

(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
2019年12月26日(26.12.2019)



(10) 国際公開番号
WO 2019/244666 A1

(51) 国際特許分類:
G10L 19/02 (2013.01) *G10L 19/022* (2013.01)
G10L 19/00 (2013.01) *H03M 7/40* (2006.01)

(21) 国際出願番号: PCT/JP2019/022681

(22) 国際出願日: 2019年6月7日(07.06.2019)

(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 2018-117635 2018年6月21日(21.06.2018) JP

(71) 出願人: ソニー株式会社 (SONY CORPORATION) [JP/JP]; 〒1080075 東京都港区港南1丁目7番1号 Tokyo (JP).

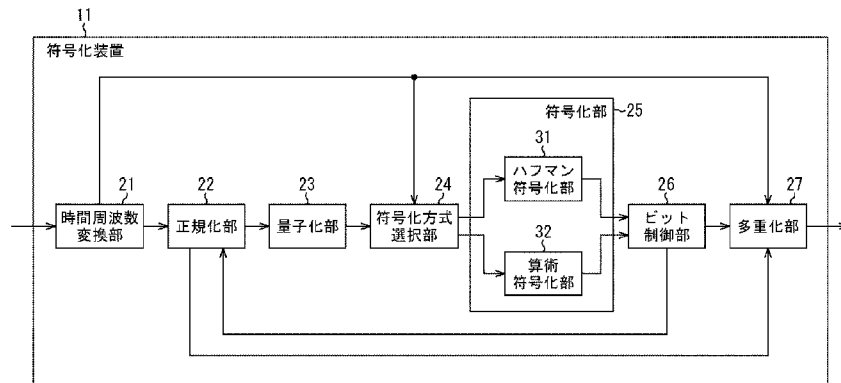
(72) 発明者: 河野 明文 (KONO Akifumi); 〒1080075 東京都港区港南1丁目7番1号 ソニー株式会社内 Tokyo (JP). 知念 徹 (CHINEN Toru); 〒1080075 東京都港区港南1丁目7番1号 ソニー株式会社内 Tokyo (JP). 本間 弘幸 (HONMA Hiroyuki); 〒1080075 東京都港区港南1丁目7番1号 ソニー株式会社内 Tokyo (JP). 及川 芳明 (OIKAWA Yoshiaki); 〒1080075 東京都港区港南1丁目7番1号 ソニー株式会社内 Tokyo (JP).

(74) 代理人: 西川 孝, 外 (NISHIKAWA Takashi et al.); 〒1600023 東京都新宿区西新宿7丁目5番25号 西新宿プライムスクエア9階 Tokyo (JP).

(54) Title: ENCODER AND ENCODING METHOD, DECODER AND DECODING METHOD, AND PROGRAM

(54) 発明の名称: 符号化装置および方法、復号装置および方法、並びにプログラム

[図19]



- 11 Encoder
- 21 Time-frequency conversion unit
- 22 Normalizing unit
- 23 Quantizing unit
- 24 Encoding system selection unit
- 25 Encoding unit
- 26 Bit control unit
- 27 Multiplexer
- 31 Huffman encoding unit
- 32 Arithmetic encoding unit

(57) Abstract: The present technology pertains to an encoder and an encoding method, a decoder and a decoding method, and a program capable of increasing an encoding efficiency. The encoder comprises: a time-frequency conversion unit for performing time-frequency conversion on an audio signal using a conversion window; and an encoding unit for performing Huffman encoding on frequency spectrum information obtained by time-frequency conversion when the conversion window is switched from a short conversion window length to a long conversion window length, and performing arithmetic



WO 2019/244666 A1

- (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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, 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, SA, 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, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類 :

- 一 国際調査報告 (条約第21条(3))

encoding on the frequency spectrum information when the conversion window is not switched from a short conversion window length to a long conversion window length. The present technology can be applied to an encoder and a decoder.

(57) 要約 : 本技術は、符号化効率を向上させることができるようにする符号化装置および方法、復号装置および方法、並びにプログラムに関する。符号化装置は、オーディオ信号に対して変換窓を用いた時間周波数変換を行う時間周波数変換部と、変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替えられた場合、時間周波数変換により得られた周波数スペクトル情報をハフマン符号化し、変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替えられなかった場合、周波数スペクトル情報を算術符号化する符号化部とを備える。本技術は符号化装置および復号装置に適用することができる。

明 細 書

発明の名称：

符号化装置および方法、復号装置および方法、並びにプログラム

技術分野

[0001] 本技術は、符号化装置および方法、復号装置および方法、並びにプログラムに関し、特に、符号化効率を向上させることができるようにした符号化装置および方法、復号装置および方法、並びにプログラムに関する。

背景技術

[0002] 例えば、オーディオ信号を符号化する方法として、国際標準規格であるMPEG(Moving Picture Experts Group)-2 AAC(Advanced Audio Coding)規格、MPEG-4 AAC規格やMPEG-D USAC(Unified Speech and Audio Coding)規格、MPEG-D USAC規格をCore CoderにしたMPEG-H 3D audio規格の符号化等が知られている(例えば、非特許文献1および非特許文献2参照)。

先行技術文献

非特許文献

[0003] 非特許文献1：INTERNATIONAL STANDARD ISO/IEC 14496-3 Fourth edition 2009-09-01 Information technology-coding of audio-visual objects-part3: Audio

非特許文献2：INTERNATIONAL STANDARD ISO/IEC 23003-3 First edition 2012-04-01 Information technology-coding of audio-visual objects-part3: Unified speech and audio coding

発明の概要

発明が解決しようとする課題

[0004] ところで、従来の7.1チャンネルサラウンド再生を超える、より高臨場感な再生や、“3D audio”で実現される多数の音素材(オブジェクト)を伝送するためには、より多くのオーディオチャンネルを圧縮効率よく高速に復号可能な符号化技術が必要となる。すなわち、符号化効率の向上が望まれている。

[0005] 本技術は、このような状況に鑑みてなされたものであり、符号化効率を向上させることができるようにするものである。

課題を解決するための手段

[0006] 本技術の第1の側面の符号化装置は、オーディオ信号に対して変換窓を用いた時間周波数変換を行う時間周波数変換部と、前記変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替えられた場合、前記時間周波数変換により得られた周波数スペクトル情報をハフマン符号化し、前記変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替えられなかった場合、前記周波数スペクトル情報を算術符号化する符号化部とを備える。

[0007] 本技術の第1の側面の符号化方法またはプログラムは、オーディオ信号に対して変換窓を用いた時間周波数変換を行い、前記変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替えられた場合、前記時間周波数変換により得られた周波数スペクトル情報をハフマン符号化し、前記変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替えられなかった場合、前記周波数スペクトル情報を算術符号化するステップを含む。

[0008] 本技術の第1の側面においては、オーディオ信号に対して変換窓を用いた時間周波数変換が行われ、前記変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替えられた場合、前記時間周波数変換により得られた周波数スペクトル情報がハフマン符号化され、前記変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替えられなかった場合、前記周波数スペクトル情報が算術符号化される。

[0009] 本技術の第2の側面の復号装置は、符号化ビットストリームを非多重化して、オーディオ信号の時間周波数変換に用いた変換窓の種類を示す変換窓情報と、前記時間周波数変換により得られた周波数スペクトル情報の符号化データとを抽出する非多重化部と、前記変換窓情報により示される前記変換窓が、短い変換窓長から長い変換窓長へと切り替えられたときに選択される前記変換窓である場合、前記符号化データをハフマン符号化に対応する復号方式で復号する復号部とを備える。

[0010] 本技術の第2の側面の復号方法またはプログラムは、符号化ビットストリームを非多重化して、オーディオ信号の時間周波数変換に用いた変換窓の種類を示す変換窓情報と、前記時間周波数変換により得られた周波数スペクトル情報の符号化データとを抽出し、前記変換窓情報により示される前記変換窓が、短い変換窓長から長い変換窓長へと切り替えられたときに選択される前記変換窓である場合、前記符号化データをハフマン符号化に対応する復号方式で復号するステップを含む。

[0011] 本技術の第2の側面においては、符号化ビットストリームが非多重化されて、オーディオ信号の時間周波数変換に用いた変換窓の種類を示す変換窓情報と、前記時間周波数変換により得られた周波数スペクトル情報の符号化データとが抽出され、前記変換窓情報により示される前記変換窓が、短い変換窓長から長い変換窓長へと切り替えられたときに選択される前記変換窓である場合、前記符号化データがハフマン符号化に対応する復号方式で復号される。

発明の効果

[0012] 本技術の第1の側面および第2の側面によれば、符号化効率を向上させることができる。

[0013] なお、ここに記載された効果は必ずしも限定されるものではなく、本開示中に記載された何れかの効果であってもよい。

図面の簡単な説明

[0014] [図1]MPEG-4 AACでの符号化について説明する図である。

[図2]MPEG-4 AACにおける変換窓の種類について説明する図である。

[図3]MPEG-D USACでの符号化について説明する図である。

[図4]MPEG-D USACにおける変換窓の種類について説明する図である。

[図5]ハフマン符号化と算術符号化の符号化効率について説明する図である。

[図6]ハフマン符号化と算術符号化の符号化効率について説明する図である。

[図7]ハフマン符号化と算術符号化の符号化効率について説明する図である。

[図8]ハフマン符号化と算術符号化の符号化効率について説明する図である。

[図9]ハフマン符号化と算術符号化の符号化効率について説明する図である。

[図10]ハフマン符号化と算術符号化の符号化効率について説明する図である

。

[図11]ハフマン符号化と算術符号化の符号化効率について説明する図である

。

[図12]ハフマン符号化と算術符号化の符号化効率について説明する図である

。

[図13]ハフマン符号化と算術符号化の符号化効率について説明する図である

。

[図14]ハフマン符号化と算術符号化の符号化効率について説明する図である

。

[図15]ハフマン符号化と算術符号化の符号化効率について説明する図である

。

[図16]ハフマン符号化と算術符号化の符号化効率について説明する図である

。

[図17]ハフマン符号化と算術符号化の符号化効率について説明する図である

。

[図18]ハフマン符号化と算術符号化の符号化効率について説明する図である

。

[図19]符号化装置の構成例を示す図である。

[図20]符号化処理を説明するフローチャートである。

[図21]復号装置の構成例を示す図である。

[図22]復号処理を説明するフローチャートである。

[図23]本技術における符号化効率について説明する図である。

[図24]本技術における符号化効率について説明する図である。

[図25]チャンネルストリームのシンタックス例を示す図である。

[図26]ics_infoのシンタックス例を示す図である。

[図27]符号化処理を説明するフローチャートである。

[図28]復号処理を説明するフローチャートである。

[図29]符号化処理を説明するフローチャートである。

[図30]コンピュータの構成例を示す図である。

発明を実施するための形態

[0015] 以下、図面を参照して、本技術を適用した実施の形態について説明する。

[0016] 〈第1の実施の形態〉

〈本技術について〉

まず、本技術の概要について説明する。なお、本技術では、符号化対象とされる信号は、オーディオ信号や画像信号など、どのような信号であってもよいが、以下では符号化対象がオーディオ信号である場合を例として説明を行う。

[0017] 例えばMPEG-4 AACでは、図1に示すようにしてオーディオ信号の符号化が行われる。

[0018] すなわち、符号化（エンコード）の処理が開始されると、最初にオーディオ信号に対してMDCT（Modified Discrete Cosine Transform）（修正離散コサイン変換）を用いた時間周波数変換が行われる。

[0019] 次に、MDCTにより得られた周波数スペクトル情報であるMDCT係数が、スケールファクターバンドごとに量子化され、その結果として量子化MDCT係数が得られる。

[0020] ここで、スケールファクターバンドとは、QMF（Quadrature Mirror Filter）分析フィルタの分解能である所定帯域幅のサブバンドを複数束ねて得られる帯域である。

[0021] 量子化により量子化MDCT係数が得られると、同一のハフマンコードブックを使用するセクションごとに、ハフマン符号が用いられて量子化MDCT係数とハフマンコードブック情報とが符号化される。すなわち、ハフマン符号化が行われる。なお、セクションは、スケールファクターバンドを複数束ねた帯域である。

[0022] 以上のようにして得られたハフマン符号、つまりハフマン符号化された量

子化MDCT係数とハフマンコードブック情報とが、オーディオ信号の符号化データとして出力される。

[0023] また、時間周波数変換においては、一般的に処理対象となるオーディオ信号の性質に応じて適切な変換窓を選択することで、単一の変換窓を用いるよりも高音質にオーディオ信号を圧縮できることが知られている。

[0024] 例えば、時間的に急激な変化を伴うアタック性の強い音楽信号（アタック性音楽信号）に対しては短い変換窓長の変換窓が適しており、時間的に急激な変化を伴わない定常性の強い音楽信号（定常性音楽信号）に対しては長い変換窓長の変換窓が適していることが知られている。

[0025] 具体的には、例えばMPEG4 AACでは図2に示すように4つのwindow sequenceのなかの適切なものに適宜切り替えられてMDCTが行われている。

[0026] 図2において、「window_sequence」はwindow sequenceを示している。ここで、window sequenceは変換窓の種類、つまりウィンドウタイプを示している。

[0027] 特に、MPEG4 AACではwindow_sequence、すなわちウィンドウタイプとして、ONLY_LONG_SEQUENCE、LONG_START_SEQUENCE、EIGHT_SHORT_SEQUENCE、およびLONG_STOP_SEQUENCEの4つの種類の変換窓が選択可能となっている。

[0028] また、図2において「num_windows」は、各ウィンドウタイプの変換窓を用いたMDCTを行うときに用いられる変換窓の数を示しており、「looks like」の欄には変換窓の形状が示されている。特に「looks like」の欄では、図中、横方向が時間方向を示しており、図中、縦方向は各サンプル位置における変換窓の大きさ、つまり各サンプルに乗算される係数の大きさを示している。

[0029] MPEG4 AACでは、オーディオ信号のMDCT時には、定常性の強いフレームではONLY_LONG_SEQUENCEが選択される。このONLY_LONG_SEQUENCEにより示される変換窓は、変換窓長が2048サンプルの変換窓となっている。

[0030] また、アタック性の強いフレームではEIGHT_SHORT_SEQUENCEが選択される。このEIGHT_SHORT_SEQUENCEにより示される変換窓は、時間方向に分割され

た8個の変換窓であり、それらの分割された各変換窓の変換窓長は256サンプルとなっている。

- [0031] EIGHT_SHORT_SEQUENCEにより示される変換窓は、LONG_STOP_SEQUENCEにより示される変換窓等の他の変換窓よりも変換窓長が短くなっている。
- [0032] window_sequenceがONLY_LONG_SEQUENCEからEIGHT_SHORT_SEQUENCEへと遷移するフレームではLONG_START_SEQUENCEが選択される。このLONG_START_SEQUENCEにより示される変換窓は、変換窓長が2048サンプルの変換窓となっている。
- [0033] window_sequenceがEIGHT_SHORT_SEQUENCEからONLY_LONG_SEQUENCEへと遷移するフレームではLONG_STOP_SEQUENCEが選択される。
- [0034] すなわち、変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替えられた場合、LONG_STOP_SEQUENCEが選択される。LONG_STOP_SEQUENCEにより示される変換窓は、変換窓長が2048サンプルの変換窓となっている。
- [0035] なお、MPEG4 AACで用いられる変換窓の詳細については、例えば「INTERNATIONAL STANDARD ISO/IEC 14496-3 Fourth edition 2009-09-01 Information technology-coding of audio-visual objects-part3:Audio」に詳細に記載されている。
- [0036] これに対して、MPEG-D USACでは、図3に示すようにしてオーディオ信号の符号化が行われる。
- [0037] すなわち、符号化（エンコード）の処理が開始されると、MPEG-4 AACにおける場合と同様に、まずオーディオ信号に対してMDCTを用いた時間周波数変換が行われる。
- [0038] そして、時間周波数変換により得られたMDCT係数がスケールファクターバンドごとに量子化され、その結果として量子化MDCT係数が得られる。
- [0039] さらに、量子化MDCT係数に対してコンテキストベースの算術符号化が行われ、算術符号化された量子化MDCT係数がオーディオ信号の符号化データとして出力される。
- [0040] コンテキストベースの算術符号化では、出現確率の高い入力ビット列に短

い符号を割り当て、出現確率の低い入力ビット列に長い符号を割り当てた出現確率テーブルが複数用意されている。

- [0041] また、符号化対象となる量子化MDCT係数に対して時間的および周波数的に近傍にある過去の量子化MDCT係数の符号化結果（コンテキスト）に基づいて、効率のよい出現確率テーブルが選択される。すなわち、時間的および周波数的に近い量子化MDCT係数の相関性が考慮されて出現確率テーブルが適切に切り替えられる。そして、選択された出現確率テーブルが用いられて量子化MDCT係数が符号化される。
- [0042] コンテキストベースの算術符号化では、複数の出現確率テーブルのなかから効率のよいものを選択して符号化を行うことで、高い符号化効率を実現することができる。
- [0043] また、算術符号化ではハフマン符号化と異なり、コードブック情報を送る必要がない。そのため、算術符号化ではハフマン符号化と比較してコードブック情報分の符号量を削減することができる。
- [0044] なお、MPEG-D USACでは図4に示すように適宜5つのwindow sequenceのなかの適切なものに切り替えられてMDCTが行われている。
- [0045] 図4において、「Window」はwindow sequenceを示しており、「num_windows」は、各ウィンドウタイプの変換窓を用いたMDCTを行うときに用いられる変換窓の数を示しており、「Window Shape」の欄には変換窓の形状が示されている。
- [0046] MPEG-D USACではwindow_sequenceとして、ONLY_LONG_SEQUENCE、LONG_START_SEQUENCE、EIGHT_SHORT_SEQUENCE、LONG_STOP_SEQUENCE、およびSTOP_START_SEQUENCEの5つの種類の変換窓が選択可能となっている。
- [0047] 特にwindow_sequence、つまりウィンドウタイプのうち、ONLY_LONG_SEQUENCE、LONG_START_SEQUENCE、EIGHT_SHORT_SEQUENCE、およびLONG_STOP_SEQUENCEについてはMPEG4 AACにおける場合と同じとなっている。
- [0048] MPEG-D USACでは、これらの4つのウィンドウタイプに加えて、さらにSTOP_START_SEQUENCEが用意されている。

- [0049] STOP_START_SEQUENCEは、window_sequenceがLONG_STOP_SEQUENCEからLONG_START_SEQUENCEへと遷移するフレームで選択される。
- [0050] このSTOP_START_SEQUENCEにより示される変換窓は、変換窓長が2048サンプルの変換窓となっている。
- [0051] なお、MPEG-D USACについては、例えば「INTERNATIONAL STANDARD ISO/IEC 23003-3 First edition 2012-04-01 Information technology-coding of audio-visual objects-part3:Unified speech and audio coding」に詳細に記載されている。
- [0052] また、以下では、MPEG4 AACを単にAACとも称し、MPEG-D USACを単にUSACとも称することとする。
- [0053] 以上において説明したAACとUSACを比較すると、現状のUSACではAACで採用しているハフマン符号化よりも圧縮効率（符号化効率）がよいとされるコンテキストベースの算術符号化が採用されている。
- [0054] しかし、全てのオーディオ信号に対してコンテキストベースの算術符号化がハフマン符号化よりも圧縮効率がよくなる（高くなる）わけではない。
- [0055] USACのコンテキストベースの算術符号化では定常性音楽信号に対しては符号が短くなり、AACのハフマン符号化よりも符号化効率が高くなりやすいが、その反面、アタック性音楽信号に対しては符号が長くなり符号化効率が低くなってしまふ。
- [0056] そのような例を図5乃至図18に示す。なお、図5乃至図18において横軸は時間、つまりオーディオ信号のフレームを示しており、縦軸はオーディオ信号を符号化したときの符号ビット数（必要ビット数）または必要ビット数の差分（差分ビット数）を示している。特に、ここでは1フレームは1024サンプルとなっている。
- [0057] 図5は、オーディオ信号としての定常性音楽信号に対してMDCTおよび量子化を行い、量子化後の同じ量子化MDCT係数に対して、AACのハフマン符号化を行った場合とUSACの算術符号化を行った場合でそれぞれ必要になる必要ビット数を示している。

- [0058] この例では、折れ線L11が各フレームにおけるUSACの算術符号化の必要ビット数を示しており、折れ線L12が各フレームにおけるAACのハフマン符号化の必要ビット数を示している。この例では、殆どのフレームにおいてAACのハフマン符号化よりも、USACの算術符号化の方が必要ビット数が少なくなっていることが分かる。
- [0059] また、図5における一部を拡大すると図6に示すようになる。なお、図6において図5における場合と対応する部分には同一の符号を付してあり、その説明は省略する。
- [0060] 図6に示す部分ではAACのハフマン符号化の必要ビット数と、USACの算術符号化の必要ビット数とでは100から150ビット程度の差があり、USACの算術符号化の方が符号化効率がよい（高い）ことが分かる。
- [0061] 図7は、図5に示した各フレームにおけるAACのハフマン符号化の必要ビット数と、USACの算術符号化の必要ビット数との差分、すなわち差分ビット数を示している。
- [0062] 図7において横軸はフレーム（時間）を示しており、縦軸は差分ビット数を示している。なお、ここでの差分ビット数は、USACの算術符号化の必要ビット数から、AACのハフマン符号化の必要ビット数を減算して得られたものである。
- [0063] 図7から分かるように、オーディオ信号が定常性音楽信号である場合、すなわちオーディオ信号が定常性を有する信号である場合、殆どのフレームにおいて差分ビット数が負の値となる。つまり、殆どのフレームにおいてAACのハフマン符号化よりも、USACの算術符号化の方が必要ビット数が少なくなっていることが分かる。
- [0064] したがって、符号化対象のオーディオ信号が定常性を有する信号である場合には、符号化方式として算術符号化を選択すると、より高い符号化効率を得ることができる。
- [0065] また、各フレームではMDCT時にwindow sequence、すなわち変換窓の種類が選択されるが、図7に示す差分ビット数のグラフを図2に示した4つのwindo

w sequenceごとに分離すると、図8乃至図11に示すようになる。

[0066] すなわち、図8は図7に示した各フレームの差分ビット数のうち、window sequenceとしてONLY_LONG_SEQUENCEが選択されたフレームの差分ビット数を示している。

[0067] 同様に、図9は図7に示した各フレームの差分ビット数のうち、window sequenceとしてLONG_START_SEQUENCEが選択されたフレームの差分ビット数を示している。図10は図7に示した各フレームの差分ビット数のうち、window sequenceとしてEIGHT_SHORT_SEQUENCEが選択されたフレームの差分ビット数を示している。

[0068] さらに図11は図7に示した各フレームの差分ビット数のうち、window sequenceとしてLONG_STOP_SEQUENCEが選択されたフレームの差分ビット数を示している。

[0069] なお、これらの図8乃至図11において横軸はフレーム（時間）を示しており、縦軸は差分ビット数を示している。

[0070] これらの図8乃至図11から分かるように、オーディオ信号が定常性音楽信号であるため、殆どのフレームにおいてONLY_LONG_SEQUENCEが選択されている。そして、残りのLONG_START_SEQUENCE、EIGHT_SHORT_SEQUENCE、およびLONG_STOP_SEQUENCEが選択されるフレームは少ないことが分かる。

[0071] ここでは、図11に示すようにLONG_STOP_SEQUENCEが選択された場合には、差分ビット数が正の値となり、AACのハフマン符号化の方が符号化効率が高くなるフレームが多くなっている。しかし、図7に示したように全体で見れば、AACのハフマン符号化よりもUSACの算術符号化の方が符号化効率が高いことが分かる。

[0072] これに対して図12乃至図18は、図5乃至図11に対応するものであり、オーディオ信号がアタック性音楽信号である場合における必要ビット数や差分ビット数を示している。

[0073] すなわち、図12は、オーディオ信号としてのアタック性音楽信号に対してMDCTおよび量子化を行い、量子化後の同じ量子化MDCT係数に対して、AACの

ハフマン符号化を行った場合とUSACの算術符号化を行った場合でそれぞれ必要になる必要ビット数を示している。

[0074] この例では、折れ線L31が各フレームにおけるUSACの算術符号化の必要ビット数を示しており、折れ線L32が各フレームにおけるAACのハフマン符号化の必要ビット数を示している。

[0075] この例では、多くのフレームではAACのハフマン符号化よりもUSACの算術符号化の方が必要ビット数が少なくなっている。しかし、AACのハフマン符号化の方が必要ビット数が少ないフレームが定常性音楽信号における場合よりも多くなっている。

[0076] また、図12における一部を拡大すると図13に示すようになる。なお、図13において図12における場合と対応する部分には同一の符号を付しており、その説明は省略する。

[0077] 図13に示す部分では、いくつかのフレームではUSACの算術符号化よりもAACのハフマン符号化の方が必要ビット数が少なくなっていることが分かる。

[0078] 図14は、図12に示した各フレームにおけるAACのハフマン符号化の必要ビット数と、USACの算術符号化の必要ビット数との差分、すなわち差分ビット数を示している。

[0079] 図14において横軸はフレーム（時間）を示しており、縦軸は差分ビット数を示している。なお、ここでの差分ビット数は、USACの算術符号化の必要ビット数から、AACのハフマン符号化の必要ビット数を減算して得られたものである。

[0080] 図14から分かるように、オーディオ信号がアタック性音楽信号である場合、すなわちオーディオ信号がアタック性を有する信号である場合、多くのフレームにおいて差分ビット数が負の値となる。

[0081] しかし、オーディオ信号が定常性音楽信号である場合と比較すると、差分ビット数が正の値となるフレームの数が多くなっていることが分かる。すなわち、AACのハフマン符号化の方が、USACの算術符号化よりも必要ビット数が少ないフレームも多くなっていることが分かる。

- [0082] また、各フレームではMDCT時にwindow sequence、すなわち変換窓の種類が選択されるが、図14に示す差分ビット数のグラフを図2に示した4つのwindow sequenceごとに分離すると、図15乃至図18に示すようになる。
- [0083] すなわち、図15は図14に示した各フレームの差分ビット数のうち、window sequenceとしてONLY_LONG_SEQUENCEが選択されたフレームの差分ビット数を示している。
- [0084] 同様に、図16は図14に示した各フレームの差分ビット数のうち、window sequenceとしてLONG_START_SEQUENCEが選択されたフレームの差分ビット数を示している。図17は図14に示した各フレームの差分ビット数のうち、window sequenceとしてEIGHT_SHORT_SEQUENCEが選択されたフレームの差分ビット数を示している。
- [0085] さらに図18は図14に示した各フレームの差分ビット数のうち、window sequenceとしてLONG_STOP_SEQUENCEが選択されたフレームの差分ビット数を示している。
- [0086] なお、これらの図15乃至図18において横軸はフレーム（時間）を示しており、縦軸は差分ビット数を示している。
- [0087] これらの図15乃至図18から分かるように、オーディオ信号がアタック性音楽信号である場合には、window sequenceとしてEIGHT_SHORT_SEQUENCE、LONG_START_SEQUENCE、LONG_STOP_SEQUENCEが選択される割合が定常性音楽信号である場合と比較して多いことが分かる。
- [0088] また、オーディオ信号がアタック性音楽信号である場合でもONLY_LONG_SEQUENCEやLONG_START_SEQUENCE、EIGHT_SHORT_SEQUENCEが選択された場合には、定常性音楽信号における場合と同様に、殆どのフレームにおいてAACのハフマン符号化よりもUSACの算術符号化の方が符号化効率が高いことが分かる。
- [0089] しかし、LONG_STOP_SEQUENCEが選択された場合には、殆どのフレームにおいてUSACの算術符号化よりもAACのハフマン符号化の方が必要ビット数が少なく、符号化効率が高いことが分かる。
- [0090] これは、USACの算術符号化において、アタック性の強いフレームと定常性

の強いフレームとの遷移でコンテキストの相関性が低くなり、効率の悪い出現確率テーブルが選択されてしまったためである。

[0091] なお、EIGHT_SHORT_SEQUENCEが選択されたフレームにおいて、USACの算術符号化の必要ビット数（符号量）が多くなるのは、量子化MDCT係数の符号化に時間方向に8分割された変換窓が用いられるためである。すなわち、量子化MDCT係数の符号化が、時間方向に8分割された256サンプル分ずつ8回に分けて行われるために、コンテキストの相関性の低減度合いが分散および軽減されるためである。

[0092] 以上のようにUSACの算術符号化では、アタック性を有するオーディオ信号において、特に短い変換窓長の変換窓を用いるフレームから、長い変換窓長の変換窓を用いるフレームに遷移したときのフレーム、すなわちLONG_STOP_SEQUENCEが選択されたフレームでAACのハフマン符号化よりも符号化効率（圧縮効率）が悪くなってしまう。

[0093] また、算術符号の符号長が長くなるということは、当然、復号時の計算量も多くなってしまうことになる。

[0094] さらに算術符号化は、1つの量子化MDCT係数の符号が全て揃わないと復号できないことや、ビット単位の演算処理が大量に発生するため、ハフマン符号化よりも計算量が多くなるという性質がある。

[0095] そこで、本技術では、オーディオ信号の符号化時に適切に符号化方式を選択することで符号化効率を向上させるとともに、復号の計算量を低減させることができるようにした。

[0096] 具体的には、例えばUSACのように時間周波数変換を用いたコーデックにおいて、短い変換窓長の変換窓で時間周波数変換を行うフレームから、その変換窓よりも長い変換窓長の変換窓で時間周波数変換を行うフレームに遷移した場合、量子化された周波数スペクトル情報がハフマン符号化されるようにする。

[0097] 例えばUSACであれば、LONG_STOP_SEQUENCEが選択されたフレームでは、符号化方式としてハフマン符号化が選択されるようにする。

[0098] また、それ以外のフレーム、つまり短い変換窓長から長い変換窓長へと遷移したときのフレーム以外の他のフレームにおいては、符号化方式としてハフマン符号化または算術符号化の何れかが選択されるようにする。

[0099] このとき、必要に応じて選択した符号化方式を識別する判定フラグを符号化ビットストリームに含めることで、復号側においてどの符号化方式が選択されたかを特定することができる。すなわち、デコーダシンタックスで判定フラグや復号方式の切り替えを規定することで、復号側において適切に復号方式を切り替えることができる。

[0100] 〈符号化装置の構成例〉

続いて、本技術を適用した符号化装置と復号装置の具体的な実施の形態について説明する。なお、以下ではMPEG-D USACをベースに符号化および復号を行う実施の形態について説明する。しかし、変換窓長が適宜切り替えられて時間周波数変換された情報に対して、コンテキストベースの算術符号化を含む複数の符号化方式のうちの何れかを選択して符号化を行うものであれば、他のどのようなコーデックでもよい。

[0101] 図19は、本技術を適用した符号化装置の構成例を示す図である。

[0102] 図19に示す符号化装置11は、時間周波数変換部21、正規化部22、量子化部23、符号化方式選択部24、符号化部25、ビット制御部26、および多重化部27を有している。

[0103] 時間周波数変換部21は、供給されたオーディオ信号のフレームごとに変換窓を選択し、選択した変換窓を用いてオーディオ信号に対する時間周波数変換を行う。

[0104] また、時間周波数変換部21は、時間周波数変換により得られた周波数スペクトル情報を正規化部22に供給するとともに、各フレームで選択した変換窓の種類(window sequence)を示す変換窓情報を符号化方式選択部24および多重化部27に供給する。

[0105] 例えば時間周波数変換部21では、時間周波数変換としてMDCTが行われ、周波数スペクトル情報としてMDCT係数が得られる。以下では、周波数スペク

トル情報がMDCT係数である場合を例として説明を続ける。

- [0106] 正規化部22は、ビット制御部26から供給された正規化のためのパラメータに基づいて、時間周波数変換部21から供給されたMDCT係数を正規化し、その結果得られた正規化MDCT係数を量子化部23に供給するとともに、正規化に関するパラメータを多重化部27に供給する。
- [0107] 量子化部23は、正規化部22から供給された正規化MDCT係数を量子化し、その結果得られた量子化MDCT係数を符号化方式選択部24に供給する。
- [0108] 符号化方式選択部24は、時間周波数変換部21から供給された変換窓情報に基づいて符号化方式を選択し、量子化部23から供給された量子化MDCT係数を、符号化方式の選択結果に応じた符号化部25のブロックへと供給する。
- [0109] 符号化部25は、符号化方式選択部24から供給された量子化MDCT係数を、符号化方式選択部24により選択（指定）された符号化方式で符号化する。符号化部25は、ハフマン符号化部31および算術符号化部32を有している。
- [0110] ハフマン符号化部31は、符号化方式選択部24から量子化MDCT係数が供給された場合、その量子化MDCT係数をハフマン符号化方式により符号化する。すなわち量子化MDCT係数がハフマン符号化される。
- [0111] ハフマン符号化部31は、ハフマン符号化により得られたMDCT符号化データおよびハフマンコードブック情報をビット制御部26に供給する。ここで、ハフマンコードブック情報は、ハフマン符号化時に用いられたハフマンコードブックを示す情報である。また、ビット制御部26へと供給されるハフマンコードブック情報はハフマン符号化されたものとなっている。
- [0112] 算術符号化部32は、符号化方式選択部24から量子化MDCT係数が供給された場合、その量子化MDCT係数を算術符号化方式により符号化する。すなわち、量子化MDCT係数に対してコンテキストベースの算術符号化が行われる。
- [0113] 算術符号化部32は、算術符号化により得られたMDCT符号化データをビット制御部26に供給する。

- [0114] ビット制御部 26 は、ハフマン符号化部 31 から MDCT 符号化データおよびハフマンコードブック情報が供給されたか、または算術符号化部 32 から MDCT 符号化データが供給されるとビット量と音質の判定を行う。
- [0115] すなわち、ビット制御部 26 は MDCT 符号化データ等のビット量（符号量）が目標使用ビット量内となっているかや、MDCT 符号化データに基づく音の音質が許容範囲内の品質であるかを判定する。
- [0116] ビット制御部 26 は、MDCT 符号化データ等のビット量が目標使用ビット量内であり、かつ音質が許容範囲内である場合、供給された MDCT 符号化データ等を多重化部 27 へと供給する。
- [0117] これに対して、ビット制御部 26 は MDCT 符号化データ等のビット量が目標使用ビット量内でない場合や、音質が許容範囲内でない場合には、正規化部 22 に供給するパラメータを再設定し、再設定後のパラメータを正規化部 22 に供給して再度の符号化を実施させる。
- [0118] 多重化部 27 は、ビット制御部 26 から供給された MDCT 符号化データやハフマンコードブック情報、時間周波数変換部 21 から供給された変換窓情報、および正規化部 22 から供給されたパラメータを多重化し、その結果得られた符号化ビットストリームを出力する。
- [0119] 〈符号化処理の説明〉
- 次に、符号化装置 11 の動作について説明する。すなわち、以下、図 20 のフローチャートを参照して、符号化装置 11 による符号化処理について説明する。なお、この符号化処理はオーディオ信号のフレームごとに行われる。
- [0120] ステップ S11 において時間周波数変換部 21 は、供給されたオーディオ信号のフレームに対して時間周波数変換を行う。
- [0121] すなわち、例えば時間周波数変換部 21 は、オーディオ信号の処理対象のフレームについて、時間的および周波数的に近傍にある MDCT 係数またはオーディオ信号の大きさや変動量などに基づいて、処理対象のフレームのアタック性や定常性の判定を行う。すなわち、MDCT 係数の大きさや変動量、オーデ

ィオ信号の大きさや変動量などから、オーディオ信号がアタック性を有するものであるか定常性を有するものであるかが特定される。

- [0122] 時間周波数変換部 21 は、アタック性や定常性の判定結果、処理対象のフレームの時間的に直前のフレームにおける変換窓の選択結果等に基づいて、処理対象のフレームについて変換窓を選択し、選択した変換窓を用いてオーディオ信号の処理対象のフレームに対する時間周波数変換を行う。時間周波数変換部 21 は、時間周波数変換により得られた MDCT 係数を正規化部 22 に供給するとともに、選択した変換窓の種類を示す変換窓情報を符号化方式選択部 24 および多重化部 27 に供給する。
- [0123] ステップ S 12 において正規化部 22 は、ビット制御部 26 から供給されたパラメータに基づいて、時間周波数変換部 21 から供給された MDCT 係数を正規化し、その結果得られた正規化 MDCT 係数を量子化部 23 に供給するとともに、正規化に関するパラメータを多重化部 27 に供給する。
- [0124] ステップ S 13 において量子化部 23 は、正規化部 22 から供給された正規化 MDCT 係数を量子化し、その結果得られた量子化 MDCT 係数を符号化方式選択部 24 に供給する。
- [0125] ステップ S 14 において符号化方式選択部 24 は、時間周波数変換部 21 から供給された変換窓情報により示される変換窓の種類、すなわち window sequence が LONG_STOP_SEQUENCE であるか否かを判定する。
- [0126] ステップ S 14 において LONG_STOP_SEQUENCE であると判定された場合、符号化方式選択部 24 は、量子化部 23 から供給された量子化 MDCT 係数を Huffman 符号化部 31 に供給し、その後、処理はステップ S 15 へと進む。
- [0127] LONG_STOP_SEQUENCE が選択されるフレームは、アタック性が強く、短い変換窓長のフレーム、つまり EIGHT_SHORT_SEQUENCE から、定常性が強く長い変換窓長のフレーム、つまり ONLY_LONG_SEQUENCE へと遷移するときのフレームである。
- [0128] このように短い変換窓長から長い変換窓長へと切り替わったフレーム、つまり LONG_STOP_SEQUENCE が選択されたフレームでは、例えば図 18 を参照し

て説明したように算術符号化よりもハフマン符号化の方が符号化効率が高い。

- [0129] そこで、そのようなフレームの符号化時には、符号化方式としてハフマン符号化方式が選択される。すなわち、MPEG4 AACと同様に同一のハフマンコードブックを使用するセクションごとに、量子化MDCT係数とハフマンコードブック情報が、ハフマン符号が用いられて符号化される。
- [0130] ステップS 15において、ハフマン符号化部31は、符号化方式選択部24から供給された量子化MDCT係数に対して、ハフマンコードブック情報を用いたハフマン符号化を行い、MDCT符号化データおよびハフマンコードブック情報をビット制御部26に供給する。
- [0131] ビット制御部26は、ハフマン符号化部31から供給されたMDCT符号化データおよびハフマンコードブック情報に基づいて目標使用ビット量や音質についての判定を行う。符号化装置11では、目標とするビット量や品質のMDCT符号化データおよびハフマンコードブック情報が得られるまで、パラメータの再設定や、正規化、量子化、ハフマン符号化の各処理が繰り返し行われる。
- [0132] そして、目標とするビット量や品質のMDCT符号化データおよびハフマンコードブック情報が得られると、ビット制御部26は、それらのMDCT符号化データおよびハフマンコードブック情報を多重化部27に供給し、処理はステップS 17へと進む。
- [0133] 一方、ステップS 14においてLONG_STOP_SEQUENCEでないと判定された場合、すなわち、短い変換窓長から長い変換窓長へと切り替えられなかった場合、その後、処理はステップS 16へと進む。この場合、符号化方式選択部24は、量子化部23から供給された量子化MDCT係数を算術符号化部32に供給する。
- [0134] ステップS 16において算術符号化部32は、符号化方式選択部24から供給された量子化MDCT係数に対してコンテキストベースの算術符号化を行い、その結果得られたMDCT符号化データをビット制御部26に供給する。換言

すれば、量子化MDCT係数が算術符号化される。

- [0135] ビット制御部26は、算術符号化部32から供給されたMDCT符号化データに基づいて目標使用ビット量や音質についての判定を行う。符号化装置11では、目標とするビット量や品質のMDCT符号化データが得られるまで、パラメータの再設定や、正規化、量子化、算術符号化の各処理が繰り返し行われる。
- [0136] そして、目標とするビット量や品質のMDCT符号化データが得られると、ビット制御部26は、そのMDCT符号化データを多重化部27に供給し、その後、処理はステップS17へと進む。
- [0137] ステップS15またはステップS16の処理が行われると、ステップS17の処理が行われる。
- [0138] すなわち、ステップS17において多重化部27は、多重化を行って符号化ビットストリームを生成し、得られた符号化ビットストリームを復号装置等に送信（出力）する。
- [0139] 例えばステップS15の処理が行われた場合には、多重化部27はビット制御部26から供給されたMDCT符号化データおよびハフマンコードブック情報と、時間周波数変換部21から供給された変換窓情報と、正規化部22から供給されたパラメータとを多重化し、符号化ビットストリームを生成する。
- [0140] また、例えばステップS16の処理が行われた場合には、多重化部27はビット制御部26から供給されたMDCT符号化データと、時間周波数変換部21から供給された変換窓情報と、正規化部22から供給されたパラメータとを多重化し、符号化ビットストリームを生成する。
- [0141] このようにして得られた符号化ビットストリームが出力されると、符号化処理は終了する。
- [0142] 以上のようにして符号化装置11は、時間周波数変換時に用いた変換窓の種類に応じて符号化方式を選択する。このようにすることで、フレームごとに適切な符号化方式を選択し、符号化効率を向上させることができる。

[0143] 〈復号装置の構成例〉

続いて、符号化装置 11 から出力された符号化ビットストリームを受信して復号を行う復号装置について説明する。

[0144] そのような復号装置は、例えば図 21 に示すように構成される。

[0145] 図 21 に示す復号装置 71 は、取得部 81、非多重化部 82、復号方式選択部 83、復号部 84、逆量子化部 85、および時間周波数逆変換部 86 を有している。

[0146] 取得部 81 は、符号化装置 11 から送信されてきた符号化ビットストリームを受信することで、符号化ビットストリームを取得し、非多重化部 82 に供給する。

[0147] 非多重化部 82 は、取得部 81 から供給された符号化ビットストリームを非多重化し、非多重化により得られた MDCT 符号化データやハフマンコードブック情報を復号方式選択部 83 に供給する。また、非多重化部 82 は、非多重化により得られた、正規化に関するパラメータを逆量子化部 85 に供給するとともに、非多重化により得られた変換窓情報を復号方式選択部 83 および時間周波数逆変換部 86 に供給する。

[0148] 復号方式選択部 83 は、非多重化部 82 から供給された変換窓情報に基づいて復号方式を選択し、非多重化部 82 から供給された MDCT 符号化データ等を、復号方式の選択結果に応じた復号部 84 のブロックへと供給する。

[0149] 復号部 84 は、復号方式選択部 83 から供給された MDCT 符号化データ等の復号を行う。復号部 84 は、ハフマン復号部 91 および算術復号部 92 を有している。

[0150] ハフマン復号部 91 は、復号方式選択部 83 から MDCT 符号化データおよびハフマンコードブック情報が供給された場合、ハフマンコードブック情報を用いて、ハフマン符号化に対応する復号方式で MDCT 符号化データを復号し、その結果得られた量子化 MDCT 係数を逆量子化部 85 に供給する。

[0151] 算術復号部 92 は、復号方式選択部 83 から MDCT 符号化データが供給された場合、算術符号化に対応する復号方式で MDCT 符号化データを復号し、その

結果得られた量子化MDCT係数を逆量子化部85に供給する。

[0152] 逆量子化部85は、非多重化部82から供給されたパラメータを用いて、ハフマン復号部91または算術復号部92から供給された量子化MDCT係数を逆量子化し、その結果得られたMDCT係数を時間周波数逆変換部86に供給する。より詳細には、例えば逆量子化部85は、量子化MDCT係数を逆量子化して得られた値に、さらに非多重化部82から供給されたパラメータ等を乗算するなどしてMDCT係数を得る。

[0153] 時間周波数逆変換部86は、非多重化部82から供給された変換窓情報に基づいて、逆量子化部85から供給されたMDCT係数に対して時間周波数逆変換を行い、その結果得られた時間信号である出力オーディオ信号を後段に出力する。

[0154] 〈復号処理の説明〉

次に、復号装置71の動作について説明する。すなわち、以下、図22のフローチャートを参照して、復号装置71による復号処理について説明する。なお、この復号処理は、取得部81により1フレーム分の符号化ビットストリームが受信されると開始される。

[0155] ステップS41において、非多重化部82は、取得部81から供給された符号化ビットストリームを非多重化し、非多重化により得られたMDCT符号化データ等を復号方式選択部83などに供給する。すなわち、符号化ビットストリームから、MDCT符号化データや変換窓情報、各種のパラメータが抽出される。

[0156] この場合、オーディオ信号（MDCT係数）がハフマン符号化されているときには、MDCT符号化データとハフマンコードブック情報が符号化ビットストリームから抽出される。これに対して、オーディオ信号が算術符号化されているときには、MDCT符号化データが符号化ビットストリームから抽出される。

[0157] また、非多重化部82は、非多重化により得られた、正規化に関するパラメータを逆量子化部85に供給するとともに、非多重化により得られた変換窓情報を復号方式選択部83および時間周波数逆変換部86に供給する。

- [0158] ステップS 4 2において復号方式選択部 8 3は、非多重化部 8 2から供給された変換窓情報により示される変換窓の種類がLONG_STOP_SEQUENCEであるか否かを判定する。
- [0159] ステップS 4 2においてLONG_STOP_SEQUENCEであると判定された場合、復号方式選択部 8 3は、非多重化部 8 2から供給されたMDCT符号化データおよびハフマンコードブック情報をハフマン復号部 9 1へと供給し、処理はステップS 4 3へと進む。
- [0160] この場合、処理対象のフレームは、短い変換窓長のフレームから長い変換窓長のフレームへと切り替わったフレームである。換言すれば、変換窓情報により示される変換窓は、短い変換窓長から長い変換窓長へと切り替えられたときに選択される変換窓である。そのため復号方式選択部 8 3では、復号方式としてハフマン符号化に対応する復号方式が選択される。
- [0161] ステップS 4 3においてハフマン復号部 9 1は、復号方式選択部 8 3から供給されたMDCT符号化データおよびハフマンコードブック情報、すなわちハフマン符号の復号を行う。具体的にはハフマン復号部 9 1は、ハフマンコードブック情報およびMDCT符号化データに基づいて量子化MDCT係数を得る。
- [0162] ハフマン復号部 9 1は、復号により得られた量子化MDCT係数を逆量子化部 8 5に供給し、その後、処理はステップS 4 5へと進む。
- [0163] これに対して、ステップS 4 2においてLONG_STOP_SEQUENCEでないと判定された場合、復号方式選択部 8 3は、非多重化部 8 2から供給されたMDCT符号化データを算術復号部 9 2へと供給し、処理はステップS 4 4へと進む。
- [0164] この場合、処理対象のフレームは、短い変換窓長のフレームから長い変換窓長のフレームへと切り替わったフレームではない。換言すれば、変換窓情報により示される変換窓は、短い変換窓長から長い変換窓長へと切り替えられたときに選択される変換窓ではない。そのため復号方式選択部 8 3では、復号方式として算術符号化に対応する復号方式が選択される。
- [0165] ステップS 4 4において算術復号部 9 2は、復号方式選択部 8 3から供給されたMDCT符号化データ、すなわち算術符号の復号を行う。

- [0166] 算術復号部 9 2 は MDCT 符号化データを復号して得られた量子化 MDCT 係数を逆量子化部 8 5 に供給し、その後、処理はステップ S 4 5 へと進む。
- [0167] ステップ S 4 3 またはステップ S 4 4 の処理が行われると、ステップ S 4 5 の処理が行われる。
- [0168] ステップ S 4 5 において逆量子化部 8 5 は、非多重化部 8 2 から供給されたパラメータを用いて、ハフマン復号部 9 1 または算術復号部 9 2 から供給された量子化 MDCT 係数を逆量子化し、その結果得られた MDCT 係数を時間周波数逆変換部 8 6 に供給する。
- [0169] ステップ S 4 6 において時間周波数逆変換部 8 6 は、非多重化部 8 2 から供給された変換窓情報に基づいて、逆量子化部 8 5 から供給された MDCT 係数に対して時間周波数逆変換を行い、その結果得られた出力オーディオ信号を後段に出力する。
- [0170] 出力オーディオ信号が出力されると、復号処理は終了する。
- [0171] 以上のようにして復号装置 7 1 は、符号化ビットストリームを非多重化して得られた変換窓情報に基づいて復号方式を選択し、選択した復号方式で復号を行う。特に、変換窓の種類が LONG_STOP_SEQUENCE である場合にはハフマン符号化に対応する復号方式が選択され、それ以外の場合には算術符号化に対応する復号方式が選択される。このようにすることで、符号化側において符号化効率が向上するだけでなく、復号側における復号時の処理量（演算量）も削減することができる。
- [0172] ところで、本技術のように LONG_STOP_SEQUENCE のフレームでハフマン符号化を行い、LONG_STOP_SEQUENCE 以外のフレームで算術符号化を行う手法をハイブリッド符号化手法と呼ぶこととする。このようなハイブリッド符号化手法によれば、符号化効率を向上させるとともに復号時の処理量も削減することができる。
- [0173] 例えば図 2 3 は、図 5 における場合と同一の定常性音楽信号を符号化するとき、USAC の LONG_STOP_SEQUENCE のフレームにハフマン符号化を用いた場合、つまりハイブリッド符号化手法により符号化を行った場合と、常に AAC の

ハフマン符号化を用いた場合との必要ビット数の差分グラフを示している。

[0174] なお、図23において横軸はフレーム（時間）を示しており、縦軸は差分ビット数を示している。ここでの差分ビット数は、ハイブリッド符号化手法の必要ビット数から、AACのハフマン符号化の必要ビット数を減算して得られたものである。

[0175] 図23に示す各フレームの差分ビット数は、図7に示した差分ビット数に対応し、これらの図23と図7を比較すると、すなわちハイブリッド符号化手法により符号化を行う場合と、常に算術符号化を行う場合とを比較すると、図23の例の方が符号化効率がよいが、それほど大きな差はないことが分かる。

[0176] これに対して、図12における場合と同一のアタック性音楽信号を符号化するとき、USACのLONG_STOP_SEQUENCEのフレームにハフマン符号化を用いた場合、つまりハイブリッド符号化手法により符号化を行った場合と、常にAACのハフマン符号化を用いた場合との必要ビット数の差分は図24に示すようになる。

[0177] なお、図24において横軸はフレーム（時間）を示しており、縦軸は差分ビット数を示している。ここでの差分ビット数は、ハイブリッド符号化手法の必要ビット数から、AACのハフマン符号化の必要ビット数を減算して得られたものである。

[0178] 図24に示す各フレームの差分ビット数は、図14に示した差分ビット数に対応し、これらの図24と図14を比較すると、すなわちハイブリッド符号化手法により符号化を行う場合と、常に算術符号化を行う場合とを比較すると、図24の例の方が大幅に差分ビット数が減少していることが分かる。つまり、大幅に符号化効率が向上していることが分かる。

[0179] さらに、ハイブリッド符号化手法では、LONG_STOP_SEQUENCEのフレームで算術符号化ではなくハフマン符号化を用いたことにより、そのフレームの復号時の処理量も低減される。

[0180] 〈第2の実施の形態〉

〈符号化方式の選択について〉

ところで、以上においてはLONG_STOP_SEQUENCEではないフレームにおいては、符号化方式として必ず算術符号化が選択されると説明した。しかし、符号化方式を選択するにあたっては、符号化効率（圧縮効率）だけでなく処理量の余裕や音質なども考慮できるとよい。

[0181] そこで、例えばLONG_STOP_SEQUENCEではないフレームにおいては、ハフマン符号化と算術符号化のうちの何れか一方が選択されるようにすることもできる。

[0182] そのような場合、例えば符号化ビットストリームには、符号化時にハフマン符号化と算術符号化のうちの何れの符号化方式が選択されたかを示す判定フラグが格納される。

[0183] ここでは、例えば判定フラグの値が「1」である場合にはハフマン符号化方式が選択されたことを示しており、判定フラグの値が「0」である場合には算術符号化方式が選択されたことを示しているとする。

[0184] このような判定フラグは、LONG_STOP_SEQUENCEではないフレーム、つまり短い変換窓長から長い変換窓長へと切り替えられなかった場合において、処理対象のフレームで選択された符号化方式を示す選択情報であるといえる。換言すれば、判定フラグは符号化方式の選択結果を示す選択情報であるといえる。

[0185] なお、LONG_STOP_SEQUENCEが選択されたフレームにおいては、必ずハフマン符号化方式が選択されるため、そのようなフレームについては符号化ビットストリームには判定フラグは含まれていない。

[0186] 例えば、適宜、判定フラグが符号化ビットストリームに格納される場合、符号化ビットストリームにおける所定のチャンネルのオーディオ信号の1フレーム分のチャンネルストリームのシンタックスは、MPEG-D USACをベースとして図25に示すようにすればよい。

[0187] 図25に示す例では、矢印Q11に示す部分、すなわち文字「ics_info()」の部分は、変換窓等に関する情報が格納されるics_infoを示している。

- [0188] また、矢印Q 1 2に示す文字「section_data()」の部分は、section_dataを示している。このsection_dataにはハフマンコードブック情報等が格納される。さらに図25における文字「ac_spectral_data」はMDCT符号化データを示している。
- [0189] また、文字「ics_info()」により示されるics_infoの部分のシンタックスは、例えば図26に示すようにされる。
- [0190] 図26に示す例では文字「window_sequence」の部分は、変換窓情報、すなわちwindow sequenceを示しており、文字「window_shape」の部分は変換窓の形状を示している。
- [0191] また、文字「huffman_coding_flag」の部分は判定フラグを示している。
- [0192] ここでは文字「window_sequence」の部分に格納される変換窓情報がLONG_STOP_SEQUENCEを示している場合、ics_infoには判定フラグは格納されていない。これに対して変換窓情報がLONG_STOP_SEQUENCE以外の種類を示している場合、ics_infoには判定フラグが格納されている。
- [0193] したがって図25に示した例では、図26の文字「window_sequence」の部分に格納される変換窓情報がLONG_STOP_SEQUENCE以外の種類を示しており、かつ図26の文字「huffman_coding_flag」の部分に値が「1」である判定フラグが格納されている場合、section_dataにハフマンコードブック情報等が格納されている。その他、図26の文字「window_sequence」の部分に格納される変換窓情報がLONG_STOP_SEQUENCEを示している場合にもsection_dataにハフマンコードブック情報等が格納されている。
- [0194] 〈符号化処理の説明〉
- 図25や図26に示した例のように、符号化ビットストリームに適宜、判定フラグが格納される場合、符号化装置11では、例えば図27に示す符号化処理が行われる。以下、図27のフローチャートを参照して、符号化装置11による符号化処理について説明する。
- [0195] なお、ステップS71乃至ステップS75の処理は、図20のステップS11乃至ステップS15の処理と同様であるので、その説明は省略する。

- [0196] ステップS 7 4 においてLONG_STOP_SEQUENCEでないとは判定された場合、ステップS 7 6 において符号化方式選択部 2 4 は、算術符号化を行うか否かを判定する。
- [0197] 例えば符号化方式選択部 2 4 は、上位の制御装置から供給された指定情報に基づいて、算術符号化を行うか否かを判定する。
- [0198] ここで、指定情報は、例えばコンテンツ制作者等が指定した符号化方式を示す情報である。例えばコンテンツ制作者は、LONG_STOP_SEQUENCEでないフレームでは、フレームごとに符号化方式としてハフマン符号化と算術符号化の何れか一方を指定することができるようになっている。
- [0199] この場合、符号化方式選択部 2 4 は、指定情報により示される符号化方式が算術符号化であるとき、ステップS 7 6 において算術符号化を行うと判定する。これに対して、符号化方式選択部 2 4 は、指定情報により示される符号化方式がハフマン符号化であるとき、ステップS 7 6 において算術符号化を行わないと判定する。
- [0200] その他、ステップS 7 6 では、復号装置 7 1 や符号化装置 1 1 のリソース、つまり処理量、符号化されるオーディオ信号のビットレート、リアルタイム性が要求されているか否か等に基づいて符号化方式が選択されるようにしてもよい。
- [0201] 具体的には、例えばオーディオ信号のビットレートが高く、十分な音質を確保することができる場合には、符号化方式選択部 2 4 は、より処理量が少ないハフマン符号化を選択し、ステップS 7 6 において算術符号化を行わないと判定するようにしてもよい。
- [0202] また、例えばリアルタイム性が要求されていたり、復号装置 7 1 のリソースが少なかったりして、音質よりも、より少ない処理量で迅速に符号化および復号の処理を行うことが重要である場合、符号化方式選択部 2 4 はハフマン符号化を選択し、ステップS 7 6 において算術符号化を行わないと判定するようにしてもよい。
- [0203] このようにリアルタイム性が要求される場合や、復号側のリソースが少な

い場合には、符号化方式としてハフマン符号化を選択することで、常に算術符号化を行うときよりも高速に処理（動作）を行うことができるようになる。

- [0204] なお、復号装置 7 1 のリソースについては、例えば復号装置 7 1 が設けられた機器の演算処理能力やメモリ量を示す情報等を、復号装置 7 1 のリソース情報として符号化処理の開始前などに予め復号装置 7 1 から取得しておけばよい。
- [0205] ステップ S 7 6 において算術符号化を行うと判定された場合、符号化方式選択部 2 4 は、量子化部 2 3 から供給された量子化 MDCT 係数を算術符号化部 3 2 に供給し、その後、ステップ S 7 7 の処理が行われる。すなわち、ステップ S 7 7 では、量子化 MDCT 係数に対してコンテキストベースの算術符号化が行われる。
- [0206] なお、ステップ S 7 7 の処理は、図 2 0 のステップ S 1 6 の処理と同様であるので、その説明は省略する。ステップ S 7 7 の処理が行われると、その後、処理はステップ S 7 9 へと進む。
- [0207] これに対して、ステップ S 7 6 において算術符号化を行わないと判定された場合、つまりハフマン符号化を行う場合、符号化方式選択部 2 4 は、量子化部 2 3 から供給された量子化 MDCT 係数をハフマン符号化部 3 1 に供給し、処理はステップ S 7 8 へと進む。
- [0208] ステップ S 7 8 では、ステップ S 7 5 と同様の処理が行われ、その結果得られた MDCT 符号化データおよびハフマンコードブック情報がハフマン符号化部 3 1 からビット制御部 2 6 へと供給される。ステップ S 7 8 の処理が行われると、その後、処理はステップ S 7 9 へと進む。
- [0209] ステップ S 7 7 またはステップ S 7 8 の処理が行われると、ステップ S 7 9 においてビット制御部 2 6 は判定フラグを生成する。
- [0210] 例えばステップ S 7 7 の処理、つまり算術符号化が行われた場合、ビット制御部 2 6 は値が「0」である判定フラグを生成し、算術符号化部 3 2 から供給された MDCT 符号化データとともに、生成した判定フラグを多重化部 2 7

へと供給する。

[0211] また、例えばステップS 7 8の処理、つまりハフマン符号化が行われた場合、ビット制御部2 6は値が「1」である判定フラグを生成し、ハフマン符号化部3 1から供給されたMDCT符号化データおよびハフマンコードブック情報とともに、生成した判定フラグを多重化部2 7へと供給する。

[0212] ステップS 7 9の処理が行われると、その後、処理はステップS 8 0へと進む。

[0213] ステップS 7 5またはステップS 7 9の処理が行われると、ステップS 8 0において多重化部2 7は、多重化を行って符号化ビットストリームを生成し、得られた符号化ビットストリームを復号装置7 1に送信する。なお、ステップS 8 0では、基本的には図2 0のステップS 1 7と同様の処理が行われる。

[0214] 例えばステップS 7 5の処理が行われた場合には、多重化部2 7はMDCT符号化データ、ハフマンコードブック情報、変換窓情報、および正規化部2 2からのパラメータが格納された符号化ビットストリームを生成する。この符号化ビットストリームには、判定フラグは含まれていない。

[0215] また、例えばステップS 7 8の処理が行われた場合、多重化部2 7は判定フラグ、MDCT符号化データ、ハフマンコードブック情報、変換窓情報、および正規化部2 2からのパラメータが格納された符号化ビットストリームを生成する。

[0216] さらに、例えばステップS 7 7の処理が行われた場合、多重化部2 7は判定フラグ、MDCT符号化データ、変換窓情報、および正規化部2 2からのパラメータが格納された符号化ビットストリームを生成する。

[0217] このようにして符号化ビットストリームが生成されて出力されると、符号化処理は終了する。

[0218] 以上のようにして符号化装置1 1は、LONG_STOP_SEQUENCEでないフレームにおいては、ハフマン符号化と算術符号化の何れか一方を選択し、選択した符号化方式により符号化を行う。このようにすることで、フレームごとに適

切な符号化方式を選択し、符号化効率を向上させるとともに、より自由度の高い符号化を実現することができる。

[0219] 〈復号処理の説明〉

また、符号化装置 11 において図 27 を参照して説明した符号化処理が行われた場合、復号装置 71 では図 28 に示す復号処理が行われる。

[0220] 以下、図 28 のフローチャートを参照して、復号装置 71 による復号処理について説明する。なお、ステップ S121 乃至ステップ S123 の処理は、図 22 のステップ S41 乃至ステップ S43 の処理と同様であるので、その説明は省略する。但し、ステップ S121 では、非多重化により符号化ビットストリームから判定フラグが抽出された場合には、その判定フラグが非多重化部 82 から復号方式選択部 83 へと供給される。

[0221] ステップ S122 において LONG_STOP_SEQUENCE でないと判定された場合、ステップ S124 において復号方式選択部 83 は、非多重化部 82 から供給された判定フラグに基づいて、MDCT 符号化データが算術符号であるか否かを判定する。すなわち、MDCT 符号化データの符号化方式が算術符号化であるか否かが判定される。

[0222] 例えば復号方式選択部 83 は、判定フラグの値が「1」である場合、算術符号でない、つまりハフマン符号であると判定し、判定フラグの値が「0」である場合、算術符号であると判定する。このように復号方式選択部 83 では、ハフマン符号化と算術符号化のうち、判定フラグにより示される符号化方式に対応する復号方式が選択される。

[0223] ステップ S124 において算術符号でない、つまりハフマン符号であると判定された場合、復号方式選択部 83 は、非多重化部 82 から供給された MDCT 符号化データおよびハフマンコードブック情報をハフマン復号部 91 へと供給し、処理はステップ S123 へと進む。そして、ステップ S123 ではハフマン符号の復号が行われる。

[0224] これに対して、ステップ S124 において算術符号であると判定された場合、復号方式選択部 83 は、非多重化部 82 から供給された MDCT 符号化デー

タを算術復号部 9 2 へと供給し、処理はステップ S 1 2 5 へと進む。

[0225] ステップ S 1 2 5 では、算術符号である MDCT 符号化データに対して、算術符号化に対応する復号方式で復号が行われるが、ステップ S 1 2 5 の処理は図 2 2 のステップ S 4 4 の処理と同様であるので、その説明は省略する。

[0226] ステップ S 1 2 3 またはステップ S 1 2 5 の処理が行われると、その後、ステップ S 1 2 6 およびステップ S 1 2 7 の処理が行われて復号処理は終了するが、これらの処理は図 2 2 のステップ S 4 5 およびステップ S 4 6 の処理と同様であるので、その説明は省略する。

[0227] 以上のようにして復号装置 7 1 は、変換窓情報と判定フラグに基づいて復号方式を選択し、復号を行う。特に、LONG_STOP_SEQUENCE でないフレームでも、判定フラグを参照することで正しい復号方式を選択することができるので、符号化効率の向上や復号側での処理量の削減だけでなく、自由度の高い符号化および復号を実現することができる。

[0228] 〈第 3 の実施の形態〉

〈符号化処理の説明〉

その他、LONG_STOP_SEQUENCE ではないフレームにおいて、ハフマン符号化と算術符号化のうちの何れか一方を選択する場合、より必要ビット数が少ない方を選択するようにしてもよい。

[0229] 例えば復号装置 7 1 や符号化装置 1 1 で処理量に余裕があり、符号化効率（圧縮効率）を優先したい場合などにおいては、LONG_STOP_SEQUENCE ではないフレームで、ハフマン符号化と算術符号化の必要ビット数を計算し、より必要ビット数が少ない方を選択するようにしてもよい。

[0230] そのような場合、符号化装置 1 1 では、例えば図 2 9 に示す符号化処理が行われる。すなわち、以下、図 2 9 のフローチャートを参照して、符号化装置 1 1 による符号化処理について説明する。

[0231] なお、ステップ S 1 5 1 乃至ステップ S 1 5 5 の処理は、図 2 0 のステップ S 1 1 乃至ステップ S 1 5 の処理と同様であるので、その説明は省略する。

- [0232] ステップS 1 5 4においてLONG_STOP_SEQUENCEでないとは判定された場合、符号化方式選択部2 4は、量子化部2 3から供給された量子化MDCT係数をハフマン符号化部3 1および算術符号化部3 2の両方に供給し、処理はステップS 1 5 6へと進む。この場合、ステップS 1 5 4の時点では、まだ何れの符号化方式が選択（採用）されるかは、決定されていない。
- [0233] ステップS 1 5 6において、算術符号化部3 2は符号化方式選択部2 4から供給された量子化MDCT係数に対してコンテキストベースの算術符号化を行い、その結果得られたMDCT符号化データをビット制御部2 6に供給する。ステップS 1 5 6では、図20のステップS 1 6と同様の処理が行われる。
- [0234] ステップS 1 5 7において、ハフマン符号化部3 1は符号化方式選択部2 4から供給された量子化MDCT係数に対してハフマン符号化を行い、その結果得られたMDCT符号化データおよびハフマンコードブック情報をビット制御部2 6に供給する。ステップS 1 5 7では、ステップS 1 5 5と同様の処理が行われる。
- [0235] ステップS 1 5 8においてビット制御部2 6は、ハフマン符号化部3 1から供給されたMDCT符号化データおよびハフマンコードブック情報のビット数と、算術符号化部3 2から供給されたMDCT符号化データのビット数とを比較して符号化方式を選択する。
- [0236] すなわち、ビット制御部2 6はハフマン符号化により得られたMDCT符号化データおよびハフマンコードブック情報のビット数（符号量）が、算術符号化により得られたMDCT符号化データのビット数よりも少ない場合、符号化方式としてハフマン符号化を選択する。
- [0237] この場合、ビット制御部2 6は、ハフマン符号化により得られたMDCT符号化データおよびハフマンコードブック情報を多重化部2 7へと供給する。
- [0238] これに対して、ビット制御部2 6は算術符号化により得られたMDCT符号化データのビット数が、ハフマン符号化により得られたMDCT符号化データおよびハフマンコードブック情報のビット数以下である場合、符号化方式として算術符号化を選択する。

- [0239] この場合、ビット制御部 26 は、算術符号化により得られた MDCT 符号化データを多重化部 27 へと供給する。
- [0240] このように、ハフマン符号化と算術符号化とで実際のビット数（符号量）、すなわちそれらの符号化方式での必要ビット数を比較することで、より必要ビット数の少ない符号化方式を確実に選択することができる。この場合、実質的にはハフマン符号化時の必要ビット数と算術符号化時の必要ビット数とに基づいて、ハフマン符号化と算術符号化のうちの何れかの符号化方式を選択し、選択した符号化方式で符号化を行ったことになる。
- [0241] ステップ S 159 においてビット制御部 26 は、ステップ S 158 における符号化方式の選択結果に応じた判定フラグを生成し、多重化部 27 に供給する。
- [0242] 例えばビット制御部 26 は、符号化方式としてハフマン符号化を選択した場合、値が「1」である判定フラグを生成し、符号化方式として算術符号化を選択した場合、値が「0」である判定フラグを生成する。
- [0243] このようにして判定フラグが生成されると、その後、処理はステップ S 160 へと進む。
- [0244] ステップ S 159 の処理が行われたか、またはステップ S 155 の処理が行われると、ステップ S 160 の処理が行われて符号化処理は終了する。なお、ステップ S 160 の処理は図 27 のステップ S 80 の処理と同様であるので、その説明は省略する。
- [0245] 以上のようにして符号化装置 11 は、LONG_STOP_SEQUENCE でないフレームにおいてはハフマン符号化と算術符号化のうちの必要ビット数の少ない方を選択し、選択した符号化方式で符号化された MDCT 符号化データを含む符号化ビットストリームを生成する。このようにすることで、フレームごとに適切な符号化方式を選択し、符号化効率を向上させるとともに、より自由度の高い符号化を実現することができる。
- [0246] また、図 29 を参照して説明した符号化処理が行われた場合、復号装置 71 では、図 28 を参照して説明した復号処理が行われる。

[0247] 以上のように本技術によれば、適切に符号化方式を選択することで算術符号化のみを用いた場合よりも符号化効率（圧縮効率）を向上させるとともに、復号時の処理量を削減することができる。

[0248] また、第2の実施の形態や第3の実施の形態では、例えばオーディオ信号のビットレートが高く音質が十分である場合や、音質よりも処理量が重要な場合などにおいても、LONG_STOP_SEQUENCEでないフレームにおいて適切な符号化方式を選択することができる。これにより、より自由度の高い符号化および復号を実現することができる。すなわち、例えば復号時の処理量をより柔軟に制御することができる。

[0249] 〈コンピュータの構成例〉

ところで、上述した一連の処理は、ハードウェアにより実行することもできるし、ソフトウェアにより実行することもできる。一連の処理をソフトウェアにより実行する場合には、そのソフトウェアを構成するプログラムが、コンピュータにインストールされる。ここで、コンピュータには、専用のハードウェアに組み込まれているコンピュータや、各種のプログラムをインストールすることで、各種の機能を実行することが可能な、例えば汎用のパーソナルコンピュータなどが含まれる。

[0250] 図30は、上述した一連の処理をプログラムにより実行するコンピュータのハードウェアの構成例を示すブロック図である。

[0251] コンピュータにおいて、CPU (Central Processing Unit) 501, ROM (Read Only Memory) 502, RAM (Random Access Memory) 503は、バス504により相互に接続されている。

[0252] バス504には、さらに、入出力インターフェース505が接続されている。入出力インターフェース505には、入力部506、出力部507、記録部508、通信部509、及びドライブ510が接続されている。

[0253] 入力部506は、キーボード、マウス、マイクロホン、撮像素子などよりなる。出力部507は、ディスプレイ、スピーカなどよりなる。記録部508は、ハードディスクや不揮発性のメモリなどよりなる。通信部509は、

ネットワークインターフェースなどよりなる。ドライブ510は、磁気ディスク、光ディスク、光磁気ディスク、又は半導体メモリなどのリムーバブル記録媒体511を駆動する。

[0254] 以上のように構成されるコンピュータでは、CPU501が、例えば、記録部508に記録されているプログラムを、入出力インターフェース505及びバス504を介して、RAM503にロードして実行することにより、上述した一連の処理が行われる。

[0255] コンピュータ（CPU501）が実行するプログラムは、例えば、パッケージメディア等としてのリムーバブル記録媒体511に記録して提供することができる。また、プログラムは、ローカルエリアネットワーク、インターネット、デジタル衛星放送といった、有線または無線の伝送媒体を介して提供することができる。

[0256] コンピュータでは、プログラムは、リムーバブル記録媒体511をドライブ510に装着することにより、入出力インターフェース505を介して、記録部508にインストールすることができる。また、プログラムは、有線または無線の伝送媒体を介して、通信部509で受信し、記録部508にインストールすることができる。その他、プログラムは、ROM502や記録部508に、あらかじめインストールしておくことができる。

[0257] なお、コンピュータが実行するプログラムは、本明細書で説明する順序に沿って時系列に処理が行われるプログラムであっても良いし、並列に、あるいは呼び出しが行われたとき等の必要なタイミングで処理が行われるプログラムであっても良い。

[0258] また、本技術の実施の形態は、上述した実施の形態に限定されるものではなく、本技術の要旨を逸脱しない範囲において種々の変更が可能である。

[0259] 例えば、本技術は、1つの機能をネットワークを介して複数の装置で分担、共同して処理するクラウドコンピューティングの構成をとることができる。

[0260] また、上述のフローチャートで説明した各ステップは、1つの装置で実行

する他、複数の装置で分担して実行することができる。

[0261] さらに、1つのステップに複数の処理が含まれる場合には、その1つのステップに含まれる複数の処理は、1つの装置で実行する他、複数の装置で分担して実行することができる。

[0262] さらに、本技術は、以下の構成とすることも可能である。

[0263] (1)

オーディオ信号に対して変換窓を用いた時間周波数変換を行う時間周波数変換部と、

前記変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替えられた場合、前記時間周波数変換により得られた周波数スペクトル情報をハフマン符号化し、前記変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替えられなかった場合、前記周波数スペクトル情報を算術符号化する符号化部と

を備える符号化装置。

(2)

前記周波数スペクトル情報の符号化データと、前記時間周波数変換に用いた前記変換窓の種類を示す変換窓情報とを多重化して符号化ビットストリームを生成する多重化部をさらに備える

(1)に記載の符号化装置。

(3)

前記符号化部は、前記変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替えられなかった場合、ハフマン符号化と算術符号化のうちの何れかの符号化方式で前記周波数スペクトル情報を符号化する

(1)または(2)に記載の符号化装置。

(4)

前記符号化部は、前記変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替えられなかった場合、符号化時の必要ビット数、前記オーディオ信号のビットレート、復号側のリソース情報、または符号化方式の指定情報

に基づいて選択された符号化方式で前記周波数スペクトル情報を符号化する

(3)に記載の符号化装置。

(5)

前記多重化部は、前記変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替えられなかった場合、前記周波数スペクトル情報の符号化方式を示す選択情報と、前記符号化データと、前記変換窓情報とを多重化して前記符号化ビットストリームを生成する

(3)または(4)に記載の符号化装置。

(6)

符号化装置が、

オーディオ信号に対して変換窓を用いた時間周波数変換を行い、

前記変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替えられた場合、前記時間周波数変換により得られた周波数スペクトル情報をハフマン符号化し、

前記変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替えられなかった場合、前記周波数スペクトル情報を算術符号化する

符号化方法。

(7)

オーディオ信号に対して変換窓を用いた時間周波数変換を行い、

前記変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替えられた場合、前記時間周波数変換により得られた周波数スペクトル情報をハフマン符号化し、

前記変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替えられなかった場合、前記周波数スペクトル情報を算術符号化する

ステップを含む処理をコンピュータに実行させるプログラム。

(8)

符号化ビットストリームを非多重化して、オーディオ信号の時間周波数変換に用いた変換窓の種類を示す変換窓情報と、前記時間周波数変換により得

られた周波数スペクトル情報の符号化データとを抽出する非多重化部と、
前記変換窓情報により示される前記変換窓が、短い変換窓長から長い変換窓長へと切り替えられたときに選択される前記変換窓である場合、前記符号化データをハフマン符号化に対応する復号方式で復号する復号部と
を備える復号装置。

(9)

前記復号部は、前記変換窓情報により示される前記変換窓が、短い変換窓長から長い変換窓長へと切り替えられたときに選択される前記変換窓でない場合、前記符号化データを算術符号化に対応する復号方式で復号する

(8)に記載の復号装置。

(10)

前記復号部は、前記変換窓情報により示される前記変換窓が、短い変換窓長から長い変換窓長へと切り替えられたときに選択される前記変換窓でない場合、ハフマン符号化と算術符号化のうち、前記符号化ビットストリームから抽出された選択情報により示される符号化方式に対応する復号方式で前記符号化データを復号する

(8)に記載の復号装置。

(11)

復号装置が、

符号化ビットストリームを非多重化して、オーディオ信号の時間周波数変換に用いた変換窓の種類を示す変換窓情報と、前記時間周波数変換により得られた周波数スペクトル情報の符号化データとを抽出し、

前記変換窓情報により示される前記変換窓が、短い変換窓長から長い変換窓長へと切り替えられたときに選択される前記変換窓である場合、前記符号化データをハフマン符号化に対応する復号方式で復号する

復号方法。

(12)

符号化ビットストリームを非多重化して、オーディオ信号の時間周波数変

換に用いた変換窓の種類を示す変換窓情報と、前記時間周波数変換により得られた周波数スペクトル情報の符号化データとを抽出し、

前記変換窓情報により示される前記変換窓が、短い変換窓長から長い変換窓長へと切り替えられたときに選択される前記変換窓である場合、前記符号化データをハフマン符号化に対応する復号方式で復号する

ステップを含む処理をコンピュータに実行させるプログラム。

符号の説明

[0264] 11 符号化装置, 21 時間周波数変換部, 24 符号化方式選択部, 26 ビット制御部, 27 多重化部, 31 ハフマン符号化部, 32 算術符号化部, 71 復号装置, 81 取得部, 82 非多重化部, 83 復号方式選択部, 91 ハフマン復号部, 92 算術復号部

請求の範囲

- [請求項1] オーディオ信号に対して変換窓を用いた時間周波数変換を行う時間周波数変換部と、
- 前記変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替えられた場合、前記時間周波数変換により得られた周波数スペクトル情報をハフマン符号化し、前記変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替えられなかった場合、前記周波数スペクトル情報を算術符号化する符号化部と
- を備える符号化装置。
- [請求項2] 前記周波数スペクトル情報の符号化データと、前記時間周波数変換に用いた前記変換窓の種類を示す変換窓情報とを多重化して符号化ビットストリームを生成する多重化部をさらに備える
- 請求項1に記載の符号化装置。
- [請求項3] 前記符号化部は、前記変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替えられなかった場合、ハフマン符号化と算術符号化のうちの何れかの符号化方式で前記周波数スペクトル情報を符号化する
- 請求項1に記載の符号化装置。
- [請求項4] 前記符号化部は、前記変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替えられなかった場合、符号化時の必要ビット数、前記オーディオ信号のビットレート、復号側のリソース情報、または符号化方式の指定情報に基づいて選択された符号化方式で前記周波数スペクトル情報を符号化する
- 請求項3に記載の符号化装置。
- [請求項5] 前記多重化部は、前記変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替えられなかった場合、前記周波数スペクトル情報の符号化方式を示す選択情報と、前記符号化データと、前記変換窓情報とを多重化して前記符号化ビットストリームを生成する

請求項3に記載の符号化装置。

[請求項6]

符号化装置が、

オーディオ信号に対して変換窓を用いた時間周波数変換を行い、
前記変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替
えられた場合、前記時間周波数変換により得られた周波数スペクトル
情報をハフマン符号化し、

前記変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替
えられなかった場合、前記周波数スペクトル情報を算術符号化する
符号化方法。

[請求項7]

オーディオ信号に対して変換窓を用いた時間周波数変換を行い、

前記変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替
えられた場合、前記時間周波数変換により得られた周波数スペクトル
情報をハフマン符号化し、

前記変換窓の変換窓長が短い変換窓長から長い変換窓長へと切り替
えられなかった場合、前記周波数スペクトル情報を算術符号化する
ステップを含む処理をコンピュータに実行させるプログラム。

[請求項8]

符号化ビットストリームを非多重化して、オーディオ信号の時間周
波数変換に用いた変換窓の種類を示す変換窓情報と、前記時間周波数
変換により得られた周波数スペクトル情報の符号化データとを抽出す
る非多重化部と、

前記変換窓情報により示される前記変換窓が、短い変換窓長から長
い変換窓長へと切り替えられたときに選択される前記変換窓である場
合、前記符号化データをハフマン符号化に対応する復号方式で復号す
る復号部と

を備える復号装置。

[請求項9]

前記復号部は、前記変換窓情報により示される前記変換窓が、短い
変換窓長から長い変換窓長へと切り替えられたときに選択される前記
変換窓でない場合、前記符号化データを算術符号化に対応する復号方

式で復号する

請求項 8 に記載の復号装置。

[請求項10] 前記復号部は、前記変換窓情報により示される前記変換窓が、短い変換窓長から長い変換窓長へと切り替えられたときに選択される前記変換窓でない場合、ハフマン符号化と算術符号化のうち、前記符号化ビットストリームから抽出された選択情報により示される符号化方式に対応する復号方式で前記符号化データを復号する

請求項 8 に記載の復号装置。

[請求項11] 復号装置が、

符号化ビットストリームを非多重化して、オーディオ信号の時間周波数変換に用いた変換窓の種類を示す変換窓情報と、前記時間周波数変換により得られた周波数スペクトル情報の符号化データとを抽出し、

前記変換窓情報により示される前記変換窓が、短い変換窓長から長い変換窓長へと切り替えられたときに選択される前記変換窓である場合、前記符号化データをハフマン符号化に対応する復号方式で復号する

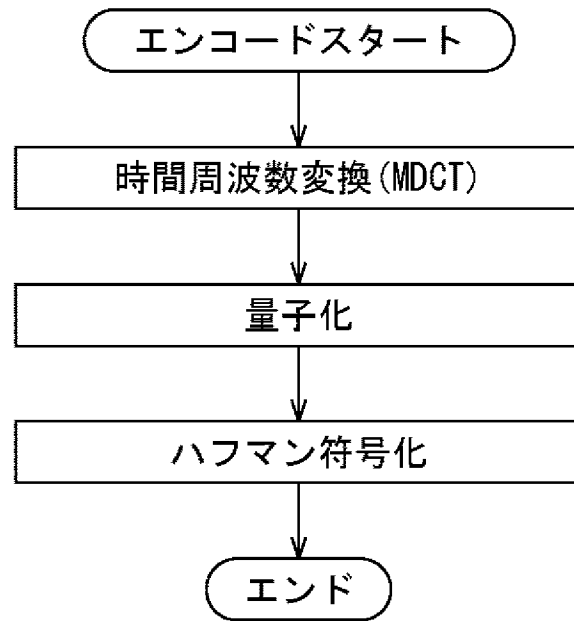
復号方法。

[請求項12] 符号化ビットストリームを非多重化して、オーディオ信号の時間周波数変換に用いた変換窓の種類を示す変換窓情報と、前記時間周波数変換により得られた周波数スペクトル情報の符号化データとを抽出し、





前記変換窓情報により示される前記変換窓が、短い変換窓長から長い変換窓長へと切り替えられたときに選択される前記変換窓である場合、前記符号化データをハフマン符号化に対応する復号方式で復号する

ステップを含む処理をコンピュータに実行させるプログラム。

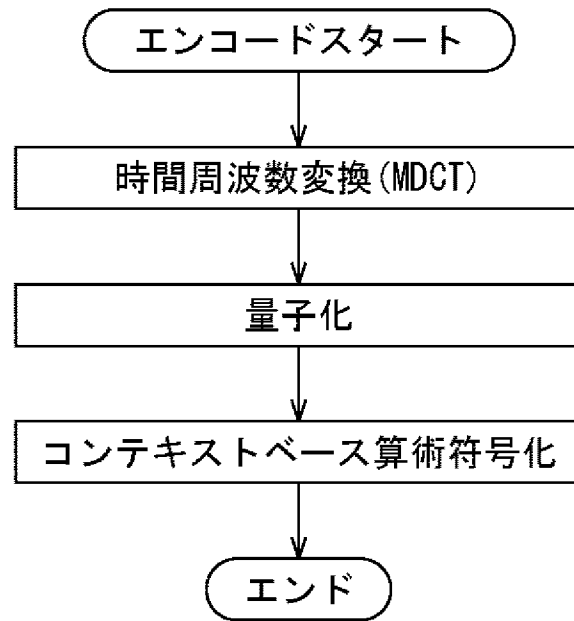
[図1]
FIG. 1



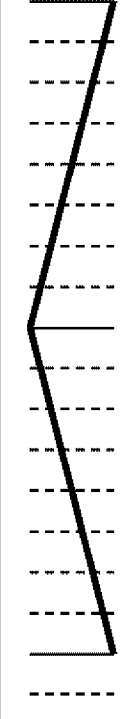
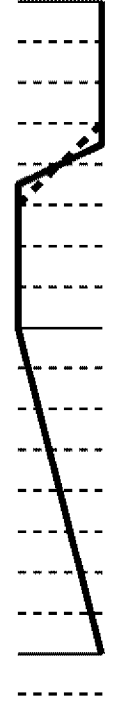
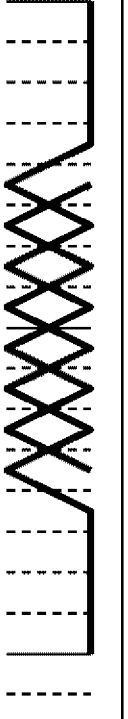
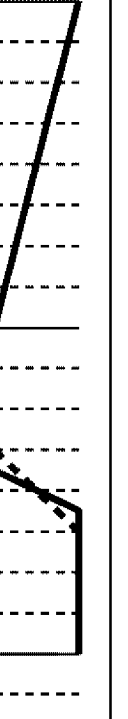

[図2]
FIG. 2

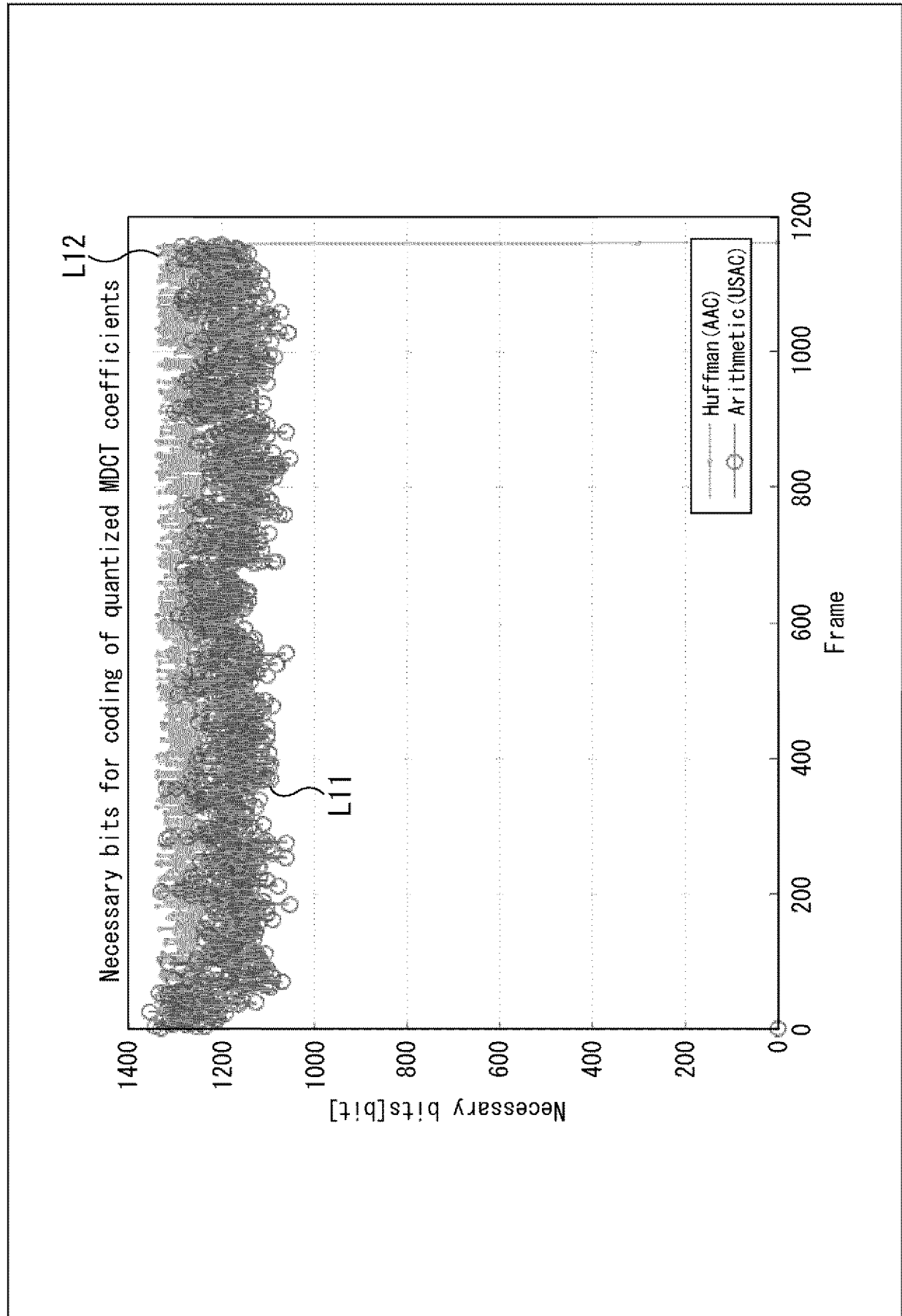
Value	window_sequence	num_ windows	looks like
0	ONLY_LONG_SEQUENCE =LONG_WINDOW	1	
1	LONG_START_SEQUENCE =LONG_START_WINDOW	1	
2	EIGHT_SHORT_SEQUENCE =8*SHORT_WINDOW	8	
3	LONG_STOP_SEQUENCE =LONG_STOP_WINDOW	1	

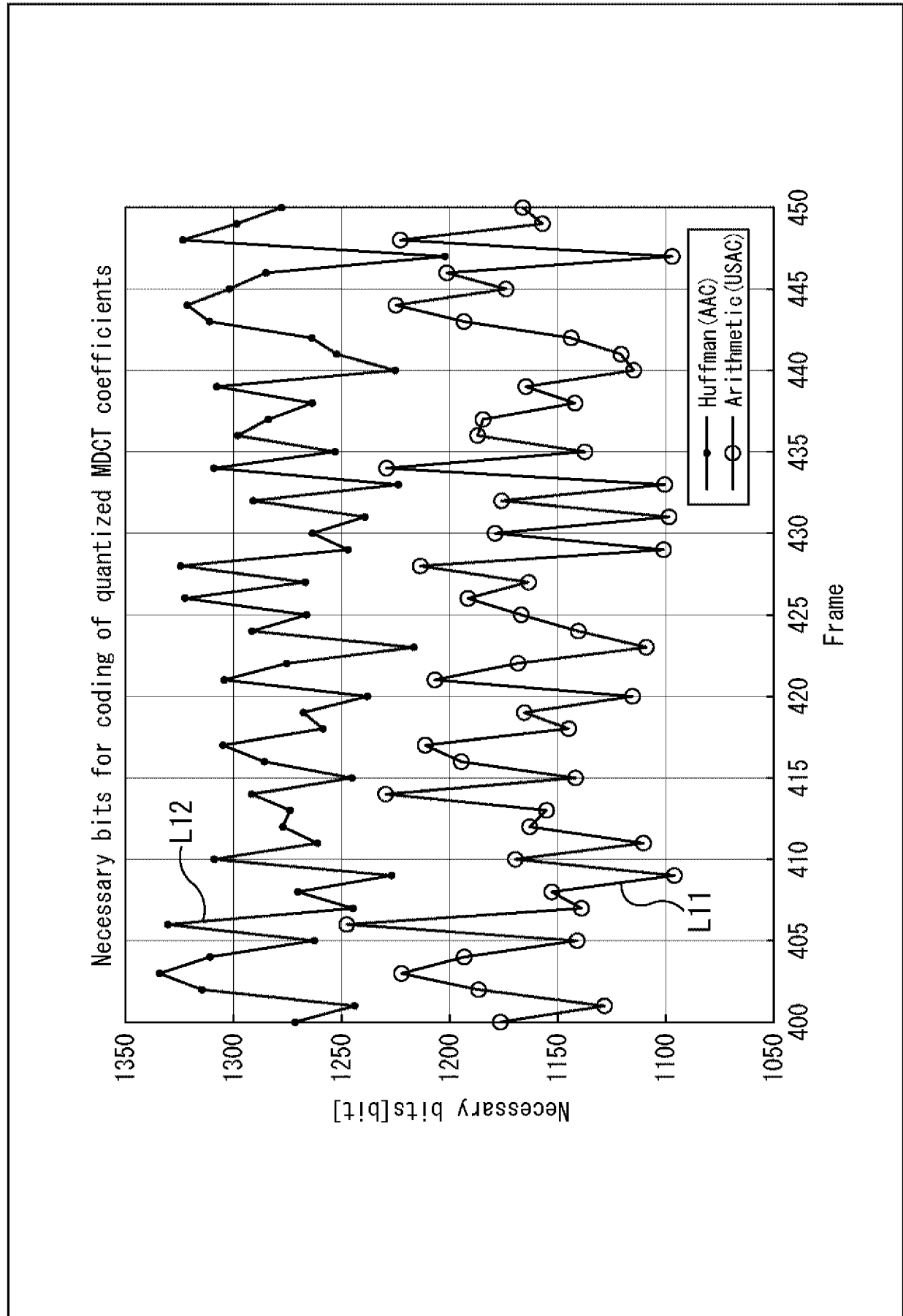
[図3]
FIG. 3



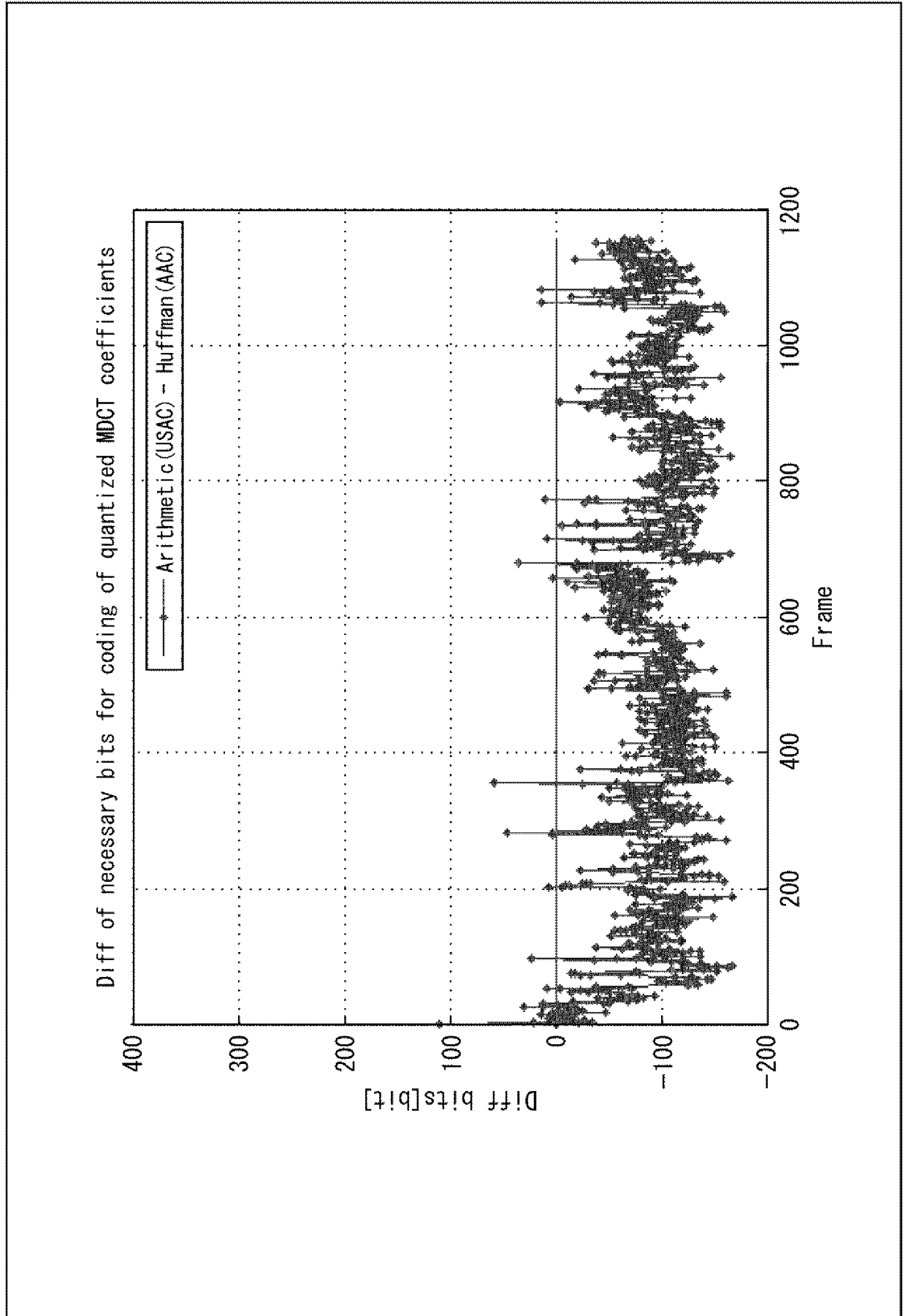
[図4]
FIG. 4

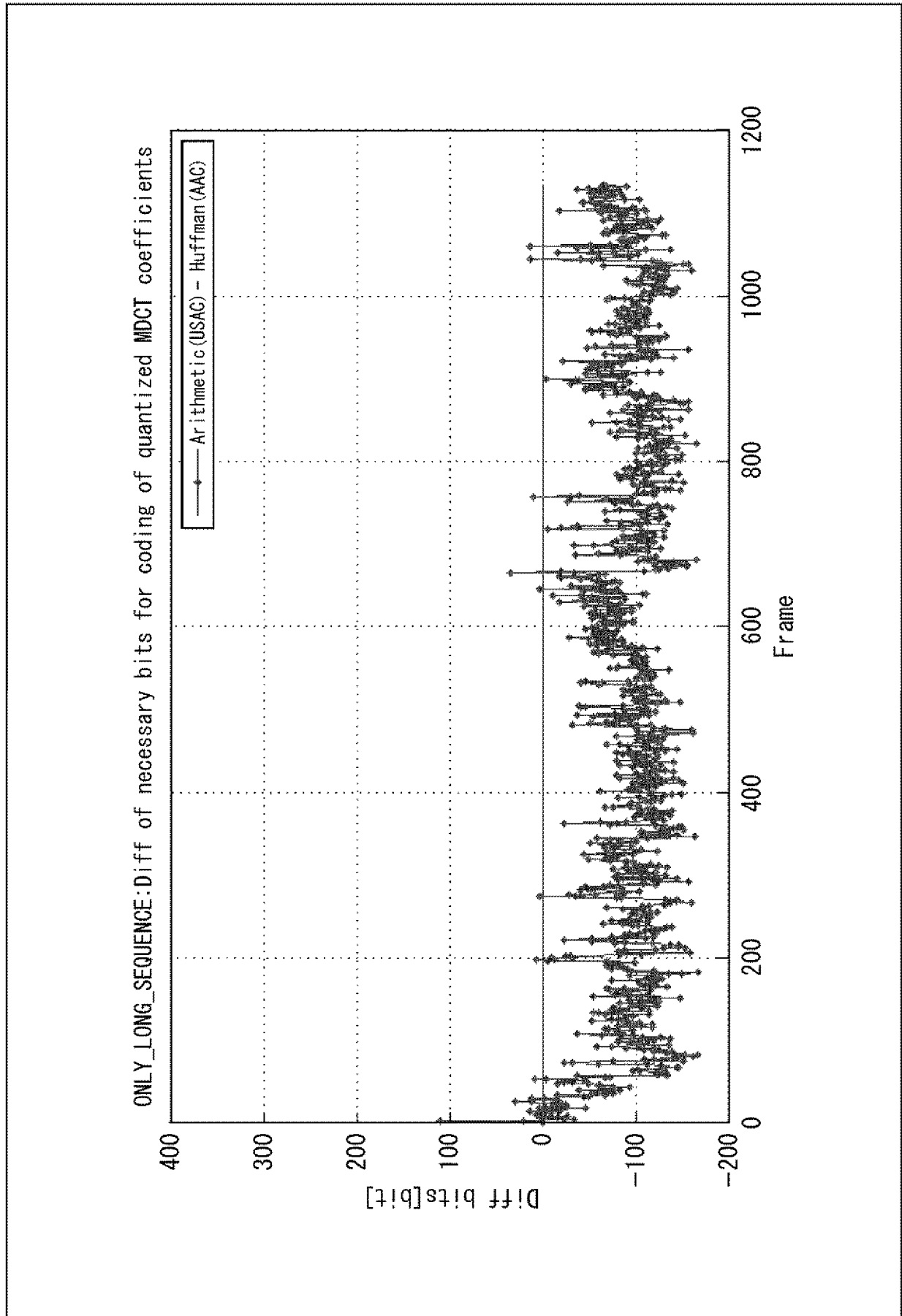
Value	Window	num_ windows	Window Shape (schematic)
0	ONLY_LONG_SEQUENCE =LONG_WINDOW	1	
1	LONG_START_SEQUENCE =LONG_START_WINDOW	1	
2	EIGHT_SHORT_SEQUENCE =8*SHORT_WINDOW	8	
3	LONG_STOP_SEQUENCE =LONG_STOP_WINDOW	1	
1	STOP_START_SEQUENCE =STOP_START_WINDOW	1	

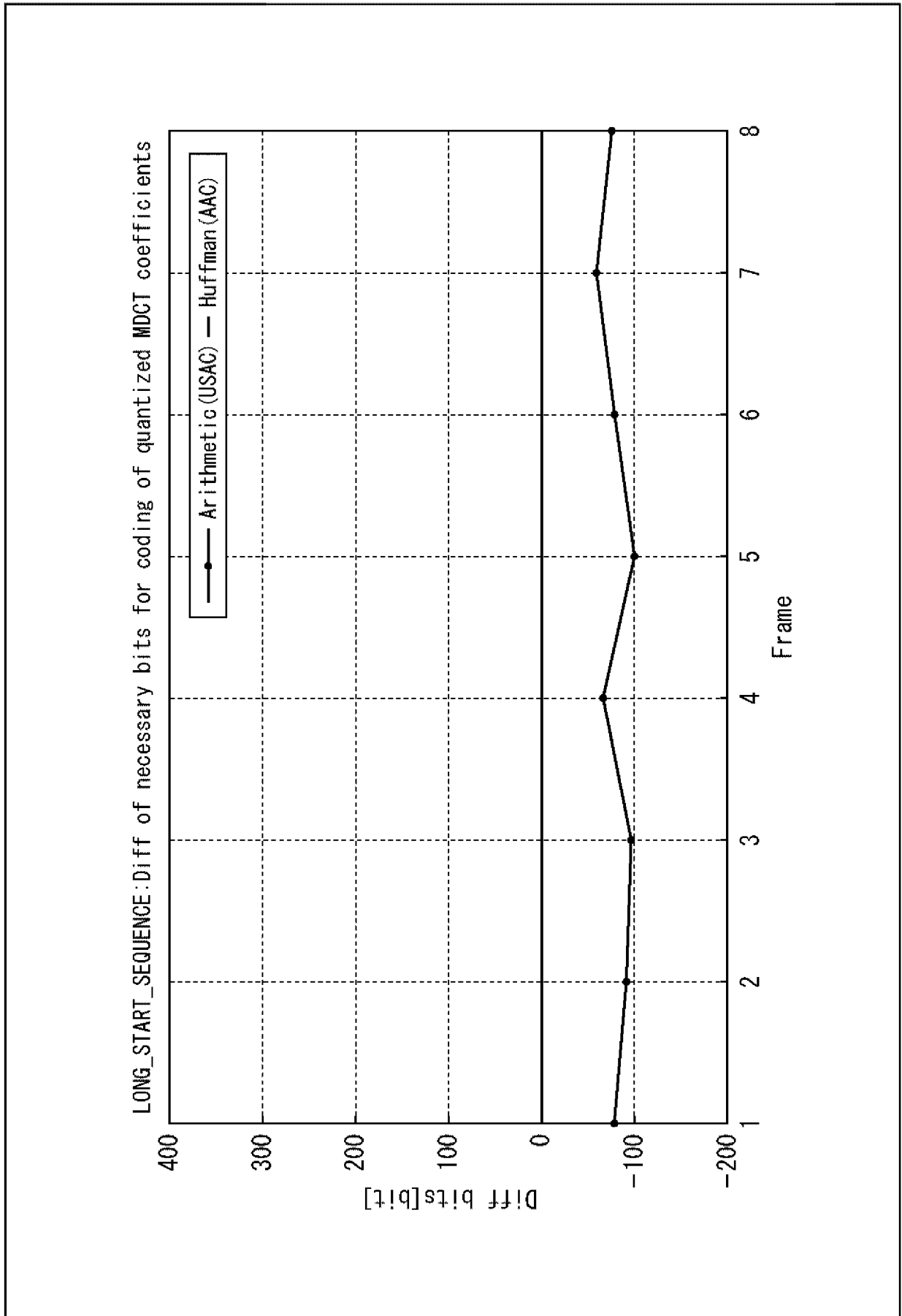
[図5]
FIG. 5

[図6]
FIG. 6

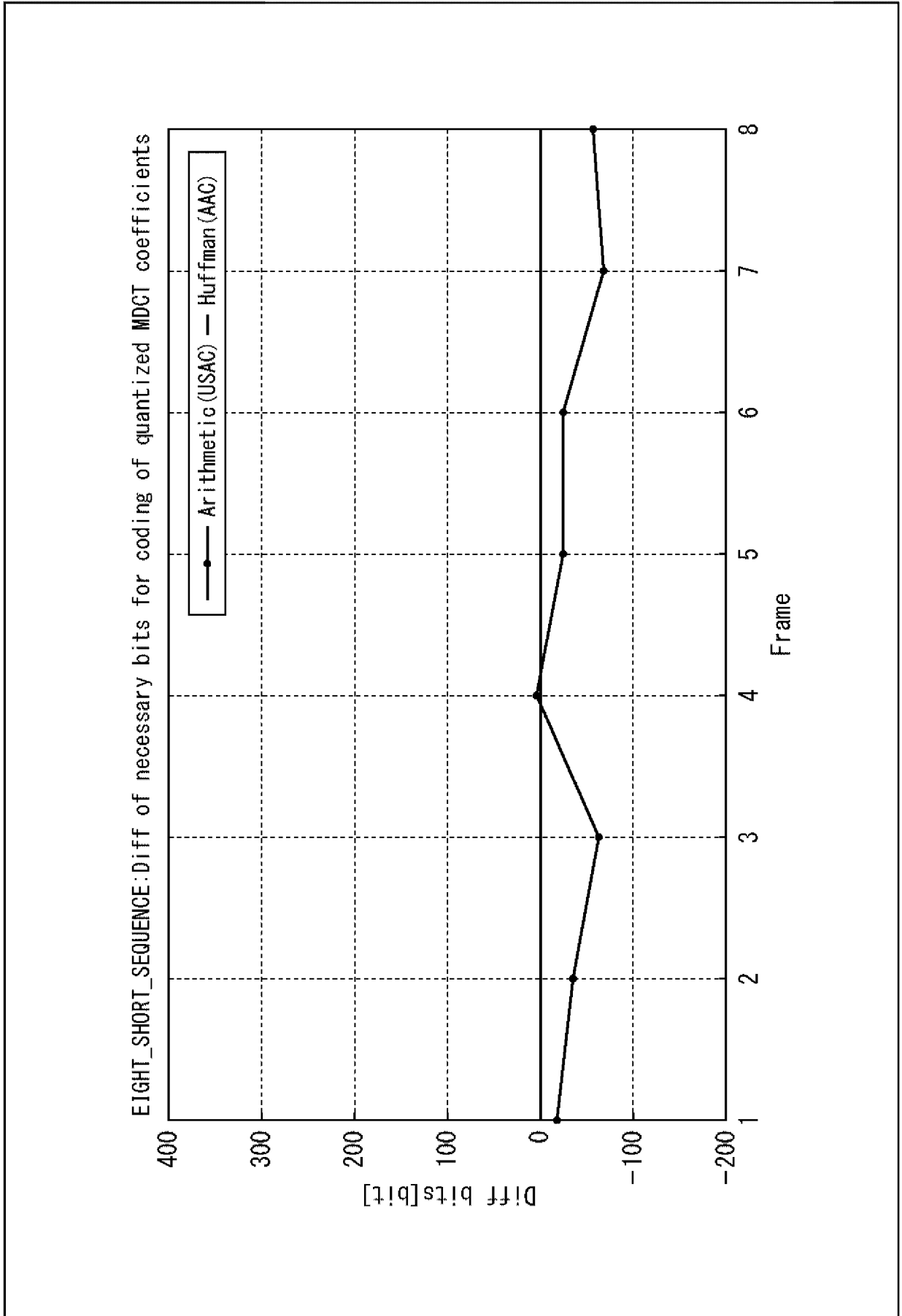
[図7]
FIG. 7



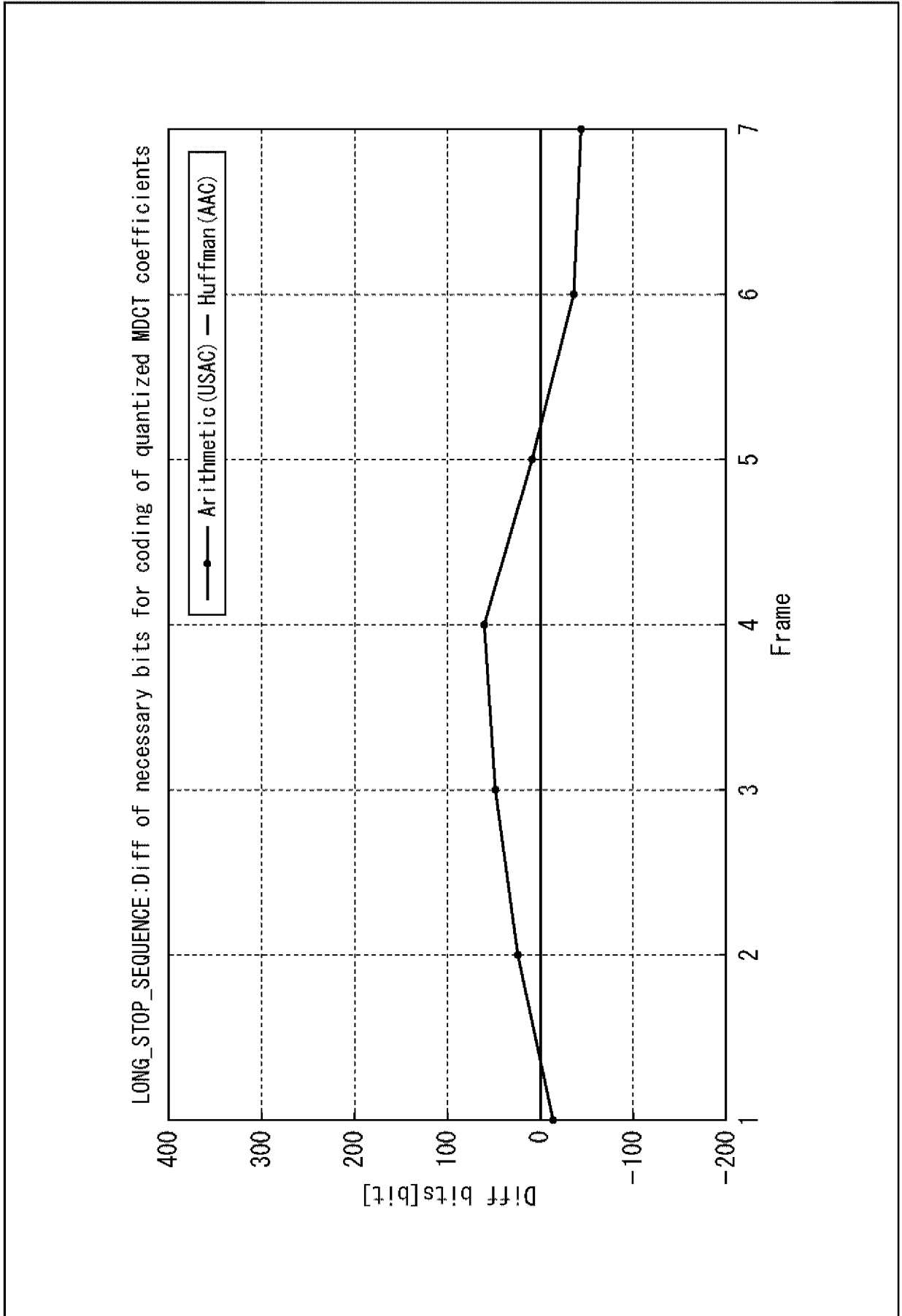
[8]
FIG. 8

[9]
FIG. 9

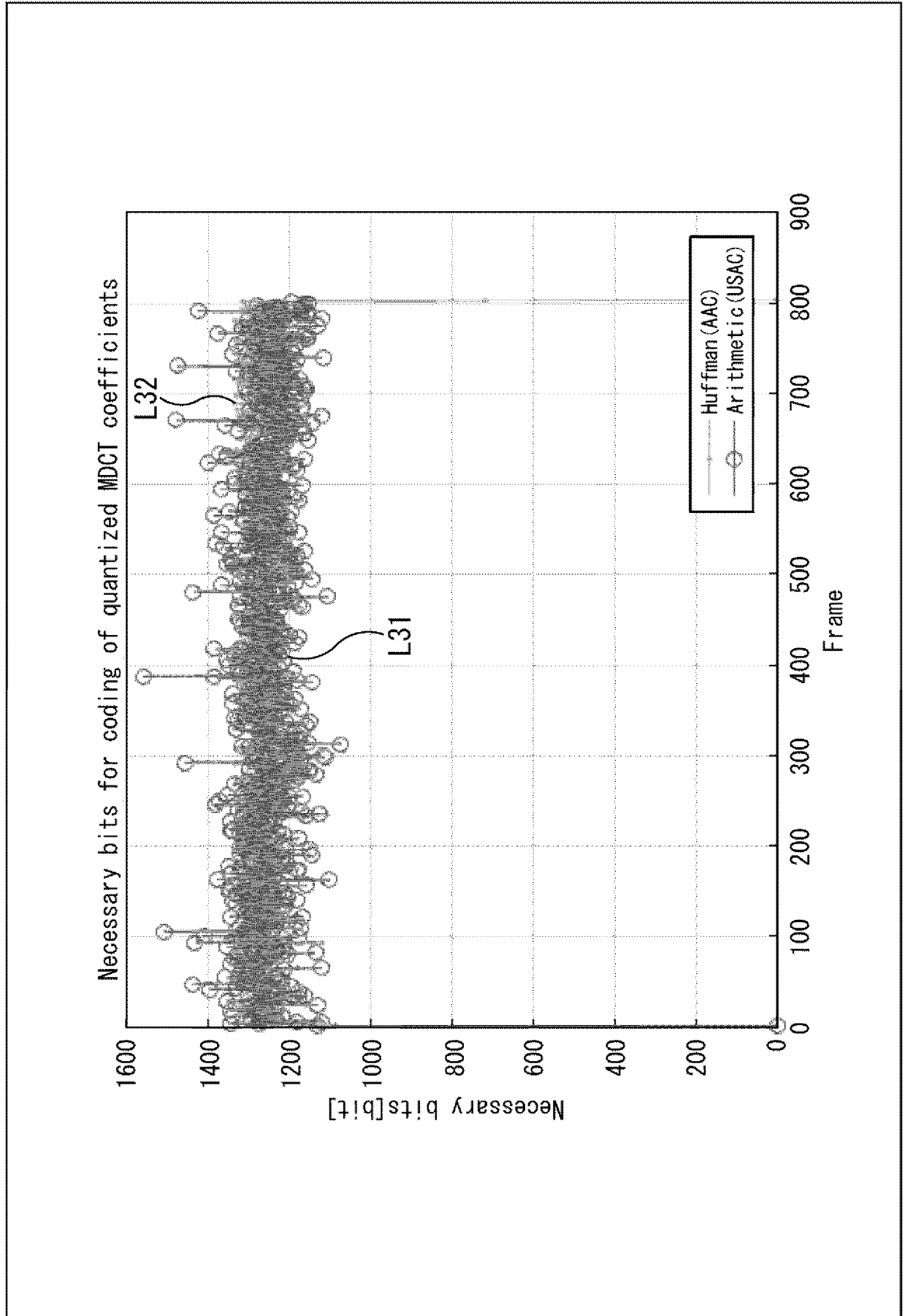
[図10]
FIG. 10



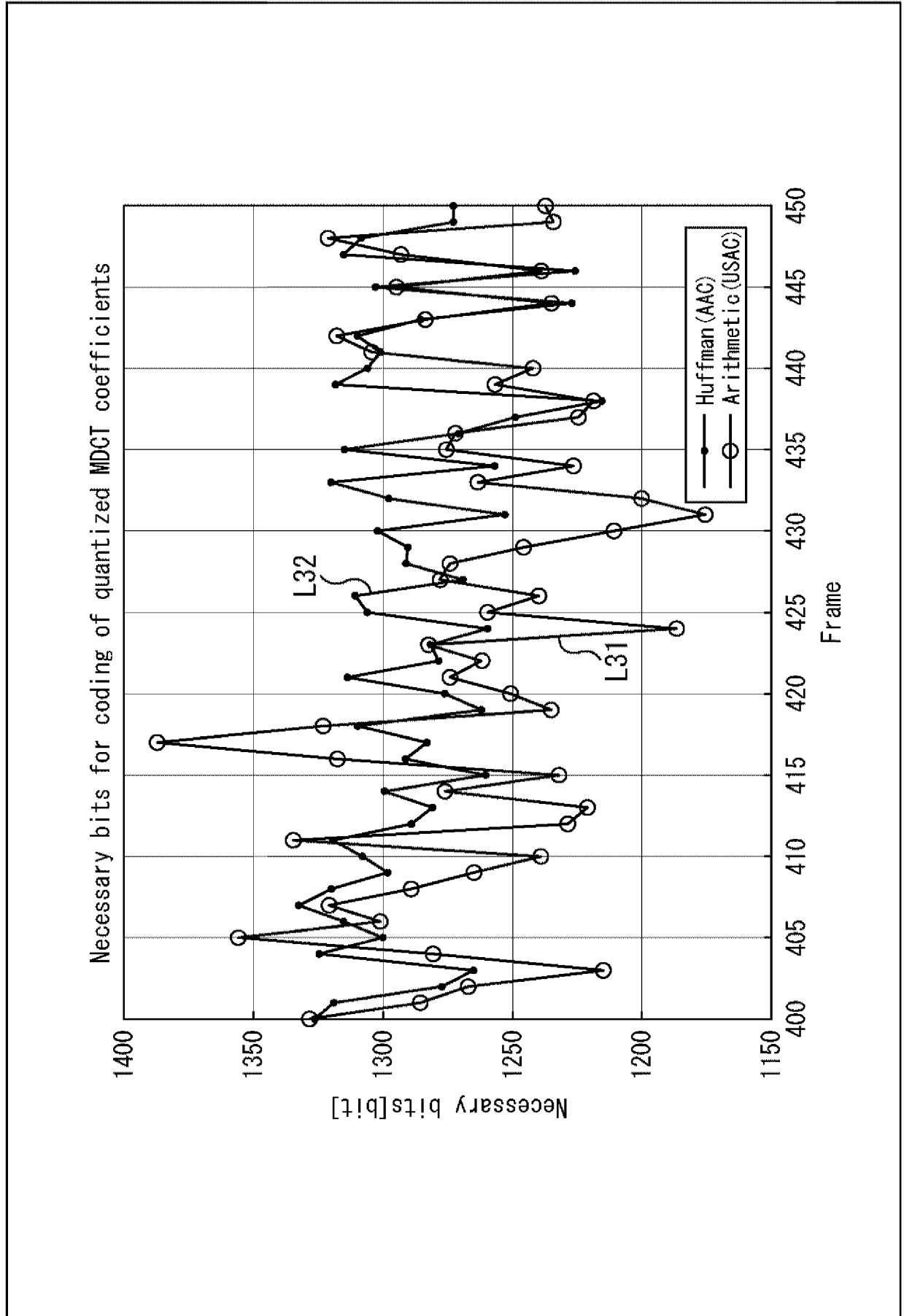
[11]
FIG. 11



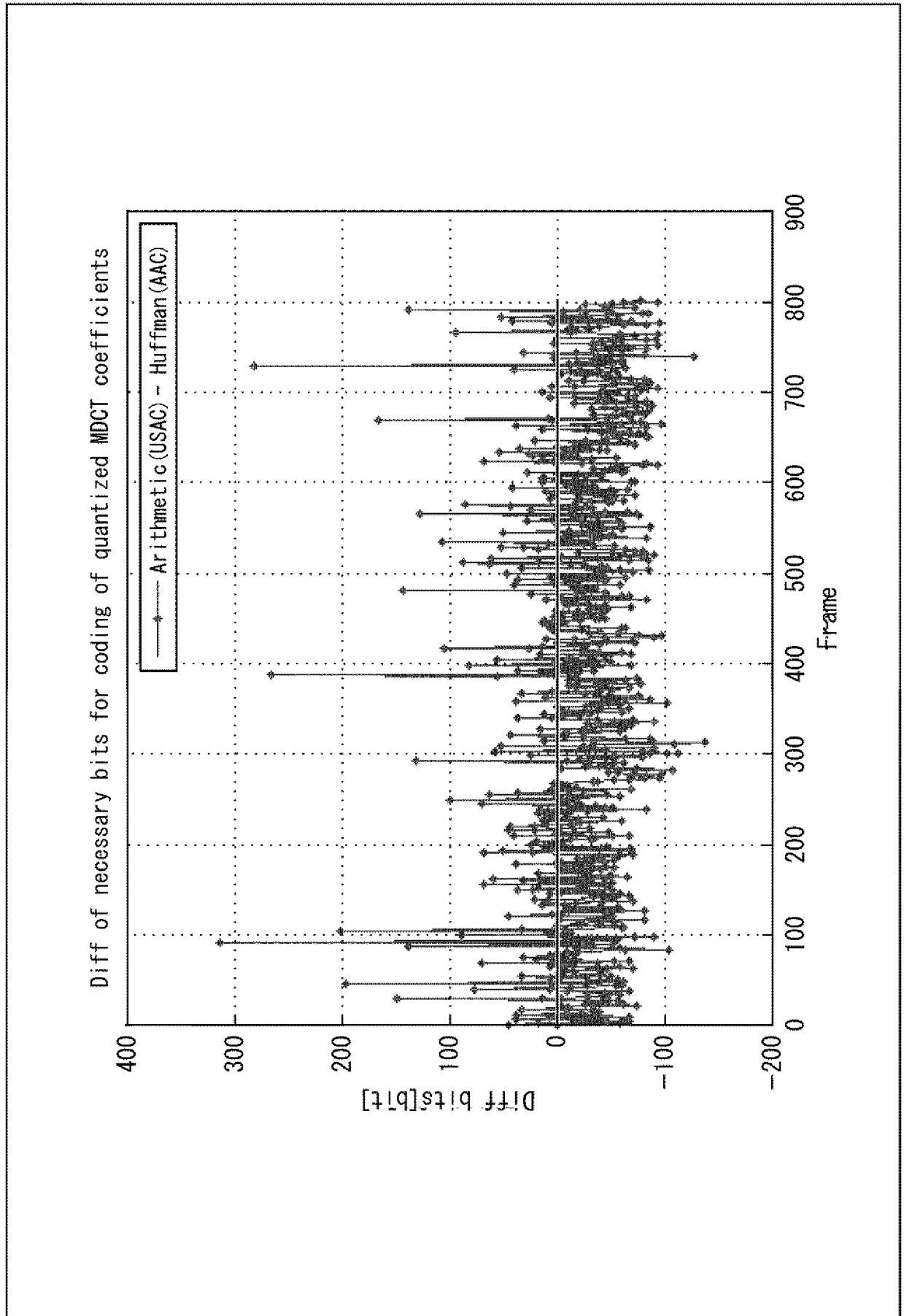
[図12]
FIG. 12



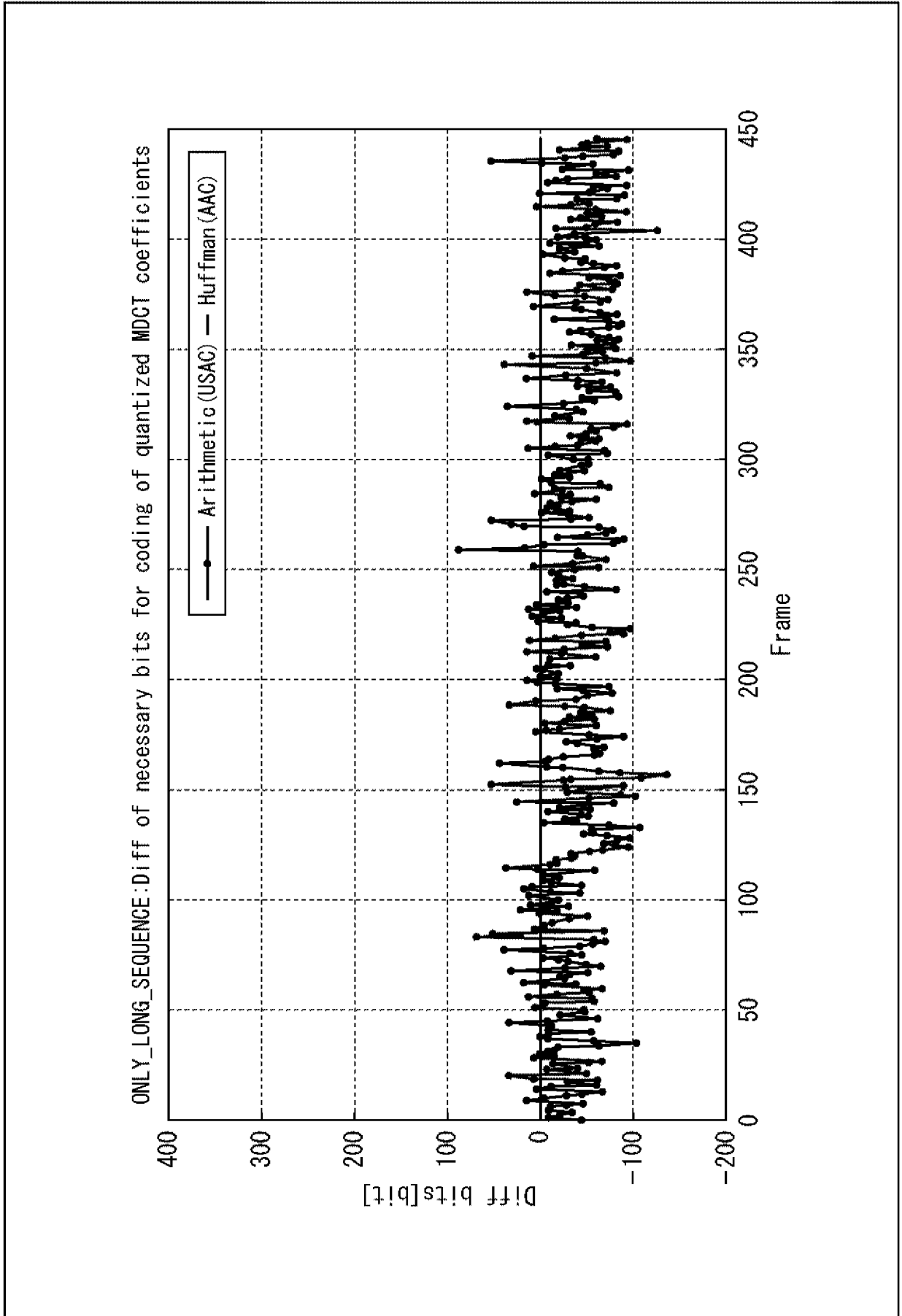
[図13]
FIG. 13



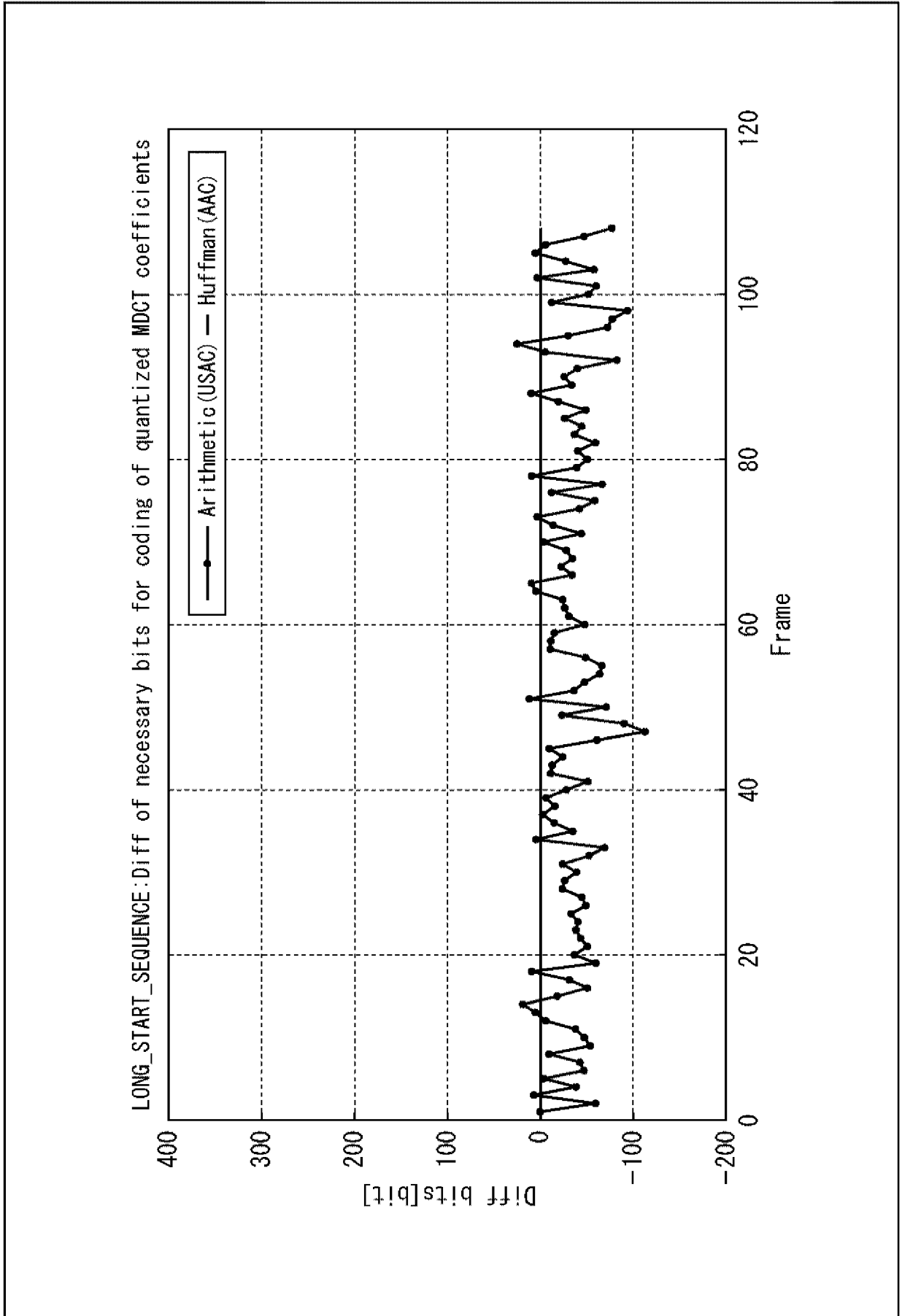
[図14]
FIG. 14



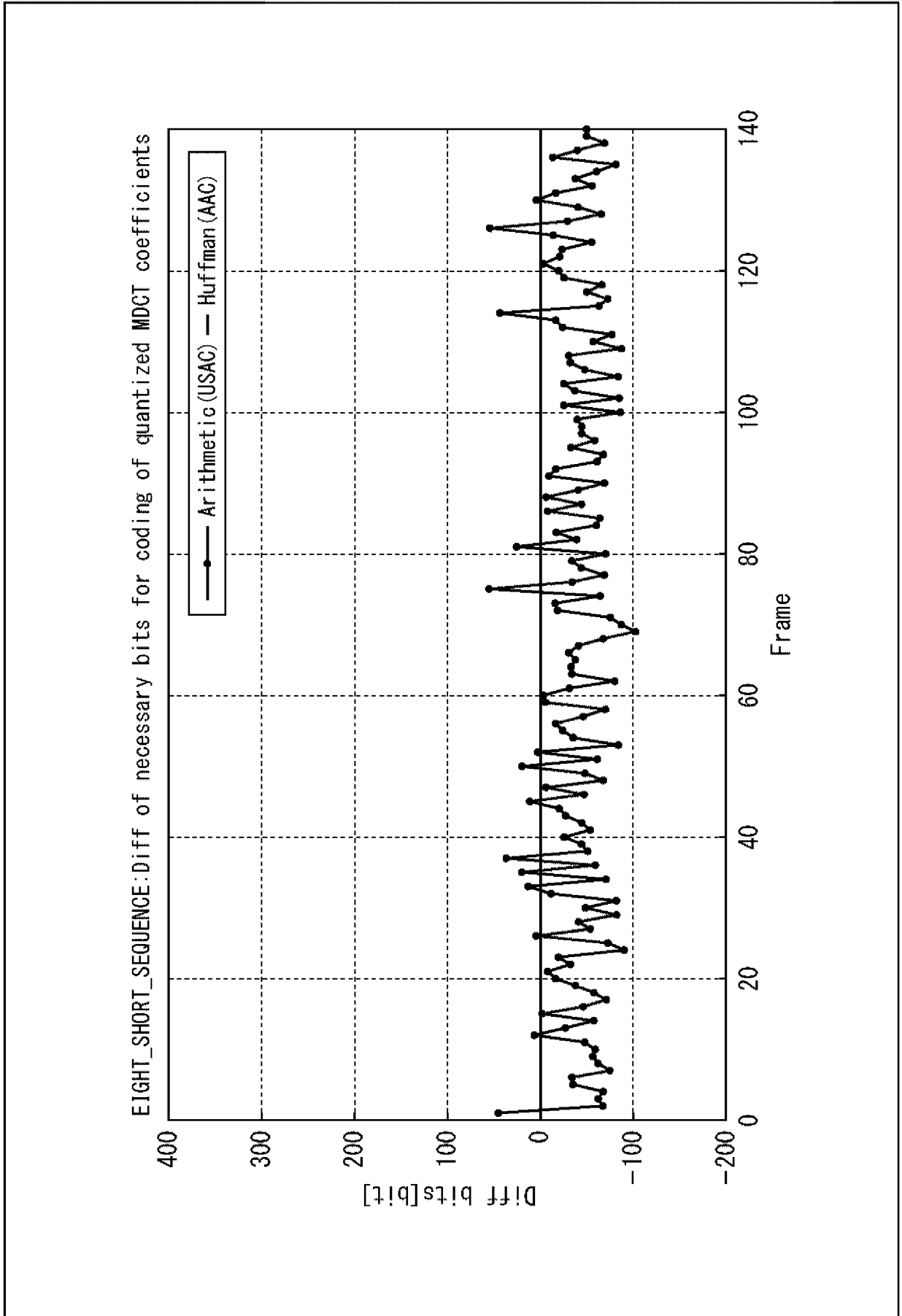
[図15]
FIG. 15




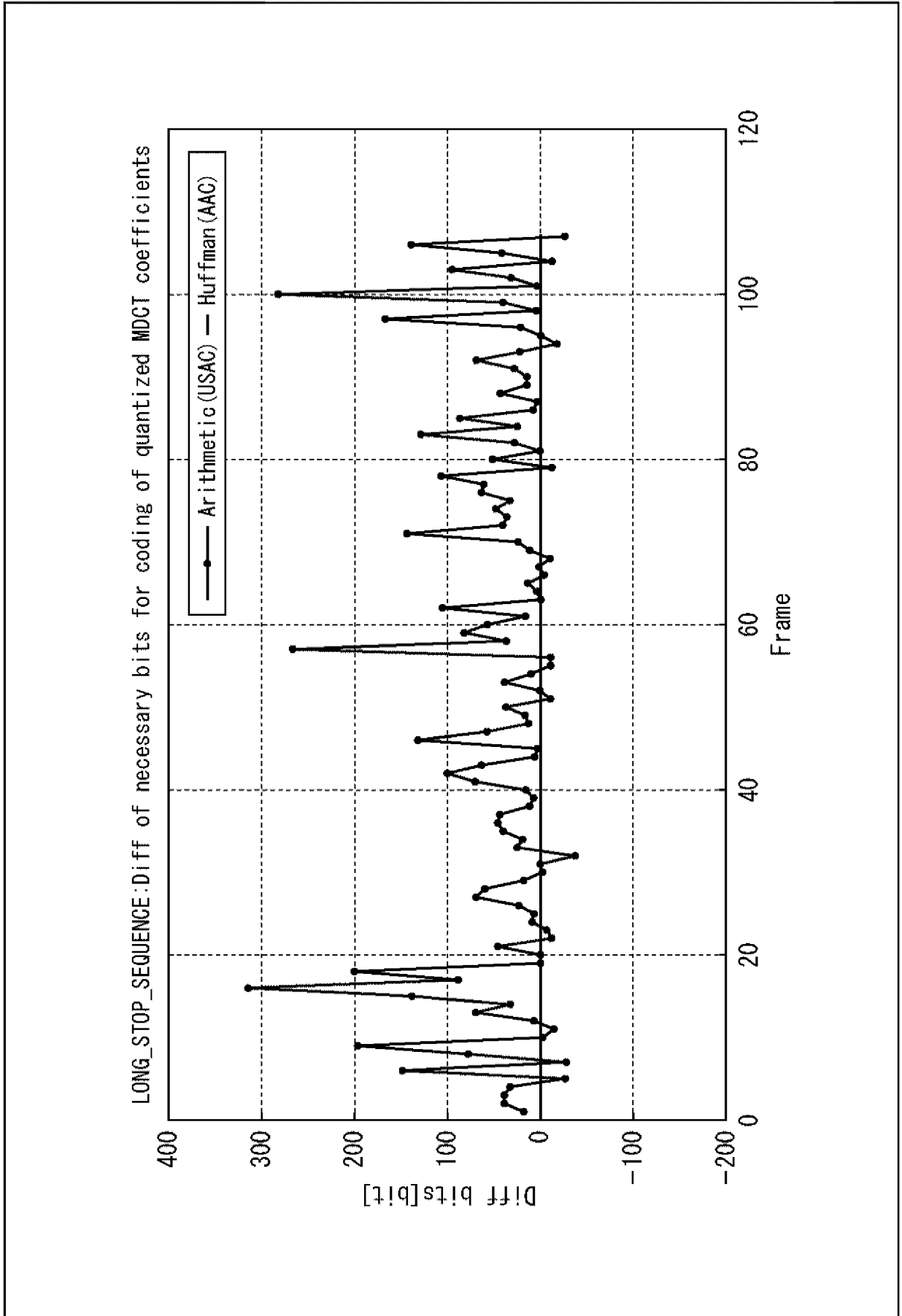
[図16]
FIG. 16



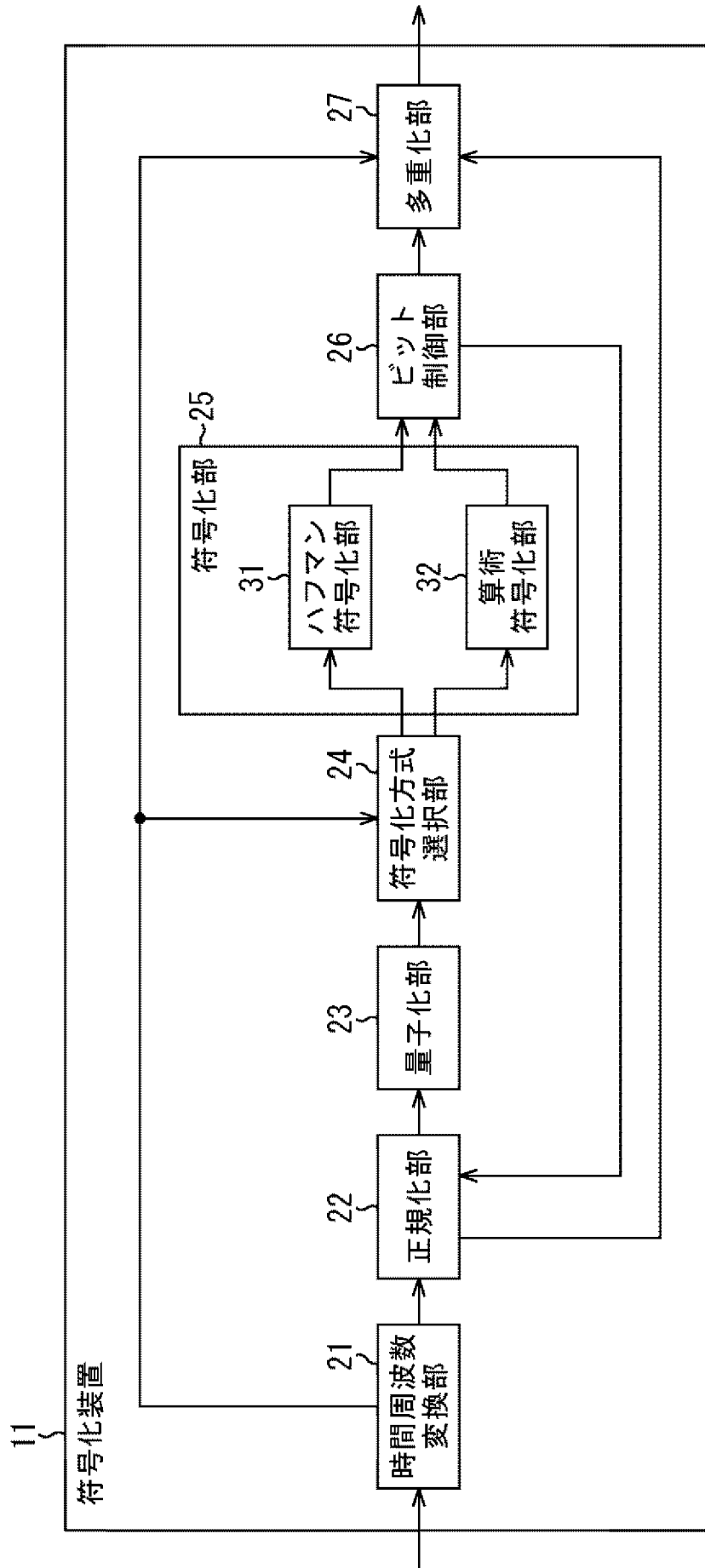
[図17]
FIG. 17



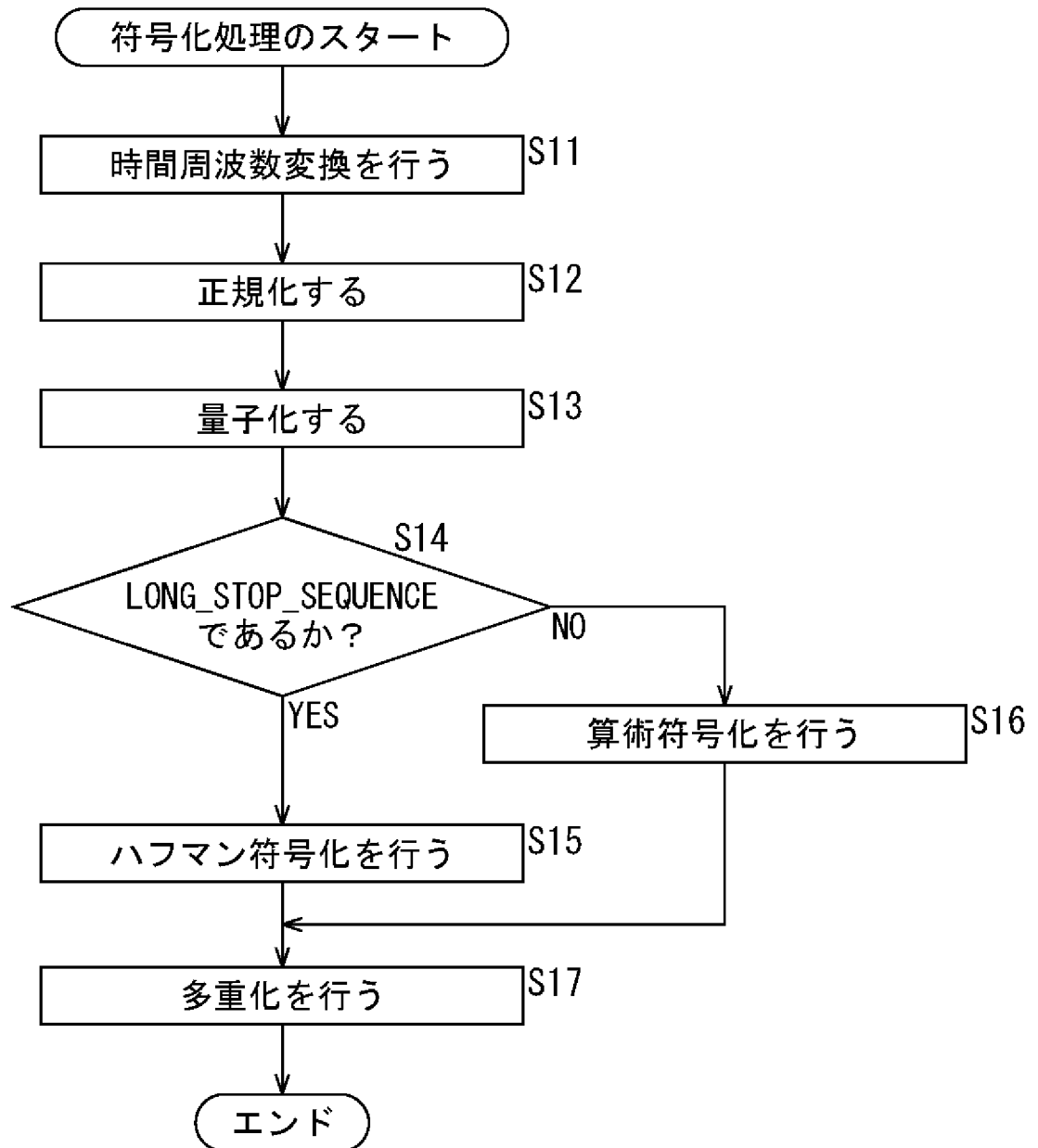
[18]
FIG. 18



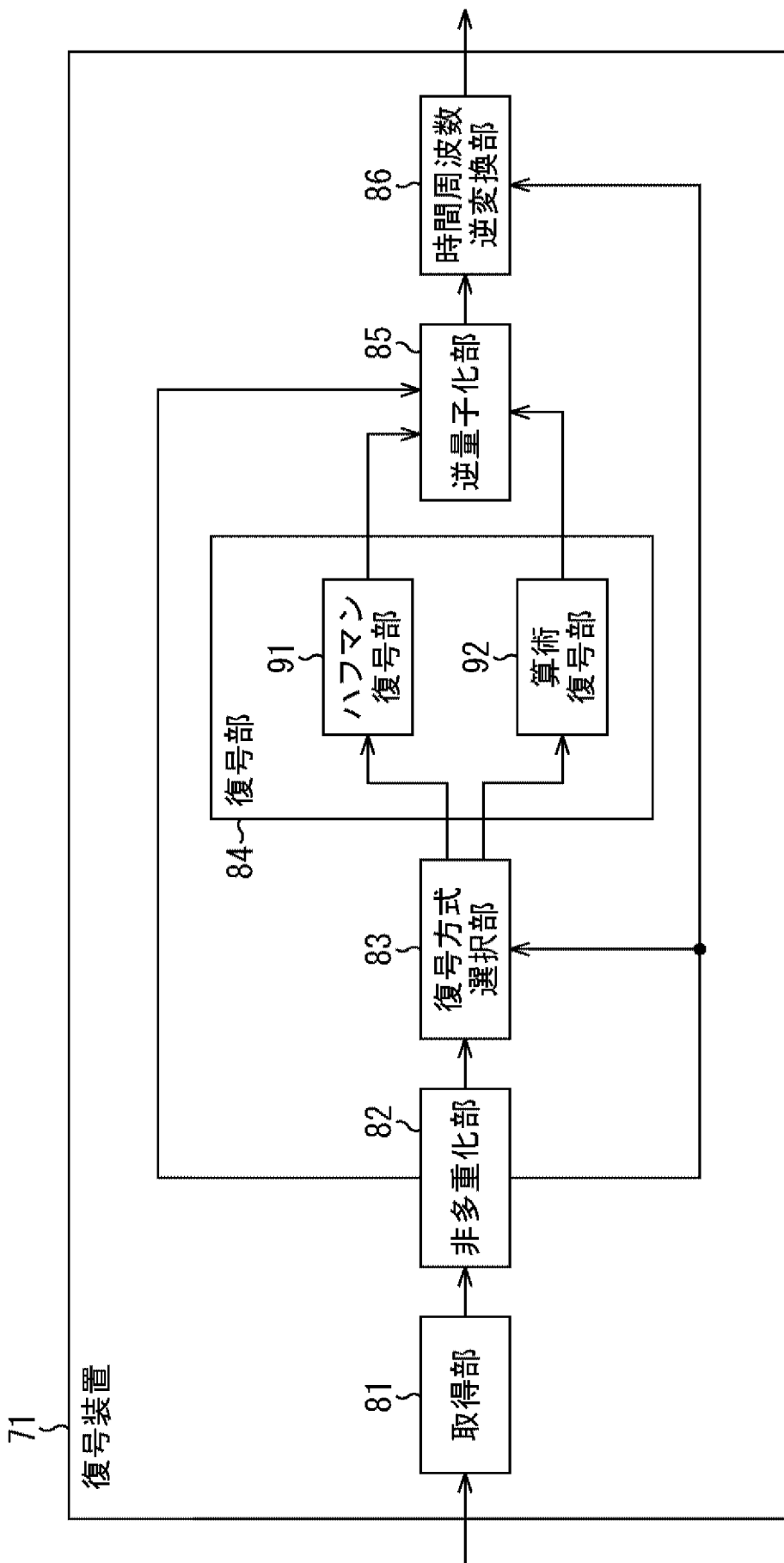
[図19]
FIG. 19



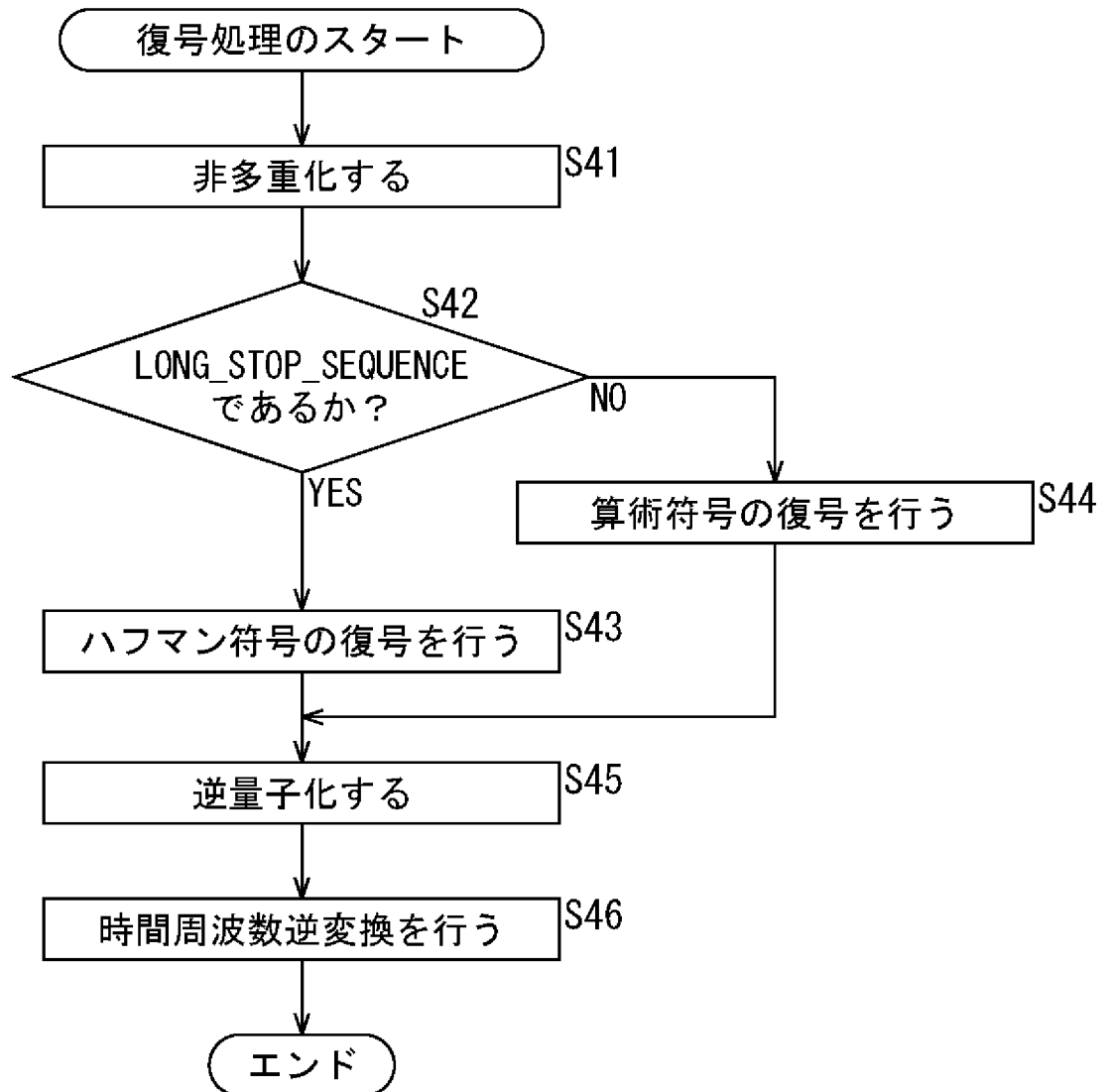
[図20]
FIG. 20



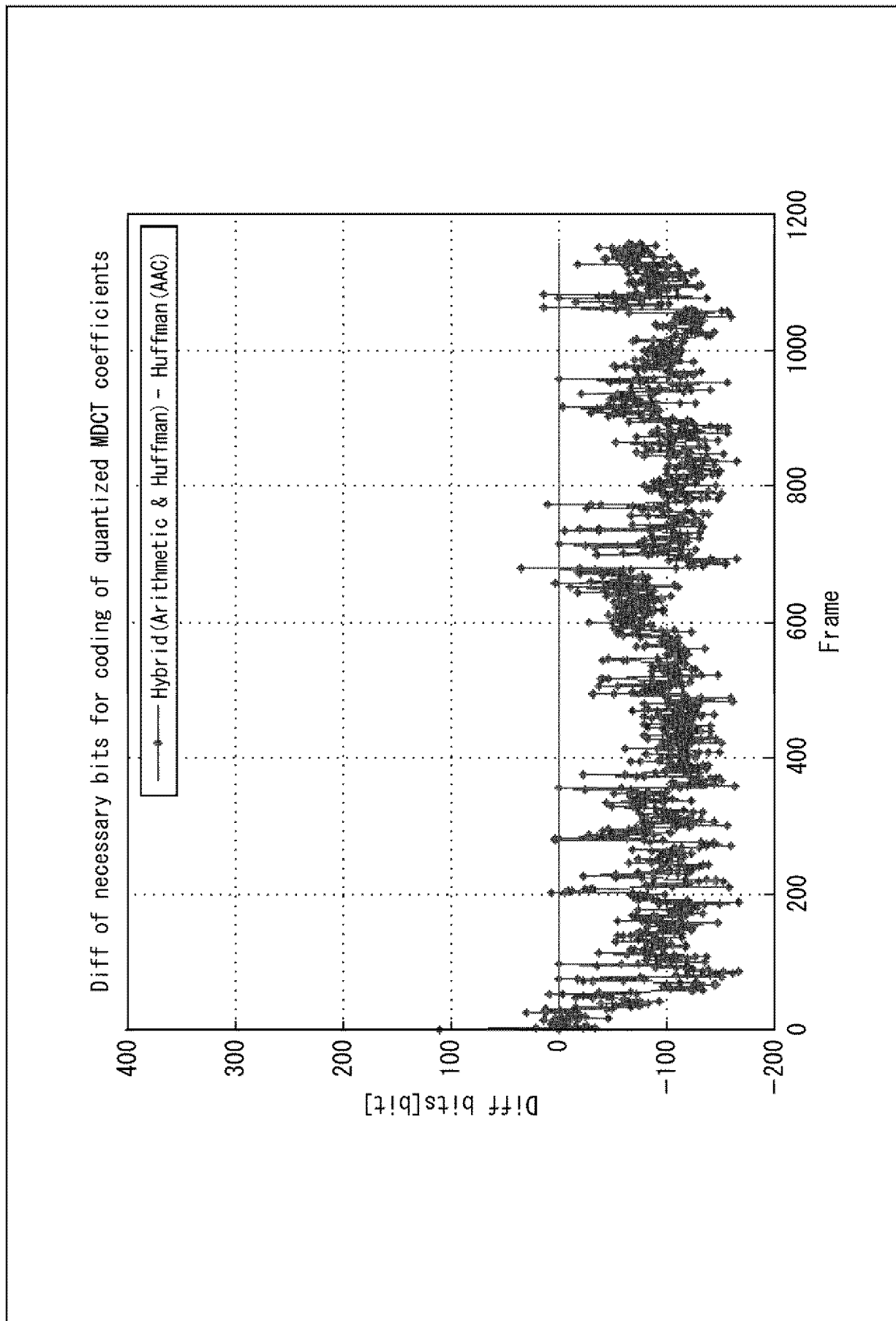
[図21]
FIG. 21



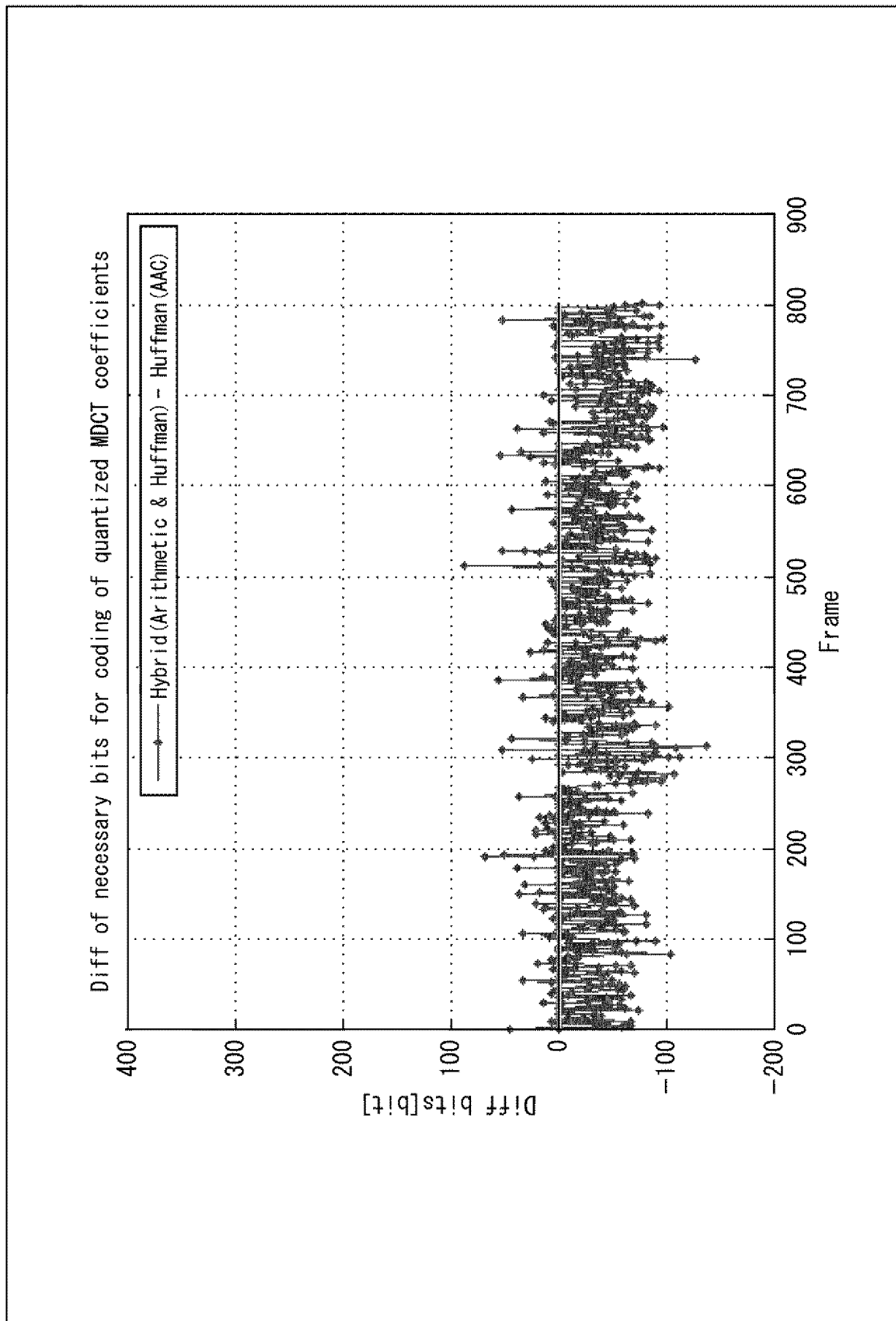
[図22]
FIG. 22



[図23]
FIG. 23



[図24]
FIG. 24



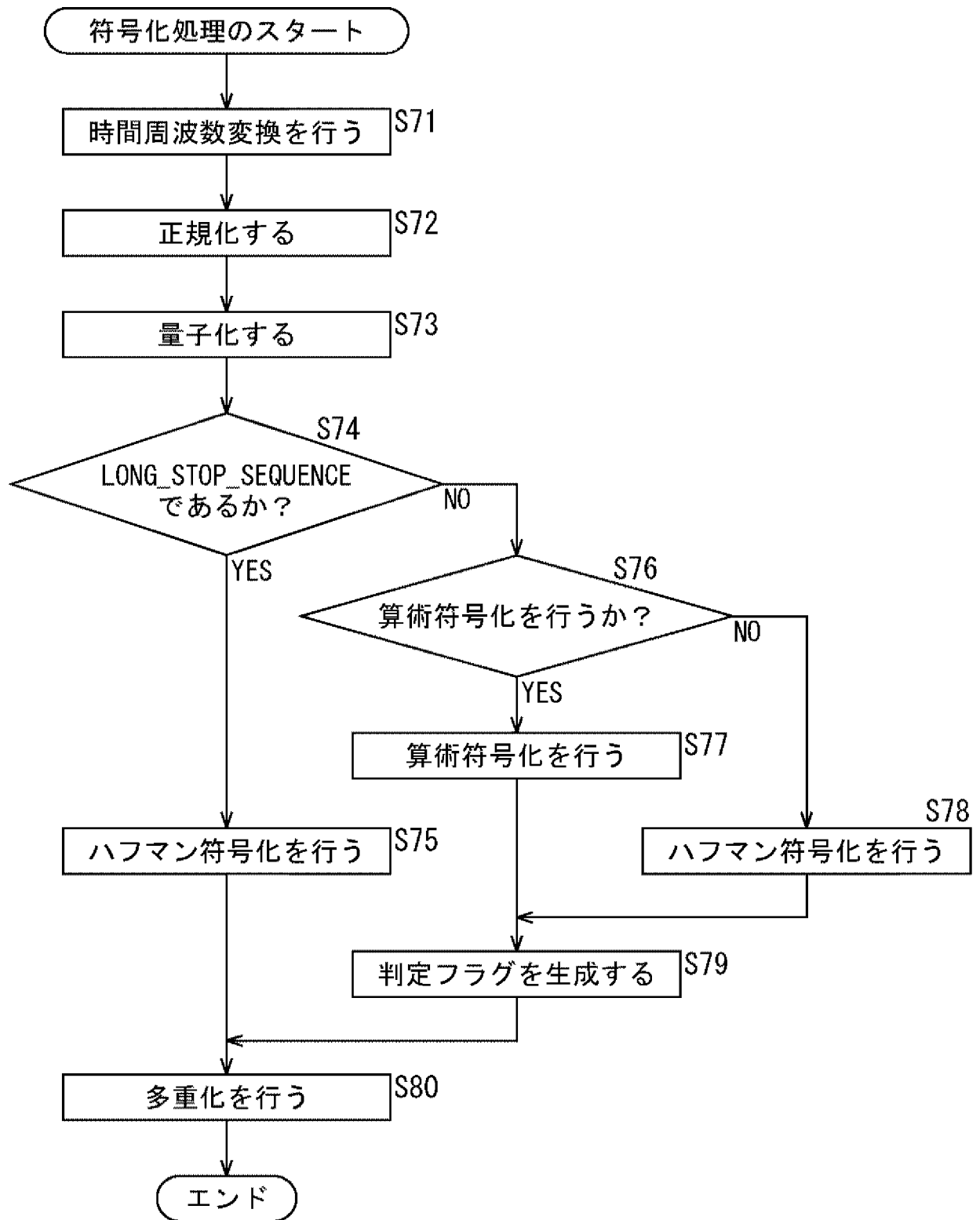
[図25]
FIG. 25

Syntax	No. of bits	Mnemonic
fd_channel_stream(common_window, common_tw, tns_data_present, noiseFilling, indepFlag)		
{		
global_gain;	8	uimsbf
if(noiseFilling){		
noise_level;	3	uimsbf
noise_offset;	5	uimsbf
}		
else{		
noise_level = 0;		
}		
if(!common_window){		
ics_info(); ← Q11		
}		
if(tw_mdct){		
if(!common_tw){		
tw_data();		
}		
}		
if((window_sequence != LONG_STOP_SEQUENCE)&&(huffman_coding_flag == 1)){		
section_data(); ← Q12		
}		
scale_factor_data();		
if(tns_data_present){		
tns_data();		
}		
ac_spectral_data(indepFlag);		
fac_data_present;	1	uimsbf
if (fac_data_present){		
fac_length = (window_sequence==EIGHT_SHORT_SEQUENCE)? ccf1/16:ccf1/8;		
fac_data(1, fac_length);		
}		
}		

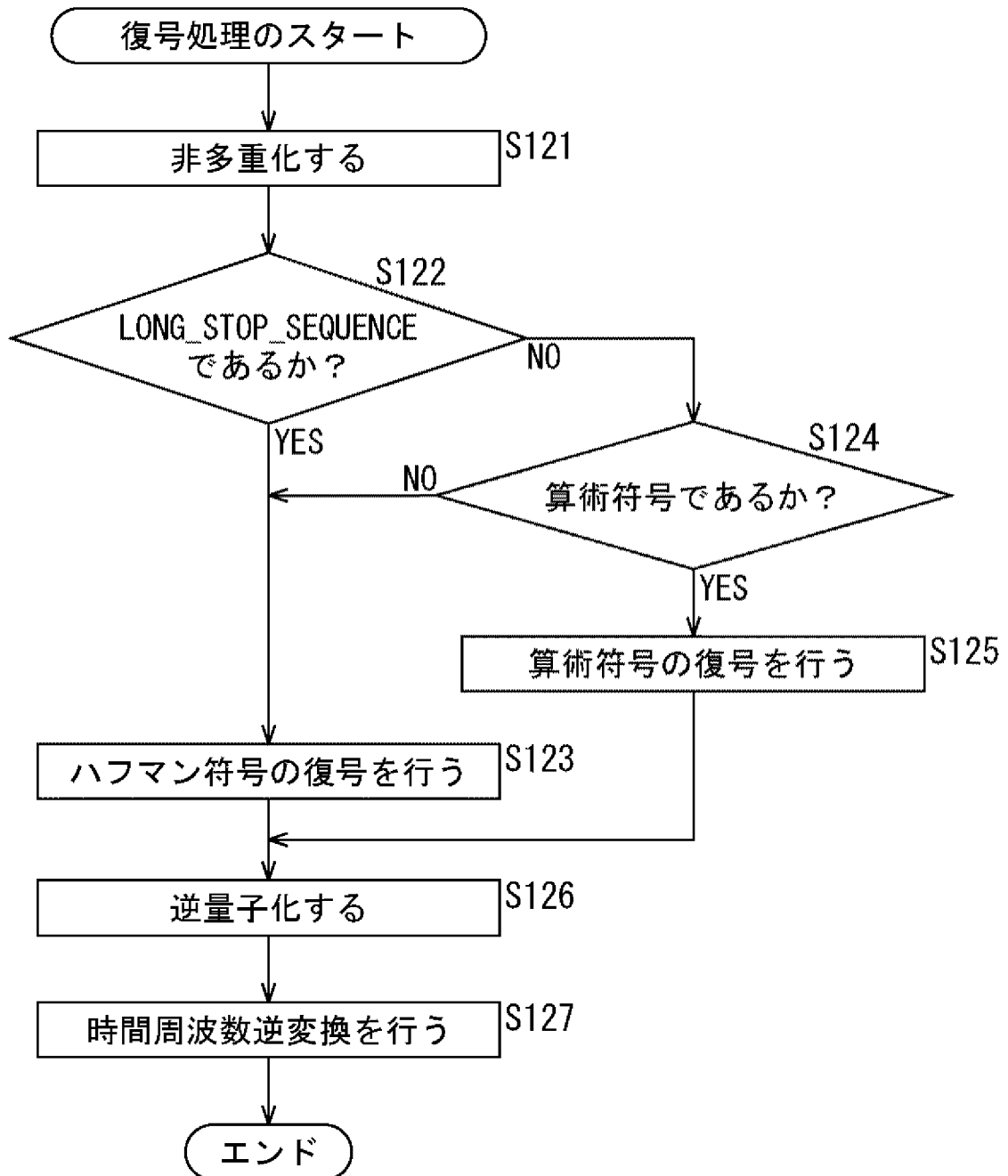
[図26]
FIG. 26

Syntax	No. of bits	Mnemonic
ics_info()		
{		
window_sequence:	2	uimsbf
window_shape	1	uimsbf
if(window_sequence == EIGHT_SHORT_SEQUENCE) {		
max_sfb:	4	uimsbf
scale_factor_grouping;	7	uimsbf
}		
else{		
max_sfb:	6	uimsbf
if(window_sequence!= LONG_STOP_SEQUENCE) {		
huffman_coding_flag	1	uimsbf
}		
}		
}		

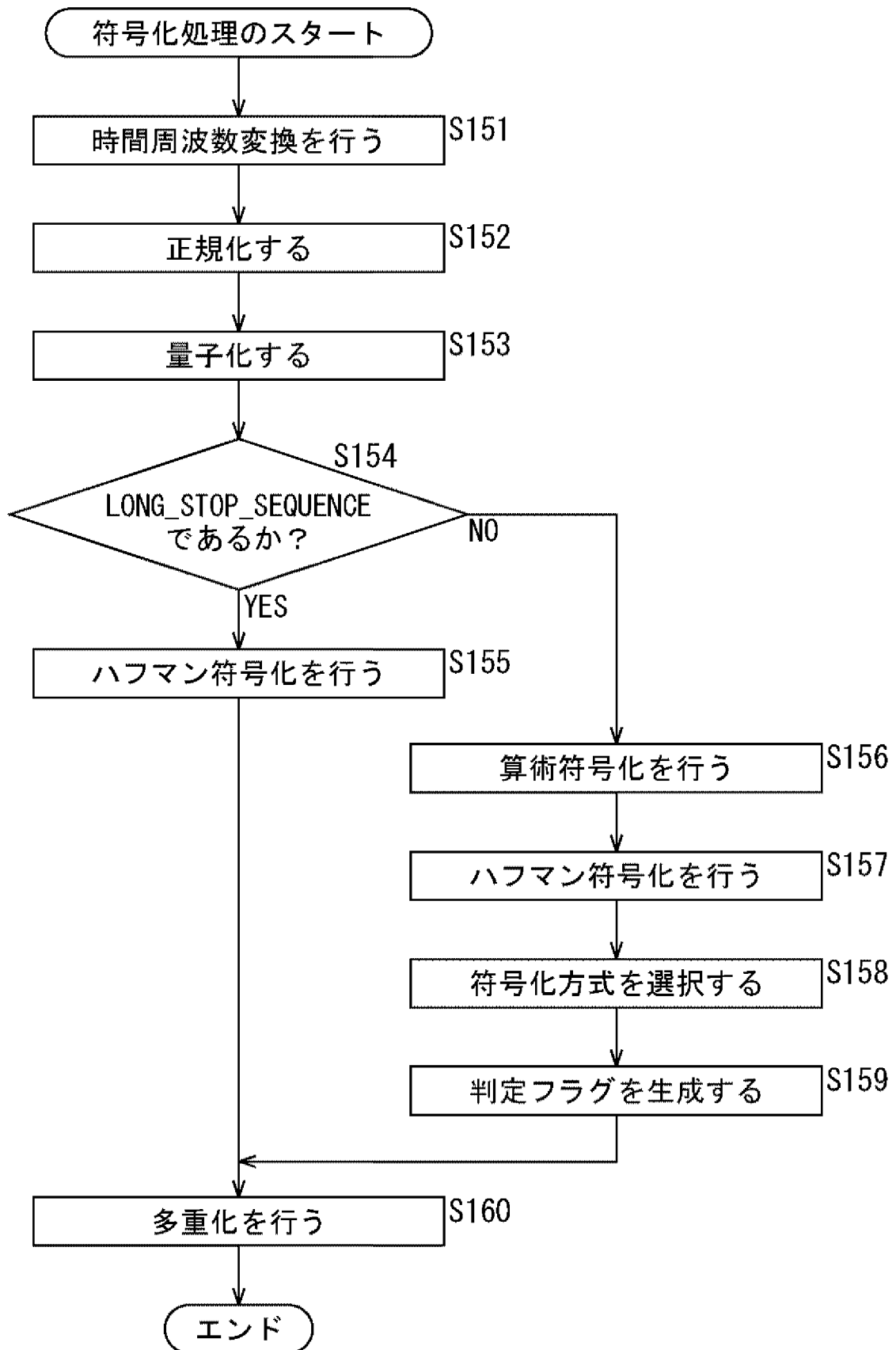
[図27]
FIG. 27



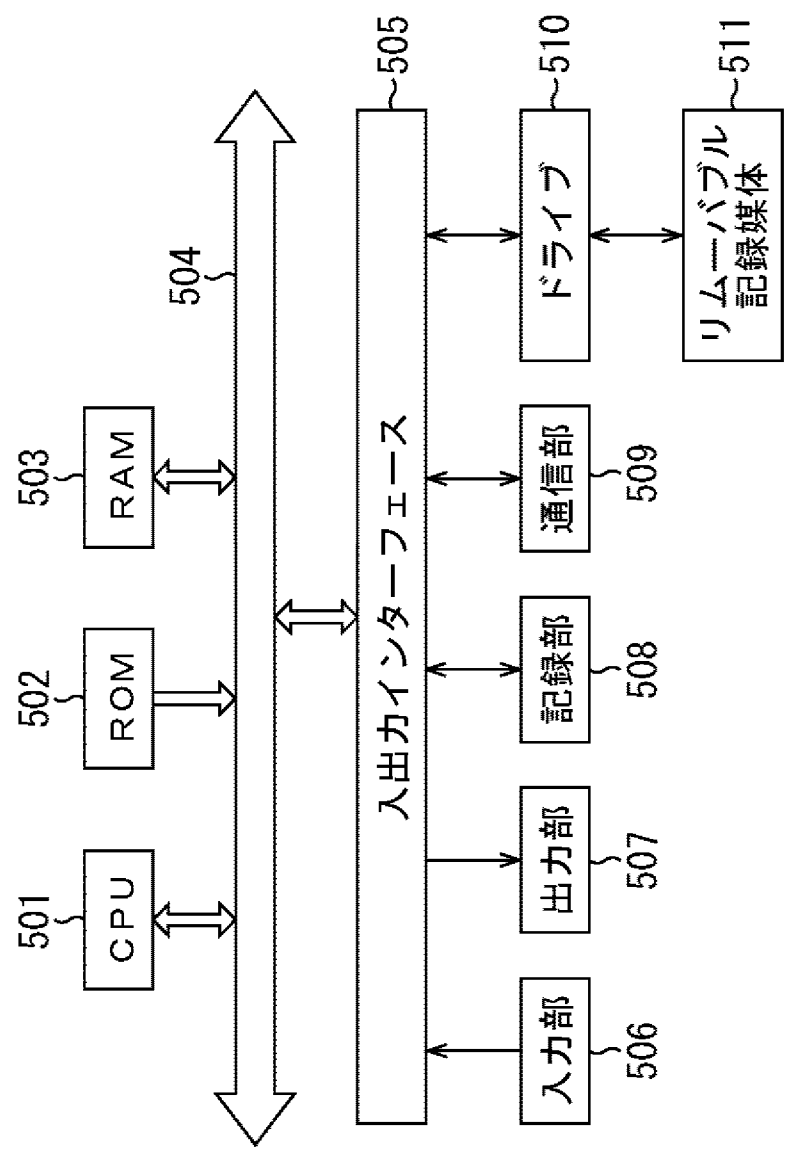
[図28]
FIG. 28



[図29]
FIG. 29



[図30]
FIG. 30



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/022681

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl. G10L19/02 (2013.01) i, G10L19/00 (2013.01) i, G10L19/022 (2013.01) i, H03M7/40 (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)

Int.Cl. G10L19/02, G10L19/00, G10L19/022, H03M7/40

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

Published examined utility model applications of Japan	1922-1996
Published unexamined utility model applications of Japan	1971-2019
Registered utility model specifications of Japan	1996-2019
Published registered utility model applications of Japan	1994-2019

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	JP 2011-022435 A (SONY CORPORATION) 03 February 2011, paragraphs [0226]-[0237] & US 2011/0015933 A1, paragraphs [0254]-[0265] & EP 2284832 A1 & CN 101958120 A	1-12
A	JP 2012-181429 A (FUJITSU LIMITED) 20 September 2012, paragraph [0024] & US 2012/0224703 A1, paragraph [0036]	1-12

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	“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
“E” earlier application or patent but published on or after the international filing date	“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
“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)	“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
“O” document referring to an oral disclosure, use, exhibition or other means	“&” document member of the same patent family
“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 01.08.2019	Date of mailing of the international search report 13.08.2019
---	--

Name and mailing address of the ISA/ Japan Patent Office 3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan	Authorized officer 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(2013.01)i, G10L19/00(2013.01)i, G10L19/022(2013.01)i, H03M7/40(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. G10L19/02, G10L19/00, G10L19/022, H03M7/40

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	1971-2019年
日本国実用新案登録公報	1996-2019年
日本国登録実用新案公報	1994-2019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	JP 2011-022435 A (ソニー株式会社) 2011.02.03, [0226] - [0237] & US 2011/0015933 A1, [0254]-[0265] & EP 2284832 A1 & CN 101958120 A	1-12
A	JP 2012-181429 A (富士通株式会社) 2012.09.20, [0024] & US 2012/0224703 A1, [0036]	1-12

☐ 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

01.08.2019

国際調査報告の発送日

13.08.2019

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)

千本 潤介

電話番号 03-3581-1101 内線 3591

5Z

3457