

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2017-9663

(P2017-9663A)

(43) 公開日 平成29年1月12日(2017.1.12)

(51) Int.Cl.	F I	テーマコード (参考)
G 1 0 L 19/00 (2013.01)	G 1 0 L 19/00 3 1 2 E	5 C 0 5 3
G 1 0 L 21/043 (2013.01)	G 1 0 L 19/00 3 1 2 F	
G 1 0 L 21/055 (2013.01)	G 1 0 L 21/043 1 0 0	
G 1 0 L 21/057 (2013.01)	G 1 0 L 21/055	
H 0 4 N 5/91 (2006.01)	G 1 0 L 21/057	

審査請求 未請求 請求項の数 20 O L (全 37 頁) 最終頁に続く

(21) 出願番号 特願2015-122214 (P2015-122214)
 (22) 出願日 平成27年6月17日 (2015.6.17)

(71) 出願人 000002185
 ソニー株式会社
 東京都港区港南1丁目7番1号
 (74) 代理人 100112955
 弁理士 丸島 敏一
 (72) 発明者 早川 知伸
 東京都港区港南1丁目7番1号 ソニー株式会社内
 Fターム(参考) 5C053 GB06 GB09 GB11 GB37 HA27
 JA01 JA22 LA02 LA06

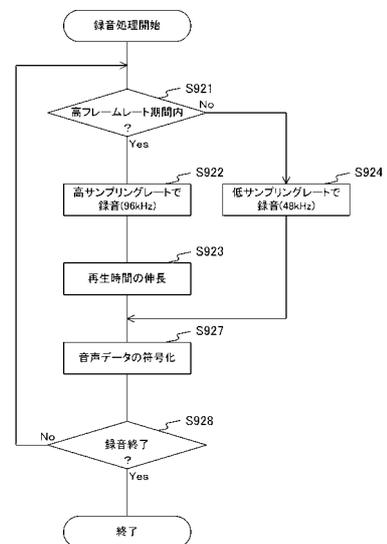
(54) 【発明の名称】 録音装置、録音システム、および、録音方法

(57) 【要約】 (修正有)

【課題】 音声を録音して再生時間を伸長する装置において音質の低下を抑制する。

【解決手段】 サンプリング処理部は、所定のサンプリングレートより高いサンプリングレートで所定の期間内において音声をサンプリングして音声データを高解像度音声データとして生成する処理と、所定のサンプリングレートで所定の期間外において音声をサンプリングして音声データを通常音声データとして生成する処理と、を行う。また、再生時間変換部は、高解像度音声データの再生時間を伸長する。

【選択図】 図 1 0



【特許請求の範囲】

【請求項 1】

所定のサンプリングレートより高いサンプリングレートで所定の期間内において音声をサンプリングして音声データを高解像度音声データとして生成する処理と前記所定のサンプリングレートで前記所定の期間外において音声をサンプリングして音声データを通常音声データとして生成する処理とを行うサンプリング処理部と、

前記高解像度音声データの再生時間を伸長する再生時間変換部とを具備する録音装置。

【請求項 2】

前記サンプリング処理部は、前記所定の期間外において前記所定のサンプリングレートで前記音声をサンプリングし、前記所定の期間内において前記所定のサンプリングレートより高いサンプリングレートにサンプリングレートを切り替えて前記音声をサンプリングする

請求項 1 記載の録音装置。

10

【請求項 3】

前記サンプリング処理部は、

前記所定のサンプリングレートより高いサンプリングレートで前記音声をサンプリングして前記高解像度音声データを生成する高解像度マイクと、

前記所定の期間外において前記高解像度音声データを前記所定のサンプリングレートでサンプリングして前記通常音声データを生成するサンプリングレート変換器と

を備える請求項 1 記載の録音装置。

20

【請求項 4】

前記サンプリング処理部は、

前記所定のサンプリングレートより高いサンプリングレートで前記音声をサンプリングして前記高解像度音声データを生成する高解像度マイクと

前記所定のサンプリングレートで前記音声をサンプリングして前記通常音声データを生成する通常マイクと、

前記所定の期間内において前記高解像度音声データを選択して出力し、前記所定の期間外において前記通常音声データを選択して出力する選択部と

を備える請求項 1 記載の録音装置。

30

【請求項 5】

前記選択部は、前記所定の期間内の一定のフェード期間において前記高解像度音声データに前記通常音声データを合成する合成処理を行う

請求項 4 記載の録音装置。

【請求項 6】

前記選択部は、前記合成処理において前記フェード期間より短い単位時間が経過するたびに前記高解像度音声データの割合を変更する

請求項 5 記載の録音装置。

【請求項 7】

所定のフレームレートより高いフレームレートで複数のフレームを撮像する撮像部と、

前記複数のフレームのうち前記所定の期間外に撮像されたフレームのフレームレートを前記所定のフレームレートに変換するフレームレート変換部と

をさらに具備する請求項 1 記載の録音装置。

40

【請求項 8】

所定のタイミングを含む期間を前記所定の期間として設定する制御部をさらに具備する請求項 7 記載の録音装置。

【請求項 9】

前記複数のフレームの中からシーンが変化したシーン変化タイミングを検出するシーン変化検出部をさらに具備し、

前記制御部は、前記シーン変化タイミングを含む期間を前記所定の期間として設定する

50

請求項 8 記載の録音装置。

【請求項 10】

所定の検知対象を検知するセンサーをさらに具備し、

前記制御部は、前記検知対象が検知されたタイミングを含む期間を前記所定の期間に設定する請求項 8 記載の録音装置。

【請求項 11】

前記再生時間が伸長された高解像度音声データに対して所定の信号処理を実行する信号処理部をさらに具備する

請求項 1 記載の録音装置。

【請求項 12】

前記信号処理部は、前記高解像度音声データを複製する

請求項 11 記載の録音装置。

【請求項 13】

前記信号処理部は、前記高解像度音声データの音量レベルを所定のゲインにより調整する

請求項 11 記載の録音装置。

【請求項 14】

前記信号処理部は、前記高解像度音声データの周波数特性を変更する

請求項 11 記載の録音装置。

【請求項 15】

所定のサンプリングレートより高いサンプリングレートで所定の期間内において音声をサンプリングして音声データを高解像度音声データとして生成する処理と前記所定のサンプリングレートで前記所定の期間外において音声をサンプリングして音声データを通常音声データとして生成する処理とを行い、前記高解像度音声データの再生時間を伸長し、前記再生時間を伸長した前記高解像度音声データに対して実行すべき信号処理を示す設定情報を含むメタデータを生成する録音装置と、

前記設定情報に従って前記信号処理を実行して前記信号処理を実行した前記高解像度音声データと前記通常音声データとを再生する再生装置と

を具備する録音システム。

【請求項 16】

前記メタデータのフォーマットは、MPEG4 - AACであり、

前記録音装置は、前記メタデータのDSE (Data Stream Element) 領域に前記設定情報を記録する

請求項 15 記載の録音システム。

【請求項 17】

前記メタデータのフォーマットは、MPEG4 - systemであり、

前記録音装置は、前記メタデータのudta領域に前記設定情報を記録する

請求項 15 記載の録音システム。

【請求項 18】

前記メタデータのフォーマットは、HMM P (Home and Mobile Multimedia Platform) であり、

前記録音装置は、前記メタデータのuuid領域に前記設定情報を記録する

請求項 15 記載の録音システム。

【請求項 19】

所定のサンプリングレートより高いサンプリングレートで所定の期間内において音声をサンプリングして音声データを高解像度音声データとして生成する処理と前記所定のサンプリングレートで前記所定の期間外において音声をサンプリングして音声データを通常音声データとして生成する処理とを行い、前記高解像度音声データの再生時間を伸長し、前記再生時間を伸長した前記高解像度音声データに対して実行すべき信号処理を示す設定情報を含むメタデータを生成する録音装置と、

10

20

30

40

50

前記設定情報を変更して当該変更した設定情報の示す前記信号処理を実行する編集装置と
を具備する録音システム。

【請求項 20】

所定のサンプリングレートより高いサンプリングレートで所定の期間内において音声をサンプリングして音声データを高解像度音声データとして生成する処理と前記所定のサンプリングレートで前記所定の期間外において音声をサンプリングして音声データを通常音声データとして生成する処理とを行うサンプリング処理手順と、

前記高解像度音声データの再生時間を伸長する再生時間変換手順と
を具備する録音方法。

【発明の詳細な説明】

【技術分野】

【0001】

本技術は、録音装置、録音システム、および、録音方法に関する。詳しくは、音声データの再生時間を変更する録音装置、録音システム、および、録音方法に関する。

【背景技術】

【0002】

従来より、音声を聞き取りやすくする目的で、音声の再生時間を引き伸ばす処理が行われている。例えば、動画をスロー再生させる際に、その動画に同期して録音された音声の再生時間を伸長する撮像装置が提案されている（例えば、特許文献1参照。）。 10

【先行技術文献】

【特許文献】

【0003】

【特許文献1】特開2010-178124号公報

【発明の概要】

【発明が解決しようとする課題】

【0004】

しかしながら、上述の撮像装置では、音声の再生時間を長くするほど音質が低下してしまう。例えば、再生時間を2倍にすると、再生時間を伸長しない場合と比較して音声の周波数は1/2に低下し、音程が約1オクターブ低下する。このように再生時間を伸長した部分と伸長していない部分とのそれぞれの音質の差異が大きくなって違和感が生じ、音声全体の再生品質が低下するという問題がある。 30

【0005】

本技術はこのような状況に鑑みて生み出されたものであり、音声を録音して再生時間を伸長する装置において音質の低下を抑制することを目的とする。

【課題を解決するための手段】

【0006】

本技術は、上述の問題点を解消するためになされたものであり、その第1の側面は、所定のサンプリングレートより高いサンプリングレートで所定の期間内において音声をサンプリングして音声データを高解像度音声データとして生成する処理と上記所定のサンプリングレートで上記所定の期間外において音声をサンプリングして音声データを通常音声データとして生成する処理とを行うサンプリング処理部と、上記高解像度音声データの再生時間を伸長する再生時間変換部とを具備する録音装置、および、録音方法である。これにより、所定のサンプリングレートより高いサンプリングレートで所定の期間内においてサンプリングされた高解像度音声データの再生時間が伸長されるという作用をもたらす。 40

【0007】

また、この第1の側面において、上記サンプリング処理部は、上記所定の期間外において上記所定のサンプリングレートで上記音声をサンプリングし、上記所定の期間内において上記所定のサンプリングレートより高いサンプリングレートにサンプリングレートを切り替えて上記音声をサンプリングしてもよい。これにより、上記所定の期間内においてサ 50

ンプリングレートが切り替えられるという作用をもたらす。

【0008】

また、この第1の側面において、上記サンプリング処理部は、上記所定のサンプリングレートより高いサンプリングレートで上記音声データをサンプリングして上記高解像度音声データを生成する高解像度マイクと、上記所定の期間外において上記高解像度音声データを上記所定のサンプリングレートでリサンプリングして上記通常音声データを生成するサンプリングレート変換器とを備えてもよい。これにより、高解像度音声データのリサンプリングにより通常音声データが生成されるという作用をもたらす。

【0009】

また、この第1の側面において、上記サンプリング処理部は、上記所定のサンプリングレートより高いサンプリングレートで上記音声データをサンプリングして上記高解像度音声データを生成する高解像度マイクと上記所定のサンプリングレートで上記音声データをサンプリングして上記通常音声データを生成する通常マイクと、上記所定の期間内において上記高解像度音声データを選択して出力し、上記所定の期間外において上記通常音声データを選択して出力する選択部とを備えてもよい。これにより、所定の期間内において高解像度音声データが選択され、所定の期間外において通常音声データが選択されるという作用をもたらす。

10

【0010】

また、この第1の側面において、上記選択部は、上記所定の期間内の一定のフェード期間において上記高解像度音声データに上記通常音声データを合成する合成処理を行ってもよい。これにより、一定のフェード期間において高解像度音声データに通常音声データが合成されるという作用をもたらす。

20

【0011】

また、この第1の側面において、上記選択部は、上記合成処理において上記フェード期間より短い単位時間が経過するたびに上記高解像度音声データの割合を変更してもよい。これにより、単位時間が経過するたびに高解像度音声データの割合が変更されるという作用をもたらす。

【0012】

また、この第1の側面において、所定のフレームレートより高いフレームレートで複数のフレームを撮像する撮像部と、上記複数のフレームのうち上記所定の期間外に撮像されたフレームのフレームレートを上記所定のフレームレートに変換するフレームレート変換部とをさらに具備してもよい。これにより、所定のフレームレートより高いフレームレートで所定の期間外に撮像された複数のフレームのフレームレートが所定のフレームレートに変換されるという作用をもたらす。

30

【0013】

また、この第1の側面において、所定のタイミングを含む期間を上記所定の期間として設定する制御部をさらに具備してもよい。これにより、所定のタイミングを含む期間が所定の期間として設定されるという作用をもたらす。

【0014】

また、この第1の側面において、上記複数のフレームの中からシーンが変化したシーン変化タイミングを検出するシーン変化検出部をさらに具備し、上記制御部は、上記シーン変化タイミングを含む期間を上記所定の期間として設定してもよい。これにより、シーン変化タイミングを含む期間が所定の期間として設定されるという作用をもたらす。

40

【0015】

また、この第1の側面において、所定の検知対象を検知するセンサーをさらに具備し、上記制御部は、上記検知対象が検知されたタイミングを含む期間を上記所定の期間に設定してもよい。これにより、検知対象が検知されたタイミングを含む期間が所定の期間として設定されるという作用をもたらす。

【0016】

また、この第1の側面において、上記再生時間が伸長された高解像度音声データに対し

50

て所定の信号処理を実行する信号処理部をさらに具備してもよい。これにより、高解像度音声データに対して所定の信号処理が実行されるという作用をもたらす。

【0017】

また、この第1の側面において、上記信号処理部は、上記高解像度音声データを複製してもよい。これにより、高解像度音声データが複製されるという作用をもたらす。

【0018】

また、この第1の側面において、上記信号処理部は、上記高解像度音声データの音量レベルを所定のゲインにより調整してもよい。これにより、高解像度音声データの音量レベルが所定のゲインにより調整されるという作用をもたらす。

【0019】

また、この第1の側面において、上記信号処理部は、上記高解像度音声データの周波数特性を変更してもよい。これにより、高解像度音声データの周波数特性が変更されるという作用をもたらす。

【0020】

また、本技術の第2の側面は、所定のサンプリングレートより高いサンプリングレートで所定の期間内において音声をサンプリングして音声データを高解像度音声データとして生成する処理と上記所定のサンプリングレートで上記所定の期間外において音声をサンプリングして音声データを通常音声データとして生成する処理とを行い、上記高解像度音声データの再生時間を伸長し、上記再生時間を伸長した上記高解像度音声データに対して実行すべき信号処理を示す設定情報を含むメタデータを生成する録音装置と、上記設定情報に従って上記信号処理を実行して上記信号処理を実行した上記高解像度音声データと上記通常音声データとを再生する再生装置とを具備する録音システムである。これにより、所定の期間内において高解像度音声データが選択され、所定の期間外において通常音声データが選択されるという作用をもたらす。

【0021】

また、この第1の側面において、上記メタデータのフォーマットは、MPEG4 - AACであり、上記録音装置は、上記メタデータのDSE (Data Stream Element) 領域に上記設定情報を記録してもよい。これにより、DSE領域に設定情報が記録されるという作用をもたらす。

【0022】

また、この第2の側面において、上記メタデータのフォーマットは、MPEG4 - systemであり、上記録音装置は、上記メタデータのudta領域に上記設定情報を記録してもよい。これにより、udta領域に上記設定情報が記録されるという作用をもたらす。

【0023】

また、この第2の側面において、上記メタデータのフォーマットは、HMMP (Home and Mobile Multimedia Platform) であり、上記録音装置は、上記メタデータのuuid領域に上記設定情報を記録してもよい。これにより、uuid領域に上記設定情報が記録されるという作用をもたらす。

【0024】

また、本技術の第2の側面は、所定のサンプリングレートより高いサンプリングレートで所定の期間内において音声をサンプリングして音声データを高解像度音声データとして生成する処理と上記所定のサンプリングレートで上記所定の期間外において音声をサンプリングして音声データを通常音声データとして生成する処理とを行い、上記高解像度音声データの再生時間を伸長し、上記再生時間を伸長した上記高解像度音声データに対して実行すべき信号処理を示す設定情報を含むメタデータを生成する録音装置と、上記設定情報を変更して当該変更した設定情報の示す上記信号処理を実行する編集装置とを具備する録音システムである。これにより、所定の期間内において高解像度音声データが選択され、所定の期間外において通常音声データが選択されるという作用をもたらす。

【発明の効果】

10

20

30

40

50

【 0 0 2 5 】

本技術によれば、音声を録音して再生時間を伸長する装置において音質の低下を抑制することができるという優れた効果を奏し得る。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いずれかの効果であってもよい。

【 図面の簡単な説明 】

【 0 0 2 6 】

【 図 1 】 第 1 の実施の形態における撮像装置の一構成例を示すブロック図である。

【 図 2 】 第 1 の実施の形態における動画取込み部の一構成例を示すブロック図である。

【 図 3 】 第 1 の実施の形態における音声取込み部の一構成例を示すブロック図である。

【 図 4 】 第 1 の実施の形態における音声処理部の一構成例を示すブロック図である。

10

【 図 5 】 第 1 の実施の形態におけるストリームの一例を示す図である。

【 図 6 】 第 1 の実施の形態における再生時間の変換前後の高解像度音声データの一例を示す図である。

【 図 7 】 第 1 の実施の形態における音声データの復元帯域の一例を示す図である。

【 図 8 】 第 1 の実施の形態におけるストリームおよびパケットのデータ構造の一例を示す図である。

【 図 9 】 第 1 の実施の形態における録画処理の一例を示すフローチャートである。

【 図 1 0 】 第 1 の実施の形態における録音処理の一例を示すフローチャートである。

【 図 1 1 】 第 1 の実施の形態の第 1 の変形例における撮像装置の一構成例を示すブロック図である。

20

【 図 1 2 】 第 1 の実施の形態の第 1 の変形例における動画取込み部の一構成例を示すブロック図である。

【 図 1 3 】 第 1 の実施の形態の第 2 の変形例における録音装置の一構成例を示すブロック図である。

【 図 1 4 】 第 2 の実施の形態における音声取込み部の一構成例を示すブロック図である。

【 図 1 5 】 第 2 の実施の形態における音声処理部の一構成例を示すブロック図である。

【 図 1 6 】 第 2 の実施の形態におけるストリームの一例を示す図である。

【 図 1 7 】 第 2 の実施の形態における周波数特性の一例を示すグラフである。

【 図 1 8 】 第 2 の実施の形態における録音処理の一例を示すフローチャートである。

【 図 1 9 】 第 3 の実施の形態における撮像システムの一構成例を示すブロック図である。

30

【 図 2 0 】 第 3 の実施の形態における再生装置の一構成例を示すブロック図である。

【 図 2 1 】 第 3 の実施の形態における M P E G 4 - A A C を用いる際の設定先のフィールドの一例を示す図である。

【 図 2 2 】 第 3 の実施の形態における M P E G 4 - s y s t e m を用いる際の設定先のフィールドの一例を示す図である。

【 図 2 3 】 第 3 の実施の形態における H M M P ファイルフォーマットを用いる際の設定先のフィールドの一例を示す図である。

【 図 2 4 】 第 3 の実施の形態における録音処理の一例を示すフローチャートである。

【 図 2 5 】 第 3 の実施の形態における再生処理の一例を示すフローチャートである。

【 図 2 6 】 第 4 の実施の形態における撮像システムの一構成例を示すブロック図である。

40

【 図 2 7 】 第 4 の実施の形態における編集処理の一例を示すフローチャートである。

【 図 2 8 】 第 5 の実施の形態における音声取込み部の一構成例を示すブロック図である。

【 図 2 9 】 第 5 の実施の形態における録音処理の一例を示すフローチャートである。

【 図 3 0 】 第 6 の実施の形態における音声取込み部の一構成例を示すブロック図である。

【 図 3 1 】 第 6 の実施の形態における合成比率の変動の一例を示すグラフである。

【 図 3 2 】 第 6 の実施の形態における録音処理の一例を示すフローチャートである。

【 発明を実施するための形態 】

【 0 0 2 7 】

以下、本技術を実施するための形態（以下、実施の形態と称する）について説明する。説明は以下の順序により行う。

50

1. 第1の実施の形態（高解像度音声データの再生時間を伸長する例）
2. 第2の実施の形態（高解像度音声データの再生時間を伸長して信号処理を行う例）
3. 第3の実施の形態（高解像度音声データの再生時間を伸長してメタデータを生成する例）
4. 第4の実施の形態（高解像度音声データの再生時間を伸長してメタデータを編集する例）
5. 第5の実施の形態（高解像度音声データの再生時間を伸長するとともに高解像度音声データのサンプリングレートを変換する例）
6. 第6の実施の形態（高解像度音声データに通常音声データを合成し、再生時間を伸長する例）

10

【0028】

< 1. 第1の実施の形態 >

[撮像装置の構成例]

図1は、第1の実施の形態における撮像装置100の一構成例を示すブロック図である。この撮像装置100は、録画および録音を行う装置であり、ユーザインターフェース部110、制御部120、動画取込み部130、動画処理部140、記録フォーマット変換部150、音声取込み部160および音声処理部170および記録部180を備える。なお、撮像装置100は、特許請求の範囲に記載の録音装置の一例である。

【0029】

ユーザインターフェース部110は、ユーザの操作に従って操作信号を生成するものである。ユーザインターフェース部110は、生成した操作信号を制御部120に供給する。

20

【0030】

制御部120は、操作信号に従って撮像装置100全体を制御するものである。この制御部120は、操作信号に従って録音および録画の動作を制御する制御信号を生成する。この制御信号は、例えば、録画および録音の開始タイミングと、それらの終了タイミングとを指示する信号を含む。この録画の開始タイミングは、録音の開始タイミングと同一のタイミングであるものとする。同様に、録画および録音の終了タイミングも同一のタイミングであるものとする。また、制御信号は、高フレームレート期間の開始タイミングおよび終了タイミングを指示する信号をさらに含む。

30

【0031】

ここで、高フレームレート期間は、再生時より高いフレームレートで撮像する期間である。例えば、ユーザが所定のボタンを押したタイミングを中心とする一定の長さ（1秒など）の期間が高フレームレート期間として設定される。高フレームレート期間外のフレームレートは、再生時と同一である。高フレームレート期間外（再生時）のフレームレートは例えば、60ヘルツ（Hz）に設定され、高フレームレート期間内のフレームレートは例えば、600ヘルツ（Hz）に設定される。

【0032】

なお、再生時のフレームレートは、60ヘルツ（Hz）に限定されず、30ヘルツ（Hz）などであってもよい。また、高フレームレート期間内のフレームレートは、再生時よりも高い値であれば、600ヘルツ（Hz）に限定されず、120ヘルツ（Hz）などであってもよい。

40

【0033】

制御部120は、上述の各タイミングを示す制御信号を動画取込み部130、音声取込み部160および音声処理部170に信号線129を介して供給する。

【0034】

動画取込み部130は、制御信号に従って、複数のビデオフレームを順に撮像するものである。この動画取込み部130は、撮像したビデオフレームを時系列順に含む動画データを信号線139を介して動画処理部140に供給する。

【0035】

50

動画処理部 140 は、動画データを符号化する処理を行うものである。この動画データは、例えば、MPEG (Moving Picture Experts Group) - 2 規格に従って符号化される。この動画処理部 140 は、符号化した動画データをビデオパッケージにパッケージ化して記録フォーマット変換部 150 に信号線 149 を介して供給する。なお、動画処理部 140 は、MPEG - 4 など、MPEG - 2 以外の規格に従って符号化してもよい。

【0036】

音声取込み部 160 は、制御信号に従って音声をサンプリングして音声データを生成するものである。この音声取込み部 160 は、所定のサンプリングレートで音声をサンプリングし、サンプリングのたびに音声の音量レベルをデジタルの音声データに量子化する。このように、サンプリングおよび量子化によりアナログ信号を A/D (Analog to Digital) 変換する方式は、PCM (Pulse Code Modulation) 方式と呼ばれる。

10

【0037】

ただし、音声取込み部 160 は、高フレームレート期間内において高フレーム期間外よりも高いサンプリングレートで音声をサンプリングする。高フレームレート期間内のサンプリングレートは例えば、96 キロヘルツ (kHz) に設定され、高フレームレート期間外のサンプリングレートは例えば、48 キロヘルツ (kHz) に設定される。以下、高フレームレート期間内にサンプリングされた音声データを「高解像度音声データ」と称し、高フレームレート期間外にサンプリングされた音声データを「通常音声データ」と称する。音声取込み部 160 は、これらの音声データを音声処理部 170 に信号線 169 を介して供給する。

20

【0038】

なお、音声取込み部 160 は、特許請求の範囲に記載のサンプリング処理部の一例である。また、高フレームレート期間外のサンプリングレートは、高フレームレート期間内より低い値であれば、48 キロヘルツ (kHz) に限定されず、44.1 キロヘルツ (kHz) などであってもよい。また、高フレームレート期間内のサンプリングレートは、高フレームレート期間外より高い値であれば、96 キロヘルツ (kHz) に限定されず、192 キロヘルツ (kHz) などであってもよい。

【0039】

音声処理部 170 は、制御信号に従って高解像度音声データの再生時間を一定の倍率 (例えば、2 倍) で伸長するものである。この音声処理部 170 は、再生時間を伸長した高解像度音声データと、通常音声データとを所定の符号化単位で符号化する。これらの音声データは、例えば、MPEG の規格に従って 20 ミリ秒 (ms) の単位で符号化される。符号化単位で符号化されたそれぞれの音声信号は、「オーディオフィーム」と呼ばれる。音声処理部 170 は、オーディオフィームをオーディオパッケージにパッケージ化して記録フォーマット変換部 150 に信号線 179 を介して供給する。

30

【0040】

記録フォーマット変換部 150 は、ビデオパッケージおよびオーディオパッケージの記録フォーマットを所定のフォーマットに変換するものである。また、記録フォーマット変換部 150 は、オーディオフィームおよびビデオフレームのそれぞれに再生時刻を設定する。これらの再生時刻として、例えば、MPEG 規格における PTS (Presentation Time Stamp) が設定される。ビデオフレームの PTS は、高フレームレート期間外におけるビデオフレームの撮像間隔と同一の間隔 (例えば、1/60 秒) で設定される。また、高解像度音声データから生成された最初のオーディオフィームの PTS は、高フレームレート期間を引き伸ばしたスロー再生期間内の所定のタイミング (開始時点や中間時点など) に設定される。そして、記録フォーマット変換部 150 は、フォーマット変換後のパッケージからなるデータをストリームとして記録部 180 に信号線 159 を介して供給する。記録部 180 は、ストリームを記録するものである。

40

【0041】

なお、動画取込み部 130、動画処理部 140、音声取込み部 160 および音声処理部 170 などの回路を 1 つの装置に設けているが、これらを複数の装置に分散して設けても

50

よい。例えば、録画を行うための回路（動画取込み部 130 および動画処理部 140 など）のみを撮像装置 100 に設け、録音を行うための回路（音声取込み部 160 および音声処理部 170 など）を録音装置に設けてもよい。

【0042】

[動画取込み部の構成例]

図 2 は、第 1 の実施の形態における動画取込み部 130 の一構成例を示すブロック図である。この動画取込み部 130 は、撮像部 131 およびフレームレート変換部 134 を備える。

【0043】

撮像部 131 は、操作信号に従って所定の垂直同期信号 SYNC_VH に同期して複数のビデオフレームを時系列順に撮像するものである。この撮像部 131 は、例えば、撮像レンズなどの光学系と撮像素子とを備える。この撮像素子として、例えば、CMOS (Complementary Metal Oxide Semiconductor) センサーや CCD (Charge Coupled Device) センサーが用いられる。また、垂直同期信号 SYNC_VH の周波数は、再生時のフレームレートより高い値であり、例えば、600 ヘルツ (Hz) である。撮像部 131 は、制御信号の示す録画の開始タイミングから終了タイミングまでの期間に亘って撮像を行い、ビデオフレームのそれぞれをフレームレート変換部 134 に供給する。

10

【0044】

フレームレート変換部 134 は、制御信号に従って、フレームレートを変換するものである。このフレームレート変換部 134 は、制御信号の示す高フレームレート期間において撮像されたビデオフレームのフレームレートを、垂直同期信号 SYNC_VL の周波数（例えば、60 ヘルツ : Hz）のフレームレートに変換する。例えば、ビデオフレームを一定数ごとに間引く処理により、フレームレートが変換される。一方、高フレームレート期間外において撮像されたビデオフレームのフレームレートは変換されない。フレームレート変換部 134 は、これらのビデオフレームを含む動画データを動画処理部 140 に供給する。

20

【0045】

[音声取込み部の構成例]

図 3 は、第 1 の実施の形態における音声取込み部 160 の一構成例を示すブロック図である。この音声取込み部 160 は、サンプリングレート可変マイク 161 を備える。

30

【0046】

サンプリングレート可変マイク 161 は、制御信号に従ってサンプリングレートを変更して音声をサンプリングするものである。このサンプリングレート可変マイク 161 は、制御信号の示す高フレームレート期間外において、一定のサンプリングレート（例えば、48 キロヘルツ）で音声をサンプリングする。一方、高フレームレート期間内においてサンプリングレート可変マイク 161 は、サンプリングレートを高フレームレート期間外より高い値（例えば、96 キロヘルツ）に切り替えて音声をサンプリングする。サンプリングレート可変マイク 161 は、高フレームレート期間外でサンプリングした通常音声データと、高フレームレート期間内でサンプリングした高解像度音声データとを音声処理部 170 に供給する。

40

【0047】

なお、サンプリングレート可変マイク 161 として、デジタルの音声データを出力するデジタルマイクをもうけているが、このデジタルマイクの代わりにアナログの音声信号を出力するアナログマイクを設けてもよい。この場合には、アナログマイクからの音声信号を AD 変換する AD 変換器がさらに設けられ、この AD 変換器がサンプリングレートを切り替えてサンプリングを行う。

【0048】

また、音声取込み部 160 は、サンプリング周波数（サンプリングレート）を切り替える際に段階的に徐々に切り替えてもよい。例えば、高フレームレート期間の開始時点から一定時間に亘って、音声取込み部 160 は、サンプリングレートを少しずつ増加させる。

50

また、高フレームレート期間の終了時点より一定時間前の時点から、その終了時点までの期間に亘って、音声取込み部 160 は、サンプリングレートをすこしずつ減少させる。これにより、サンプリングレートが切り替わる部分の違和感を軽減することができる。

【0049】

[音声処理部の構成例]

図 4 は、第 1 の実施の形態における音声処理部 170 の一構成例を示すブロック図である。この音声処理部 170 は、バッファ 171、再生時間変換部 172 および音声符号化部 177 を備える。

【0050】

バッファ 171 は、一定のデータ量の音声データを保持するものである。再生時間変換部 172 は、高解像度音声データの再生時間を変換するものである。この再生時間変換部 172 は、制御信号の示す高フレームレート期間内にサンプリングされた音声データ（すなわち、高解像度音声データ）をバッファ 171 から読み出し、その再生時間を一定の倍率で伸長して音声符号化部 177 に供給する。一方、高フレームレート期間外にサンプリングされた音声データは、再生時間が変更されずにそのまま音声符号化部 177 に供給される。

10

【0051】

音声符号化部 177 は、音声データをオーディオフレームに符号化するものである。この音声符号化部 177 は、オーディオフレームをオーディオパケットにパケット化して記録フォーマット変換部 150 に信号線 179 を介して供給する。

20

【0052】

図 5 は、第 1 の実施の形態におけるストリームの一例を示す図である。同図における a は、垂直同期信号 SYNC_VH に同期して撮像されたビデオフレームの一例を示す図である。垂直同期信号 SYNC_VH の周波数が 600 ヘルツ (Hz) である場合、1/600 秒ごとに複数のビデオフレームが撮像される。

【0053】

図 5 における b は、フレームレート変換後のフレームの一例を示す図である。ユーザの操作に従って高フレームレート期間が設定され、その高フレームレート期間外において、フレームレートが例えば、60 ヘルツ (Hz) の低フレームレートに変換される。同図における c の太線で囲ったビデオフレームは、高フレームレート期間内のビデオフレームを示す。

30

【0054】

図 5 における c は、サンプリングされた音声データの一例を示す図である。例えば、サンプリングにより音声データ Sa1、Sa2 および Sa3 が順に生成される。ここで、音声データ Sa1 は、高フレームレート期間の開始タイミング Ts より前に、比較的低いサンプリングレート（例えば、48 キロヘルツ）でサンプリングされた通常音声データである。また、音声データ Sa2 は、高フレームレート期間に亘って比較的高いサンプリングレート（例えば、96 キロヘルツ）でサンプリングされた高解像度音声データである。また、音声データ Sa3 は、高フレームレート期間の終了タイミング Te の後に、比較的低いサンプリングレートでサンプリングされた通常音声データである。

40

【0055】

図 5 における d は、再生時刻が設定されたビデオフレームの一例を示すものである。検出フレームを含むビデオフレームのそれぞれに、例えば、60 ヘルツ (Hz) の低フレームレートで再生するための再生時刻が設定される。この再生時刻に従って、高フレームレート期間に撮像された動体は、非常にゆっくりとした動きで再生される。例えば、高フレームレート期間のフレームレートが 600 ヘルツ (Hz) で、再生時のフレームレートが 60 ヘルツ (Hz) である場合、スロー再生期間は、高フレームレート期間の 10 倍に引き伸ばされ、動体の動作速度は 1/10 に低下する。

【0056】

図 5 における e は、再生時間変換後の音声データの一例を示す図である。音声処理部 1

50

70は、高解像度音声データ(S a 2)の再生時間を伸長して音声データS a 2'を生成する。この音声データS a 2'の再生時刻は、例えば、スロー再生期間の中間時点T c'に設定される。なお、変換後の音声データS a 2'の再生時刻は、スロー再生期間の中間時点T c'に限定されず、例えば、スロー再生期間の開始タイミングT sであってもよい。また、スロー再生期間の開始タイミングT sの直前に、連続した無音の期間がある場合には、撮像装置100は、その期間の開始タイミングを音声データS a 2'の再生時刻としてもよい。

【0057】

図6は、第1の実施の形態における再生時間の変換前後の高解像度音声データの一例を示す図である。同図におけるaは、再生時間を変換する前の高解像度音声データの一例を示す図であり、同図におけるbは、再生時間を変換した後の高解像度音声データの一例を示す図である。また、同図において縦軸は、音量レベルを示し、横軸は時間を示す。また、点線は、音声データをD A (Digital to Analog)変換した際に復元されるアナログの音声信号の波形を示す。

10

【0058】

図6におけるaに例示するように、高フレームレート期間において、比較的高いサンプリングレートで音声データ502および504などの音声データがサンプリングされる。例えば、サンプリングレートが96キロヘルツ(kHz)の際には、1秒当たり96×1000個の音声データが生成される。量子化ビット数を24ビットとすると、モノラルの音声データの1秒当たりのデータ量は、96×1000×24ビットである。なお、量子化ビット数は24ビットに限定されず、16ビットなどであってもよい。

20

【0059】

そして、図6におけるbに例示するように、変換前の音声データ(502や504)の間に、音声データ503などのデータが補間される。同図において斜線を引いた部分は、補間された音声データを示す。音声処理部170は、例えば、隣り合う音声データ502および504のそれぞれの音量レベルの中間の音量レベルを示すデータを音声データ503として補間する。この補間により、音声データの再生時間が引き伸ばされる。例えば、補間により音声データの個数を2倍にした場合には、再生時間は2倍に伸長される。このように、再生時間を伸長する処理は、タイムストレッチや話速変換と呼ばれる。

30

【0060】

ここで、サンプリングされた音声データにより復元することができるアナログの音声の最大周波数は、標本化定理より、サンプリング周波数 f_s の半分の値である。このサンプリング周波数の半分の値($f_s/2$)は、ナイキスト周波数と呼ばれる。このナイキスト周波数(復元可能な周波数)は、再生時間の伸長により低下する。例えば、再生時間を2倍にした場合には、ナイキスト周波数は半分に低下してしまう。

【0061】

なお、音声処理部170は、波形をそのまま引き伸ばす処理によりタイムストレッチを行っているが、再生時間を伸長することができるのであれば、この方式に限定されない。例えば、音声処理部170は、音声波形を複数に分割し、それらの一部を複製して挿入する処理により、再生時間を引き伸ばしてもよい。この処理によれば、周波数をほとんど変えずに再生時間を伸長することができる。この場合であっても、再生時間を変換した部分で若干の音質低下が生じるため、高解像度音声データを録音することにより、その音質低下を抑制することができる。

40

【0062】

図7は、第1の実施の形態における音声データの復元帯域の一例を示す図である。ここで、高解像度音声データのサンプリングレートは96キロヘルツ(kHz)とし、通常音声データのサンプリングレートは48キロヘルツ(kHz)とする。同図に例示するように、高解像度音声データのD A変換により復元される音声の周波数帯域(以下、「復元帯域」と称する。)は、再生時間を変換する前においては、標本化定理より0乃至48キロヘルツ(kHz)である。この高解像度音声データの再生時間を2倍に伸長すると、復元

50

帯域は伸長前の半分、すなわち 0 乃至 24 キロヘルツ (kHz) となる。一方、通常音声データの復元帯域は、標本化定理より 0 乃至 24 キロヘルツ (kHz) となる。

【0063】

ここで、一般的な人間の可聴範囲は、20 ヘルツ (Hz) 乃至 20 キロヘルツ (kHz) であり、再生時間変更後の復元帯域より狭い。したがって、再生時間を変更しても、ユーザが音質の低下を感じる事がなくなる。また、再生時間変更後の高解像度音声データの復元帯域は、通常音声データと同一であるから、再生時間を引き伸ばしたスロー再生期間の音質は、引き伸ばしていない期間と変わらなくなる。

【0064】

これに対して、特許文献 1 に記載の装置は、高フレームレート期間においても、サンプリングレートを変えずに録音している。この構成では、高フレームレート期間にサンプリングした通音声データの再生時間を引き伸ばすと、引き伸ばしていない期間と比較して復元帯域が狭くなり、音質が低下してしまう。

10

【0065】

なお、上述の標本化定理に基づいて、通常音声データのサンプリングレートは、可聴範囲の最大周波数 (約 20 キロヘルツ) の 2 倍より高い値が望ましい。また、高解像度音声データのサンプリングレートは、可聴範囲の最大周波数の 2 倍に、再生時間に対する倍率 (2 倍など) を乗じた値より高い値が望ましい。

【0066】

図 8 は、第 1 の実施の形態におけるストリームおよびパケットのデータ構造の一例を示す図である。同図における a は、ストリームのデータ構造の一例を示す図である。MPEG-2 TS 規格においてストリームは、例えば、補助的なデータを含むパケット ARI_PACKET と、ビデオパケット V_PACKET と、オーディオパケット A_PACKET とを含む。ビデオフレームは、1 つ以上のビデオパケット V_PACKET に格納され、オーディオフレームは、1 つ以上のオーディオパケット A_PACKET に格納される。

20

【0067】

図 8 における b は、ビデオパケット V_PACKET のデータ構造の一例を示す図である。MPEG-2 TS 規格においてビデオパケット V_PACKET には、パケット開始コードと、パケット長と、「10」のコードと、フラグおよび制御と、PES ヘッダ長と、コンディショナル・コーディングと、パケット・データとが格納される。なお、オーディオパケットのデータ構造は、ビデオパケットと同様である。

30

【0068】

パケット開始コードのフィールドには、パケットの先頭を示す先頭開始コードと、ストリームを識別するためのストリーム ID とが格納される。コンディショナル・コーディングのフィールドには、再生時刻を示す PTS や、復号時刻を示す DTS (Decoding Time Stamp) とが格納される。

【0069】

[撮像装置の動作例]

図 9 は、第 1 の実施の形態における録画処理の一例を示すフローチャートである。この録画処理は、例えば、録画を開始するための操作 (録画ボタンの押下など) が行われたときに開始する。撮像装置 100 は、ビデオフレームを 600 ヘルツ (Hz) の高フレームレートで生成する (ステップ S901)。また、撮像装置 100 は、シーン変化を検出したときに、高フレームレート期間を設定し (ステップ S902)、高フレームレート期間外においてフレームレートを 60 ヘルツ (Hz) の低フレームレートに変換する (ステップ S903)。

40

【0070】

そして、撮像装置 100 は、ビデオフレームを符号化する (ステップ S904)。撮像装置 100 は、録画を終了するための操作 (停止ボタンの押下など) が行われたか否かを判断する (ステップ S905)。録画を終了するための操作が行われていない場合には (ステップ S905: No)、撮像装置 100 は、ステップ S901 以降を繰り返す。一方

50

、録画を終了するための操作が行われた場合には（ステップS905：Yes）、撮像装置100は、録画処理を終了する。

【0071】

図10は、第1の実施の形態における録音処理の一例を示すフローチャートである。この録音処理は、例えば、録画を開始するための操作（録画ボタンの押下など）が行われたときに開始する。

【0072】

撮像装置100は、現在時刻が高フレームレート期間内であるか否かを判断する（ステップS921）。高フレームレート期間内である場合に（ステップS921：Yes）撮像装置100は、96キロヘルツ（kHz）の高サンプリングレートで録音を行い（ステップS922）、生成した高解像度音声データの再生時間を伸長する（ステップS923）。一方、高フレームレート期間内でない場合に（ステップS921：No）撮像装置100は、48キロヘルツ（kHz）の低サンプリングレートで録音を行う（ステップS924）。

10

【0073】

ステップS923またはS924の後、撮像装置100は、音声データを符号化してオーディオフレームを生成し（ステップS927）、録音を終了するための操作（停止ボタンの押下など）が行われたか否かを判断する（ステップS928）。録音を終了するための操作が行われていない場合には（ステップS928：No）、撮像装置100は、ステップS921以降を繰り返す。一方、録音を終了するための操作が行われた場合には（ステップS928：Yes）、撮像装置100は、録音処理を終了する。

20

【0074】

このように、本技術の第1の実施の形態によれば、高フレームレート期間内に比較的高いサンプリングレートでサンプリングして高解像度音声データを生成し、そのデータの再生時間を伸長するため、再生時間の伸長による音質の低下を抑制することができる。

【0075】

[第1の変形例]

第1の実施の形態では、高フレームレート期間を手動で設定していたが、手動とすると、操作ミスにより、その期間の開始タイミングがずれるおそれがある。また、高フレームレート期間の設定を手動にすると、撮像装置100の操作が複雑になり、撮像装置100の利便性が低下するおそれがある。この第1の実施の形態の第1の変形例の撮像装置100は、ユーザの操作によらずに、高フレームレート期間を設定する点において第1の実施の形態と異なる。

30

【0076】

図11は、第1の実施の形態の第1の変形例における撮像装置100の一構成例を示すブロック図である。この第1の変形例の動画取込み部130は、シーンが変化するタイミングをシーン変化タイミングとして検出する点において第1の実施の形態と異なる。動画取込み部130は、検出したシーン変化タイミングを制御部120に供給する。

【0077】

第1の変形例の制御部120は、検出されたシーン変化タイミングを含む一定の期間を高フレームレート期間に設定する点において第1の実施の形態と異なる。例えば、シーン変化タイミングを中心とする一定の長さ（1秒など）の期間が高フレームレート期間に設定される。

40

【0078】

なお、制御部120は、ユーザの操作に従って録音の開始タイミングおよび終了タイミングのみを設定する構成としているが、ユーザの操作に従って、さらに高フレームレート期間を設定してもよい。例えば、制御部120は、シーン変化タイミングとユーザの指定したタイミングとのいずれかを含む期間を高フレームレート期間として設定すればよい。

【0079】

図12は、第1の実施の形態の第1の変形例における動画取込み部130の一構成例を

50

示すブロック図である。この第1の変形例の動画取込み部130は、バッファ132およびシーン変化検出部133をさらに備える点において第1の実施の形態と異なる。

【0080】

バッファ132は、撮像部131により撮像されたビデオフレームを保持するものである。

【0081】

シーン変化検出部133は、シーンが変化したときのビデオフレームを検出するものである。このシーン変化検出部133は、撮像部131からビデオフレームを現ビデオフレームとして取得し、その現ビデオフレームより前のビデオフレームをバッファ132から前ビデオフレームとして取得する。そして、シーン変化検出部133は、現ビデオフレームと前ビデオフレームとを比較し、その比較結果に基づいてシーンの変化の有無を検出する。シーン変化検出部133は、シーンの変化が生じると、そのときのフレームの撮像時刻をシーン変化タイミングとして制御部120に供給する。

10

【0082】

このように、本技術の第1の実施の形態の第1の変形例によれば、撮像装置100がシーン変化タイミングを検出して、そのタイミングを含む一定期間を高フレームレート期間に設定するため、高フレームレート期間を手動で設定する必要がなくなる。

【0083】

[第2の変形例]

第1の実施の形態では、高解像度音声データを録音する期間を手動で設定していたが、手動にすると、操作ミスが生じるおそれや、撮像装置100の利便性が低下するおそれがある。また、第1の実施の形態では、録音および録画の両方を行っていたが、録画時間が長くなるほど動画データのデータ量が多くなり、記録部180の記憶容量が不足するおそれがある。この第1の実施の形態の第2の変形例の装置は、ユーザの操作によらずに高フレームレート期間を設定する点と、録画を行わない点とにおいて第1の実施の形態と異なる。

20

【0084】

図13は、第1の実施の形態の第2の変形例における録音装置101の一構成例を示すブロック図である。この録音装置101は、ユーザインターフェース部110、センサー115、制御部125、音声取込み部160、音声処理部170、記録フォーマット変換部150および記録部180を備える。

30

【0085】

センサー115は、不審者などの検知対象を検知するものである。例えば、圧電体に加えられた力を電気信号に変換する圧電センサーや、赤外線の光量を電気信号に変換する赤外線センサーがセンサー115として用いられる。このセンサー115は、検知対象を検知したか否かを示す検知信号を制御部125に供給する。

【0086】

制御部125は、操作信号に従って音声取込み部160に録音を開始させ、操作信号に従って録音を終了させる。また、録音中において、センサーにより検知対象が検知されると、制御部125は、検知対象の検知タイミングを含む一定の期間を高解像度録音期間として設定し、その期間において音声取込み部160に比較的高いサンプリングレートでサンプリングを行わせる。

40

【0087】

第2の変形例の音声取込み部160は、高フレームレート期間の代わりに、高解像度期間において高解像度音声データを生成する点以外は、第1の実施の形態と同様である。また、第2の変形例の音声処理部170は、高フレームレート期間の代わりに、高解像度期間において高解像度音声データの再生時間を変換する点以外は、第1の実施の形態と同様である。

【0088】

なお、制御部120は、ユーザの操作に従って録音の開始タイミングおよび終了タイミ

50

ングのみを設定する構成としているが、ユーザの操作に従って、さらに高解像度期間を設定してもよい。例えば、制御部120は、センサー115の検知タイミングとユーザの指定したタイミングとのいずれかを含む期間を高解像度期間として設定すればよい。

【0089】

また、動画取込み部130をさらに設け、第1の変形例と同様に、シーン変化タイミングを検出してもよい。この場合に制御部120は、例えば、センサー115の検知タイミングとシーン変化タイミングとのいずれかを含む期間を高解像度期間として設定すればよい。さらに、制御部120は、センサー115の検知タイミングとシーン変化タイミングとユーザの指定したタイミングとのいずれかを含む期間を高解像度期間として設定してもよい。

10

【0090】

このように、本技術の第1の実施の形態の第2の変形例によれば、検知対象の検知タイミングを含む一定期間を録音装置101が高解像度録音期間として設定するため、高解像度音声データを録音する期間を手動で設定する必要がなくなる。また、録音装置101は、録画を行わずに録音のみを行うため、記録部180に記録するデータのデータ量を少なくすることができる。

【0091】

< 2 . 第2の実施の形態 >

第1の実施の形態では、高解像度音声データの再生時間を伸長していたが、スロー再生期間に対して伸長後の再生時間が短いと、無音の時間が生じて再生品質が低下するおそれがある。第2の実施の形態の撮像装置100は、無音の時間を短くした点において第1の実施の形態と異なる。

20

【0092】

図14は、第2の実施の形態における音声取込み部160の一構成例を示すブロック図である。第2の実施の形態の音声取込み部160は、付加情報生成部162をさらに備える点において第1の実施の形態と異なる。

【0093】

付加情報生成部162は、高解像度音声データに対して実行すべき信号処理を示す付加情報を生成するものである。例えば、高解像度音声データの複製処理と、音量レベルの調整処理と、周波数特性を変更するイコライザ処理とを含む信号処理が実行される。付加情報生成部162は、それらの信号処理の設定内容を含む付加情報を生成する。この設定内容は、高解像度音声データの再生時刻や、その高解像度音声データの複製処理における複製回数などを含む。この複製回数は、例えば、次の式により設定される。

30

$$\text{複製回数} = \text{SYNC_VH} / (\text{SYNC_VL} \times n) \quad \dots \text{式3}$$

上式において、nは、高解像度音声データの再生時間を伸長する倍率である。なお、複製回数が整数にならない場合には、端数切捨てなどの端数処理が行われる。

【0094】

例えば、SYNC_VHが600ヘルツ(Hz)、SYNC_LHが60ヘルツ(Hz)で、再生時間に対する倍率nが2倍である場合、「5」が複製回数として設定される。付加情報生成部162は、生成した付加情報を音声データに付加して音声処理部170に供給する。

40

【0095】

図15は、第2の実施の形態における音声処理部170の一構成例を示すブロック図である。この第2の実施の形態の音声処理部170は、複製部173およびエフェクト処理部174をさらに備える点において第1の実施の形態と異なる。エフェクト処理部174は、ゲイン調整部175およびイコライザ処理部176を備える。なお、複製部173およびエフェクト処理部174を含む回路は、特許請求の範囲に記載の信号処理部の一例である。

【0096】

複製部173は、高解像度音声データを複製するものである。この複製部173は、付

50

加情報の示す複製回数に亘って複製を行い、生成した複製の音声データのそれぞれを複製音声データとしてゲイン調整部 175 に供給する。

【0097】

ゲイン調整部 175 は、複製音声データの音量レベルをゲインにより調整するものである。このゲイン調整部 175 は、付加情報に従って、例えば、複製音声データごとに、異なるゲインにより音量レベルを調整する。付加情報には、例えば、複製音声データごとのゲインの変化量などが設定される。ゲイン調整部 175 は、音量レベルを調整した複製音声データをイコライザ処理部 176 に供給する。

【0098】

イコライザ処理部 176 は、複製音声データの周波数特性を互いに異なる特性に変更するイコライザ処理を行うものである。例えば、イコライザ処理部 176 は、所定の閾値より低い低周波数領域に対するゲインを、その閾値より高い高周波数領域に対するゲインよりも相対的に高くする処理を複製音声データ毎に行い、再生時刻が遅いほど、その閾値を低くする。このような周波数特性の変更により、時間の経過に伴って、低周波数領域の音声の強調の度合いが徐々に大きくなる音響効果を得ることができる。ここで、付加情報にはイコライザ値が設定される。イコライザ値は、ゲインを制御する帯域や、ゲインの制御量などを含む。イコライザ処理部 176 は、イコライザ処理後の音声信号を音声符号化部 177 に供給する。

10

【0099】

なお、周波数特性の変更の方法は、低周波数領域の強調に限定されない。イコライザ処理部 176 は、高周波数領域を徐々に強調してもよいし、所定の中心周波数を中心とする一定の帯域に対するゲインを変更してもよい。

20

【0100】

また、音声処理部 170 は、複製処理と音量レベルの調整処理とイコライザ処理とを全て実行しているが、この構成に限定されず、これらの処理の一部（複製処理のみ等）を実行する構成であってもよい。また、音声処理部 170 は、複製処理、音量レベルの調整処理およびイコライザ処理などに加えて、これらと異なる信号処理（ノイズ除去処理など）をさらに実行してもよい。

【0101】

また、音声処理部 170 は、再生時間の変更、複製、音量レベルの調整、イコライザ処理の順に各処理を実行しているが、この順と異なる順序で、それぞれの処理を実行してもよい。例えば、音声処理部 170 は、複製した後に再生時間の変更を行ってもよいし、音量レベルを調整した後に複製してもよい。

30

【0102】

図 16 は、第 2 の実施の形態におけるストリームの一例を示す図である。同図における a は、垂直同期信号 SYNC_VH に同期して撮像されたビデオフレームの一例を示す図である。同図における b は、フレームレート変換後のフレームの一例を示す図である。同図における c は、サンプリングされた音声データの一例を示す図である。同図における d は、再生時刻が設定されたビデオフレームの一例を示すものである。

【0103】

図 16 における e は、複製音声データの一例を示す図である。音声処理部 170 は、音声データ Sa2（高解像度音声データ）の再生時間を伸長し、伸長後の Sa2 を複製して複製音声データ Sa2' - 1 乃至 Sa2' - m（m は 2 以上の整数）の m 個の音声データを生成する。最初の複製音声データ Sa2' - 1 の再生時刻は、例えば、スロー再生期間の開始タイミングに設定される。これらの複製音声データは、複製元の音声データと同一であるため、スロー再生期間において同じ音声が続く再生される。これにより、繰り返して再生しない場合と比較して、スロー再生期間内の無音の時間を短くすることができる。

40

【0104】

図 16 における f は、複製音声データのそれぞれの音量調整に用いられるゲインの一例

50

を示す図である。同図における f の縦軸は、ゲインを示し、横軸は時間を示す。最初に再生される複製音声データ $S a 2' - 1$ に対しては、「0」デシベル (dB) のゲインが設定される。以降の複製音声データ $S a 2' - 2$ 乃至 $S a 2' - m$ に対しては、再生時刻が遅いほど小さなゲインが設定される。これにより、繰り返し再生される音声の音量レベルが徐々に小さくなる。

【0105】

図17は、第2の実施の形態における周波数特性の一例を示すグラフである。同図における縦軸はゲインであり、横軸は周波数である。また、点線の曲線は、スロー再生期間において最初に再生される複製音声データ $S a 2' - 1$ の特性を示し、一点鎖線の曲線は、その次に再生される複製音声データ $S a 2' - 2$ の特性を示す。実線の曲線は、スロー再生期間において最後に再生される複製音声データ $S a 2' - m$ の特性を示す。複製音声データ $S a 2' - 1$ では、閾値 $T h 1$ に対して高周波数領域のゲインが相対的に低くなるように調整され、複製音声データ $S a 2' - 2$ では、閾値 $T h 1$ より低い閾値 $T h 2$ に対して高周波数領域のゲインが相対的に低くなるように調整される。また、複製音声データ $S a 2' - m$ では、それらよりも低い閾値 $T h m$ に対して高周波数領域のゲインが相対的に低くなるように調整される。このような周波数特性の変更により、時間の経過に伴い、低周波数領域の音声に対する強調の度合いが大きくなる音響効果が得られる。

10

【0106】

図18は、第2の実施の形態における録音処理の一例を示すフローチャートである。第2の実施の形態の録音処理は、ステップ $S 9 2 4$ および $S 9 2 5$ をさらに実行する点において第1の実施の形態と異なる。

20

【0107】

撮像装置100は、再生時間の変更(ステップ $S 9 2 3$)の後、高解像度音声データを複製し(ステップ $S 9 2 4$)、音量レベルの調整やイコライザ処理などのエフェクト処理を実行する(ステップ $S 9 2 5$)。ステップ $S 9 2 4$ または $S 9 2 5$ の後、撮像装置100は、音声データの符号化を行う(ステップ $S 9 2 7$)。

【0108】

このように、本技術の第2の実施の形態によれば、音声処理部170は、高解像度音声データに対して複製処理を実行するため、同一の音声を繰り返し、再生させることができる。また、音声処理部170は、高解像度音声データに対して周波数特性を変更させるイコライザ処理を実行するため、低周波数領域の音声に対する強調の度合いが大きくなるなどの音響効果を生じさせることができる。これらの繰り返し再生や音響効果により、臨場感を向上させることができる。

30

【0109】

< 3 . 第3の実施の形態 >

上述の第1の実施の形態では、音声データの複製などの信号処理を撮像装置100が記録時に行っていたが、複製等を記録時に行うとストリームの記録に必要なデータサイズが増加してしまう。データサイズを削減する観点から、複製等の信号処理を再生時に行うことが望ましい。この第2の実施の形態の撮像装置100は、音声データの信号処理を再生時に行う点において第1の実施の形態と異なる。

40

【0110】

図19は、第2の実施の形態における撮像システムの一構成例を示すブロック図である。この撮像システムは、撮像装置100および再生装置200を備える。第2の実施の形態の撮像装置100は、メタデータ生成部190をさらに備える点において第1の実施の形態と異なる。

【0111】

メタデータ生成部190は、高解像度音声データの再生時刻と信号処理の内容(複製回数など)とを示す詳細設定データを付加情報から生成してメタデータに格納し、そのメタデータを記録フォーマット変換部150に供給する。第2の実施の形態の記録フォーマット変換部150は、メタデータをストリームに付加して再生装置200に供給する。再生

50

装置 200 は、ストリームを再生する装置である。

【0112】

図 20 は、第 3 の実施の形態における再生装置 200 の一構成例を示すブロック図である。この再生装置 200 は、ユーザインターフェース部 210、メタデータ分離部 220、再生制御部 230、復号部 240、複製部 250、エフェクト処理部 260、表示部 270 およびスピーカ 280 を備える。

【0113】

ユーザインターフェース部 210 は、ユーザの操作に従って操作信号を生成するものである。操作信号は、例えば、ストリームの再生開始や再生停止を指示する信号を含む。ユーザインターフェース部 210 は、操作信号をメタデータ分離部 220 に供給する。

10

【0114】

メタデータ分離部 220 は、操作信号に従ってストリームを取得し、そのストリームをメタデータと符号化データ（ビデオパケットやオーディオパケット）とに分離するものである。メタデータ分離部 220 は、分離したメタデータを再生制御部 230 に供給し、符号化データを復号部 240 に供給する。

【0115】

復号部 240 は、符号化データを元の音声データおよび動画データに復号するものである。復号部 240 は、復号により得られた音声データを複製部 250 に供給し、動画データを表示部 270 に供給する。表示部 270 は、動画データを表示するものである。

20

【0116】

再生制御部 230 は、複製部 250 およびエフェクト処理部 260 を制御するものである。この再生制御部 230 は、メタデータから、高解像度音声データの再生時刻と複製回数と、エフェクト処理の設定内容とを取得し、その音声再生時刻および複製回数を複製部 250 に供給する。また、再生制御部 230 は、エフェクト処理の設定内容をエフェクト処理部 260 に供給する。

【0117】

複製部 250 は、再生制御部 230 の制御に従って音声データを複製するものである。この複製部 250 は、復号部 240 からの音声データが供給されるたびに、その再生時刻と、再生制御部 230 からの再生時刻とが一致するか否かを判断する。一致する場合に複製部 250 は、再生制御部 230 により設定された複製回数により、音声データを複製してエフェクト処理部 260 に供給する。一方、再生時刻が一致しない場合に複製部 250 は、音声データを複製せずに、スピーカ 280 に供給する。

30

【0118】

エフェクト処理部 260 は、再生制御部 230 の制御に従って複製音声データのそれぞれに対して異なる信号処理を実行するものである。このエフェクト処理部 260 は、ゲインの調整処理やイコライザ処理などを実行し、処理後の複製音声データをスピーカ 280 に供給する。スピーカ 280 は、音声データを物理的な振動に変換して音声を再生するものである。

【0119】

なお、再生装置 200 を撮像装置 100 の外部に設ける構成としているが、再生装置 200 の機能を撮像装置 100 に設けてもよい。

40

【0120】

図 21 は、第 3 の実施の形態における M P E G 4 - A A C を用いる際の設定先のフィールドの一例を示す図である。同図に例示するように M P E G 4 - A A C 規格のメタデータ 510 内においては、例えば、D S E (Data Stream Element) 領域 511 に詳細設定データが格納される。

【0121】

図 22 は、第 2 の実施の形態における M P E G 4 - s y s t e m を用いる際の設定先のフィールドの一例を示す図である。同図に例示するように、M P E G 4 - s y s t e m 規格のメタデータ 520 内においては、例えば、u d t a 領域 521 に詳細設定データが格

50

納される。

【0122】

図23は、第2の実施の形態におけるHMPファイルフォーマットを用いる際の設定先のフィールドの一例を示す図である。同図に例示するように、HMP規格のメタデータ530内においては、例えば、uid領域531に詳細設定データが格納される。

【0123】

図24は、第3の実施の形態における録音処理の一例を示すフローチャートである。第3の実施の形態の録音処理は、ステップS926をさらに実行する点において第1の実施の形態と異なる。

【0124】

撮像装置100は、ステップS923またはS924の後、設定内容を格納したメタデータを生成し(ステップS926)、ステップS927以降を実行する。

【0125】

図25は、第3の実施の形態における再生処理の一例を示すフローチャートである。この動作は、例えば、ストリームを再生するための操作(再生ボタンの押下など)が行われたときに開始する。

【0126】

再生装置200は、符号化データを復号する復号処理を行い(ステップS951)、メタデータを参照して、復号された音声データが複製対象の高解像度音声データであるか否かを判断する(ステップS952)。複製対象である場合に(ステップS952:Yes)、再生装置200は、その高解像度音声データを複製し(ステップS953)、音量の調整やイコライザ処理などのエフェクト処理を実行する(ステップS954)。

【0127】

複製対象でない場合(ステップS952:No)またはステップS954の後に、再生装置200は、動画および音声を表示部やスピーカにより再生する(ステップS955)。そして、再生装置200は、再生終了時刻であるか否かを判断する(ステップS956)。再生終了時刻でない場合には(ステップS956:No)、撮像装置100は、ステップS951以降を繰り返す。一方、再生終了時刻である場合には(ステップS956:Yes)、再生装置200は、再生処理を終了する。

【0128】

このように、本技術の第3の実施の形態によれば、高解像度音声データの複製処理などを再生装置200が行うため、撮像装置100が記録時に音声データを複製する必要がなくなり、ストリームの記録に必要なデータサイズを小さくすることができる。

【0129】

<4. 第4の実施の形態>

上述の第3の実施の形態では、撮像装置100は、信号処理の設定内容(複製回数など)を示すメタデータを生成していたが、この設定内容をユーザの操作に従って変更してもよい。この第4の実施の形態の撮像システムは、ユーザの操作に従って複製回数などの設定内容を変更する点において第3の実施の形態と異なる。

【0130】

図26は、第4の実施の形態における撮像システムの一構成例を示すブロック図である。この第4の実施の形態の撮像システムは、再生装置200の代わりに編集装置300を備える点において第3の実施の形態と異なる。

【0131】

編集装置300は、ユーザインターフェース部310、メタデータ分離部320、編集制御部330、復号部340、再生時間変換部350、複製部360、エフェクト処理部370および再符号化部380を備える。

【0132】

ユーザインターフェース部310は、ユーザの操作に従って操作信号を生成するものである。例えば、メタデータにおける設定内容の変更を指示する操作信号が生成される。ユ

10

20

30

40

50

ーザインターフェース部 310 は、生成した操作信号を編集制御部 330 に供給する。

【0133】

メタデータ分離部 320 は、操作信号に従ってストリームをメタデータと符号化データとに分離するものである。メタデータ分離部 320 は、分離したメタデータを編集制御部 330 に供給し、符号化データを復号部 340 に供給する。

【0134】

編集制御部 330 は、操作信号に従って、メタデータの設定内容を変更するものである。複製回数および倍率の一方がユーザにより変更された場合に編集制御部 330 は、式 3 を満たすように、他方を変更する。この編集制御部 330 は、複製対象の再生時刻を復号部 340 に供給する。また、編集制御部 330 は、変更後における倍率を再生時間変換部 350 に供給し、複製回数を複製部 360 に供給し、エフェクト処理の設定内容をエフェクト処理部 370 に供給する。

【0135】

復号部 340 は、符号化データを復号するものである。この復号部 340 は、復号した高解像度音声データを再生時間変換部 350 に供給する。通常音声データと動画データについては、復号されずに撮像装置 100 に供給される。

【0136】

再生時間変換部 350 は、編集制御部 330 の制御に従って復号部 340 からの音声データの再生時間を伸長するものである。この再生時間変換部 350 は、再生時間を伸長した音声データを複製部 360 に供給する。複製部 360 は、編集制御部 330 の制御に従って高解像度音声データを複製するものである。この複製部 360 は、高解像度音声データを複製してエフェクト処理部 370 に供給する。

【0137】

エフェクト処理部 370 は、編集制御部 330 の制御に従って、音量レベルの調整処理やイコライザ処理などの信号処理を実行するものである。このエフェクト処理部 370 は、信号処理後の複製音声データを再符号化部 380 に供給する。

【0138】

再符号化部 380 は、高解像度音声データを再度符号化するものである。再符号化部 380 は、符号化により生成したストリームを撮像装置 100 に供給する。

【0139】

このように、ユーザが再生時間や複製回数を微調整することにより、再生品質をさらに向上させることができる。例えば、再生時間を「2」倍にして、「5」回複製する設定について、ユーザが、現状の再生速度は遅すぎると感じた場合には、倍率を「1.5」倍などに変更する。倍率が「1.5」倍に変更されると、編集装置 300 は、式 3 を使用して複製回数を変更する。

【0140】

なお、編集装置 300 を撮像装置 100 の外部に設ける構成としているが、編集装置 300 内の各回路を撮像装置 100 の内部に設けてもよい。

【0141】

図 27 は、第 3 の実施の形態における編集処理の一例を示すフローチャートである。この編集処理は、例えば、メタデータの編集を行うためのアプリケーションが実行されたときに開始する。

【0142】

編集装置 300 は、メタデータを分離し、ユーザの操作に従ってメタデータ内の複製回数や倍率などを変更する（ステップ S971）。また、編集装置 300 は、符号化データを復号する復号処理を実行し（ステップ S972）、複製対象の音声信号の再生時間を変更する（ステップ S973）。そして、編集装置 300 は、再生時間を変更した高解像度音声データを複製し（ステップ S974）、複製音声信号のそれぞれに対してエフェクト処理を実行する（ステップ S975）。編集装置 300 は、複製音声データを再符号化し（ステップ S976）、編集を終了するための操作が行われたか否かを判断する（ステ

10

20

30

40

50

ップS 9 7 7)。編集を終了するための操作が行われていない場合には(ステップS 9 7 7: No)、撮像装置1 0 0は、ステップS 9 7 1以降を繰り返す。一方、編集を終了するための操作が行われた場合に(ステップS 9 7 7: Yes)、撮像装置1 0 0は、編集処理を終了する。

【0 1 4 3】

このように、本技術の第4の実施の形態によれば、編集装置3 0 0は、ユーザの操作に従ってメタデータ内の複製回数や再生時間などの設定を変更するため、複製回数や再生時間を微調整して再生品質をさらに向上させることができる。

【0 1 4 4】

< 5 . 第5の実施の形態 >

上述の第1の実施の形態では、音声取込み部1 6 0は、サンプリングレート可変マイク1 6 1により、サンプリングレートを切り替えて通常音声データをサンプリングしていた。しかし、音声取込み部1 6 0は、サンプリングレートが固定のマイクによりサンプリングした高解像度音声データをリサンプリングして通常音声データを生成することもできる。この第5の実施の形態の音声取込み部1 6 0は、高解像度音声データをリサンプリングして通常音声データを生成する点において第1の実施の形態と異なる。

【0 1 4 5】

図2 8は、第5の実施の形態における音声取込み部1 6 0の一構成例を示すブロック図である。第5の実施の形態の音声取込み部1 6 0は、サンプリングレート可変マイク1 6 1の代わりに、高解像度マイク1 6 3およびサンプリングレート変換器1 6 4を備える点において第1の実施の形態と異なる。

【0 1 4 6】

高解像度マイク1 6 3は、制御信号に従って、所定のサンプリングレート(4 8キロヘルツなど)より高いサンプリングレート(9 6キロヘルツなど)で音声をサンプリングして高解像度音声データを生成するものである。この高解像度マイク1 6 3は、録音開始タイミングから録音終了タイミングまでの期間に亘って高解像度音声データを生成し、サンプリングレート変換器1 6 4に供給する。

【0 1 4 7】

サンプリングレート変換器1 6 4は、制御信号の示す高フレームレート期間外において、所定のサンプリングレート(4 8キロヘルツなど)で高解像度音声データをリサンプリングするものである。サンプリングレート変換器1 6 4は、サンプリングレート変換後の音声データを通常音声データとして音声処理部1 7 0に供給する。一方、高フレームレート期間内の高解像度音声データは、そのまま音声処理部1 7 0に供給される。

【0 1 4 8】

なお、高解像度マイク1 6 3として、デジタルの音声データを出力するデジタルマイクをもうけているが、このデジタルマイクの代わりにアナログの音声信号を出力するアナログマイクを設けてもよい。この場合には、アナログマイクからの音声信号をAD変換するAD変換器がアナログマイクとサンプリングレート変換部1 6 4との間にさらに設けられ、このAD変換器が高サンプリングレートでサンプリングを行う。

【0 1 4 9】

また、サンプリングレート変換器1 6 4は、サンプリングレートを変換する際に段階的に徐々に変換してもよい。例えば、高フレームレート期間の開始時点から一定時間に亘って、サンプリングレート変換器1 6 4は、リサンプリング時のサンプリングレートを少しずつ増加させる。また、高フレームレート期間の終了時点より一定時間前の時点から、その終了時点までの期間に亘って、音声取込み部1 6 0は、サンプリングレートをすこしずつ減少させる。

【0 1 5 0】

また、音声取込み部1 6 0は、サンプリングレートが変わる部分の違和感を軽減する目的で、サンプリングレート変換部1 6 4の後段にイコライザ処理部を追加してもよい。このイコライザ処理部は、周波数が一定値より高い高周波数帯域の音量レベルを段階的にゲ

10

20

30

40

50

インにより徐々に調整する。例えば、イコライザ処理部は、高フレームレート期間の開始時点から一定時間に亘って、高周波数帯域の音量レベルを徐々に上昇させる。また、高フレームレート期間の終了時点より一定時間前の時点から、その終了時点までの期間に亘って、イコライザ処理部は、高周波数帯域の音量レベルを徐々に低下させる。

【0151】

図29は、第5の実施の形態における録音処理の一例を示すフローチャートである。この第5の実施の形態の録音処理は、ステップS922およびS924の代わりにステップS931およびS932を実行する点において第1の実施の形態と異なる。

【0152】

録画を開始するための操作が行われると、撮像装置100は、96キロヘルツ(kHz)の高サンプリングレートで録音を行い(ステップS931)、現在時刻が高フレームレート期間内であるか否かを判断する(ステップS912)。高フレームレート期間内である場合に(ステップS921:Yes)撮像装置100は、生成した高解像度音声データの再生時間を変換する(ステップS923)。一方、高フレームレート期間内でない場合に(ステップS921:No)撮像装置100は、48キロヘルツ(kHz)の低サンプリングレートにサンプリングレートを変換する(ステップS932)。ステップS923またはS932の後に、撮像装置100は、音声データを符号化する(ステップS927)。

10

【0153】

このように、本技術の第5の実施の形態によれば、高フレームレート期間外において高解像度音声データをリサンプリングして通常音声データを生成するため、サンプリングレート可変マイク161を用いることなく通常音声データを生成することができる。

20

【0154】

<6. 第6の実施の形態>

上述の第1の実施の形態では、音声取込み部160は、1つのサンプリングレート可変マイク161により、サンプリングレートを切り替えて通常音声データをサンプリングしていた。しかし、音声取込み部160は、サンプリングレートが異なる2つのマイクにより高解像度音声データおよび通常音声データを生成することもできる。この第6の実施の形態の音声取込み部160は、サンプリングレートが異なる2つのマイクにより高解像度音声データおよび通常音声データを生成する点において第1の実施の形態と異なる。

30

【0155】

図30は、第6の実施の形態における音声取込み部160の一構成例を示すブロック図である。この第6の実施の形態の音声取込み部160は、高解像度マイク163、通常マイク165、同期出力部166、サンプリングレート変換器164および合成部167を備える。

【0156】

通常マイク165は、制御信号に従って、所定のサンプリングレート(48キロヘルツなど)で音声データをサンプリングして通常音声データを生成するものである。この通常マイク165は、録音開始タイミングから録音終了タイミングまでの期間に亘って通常音声データを生成し、同期出力部166に供給する。

40

【0157】

高解像度マイク163は、制御信号に従って、所定のサンプリングレートより高いサンプリングレート(96キロヘルツなど)で音声データをサンプリングして高解像度音声データを生成するものである。この高解像度マイク163は、録音開始タイミングから録音終了タイミングまでの期間に亘って高解像度音声データを生成し、同期出力部166に供給する。

【0158】

なお、高解像度マイク163および通常マイク165として、デジタルの音声データを出力するデジタルマイクをもうけているが、これらのデジタルマイクの代わりにアナログの音声信号を出力するアナログマイクを1つ設けてもよい。この場合には、アナログマイ

50

クからの音声信号をAD変換する2つのAD変換器がさらに設けられ、これらのAD変換器が互いに異なるサンプリングレートでサンプリングを行う。

【0159】

同期出力部166は、高解像度音声データおよび通常音声データを所定の同期信号に同期して出力するものである。この同期出力部166は、通常音声データをサンプリングレート変換器164に出力し、高解像度音声データを合成部167に出力する。

【0160】

サンプリングレート変換器164は、通常音声データのサンプリングレートを、必要に応じて、より高いサンプリングレートに変換するものである。このサンプリングレート変換器164は、変換後の通常音声データを合成部167に供給する。

10

【0161】

合成部167は、通常音声データおよび高解像度音声データを合成するものである。合成部167は、高フレームレート期間外において高解像度音声データの比率を「0」に設定して、通常音声データのみを選択して出力する。一方、高フレームレート期間内において合成部167は、フェード期間内であれば、高解像度音声データに通常音声データを合成した音声データを出力し、フェード期間外であれば、高解像度音声データの比率を「1」に設定して高解像度音声データを選択して出力する。なお、合成部167は、特許請求の範囲に記載の選択部の一例である。

【0162】

ここで、フェード期間は、フェードイン期間と、フェードアウト期間とからなる。フェードイン期間は、高フレームレート期間の開始タイミングから一定時間が経過するまでの期間である。一方、フェードアウト期間は、高フレームレート期間の終了タイミングより一定時間前のタイミングから、その終了タイミングまでの期間である。

20

【0163】

合成部167は、フェードイン期間において、フェードイン期間より短い単位時間が経過するたびに、合成における高解像度音声データの比率を高くする。これにより、高解像度音声データの割合が徐々に大きくなる。一方、合成部167は、フェードアウト期間において、フェードアウト期間より短い単位時間が経過するたびに、合成における高解像度音声データの比率を低くする。これにより、高解像度音声データの割合が徐々に小さくなる。このように、フェードインおよびフェードアウトによりデータの割合を徐々に変更する処理は、クロスフェード処理と呼ばれる。このクロスフェード処理により、通常音声データおよび高解像度音声データの一方から他方への切り替えの部分の違和感を軽減することができる。

30

【0164】

なお、音声取込み部160は、合成部167によりクロスフェード処理を行っているが、この構成に限定されない。音声取込み部160は、合成部167の代わりにセクタなどを備え、クロスフェード処理を行わずに音声データをセクタで切り替えて出力してもよい。この場合に、違和感を軽減する目的で、セクタの後段にイコライザ処理部を追加し、イコライザ処理部が高周波数帯域の音量レベルを段階的にゲインにより徐々に調整してもよい。例えば、イコライザ処理部は、フェードイン期間に相当する期間において、高周波数帯域の音量レベルを徐々に上昇させ、フェードアウト期間に相当する期間において、高周波数帯域の音量レベルを徐々に低下させればよい。

40

【0165】

また、音声取込み部160は、2つのマイクにより音声をサンプリングしているが、3つ以上のマイクによりサンプリングしてもよい。例えば、サンプリングレートが48、96および192キロヘルツ(kHz)の3つのマイクを備え、合成部167が、それらを合成してもよい。例えば、フェードイン期間内の開始時点から、ある時点までにおいて合成部167が48キロヘルツの音声と96キロヘルツの音声とを合成して、96キロヘルツの割合を徐々に高くする。そして、その時点からフェードイン期間の終了時点までにおいて合成部167が96キロヘルツの音声と192キロヘルツの音声とを合成して、19

50

2キロヘルツの割合を徐々に高くする。フェードアウト期間では、合成部167はフェードイン期間と対照的な処理を行えばよい。

【0166】

図31は、第6の実施の形態における合成比率の変動の一例を示すグラフである。同図において縦軸は高解像度音声データの合成比率を示し、横軸は時間を示す。

【0167】

同図に例示するように、高フレームレート期間外では高解像度フレームレートの合成比率は、「0」に設定される。この合成比率により、合成部167は、通常音声データを選択して出力する。

【0168】

高フレームレート期間内のフェードイン期間において、合成部167は、高解像度音声データの比率を徐々に高くして合成を行う。また、高フレームレート期間内のフェードアウト期間において、合成部167は、高解像度音声データの比率を徐々に低くして合成を行う。また、高フレームレート期間のうちフェード期間外において高解像度音声データの比率は「1」に設定される。この合成比率により、合成部167は、高解像度音声データを選択して出力する。

【0169】

図32は、第6の実施の形態における録音処理の一例を示すフローチャートである。撮像装置100は、まず、高サンプリングレート(96キロヘルツなど)および低サンプリングレート(48キロヘルツなど)のそれぞれで録音する(ステップS941)。そして、撮像装置100は、現在時刻が高フレームレート期間内であるか否かを判断する(ステップS942)。

【0170】

高フレームレート期間内である場合に(ステップS942: Yes)、撮像装置100は、通常音声データのサンプリングレートを必要に応じて変換する(ステップS943)。そして、撮像装置100は、クロスフェード処理を行う(ステップS944)。一方、高フレームレート期間外である場合に(ステップS942: No)、撮像装置100は、通常音声データを選択する(ステップS945)。

【0171】

ステップS944またはS945の後に撮像装置100は、音声データを符号化し(ステップS946)、録音を終了するための操作(停止ボタンの押下など)が行われたか否かを判断する(ステップS947)。録音を終了するための操作が行われていない場合には(ステップS947: No)、撮像装置100は、ステップS941以降を繰り返す。一方、録画を終了するための操作が行われた場合には(ステップS947: Yes)、撮像装置100は、録音処理を終了する。

【0172】

このように、本技術の第6の実施の形態によれば、撮像装置100が、高解像度音声データおよび通常音声データを別々に生成し、いずれかを選択して出力するため、それらの音声データをサンプリングレート可変マイク161を用いずに生成することができる。

【0173】

なお、上述の実施の形態は本技術を具現化するための一例を示したものであり、実施の形態における事項と、特許請求の範囲における発明特定事項とはそれぞれ対応関係を有する。同様に、特許請求の範囲における発明特定事項と、これと同一名称を付した本技術の実施の形態における事項とはそれぞれ対応関係を有する。ただし、本技術は実施の形態に限定されるものではなく、その要旨を逸脱しない範囲において実施の形態に種々の変形を施すことにより具現化することができる。

【0174】

また、上述の実施の形態において説明した処理手順は、これら一連の手順を有する方法として捉えてもよく、また、これら一連の手順をコンピュータに実行させるためのプログラム乃至そのプログラムを記憶する記録媒体として捉えてもよい。この記録媒体として、

10

20

30

40

50

例えば、C D (Compact Disc)、M D (MiniDisc)、D V D (Digital Versatile Disc)、メモ리카ード、ブルーレイディスク (Blu-ray (登録商標) Disc) 等を用いることができる。

【0175】

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いずれかの効果であってもよい。

【0176】

なお、本技術は以下のような構成もとることができる。

(1) 所定のサンプリングレートより高いサンプリングレートで所定の期間内において音声をサンプリングして音声データを高解像度音声データとして生成する処理と前記所定のサンプリングレートで前記所定の期間外において音声をサンプリングして音声データを通常音声データとして生成する処理とを行うサンプリング処理部と、

前記高解像度音声データの再生時間を伸長する再生時間変換部とを具備する録音装置。

(2) 前記サンプリング処理部は、前記所定の期間外において前記所定のサンプリングレートで前記音声をサンプリングし、前記所定の期間内において前記所定のサンプリングレートより高いサンプリングレートにサンプリングレートを切り替えて前記音声をサンプリングする

前記(1)記載の録音装置。

(3) 前記サンプリング処理部は、

前記所定のサンプリングレートより高いサンプリングレートで前記音声をサンプリングして前記高解像度音声データを生成する高解像度マイクと、

前記所定の期間外において前記高解像度音声データを前記所定のサンプリングレートでサンプリングして前記通常音声データを生成するサンプリングレート変換器と

を備える前記(1)記載の録音装置。

(4) 前記サンプリング処理部は、

前記所定のサンプリングレートより高いサンプリングレートで前記音声をサンプリングして前記高解像度音声データを生成する高解像度マイクと

前記所定のサンプリングレートで前記音声をサンプリングして前記通常音声データを生成する通常マイクと、

前記所定の期間内において前記高解像度音声データを選択して出力し、前記所定の期間外において前記通常音声データを選択して出力する選択部と

を備える請求項1記載の録音装置。

(5) 前記選択部は、前記所定の期間内の一定のフェード期間において前記高解像度音声データに前記通常音声データを合成する合成処理を行う

前記(4)記載の録音装置。

(6) 前記選択部は、前記合成処理において前記フェード期間より短い単位時間が経過するたびに前記高解像度音声データの割合を変更する

前記(5)記載の録音装置。

(7) 所定のフレームレートより高いフレームレートで複数のフレームを撮像する撮像部と、

前記複数のフレームのうち前記所定の期間外に撮像されたフレームのフレームレートを前記所定のフレームレートに変換するフレームレート変換部と

をさらに具備する前記(1)から(7)のいずれかに記載の録音装置。

(8) 所定のタイミングを含む期間を前記所定の期間として設定する制御部をさらに具備する

前記(7)記載の録音装置。

(9) 前記複数のフレームの中からシーンが変化したシーン変化タイミングを検出するシーン変化検出部をさらに具備し、

前記制御部は、前記シーン変化タイミングを含む期間を前記所定の期間として設定する

前記(8)記載の録音装置。

(10) 所定の検知対象を検知するセンサーをさらに具備し、

前記制御部は、前記検知対象が検知されたタイミングを含む期間を前記所定の期間に設定する前記(8)または(9)に記載の録音装置。

(11) 前記再生時間が伸長された高解像度音声データに対して所定の信号処理を実行する信号処理部をさらに具備する

前記(1)から(10)のいずれかに記載の録音装置。

(12) 前記信号処理部は、前記高解像度音声データを複製する

前記(11)記載の録音装置。

(13) 前記信号処理部は、前記高解像度音声データの音量レベルを所定のゲインにより調整する

前記(11)または(12)に記載の録音装置。

(14) 前記信号処理部は、前記高解像度音声データの周波数特性を変更する

前記(11)から(13)のいずれかに記載の録音装置。

(15) 所定のサンプリングレートより高いサンプリングレートで所定の期間内において音声データをサンプリングして音声データを高解像度音声データとして生成する処理と前記所定のサンプリングレートで前記所定の期間外において音声データをサンプリングして音声データを通常音声データとして生成する処理とを行い、前記高解像度音声データの再生時間を伸長し、前記再生時間を伸長した前記高解像度音声データに対して実行すべき信号処理を示す設定情報を含むメタデータを生成する録音装置と、

前記設定情報に従って前記信号処理を実行して前記信号処理を実行した前記高解像度音声データと前記通常音声データとを再生する再生装置とを具備する録音システム。

(16) 前記メタデータのフォーマットは、MPEG4-AACであり、

前記録音装置は、前記メタデータのDSE(Data Stream Element)領域に前記設定情報を記録する

前記(15)記載の録音システム。

(17) 前記メタデータのフォーマットは、MPEG4-systemであり、

前記録音装置は、前記メタデータのudta領域に前記設定情報を記録する

前記(15)記載の録音システム。

(18) 前記メタデータのフォーマットは、HMM P(Home and Mobile Multimedia Platform)であり、

前記録音装置は、前記メタデータのuuid領域に前記設定情報を記録する

前記(15)記載の録音システム。

(19) 所定のサンプリングレートより高いサンプリングレートで所定の期間内において音声データをサンプリングして音声データを高解像度音声データとして生成する処理と前記所定のサンプリングレートで前記所定の期間外において音声データをサンプリングして音声データを通常音声データとして生成する処理とを行い、前記高解像度音声データの再生時間を伸長し、前記再生時間を伸長した前記高解像度音声データに対して実行すべき信号処理を示す設定情報を含むメタデータを生成する録音装置と、

前記設定情報を変更して当該変更した設定情報の示す前記信号処理を実行する編集装置と

を具備する録音システム。

(10) 所定のサンプリングレートより高いサンプリングレートで所定の期間内において音声データをサンプリングして音声データを高解像度音声データとして生成する処理と前記所定のサンプリングレートで前記所定の期間外において音声データをサンプリングして音声データを通常音声データとして生成する処理とを行うサンプリング処理手順と、

前記高解像度音声データの再生時間を伸長する再生時間変換手順とを具備する録音方法。

【符号の説明】

10

20

30

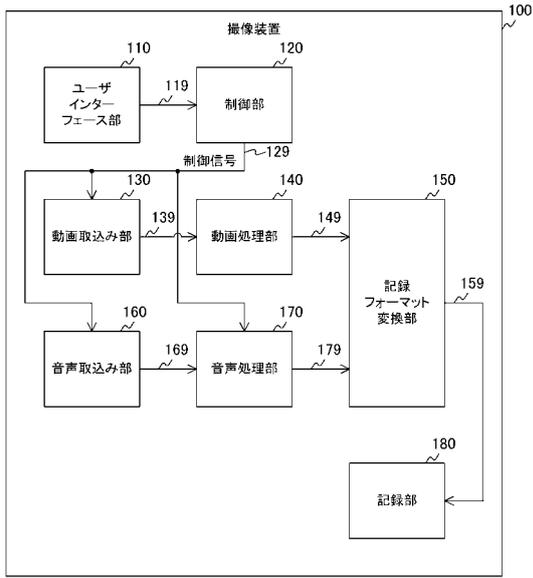
40

50

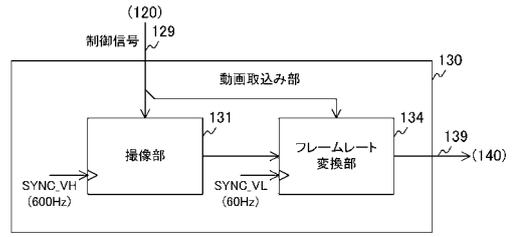
【 0 1 7 7 】

1 0 0	撮像装置	
1 0 1	録音装置	
1 1 0、2 1 0、3 1 0	ユーザインターフェース部	
1 1 5	センサー	
1 2 0、1 2 5	制御部	
1 3 0	動画取込み部	
1 3 1	撮像部	
1 3 2、1 7 1	バッファ	
1 3 3	シーン変化検出部	10
1 3 4	フレームレート変換部	
1 4 0	動画処理部	
1 5 0	記録フォーマット変換部	
1 6 0	音声取込み部	
1 6 1	サンプリングレート可変マイク	
1 6 2	付加情報生成部	
1 6 3	高解像度マイク	
1 6 4	サンプリングレート変換器	
1 6 5	通常マイク	
1 6 6	同期出力部	20
1 6 7	合成部	
1 7 0	音声処理部	
1 7 2、3 5 0	再生時間変換部	
1 7 3、2 5 0、3 6 0	複製部	
1 7 4、2 6 0、3 7 0	エフェクト処理部	
1 7 5	ゲイン調整部	
1 7 6	イコライザ処理部	
1 7 7	音声符号化部	
1 8 0	記録部	
1 9 0	メタデータ生成部	30
2 0 0	再生装置	
2 2 0、3 2 0	メタデータ分離部	
2 3 0	再生制御部	
2 4 0、3 4 0	復号部	
2 7 0	表示部	
2 8 0	スピーカ	
3 0 0	編集装置	
3 3 0	編集制御部	
3 8 0	再符号化部	

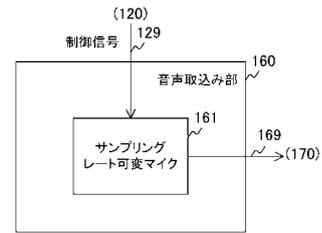
【図1】



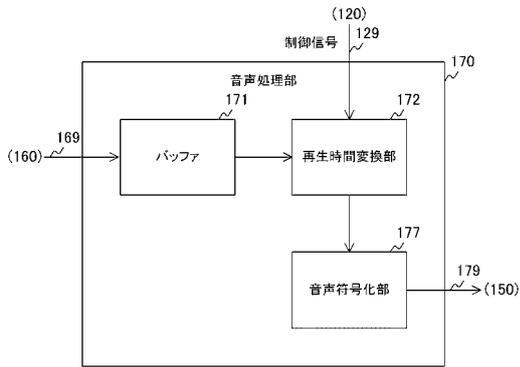
【図2】



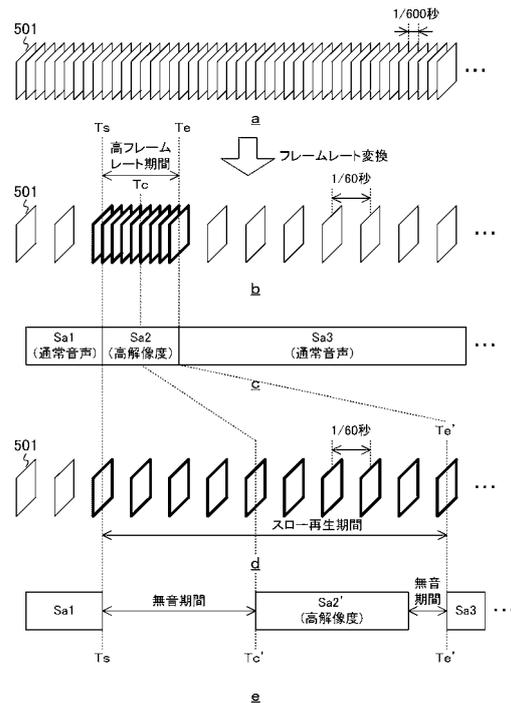
【図3】



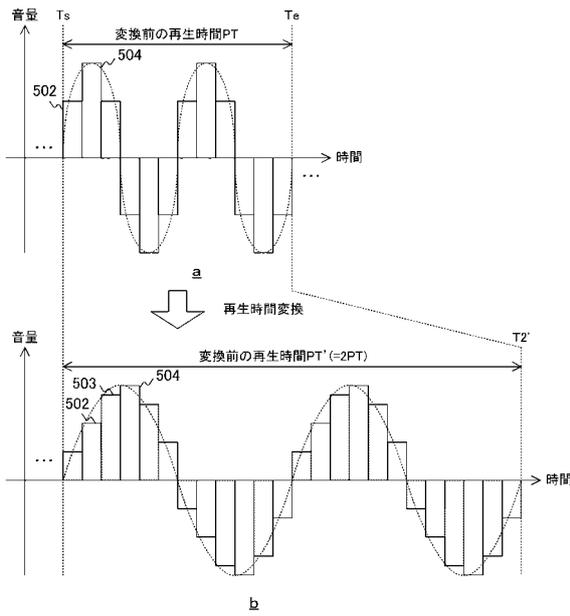
【図4】



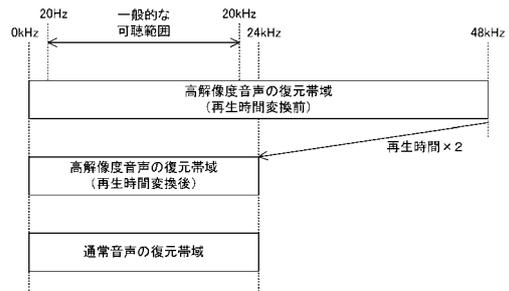
【図5】



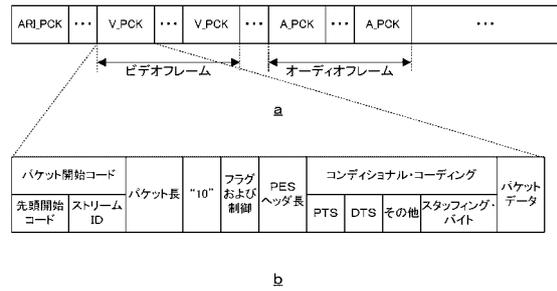
【 図 6 】



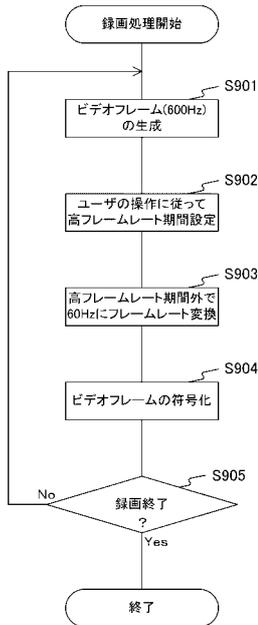
【 図 7 】



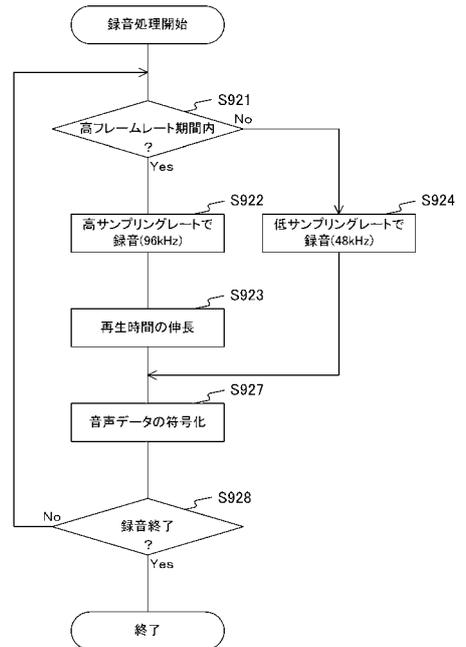
【 図 8 】



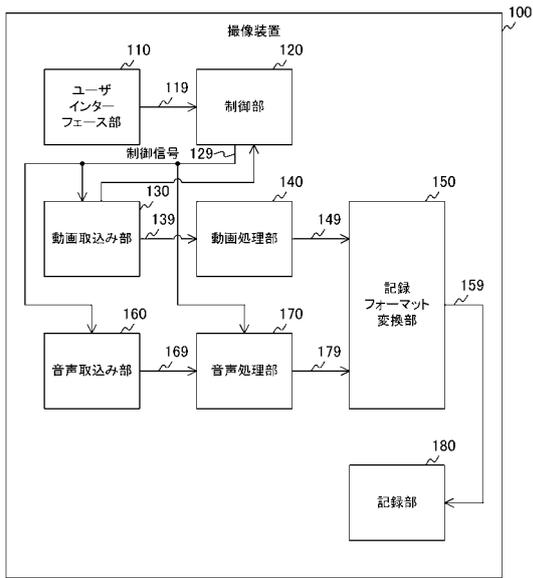
【 図 9 】



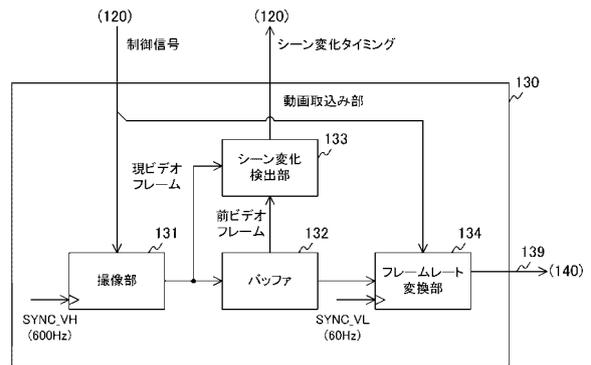
【 図 10 】



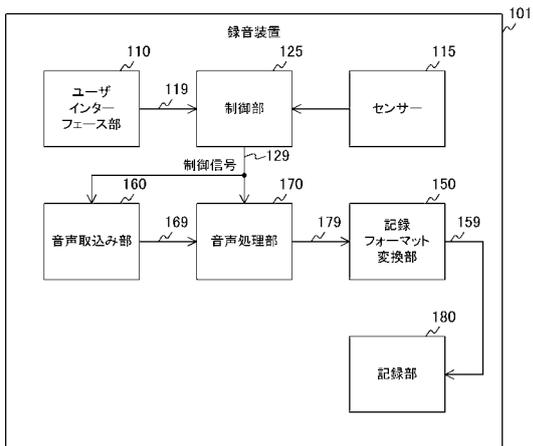
【図 1 1】



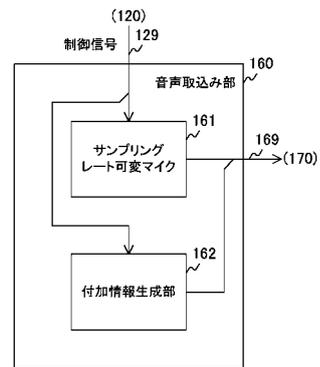
【図 1 2】



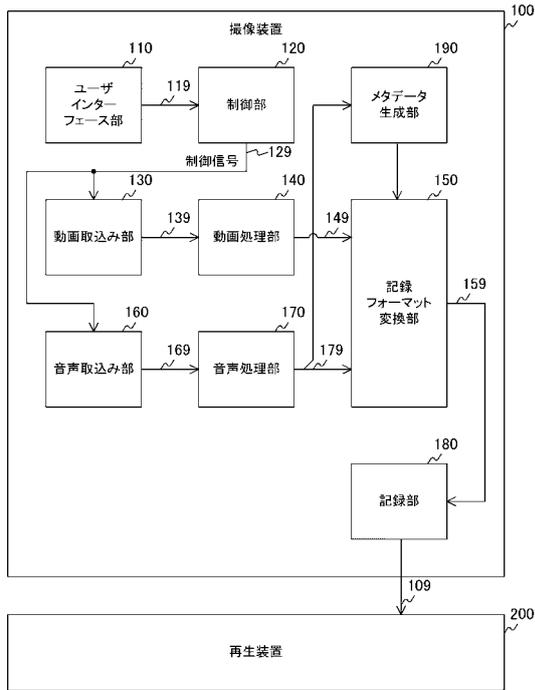
【図 1 3】



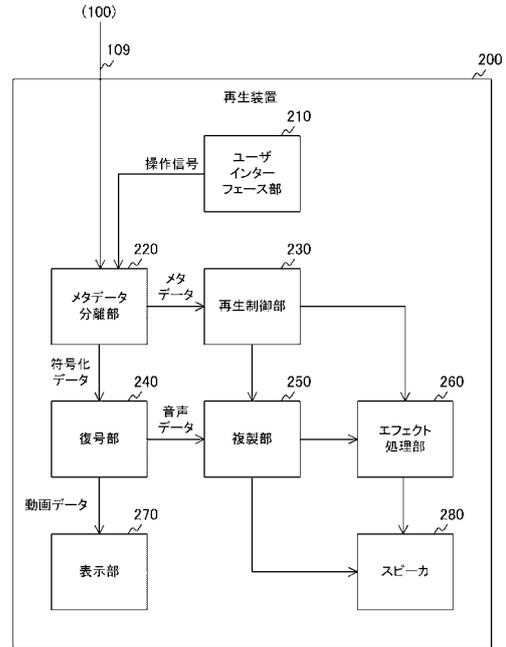
【図 1 4】



【図 19】



【図 20】



【図 21】

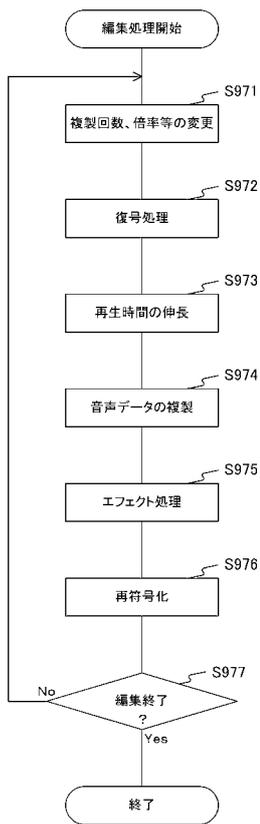
```

Syntax                                     No. of bits  Mnemonic
raw_data_block()
{
  while((id=id_syn_ele) != ID_END){
    switch(id)
    {
      case ID_SCE: single_channel_element(); 3      uimsbf
      break;
      case ID_CPE: channel_pair_element();
      break;
      case ID_CCE: coupling_channel_element();
      break;
      case ID_LFE: lfe_channel_element(); 511
      break;
      case ID_DSE: data_stream_element();
      break;
      case ID_PCE: program_config_element();
      break;
      case ID_FIL: fill_element();
    }
  }
  byte_alignment();
}
    
```

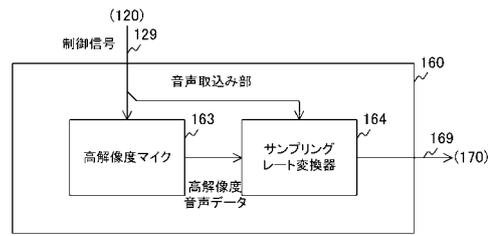
【図 22】

Atom	Code	Length	Value	Description
Moov			*	13.2.3.1 container for all the meta-data
mvhd			*	13.2.3.3 movie header, overall declarations
oods			*	13.2.3.4 object descriptor
trak			*	13.2.3.4.2 container for an individual track or stream
tkhd			*	13.2.3.6 track header, overall information about the track
trf				13.2.3.7 track reference container
edts				13.2.3.25 edit list container
elst				13.2.3.26 an edit list
mdia			*	13.2.3.8 container for the media information in a track
mdhd			*	13.2.3.9 media header, overall information about the media
hdlr				13.2.3.10 handler at this level, the media (handler) type
minf			*	13.2.3.11 media information container
vmhd				13.2.3.12.1 video media header, overall information (video track only)
smhd				13.2.3.12.2 sound media header, overall information (sound track only)
hmhd				13.2.3.12.3 hint media header, overall information (hint track only)
<mpeg>				13.2.3.12.4 mpeg stream headers
dinf			*	13.2.3.13 data information atom, container
dref			*	13.2.3.14 data reference atom, declares source(s) of media in track
stbl			*	13.2.3.15 sample table atom, container for the time/space map
stts			*	13.2.3.16.1 (decoding) time-to-sample
:	:	:	:	:
mdat				13.2.3.2 Media data container
free				13.2.3.24 free space
skip				13.2.3.24 free space
udta				13.2.3.27 user-data, copyright etc.

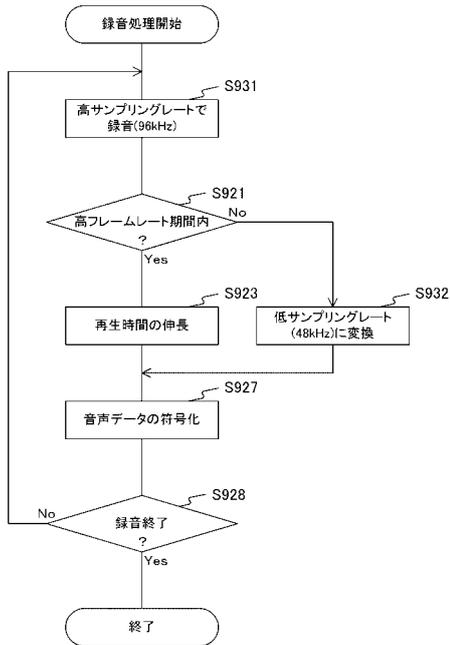
【 図 2 7 】



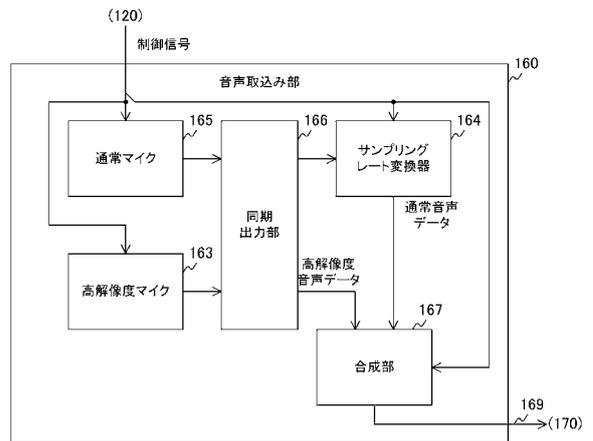
【 図 2 8 】



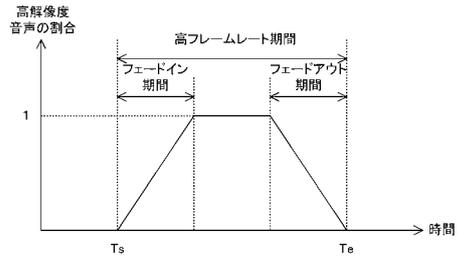
【 図 2 9 】



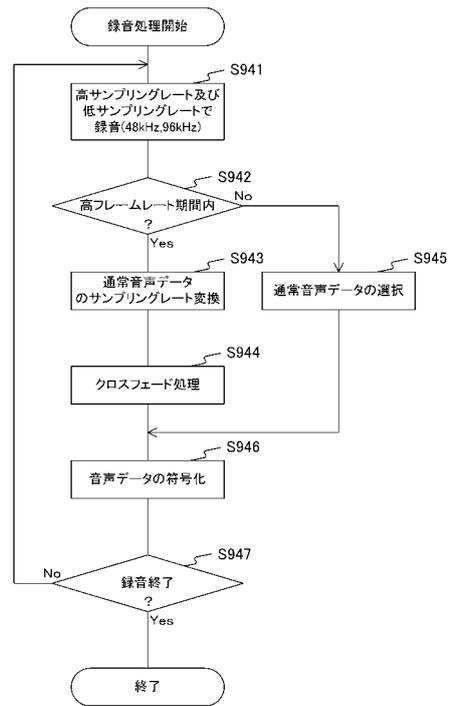
【 図 3 0 】



【 図 3 1 】



【 図 3 2 】



フロントページの続き

(51)Int.Cl.

F I

H 0 4 N 5/91

C

テーマコード(参考)