

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4367459号  
(P4367459)

(45) 発行日 平成21年11月18日(2009.11.18)

(24) 登録日 平成21年9月4日(2009.9.4)

|                         |                       |
|-------------------------|-----------------------|
| (51) Int.Cl.            | F I                   |
| G 1 0 L 19/00 (2006.01) | G 1 0 L 19/00 2 2 0 F |
| G 1 0 L 19/02 (2006.01) | G 1 0 L 19/02 1 5 0   |
| G 1 0 L 19/06 (2006.01) | G 1 0 L 19/06 B       |

請求項の数 2 (全 19 頁)

|            |                                     |           |                       |
|------------|-------------------------------------|-----------|-----------------------|
| (21) 出願番号  | 特願2006-187733 (P2006-187733)        | (73) 特許権者 | 000004329             |
| (22) 出願日   | 平成18年7月7日(2006.7.7)                 |           | 日本ビクター株式会社            |
| (62) 分割の表示 | 特願2006-184164 (P2006-184164)<br>の分割 |           | 神奈川県横浜市神奈川区守屋町3丁目12番地 |
| 原出願日       | 平成10年11月16日(1998.11.16)             | (72) 発明者  | 田中 美昭                 |
| (65) 公開番号  | 特開2007-11381 (P2007-11381A)         |           | 神奈川県横浜市神奈川区守屋町3丁目12番地 |
| (43) 公開日   | 平成19年1月18日(2007.1.18)               |           | 日本ビクター株式会社内           |
| 審査請求日      | 平成18年7月31日(2006.7.31)               | (72) 発明者  | 植野 昭治                 |
|            |                                     |           | 神奈川県横浜市神奈川区守屋町3丁目12番地 |
|            |                                     |           | 日本ビクター株式会社内           |
|            |                                     | (72) 発明者  | 淵上 徳彦                 |
|            |                                     |           | 神奈川県横浜市神奈川区守屋町3丁目12番地 |
|            |                                     |           | 日本ビクター株式会社内           |
|            |                                     | 審査官       | 山下 剛史                 |

最終頁に続く

(54) 【発明の名称】 音声符号化方法及び音声復号化方法

(57) 【特許請求の範囲】

【請求項1】

ステレオ2チャンネルの音声信号を、そのままのチャンネル又は互いに相関をとったチャンネル毎に、入力される音声信号にตอบสนองして先頭サンプル値を所定時間のフレーム単位で得ると共に、特性が異なる複数の線形予測方法により時間領域の過去から現在の信号の線形予測値がそれぞれ予測され、その予測される線形予測値と前記音声信号とから得られる予測残差が最小となるような線形予測方法を選択して予測符号化するステップと、

ヘッダ情報と、圧縮PCMアクセスユニットを含むユーザデータと、を含んだデータ構造にすると共に、前記圧縮PCMアクセスユニットは前記フレーム中に複数設けられ、前記ステップにより選択された各チャンネルの線形予測方法と予測残差を含む予測符号化データを、前記圧縮PCMアクセスユニット内に配置されるサブパッケージに格納する場合に前記予測残差をビット情報に基づいたビット数でパッキングして格納し、前記音声信号のサンプリング周波数に応じたサンプル数により格納し、前記圧縮PCMアクセスユニットが前記フレーム中の先頭のものである場合には、さらに、前記サブパッケージにリスタートヘッダを設けると共に前記先頭サンプル値を収納し、また、さらに、前記各圧縮PCMアクセスユニット内に前記サブパッケージ内のデータが可変ビットレート圧縮された圧縮データであることを示すVBR識別子と、再生側において元のアナログ音声信号に復元される際に用いられるサンプリング周波数及び量子化ビット数とを含む同期情報部を設けるステップと、

からなる音声符号化方法。

## 【請求項 2】

請求項 1 記載の音声符号化方法により符号化されたデータから元の音声信号を復号する音声復号方法であって、

前記ユーザデータから圧縮 P C M アクセスユニットを抽出するステップと、

前記圧縮 P C M アクセスユニットから V B R 識別子を抽出するステップと、

前記抽出された識別子に基づいて前記リスタートヘッダを有する前記圧縮 P C M アクセスユニットから前記先頭サンプル値を取り出すと共に、前記各圧縮 P C M アクセスユニットからサンプリング周波数に応じたサンプル数の予測残差および線形予測方法を含む予測符号化データを取り出すステップと、

前記予測残差をビット情報に基づいたビット数で復号し、この復号した予測残差と前記先頭サンプル値と線形予測方法とに基づいて予測値を算出するステップと、

この算出された予測値から前記ステレオ 2 チャンルの音声信号を復元するステップと、

前記復元されたステレオ 2 チャンルの音声データを前記同期情報部内のサンプリング周波数及び量子化ビット数に基づいてアナログ音声信号に変換するステップと、  
からなる音声復号方法。

## 【発明の詳細な説明】

## 【技術分野】

## 【0001】

本発明は、音声信号の音声符号化方法及び音声復号化方法に関する。

## 【背景技術】

## 【0002】

音声信号を可変長で圧縮する方法として、本発明者は先の出願（特願平 9 - 289159 号）において 1 チャンルの原デジタル音声信号に対して、特性が異なる複数の予測器により時間領域における過去の信号から現在の信号の複数の線形予測値を算出し、原デジタル音声信号と、この複数の線形予測値から予測器毎の予測残差を算出し、予測残差の最小値を選択する予測符号化方法を提案している。

## 【0003】

なお、上記方法では原デジタル音声信号がサンプリング周波数 = 96 kHz、量子化ビット数 = 20 ビット程度の場合にある程度の圧縮効果を得ることができるが、近年の DVD オーディオディスクではこの 2 倍のサンプリング周波数（= 192 kHz）が使用され、また、量子化ビット数も 24 ビットが使用される傾向がある。また、マルチチャンネルにおけるサンプリング周波数と量子化ビット数はチャンネル毎に異なることもある。

## 【発明の開示】

## 【発明が解決しようとする課題】

## 【0004】

ところで、マルチチャンネルの音声信号を伝送する場合、著作権者がオーディオソースに依っては圧縮を希望するものとそうでないものがあり、また、ユーザがマルチチャンネルをステレオ 2 チャンネルにダウンミクスして再生することを望まないものとそうでないものの 2 通りがある。したがって、このように圧縮又は非圧縮で選択的に伝送する 2 通りと、再生側のダウンミクスを選択的に許可、禁止する 2 通りの合計 4 通りで伝送した場合には、再生側でこれを識別して選択的に再生する必要がある。

また、サンプリング周波数や量子化ビット数が高くなることから、従来以上に圧縮率を改善する必要がある。

## 【0005】

そこで本発明は、より圧縮率を改善したり、再生側のダウンミクスを選択的に許可又は禁止しても再生側が正常に再生することができる音声符号化方法及び音声復号化方法を提供することを目的とする。

## 【課題を解決するための手段】

## 【 0 0 0 6 】

本発明は上記目的を達成するために、以下の 1 ) 及び 2 ) の手段より成る。

すなわち、

1 ) ステレオ 2 チャンネルの音声信号を、そのままのチャンネル又は互いに相関をとったチャンネル毎に、入力される音声信号に応答して先頭サンプル値を所定時間のフレーム単位で得ると共に、特性が異なる複数の線形予測方法により時間領域の過去から現在の信号の線形予測値がそれぞれ予測され、その予測される線形予測値と前記音声信号とから得られる予測残差が最小となるような線形予測方法を選択して予測符号化するステップと、

ヘッダ情報と、圧縮 P C M アクセスユニットを含むユーザデータと、を含んだデータ構造にすると共に、前記圧縮 P C M アクセスユニットは前記フレーム中に複数設けられ、前記ステップにより選択された各チャンネルの線形予測方法と予測残差を含む予測符号化データを、前記圧縮 P C M アクセスユニット内に配置されるサブパケットに格納する場合に前記予測残差をビット情報に基づいたビット数でパッキングして格納し、前記音声信号のサンプリング周波数に応じたサンプル数により格納し、前記圧縮 P C M アクセスユニットが前記フレーム中の先頭のものである場合には、さらに、前記サブパケットにリスタートヘッダを設けると共に前記先頭サンプル値を収納し、また、さらに、前記各圧縮 P C M アクセスユニット内に前記サブパケット内のデータが可変ビットレート圧縮された圧縮データであることを示す V B R 識別子と、再生側において元のアナログ音声信号に復元される際に用いられるサンプリング周波数及び量子化ビット数とを含む同期情報部を設けるステップと、

からなる音声符号化方法。

2 ) 請求項 1 記載の音声符号化方法により符号化されたデータから元の音声信号を復号する音声復号方法であって、

前記ユーザデータから圧縮 P C M アクセスユニットを抽出するステップと、

前記圧縮 P C M アクセスユニットから V B R 識別子を抽出するステップと、

前記抽出された識別子に基づいて前記リスタートヘッダを有する前記圧縮 P C M アクセスユニットから前記先頭サンプル値を取り出すと共に、前記各圧縮 P C M アクセスユニットからサンプリング周波数に応じたサンプル数の予測残差および線形予測方法を含む予測符号化データを取り出すステップと、

前記予測残差をビット情報に基づいたビット数で復号し、この復号した予測残差と前記先頭サンプル値と線形予測方法とに基づいて予測値を算出するステップと、

この算出された予測値から前記ステレオ 2 チャンネルの音声信号を復元するステップと、

前記復元されたステレオ 2 チャンネルの音声データを前記同期情報部内のサンプリング周波数及び量子化ビット数に基づいてアナログ音声信号に変換するステップと、

からなる音声復号方法。

## 【 発明の効果 】

## 【 0 0 0 7 】

以上説明したように本発明によれば、従来以上に圧縮率を改善した音声符号化ができると共に、これにより符号化された音声信号を不都合なく音声復号できる。

## 【 発明を実施するための最良の形態 】

## 【 0 0 0 8 】

以下、図面を参照して本発明の実施の形態を説明する。図 1 ~ 図 4 は本発明が適用されるマルチチャンネル伝送形態を実現する音声符号化装置の処理を示す説明図である。

## 【 0 0 0 9 】

ここで、マルチチャンネル方式としては、例えば次の 4 つの方式が知られている。

( 1 ) 4 チャンネル方式          ドルビーサラウンド方式のように、前方 L、C、R の 3 チャンネル + 後方 S の 1 チャンネルの合計 4 チャンネル

10

20

30

40

50

(2) 5チャンネル方式      ドルビーAC-3方式のSWチャンネルなしのように、前方L、C、Rの3チャンネル+後方SL、SRの2チャンネルの合計5チャンネル

(3) 6チャンネル方式      DTS (Digital Theater System) 方式や、ドルビーAC-3方式のように6チャンネル(L、C、R、SW(Lfe)、SL、SR)

(4) 8チャンネル方式      SDDS (Sony Dynamic Digital Sound) 方式のように、前方L、LC、C、RC、R、SWの6チャンネル+後方SL、SRの2チャンネルの合計8チャンネル

【0010】

図1は第1の例の伝送形態として、マルチチャンネルを圧縮するとともに再生側のダウンミクスを禁止する場合を示している。符号化側の6チャンネル(ch)ミクス&マトリクス回路1'は、マルチチャンネル信号の一例としてフロントレフト(Lf)、センタ(C)、フロントライト(Rf)、サラウンドレフト(Ls)、サラウンドライト(Rs)及びLfe (Low Frequency Effect) の6chのPCMデータを次式(1-1)により6ch「1」~「6」分の相関信号に変換し、符号化部2'に出力する。

$$「1」 = Lf + Rf - C$$

$$「2」 = Lf - Rf - C$$

$$「3」 = C - (Ls + Rs) / 2$$

$$「4」 = Ls + Rs$$

$$「5」 = Ls - Rs$$

$$「6」 = Lfe - a \times C$$

ただし、 $0 < a < 1$       ... (1-1)

このような6チャンネル(ch)ミクス&マトリクス回路1'による相関式と符号化部2'の符号化方式は選択手段7'で選択される。以下説明する図2、図3、図4、図5及び図6でも同様であるので、これらの図では選択手段7'を略すことにする。

【0011】

第1と第2の符号化部2'-1、2'-2を有する符号化部2'は図7に詳しく示すようにこの6ch「1」~「6」のPCMデータを予測符号化し、予測符号化データを図8に示すようなビットストリームで記録媒体5や通信媒体6を介して復号側に伝送する。復号側では第1と第2の復号化部3'-1、3'-2を有する復号化部3'により、図14に詳しく示すように6ch「1」~「6」の予測符号化データをPCMデータに復号し、次いでミクス&マトリクス回路4'により式(1-1)に基づいて元の6ch(Lf、C、Rf、Ls、Rs、Lfe)のみを復元する。

【0012】

図2は第2の例の伝送形態として、マルチチャンネルを圧縮するとともに再生側のダウンミクスを許可する場合を示している。符号化側の6chミクス&マトリクス回路1'は、元の6ch(Lf、C、Rf、Ls、Rs、Lfe)と係数 $m_{ij}$  ( $i = 1, 2, j = 1, 2 \sim 6$ )により次式(2)のようにステレオ2chデータ(L、R)を生成(ダウンミクス)する。

$$L = m_{11} \cdot Lf + m_{12} \cdot Rf + m_{13} \cdot C \\ + m_{14} \cdot Ls + m_{15} \cdot Rs + m_{16} \cdot Lfe$$

$$R = m_{21} \cdot Lf + m_{22} \cdot Rf + m_{23} \cdot C$$

$$+ m_{24} \cdot Ls + m_{25} \cdot Rs + m_{26} \cdot Lfe \quad \dots (2)$$

【0013】

そして、式(2)と次式(1-2)により次のような第1グループの2チャンネル分の相関信号「1」、「2」と第2グループの4チャンネル分の相関信号「3」~「6」に変換し、それぞれ第1符号化部2'-1、第2符号化部2'-2に出

10

20

30

40

50

力する。

$$「1」 = L + R$$

$$「2」 = L - R$$

$$「3」 \sim 「6」 \text{は式} (1 - 1) \text{と同じ} \quad \dots (1 - 2)$$

#### 【0014】

第1、第2符号化部2'-1、2'-2はそれぞれ第1グループチャンネル「1」、「2」と第2グループチャンネル「3」～「6」のPCMデータを予測符号化し、各チャンネルの予測符号化データを記録媒体5や通信媒体6を介して復号側に伝送する。復号側では第1、第2復号化部3'-1、3'-2により、それぞれ第1グループチャンネル「1」、「2」と第2グループチャンネル「3」～「6」の予測符号化データをPCMデータに復号し、次いでミクス&マトリクス回路4'により式(1-2)、(2)に基づいて元の6ch(Lf、C、Rf、Ls、Rs、Lfe)を復元するとともに、第1グループチャンネル「1」、「2」を加算、減算することによりそれぞれステレオ2chデータ(L、R)を生成する。

10

#### 【0015】

図3は第3の例の伝送形態として、マルチチャンネルを圧縮しないで伝送するとともに再生側のダウンミクスを禁止する場合を示している。この場合には、非圧縮であるので、符号化側では相関信号も生成することなく元の6ch(Lf、C、Rf、Ls、Rs、Lfe)のPCMデータをそのまま伝送し(ただし、フォーマット化する)、復号化側ではデフォーマット化した後、元の6ch(Lf、C、Rf、Ls、Rs、Lfe)のみを復元する。

20

#### 【0016】

図4は第4の例の伝送形態として、マルチチャンネルを圧縮しないで伝送するとともに再生側のダウンミクスを許可する場合を示している。この場合にも、非圧縮であるので、符号化側では圧縮率を高めるための相関信号も生成することなく元の6ch(Lf、C、Rf、Ls、Rs、Lfe)のPCMデータをそのまま伝送する(ただし、フォーマット化する)。復号化側ではデフォーマット化した後、元の6ch(Lf、C、Rf、Ls、Rs、Lfe)を復元するとともに、式(2)によりステレオ2chデータ(L、R)を生成(ダウンミクス)する。

30

#### 【0017】

図5は図1においてマルチチャンネルを圧縮するとともに再生側のダウンミクスを禁止する場合の変形例を示している。この場合には、符号化側では次式(1-3)により6ch(1)～(6)分の相関信号に変換し、符号化部2'はこれを予測符号化する。そして、復号化側では式(1-2)により元の6ch(Lf、C、Rf、Ls、Rs、Lfe)のみを復元する。

$$「1」 = Lf - C$$

$$「2」 = Rf - C$$

$$「3」 \sim 「6」 \text{は式} (1 - 1) \text{と同じ} \quad \dots (1 - 3)$$

このように再生側のダウンミクスを禁止する場合は、これに対応して式(2)のダウンミクス係数を符号化に加えないとともに、符号化側で式(2)によりステレオ2chデータ(L、R)を生成(ダウンミクス)することが禁じられる。

40

#### 【0018】

図6は図2においてマルチチャンネルを圧縮するとともに再生側のダウンミクスを許可する場合の変形例を示している。この場合には、符号化側では式(2)によりステレオ2chデータ(L、R)を生成(ダウンミクス)し、次いで次式(1-4)により次のような第1グループの2チャンネル「1」、「2」と第2グループの4チャンネル分の相関信号「3」～「6」に変換し、第1、第2符号化部2'-1、2'-2はこの各グループチャンネルを予測符号化する。そして、復号化側では式(1-4)、(2)により元の6ch(Lf、C、Rf、Ls、Rs、Lfe)を復元するとともにステレオ2chデータ(L、R)をそのまま出力する。

50

「 1 」 = L

「 2 」 = R

「 3 」 ~ 「 6 」 は式 ( 1 - 1 ) と同じ

... ( 1 - 4 )

#### 【 0 0 1 9 】

図 7 を参照して符号化部 2 ' - 1、2 ' - 2 について詳しく説明する。各 ch 「 1 」 ~ 「 6 」 の P C M データは 1 フレーム毎に 1 フレームバッファ 1 0 に格納される。そして、1 フレームの各 ch 「 1 」 ~ 「 6 」 のサンプルデータがそれぞれ予測回路 1 3 D 1、1 3 D 2、1 5 D 1 ~ 1 5 D 4 に印加されるとともに、各 ch 「 1 」 ~ 「 6 」 の各フレームの先頭サンプルデータがフォーマット化回路 1 9 に印加される。予測回路 1 3 D 1、1 3 D 2、1 5 D 1 ~ 1 5 D 4 はそれぞれ、各 ch 「 1 」 ~ 「 6 」 の P C M データに対して、特性が異なる複数の予測器 ( 不図示 ) により時間領域における過去の信号から現在の信号の複数の線形予測値を算出し、次いで原 P C M データと、この複数の線形予測値から予測器毎の予測残差を算出する。続くバッファ・選択器 1 4 D 1、1 4 D 2、1 6 D 1 ~ 1 6 D 4 はそれぞれ、予測回路 1 3 D 1、1 3 D 2、1 5 D 1 ~ 1 5 D 4 により算出された各予測残差を一時記憶して、選択信号 / D T S ( デコーディング・タイム・スタンプ ) 生成器 1 7 により指定されたサブフレーム毎に予測残差の最小値を選択する。

#### 【 0 0 2 0 】

選択信号 / D T S 生成器 1 7 は予測残差のビット数フラグをパッキング回路 1 8 とフォーマット化回路 1 9 に対して印加し、また、予測残差が最小の予測器を示す予測器選択フラグと、相関係数 a と、復号化側が入力バッファ 2 2 a ( 図 1 4 ) からストリームデータを取り出す時間を示す D T S をフォーマット化回路 1 9 に対して印加する。パッキング回路 1 8 はバッファ・選択器 1 4 D 1、1 4 D 2、1 6 D 1 ~ 1 6 D 4 により選択された 6 ch 分の予測残差を、選択信号 / D T S 生成器 1 7 により指定されたビット数フラグに基づいて指定ビット数でパッキングする。また P T S 生成器 1 7 c は、復号化側が出力バッファ 1 1 0 ( 図 1 4 ) から P C M データを取り出す時間を示す P T S ( プレゼンテーション・タイム・スタンプ ) を生成してフォーマット化回路 1 9 に出力する。フォーマット化回路 1 9 にはまた、圧縮 / 非圧縮などを示す符号化モードと、ダウンミクス許可 / 禁止を示す識別子が印加される。

#### 【 0 0 2 1 】

続くフォーマット化回路 1 9 は図 8 ~ 図 1 3 に示すようなユーザデータにフォーマット化する。図 8 に示すユーザデータ ( サブパケット ) は、前方グループに関する 2 ch 「 1 」、 「 2 」 の予測符号化データを含む可変レートビットストリーム ( サブストリーム ) B S 0 と、他のグループに関する 4 ch 「 3 」 ~ 「 6 」 の予測符号化データを含む可変レートビットストリーム ( サブストリーム ) B S 1 と、サブストリーム B S 0、B S 1 の前に設けられたビットストリームヘッダ ( リスタートヘッダ ) により構成されている。

#### 【 0 0 2 2 】

また、サブストリーム B S 0、B S 1 の 1 フレーム分は

- ・ フレームヘッダと、
- ・ 各 ch 「 1 」 ~ 「 6 」 の 1 フレームの先頭サンプルデータと、
- ・ 各 ch 「 1 」 ~ 「 6 」 のサブフレーム毎の予測器選択フラグと、
- ・ 各 ch 「 1 」 ~ 「 6 」 のサブフレーム毎のビット数フラグと、
- ・ 各 ch 「 1 」 ~ 「 6 」 の予測残差データ列 ( 可変ビット数 ) と、
- ・ ch 「 6 」 の係数 a とが、

多重化されている。このような予測符号化によれば、原信号が例えばサンプリング周波数 = 9 6 k H z、量子化ビット数 = 2 4 ビット、6 チャンネルの場合、7 1 % の圧縮率を実現することができる。

#### 【 0 0 2 3 】

図 7 に示す符号化部 2' - 1、2' - 2 により予測符号化された可変レートビットストリームデータを、記録媒体の一例として DVD オーディオディスクに記録する場合には、図 9 に示すオーディオ (A) パックにパッキングされる。このパックは 2034 バイトのユーザデータ (A パケット、V パケット) に対して 4 バイトのパックスタート情報と、6 バイトの SCR (System Clock Reference : システム時刻基準参照値) 情報と、3 バイトの Mux レート (rate) 情報と 1 バイトのスタッフィングの合計 14 バイトのパックヘッダが付加されて構成されている (1 パック = 合計 2048 バイト)。この場合、タイムスタンプである SCR 情報を、先頭パックでは「1」として同一タイトル内で連続とすることにより同一タイトル内の A パックの時間を管理することができる。

10

#### 【0024】

圧縮 PCM の A パケットは図 10 に詳しく示すように、19 又は 14 バイトのパケットヘッダと、圧縮 PCM のプライベートヘッダと、図 11 に示すフォーマットの 1 ないし 2011 バイトのオーディオデータ (圧縮 PCM) により構成されている。そして、DTS と PTS は図 5 のパケットヘッダ内に (具体的にはパケットヘッダの 10 ~ 14 バイト目に PTS が、15 ~ 19 バイト目に DTS が) セットされる。圧縮 PCM のプライベートヘッダは、

- ・ 1 バイトのサブストリーム ID と、
  - ・ 2 バイトの UPC / EAN - ISRC (Universal Product Code/European Article Number-International Standard Recording Code) 番号、及び UPC / EAN - ISRC データと、
  - ・ 1 バイトのプライベートヘッダ長と、
  - ・ 2 バイトの第 1 アクセスユニットポイントと、
  - ・ 8 バイトのオーディオデータ情報 (ADI) と、
  - ・ 0 ~ 7 バイトのスタッフィングバイトとに、
- より構成されている。

20

#### 【0025】

また、ADI 内に 1 秒後のアクセスユニットをサーチするための前方アクセスユニット・サーチポイントと、1 秒前のアクセスユニットをサーチするための後方アクセスユニット・サーチポイントがともに 1 バイトでセットされる。具体的には ADI の 7 バイト目に前方アクセスユニット・サーチポイントが、8 バイト目に後方アクセスユニット・サーチポイントがセットされる。

30

#### 【0026】

図 10 に示す圧縮 PCM (PPCM ともいう) のオーディオパケットにおけるオーディオデータエリアは、図 11 に示すようにサブパケットと複数の PPCM アクセスユニットにより構成され、PPCM アクセスユニットは PPCM シンク情報とサブパケットにより構成されている。最初の PPCM アクセスユニット内のサブパケットは、ディレクトリと、サブストリーム「0」と、CRC と、サブストリーム「1」と、CRC とエクストラ情報により構成され、サブストリーム「0」、「1」は PPCM ブロックのみにより構成されている。2 番目以降の PPCM アクセスユニット内のサブパケットは、ディレクトリを除いてサブストリーム「0」と、CRC と、サブストリーム「1」と、CRC とエクストラ情報により構成され、サブストリーム「0」、「1」はリスタートヘッダと PPCM ブロックにより構成されている。

40

#### 【0027】

PPCM シンク情報 (以下、同期情報ともいう) は次の情報を含む。

- ・ 1 パケット当たりのサンプル数：サンプリング周波数  $f_s$  に応じて 40、80 又は 160 が選択される。
- ・ データレート：VBR の場合には「0」(サブパケット内のデータが圧縮データであることを示す識別子)

50

- ・ サンプリング周波数  $f_s$  及び量子化ビット数  $Q_b$
- ・ チャンネル割り当て情報

## 【 0 0 2 8 】

フォーマット化回路 19 はまた、図 8 ~ 図 11 に示すオーディオパックを管理するために図 12、図 13 に示すような管理情報を含む A T S I (オーディオ・タイトル・セット・インフォメーション) をフォーマット化する。図 12 は A O T T - A O B - A T R (オーディオオンリタイトル・オーディオオブジェクトセット・アトリビュート) を示し、この A O T T - A O B - A T R ( b 1 2 7 ~ b 0 ) は、M S B 側から順に

- ・ 8 ビット ( b 1 2 7 ~ b 1 2 0 ) のオーディオ符号化モードと、
- ・ 8 ビット ( b 1 1 9 ~ b 1 1 2 ) の保留領域と、
- ・ 4 ビット ( b 1 1 1 ~ b 1 0 8 ) のチャンネルグループ「1」の量子化ビット数  $Q_1$  と、
- ・ 4 ビット ( b 1 0 7 ~ b 1 0 4 ) のチャンネルグループ「2」の量子化ビット数  $Q_2$  と、
- ・ 4 ビット ( b 1 0 3 ~ b 1 0 0 ) のチャンネルグループ「1」のサンプリング周波数  $f_{s1}$  と、
- ・ 4 ビット ( b 9 9 ~ b 9 6 ) のチャンネルグループ「2」のサンプリング周波数  $f_{s2}$  と、
- ・ 3 ビット ( b 9 5 ~ b 9 3 ) のマルチチャンネル構造のタイプと、
- ・ 5 ビット ( b 9 2 ~ b 8 8 ) のチャンネル割り当てと、
- ・ 8 ビット  $\times$  11 ( b 8 7 ~ b 0 ) の保留領域により構成されている。

## 【 0 0 2 9 】

上記データを以下に詳しく示す。

- ( 1 ) オーディオ符号化モード ( b 1 2 7 ~ b 1 2 0 )

0 0 0 0 0 0 0 0 b : リニア P C M モード

0 0 0 0 0 0 0 1 b : 圧縮 P C M モード

その他 : その他の符号化モード用に保留

## 【 0 0 3 0 】

- ( 2 ) チャンネルグループ 1 の量子化ビット数  $Q_1$  ( b 1 1 1 ~ b 1 0 8 )

0 0 0 0 b : 16 ビット

0 0 0 1 b : 20 ビット

0 0 1 0 b : 24 ビット

その他 : 保留

- ( 3 ) チャンネルグループ 2 の量子化ビット数  $Q_2$  ( b 1 0 7 ~ b 1 0 4 )

・ チャンネルグループ 1 の量子化ビット数  $Q_1$  が「 0 0 0 0 b 」の場合には「 0 0 0 0 b 」

・ チャンネルグループ 1 の量子化ビット数  $Q_1$  が「 0 0 0 1 b 」の場合には「 0 0 0 0 b 」又は「 0 0 0 1 b 」

・ チャンネルグループ 1 の量子化ビット数  $Q_1$  が「 0 0 1 0 b 」の場合には「 0 0 0 0 b 」、「 0 0 0 1 b 」又は「 0 0 1 0 b 」

ただし、 0 0 0 0 b : 16 ビット

0 0 0 1 b : 20 ビット

0 0 1 0 b : 24 ビット

その他 : 保留

## 【 0 0 3 1 】

- ( 4 ) チャンネルグループ 1 のサンプリング周波数  $f_{s1}$  ( b 1 0 3 ~ b 1 0 0 )

0 0 0 0 b : 48 k H z

0 0 0 1 b : 96 k H z

0 0 1 0 b : 192 k H z

10

20

30

40

50

1 0 0 0 b : 4 4 . 1 k H z  
 1 0 0 1 b : 8 8 . 2 k H z  
 1 0 1 0 b : 1 7 6 . 4 k H z  
 その他 : 保留

# 【 0 0 3 2 】

( 5 ) チャンネルグループ 2 のサンプリング周波数  $f_{s2}$  ( b 9 9 ~ b 9 6 )

・チャンネルグループ 1 のサンプリング周波数  $f_{s1}$  が「 0 0 0 0 b 」の場合には「 0 0 0 0 b 」

・チャンネルグループ 1 のサンプリング周波数  $f_{s1}$  が「 0 0 0 1 b 」の場合には「 0 0 0 0 b 」又は「 0 0 0 1 b 」

・チャンネルグループ 1 のサンプリング周波数  $f_{s1}$  が「 0 0 1 0 b 」の場合には「 0 0 0 0 b 」、 「 0 0 0 1 b 」又は「 0 0 1 0 b 」

・チャンネルグループ 1 のサンプリング周波数  $f_{s1}$  が「 1 0 0 0 b 」の場合には「 1 0 0 0 b 」

・チャンネルグループ 1 のサンプリング周波数  $f_{s1}$  が「 1 0 0 1 b 」の場合には「 1 0 0 0 b 」又は「 1 0 0 1 b 」

・チャンネルグループ 1 のサンプリング周波数  $f_{s1}$  が「 1 0 1 0 b 」の場合には「 1 0 0 0 b 」、 「 1 0 0 1 b 」又は「 1 0 1 0 b 」

# 【 0 0 3 3 】

( 6 ) マルチチャンネル構造のタイプ ( b 9 5 ~ b 9 3 )

0 0 0 b : タイプ 1

その他 : 保留

( 7 ) チャンネル割り当て ( b 9 2 ~ b 8 8 )

1 チャンネル ( モノラル ) から 6 チャンネルまでのグループ「 1 」、 「 2 」のチャンネル割り当て情報

# 【 0 0 3 4 】

図 1 3 は A T S - P G - C N T ( オーディオタイトルセット・プログラム・コンテンツ ) を示し、これは先頭から順に

・ 1 ビット ( b 3 1 ) の、前回と今回の P G の関係 ( R / A ) と、

・ 1 ビット ( b 3 0 ) の S T C 不連続性フラグ ( S T C - F ) と、

・ 3 ビット ( b 2 9 ~ b 2 7 ) のアトリビュート数 ( A T R N ) と、

・ 3 ビット ( b 2 6 ~ b 2 4 ) のチャンネルグループ ( C h G r ) 「 2 」のビットシフトデータと、

・ 2 ビット ( b 2 3 、 b 2 2 ) の保留領域と、

・ 1 ビット ( b 2 1 ) のダウンミックスモード ( D - M ) と、

・ 1 ビット ( b 2 0 ) のダウンミックス係数の有効性 ( 図示 ) と、

・ 4 ビット ( b 1 9 ~ b 1 6 ) のダウンミックス係数テーブル番号 ( D M - C O E F F T N ) と、

・ 各々が 1 ビット、合計 1 6 ビット ( b 1 5 ~ b 0 ) の R T I フラグ F 1 5 ~ F 0 により構成されている。

そして、ビット ( b 2 1 ) のダウンミックスモード ( D - M ) が「 1 」の場合に「ダウンミックス禁止」、 「 0 」の場合に「ダウンミックス許可」を表す。

# 【 0 0 3 5 】

次に図 1 4 を参照して復号化部 3 ' ( 3 ' - 1 、 3 ' - 2 ) について説明する。なお、この復号化部 3 ' ( 3 ' - 1 、 3 ' - 2 ) とミクス & マトリクス回路 4 ' は、ハードウェアの他にコンピュータプログラムによっても実現することができる。上記フォーマットの可変レートビットストリームデータ B S 0 、 B S 1 は、デフォーマット化回路 2 1 により分離される。そして、各 c h 「 1 」 ~ 「 6 」の 1 フレームの先頭サンプルデータと予測器選択フラグはそれぞれ予測回路 2 4 D 1 、 2 4 D 2 、 2 3 D 1 ~ 2 3 D 4 に印加され、各 c h 「 1 」 ~ 「 6 」のビット

数フラグはアンパッキング回路 2 2 に印加される。また、S C R と、D T S と予測残差データ列は入力バッファ 2 2 a に印加され、P T S は出力バッファ 1 1 0 に印加される。また、圧縮 / 非圧縮などを示す符号化モードと、ダウンミクス許可 / 禁止を示す識別子は制御部 1 0 0 に印加され、サンプリング周波数  $f_s$  及び量子化ビット数  $Q_b$  は D / A 変換器 1 0 2 に印加される。ここで、予測回路 2 4 D 1、2 4 D 2、2 3 D 1 ~ 2 3 D 4 内の複数の予測器 (不図示) はそれぞれ、符号化側の予測回路 1 3 D 1、1 3 D 2、1 5 D 1 ~ 1 5 D 4 内の複数の予測器と同一の特性であり、予測器選択フラグにより同一特性のものが選択される。

【 0 0 3 6 】

デフォーマット化回路 2 1 により分離されたストリームデータ (予測残差データ列) は、図 1 5 に示すように S C R によりアクセスユニット毎に入力バッファ 2 2 a に取り込まれて蓄積される。ここで、1 つのアクセスユニットのデータ量は、例えば  $f_s = 96 \text{ kHz}$  の場合には  $(1 / 96 \text{ kHz})$  秒分であるが、図 1 6、図 1 7 ( a ) に詳しく示すように可変長である。そして、入力バッファ 2 2 a に蓄積されたストリームデータは D T S に基づいて F I F O で読み出されてアンパッキング回路 2 2 に印加される。

【 0 0 3 7 】

アンパッキング回路 2 2 は各 c h 「 1 」 ~ 「 6 」 の予測残差データ列をビット数フラグ毎に基づいて分離してそれぞれ予測回路 2 4 D 1、2 4 D 2、2 3 D 1 ~ 2 3 D 4 に出力する。予測回路 2 4 D 1、2 4 D 2、2 3 D 1 ~ 2 3 D 4 ではそれぞれ、アンパッキング回路 2 2 からの各 c h 「 1 」 ~ 「 6 」 の今回の予測残差データと、内部の複数の予測器の内、予測器選択フラグにより選択された各 1 つにより予測された前回の予測値が加算されて今回の予測値が算出され、次いで 1 フレームの先頭サンプルデータを基準として各サンプルの P C M データが算出されて出力バッファ 1 1 0 に蓄積される。出力バッファ 1 1 0 に蓄積された P C M データは P T S に基づいて読み出されて出力され、したがって、図 1 7 ( a ) に示す可変長のアクセスユニットが伸長されて、図 1 7 ( b ) に示す一定長のプレゼンテーションユニットが出力される。

【 0 0 3 8 】

また、P P C M シンク情報内のサンプリング周波数  $f_s$  及び量子化ビット数  $Q_b$  に基づいて、P C M データが D / A 変換器 1 0 2 によりアナログ信号に変換される。ここで、操作部 1 0 1 を介してサーチ再生が指示された場合には、制御部 1 0 0 により図 5 に示す前方アクセスユニット・サーチポインタ ( 1 秒先 ) と後方アクセスユニット・サーチポインタ ( 1 秒前 ) に基づいてアクセスユニットを再生する。このサーチポインタとしては、1 秒先、1 秒前の代わりに 2 秒先、2 秒前のものでよい。

【 0 0 3 9 】

符号化部 2 ' ( 2 ' - 1、2 ' - 2 ) により予測符号化された可変レートビットストリームデータをネットワークを介して伝送する場合には、符号化側では図 1 8 に示すように伝送用にパケット化し (ステップ S 4 1)、次いでパケットヘッダを付与し (ステップ S 4 2)、次いでこのパケットをネットワーク上に送り出す (ステップ S 4 3)。

【 0 0 4 0 】

復号側では図 1 9 ( A ) に示すようにヘッダを除去し (ステップ S 5 1)、次いでデータを復元し (ステップ S 5 2)、次いでこのデータをメモリに格納して復号を待つ (ステップ S 5 3)。そして、復号を行う場合には図 1 9 ( B ) に示すように、デフォーマット化を行い (ステップ S 6 1)、次いで入力バッファ 2 2 a の入出力制御を行い (ステップ S 6 2)、次いでアンパッキングを行う (ステップ S 6 3)。なお、このとき、サーチ再生指示がある場合にはサーチポインタをデコードする。次いで予測器をフラグに基づいて選択してデコードを行い (

10

20

30

40

50

ステップ S 6 4 )、次いで出力バッファ 1 1 0 の入出力制御を行い ( ステップ S 6 5 )、次いで元のマルチチャネルを復元し ( ステップ S 6 6 )、次いでこれを出力し ( ステップ S 6 7 )、以下、これを繰り返す。

#### 【 0 0 4 1 】

次に図 2 0、図 2 1 を参照して第 2 の実施形態について説明する。上記の実施形態では、1 グループの相関性の信号「1」～「6」を予測符号化するように構成されているが、この第 4 の実施形態では複数グループの相関性のある信号を生成して予測符号化し、圧縮率が最も高いグループの予測符号化データを選択するように構成されている。このため図 2 0 に示す符号化部では、第 1 ～ 第 n の相関回路 1 - 1 ～ 1 - n が設けられ、この n 個の相関回路 1 - 1 ～ 1 - n は例えば 6 ch ( L f、C、R f、L s、R s、L f e ) の P C M データを、相関性が異なる n 種類の 6 ch 信号「1」～「6」に変換する。

10

#### 【 0 0 4 2 】

例えば第 1 の相関回路 1 - 1 は以下のように変換し、

$$( 1 ) = L f$$

$$( 2 ) = C - ( L s + R s ) / 2$$

$$( 3 ) = R f - L f$$

$$( 4 ) = L s - a \times L f e$$

$$( 5 ) = R s - b \times R f$$

$$( 6 ) = L f e$$

20

また、第 n の相関回路 1 - n は以下のように変換する。

$$( 1 ) = L f + R f$$

$$( 2 ) = C - L f$$

$$( 3 ) = R f - L f$$

$$( 4 ) = L s - L f$$

$$( 5 ) = R s - L f$$

$$( 6 ) = L f e - C$$

#### 【 0 0 4 3 】

また、相関回路 1 - 1 ～ 1 - n 毎に予測回路 1 5 とバッファ・選択器 1 6 が設けられ、グループ毎の予測残差の最小値のデータ量に基づいて圧縮率が最も高いグループが相関選択信号生成器 1 7 b により選択される。このとき、フォーマット化回路 1 9 はその選択フラグ ( 相関回路選択フラグ、その相関回路の相関係数 a、b ) を追加して多重化する。

30

#### 【 0 0 4 4 】

また、図 2 1 に示す復号化側では、符号化側の相関回路 1 - 1 ～ 1 - n に対して n 個の相関回路 4 - 1 ～ 4 - n ( 又は係数 a、b が変更可能な 1 つの相関回路 4 ) が設けられる。なお、図 2 0 に示す n グループの予測回路が同一の構成である場合、復号装置では図 2 1 に示すように n グループ分の予測回路を設ける必要はなく、1 つのグループ分の予測回路でよい。そして、符号化装置から伝送された選択フラグに基づいて相関回路 4 - 1 ～ 4 - n の 1 つを選択、又は係数 a、b を設定して元の 6 ch ( L f、C、R f、L s、R s、L f e ) を復元し、また、式 ( 2 ) によりマルチチャネルをダウンミクスしてステレオ 2 ch データ ( L、R ) を生成する。

40

#### 【 0 0 4 5 】

また、上記の第 1 の実施形態では、1 種類の相関性の信号「1」～「6」を予測符号化するように構成されているが、この信号「1」～「6」のグループと原信号 ( L f、C、R f、L s、R s、L f e ) のグループを予測符号化し、圧縮率が高い方のグループを選択するようにしてもよい。

本発明によれば、特許請求の範囲に記載した発明の他に、次のような発明が提供される。

50

マルチチャネルの音声信号が圧縮されたデータ又は圧縮されないデータを選択的にオーディオパケットに配置するフォーマット化手段と、

前記オーディオパケット内のマルチチャネルデータが圧縮されているか否か、あるいは、前記オーディオパケット内のマルチチャネルデータをステレオ2チャネルにダウンミクスすることを許可するか又は禁止するかによってあらかじめダウンミクスして符号化するか否か、あるいはダウンミクス係数を符号化するか否かを選択する手段とを、

有する音声符号化装置。

【図面の簡単な説明】

【0046】

10

【図1】本発明が適用されるマルチチャネルの伝送形態の第1の例を示す説明図である。

【図2】本発明が適用されるマルチチャネルの伝送形態の第2の例を示す説明図である。

【図3】本発明が適用されるマルチチャネルの伝送形態の第3の例を示す説明図である。

【図4】本発明が適用されるマルチチャネルの伝送形態の第4の例を示す説明図である。

【図5】図1の変形例を示す説明図である。

【図6】図2の変形例を示す説明図である。

【図7】図1の符号化部を詳しく示すブロック図である。

【図8】図1、図7の符号化部により符号化されたビットストリームを示す説明図である。

。

【図9】DVDのパックのフォーマットを示す説明図である。

20

【図10】DVDのオーディオパックのフォーマットを示す説明図である。

【図11】図10のオーディオデータエリアのフォーマットを詳しく示す説明図である。

【図12】DVDオーディオのAOTT-AOB-ATR（オーディオオンリタイトル・オーディオオブジェクトセット・アトリビュート）を示す説明図である。

【図13】DVDオーディオのATS-PG-CNT（オーディオタイトルセット・プログラム・コンテンツ）を示す説明図である。

【図14】図1の復号化部を詳しく示すブロック図である。

【図15】図14の入力バッファの書き込み／読み出しタイミングを示すタイミングチャートである。

【図16】アクセスユニット毎の圧縮データ量を示す説明図である。

30

【図17】アクセスユニットとプレゼンテーションユニットを示す説明図である。

【図18】音声伝送方法を示すフローチャートである。

【図19】音声伝送方法を示すフローチャートである。

【図20】第2の実施形態の音声符号化装置を示すブロック図である。

【図21】第2の実施形態の音声復号装置を示すブロック図である。

【符号の説明】

【0047】

1' 6chミクス&マトリクス回路

13D1, 13D2, 15D1~15D4 予測回路（バッファ・選択器14D1, 14D2, 16D1~16D4と共に圧縮手段を構成する。）

40

14D1, 14D2, 16D1~16D4 バッファ・選択器

17 選択信号／DTS生成器

17c PTS生成器

19 フォーマット化回路

21 デフォーマット化回路（分離手段）

22 アンパッキング回路

22a 入力バッファ

24D1, 24D2, 23D1~23D4 予測回路（伸長手段）

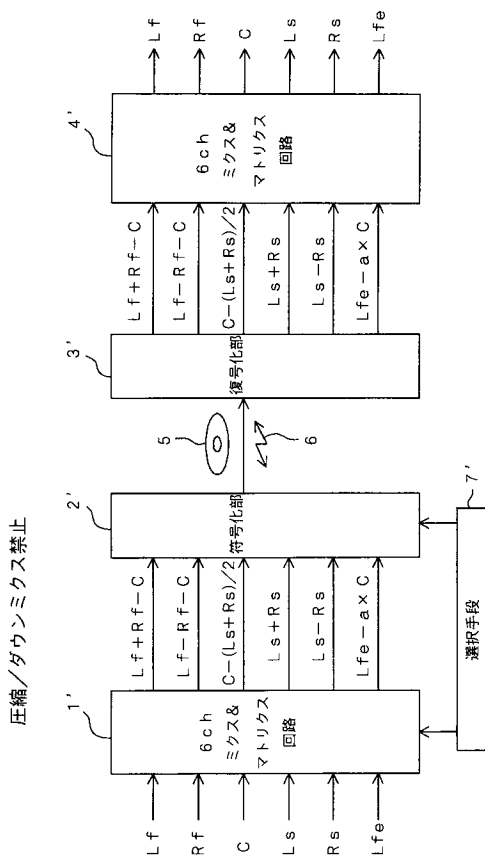
100 制御部（再生手段）

102 D/A変換器

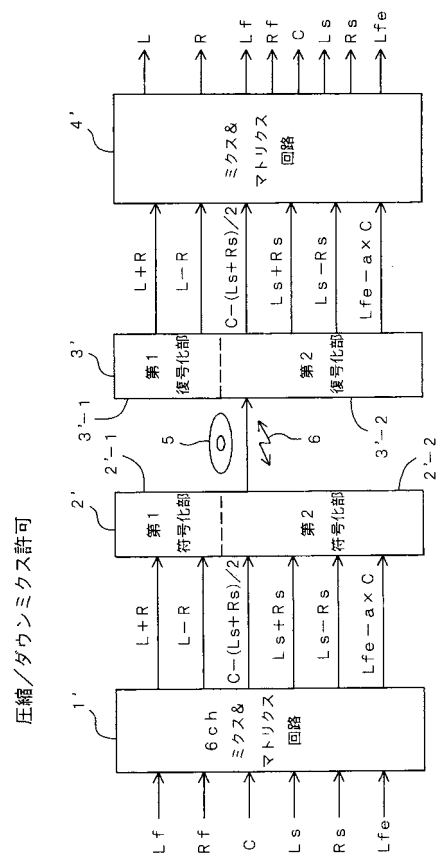
50

1 1 0 出力バッファ

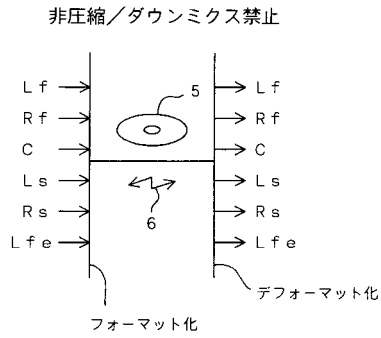
【図 1】



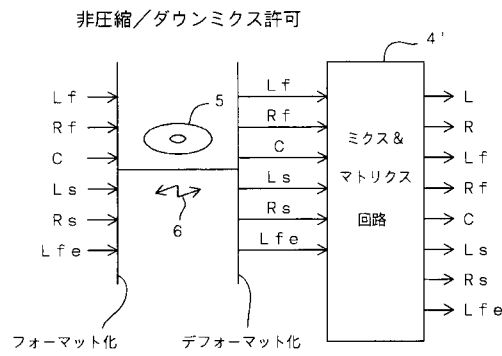
【図 2】



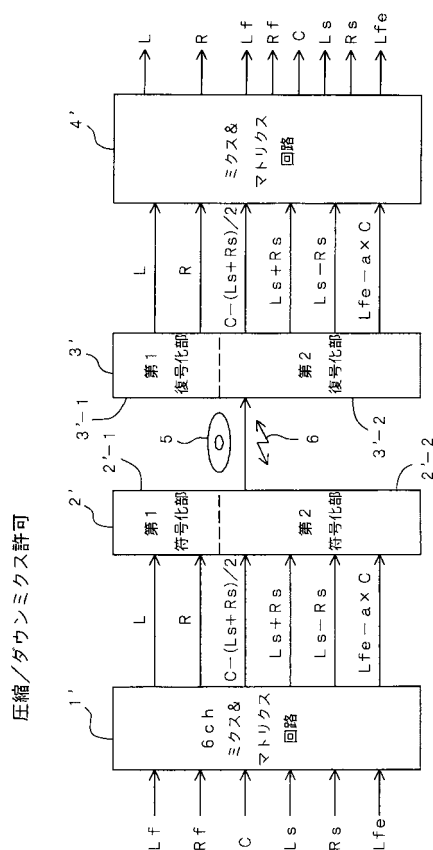
【図 3】



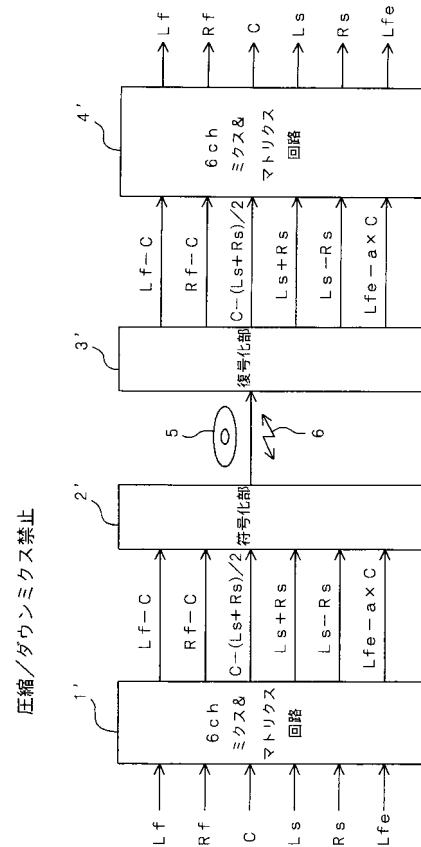
【図 4】



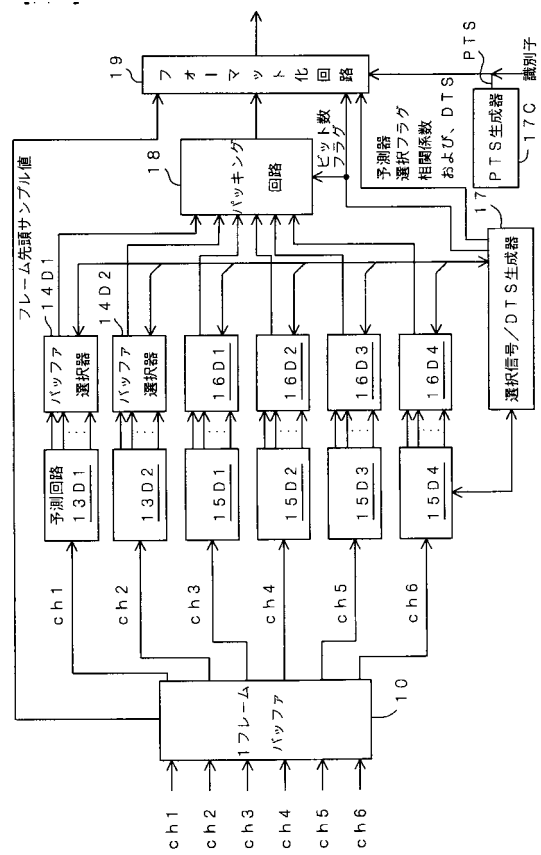
【図 6】



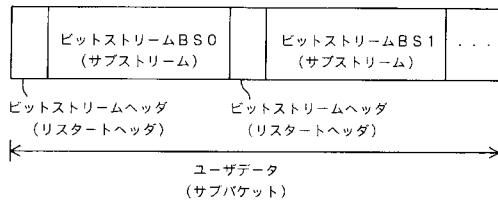
【図 5】



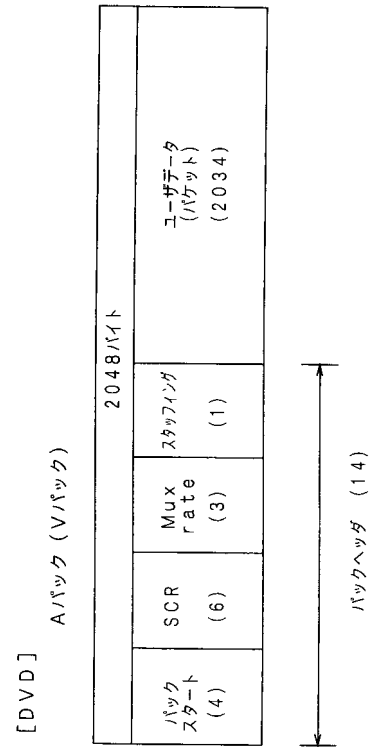
【図 7】



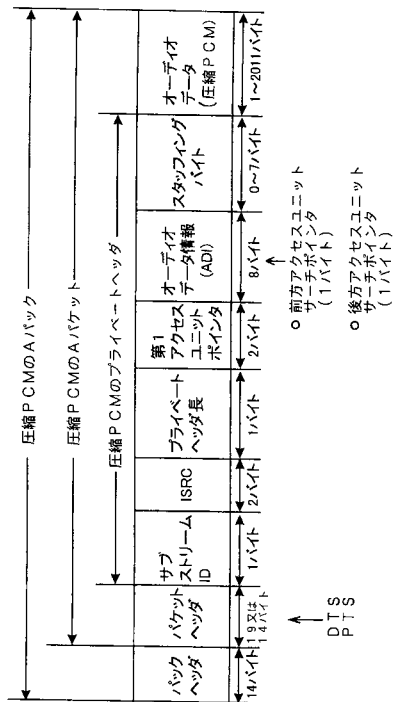
【図 8】



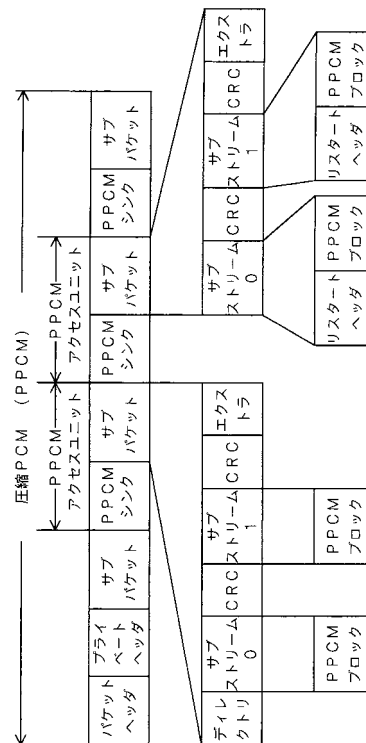
【図 9】



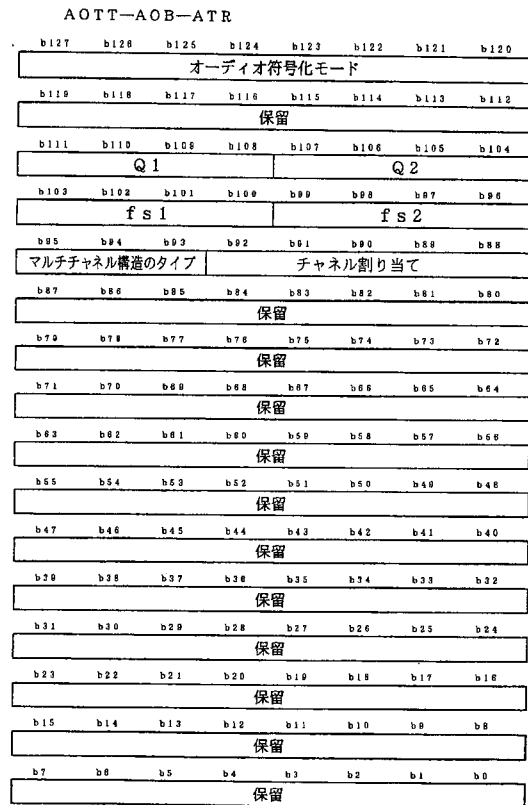
【図 10】



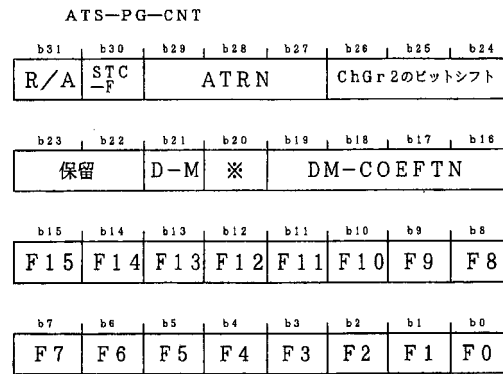
【図 11】



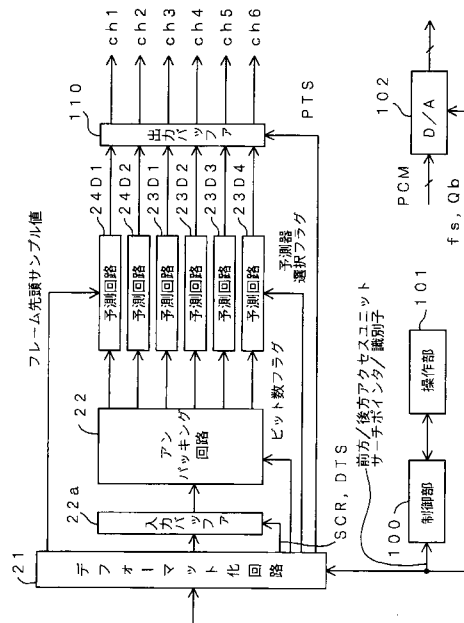
【図 12】



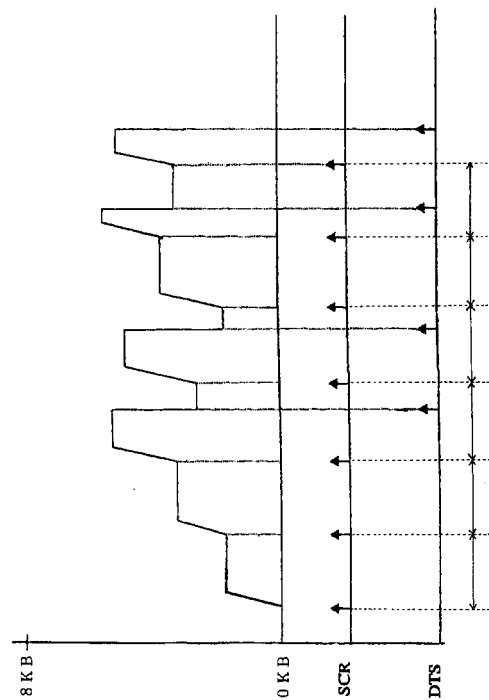
【図 13】



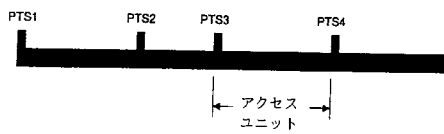
【図 14】



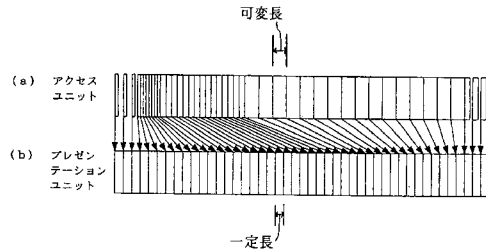
【図 15】



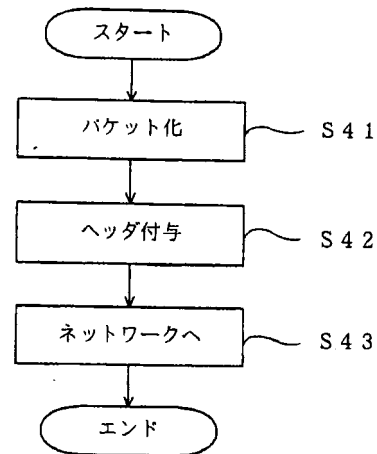
【図 16】



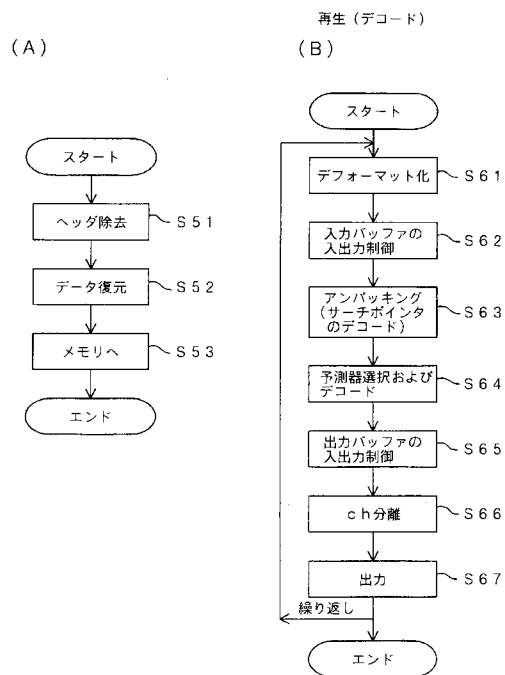
【図 17】



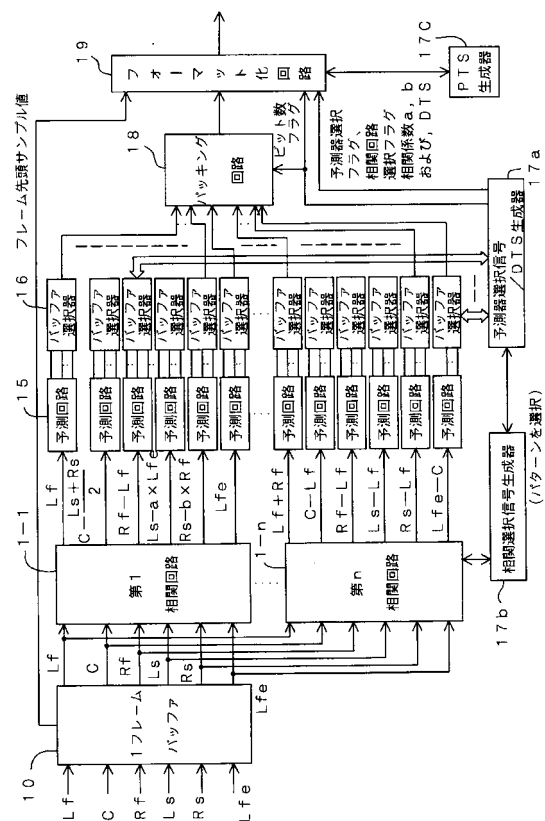
【図 18】



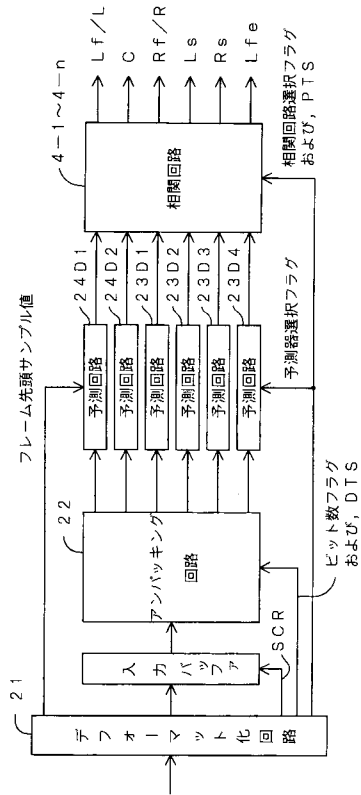
【図 19】



【図 20】



【図 21】



---

フロントページの続き

(58)調査した分野(Int.Cl. , D B 名)

|         |                       |
|---------|-----------------------|
| G 1 0 L | 1 9 / 0 0 - 1 9 / 1 4 |
| G 1 1 B | 2 0 / 1 0 - 2 0 / 2 0 |
| H 0 3 M | 7 / 3 0 - 7 / 3 8     |
| H 0 4 S | 3 / 0 0 - 5 / 0 2     |