

(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1月31日(31.01.2013)



(10) 国際公開番号

WO 2013/015011 A1

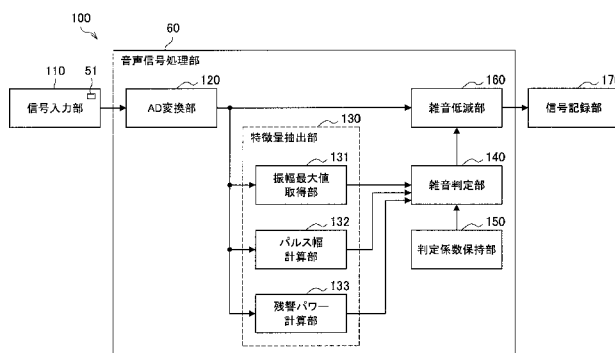
- (51) 国際特許分類:
G10L 11/00 (2006.01) H04N 5/225 (2006.01)
G10L 21/02 (2013.01)
- (21) 国際出願番号: PCT/JP2012/063753
- (22) 国際出願日: 2012年5月29日(29.05.2012)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2011-163045 2011年7月26日(26.07.2011) JP
- (71) 出願人 (米国を除く全ての指定国について): ソニー株式会社 (SONY CORPORATION) [JP/JP]; 〒1080075 東京都港区港南1丁目7番1号 Tokyo (JP).
- (72) 発明者: および
- (75) 発明者/出願人 (米国についてのみ): 大迫 慶一 (OSAKO, Keiichi) [JP/JP]; 〒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). 関矢 俊之 (SEKIYA, Toshiyuki) [JP/JP]; 〒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). 安部 素嗣 (ABE, Mototsugu) [JP/JP]; 〒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: 亀谷 美明, 外 (KAMEYA, Yoshiaki et al.); 〒1600004 東京都新宿区四谷3-1-3 第一富澤ビル はづき国際特許事務所 四谷オフィス Tokyo (JP).
- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, 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, 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, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[続葉有]

(54) Title: AUDIO-SIGNAL-PROCESSING APPARATUS, IMAGING APPARATUS, AUDIO-SIGNAL-PROCESSING METHOD, PROGRAM, AND RECORDING MEDIUM

(54) 発明の名称: 音声信号処理装置、撮像装置、音声信号処理方法、プログラム及び記録媒体

[図2]



- 60 Audio-signal-processing unit
110 Signal unit
120 AD converter unit
130 Feature-value extractor
131 Amplitude maximum value acquisition unit
132 Pulse width calculation unit
133 Reverberation power calculation unit
140 Noise-assessment unit
150 Assessment factor retention portion
160 Noise reduction unit
170 Signal recording

(57) Abstract: [Problem] To detect, with a high degree of precision, a noise interval including operational noise in a pulse state, from an audio signal only, collected without using drive-control information from a drive apparatus; and to improve the precision of noise reduction. [Solution] An audio-signal-processing apparatus is provided with: a feature-value extractor for extracting from an audio signal obtained from a noise collector a feature value showing operation noise in a pulse state generated from a noise generator installed in the same cabinet with the noise collector; and a noise-assessment unit for assessing, on the basis of the feature value, whether the operation noise is included in the audio signal.

(57) 要約: 【課題】駆動装置の駆動制御情報を用いずに、收音された音声信号のみから、パルス状の作動音が含まれる雑音区間を高精度に検出して、雑音低減精度を向上する。【解決手段】收音部により得られる音声信号から、前記收音部と同一の筐体に設けられた発音部から発生するパルス状の作動音を表す特徴量を抽出する特徴量抽出部と、前記特徴量に基づいて、前記音声信号に前記作動音が含まれるか否かを判定する雑音判定部と、を備える、音声信号処理装置が提供される。



添付公開書類:

— 国際調査報告 (条約第 21 条(3))

明 細 書

発明の名称：

音声信号処理装置、撮像装置、音声信号処理方法、プログラム及び記録媒体

技術分野

[0001] 本開示は、音声信号処理装置、撮像装置、音声信号処理方法、プログラム及び記録媒体に関する。

背景技術

[0002] 外部音声を受音して記録する音声記録装置として、ビデオカメラや、動画撮像機能付きのデジタルカメラ、ＩＣレコーダ等が知られている。これら装置を動作させるときには、装置本体から発生するパルス状の作動音が、記録音声に混入してしまう場合がある。

[0003] 例えば、動画撮像機能を備えた撮像装置は、動画撮像中にマイクロホンにより装置周辺の外部音声を受音し、当該音声動画をともに記録する。かかる動画撮像時には、撮像装置に筐体内で、ズーム駆動音、絞り駆動音、オートフォーカス駆動音、操作ボタンの押下音などのパルス状の作動音が発生する。特に、撮像光学系を駆動させる駆動装置（ズームモータ、絞り機構、フォーカスモータ等）の動作開始時又は終了時には、例えば、モータとギアが噛み合うときの「カチッ」というパルス状の機械駆動音が発生する。

[0004] かかるパルス状の作動音は、ユーザが録音を所望する外部音声に雑音として混入・記録されてしまうと、非常に耳障りである。このため、音声記録装置では、記録時にパルス状の作動音を低減するための静音対策や雑音除去対策が必要となる。

[0005] パルス状の機械駆動音を低減する方法として、これまでもいくつかの方法が提案されている。例えば、特許文献１には、駆動装置から発生する機械駆動音を低減する際に、低減処理の負荷を低減するために、駆動装置を制御するための駆動信号の送信に応じて、低減処理を実行することが提案されてい

る。また、特許文献2には、レンズ駆動部を制御するための駆動信号（制御信号）に基づいて、レンズ駆動部からの雑音が大きい期間に対応する入力音声データを、当該期間の前後の音声データで補間することが提案されている。

先行技術文献

特許文献

[0006] 特許文献1：特開2011-2723号公報

特許文献2：特開平8-124299号公報

発明の概要

発明が解決しようとする課題

[0007] しかしながら、上記特許文献1、2の技術では、駆動装置を制御するための駆動信号の送信タイミングに依存して雑音低減処理を行っているので、パルス状の雑音の低減精度が低く、原音の音質が劣化してしまう。即ち、実際の撮像装置等では、制御装置が駆動信号を送信してから駆動装置は動作するまでの間に遅延が生じる場合がある。しかも、特許文献1の方法では、機械駆動音が発生してからマイクロホンに到達するまでの時間を考慮していない。従って、特許文献1のように駆動信号の送信に応じて低減処理を行うと、実際の雑音区間と雑音低減処理区間とがずれてしまう。このため、非雑音区間でも低減処理を行うことになるので、原音の忠実性の著しい低下や、音質劣化を招いてしまう。

[0008] また、特許文献2の技術では、駆動信号に対する発生雑音の遅延を予め考慮して、駆動信号の送信から一定の遅延時間の経過後に、雑音低減処理を実行している。しかし、設定される遅延時間は固定値であり、実際の遅延時間がばらつく場合は考慮されていないので、雑音区間を正確に特定できない。従って、雑音低減処理区間を実際の雑音区間に正確に合わせることができないので、高精度の雑音低減処理を実現できず、音質劣化の恐れがある。

[0009] さらに、特許文献2では、駆動信号から得られるタイミングで、入力音声

データと補間音声データのレベル（音量）によって、雑音低減処理を行うか否かを判定している。これにより、装置周囲の外部音声が大きい場合には、雑音が相対的に目立たなくなるため、低減処理を行わず、逆に、装置周囲が静かであり、雑音が目立つ場合には、低減処理を行って当該雑音を除去することができる。しかしながら、雑音が人間の耳に知覚される要因は音量だけではない。従って、特許文献2のように音量のみで雑音区間を判定する方法では、雑音区間の判定精度が不安定であり、雑音が知覚されない区間に低減処理を実行したり、逆に、雑音が知覚される区間で低減処理を実行しなかったりする可能性がある。

[0010] そこで、上記事情に鑑みて、駆動装置の駆動制御情報を用いずに、收音された音声信号のみから、パルス状の作動音が含まれる雑音区間を高精度に検出して、雑音低減精度を向上することが可能な手法が求められていた。

課題を解決するための手段

[0011] 本開示によれば、收音部により得られる音声信号から、前記收音部と同一の筐体に設けられた発音部から発生するパルス状の作動音を表す特徴量を抽出する特徴量抽出部と、前記特徴量に基づいて、前記音声信号に前記作動音が含まれるか否かを判定する雑音判定部と、を備える、音声信号処理装置が提供される。

[0012] また、本開示によれば、外部音声を音声信号に変換する收音部と、前記收音部と同一の筐体に設けられ、パルス状の作動音を発生させる発音部と、前記收音部から出力された前記音声信号から、前記作動音を表す特徴量を抽出する特徴量抽出部と、前記特徴量に基づいて、前記音声信号に前記作動音が含まれるか否かを判定する雑音判定部と、前記雑音判定部により前記作動音が含まれると判定された場合に、前記音声信号に対して雑音低減処理を行う雑音低減部と、を備える、撮像装置が提供される。

[0013] また、本開示によれば、收音部により得られる音声信号から、前記收音部と同一の筐体に設けられた発音部から発生するパルス状の作動音を表す特徴量を抽出することと、前記特徴量に基づいて、前記音声信号に前記作動音が

含まれるか否かを判定することと、を含む、音声信号処理方法が提供される。

[0014] また、本開示によれば、收音部により得られる音声信号から、前記收音部と同一の筐体に設けられた発音部から発生するパルス状の作動音を表す特徴量を抽出することと、前記特徴量に基づいて、前記音声信号に前記作動音が含まれるか否かを判定することと、をコンピュータに実行させるためのプログラムが提供される。

[0015] また、本開示によれば、收音部により得られる音声信号から、前記收音部と同一の筐体に設けられた発音部から発生するパルス状の作動音を表す特徴量を抽出することと、前記特徴量に基づいて、前記音声信号に前記作動音が含まれるか否かを判定することと、をコンピュータに実行させるためのプログラムが記録された、コンピュータ読み取り可能な記録媒体が提供される。

[0016] 上記構成により、收音部により得られる音声信号から、前記收音部と同一の筐体に設けられた発音部から発生するパルス状の作動音を表す特徴量が抽出され、前記特徴量に基づいて、前記音声信号に前記作動音が含まれるか否かが判定される。これにより、音声信号から抽出される特徴量を用いて、パルス状の作動音の有無を高精度で判定できる。

発明の効果

[0017] 以上説明したように本開示によれば、駆動装置の駆動制御情報を用いずに、收音された音声信号のみから、パルス状の作動音の雑音区間を高精度に検出して、雑音低減精度を向上することができる。

図面の簡単な説明

[0018] [図1]本開示の第1の実施形態に係る音声信号処理装置が適用されたデジタルカメラのハードウェア構成を示すブロック図である。

[図2]同実施形態に係る音声信号処理装置の機能構成を示すブロック図である。

[図3]同実施形態に係るパルス機械音を含む音声信号の波形図である。

[図4]同実施形態に係るパルス機械音の判定方法を示す説明図である。

[図5]同実施形態に係る音声信号処理方法を示すフローチャートである。

[図6]本開示の第2の実施形態に係る音声信号処理装置の機能構成を示すブロック図である。

[図7]同実施形態に係るパルス機械音の判定方法を示す説明図である。

[図8]本開示の第3の実施形態に係る音声信号処理装置の機能構成を示すブロック図である。

[図9]本開示の第4の実施形態に係る音声信号処理装置の機能構成を示すブロック図である。

[図10]本開示の第5の実施形態に係る音声信号処理装置の機能構成を示すブロック図である。

[図11]同実施形態に係る音声信号処理方法を示すフローチャートである。

[図12]本開示の第6の実施形態に係る音声信号処理装置の機能構成を示すブロック図である。

[図13]本開示の第6の実施形態の変更例に係る音声信号処理装置の機能構成を示すブロック図である。

[図14]本開示の第7の実施形態に係る音声信号処理装置の機能構成を示すブロック図である。

発明を実施するための形態

[0019] 以下に添付図面を参照しながら、本開示の好適な実施の形態について詳細に説明する。なお、本明細書及び図面において、実質的に同一の機能構成を有する構成要素については、同一の符号を付することにより重複説明を省略する。

[0020] なお、説明は以下の順序で行うものとする。

1. 第1の実施の形態

1. 1. 機械音低減方法の概要

1. 2. 音声信号処理装置の構成

1. 2. 1. 音声信号処理装置のハードウェア構成

1. 2. 2. 音声信号処理装置の機能構成

1. 3. パルス機械音の判定方法の具体例
 1. 3. 1. 統計的識別法を用いた判定方法
 1. 3. 2. テーブル判定を用いた判定方法
1. 4. 音声信号処理方法
1. 5. 効果
2. 第2の実施の形態
 2. 1. 音声信号処理装置の機能構成
 2. 2. パルス機械音の判定方法の具体例
 2. 2. 1. 統計的識別法を用いた判定方法
 2. 2. 2. テーブル判定を用いた判定方法
 2. 3. 効果
3. 第3の実施の形態
 3. 1. 音声信号処理装置の機能構成
 3. 2. パルス機械音の判定方法の具体例
 3. 2. 1. 統計的識別法を用いた判定方法
 3. 2. 2. テーブル判定を用いた判定方法
 3. 3. 効果
4. 第4の実施の形態
 4. 1. 音声信号処理装置の機能構成
 4. 2. 効果
5. 第5の実施の形態
 5. 1. 音声信号処理装置の機能構成
 5. 2. 音声信号処理方法
 5. 3. 効果
6. 第6の実施の形態
 6. 1. 雑音判定方法の概要
 6. 2. 音声信号処理装置の機能構成
 6. 3. 音声信号処理装置の変更例

6. 4. 効果

7. 第7の実施の形態

7. 1. 音声信号処理装置の機能構成

7. 2. 効果

8. まとめ

[0021] <1. 第1の実施の形態>

[1. 1. 機械音低減方法の概要]

まず、本開示の第1の実施形態に係る音声信号処理装置及び方法を用いた機械音低減方法の概要について説明する。

[0022] 本実施形態に係る音声信号処理装置及び方法は、音声記録装置又は音声再生装置において、当該装置の筐体内に設置された発音部（例えば駆動装置）から発生するパルス状の作動音（雑音）を低減する技術に関する。特に、本実施形態では、動画撮像機能を有する撮像装置において、動画を撮像しながら周辺音声を録音するときに、撮像装置に内蔵された駆動装置の動作開始時或いは動作終了時に発生するパルス状の機械駆動音を低減対象とする。

[0023] ここで、駆動装置は、撮像光学系を用いた撮像動作を行うために撮像装置に内蔵された駆動装置であり、例えば、ズームレンズを移動させるズームモータや、フォーカスレンズを移動させるフォーカスモータ、絞り機構又はシャッターを制御する駆動機構などを含む。これら駆動装置は、撮像装置の収音部と同一の筐体内に設けられる。また、パルス状の機械駆動音（以下、「パルス機械音」という。）は、例えば、上記各種の駆動装置が動作開始又は動作終了するときに発生する瞬間的な雑音（例えば、ズームモータの駆動音、フォーカスモータの駆動音、絞り機構の駆動音、シャッター音、操作ボタンの押下音など）である。例えば、パルス機械音は、ズームモータ等の動作開始時又は動作終了時に、モータとギアが噛み合うことにより発生する「カチッ」又は「パチッ」という音などである。

[0024] 以下では、音声信号処理装置が、動画撮像機能を有するデジタルカメラであり、除去対象雑音であるパルス機械音が、該デジタルカメラにおける光学

ズーム動作開始時に発生するズーム開始音である例について説明する。しかし、本開示の音声信号処理装置や、パルス機械音は、かかる例に限定されない。

[0025] デジタルカメラによる撮像及び録音中に、ユーザがズーム操作を行うと、該カメラの内部でズームモータが駆動して、ズームレンズを駆動させるギアと係合して、瞬間的に大きいパルス機械音（ズーム開始音）が発生する。すると、デジタルカメラのマイクロホンは、ユーザが録音を所望するカメラ周囲の外部音声（例えば、環境音、人の話し声など、マイクロホンに收音される任意の音声を含む。以下「所望音」又は「背景音」という。）のみならず、カメラ内部で発生したパルス機械音も收音してしまう。このため、所望音にパルス機械音が雑音として混入した状態で録音されてしまうので、当該録音された音声を再生したときに、所望音に混入したパルス機械音がユーザにとって耳障りとなる。例えば、パルス機械音は200Hz以下の筐体の振動を伴い、マイクロホン近傍で発生するため、所望音に比べて大きな音量で收音される。このようにパルス機械音と所望音で音量差があるため、所望音に機械音が混入していると、録音音声の再生時にパルス機械音が目立ってしまう。従って、動画及び音声の記録時又は再生時に、上記ズーム開始音等のパルス機械音を適切に除去した上で所望音のみを記録可能な技術が希求されていた。

[0026] 従来の雑音低減技術では、上記特許文献1、2記載のように、駆動装置を制御するための駆動信号の送信タイミングによって、機械駆動音の発生区間（雑音区間）を推定し、当該雑音区間に対して雑音低減処理を施していた。しかし、駆動信号に対する実際に収録される雑音の遅延等が原因で、駆動信号の送信タイミングから雑音区間を正確に推定することは困難である。また、特許文献2記載のように音声信号のレベル（音量）のみから雑音の有無を判定する方法では、人間の耳が知覚する実際の雑音の発生区間を正確に検出することができなかった。

[0027] そこで、本実施形態では、撮像装置の收音部により收音された音声信号が

ら、上記パルス機械音を表す特徴量を抽出し、駆動信号を用いずに当該特徴量のみを用いてパルス機械音の有無を判定して、パルス機械音が含まれる雑音区間を検出する。ここで、パルス機械音を表す特徴量は、（１）パルス機械音のパルス成分を表すパルス特徴量（第１の特徴量に相当する。）と、（２）パルス機械音の残響成分を表す残響特徴量（第２の特徴量に相当する。）を含むことが好ましい。このようにパルス機械音のパルス成分と残響成分を表す２種類の特徴量を用いることで、音声信号に含まれるパルス機械音を正確に検出することが可能となる。

[0028] さらに、上記（１）パルス特徴量は、（a）パルス機械音のパルス成分の振幅最大値 A と、（b）当該パルス成分のパルス幅 W を含むことが好ましく、（２）残響特徴量は、（c）パルス機械音の残響成分を表す狭帯域信号のパワー P を含むことが好ましい。かかる（a）振幅最大値 A 及び（b）パルス幅 W は、上記パルス特徴量の代表的なパラメータである。また、（c）パルス機械音の残響成分を表す狭帯域信号のパワー P （以下、「残響成分パワー P 」という。）は、上記残響特徴量の代表的なパラメータである。なお、狭帯域信号は、音声信号のうち、パルス機械音の残響成分に特有の低周波数帯域（例えば 200 Hz 以下）の信号であり、残響成分パワー P は、例えば、当該信号の所定区間における振幅（信号レベル）の積算値である。

[0029] 本実施形態に係る音声信号処理装置の特徴量抽出部は、上記（a）パルス機械音のパルス成分の振幅最大値 A 、（b）当該パルス成分のパルス幅 W 、及び（c）残響成分パワー P という３つの特徴量を、音声信号から抽出する。そして、音声信号処理装置の雑音判定部は、抽出された３つの特徴量のベクトル $x(A, W, P)$ と判定係数を用いて、音声信号にパルス機械音が含まれるか否かを総合的に判定する。なお、この判定に必要な判定係数は、予め事前学習又は事前設定によって得られている。このように、上記３つの特徴量（ A, W, P ）を用いることで、音声信号に含まれるパルス機械音を、更に正確に検出することが可能となる。以下に、本実施形態に係るパルス機械音の低減方法について詳述する。

[0030] [1. 2. 音声信号処理装置の構成]

[1. 2. 1. 音声信号処理装置のハードウェア構成]

まず、図1を参照して、本実施形態に係る音声信号処理装置が適用されたデジタルカメラのハードウェア構成例について説明する。図1は、本実施形態に係る音声信号処理装置が適用されたデジタルカメラ1のハードウェア構成を示すブロック図である。

[0031] 本実施形態に係るデジタルカメラ1は、例えば、動画撮像中に動画と共に音声も記録可能な撮像装置である。このデジタルカメラ1は、被写体を撮像して、当該撮像により得られた撮像画像（静止画又は動画のいずれでもよい。）をデジタル方式の画像データに変換し、音声とともに記録媒体に記録する。

[0032] 図1に示すように、本実施形態に係るデジタルカメラ1は、概略的には、撮像部10と、画像信号処理部20と、表示部30と、記録媒体40と、收音部50と、音声信号処理部60と、制御部70と、操作部80とを備える。

[0033] 撮像部10は、被写体を撮像して、撮像画像を表すアナログ画像信号を出力する。撮像部10は、撮像光学系11と、撮像素子12と、タイミングジェネレータ13と、駆動装置14とを備える。

[0034] 撮像光学系11は、フォーカスレンズ、ズームレンズ、補正レンズ等の各種レンズや、不要な波長を除去する光学フィルタ、シャッター、絞り等の光学部品からなる。被写体から入射された光学像（被写体像）は、撮像光学系11における各光学部品を介して、撮像素子12の露光面に結像される。撮像素子12（イメージセンサ）は、例えば、CCD（Charge Coupled Device）又はCMOS（Complementary Metal Oxide Semiconductor）などの固体撮像素子で構成される。この撮像素子12は、撮像光学系11から導かれた光学像を光電変換し、撮像画像を表す電気信号（アナログ画像信号）を出力する。

[0035] 撮像光学系11には、該撮像光学系11の光学部品を駆動するための駆動

装置 14 が機械的に接続されている。この駆動装置 14 は、例えば、ズームモータ 15、フォーカスモータ 16、絞り機構（図示せず。）などを含む。駆動装置 14 は、後述する制御部 70 の指示に従って、撮像光学系 11 の光学部品を駆動させ、ズームレンズ、フォーカスレンズを移動させたり、絞りを調整したりする。例えば、ズームモータ 15 は、ズームレンズをテレ／ワイド方向に移動させることで、画角を調整するズーム動作を行う。また、フォーカスモータ 16 は、フォーカスレンズを移動させることで、被写体に焦点を合わせるフォーカス動作を行う。

[0036] また、タイミングジェネレータ 13（以下、TG13 という。）は、制御部 70 の指示に従って、撮像素子 12 に必要な動作パルスを生成する。例えば、TG13 は、垂直転送のための 4 相パルス、フィールドシフトパルス、水平転送のための 2 相パルス、シャッタパルスなどの各種パルスを生成し、撮像素子 12 に供給する。この TG13 により撮像素子 12 を駆動させることで、被写体像が撮像される。また、TG13 が、撮像素子 12 のシャッタースピードを調整することで、撮像画像の露光量や露光期間が制御される（電子シャッター機能）。上記の撮像素子 12 が出力した画像信号は画像信号処理部 20 に入力される。

[0037] 画像信号処理部 20 は、マイクロコントローラなどの電子回路で構成され、撮像素子 12 から出力される画像信号に対して所定の画像処理を施し、当該画像処理後の画像信号を表示部 30 や制御部 70 に出力する。画像信号処理部 20 は、アナログ信号処理部 21、アナログ／デジタル（A／D）変換部 22、デジタル信号処理部 23 を備える。

[0038] アナログ信号処理部 21 は、画像信号を前処理する所謂アナログフロントエンドである。該アナログ信号処理部 21 は、例えば、撮像素子 12 から出力される画像信号に対して、CDS（correlated double sampling：相関 2 重サンプリング）処理、プログラマブルゲインアンプ（PGA）によるゲイン処理などを行う。A／D 変換部 22 は、アナログ信号処理部 21 から入力されたアナログ画像信号をデジタル画像信号に

変換して、デジタル信号処理部23に出力する。デジタル信号処理部23は、入力されたデジタル画像信号に対して、例えば、ノイズ除去、ホワイトバランス調整、色補正、エッジ強調、ガンマ補正等のデジタル信号処理を行って、表示部30や制御部70等に出力する。

[0039] 表示部30は、例えば、液晶ディスプレイ（LCD: Liquid Crystal Display）、有機ELディスプレイなどの表示装置で構成される。表示部30は、制御部70による制御に従って、入力された各種の画像データを表示する。例えば、表示部30は、撮像中に画像信号処理部20からリアルタイムで入力される撮像画像（スルー画像）を表示する。これにより、ユーザは、デジタルカメラ1で撮像中のスルー画像を見ながら、デジタルカメラ1を操作することができる。また、記録媒体40に記録されている撮像画像を再生したときに、表示部30は、当該再生画像を表示する。これにより、ユーザは、記録媒体40に記録されている撮像画像の内容を確認することができる。

[0040] 記録媒体40は、上記撮像画像のデータ、音声データ、それらのメタデータなどの各種のデータを記憶する。記録媒体40は、例えば、メモリカード等の半導体メモリ、又は、光ディスク、ハードディスク等のディスク状記録媒体などを使用できる。なお、光ディスクは、例えば、ブルーレイディスク（Blu-ray Disc）、DVD（Digital Versatile Disc）又はCD（Compact Disc）等を含む。なお、記録媒体40は、デジタルカメラ1に内蔵されてもよいし、デジタルカメラ1に着脱可能なリムーバブルメディアであってもよい。

[0041] 收音部50は、デジタルカメラ1周辺の外部音声を收音する。本実施形態に係る收音部50は、1つの外部音声収録用のマイクロホン51からなるモノラルマイクロホンであるが、2つのマイクロホンからなるステレオマイクロホンで構成されてもよい。マイクロホン51は、外部音声を收音して得られた音声信号をそれぞれ出力する。かかる收音部50により、動画撮像中に外部音声を收音して、動画と共に記録できるようになる。かかるマイクロホ

ン５１は、外部音声（所望音）を收音するためにデジタルカメラ１の筐体に設けられているが、当該筐体内に設けられた発音部（上記駆動装置１４）の機械駆動音も雑音として收音してしまう。

[0042] 音声信号処理部６０は、マイクロコントローラなどの電子回路で構成され、音声信号に対して所定の音声処理を施して、記録用の音声信号を出力する。この音声処理は、例えば、ＡＤ変換処理、雑音低減処理などを含む。本実施形態は、この音声信号処理部６０による雑音低減処理を特徴としているが、その詳細説明は後述する。

[0043] 制御部７０は、マイクロコントローラなどの電子回路で構成され、デジタルカメラ１の全体の動作を制御する。制御部７０は、例えば、ＣＰＵ７１、ＥＥＰＲＯＭ（Ｅｌｅｃｔｒｉｃａｌｌｙ Ｅｒａｓａｂｌｅ Ｐｒｏｇｒａｍｍａｂｌｅ ＲＯＭ）７２、ＲＯＭ（Ｒｅａｄ Ｏｎｌｙ Ｍｅｍｏｒｙ）７３、ＲＡＭ（Ｒａｎｄｏｍ Ａｃｃｅｓｓ Ｍｅｍｏｒｙ）７４を備える。かかる制御部７０は、デジタルカメラ１内の各部を制御する。例えば、制御部７０は、マイクロホン５１により收音された音声信号から、駆動装置１４で発生した機械音を雑音として低減するに、音声信号処理部６０の動作を制御する。

[0044] 制御部７０におけるＲＯＭ７３には、ＣＰＵ７１に各種の制御処理を実行させるためのプログラムが格納されている。ＣＰＵ７１は、該プログラムに基づいて動作して、ＲＡＭ７４を用いながら、上記各制御のための必要な演算・制御処理を実行する。該プログラムは、デジタルカメラ１に内蔵された記憶装置（例えば、ＥＥＰＲＯＭ７２、ＲＯＭ７３等）に予め格納しておくことができる。また、当該プログラムは、ディスク状記録媒体、メモリカードなどのリムーバブル記録媒体に格納されて、デジタルカメラ１に提供されてもよいし、ＬＡＮ、インターネット等のネットワークを介してデジタルカメラ１にダウンロードされてもよい。

[0045] ここで、制御部７０による制御の具体例について説明する。制御部７０は、上記撮像部１０のＴＧ１３や駆動装置１４を制御して、撮像部１０による

撮像処理を制御する。例えば、制御部 70 は、上記撮像光学系 11 の絞りの調整、撮像素子 12 の電子シャッタースピードの設定、アナログ信号処理部 21 の A G C のゲイン設定などにより、自動露光制御を行う（A E 機能）。また、制御部 70 は、上記撮像光学系 11 のフォーカスレンズを移動させて、フォーカスポジションを変更することで、特定の被写体に対して撮像光学系 11 の焦点を自動的に合わせるオートフォーカス制御を行う（A F 機能）。また、制御部 70 は、上記撮像光学系 11 のズームレンズを移動させて、ズームポジションを変更することで、撮像画像の画角を調整する。また、制御部 70 は、記録媒体 40 に対して撮像画像、メタデータなどの各種のデータを記録し、また、記録媒体 40 に記録されているデータを読み出して再生する。さらに、制御部 70 は、表示部 30 に表示するための各種の表示画像を生成し、表示部 30 を制御して該表示画像を表示させる。

[0046] 操作部 80、表示部 30 は、ユーザがデジタルカメラ 1 の動作を操作するためのユーザインターフェースとして機能する。操作部 80 は、ボタン、レバー等の各種の操作キー、又はタッチパネル等で構成され、例えば、ズームボタン、シャッターボタン、電源ボタンなどを含む。操作部 80 は、ユーザ操作に応じて、各種の撮像動作を指示するための指示情報を制御部 70 に出力する。

[0047] [1. 2. 2. 音声信号処理装置の機能構成]

次に、図 2 を参照して、本実施形態に係るデジタルカメラ 1 に適用された音声信号処理装置の機能構成例について説明する。図 2 は、本実施形態に係る音声信号処理装置 100 の機能構成を示すブロック図である。

[0048] 図 2 に示すように、音声信号処理装置 100 は、信号入力部 110 と、A D 変換部 120 と、特徴量抽出部 130 と、雑音判定部 140 と、判定係数保持部 150 と、雑音低減部 160 と、信号記録部 170 とを備える。信号入力部 110 は、上記図 1 のマイクロホン 51 を備える。また、A D 変換部 120、特徴量抽出部 130、雑音判定部 140、判定係数保持部 150 及び雑音低減部 160 は、上記図 1 の音声信号処理部 60 を構成する。信号記

録部 170 は、上記図 1 の制御部 70 及び記録媒体 40 で構成される。

[0049] これら音声信号処理装置 100 の各部は、専用のハードウェアで構成されてもよいし、ソフトウェアで構成されてもよい。ソフトウェアを用いる場合、音声信号処理装置 100 のプロセッサが、以下に説明する各機能部の機能を実現するためのプログラムを実行すればよい。当該プログラムは、コンピュータ読み取り可能な記録媒体（例えば、光ディスク、ハードディスク、半導体メモリ等）を介して音声信号処理装置 100 に提供されてもよいし、又は各種の通信手段を介して提供されてもよい。以下に、音声信号処理装置 100 の各部について説明する。

[0050] 信号入力部 110 は、デジタルカメラ 1 の筐体に設置されたマイクロホン 51 で構成される。信号入力部 110 は、マイクロホン 51 は、デジタルカメラ 1 の周囲の所望音（録音対象の音声）を收音し、当該外部音声を音声信号に変換して出力する。この音声信号には、所望音のみならず、デジタルカメラ 1 の駆動装置 14 で発生するパルス機械音やその他の機械駆動音などの雑音が混入する。

[0051] AD変換部 120 は、上記マイクロホン 51 から出力されたアナログ音声信号を、デジタル音声信号に変換する。当該デジタル音声信号は、特徴量抽出部 130 及び雑音低減部 160 に入力される。

[0052] 特徴量抽出部 130 は、上記マイクロホン 51 から出力された音声信号から、上記パルス機械音を表す特徴量を抽出する。特徴量抽出部 130 は、上述した（１）パルス機械音のパルス成分を表すパルス特徴量を抽出するパルス特徴量抽出部、及び（２）パルス機械音の残響成分を表す残響特徴量を抽出する残響特徴量抽出部として機能する。具体的には、特徴量抽出部 130 は、パルス特徴量抽出部として、振幅最大値取得部 131 及びパルス幅計算部 132 を備え、残響特徴量抽出部として残響パワー計算部 133 を備える。

[0053] 振幅最大値取得部 131 は、音声信号のパルス成分の振幅（信号レベル）の最大値を取得し、当該振幅最大値 A をパルス特徴量として出力する。パル

ス幅計算部 132 は、音声信号のパルス成分のパルス幅 W を計算し、当該パルス幅 W をパルス特徴量として出力する。残響パワー計算部 133 は、音声信号の残響成分を表す狭帯域信号のパワー P を計算し、当該残響パワー P を残響特徴量として出力する。

[0054] ここで、図 3 を参照して、パルス機械音について説明する。図 3 は、本実施形態に係るパルス機械音を含む音声信号の波形図である。

[0055] 図 3 A に示すように、デジタルカメラ 1 による録音中に駆動装置 14 が動作開始すると、駆動装置 14 からパルス機械音が発生し、外部音声（背景音）の音声信号にパルス機械音が雑音として混入する。パルス機械音は、振幅が瞬間的に立ち上がるパルス成分と、当該パルス成分に続く残響成分とを含む。

[0056] パルス機械音のパルス成分は、パルス機械音の発生当初に信号レベル（振幅）がパルス状に立ち上がる部分である。このパルス成分の信号レベル（振幅）の最大値が振幅最大値 A である（図 3 C 参照。）。また、パルス幅 W は、パルス成分の時間幅であり、例えば、信号レベル（振幅）が上記振幅最大値 A から半分になるまでの時間幅である（図 3 C 参照。）。

[0057] 一方、パルス機械音の残響成分は、パルス成分に後続する特定の低周波数帯域の狭帯域信号からなる部分であり、パルス機械音の残響に応じた信号特性を有する。デジタルカメラ 1 の駆動装置 14 等から発生する駆動雑音は、デジタルカメラ 1 の本体の共振成分を多く含み、駆動装置 14 の動作開始時に上記パルス成分が発生した後に、当該パルス成分の共振成分等からなる残響成分が発生する。このようなパルス機械音の残響成分は、一般的な背景音とは異なる特定の周波数帯域、例えば 200 Hz 以下の低周波数帯域にパワーが集中する。この残響成分を表す狭帯域信号のパワーが残響成分パワー P である。

[0058] このように、パルス機械音はパルス成分及び残響成分という特徴的な成分を含むため、これら 2 種類の成分を検出することができれば、パルス機械音の有無を正確に検出できる。そこで、特徴量抽出部 130 は、入力された音

声信号から上記3つの特徴量（A，W，P）を計算する。この際、デジタル音声信号の1サンプルデータごとに特徴量を計算することは困難であるので、特徴量抽出部130は、所定量（例えば1フレーム）の音声信号のデータをバッファ（図示せず。）に保存し、当該1フレームの音声信号のデータの特徴量を計算する。ここで、音声信号の1フレームは、例えば、512サンプル又は1024サンプルなどであってよいが、かかる例に限定されず、任意のデータ数であってよい。

[0059] 振幅最大値取得部131は、音声信号の1フレーム中のデータの信号レベル（振幅）の絶対値を計算し、その中から最大値を抽出することで、振幅最大値Aを求める。パルス機械音のパルス成分の音量が大きいほど、振幅最大値Aも大きくなる。パルス幅計算部132は、音声信号の信号レベル（振幅）が上記振幅最大値Aから例えば半分になるまでの時間幅を計算することで、パルス幅Wを求める。パルス機械音のパルス成分が急峻であるほど、パルス幅Wは小さくなる。

[0060] 残響パワー計算部133は、特定の周波数帯域を通過させるバンドパスフィルタ（BPF）を用いて、パルス機械音の残響成分を表す狭帯域信号のパワーの値（残響成分パワーP）を計算する。例えば、残響パワー計算部133は、音声信号から200Hz以下の低周波数帯域の信号を取り出し、この信号の信号レベル値を積算することで、残響成分パワーPの値を求める。なお、残響パワー計算部133は、雑音の特性に応じて、当該200Hz以下の帯域とは異なる周波数帯域から残響成分パワーPを計算してもよい。また、残響パワー計算部133は、複数の周波数帯域から残響成分パワーPの値を計算してもよい。また、高速フーリエ変換（FFT）アルゴリズムを用いて、音声信号のパワースペクトルを算出することで、残響成分パワーPの値を計算してもよい。

[0061] 上記のようにして、特徴量抽出部130により得られた3つの特徴量（振幅最大値A、パルス幅W、残響成分パワーP）は、雑音判定部140に入力される。

[0062] 図2に戻り、雑音判定部140は、上記パルス機械音を表す特徴量（振幅最大値A、パルス幅W、残響成分パワーP）に基づいて、音声信号にパルス機械音が含まれるか否かを判定する。判定係数保持部150は、メモリ等の記憶装置で構成され、雑音判定部140によるパルス機械音の有無の判定に使用される判定係数を保持する。この判定係数は、デジタルカメラ1又は駆動装置14の特性、マイクロホン51と駆動装置14の位置関係等に応じて異なるため、機器ごとに発生するパルス機械音を予め学習・設定された判定係数が、判定係数保持部150に記憶される。

[0063] 雑音判定部140は、特徴量抽出部130により得られた3つの特徴量からなる特徴量ベクトル $\times (A, W, P)$ と、判定係数保持部に保持された判定係数を用いて、音声信号におけるパルス機械音の有無を総合的に判定する。この雑音判定部140により、音声信号にパルス機械音が含まれているか否かが判定され、音声信号におけるパルス機械音が含まれている区間が特定される。この判定方法の詳細については後述する。

[0064] 雑音低減部160は、上記雑音判定部140による判定結果に応じて、音声信号に対して雑音低減処理を行い、音声信号からパルス機械音等の雑音を除去する。具体的には、パルス機械音が含まれると判定された場合に、雑音低減部160は、当該パルス機械音が含まれる区間の音声信号に対して雑音低減処理を行う。一方、パルス機械音が含まれていないと判定された場合に、雑音低減部160は、雑音低減処理を行わない。このように、パルス機械音が含まれる場合にのみ、当該パルス機械音が含まれる区間（雑音区間）の音声信号に対して雑音低減処理を行うことで、雑音低減処理の処理効率を向上し、無駄な処理負荷を軽減できる。

[0065] この雑音低減部160による雑音低減処理方法としては、公知の任意の方法を用いることができるが、例えば次の処理方法がある。（1）雑音低減部160は、雑音区間の音量（信号レベル）を低減するゲインコントロールを行ってもよい。（2）また、雑音低減部160は、予め作成されたパルス機械音のモデル信号を、入力された実際の音声信号から減算してもよい。（3

）また、雑音低減部 160 は、バンドストップフィルタにより、パルス機械音の周波数帯域（雑音帯域）を抑制してもよい。（４）また、雑音低減部 160 は、

パルス機械音が含まれる雑音期間の前後の音声信号から、当該雑音区間の背景音の信号波形を推定し、推定した信号を用いて、雑音区間の信号を補間してもよい。

[0066] 信号記録部 170 は、上記雑音低減部 160 によりパルス機械音が低減されたデジタル音声信号を記録媒体（図示せず。）に記録する。記録媒体は、ハードディスク、磁気テープ等の磁気記録媒体、DVD、ブルーレイディスク等の光記録媒体、フラッシュメモリ、USBメモリ等の半導体メモリなど、任意の記録媒体であってよい。

[0067] [1. 3. パルス機械音の判定方法の具体例]

次に、本実施形態に係る音声信号処理装置 100 の雑音判定部 140 によるパルス機械音の判定方法の具体例について説明する。以下では、統計的識別法を用いた判定方法と、テーブル判定を用いた判定方法を例示する。

[0068] [1. 3. 1. 統計的識別法を用いた判定方法]

雑音判定部 140 は、線形識別器又は人工ニューラルネットワークなどを用いた統計的識別法により、音声信号におけるパルス機械音の有無を判定することができる。例えば、線形識別器を用いた統計的識別法では、以下の式（１）を用いて、判定値 y として上記特徴量の加重平均を計算する。

[0069] [数1]

$$y = \sum_{i=0}^{I-1} w_i \cdot x_i + w_I \quad \cdots (1)$$

[0070] ここで、 i は、次元を表すインデックスであり、 I は、荷重係数ベクトル w_i の総次元数である。 x_i は、特徴量ベクトルを表す（例えば、上記特徴量抽出部 130 により抽出された３つの特徴量ベクトルは、 $x_0 = A$ 、 $x_1 = W$

、 $x_2 = P$ 、である。）。 w_i は、予め設定される荷重（上記判定係数の一例）であり、 w_1 は判定の閾値を決める定数である。

[0071] 雑音判定部140は、上記式(1)で求めた判定値 y を用いて、音声信号の1フレーム中におけるパルス機械音の有無を判定する。例えば、雑音判定部140は、判定値 y が所定の閾値 y_0 （例えば、 $y_0 = 0$ ）以上であるか否かを判定する。この結果、雑音判定部140は、 $y \geq y_0$ であれば、当該1フレーム中にパルス機械音があると判定し、 $y < y_0$ であれば、当該1フレーム中にパルス機械音が無いと判定する。

[0072] なお、荷重 w_i や閾値 y_0 は、上記判定係数に相当し、判定係数保持部150に予め保持される。荷重 w_i は、事前に収録された多数の学習用の雑音サンプルに基づいて作成される。例えば、当該多数の雑音サンプルに対して、雑音／非雑音の正解付与を行い、最急降下法又はニュートン法などの最適化法を用いることによって、雑音と非雑音の識別性能が最も高い荷重 w_i を設定することができる。

[0073] [1. 3. 2. テーブル判定を用いた判定方法]

雑音判定部140は、予め設定された特徴量の閾値を表すテーブルを用いて、音声信号におけるパルス機械音の有無を判定することができる。雑音判定部140は、例えば図4に示すようなテーブル142を用いて、パルス機械音の有無を判定する。

[0074] 図4に示すように、テーブル142には、各特徴量 A 、 W 、 P 、・・・の閾値 A_0 、 W_0 、 P_0 、・・・と、判定基準に関する情報が保存されている。例えば、振幅最大値 A の閾値 A_0 は25000であり、実際の振幅最大値 A が A_0 を超であれば、振幅最大値 A に関する判定が正（True）となるように設定されている。雑音判定部140は、特徴量抽出部130により計算された実際の振幅最大値 A 、パルス幅 W 、残響成分パワー P と、テーブル142に保持された閾値 A_0 、 W_0 、 P_0 を比較し、各特徴量の項目について正否を判定する。図示の例では、振幅最大値 A 及び残響成分パワー P の項目が正（True）と判定され、パルス幅 W の項目が否（False）と判定されている。

[0075] さらに、雑音判定部 140 は、正判定の項目数 z と、その閾値 z_0 を比較する。この結果、雑音判定部 140 は、 $z \geq z_0$ であれば、音声信号の 1 フレーム中にパルス機械音が有ると判定し、 $z < z_0$ であれば、当該 1 フレーム中にパルス機械音が無いと判定する。例えば $z_0 = 3$ である場合、図示の例では正と判定された項目数 z が 2 であるので、音声信号の 1 フレーム中にパルス機械音が無いと判定される。

[0076] 以上のような閾値 A_0 、 W_0 、 P_0 、 $\dots$ や、正判定の項目数 z の閾値 z_0 等は、上記判定係数に相当する。これらの判定係数や判定基準を含むテーブル 142 は、判定係数保持部 150 に予め保持される。テーブル 142 内の判定係数は、事前に収録された雑音サンプルに基づいて作成される。例えば、音声信号処理装置 100 の設計者は、実際のパルス機械音の特性等に応じて、これらの判定係数を任意に設定することができる。

[0077] [1. 4. 音声信号処理方法]

次に、図 5 を参照して、上記の音声信号処理装置 100 を用いた音声信号処理方法（機械音低減方法）について説明する。図 5 は、本実施形態に係る音声信号処理方法を示すフローチャートである。

[0078] 図 5 に示すように、本実施形態に係る音声信号処理装置 100 を具備するデジタルカメラ 1 による撮像及び録音中には、その周囲の外部音声マイクロホン 51 により收音され、音声信号が出力される。すると、音声信号処理装置 100 の AD 変換部 120 は、マイクロホン 51 から入力されたアナログ音声信号を、デジタル音声信号に変換する（S10）。

[0079] 次いで、特徴量抽出部 130 は、S10 で AD 変換された入力音声信号を読み込み、バッファに蓄積する（S12）。そして、雑音判定処理の処理単位である 1 フレーム分の音声信号のデータがバッファに蓄積された後に（S14）、特徴量抽出部 130 は、当該 1 フレーム分の音声信号のデータから、パルス機械音を表す特徴量を算出する（S16）。具体的には、特徴量抽出部 130 は、パルス特徴量として、パルス機械音のパルス成分の振幅最大値 A 及びパルス幅 W を算出し、残響特徴量として、パルス機械音の残響成分

パワーPを算出する。

[0080] さらに、雑音判定部140は、S16で得られた特徴量A、W、Pと、上記判定係数を用いて、入力音声信号の1フレーム中にパルス機械音が含まれるか否かを判定する(S18)。この結果、パルス機械音が含まれると判定された場合、雑音低減部160は、入力音声信号の当該1フレームに対して雑音低減処理を実行して、パルス機械音を含む雑音を除去する(S20)。その後、信号記録部170は、S20で雑音低減された音声信号を記録媒体に記録する(S22)。

[0081] 一方、S18でパルス機械音が含まれないと判定された場合、雑音低減部160により入力音声信号の1フレームに対して雑音低減処理を実行することなく、信号記録部170は、当該入力音声信号の1フレームを記録媒体に記録する(S22)。

[0082] その後、デジタルカメラ1による撮像及び録音動作が終了(S24)するまで、入力音声信号の次の1フレームに対して、上記S10～S22の処理が繰り返される。

[0083] [1.5. 効果]

以上、本開示の第1の実施形態に係る音声信号処理装置100の構成と、これを用いた音声信号処理方法について説明した。本実施形態によれば、音声信号処理装置100は、マイクロホン51から入力された音声信号から、パルス機械音を表す3つの特徴量を抽出し、当該特徴量を用いてパルス機械音の有無を判定する。これにより、雑音発生源である駆動装置14の駆動制御情報を用いることなく、実際に収音された入力音声信号のみを用いて、パルス機械音の有無やその区間を正確に検出することができる。

[0084] そして、音声信号処理装置100は、音声信号のうち上記パルス機械音が検出された雑音区間のみに対して、雑音低減処理を施す。これにより、実際のパルス機械音が含まれる雑音区間に雑音低減処理区間を正確に合わせることができるので、非雑音区間で低減処理を行ったり、逆に、雑音区間で低減処理を行わなかったりすることがない。従って、音声信号の非雑音区間にお

ける音質劣化の恐れのない、高精度の雑音低減処理を実現でき、收音された外部音声（所望音）の忠実性を向上できる。

[0085] さらに、上記雑音判定処理では、音声信号の音量（信号レベル）に関する特徴量（振幅最大値 A ）だけでなく、パルス機械音のパルス幅 W や、パルス機械音の残響成分パワー P を用いることで、人間の耳に知覚されうるパルス機械音を高精度で検出できる。従って、雑音が人間の耳に知覚されない非雑音区間に対して雑音低減処理を実行したり、逆に、雑音が知覚される雑音区間に対して低減処理を実行しなかったりすることがない。よって、外部音声の所望音（背景音）の品質を劣化させることなく、耳障りなパルス機械音等の雑音のみを好適に除去することが可能となる。

[0086] 加えて、雑音発生源である駆動装置 14 の駆動制御情報を使用しないので、音声信号処理装置 100 及びデジタルカメラ 1 のシステム構成を簡素化できる。また、当該駆動制御情報を使用せずに、記録された音声信号のデータのみから雑音（パルス機械音）を検出するので、音声信号記録装置による音声信号の記録時のみならず、音声信号再生装置による音声信号の再生時にも、雑音の有無の判定及び低減処理を実行できる。

[0087] < 2. 第 2 の実施の形態 >

次に、本開示の第 2 の実施形態に係る音声信号処理装置及び音声信号処理方法について説明する。第 2 の実施形態に係る音声信号処理装置は、残響特徴量（第 2 の特徴量）として、残響成分パワー P の代わりに、残響成分を表す信号の零交差回数 N を用いることを特徴としている。なお、第 2 の実施形態のその他の機能構成は、上記第 1 の実施形態と実質的に同一であるので、その詳細説明は省略する。

[0088] [2. 1. 音声信号処理装置の機能構成]

まず、図 6 を参照して、第 2 の実施形態に係る音声信号処理装置 100 の機能構成について説明する。図 6 は、第 2 の実施形態に係る音声信号処理装置 100 の機能構成を示すブロック図である。

[0089] 図 6 に示すように、音声信号処理装置 100 は、第 1 の実施形態と同様、

信号入力部 110 と、AD 変換部 120 と、特徴量抽出部 130 と、雑音判定部 140 と、判定係数保持部 150 と、雑音低減部 160 と、信号記録部 170 とを備える。そして、特徴量抽出部 130 は、パルス特徴量抽出部として、第 1 の実施形態と同様に振幅最大値取得部 131 及びパルス幅計算部 132 を備えるが、残響特徴量抽出部として零交差回数計算部 134 を備えることを特徴としている。

[0090] 零交差回数計算部 134 は、パルス機械音の残響成分を表す狭帯域信号の零交差回数 N を計算する。ここで、零交差回数 N は、当該狭帯域信号の零交差点（ゼロクロスポイント）の数である。零交差点は、音声信号の時間波形において信号値が正の値から負の値に、又は負の値から正の値に変化する点を示す。零交差点では、音声信号の信号値はゼロとなる。この零交差点の数が多いほど、音声信号がより高い周波数成分を有することになり、零交差点の数が少ないほど、音声信号がより低い周波数成分を有することになる。このように、零交差点の数は、音声信号に含まれる低周波数成分を示すパラメータとなる。

[0091] 上記のように、パルス機械音の残響成分を表す狭帯域信号は、特有の低周波数帯域（例えば 200 Hz 以下）の信号であり、通常背景音とは周波数帯域が相違する。従って、所定区間における音声信号の零交差回数 N をカウントすれば、当該零交差回数 N により、その区間にパルス機械音の残響成分が含まれるか否かを判定することができる。

[0092] そこで、零交差回数計算部 134 は、例えば、音声信号の所定区間（例えば 1 フレーム）から例えば 200 Hz 以下の低周波数帯域の信号を取り出し、この信号の零交差点の数をカウントすることで、零交差回数 N を求める。そして、零交差回数計算部 134 は、当該零交差回数 N を残響特徴量（第 2 の特徴量）として雑音判定部 140 に出力する。

[0093] なお、零交差回数計算部 134 は、雑音の特性に応じて、上記 200 Hz 以下の帯域とは異なる周波数帯域の信号の零交差回数 N を計算してもよい。また、零交差回数計算部 134 は、複数の周波数帯域の信号の零交差回数 N

を計算してもよい。また、零交差回数計算部 134 は、上記零交差回数 N に代えて、零交差回数 N を音声信号の 1 フレーム中のサンプル数で除算した値を、残響特徴量として用いてもよい。

[0094] 雑音判定部 140 は、上記パルス機械音を表す特徴量（振幅最大値 A 、パルス幅 W 、零交差回数 N ）に基づいて、音声信号にパルス機械音が含まれるか否かを判定する。このとき、雑音判定部 140 は、上記零交差回数計算部 134 により得られた零交差回数 N と、予め設定された零交差回数 N の閾値 N_0 とを比較し、 $N_0 > N$ であれば、その区間における音声信号にパルス機械音の残響成分が存在すると判定する。

[0095] [2. 2. パルス機械音の判定方法の具体例]

次に、第 2 の実施形態に係る音声信号処理装置 100 の雑音判定部 140 によるパルス機械音の判定方法の具体例について説明する。第 2 の実施形態でも、第 1 の実施形態と同様に、例えば、統計的識別法を用いた判定方法、又はテーブル判定を用いた判定方法等を用いることができる。

[0096] [2. 2. 1. 統計的識別法を用いた判定方法]

雑音判定部 140 は、線形識別器又は人工ニューラルネットワークなどを用いた統計的識別法により、音声信号におけるパルス機械音の有無を判定することができる。例えば、線形識別器を用いた統計的識別法では、上述した式 (1) を用いて、判定値 y として上記特徴量の加重平均を計算する。このとき、式 (1) 中の x_i としては、上記特徴量抽出部 130 により抽出された 3 つの特徴量のベクトルである $\{x_i = (A, W, N)\}$ を用いる。その他の点は、第 1 の実施形態と同様であるので、詳細説明を省略する。

[0097] [2. 2. 2. テーブル判定を用いた判定方法]

雑音判定部 140 は、予め設定された特徴量の閾値を表すテーブルを用いて、音声信号におけるパルス機械音の有無を判定することができる。雑音判定部 140 は、例えば図 7 に示すようなテーブル 144 を用いて、パルス機械音の有無を判定する。

[0098] 図 7 に示すように、テーブル 144 には、各特徴量 A 、 W 、 N 、・・・の

閾値 A_0 、 W_0 、 N_0 、・・・と、判定基準に関する情報が保存されている。例えば、零交差回数 N の閾値 N_0 は 50 であり、実際の零交差回数 N が N_0 未満であれば、零交差回数 N に関する判定が正 (True) となるように設定されている。雑音判定部 140 は、特徴量抽出部 130 により計算された実際の振幅最大値 A 、パルス幅 W 、零交差回数 N と、テーブル 144 に保持された閾値 A_0 、 W_0 、 N_0 を比較し、各特徴量の項目について正否を判定する。図示の例では、振幅最大値 A 及び零交差回数 N の項目が正 (True) と判定され、パルス幅 W の項目が否 (False) と判定されている。

[0099] さらに、第 1 の実施形態と同様に、雑音判定部 140 は、正判定の項目数 z と、その閾値 z_0 を比較し、その比較結果に基づいて、音声信号の 1 フレーム中にパルス機械音が含まれるか否かを判定する。なお、閾値 A_0 、 W_0 、 N_0 、・・・や、正判定の項目数 z の閾値 z_0 等は、パルス機械音の有無を判定するための判定係数に相当する。

[0100] [2. 3. 効果]

以上、本開示の第 2 の実施形態に係る音声信号処理装置 100 の構成と、これを用いた雑音判定方法について説明した。なお、音声信号処理装置 100 を用いた音声信号処理方法の手順は第 1 の実施形態と同様である (図 5 参照)。

[0101] 第 2 の実施形態によれば、残響特徴量として零交差回数 N を用いることで、パルス機械音の残響成分の有無を適切に判定できるので、音声信号にパルス機械音が含まれるか否かを正確に検出できる。従って、第 2 の実施形態でも、上述した第 1 の実施形態の効果と同様の効果が得られる。

[0102] <3. 第 3 の実施の形態>

次に、本開示の第 3 の実施形態に係る音声信号処理装置及び音声信号処理方法について説明する。第 3 の実施形態に係る音声信号処理装置は、上記 3 つの特徴量 A 、 W 、 P だけでなく、その他の追加特徴量 (第 3 の特徴量) をも用いることを特徴としている。なお、第 3 の実施形態のその他の機能構成は、上記第 1 の実施形態と実質的に同一であるので、その詳細説明は省略す

る。

[0103] [3. 1. 音声信号処理装置の機能構成]

まず、図8を参照して、第3の実施形態に係る音声信号処理装置100の機能構成について説明する。図8は、第3の実施形態に係る音声信号処理装置100の機能構成を示すブロック図である。

[0104] 図8に示すように、音声信号処理装置100は、第1の実施形態と同様、信号入力部110と、AD変換部120と、特徴量抽出部130と、雑音判定部140と、判定係数保持部150と、雑音低減部160と、信号記録部170とを備える。そして、特徴量抽出部130は、振幅最大値取得部131、パルス幅計算部132及び残響パワー計算部133に加えて、1又は2以上の追加特徴量計算部135-1~N（Nは任意の自然数）を備えることを特徴としている。

[0105] 追加特徴量計算部135-1~N（以下、追加特徴量計算部135と総称する。）は、上記第1の実施形態で用いた基本特徴量（振幅最大値A、パルス幅W、残響成分パワーP）以外にも、追加特徴量Bをそれぞれ計算する。

[0106] ここで、追加特徴量Bは、例えば、（a）音声信号の二乗平均平方根（RMS: Root Mean Square）、（b）当該RMSの変化値、（c）上記パルス機械音の残響成分を表す狭帯域信号の振幅変化値、（d）当該狭帯域信号の零交差回数N、（e）零交差回数Nの変化値などである。追加特徴量計算部135は、これら各種の追加特徴量の一部のみを計算してもよいし、全部を計算してもよい。

[0107] （a）音声信号のRMSは、音声信号のある区間内の平均エネルギーを表し、パルス機械音のパルス成分を表すパルス特徴量に相当する。音声信号に雑音が含まれる場合には、RMSは、雑音やその近傍の音声信号のエネルギーを表すことになる。従って、音声信号にパルス機械音が含まれる場合には、そのパルス成分の区間のRMSが大きくなり、その前後のRMSが小さくなるので、当該RSMを用いて、パルス成分の有無を検出することができる。

- [0108] (b) 音声信号のRMSの変化値は、上記RMSの時間推移を表し、パルス機械音のパルス成分を表すパルス特徴量に相当する。このRMSの変化値によって、パルス機械音が含まれる区間とその前後の区間との間で音声信号のエネルギー変化を比較して、パルス機械音の有無をより正確に検出したり、パルス機械音のパルス成分の信号波形などを細かく分析したりすることができる。
- [0109] (c) 狭帯域信号の振幅変化値は、パルス機械音の残響成分の振幅の時間推移を表し、パルス機械音の残響成分を表す残響特徴量に相当する。この狭帯域信号の振幅変化値によって、パルス機械音の残響成分が含まれる区間とその前後の区間との間で音声信号の振幅変化を比較して、パルス機械音の有無をより正確に検出したり、パルス機械音の残響成分の信号波形などを細かく分析したりすることができる。
- [0110] (d) 狭帯域信号の零交差回数Nは、パルス機械音の残響成分のエネルギーを表し、パルス機械音の残響成分を表す残響特徴量に相当する。第2の実施形態で説明したように、零交差回数Nによって、パルス機械音の残響成分の有無をより正確に検出可能となる。
- [0111] (e) 零交差回数Nの変化値は、パルス機械音の残響成分を表す狭帯域信号の零交差回数Nの時間推移を表し、パルス機械音の残響成分を表す残響特徴量に相当する。この零交差回数Nの変化値によって、パルス機械音の残響成分が含まれる区間とその前後の区間との間で音声信号の零交差点の数を比較して、パルス機械音の有無をより正確に検出したり、パルス機械音の残響成分の信号波形などを細かく分析したりすることができる。
- [0112] このように、追加特徴量計算部135は、各種の追加特徴量Bを算出し、算出した追加特徴量Bを雑音判定部140に出力する。雑音判定部140は、上記3つの基本特徴量（振幅最大値A、パルス幅W、残響成分パワーP）及び少なくとも1つの追加特徴量B（第1、第2及び第3の特徴量に相当する。）を用いて、パルス機械音の有無を判定する。
- [0113] [3. 2. パルス機械音の判定方法の具体例]

次に、第3の実施形態に係る音声信号処理装置100の雑音判定部140によるパルス機械音の判定方法の具体例について説明する。第3の実施形態でも、第1の実施形態と同様に、例えば、統計的識別法を用いた判定方法、又はテーブル判定を用いた判定方法等を用いることができる。

[0114] [3. 2. 1. 統計的識別法を用いた判定方法]

統計的識別法を用いた判定方法では、上記式(1)中の x_i として、上記3つの基本特徴量 A 、 W 、 P と、少なくとも1つの追加特徴量 B を含む4以上の特徴量のベクトル $x_i = (A, W, P, B_1, B_2, \dots, B_N)$ を用いる。その他の点は、第1の実施形態と同様であるので、詳細説明を省略する。

[0115] なお、各々の特徴量の重要度や、パルス機械音の特性に応じて、式(1)における各特徴量の荷重 w_i を変更してもよい。例えば、追加特徴量 B の荷重 w_i よりも基本特徴量 A 、 W 、 P の荷重 w_i を大きくしてもよい。また、複数の追加特徴量 B_1 、 B_2 、 $\dots$ 、 B_N を用いる場合には、重要な追加特徴量 B (例えば零交差回数 N)ほど、荷重 w_i を大きくしてもよい。これにより、パルス機械音の有無の判定精度を向上することができる。

[0116] [3. 2. 2. テーブル判定を用いた判定方法]

また、テーブル判定を用いた判定方法では、第1の実施形態と同様に、予め設定された特徴量の閾値を表すテーブル(図4参照。)を用いて、音声信号におけるパルス機械音の有無を判定する。このテーブルには、各特徴量 A 、 W 、 P 、 B_1 、 B_2 、 $\dots$ 、 B_N の閾値 A_0 、 W_0 、 N_0 、 B_{1_0} 、 B_{2_0} 、 $\dots$ 、 B_{N_0} と、判定基準に関する情報が保存されている。雑音判定部140は、実際の特徴量 A 、 W 、 P 、 B_1 、 B_2 、 $\dots$ 、 B_N と、テーブル144に保持された閾値 A_0 、 W_0 、 N_0 、 B_{1_0} 、 B_{2_0} 、 $\dots$ 、 B_{N_0} をそれぞれ比較し、各特徴量の項目について正否を判定する。さらに、雑音判定部140は、正判定の項目数 z と、その閾値 z_0 を比較し、その比較結果に基づいて、音声信号の1フレーム中にパルス機械音が含まれるか否かを判定する。この判定方法は、第1の実施形態と同様であるので、詳細説明を省略する。

[0117] [3. 3. 効果]

以上、本開示の第3の実施形態に係る音声信号処理装置100の構成と、これを用いた雑音判定方法について説明した。なお、音声信号処理装置100を用いた音声信号処理方法の手順は第1の実施形態と同様である（図5参照。）。

[0118] 第3の実施形態によれば、基本特徴量A、W、Pに加えて、追加特徴量Bを用いて、パルス機械音の有無を判定する。従って、第3の実施形態では、上述した第1の実施形態の効果に加え、パルス機械音の有無をより高精度で検出でき、パルス機械音の信号波形等もより正確に検出できる効果がある。

[0119] なお、上記第3の実施形態では、基本特徴量として残響成分パワーPを用い、追加特徴量として零交差回数Nを用いる例を説明したが、本技術はかかる例に限定されない。例えば、基本特徴量として、残響成分パワーPに代えて零交差回数N又はその変化値等を用い、追加特徴量として残響成分パワーP又はその変化値等を用いてもよい。

[0120] <4. 第4の実施の形態>

次に、本開示の第4の実施形態に係る音声信号処理装置及び音声信号処理方法について説明する。第4の実施形態に係る音声信号処理装置は、音声信号の複数の区間から複数の特徴量を抽出し、当該複数の特徴量に基づいて、パルス機械音の有無を判定することを特徴としている。この際、上記複数の特徴量を演算し（例えば、加算和又は平均値を計算し）、この演算値を新たな特徴量として用いて、パルス機械音の有無を判定してもよい。なお、第4の実施形態のその他の機能構成は、上記第1又は第3の実施形態と実質的に同一であるので、その詳細説明は省略する。

[0121] [4. 1. 音声信号処理装置の機能構成]

まず、図9を参照して、第4の実施形態に係る音声信号処理装置100の機能構成について説明する。図9は、第4の実施形態に係る音声信号処理装置100の機能構成を示すブロック図である。

[0122] 図9に示すように、音声信号処理装置100は、第1の実施形態と同様、

信号入力部 110 と、AD 変換部 120 と、特徴量抽出部 130 と、雑音判定部 140 と、判定係数保持部 150 と、雑音低減部 160 と、信号記録部 170 とを備える。さらに、音声信号処理装置 100 は、特徴量抽出部 130 及び雑音判定部 140 に接続された特徴量保持部 152 を備えることを特徴としている。

[0123] 第 4 の実施形態では、特徴量抽出部 130 は、特徴量を計算するための音声信号の区間（フレーム）を任意の位置に設定可能である。そして、特徴量抽出部 130 は、音声信号の任意の位置にある複数の区間（複数のフレーム）からそれぞれ特徴量を算出し、雑音判定部 140 に出力する。雑音判定部 140 は、当該複数の特徴量を特徴量保持部 152 に保存する。特徴量保持部 152 は、半導体メモリ等の記憶装置からなり、特徴量抽出部 130 により算出された特徴量を一時的に保持する。特徴量抽出部 130 が特徴量を保持する期間は任意であり、記憶データ量を抑制するために、古い特徴量のデータを順次削除してもよい。

[0124] なお、特徴量抽出部 130 は、上記各種の特徴量（例えば、振幅最大値 A、パルス幅 W、残響成分パワー P 及び追加特徴量 B）の全てについて、音声信号の複数の区間から計算してもよい。或いは、特徴量抽出部 130 は、ある特徴量（例えば、振幅最大値 A 及びパルス幅 W）については、音声信号の複数の区間から計算し、他の特徴量（例えば、残響成分パワー P 及び追加特徴量 B）については、音声信号の 1 つの区間から計算してもよい。

[0125] このようにして、音声信号の複数の区間（フレーム）からそれぞれ算出した特徴量は、特徴量保持部 152 に保持される。雑音判定部 140 は、音声信号の複数の区間から算出した特徴量を特徴量保持部 152 から読み出し、当該複数の特徴量に基づいて、音声信号の有無を判定する。この際、雑音判定部 140 は、当該複数の特徴量の加算和又は平均値等を演算し、当該演算値に基づいて、音声信号の有無を判定してもよい。

[0126] なお、第 4 の実施形態でも、第 1 の実施形態と同様に、雑音判定部 140 によりパルス機械音の有無を判定するために、例えば、統計的識別法を用い

た判定方法、又はテーブル判定を用いた判定方法等を用いればよい。これらの判定方法は、上記第1又は第3の実施形態と同様であるので、詳細説明を省略する。

[0127] [4. 2. 効果]

以上、本開示の第4の実施形態に係る音声信号処理装置100の構成について説明した。なお、音声信号処理装置100を用いた音声信号処理方法の手順は第1の実施形態と同様である（図5参照。）。

[0128] 第4の実施形態によれば、音声信号の複数の区間から算出した複数の特徴量を用いて、音声信号の有無を判定する。特に、これら複数の特徴量の加算和又は平均値等を演算することで、当該演算値を新たな特徴量とし、当該特徴量に基づいて音声信号の有無を判定することが好ましい。

[0129] これにより、判定の安定性を高めて、判定精度を向上できる。例えば、音声信号にパルス機械音以外の他の雑音が含まれるときや、背景音が変化するときなどには、音声信号の1つの区間からのみでは、適正な特徴量を算出できない場合もありうる。これに対し、本実施形態のように複数の区間の音声信号から複数の特徴量を算出して演算することにより、音声信号の異常値や特異値の影響を抑制できる。従って、適正な特徴量を安定的に算出できるので、パルス機械音の有無の判定精度を向上できる。

[0130] なお、上記第4の実施形態では、特徴量抽出部130の後段に特徴量保持部152を設け、計算後の特徴量を保持する構成であったが、本技術はかかる例に限定されない。例えば、特徴量抽出部130の前段に、AD変換後の音声信号の一定区間のデータを保持する信号保持部を設け、特徴量抽出部130は、当該信号保持部から音声信号の任意の区間のデータを読み出すことで、音声信号の複数の区間の特徴量を算出してもよい。これによっても、雑音判定部140は、音声信号の複数の区間の特徴量に基づいて、パルス機械音の有無を判定することが可能となる。

[0131] <5. 第5の実施の形態>

次に、本開示の第5の実施形態に係る音声信号処理装置及び音声信号処理

方法について説明する。第5の実施形態に係る音声信号処理装置は、パルス機械音の発生源である駆動装置14の制御信号を取得し、駆動装置14の駆動期間中にだけ、上記特徴量抽出処理と雑音判定処理を行うことを特徴としている。なお、第5の実施形態のその他の機能構成は、上記第1の実施形態と実質的に同一であるので、その詳細説明は省略する。

[0132] [5. 1. 音声信号処理装置の機能構成]

まず、図10を参照して、第5の実施形態に係る音声信号処理装置100の機能構成について説明する。図10は、第5の実施形態に係る音声信号処理装置100の機能構成を示すブロック図である。

[0133] 図10に示すように、音声信号処理装置100は、第1の実施形態と同様、信号入力部110と、AD変換部120と、特徴量抽出部130と、雑音判定部140と、判定係数保持部150と、雑音低減部160と、信号記録部170とを備える。さらに、音声信号処理装置100は、特徴量抽出部130に接続された制御信号取得部180を備えることを特徴としている。

[0134] パルス機械音の発生源である駆動装置14は、デジタルカメラ1の制御部70（図1参照。）が生成する制御信号により制御される。この制御信号は、駆動装置14の駆動を制御するための駆動制御情報に相当し、例えば、デジタルカメラ1のズームモータ15又はフォーカスモータ16等を駆動させるための制御信号である。この制御信号により、駆動装置14の駆動の有無を判断し、当該駆動装置14からパルス機械音が発生し得る期間を検出することができる。

[0135] 駆動装置14が動作していないとき（非駆動期間）には、パルス機械音が発生することがないので、上述した特徴量抽出処理や雑音判定処理を実行する必要がない。一方、駆動装置14が動作しているとき（駆動期間）には、パルス機械音が発生するので、上述した特徴量抽出処理や雑音判定処理を実行する必要がある。そこで、制御信号を取得して駆動装置14の駆動期間を検出できれば、その駆動期間にのみこれらの処理を実行することで、演算量を節減することが可能となる。

[0136] そこで、第5の実施形態に係る音声信号処理装置100は、制御部70等から上記制御信号を取得して、当該制御信号に基づいて、駆動装置14の駆動期間にのみ前述した特徴量抽出処理及び雑音判定処理を実行する。詳細には、図10に示すように、音声信号処理装置100は、駆動装置14を制御するための制御信号を取得する制御信号取得部180を備える。この制御信号取得部180は、上記制御部70から送信された制御信号を受信して、特徴量抽出部130に出力する。

[0137] 特徴量抽出部130は、不図示のバッファメモリ等の記憶部を具備しており、マイクロホン51から入力される音声信号をバッファリングしている。また、特徴量抽出部130は、制御信号取得部180により制御信号を取得したか否かを監視している。そして、特徴量抽出部130は、駆動装置14の駆動開始を表す制御信号が取得された場合に、バッファメモリに保存された音声信号を読み出して、上記各種の特徴量を算出する処理を実行開始し、その後、駆動装置14の駆動終了を表す制御信号が取得された場合に、当該処理の実行を終了する。このように、特徴量抽出部130は、制御信号を用いて、駆動装置14が駆動する期間を判断しており、当該駆動装置14の駆動期間にのみ特徴量抽出処理を実行する。同様に、雑音判定部140も、制御信号により駆動装置14の駆動期間を判断しており、当該駆動装置14の駆動期間にのみ、上記特徴量抽出部130から得られる特徴量を用いて、上述した雑音判定処理を実行する。

[0138] [5. 2. 音声信号処理方法]

次に、図11を参照して、第5の実施形態に係る音声信号処理装置100を用いた音声信号処理方法（機械音低減方法）について説明する。図11は、本実施形態に係る音声信号処理方法を示すフローチャートである。

[0139] 図11に示すように、第5の実施形態に係る音声信号処理装置100は、まず、マイクロホン51から入力された音声信号をAD変換し（S10）、バッファに蓄積する（S12）。このS10～S12の処理は第1の実施形態と同様である。

[0140] 次いで、雑音判定処理の処理単位である１フレーム分の音声信号のデータがバッファに蓄積された後に（Ｓ１４）、特徴量抽出部１３０は、雑音発生源である駆動装置１４が動作中であるか否か、即ち、駆動期間であるか否かを判断する（Ｓ１５）。例えば、Ｓ１５にて、特徴量抽出部１３０は、デジタルカメラ１本体の制御部７０から、駆動装置１４の制御信号を受信しているか否かに基づいて、駆動期間であるか否かを判定する。制御信号を受信している期間は、駆動装置１４が動作しており、駆動期間であり、一方、制御信号を受信していない期間は、駆動装置１４が動作しておらず、駆動期間でないと判定する。

[0141] この結果、駆動装置１４の駆動期間であると判定された場合には、第１の実施形態と同様に、特徴量抽出部１３０は、上記入力音声信号の１フレーム分の音声信号のデータから、パルス機械音を表す特徴量を算出し（Ｓ１６）、雑音判定部１４０は、Ｓ１６で得られた特徴量Ａ、Ｗ、Ｐと、上記判定係数を用いて、当該１フレーム中にパルス機械音が含まれるか否かを判定する（Ｓ１８）。この結果、パルス機械音が含まれると判定された場合、雑音低減部１６０は、入力音声信号の当該１フレームに対して雑音低減処理を実行して、パルス機械音を含む雑音を除去する（Ｓ２０）。その後、信号記録部１７０は、Ｓ２０で雑音低減された音声信号を記録媒体に記録する（Ｓ２２）。

[0142] 一方、Ｓ１５で駆動装置１４の駆動期間でないと判定された場合には、第１の実施形態とは異なり、雑音判定部１４０による特徴量の算出処理や、雑音判定部１４０によるパルス機械音の有無の判定処理を行わない。そして、信号記録部１７０は、入力音声信号をそのまま記録媒体に記録する（Ｓ２２）。

[0143] その後、デジタルカメラ１による撮像及び録音動作が終了（Ｓ２４）するまで、入力音声信号の次の１フレームに対して、上記Ｓ１０～Ｓ２２の処理が繰り返される。

[0144] [５．３．効果]

以上、本開示の第5の実施形態に係る音声信号処理装置100の構成と、それを用いた音声信号処理方法について説明した。

[0145] 第5の実施形態によれば、駆動装置14の制御情報を用いて、駆動装置14の駆動の有無を判定し、駆動装置14の動作期間にのみ、特徴量抽出処理及び雑音判定処理を実行する。これにより、駆動装置14の非駆動期間中には、これらの演算処理を行わなくてすむので、その分の演算量を節減することができる。特に、デジタルカメラ1の撮像動作期間中に、ズームモータ15等の駆動装置14が頻繁に動作することは稀であり、撮像動作期間全体のうち、駆動装置14の動作期間は僅かな時間に限られる。従って、上記の処理を駆動期間にのみ実施し、非駆動期間に実施しないことで、演算量を大幅に削減可能である。

[0146] また、駆動装置14の非駆動期間に上記処理を行うことで、パルス機械音が存在しないのに存在すると判定してしまう雑音判定エラーが生じる恐れがある。しかし、第5の実施形態によれば、かかる雑音判定エラーも低減できるので、原音に忠実な高品質な音声信号を記録することができる。

[0147] <6. 第6の実施の形態>

次に、本開示の第6の実施形態に係る音声信号処理装置及び音声信号処理方法について説明する。第6の実施形態に係る音声信号処理装置は、雑音発生源である駆動装置14の特性に応じて上記判定係数を変更し、当該判定係数及び特徴量を用いて、当該駆動装置14に特有のパルス機械音の有無を判定することを特徴としている。なお、第6の実施形態のその他の機能構成は、上記第1の実施形態と実質的に同一であるので、その詳細説明は省略する。

[0148] [6. 1. 雑音判定方法の概要]

上記パルス機械音の特性は、当該パルス機械音を発生させる発音部（駆動装置14等）の特性に応じて変化する。例えば、一眼レフカメラ等では、カメラ本体に対して複数種類の交換式のレンズ装置（例えば、図1の撮像部10に相当する。）を着脱可能である。この場合、装着されるレンズ装置の特

性（例えば、レンズを駆動させる駆動装置 14 の種類、配置等）に応じて、デジタルカメラ 1 で収音されるパルス機械音の特性が変化する。従って、デジタルカメラ 1 に実際に装着されたレンズ装置の特性に応じて、上述した雑音判定処理に用いる各種の係数、テーブル、特徴量の閾値等のパラメータ（以下「判定係数」という。）を設定することが好ましい。

[0149] そこで、第 6 の実施形態では、交換式のレンズ装置ごとに、当該各レンズ装置の特性に対応する判定係数を予め設定しておく。そして、実際の雑音判定処理では、デジタルカメラ 1 に装着されたレンズ装置に対応する判定係数を用いて、当該レンズ装置に特有の雑音（パルス機械音）が、入力音声信号に含まれるか否かを判定する。このように、雑音判定に用いる判定係数を変更可能とすることで、各レンズ装置に特有の雑音を高精度で検出可能となる。

[0150] ここで、判定係数は、パルス機械音の有無を判定するために用いられるパラメータである。例えば、判定係数は、上述した統計的識別法における数式（1）中の荷重 w_i 、 w_l 、閾値 y_0 や、テーブル判定方法におけるテーブル 142（図 4 等参照。）中の各特徴量の閾値 A_0 、 W_0 、 P_0 ・・・、判定の項目数 z の閾値 z_0 などを含む。各種のレンズ装置の駆動装置 14 から発生するパルス機械音を事前に分析し、この分析結果に応じて上記各種の判定係数を設定しておけば、レンズ装置ごとに適切な雑音低減処理を実現できる。

[0151] [6. 2. 音声信号処理装置の機能構成]

次に、図 12 を参照して、第 6 の実施形態に係る音声信号処理装置 100 の機能構成について説明する。図 12 は、第 6 の実施形態に係る音声信号処理装置 100 の機能構成を示すブロック図である。

[0152] 図 12 に示すように、音声信号処理装置 100 は、第 1 の実施形態と同様、信号入力部 110 と、AD 変換部 120 と、特徴量抽出部 130 と、雑音判定部 140 と、判定係数保持部 150 と、雑音低減部 160 と、信号記録部 170 とを備える。さらに、音声信号処理装置 100 は、複数の判定係数の中から、デジタルカメラ 1 に接続されたレンズ装置に対応する判定係数を

選択する判定係数選択部 154 を備えることを特徴としている。

[0153] 判定係数選択部 154 は、各種のレンズ装置ごとに予め設定された複数の判定係数を保持する。この判定係数選択部 154 に保持されている判定係数は、当該各レンズ装置の駆動装置 14 の特性や、当該駆動装置 14 が発生するパルス機械音の特性に応じて、予め設定されている。パルス機械音のパルス成分の振幅最大値やパルス幅、残響成分の周波数帯域、零交差回数等は、レンズ装置ごとに相異なる。そこで、各レンズ装置で発生するパルス機械音に特有のパルス成分や残響成分の特性に合わせて、判定係数の値が設定される。また、判定係数選択部 154 は、これらの各レンズ装置に対応する判定係数を、当該各レンズ装置の識別情報に関連づけて保持している。

[0154] また、判定係数選択部 154 は、実際にデジタルカメラ 1 に接続された 1 つのレンズ装置を表す情報（接続レンズ情報）を、当該レンズ装置又は制御部 70 等から取得する。この接続レンズ情報は、レンズ装置の種類ごとに固有に付与される識別情報（例えば、レンズ装置の型番、名称等）である。判定係数選択部 154 は、当該接続レンズ情報に基づいて、デジタルカメラ 1 に接続されたレンズ装置を特定し、判定係数保持部 150 に保持されている複数の判定係数の中から、当該レンズ装置に対応する 1 つの判定係数を選択する。

[0155] そして、雑音判定部 140 は、判定係数選択部 154 により選択された判定係数と、特徴量抽出部 130 により得られた特徴量とを用いて、上記雑音判定処理を行い、当該レンズ装置に特有のパルス機械音の有無を判定する。この雑音判定処理では、上記第 1 の実施形態と同様に、例えば、統計的識別法を用いた判定方法、又はテーブル判定を用いた判定方法等を用いることができる。

[0156] このように図 12 に示す構成例では、音声信号処理装置 100 内の判定係数保持部 150 に、デジタルカメラ 1 に装着可能な複数種類の交換式レンズ装置にそれぞれ対応する複数の判定係数を予め保持しておく。そして、実際の雑音判定処理時には、判定係数保持部 150 内の複数の判定係数の中から

、デジタルカメラ１に実際に装着された１つのレンズ装置に対応する判定係数を選択し、雑音判定処理に利用する。

[0157] [6. 3. 音声信号処理装置の変更例]

次に、図１３を参照して、第６の実施形態に係る音声信号処理装置１００の変更例について説明する。図１３は、第６の実施形態に係る音声信号処理装置１００の変更例を示すブロック図である。

[0158] 図１３に示す例では、音声信号処理装置１００は、上記図１２の例のような判定係数保持部１５０及び判定係数選択部１５４を備えておらず、雑音判定に必要な判定係数は、音声信号処理装置１００内ではなく、各々の交換式レンズ装置１５６Ａ～１５６Ｎ内に保持されている。この各レンズ装置１５６Ａ～１５６Ｎに保持されている判定係数は、上記と同様に、当該各レンズ装置１５６Ａ～１５６Ｎの特性や、そのパルス機械音の特性等に応じて予め設定されている。

[0159] そして、デジタルカメラ１にレンズ装置１５６Ａが接続されたときには、音声信号処理装置１００の雑音判定部１４０は、当該レンズ装置１５６Ａから、当該レンズ装置１５６Ａに対応する判定係数を取得する。例えば、デジタルカメラ１の制御部７０がレンズ装置１５６Ａの識別情報とともに判定係数を取得し、雑音判定部１４０は、当該制御部７０から送信された判定係数を受信する。さらに、雑音判定部１４０は、上記接続されたレンズ装置１５６Ａから取得した判定係数と、特徴量抽出部１３０により得られた特徴量を用いて、当該レンズ装置１５６Ａに特有のパルス機械音の有無を判定する。

[0160] このように、図１３の構成例では、雑音判定に必要な判定係数を各々のレンズ装置１５６内に保持しておく。そして、実際の雑音判定処理時には、デジタルカメラ１に実際に接続されたレンズ装置１５６Ａから判定係数を取得して、雑音判定処理に利用する。

[0161] [6. 4. 効果]

以上、本開示の第６の実施形態に係る音声信号処理装置１００の構成について説明した。なお、音声信号処理装置１００を用いた音声信号処理方法の

手順は第 1 の実施形態と同様である（図 5 参照。）。

[0162] 第 6 の実施形態によれば、デジタルカメラ 1 に接続される交換式のレンズ装置の特性に応じて、雑音判定処理に用いる判定係数を変更する。これにより、当該接続されたレンズ装置に特有のパルス機械音を高精度で判定及び検出できるので、雑音低減精度を大幅に向上できる。特に、一眼レフカメラなど、相異なる種類のレンズ装置を装着可能な機器では、レンズ装置ごとに異なるパルス機械音を適切に低減した上で、高品質の背景音を記録できる。

[0163] さらに、図 1 2 の示した例では、音声信号処理装置 1 0 0 自身の判定係数保持部 1 5 0 に、相異なる複数のレンズ装置に対応する判定係数を保持することで、これらレンズ装置の全てから発生するパルス機械音に柔軟に対応可能となる。また、判定係数保持部 1 5 0 に保持される判定係数を更新（追加、削除又は変更）することで、新たなタイプのレンズ装置にも対応可能なる。

[0164] さらに、図 1 3 の示した例では、各レンズ装置に適切な判定係数を保持させ、デジタルカメラ 1 に接続されたレンズ装置から判定係数を取得する。これにより、新たなタイプのレンズ装置 1 5 6 に対しても、既存のデジタルカメラ 1 に搭載された音声信号処理装置 1 0 0 の記憶部（ROM 等）に保持された判定係数を書き替えることなく、図 1 2 と同様に、レンズ装置ごとに判定係数を変更することが可能にある。このように、図 1 3 の例では、既存のデジタルカメラ 1 の内部情報を更新せずとも、交換式レンズ装置 1 5 6 の追加、削除又は変更等に柔軟に対応できるという利点がある。

[0165] なお、上記説明では、デジタルカメラ 1 に交換式のレンズ装置を着脱する場合に、判定係数を変更する例について説明したが、本技術はかかる例に限定されない。例えば、音声信号処理装置が搭載される機器に、相異なる特性のパルス機械音を発生させる複数の発音部（複数の駆動装置 1 4 等）が存在する場合や、当該機器に当該複数の発音部を選択的に着脱する場合などにも、上記と同様にして判定係数を変更すればよい。

[0166] < 7. 第 7 の実施の形態 >

次に、本開示の第 7 の実施形態に係る音声信号処理装置及び音声信号処理方法について説明する。上記では、音声信号処理装置 100 をデジタルカメラ 1 等の記録装置に設け、音声信号の記録時に雑音低減する例について説明した。これに対し、第 7 の実施形態では、音声信号処理装置を再生装置に設け、記録された音声信号の再生時に雑音低減することを特徴としている。なお、第 7 の実施形態のその他の機能構成は、上記第 1 の実施形態と実質的に同一であるので、その詳細説明は省略する。

[0167] [7. 1. 音声信号処理装置の機能構成]

次に、図 14 を参照して、第 7 の実施形態に係る音声信号処理装置 100 の機能構成について説明する。図 14 は、第 7 の実施形態に係る音声信号処理装置 100 の機能構成を示すブロック図である。

[0168] 図 14 に示すように、音声信号処理装置 100 は、信号記録部 190 及び信号再生部 200 を備える再生装置に適用され、信号記録部 190 から読み出された再生対象の音声信号に対して、特徴量抽出、雑音判定、雑音低減等の各種処理を実行することを特徴としている。音声信号処理装置 100 は、第 1 の実施形態と同様、特徴量抽出部 130 と、雑音判定部 140 と、判定係数保持部 150 と、雑音低減部 160 とを備える。さらに、音声信号処理装置 100 は、信号記録部 190 から、再生対象の音声信号を取得する信号取得部 192 を備える。

[0169] 信号記録部 190 は、ハードディスク、光ディスク、半導体メモリ、テープ等の記録媒体（図示せず。）を具備している。この記録媒体には、上記デジタルカメラ 1 等による撮像・録音時に、音声信号のデータが記録されている。なお、ズームモータ 15 又はフォーカスモータ 16 等の制御信号は、音声信号に関連付けて保存される必要はない。

[0170] そして、信号記録部 190 は、不図示の制御部からの指示に基づき、再生対象の音声信号を記録媒体から読み出す。信号取得部 192 は、信号記録部 190 から、再生対象の音声信号のデータを取得して、特徴量抽出部 130 及び雑音低減部 160 に出力する。このとき、信号取得部 192 は、例えば

、５１２サンプル又は１０２４サンプル程度の区間のデータを、１フレームとして取得する

[0171] その後、上記第１の実施形態と同様にして、特徴量抽出部１３０による特徴量の算出処理、雑音判定部１４０による雑音判定処理が行われる。この際、雑音判定に用いる判定係数は、判定係数保持部１５０に保持されていてもよいし、信号記録部１９０の記録媒体中に音声信号に関連づけて記録されていてもよい。

[0172] 上記の判定結果、再生対象の音声信号にパルス機械音が含まれていと判定された場合には、雑音低減部１６０により、音声信号からパルス機械音が低減される。かかるパルス機械音の低減後の音声信号は、信号再生部２００に出力される。信号再生部２００は、例えば、スピーカ、又はヘッドホン等の音声再生装置で構成され、上記雑音が低減された音声信号を再生する。

[0173] [７．２．効果]

以上、本開示の第７の実施形態に係る音声信号処理装置１００の構成について説明した。なお、音声信号処理装置１００を用いた音声信号処理方法の手順は第１の実施形態と同様である（図５参照。）。ただし、第７の実施形態では、音声再生時に雑音低減を行うため、図５のＡＤ変換処理（Ｓ１０）は不要であり、Ｓ２２では音声信号を再生し、再生対象の全ての音声信号のデータの再生が終了するまで（Ｓ２４）、Ｓ１２～Ｓ２２の処理を繰り返す。

[0174] 第７の実施形態によれば、音声信号の記録時（撮像、録音等）時だけでなく、記録された音声信号を再生するときにも、再生対象の音声信号に含まれるパルス機械音を適切に低減することができる。

[0175] < ８．まとめ >

以上、本開示の好適な実施形態に係る音声信号処理装置及び音声信号処理方法について説明した。本実施形態によれば、駆動装置１４の動作開始又は終了時にパルス機械音が発生した場合であっても、マイクロホン５１により収音された音声信号から、当該パルス機械音とその区間（雑音区間）を高精

度で検出することができる。そして、検出された雑音区間に雑音低減処理を施すことにより、非雑音区間における音質劣化の恐れのない、高品質な音声を提供することができる。

[0176] この際、駆動装置 14 の制御信号を用いずとも、実際に入力された音声信号のみを用いてパルス機械音を検出できる。従って、実際にパルス機械音が含まれる雑音期間と、雑音低減処理期間とを正確に合わせて、雑音低減処理を実現できる。よって、上述した制御信号と雑音区間の間の遅延による雑音低減精度の低下、音質の劣化の問題を解消できる。

[0177] さらに、パルス機械音に特有の各種の特徴量（振幅最大値 A、パルス幅 W、残響成分パワー P、零交差回数 N 等）を組み合わせ、総合的にパルス機械音の有無を判定することにより、人間の耳に知覚されるパルス機械音を正確に検出できる。従って、人間の耳に雑音が知覚されない非雑音区間において低減処理を行ったり、雑音が知覚される雑音区間において低減処理を行わないといった不具合を解消でき、高精度の雑音低減処理を実現できる。

[0178] また、駆動装置 14 の駆動制御情報（制御信号）を使用せずに、音声信号のみからパルス機械音を判定できるので、音声信号処理装置 100 が適用される機器のシステム構成を簡素化できる。また、音声信号の収音・記録時のみならず、記録媒体に記録された音声信号の再生時にも、雑音の有無の判定及び低減処理を実行できる。

[0179] 以上、添付図面を参照しながら本開示の好適な実施形態について詳細に説明したが、本技術はかかる例に限定されない。本開示の技術分野における通常の知識を有する者であれば、特許請求の範囲に記載された技術的思想の範疇内において、各種の変更例または修正例に想到し得ることは明らかであり、これらについても、当然に本開示の技術的範囲に属するものと了解される。

[0180] 例えば、上記実施形態では主に、音声信号処理装置としてデジタルカメラ 1 を例示し、動画撮像と共に録音する時に機械音を低減する例について説明したが、本技術はかかる例に限定されない。本技術の音声信号処理装置は、

各種の音声信号記録装置又は音声信号再生装置等の電子機器に適用できる。
例えば、音声信号処理装置は、記録再生装置（例えば、ブルーレイディスク／DVDレコーダ）、テレビジョン受像器、システムステレオ装置、撮像装置（例えば、デジタルカメラ、デジタルビデオカメラ）、携帯端末（例えば、携帯型音楽／映像プレーヤ、携帯型ゲーム機、ICレコーダ）、パーソナルコンピュータ、ゲーム機、カーナビゲーション装置、デジタルフォトフレーム、家庭電化製品、自動販売機、ATM、キオスク端末など、任意の電子機器に適用できる。

[0181] なお、本技術は以下のような構成も取ることができる。

[0182] （１）收音部により得られる音声信号から、前記收音部と同一の筐体に設けられた発音部から発生するパルス状の作動音を表す特徴量を抽出する特徴量抽出部と、

前記特徴量に基づいて、前記音声信号に前記作動音が含まれるか否かを判定する雑音判定部と、
を備える、音声信号処理装置。

[0183] （２）前記発音部は、前記收音部と同一の筐体に設けられた駆動装置であり、

前記作動音は、前記駆動装置の動作開始時又は動作終了時に発生するパルス状の機械駆動音である、前記（１）に記載の音声信号処理装置。

[0184] （３）前記特徴量抽出部は、前記特徴量として、前記作動音のパルス成分を表す第１の特徴量と、前記作動音の残響成分を表す第２の特徴量とを抽出し、

前記雑音判定部は、前記第１及び第２の特徴量に基づいて、前記音声信号に前記作動音が含まれるか否かを判定する、前記（１）又は（２）に記載の音声信号処理装置。

[0185] （４）前記第１の特徴量は、前記作動音のパルス成分の振幅最大値及びパルス幅を含み、

前記第２の特徴量は、前記作動音の残響成分を表す狭帯域信号のパワーを

含む、前記（３）に記載の音声信号処理装置。

[0186] （５）前記第１の特徴量は、前記作動音のパルス成分の振幅最大値及びパルス幅を含み、

前記第２の特徴量は、前記作動音の残響成分を表す狭帯域信号の零交差回数を含む、前記（３）に記載の音声信号処理装置。

[0187] （６）前記特徴量抽出部は、前記特徴量として、前記音声信号のＲＭＳ、前記ＲＭＳの変化値、前記作動音の残響成分を表す狭帯域信号の振幅変化値、前記狭帯域信号の零交差回数、又は前記零交差回数の変化値のうち一部又は全部を含む第３の特徴量をさらに抽出し、

前記雑音判定部は、前記第１、第２及び第３の特徴量に基づいて、前記音声信号に前記作動音が含まれるか否かを判定する、前記（３）～（５）のいずれか一項に記載の音声信号処理装置。

[0188] （７）前記特徴量抽出部は、前記音声信号の複数の区間から複数の前記特徴量を抽出し、

前記雑音判定部は、前記複数の区間から抽出された前記複数の特徴量に基づいて、前記音声信号に前記作動音が含まれるか否かを判定する、前記（１）～（６）のいずれか一項に記載の音声信号処理装置。

[0189] （８）前記駆動装置を制御するための制御信号を取得する制御信号取得部をさらに含み、

前記制御信号に基づいて前記駆動装置の駆動期間を判断し、前記駆動装置の駆動期間中にのみ、前記特徴量抽出部による前記特徴量の抽出処理と、前記雑音判定部による判定処理を実行する、前記（２）に記載の音声信号処理装置。

[0190] （９）前記雑音判定部は、

前記発音部の特性に応じて設定された判定係数、及び前記特徴量を用いて、前記発音部に特有の前記作動音が前記音声信号に含まれるか否かを判定する、前記（１）～（８）のいずれか一項に記載の音声信号処理装置。

[0191] （１０）前記音声信号処理装置は、撮像装置に設けられ、

前記発音部は、前記撮像装置の本体に着脱可能な交換式のレンズ装置であり、

前記音声信号処理装置は、

前記レンズ装置ごとに設定された複数の判定係数を保持する判定係数保持部と、

前記撮像装置に接続された前記レンズ装置を表す情報に基づいて、前記複数の判定係数の中から当該レンズ装置に対応する判定係数を選択する判定係数選択部と、

を更に備え、

前記雑音判定部は、

前記判定係数選択部により選択された前記判定係数、及び前記特徴量を用いて、前記撮像装置に接続された前記レンズ装置に特有の前記作動音が前記音声信号に含まれるか否かを判定する、前記（９）に記載に記載の音声信号処理装置。

[0192] （１１）前記音声信号処理装置は、撮像装置に設けられ、

前記発音部は、前記撮像装置の本体に着脱可能な交換式のレンズ装置であり、

前記レンズ装置は、当該レンズ装置の特性に応じて設定された判定係数を保持しており、

前記雑音判定部は、前記撮像装置に接続された前記レンズ装置から前記判定係数を取得し、前記取得した判定係数及び前記特徴量を用いて、前記撮像装置に接続された前記レンズ装置に特有の前記作動音が前記音声信号に含まれるか否かを判定する、前記（９）に記載に記載の音声信号処理装置。

[0193] （１２）前記音声信号処理装置は、前記音声信号を再生する再生装置に設けられ、

前記特徴量抽出部は、前記音声信号の再生時に前記音声信号から前記特徴量を抽出し、

前記雑音判定部は、前記特徴量に基づいて、前記音声信号に前記作動音が

含まれるか否かを判定する、前記（１）～（９）のいずれか一項に記載に記載の音声信号処理装置。

[0194] （１３）外部音声を音声信号に変換する收音部と、

前記收音部と同一の筐体に設けられ、パルス状の作動音を発生させる発音部と、

前記收音部から出力された前記音声信号から、前記作動音を表す特徴量を抽出する特徴量抽出部と、

前記特徴量に基づいて、前記音声信号に前記作動音が含まれるか否かを判定する雑音判定部と、

前記雑音判定部により前記作動音が含まれると判定された場合に、前記音声信号に対して雑音低減処理を行う雑音低減部と、
を備える、撮像装置。

[0195] （１４）收音部により得られる音声信号から、前記收音部と同一の筐体に設けられた発音部から発生するパルス状の作動音を表す特徴量を抽出することと、

前記特徴量に基づいて、前記音声信号に前記作動音が含まれるか否かを判定することと、

を含む、音声信号処理方法。

[0196] （１５）收音部により得られる音声信号から、前記收音部と同一の筐体に設けられた発音部から発生するパルス状の作動音を表す特徴量を抽出することと、

前記特徴量に基づいて、前記音声信号に前記作動音が含まれるか否かを判定することと、

をコンピュータに実行させるためのプログラム。

[0197] （１６）收音部により得られる音声信号から、前記收音部と同一の筐体に設けられた発音部から発生するパルス状の作動音を表す特徴量を抽出することと、

前記特徴量に基づいて、前記音声信号に前記作動音が含まれるか否かを判

定することと、

をコンピュータに実行させるためのプログラムが記録された、コンピュータ読み取り可能な記録媒体。

符号の説明

- [0198]
- 1 デジタルカメラ
 - 2 筐体
 - 3 レンズ部
 - 10 撮像部
 - 14 駆動装置
 - 15 ズームモータ
 - 16 フォーカスモータ
 - 51 マイクロホン
 - 60 音声信号処理部
 - 70 制御部
 - 100 音声信号処理装置
 - 110 信号入力部
 - 120 A/D変換部
 - 130 特徴量抽出部
 - 131 振幅最大値取得部
 - 132 パルス幅計算部
 - 133 残響パワー計算部
 - 134 零交差回数計算部
 - 135 追加特徴量計算部
 - 140 雑音判定部
 - 142、144 テーブル
 - 150 判定係数保持部
 - 152 特徴量保持部
 - 154 判定係数選択部

- 1 5 6 レンズ装置
- 1 6 0 雑音低減部
- 1 7 0、1 9 0 信号記録部
- 1 8 0 制御信号取得部
- 1 9 2 信号取得部
- 2 0 0 信号再生部

請求の範囲

- [請求項1] 收音部により得られる音声信号から、前記收音部と同一の筐体に設けられた発音部から発生するパルス状の作動音を表す特徴量を抽出する特徴量抽出部と、
- 前記特徴量に基づいて、前記音声信号に前記作動音が含まれるか否かを判定する雑音判定部と、
- を備える、音声信号処理装置。
- [請求項2] 前記発音部は、前記收音部と同一の筐体に設けられた駆動装置であり、
- 前記作動音は、前記駆動装置の動作開始時又は動作終了時に発生するパルス状の機械駆動音である、請求項 1 に記載の音声信号処理装置。
- [請求項3] 前記特徴量抽出部は、前記特徴量として、前記作動音のパルス成分を表す第 1 の特徴量と、前記作動音の残響成分を表す第 2 の特徴量とを抽出し、
- 前記雑音判定部は、前記第 1 及び第 2 の特徴量に基づいて、前記音声信号に前記作動音が含まれるか否かを判定する、請求項 1 に記載の音声信号処理装置。
- [請求項4] 前記第 1 の特徴量は、前記作動音のパルス成分の振幅最大値及びパルス幅を含み、
- 前記第 2 の特徴量は、前記作動音の残響成分を表す狭帯域信号のパワーを含む、請求項 3 に記載の音声信号処理装置。
- [請求項5] 前記第 1 の特徴量は、前記作動音のパルス成分の振幅最大値及びパルス幅を含み、
- 前記第 2 の特徴量は、前記作動音の残響成分を表す狭帯域信号の零交差回数を含む、請求項 3 に記載の音声信号処理装置。
- [請求項6] 前記特徴量抽出部は、前記特徴量として、前記音声信号の R M S、前記 R M S の変化値、前記作動音の残響成分を表す狭帯域信号の振幅

変化値、前記狭帯域信号の零交差回数、又は前記零交差回数の変化値のうち一部又は全部を含む第3の特徴量をさらに抽出し、

前記雑音判定部は、前記第1、第2及び第3の特徴量に基づいて、前記音声信号に前記作動音が含まれるか否かを判定する、請求項3に記載の音声信号処理装置。

[請求項7] 前記特徴量抽出部は、前記音声信号の複数の区間から複数の前記特徴量を抽出し、

前記雑音判定部は、前記複数の区間から抽出された前記複数の特徴量に基づいて、前記音声信号に前記作動音が含まれるか否かを判定する、請求項1に記載の音声信号処理装置。

[請求項8] 前記駆動装置を制御するための制御信号を取得する制御信号取得部をさらに含み、

前記制御信号に基づいて前記駆動装置の駆動期間を判断し、前記駆動装置の駆動期間中にのみ、前記特徴量抽出部による前記特徴量の抽出処理と、前記雑音判定部による判定処理を実行する、請求項2に記載の音声信号処理装置。

[請求項9] 前記雑音判定部は、

前記発音部の特性に応じて設定された判定係数、及び前記特徴量を用いて、前記発音部に特有の前記作動音が前記音声信号に含まれるか否かを判定する、請求項1に記載の音声信号処理装置。

[請求項10] 前記音声信号処理装置は、撮像装置に設けられ、

前記発音部は、前記撮像装置の本体に着脱可能な交換式のレンズ装置であり、

前記音声信号処理装置は、

前記レンズ装置ごとに設定された複数の判定係数を保持する判定係数保持部と、

前記撮像装置に接続された前記レンズ装置を表す情報に基づいて、前記複数の判定係数の中から当該レンズ装置に対応する判定係数を選

択する判定係数選択部と、

を更に備え、

前記雑音判定部は、

前記判定係数選択部により選択された前記判定係数、及び前記特徴量を用いて、前記撮像装置に接続された前記レンズ装置に特有の前記作動音が前記音声信号に含まれるか否かを判定する、請求項 9 に記載に記載の音声信号処理装置。

[請求項11]

前記音声信号処理装置は、撮像装置に設けられ、

前記発音部は、前記撮像装置の本体に着脱可能な交換式のレンズ装置であり、

前記レンズ装置は、当該レンズ装置の特性に応じて設定された判定係数を保持しており、

前記雑音判定部は、前記撮像装置に接続された前記レンズ装置から前記判定係数を取得し、前記取得した判定係数及び前記特徴量を用いて、前記撮像装置に接続された前記レンズ装置に特有の前記作動音が前記音声信号に含まれるか否かを判定する、請求項 9 に記載に記載の音声信号処理装置。

[請求項12]

前記音声信号処理装置は、前記音声信号を再生する再生装置に設けられ、

前記特徴量抽出部は、前記音声信号の再生時に前記音声信号から前記特徴量を抽出し、

前記雑音判定部は、前記特徴量に基づいて、前記音声信号に前記作動音が含まれるか否かを判定する、請求項 1 に記載の音声信号処理装置。

[請求項13]

外部音声を音声信号に変換する收音部と、

前記收音部と同一の筐体に設けられ、パルス状の作動音を発生させる発音部と、

前記收音部から出力された前記音声信号から、前記作動音を表す特

微量を抽出する特徴量抽出部と、

前記特徴量に基づいて、前記音声信号に前記作動音が含まれるか否かを判定する雑音判定部と、

前記雑音判定部により前記作動音が含まれると判定された場合に、前記音声信号に対して雑音低減処理を行う雑音低減部と、
を備える、撮像装置。

[請求項14] 収音部により得られる音声信号から、前記収音部と同一の筐体に設けられた発音部から発生するパルス状の作動音を表す特徴量を抽出することと、

前記特徴量に基づいて、前記音声信号に前記作動音が含まれるか否かを判定することと、
を含む、音声信号処理方法。

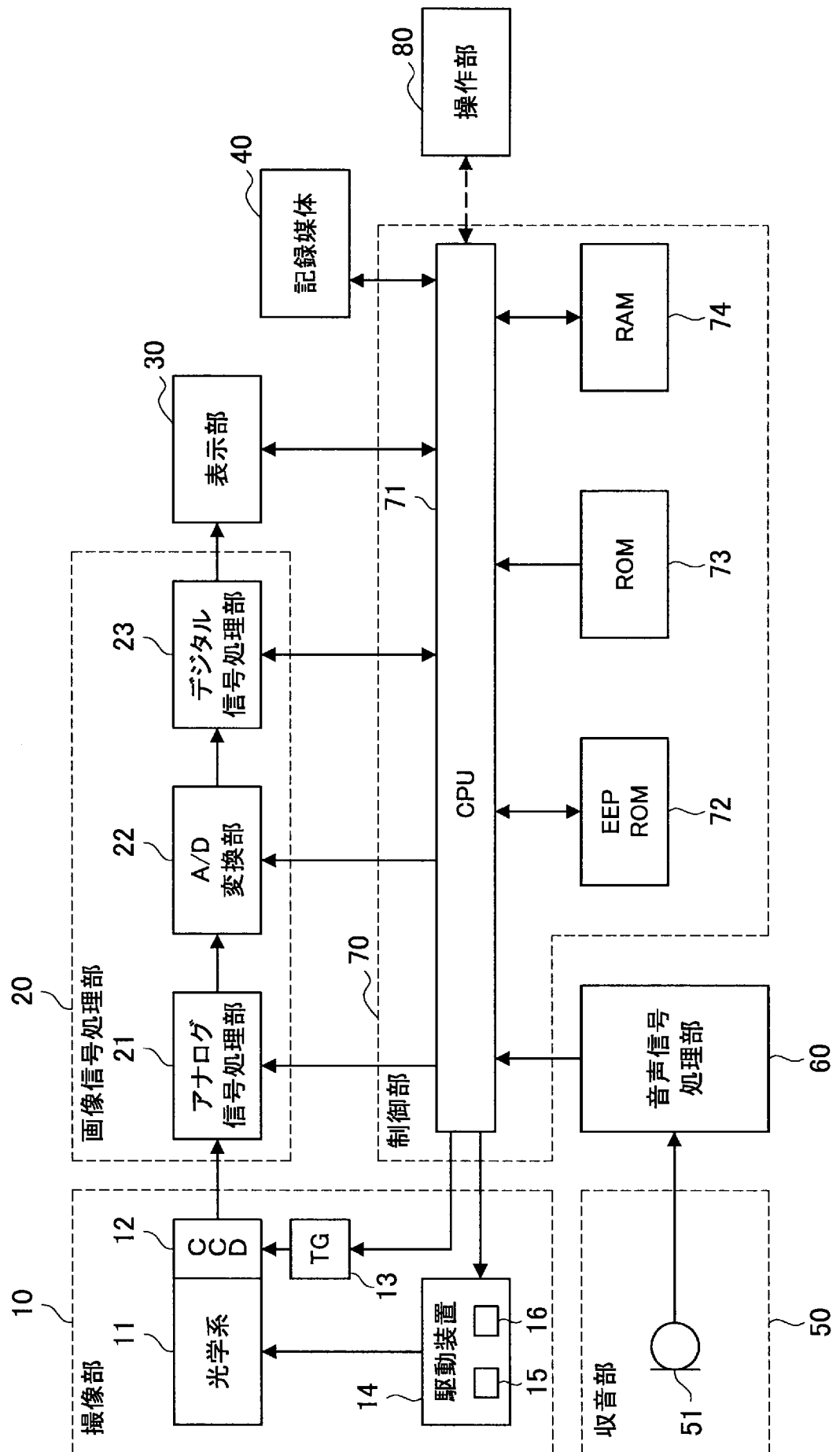
[請求項15] 収音部により得られる音声信号から、前記収音部と同一の筐体に設けられた発音部から発生するパルス状の作動音を表す特徴量を抽出することと、

前記特徴量に基づいて、前記音声信号に前記作動音が含まれるか否かを判定することと、
をコンピュータに実行させるためのプログラム。

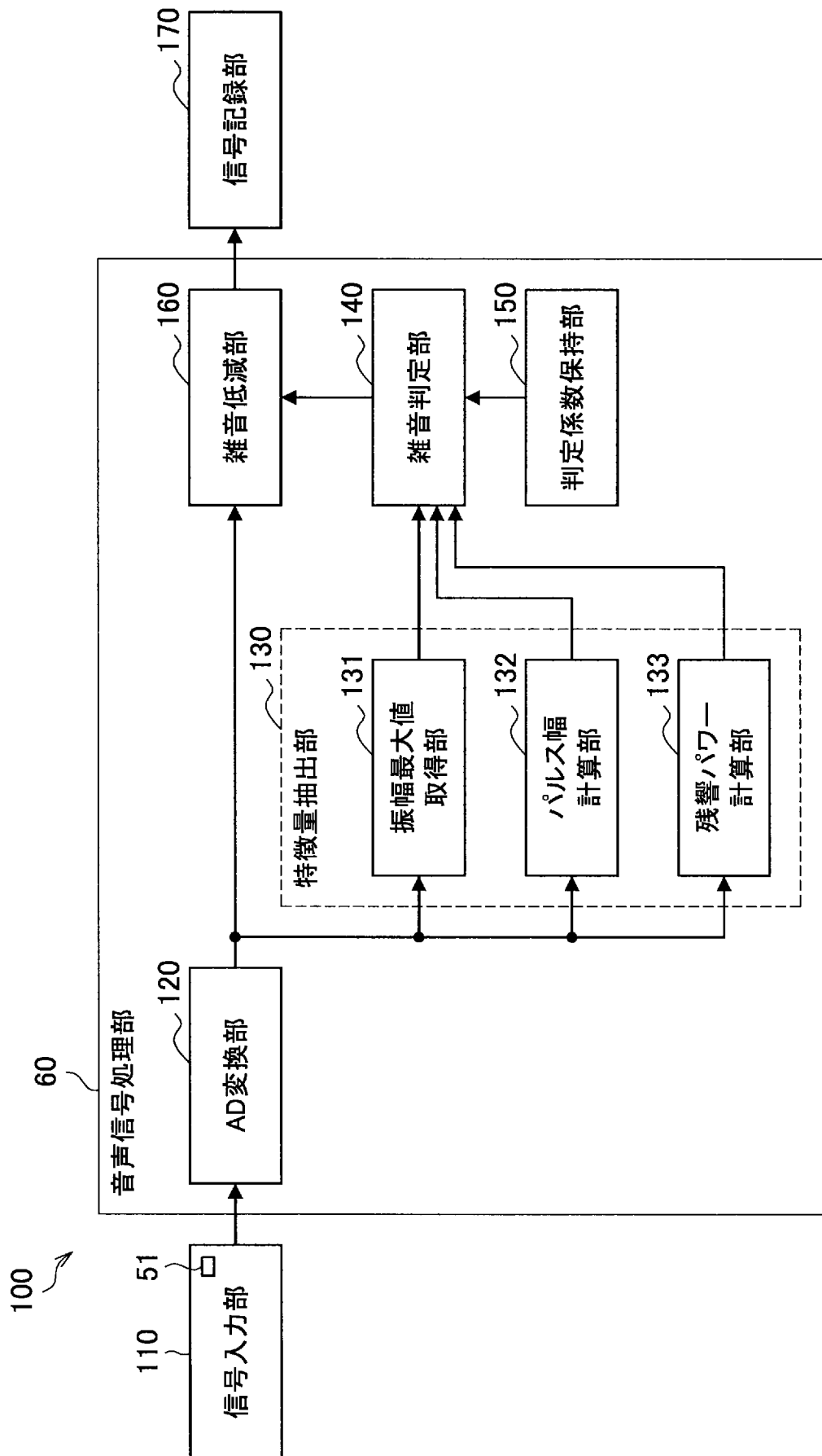
[請求項16] 収音部により得られる音声信号から、前記収音部と同一の筐体に設けられた発音部から発生するパルス状の作動音を表す特徴量を抽出することと、

前記特徴量に基づいて、前記音声信号に前記作動音が含まれるか否かを判定することと、
をコンピュータに実行させるためのプログラムが記録された、コンピュータ読み取り可能な記録媒体。

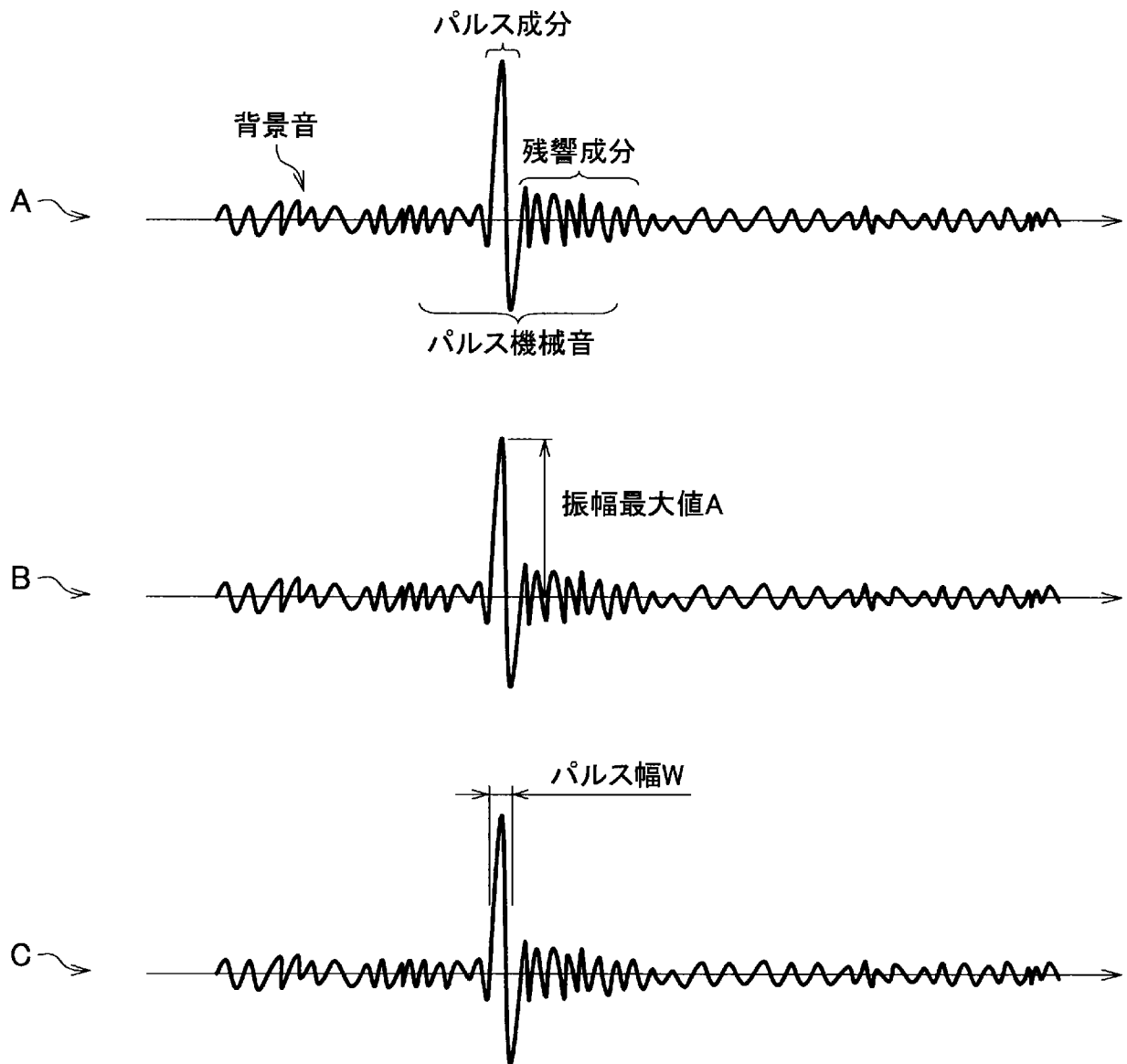
[図1]



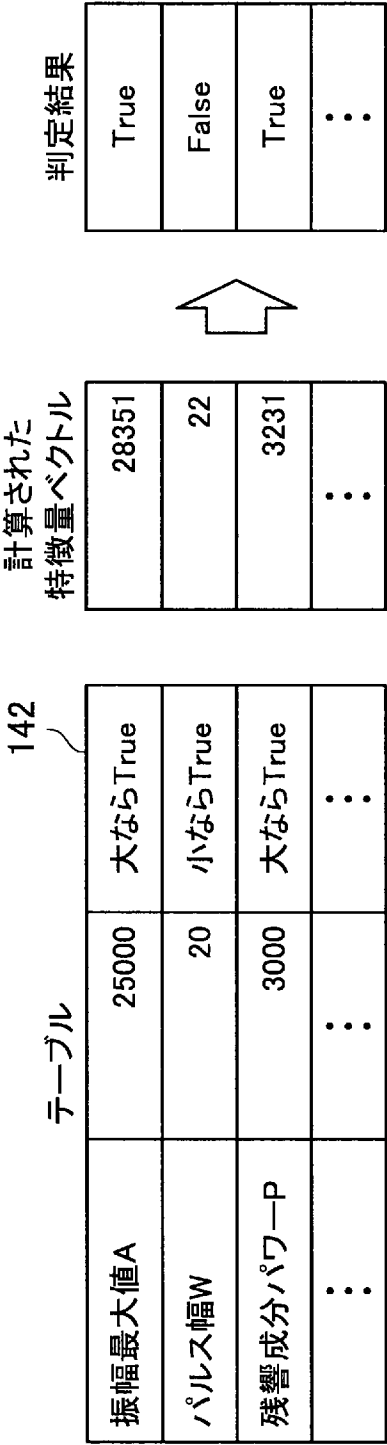
[図2]



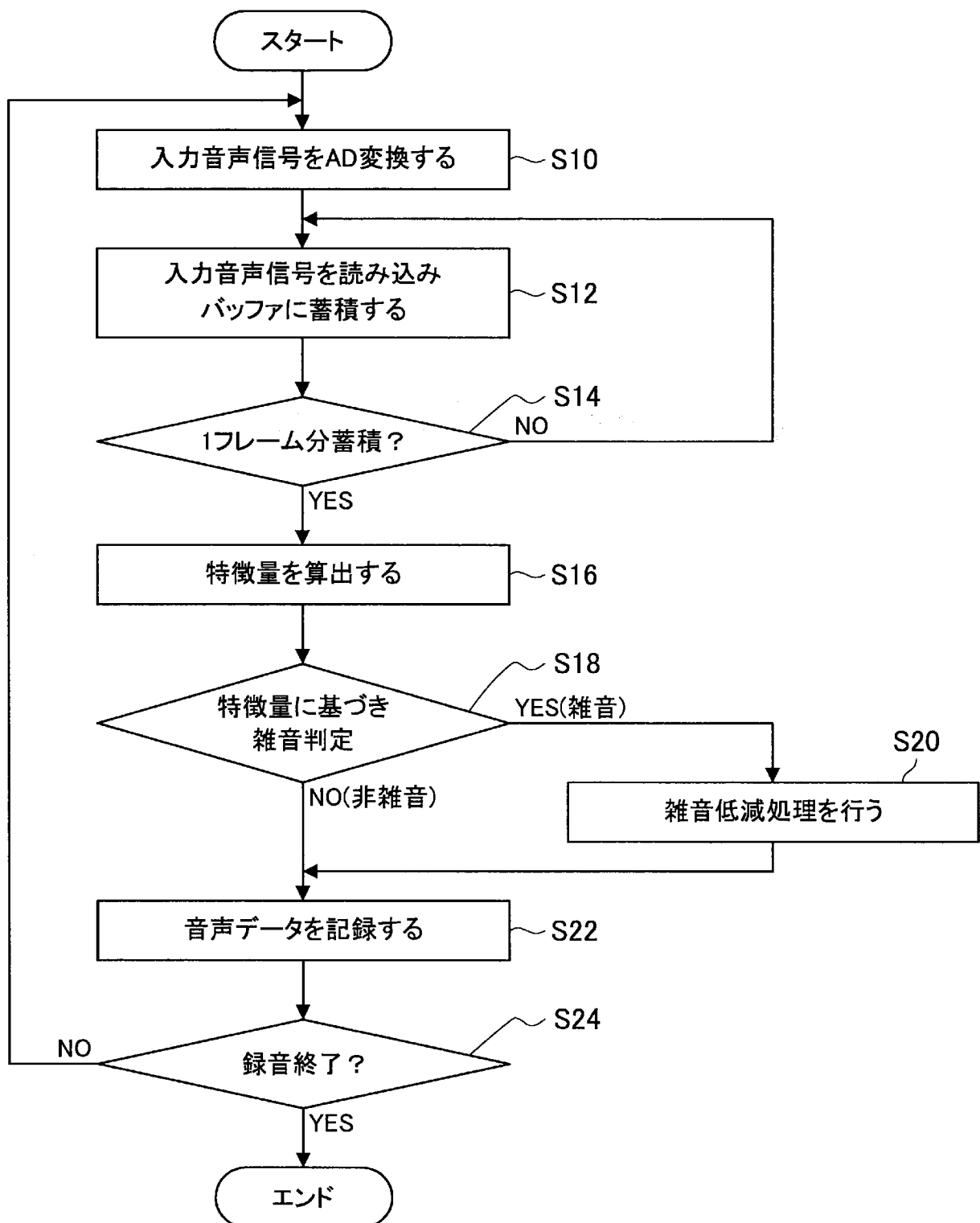
[図3]



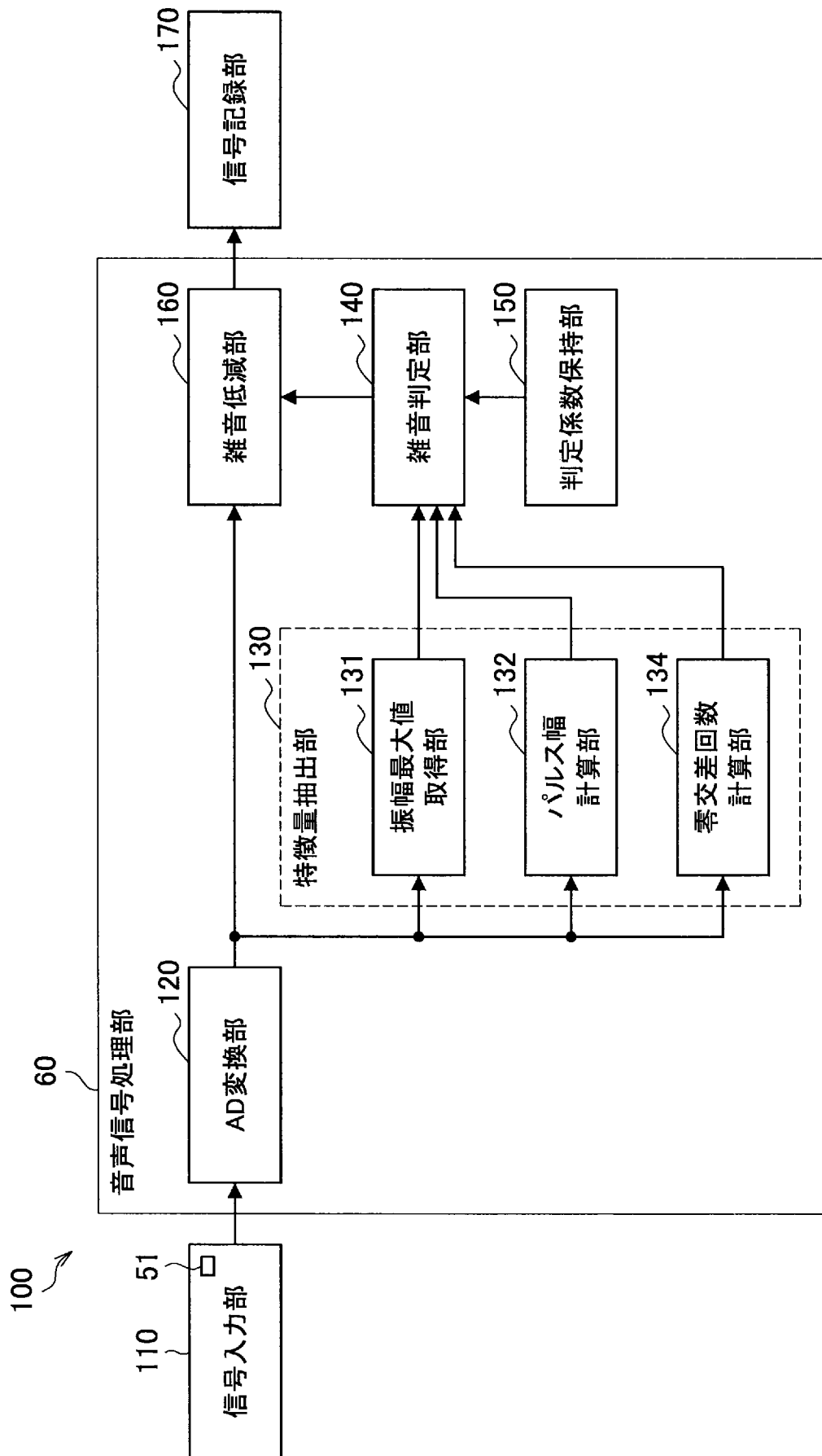
[図4]



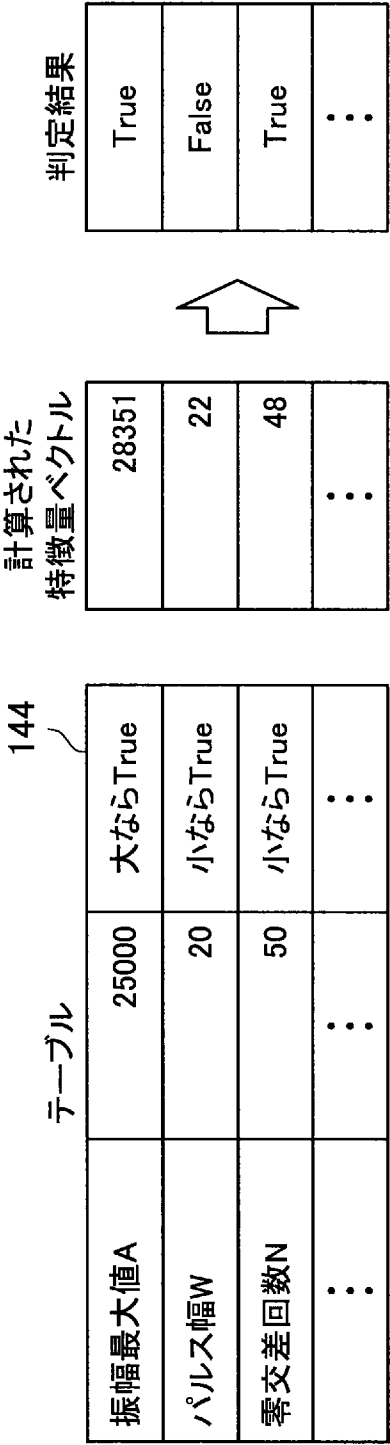
[図5]



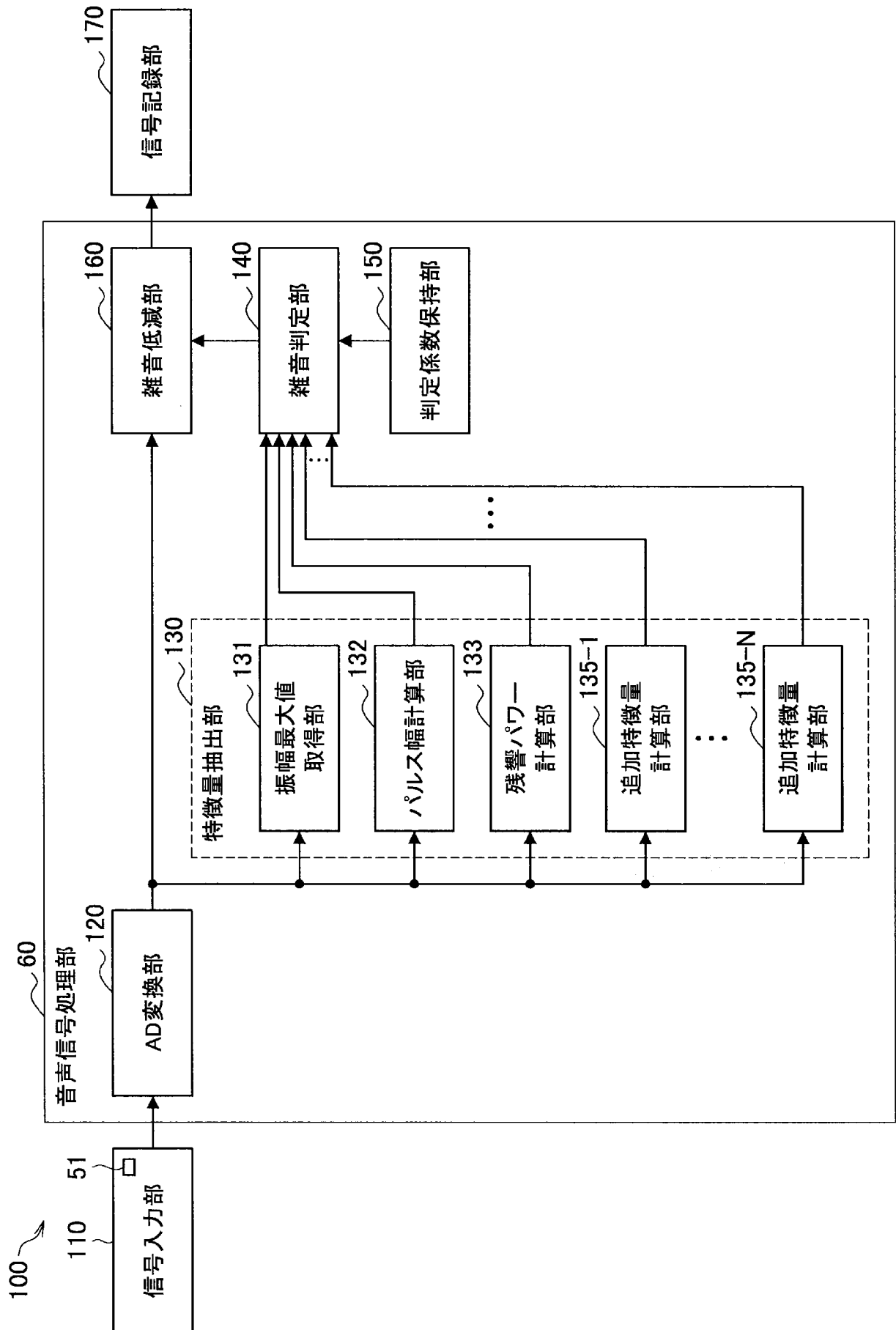
[図6]



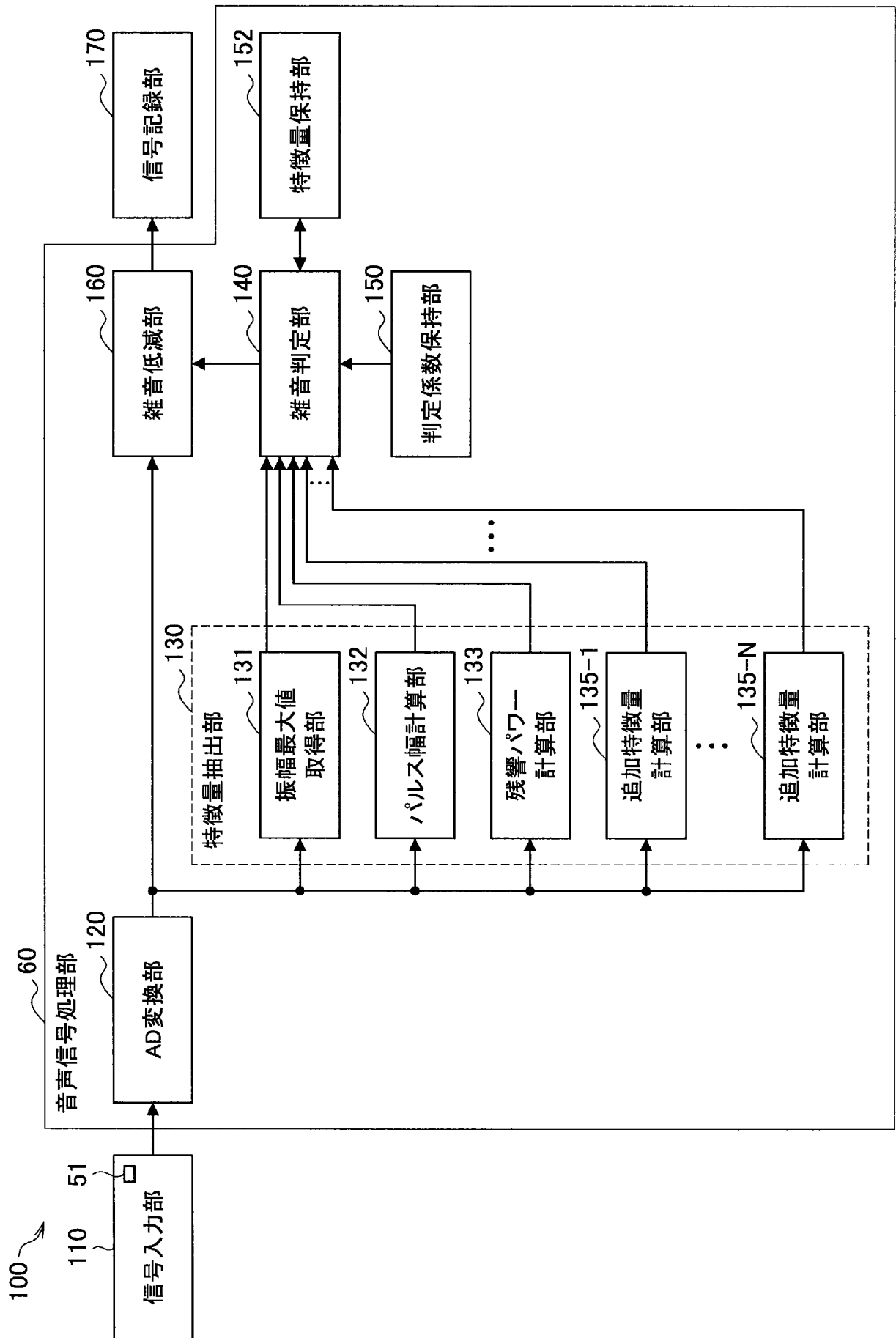
[図7]



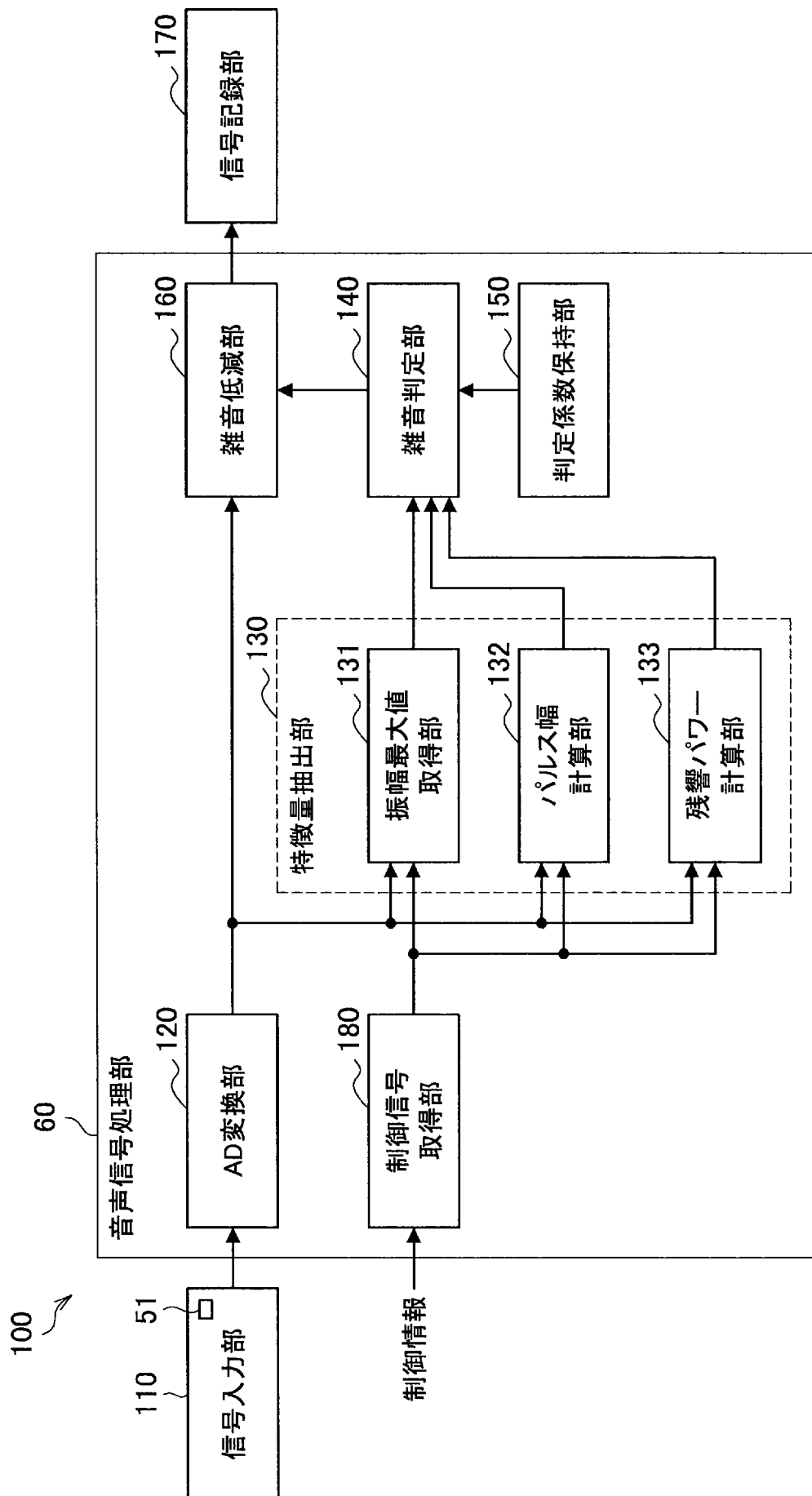
[図8]



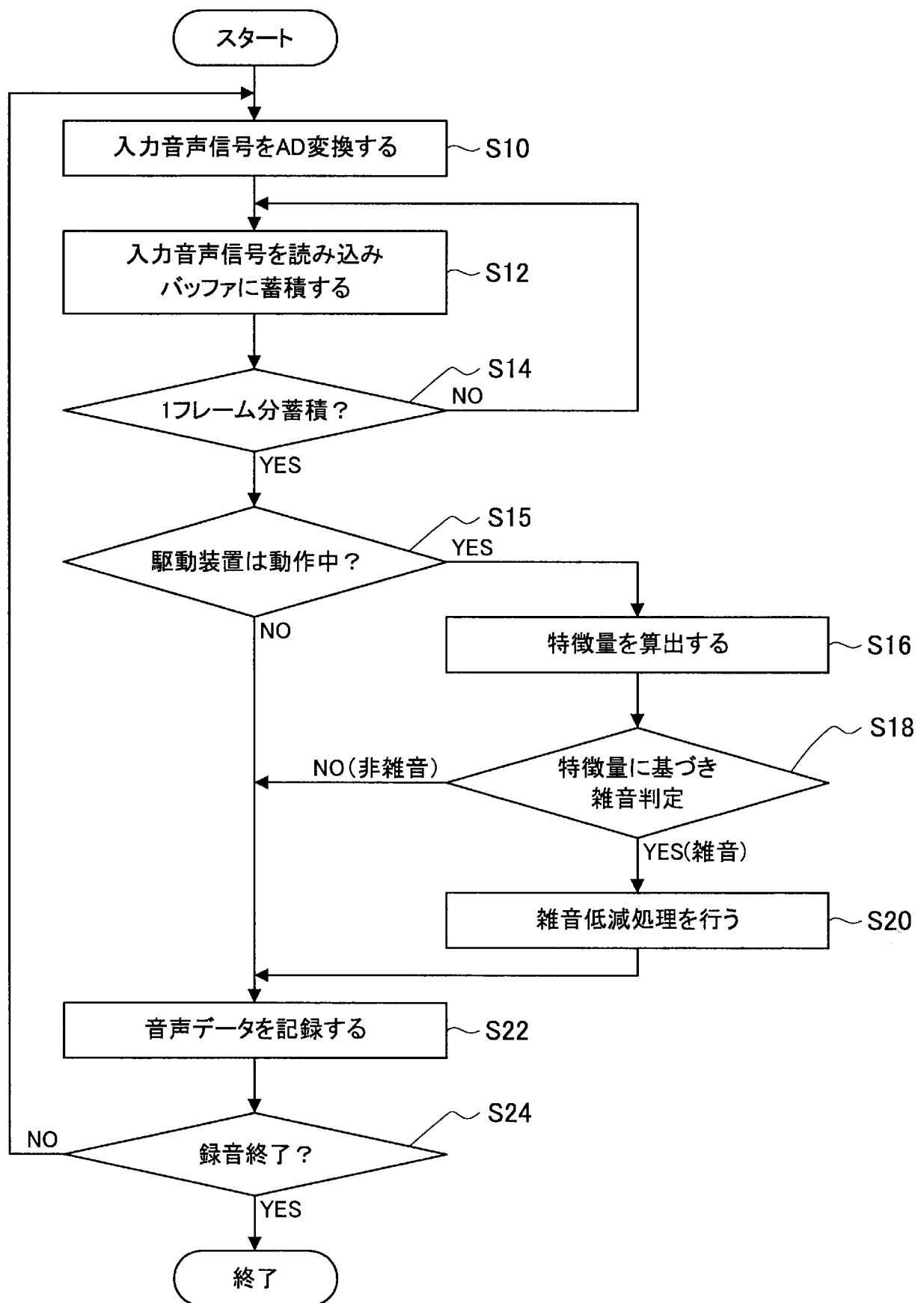
[図9]



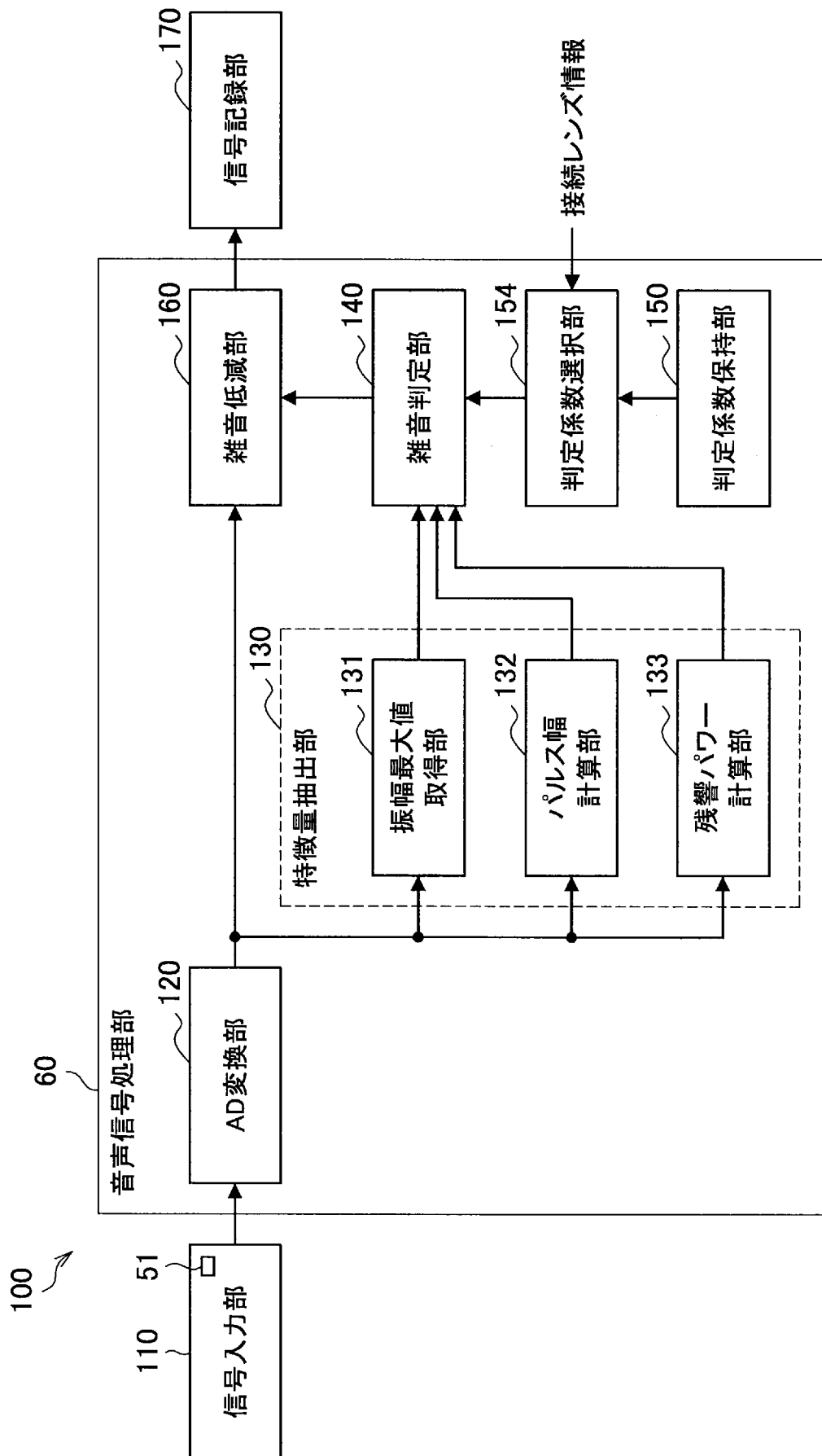
[図10]



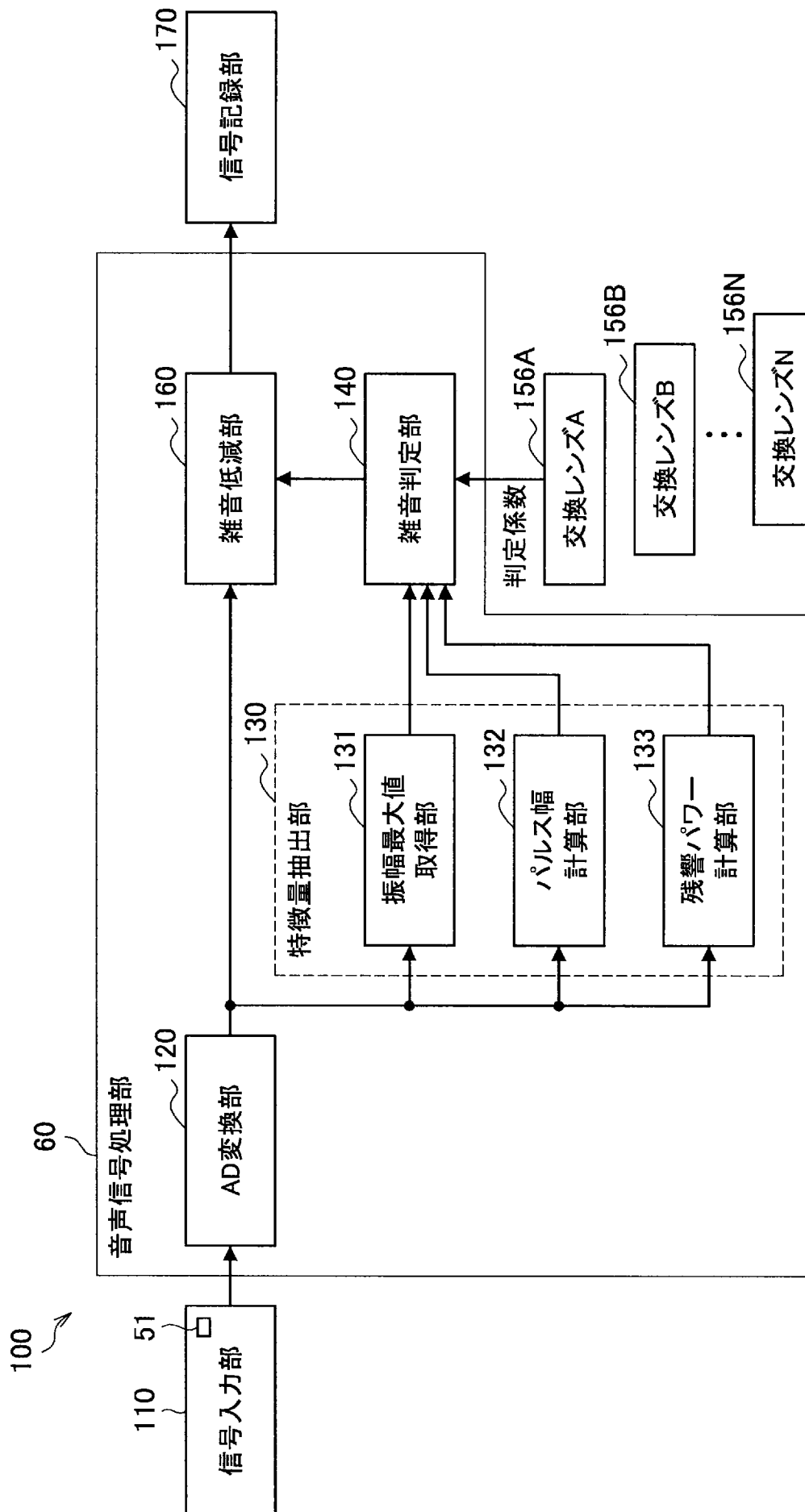
[図11]



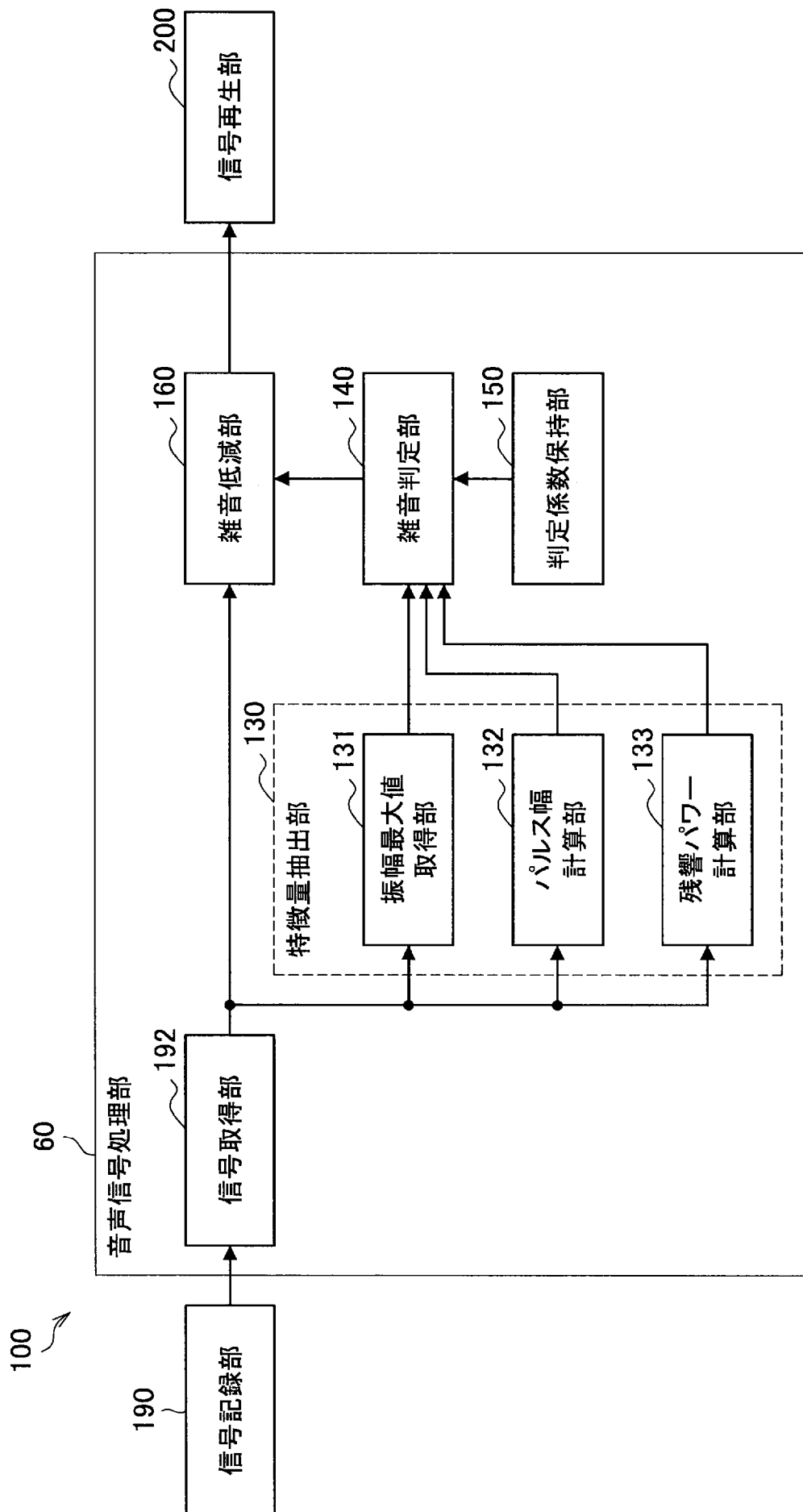
[図12]



[図13]



[図14]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2012/063753

A. CLASSIFICATION OF SUBJECT MATTER

G10L11/00(2006.01)i, G10L21/02(2006.01)i, H04N5/225(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)

G10L11/00, G10L21/02, H04N5/225

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.
X Y A	JP 2011-77604 A (Sony Corp.), 14 April 2011 (14.04.2011), entire text; all drawings (Family: none)	1-2, 7-16 3-5 6
Y A	JP 2011-128391 A (Toshiba Corp.), 30 June 2011 (30.06.2011), entire text; all drawings (Family: none)	3-5 1-2, 6-16
P, A	JP 2012-27186 A (Sony Corp.), 09 February 2012 (09.02.2012), entire text; all drawings (Family: none)	1-16

☒ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

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

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

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search
05 July, 2012 (05.07.12)

Date of mailing of the international search report
17 July, 2012 (17.07.12)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2012/063753

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2011-2723 A (Canon Inc.), 06 January 2011 (06.01.2011), entire text; all drawings (Family: none)	1-16

A. 発明の属する分野の分類 (国際特許分類 (I P C))

Int.Cl. G10L11/00(2006.01)i, G10L21/02(2006.01)i, H04N5/225(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 (I P C))

Int.Cl. G10L11/00, G10L21/02, H04N5/225

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	1 9 2 2 - 1 9 9 6 年
日本国公開実用新案公報	1 9 7 1 - 2 0 1 2 年
日本国実用新案登録公報	1 9 9 6 - 2 0 1 2 年
日本国登録実用新案公報	1 9 9 4 - 2 0 1 2 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
X Y A	JP 2011-77604 A (ソニー株式会社) 2011.04.14, 全文、全図 (ファミリーなし)	1-2, 7-16 3-5 6
Y A	JP 2011-128391 A (株式会社東芝) 2011.06.30, 全文、全図 (ファミリーなし)	3-5 1-2, 6-16
P, A	JP 2012-27186 A (ソニー株式会社) 2012.02.09, 全文、全図 (ファミリーなし)	1-16

☒ 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

0 5 . 0 7 . 2 0 1 2

国際調査報告の発送日

1 7 . 0 7 . 2 0 1 2

国際調査機関の名称及びあて先

日本国特許庁 (I S A / J P)

郵便番号 1 0 0 - 8 9 1 5

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)

井上 健一

5 Z

9 3 7 3

電話番号 0 3 - 3 5 8 1 - 1 1 0 1 内線 3 5 9 1

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	JP 2011-2723 A (キヤノン株式会社) 2011.01.06, 全文、全図 (ファミリーなし)	1-16