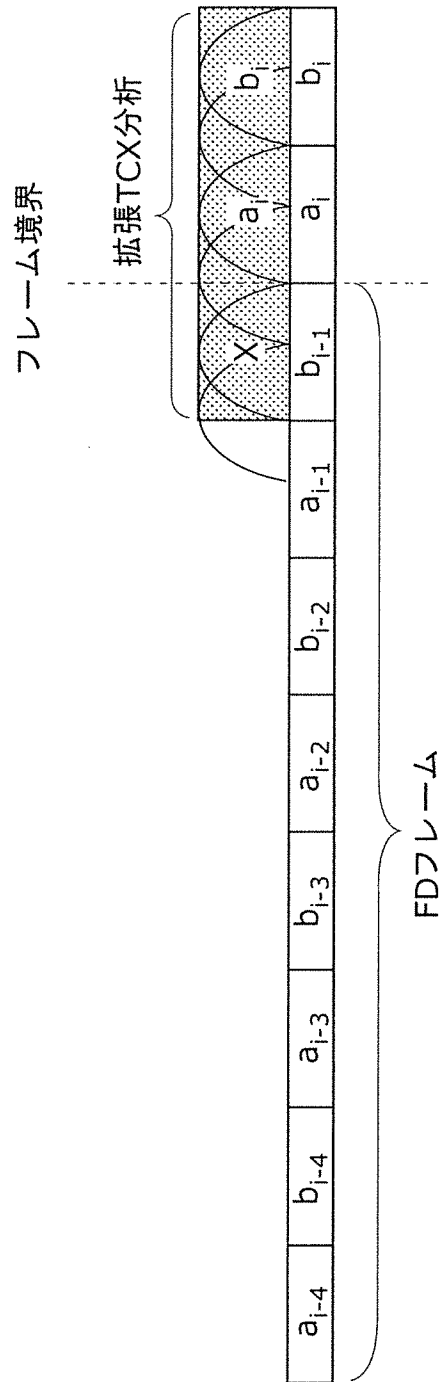
[図25]



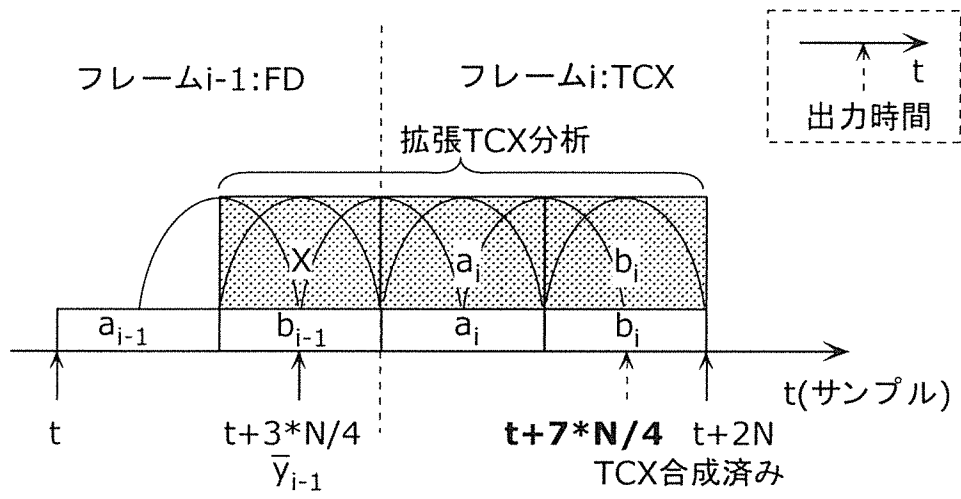
[図26]



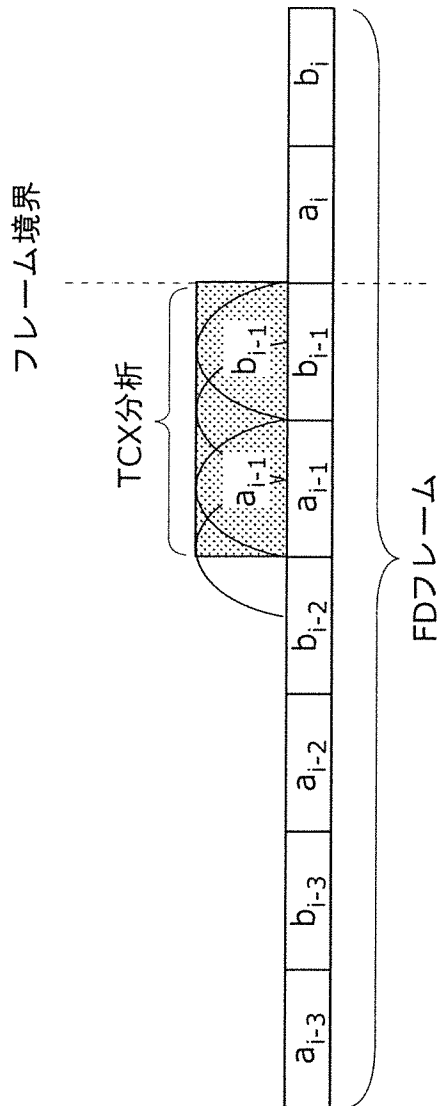
[図27]



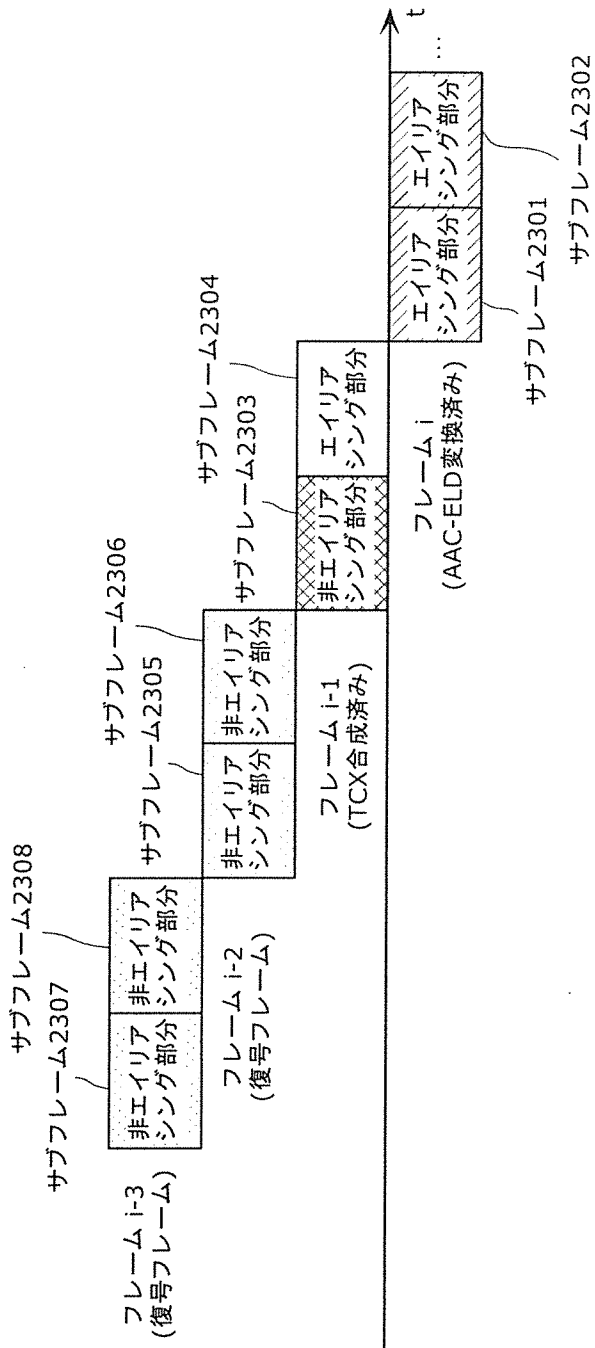
[図29]



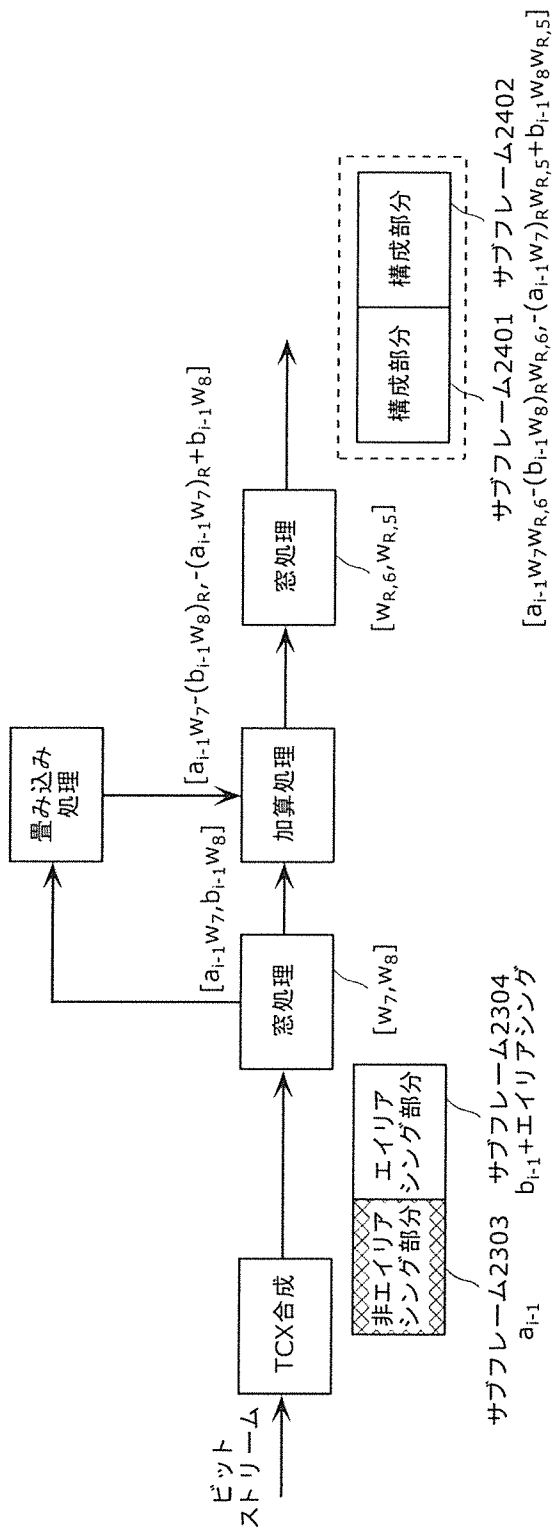
[図30]



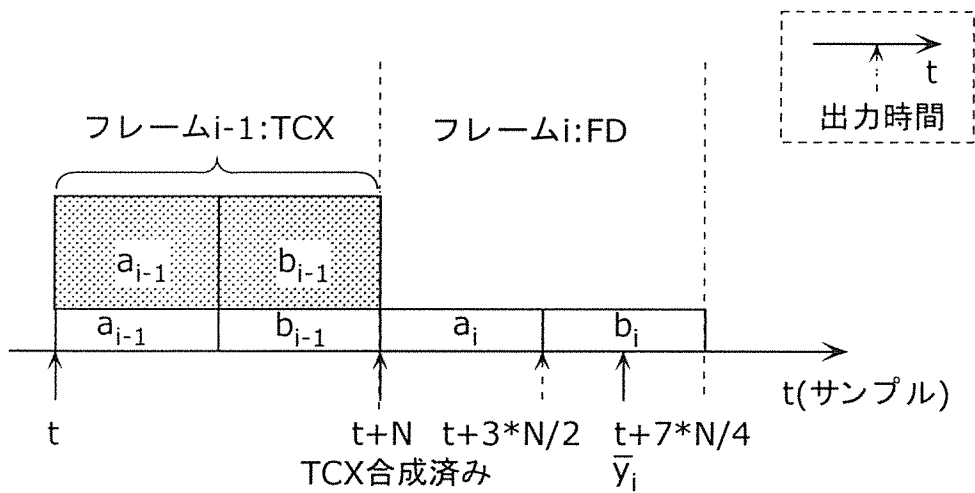
[図31]



[図32]



[図33]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2012/006802

A. CLASSIFICATION OF SUBJECT MATTER

G10L19/14(2006.01) i, G10L19/02(2006.01) 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)

G10L19/14, G10L19/02

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	1996-2012
Kokai Jitsuyo Shinan Koho	1971-2012	Toroku Jitsuyo Shinan Koho	1994-2012

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.
A	WO 2010/003532 A1 (FRAUNHOFER-GESELLSCHAFT), 14 January 2010 (14.01.2010), column 22, line 37 to column 42, line 13; all drawings & JP 2011-527446 A & EP 2301020 A & AR 72479 A & AU 2009267518 A & CA 2730355 A & CN 102089812 A & MX 2011000373 A & TW 201011737 A & KR 10-2011-0040823 A & CO 6341672 A & RU 2011102426 A	1-20

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

* Special categories of cited documents:	"I" 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 14 November, 2012 (14.11.12)	Date of mailing of the international search report 27 November, 2012 (27.11.12)
Name and mailing address of the ISA/ Japanese Patent Office	Authorized officer
Facsimile No.	Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2012/006802

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 2010/003663 A1 (FRAUNHOFER-GESELLSCHAFT), 14 January 2010 (14.01.2010), column 5, line 20 to column 14, line 16; fig. 1 to 5 & JP 2011-527459 A & EP 2311034 A & AR 72556 A & AU 2009267394 A & CA 2730315 A & CN 102105930 A & TW 201009815 A & MX 2011000369 A & KR 10-2011-0052622 A	1-20
A	JP 2010-210680 A (NTT Docomo Inc.), 24 September 2010 (24.09.2010), paragraphs [0008] to [0009]; fig. 3 & EP 2405426 A1 & WO 2010/101190 A1 & CA 2754404 A & TW 201126513 A & KR 10-2011-0124279 A & AU 2010219643 A & SG 174241 A & MX 2011009333 A & CN 102341851 A	1-20
A	JP 08-263098 A (Nippon Telegraph and Telephone Corp.), 11 October 1996 (11.10.1996), entire text; all drawings (Family: none)	1-20
P,A	WO 2011/158485 A1 (Panasonic Corp.), 22 December 2011 (22.12.2011), entire text; all drawings (Family: none)	1-20

A. 発明の属する分野の分類 (国際特許分類 (IPC))
 Int.Cl. G10L19/14(2006.01)i, G10L19/02(2006.01)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. G10L19/14, G10L19/02

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 1971-2012年
 日本国実用新案登録公報 1996-2012年
 日本国登録実用新案公報 1994-2012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
A	WO 2010/003532 A1 (FRAUNHOFER-GESELLSCHAFT) 2010.01.14, 第22欄第37行~第42欄第13行, 全図 & JP 2011-527446 A & EP 2301020 A & AR 72479 A & AU 2009267518 A & CA 2730355 A & CN 102089812 A & MX 2011000373 A & TW 201011737 A & KR 10-2011-0040823 A & CO 6341672 A & RU 2011102426 A	1-20

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 14.11.2012	国際調査報告の発送日 27.11.2012		
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) 菊地 陽一	5 Z	3 2 5 0
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	WO 2010/003663 A1 (FRAUNHOFER-GESELLSCHAFT) 2010.01.14, 第5欄第20行～第14欄第16行, 第1-5図 & JP 2011-527459 A & EP 2311034 A & AR 72556 A & AU 2009267394 A & CA 2730315 A & CN 102105930 A & TW 201009815 A & MX 2011000369 A & KR 10-2011-0052622 A	1-20
A	JP 2010-210680 A (株式会社エヌ・ティ・ティ・ドコモ) 2010.09.24, [0008]-[0009], 第3図 & EP 2405426 A1 & WO 2010/101190 A1 & CA 2754404 A & TW 201126513 A & KR 10-2011-0124279 A & AU 2010219643 A & SG 174241 A & MX 2011009333 A & CN 102341851 A	1-20
A	JP 08-263098 A (日本電信電話株式会社) 1996.10.11, 全文, 全図 (ファミリーなし)	1-20
P, A	WO 2011/158485 A1 (パナソニック株式会社) 2011.12.22, 全文, 全図 (ファミリーなし)	1-20