

(19) 日本国特許庁(JP)

## (12) 特許公報(B2)

(11) 特許番号

特許第4630805号  
(P4630805)

(45) 発行日 平成23年2月9日(2011.2.9)

(24) 登録日 平成22年11月19日(2010.11.19)

(51) Int.Cl.	F 1
HO4N 5/91 (2006.01)	HO4N 5/91 Z
HO4N 5/92 (2006.01)	HO4N 5/92 H
HO4N 7/26 (2006.01)	HO4N 7/13 Z

請求項の数 6 (全 24 頁)

(21) 出願番号 特願2005-355308 (P2005-355308)  
 (22) 出願日 平成17年12月8日 (2005.12.8)  
 (65) 公開番号 特開2007-159058 (P2007-159058A)  
 (43) 公開日 平成19年6月21日 (2007.6.21)  
 審査請求日 平成20年12月1日 (2008.12.1)

(73) 特許権者 000001007  
 キヤノン株式会社  
 東京都大田区下丸子3丁目30番2号  
 (74) 代理人 100076428  
 弁理士 大塚 康徳  
 (74) 代理人 100112508  
 弁理士 高柳 司郎  
 (74) 代理人 100115071  
 弁理士 大塚 康弘  
 (74) 代理人 100116894  
 弁理士 木村 秀二  
 (72) 発明者 藤井 昭雄  
 東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

最終頁に続く

(54) 【発明の名称】記録装置及び記録方法

## (57) 【特許請求の範囲】

## 【請求項 1】

入力された動画像データを、フレーム内符号化とフレーム間予測符号化とを用いて符号化し、複数の動画符号化単位からなる符号化された動画像データを出力する動画符号化手段と、

入力された音声データを符号化し、複数の音声符号化単位からなる符号化された音声データを出力する音声符号化手段と、

所定数の前記動画符号化単位の最大符号量と、所定数の前記音声符号化単位の最大符号量とを設定する設定手段と、

前記所定数の動画符号化単位の符号量が前記設定手段により設定された動画符号化単位の最大符号量を超えないように前記動画像データを符号化するよう前記動画符号化手段を制御する動画符号量制御手段と、

前記所定数の音声符号化単位の符号量が前記設定手段により設定された音声符号化単位の最大符号量を超えないように前記音声データを符号化するよう前記音声符号化手段を制御する音声符号量制御手段と、

それぞれが前記所定数の動画符号化単位の符号化された動画像データを含む複数の動画チャンクと、それぞれが前記所定数の音声符号化単位の符号化された音声データを含む複数の音声チャンクとを多重化して記録媒体に記録すると共に複数の前記動画チャンク及び音声チャンクの各々にアクセスするためのオフセット値の情報を前記記録媒体に記録する記録手段とを備え、

10

20

前記記録手段は、前記動画チャンクと前記音声チャンクを交互に多重化し、一つの前記動画チャンクと一つの前記音声チャンクの合計のサイズが前記動画符号化単位の最大符号量と前記音声符号化単位の最大符号量の合計に等しくなるように前記記録媒体に記録すると共に、複数の前記動画チャンクと音声チャンク及び前記オフセット値の情報とは別に、前記合計のサイズの情報を前記記録媒体に記録することを特徴とする記録装置。

**【請求項 2】**

前記記録手段は、更に、前記所定数の動画符号化単位の最大符号量の情報と、前記所定数の音声符号化単位の最大符号量の情報を、複数の前記動画チャンクと音声チャンク及び前記オフセット値の情報とは別に前記記録媒体に記録することを特徴とする請求項 1 に記載の記録装置。

10

**【請求項 3】**

前記記録手段は、更に、前記動画チャンクに含まれる動画像データのフレーム数の情報と、前記音声チャンクに含まれる音声データのフレーム数の情報を、複数の前記動画チャンクと音声チャンク及び前記オフセット値の情報とは別に前記記録媒体に記録することを特徴とする請求項 1 または 2 に記載の記録装置。

**【請求項 4】**

前記記録媒体から複数の前記動画チャンクと音声チャンク及び前記合計のサイズの情報を再生する再生手段を備え、前記再生手段は前記記録媒体から再生された前記合計のサイズの情報に基づいて前記動画チャンクと音声チャンクを特殊再生することを特徴とする請求項 1 から 3 の何れか 1 項に記載の記録装置。

20

**【請求項 5】**

前記記録手段は、複数の前記動画チャンクと音声チャンク、前記オフセット値の情報及び前記合計のサイズの情報を一つの動画ファイルの中に記録することを特徴とする請求項 1 から 4 のいずれか 1 項に記載の記録装置。

**【請求項 6】**

動画像データおよび音声データを符号化して記録する記録装置における記録方法であつて、

動画符号化手段が、入力された動画像データを、フレーム内符号化とフレーム間予測符号化とを用いて符号化し、複数の動画符号化単位からなる符号化された動画像データを出力する動画符号化工程と、

30

音声符号化手段が、入力された音声データを符号化し、複数の音声符号化単位からなる符号化された音声データを出力する音声符号化工程と、

設定手段が、所定数の前記動画符号化単位の最大符号量と、所定数の前記音声符号化単位の最大符号量とを設定する設定工程と、

動画符号量制御手段が、前記所定数の動画符号化単位の符号量が前記設定工程において設定された動画符号化単位の最大符号量を超えないように前記動画像データを符号化するよう前記動画符号化手段を制御する動画符号量制御工程と、

音声符号量制御手段が、前記所定数の音声符号化単位の符号量が前記設定工程において設定された音声符号化単位の最大符号量を超えないように前記音声データを符号化するよう前記音声符号化手段を制御する音声符号量制御工程と、

40

記録手段が、それぞれが前記所定数の動画符号化単位の符号化された動画像データを含む複数の動画チャンクと、それぞれが前記所定数の音声符号化単位の符号化された音声データを含む複数の音声チャンクとを多重化して記録媒体に記録すると共に複数の前記動画チャンク及び音声チャンクの各々にアクセスするためのオフセット値の情報を前記記録媒体に記録する記録工程とを有し、

前記記録工程では、前記動画チャンクと前記音声チャンクを交互に多重化し、一つの前記動画チャンクと一つの前記音声チャンクの合計のサイズが前記動画符号化単位の最大符号量と前記音声符号化単位の最大符号量の合計に等しくなるように前記記録媒体に記録すると共に、複数の前記動画チャンクと音声チャンク及び前記オフセット値の情報とは別に、前記合計のサイズの情報を前記記録媒体に記録することを特徴とする記録方法。

50

**【発明の詳細な説明】****【技術分野】****【0001】**

本発明は、動画像データや音声データを符号化して記録する技術に関するものである。

**【背景技術】****【0002】**

近年、デジタルビデオカメラやデジタルカメラ等、動画像データをMPEG等のフレーム間予測符号化を用いて符号化して記録再生する装置が普及している。

**【0003】**

また、デジタルビデオカメラやデジタルカメラ等で撮影されたMPEG (MPEG2 / MPEG4) の動画像データやオーディオデータはMP4という汎用ファイルフォーマットで記録されている。これらのデータをMP4ファイルとして記録することで、他の機器で再生できる互換性が保証される。尚、MP4は、MPEG audio Layer-4の略称である。

**【0004】**

MP4ファイルは、符号化された画像データ及びオーディオデータのストリームが入っているmdat boxと各ストリームに関する情報が入っているmoov boxとからなる。mdat boxは更にチャンク(chunk)で構成され、例えば、動画像のチャンクとオーディオのチャンクとから構成される。

**【0005】**

moov boxには、ファイルの各先頭から各チャンクへのオフセットバイト数や、サイズ、サンプル数等が格納される。

**【0006】**

MP4ファイルを再生する場合、記録媒体からMP4ファイルのmoov boxを読み込み、そのmoov boxに基づいて動画像データのチャンク、オーディオデータのチャンクへのアクセスができるようになる。

**【0007】**

しかし、moov boxのデータはチャンク数に比例してそのサイズが増大する。つまり、記録時間が長くなればそれに比例してサイズが増大するものであり、ストリームへのアクセス情報であるmoov boxデータを全てデジタルビデオカメラやデジタルカメラ等の携帯型小型撮影機器に読み込むためには大きなメモリを必要とする。

**【0008】**

このような問題に対処するために、moov boxのデータとは別にストリームにアクセスする情報を選別して別データとし、MP4ファイルにリンクする情報と共に別ファイルとして記録しておく技術が提案されている。この技術によれば、再生時にその別ファイルからMP4のストリームデータへアクセスすることで、再生時にストリームへのアクセス情報を展開するメモリを削減している(例えば、特許文献1参照)。

**【0009】**

ところが、上記のようにmdat box内のストリームデータへのアクセス情報を選別して、アクセス情報の縮小を図っても、記録時間の増大に伴って大きくなることには変わりない。また、アクセスデータを別ファイルとして記録しておくので、MP4ファイルを別の記録媒体にコピーした場合には使用できない。

**【特許文献1】特開2004-128938号公報****【発明の開示】****【発明が解決しようとする課題】****【0010】**

上述したようにMP4ファイルを再生する場合、データストリームへのアクセス情報を格納するためには大きなメモリ容量を必要とするという問題がある。

**【0011】**

また、記録媒体に記録されている1つのMP4ファイルからmoov box内のデータストリームへのアクセス情報を一度に機器内に取り込めない場合には、以下のように対処しなけ

10

20

30

40

50

ればならない。即ち、MP4ファイルデータを再生していくに従い、読み込んだアクセス情報の切れ目が近づくと、最初に読み込んだアクセス情報を破棄し、切れ目に続くアクセス情報をファイルの先頭にシークしてmoov boxから読み込んで取得しなければならない。しかし、記録レートの大きいファイルや、記録時間の長いファイルの後ろのほうを再生している場合や、特殊再生（サーチ）している場合等でリアルタイムにシームレスに再生できなくなるという問題がある。

#### 【0012】

本発明は、上記課題に鑑みてなされ、その目的は、ファイルとして記録された動画像データや音声データの再生時やサーチ時に必要なメモリ容量を削減し、データアクセスを高速化できる技術を実現することである。

10

#### 【課題を解決するための手段】

#### 【0013】

上記課題を解決し、目的を達成するために、本発明の記録装置は、入力された動画像データを、フレーム内符号化とフレーム間予測符号化とを用いて符号化し、複数の動画符号化単位からなる符号化された動画像データを出力する動画符号化手段と、入力された音声データを符号化し、複数の音声符号化単位からなる符号化された音声データを出力する音声符号化手段と、所定数の前記動画符号化単位の最大符号量と、所定数の前記音声符号化単位の最大符号量とを設定する設定手段と、前記所定数の動画符号化単位の符号量が前記設定手段により設定された動画符号化単位の最大符号量を超えないように前記動画像データを符号化するよう前記動画符号化手段を制御する動画符号量制御手段と、前記所定数の音声符号化単位の符号量が前記設定手段により設定された音声符号化単位の最大符号量を超えないように前記音声データを符号化するよう前記音声符号化手段を制御する音声符号量制御手段と、それぞれが前記所定数の動画符号化単位の符号化された動画像データを含む複数の動画チャンクと、それぞれが前記所定数の音声符号化単位の符号化された音声データを含む複数の音声チャンクとを多重化して記録媒体に記録すると共に複数の前記動画チャンク及び音声チャンクの各々にアクセスするためのオフセット値の情報を前記記録媒体に記録する記録手段とを備え、前記記録手段は、前記動画チャンクと前記音声チャンクを交互に多重化し、一つの前記動画チャンクと一つの前記音声チャンクの合計のサイズが前記動画符号化単位の最大符号量と前記音声符号化単位の最大符号量の合計に等しくなるように前記記録媒体に記録すると共に、複数の前記動画チャンクと音声チャンク及び前記オフセット値の情報を別に、前記合計のサイズの情報を前記記録媒体に記録する。

20

#### 【0014】

また、本発明の記録方法は、動画像データおよび音声データを符号化して記録する記録装置における記録方法であって、動画符号化手段が、入力された動画像データを、フレーム内符号化とフレーム間予測符号化とを用いて符号化し、複数の動画符号化単位からなる符号化された動画像データを出力する動画符号化工程と、音声符号化手段が、入力された音声データを符号化し、複数の音声符号化単位からなる符号化された音声データを出力する音声符号化工程と、設定手段が、所定数の前記動画符号化単位の最大符号量と、所定数の前記音声符号化単位の最大符号量とを設定する設定工程と、動画符号量制御手段が、前記所定数の動画符号化単位の符号量が前記設定工程において設定された動画符号化単位の最大符号量を超えないように前記動画像データを符号化するよう前記動画符号化手段を制御する動画符号量制御工程と、音声符号量制御手段が、前記所定数の音声符号化単位の符号量が前記設定工程において設定された音声符号化単位の最大符号量を超えないように前記音声データを符号化するよう前記音声符号化手段を制御する音声符号量制御工程と、記録手段が、それぞれが前記所定数の動画符号化単位の符号化された動画像データを含む複数の動画チャンクと、それぞれが前記所定数の音声符号化単位の符号化された音声データを含む複数の音声チャンクとを多重化して記録媒体に記録すると共に複数の前記動画チャンク及び音声チャンクの各々にアクセスするためのオフセット値の情報を前記記録媒体に記録する記録工程とを有し、前記記録工程では、前記動画チャンクと前記音声チャンクを交互に多重化し、一つの前記動画チャンクと一つの前記音声チャンクの合計のサイズが前記合計のサイズの情報を前記記録媒体に記録する。

40

50

記動画符号化単位の最大符号量と前記音声符号化単位の最大符号量の合計に等しくなるように前記記録媒体に記録すると共に、複数の前記動画チャンクと音声チャンク及び前記オフセット値の情報とは別に、前記合計のサイズの情報を前記記録媒体に記録する。

【発明の効果】

【0017】

本発明によれば、ファイルとして記録された動画像データや音声データの再生時やサーチ時に必要なメモリ容量を削減し、データアクセスを高速化できる。

【発明を実施するための最良の形態】

【0018】

以下に、添付図面を参照して本発明を実施するための最良の形態について詳細に説明する。 10

【0019】

尚、以下に説明する実施の形態は、本発明の実現手段としての一例であり、本発明が適用される装置の構成や各種条件によって適宜修正又は変更されるべきものであり、本発明は以下の実施の形態に限定されるものではない。

【0020】

〔記録装置の説明〕

図3は、本発明に係る実施形態の記録装置の構成を示すブロック図である。

【0021】

図3において、301は画像データの入力端子である。302は画像符号化回路である。303は画像とデータとオーディオデータのストリームを多重化する多重化回路である。304は記録回路である。305はオーディオデータの入力端子である。306はオーディオの符号化回路である。307はファイルのヘッダの付加回路である。308は記録時の画質等のパラメータの入力端子である。309は符号量設定回路である。310はチャンク構成設定回路である。311は記録媒体である。 20

【0022】

本実施形態では、動画像データ及び音声データをMP4ファイル形式で記録再生する。

【0023】

ここで、MP4ファイルの構成について説明する。

【0024】

MP4ファイルはQuicktime(登録商標)ファイルフォーマットに基づいている。Quicktime(登録商標)ファイルフォーマットは、図9に示すように、box(又はatom)と呼ばれる基本単位で構成される。 30

【0025】

boxは、図9に示すようにboxのサイズを示す4バイトのsizeフィールドと、boxのタイプを示す4バイトのtypeフィールドと、それに続くdataフィールドとからなる。

【0026】

MP4ファイルは基本的には図7(a)に示すように、符号化された画像データ及びオーディオデータのストリームが入っているmdat boxと、各ストリームに関する情報が入っているmoov boxとからなる。 40

【0027】

mdat boxの中は更に、図7(b)に示すようにチャンク(chunk)で構成され、この例の場合には動画像のチャンクvideo chunk v(n)とオーディオのチャンクaudio chunk a(n)とから構成される。

【0028】

動画像及びオーディオの各チャンクは、更に図7(d)、(f)に示すようにサンプル(sample)から構成される。図7(d)は動画像のchunk v(1)がsample sv(1)、sv(2)、・・・、sv(K)から構成されることを示している。また、図7(f)はオーディオのchunk a(1)がsample sa(1)、sa(2)、・・・、sa(M)から構成されることを示している。

【0029】

50

上記各sampleは、例えば動画像の場合、図7(e)に示すように、sample sv(1)、sv(2)、sv(3)、sv(4)、…に対しても $I_n$ 、 $B_{-2}$ 、 $B_{-1}$ 、 $P_3$ 、…の符号化されたMPEGストリームが対応する。ここで $I_n$ はイントラ符号化(フレーム内符号化)されたフレーム画像データである。 $B_n$ は双方向から参照して符号化(フレーム間符号化)されるフレーム画像データである。 $P_n$ は一方向(順方向)から参照して符号化(フレーム間符号化)されるフレーム画像データである。上記画像データはいずれも可変長データである。

#### 【0030】

オーディオデータの各sampleは、図7(g)に示すように、オーディオデータの符号化的単位であるフレームに対応付けられる。

#### 【0031】

図7(a)のmoov boxは、図7(c)に示すように、作成日時等を入れるヘッダ情報からなるmvhd boxを有する。また、moov boxは、mdat boxに格納されるストリームデータの情報を入れる動画像及びオーディオの各trak box(video)とtrak box(audio)を有する。

#### 【0032】

trak boxに格納されるboxは、図8(a)に示すようなboxであるが、ここでは本発明に関係するstsc、stsز、stcoの各boxについてのみ説明する。

#### 【0033】

stco boxは、図8(b)に示すように、ストリームの各チャンクへのオフセットアドレス値を格納する。図8(a)、(b)は動画像のstco boxのみを示しているが、stco boxは動画像、オーディオの各trackに1つずつ存在する。

#### 【0034】

stsز boxは、図8(c)に示すように、ストリームの各サンプルのサイズを示す情報を格納する。図8(a)、(c)は動画像のstsز boxのみを示しているが、stsز boxは動画像とオーディオの各trackに1つずつ存在する。

#### 【0035】

stsc boxは、図8(d)に示すように、各チャンクのサンプル数を示す情報を格納する。図8(a)、(d)は動画像のstsc boxのみを示しているが、stsc boxは動画像とオーディオの各trackに1つずつ存在する。

#### 【0036】

尚、上記stco、stsز、stscの各boxに格納されるデータは記録時間に伴って増大していく。

#### 【0037】

例えば秒間30フレーム/秒の画像を15フレーム毎に1チャンクに格納するようにしたとして、1時間記録したとすると、動画像の総チャンク数がおよそ、 $2 \times 60 \times 60 = 7200$ となる。

#### 【0038】

各チャンクへのオフセットアドレス情報を格納するstco boxは、 $2 \times 60 \times 60 \times 4 = 28800$ (byte)となり、およそ28800バイト(約28kbyte)ほどのデータエリアを必要とする。

#### 【0039】

また、各サンプルのサイズを示す情報を格納するstsز boxは、 $30 \times 60 \times 60 \times 4 = 432000$ (byte)となり、およそ432000バイト(約422kbyte)ほどのデータエリアを必要とする。

#### 【0040】

また、各チャンクのサンプル数を示す情報を格納するstsc boxは、 $2 \times 60 \times 60 \times 12 = 86400$ (byte)となり、およそ86400バイト(約84kbyte)ほどのデータエリアを必要とする。

#### 【0041】

オーディオデータに関しては、例えばAACを用いてオーディオデータを圧縮する。サンプリング周波数を48kHz、フレームサイズ(サンプルサイズ)を1024オーダー

10

20

30

40

50

ディオサンプルとする。そして、1チャンク当たり46フレーム(サンプル)(約1秒)とすると、1時間記録した時のオーディオの総チャンク数はおよそ、 $48000 \times 60 \times 60 / (1024 \times 46) = 3668$ となる。

#### 【0042】

stco boxは、 $3668 \times 4 = 14672$ (byte)

stsز boxは、 $46 \times 3668 \times 4 = 674912$ (byte)

stsc boxは、 $3668 \times 12 = 44016$ (byte)となり、それぞれ約14kbyte、659kbyte、42kbyteほどになる。

#### 【0043】

上記からMP4ファイルとして記録していく場合、1時間で1.2Mbyte以上のおも10ov boxの容量を必要とする。

#### 【0044】

[記録動作及び再生動作の説明]

先ず、具体的な記録動作について説明する。

#### 【0045】

操作者の設定により、入力端子308から画質の記録パラメータ、若しくは記録媒体311への所望する記録時間のパラメータが符号量設定回路309に入力される。

#### 【0046】

更に、チャンク構成設定回路310から画像とオーディオデータのチャンク構成の情報が符号量設定回路309、画像符号化回路302、オーディオ符号化回路306、ヘッダ付加回路307に送られる。

#### 【0047】

ここでチャンク構成設定回路310は例えば図7(d)、(f)に示すように、動画像データをMPEG2若しくはMPEG4で圧縮し、Kフレームで動画像データの1チャンクとする。また、オーディオデータはMPEG2 AACで圧縮し、Mフレームで1チャンクを構成するように符号量設定回路309、画像符号化回路302、オーディオ符号化回路306、ヘッダ付加回路307に設定情報を送る。本実施形態では、動画像データの1チャンクをMPEG符号化におけるGOP(Group of Picture)の整数倍のフレーム数とする。

#### 【0048】

符号量設定回路309は、入力端子308から入力されたパラメータとチャンク構成設定回路310から供給されるチャンクの構成情報とから画像及びオーディオデータの1チャンク当たりの符号量を決める。

#### 【0049】

例えば図5(a)に示すように、画像データ1チャンクの符号量をVL、オーディオデータ1チャンクの符号量をALとし、画像データ1チャンクとオーディオデータ1チャンクの合計の符号量 $VL + AL = L_{fix}$ と決める。図5(a)では、1GOPを15フレームとし、1チャンクを2GOP分、即ち、 $K = 30$ フレームとしている。

#### 【0050】

$L_{fix}$ は少なくとも記録開始から記録停止まで、固定の値として設定される。

#### 【0051】

符号量設定回路309は決定した1チャンク当たりの画像データ符号量VL、オーディオデータ符号量ALを画像符号化回路302、オーディオ符号化回路306、ヘッダ付加回路307に設定する。

#### 【0052】

次に記録が開始されると、画像符号化回路302、オーディオ符号化回路306は画像データ及びオーディオデータの符号化を開始する。

#### 【0053】

ここで、図1を用いて画像符号化回路302の詳細について説明する。

#### 【0054】

10

20

30

40

50

図1において、101は画像データの入力端子である。102はカメラ信号処理回路である。103は符号化のためのフレーム順の並べ替えを行う並べ替え回路である。104は加算回路である。105はスイッチ回路である。106はDCT回路である。107は量子化回路である。108は逆量子化回路である。109は逆DCT回路である。110は加算回路である。111はメモリである。112は動き補償回路である。113はスイッチ回路である。114は可変長符号化回路である。115はストリーム生成回路である。116はバッファである。117は符号量制御回路である。118はストリームデータの出力端子である。119はストリームパラメータ付加回路である。120はヘッダ情報生成回路である。121はヘッダに付加する情報の出力端子である。122は符号量設定回路である。123は記録制御回路である。124は記録制御回路への制御データ入力端子である。

#### 【0055】

入力端子124から符号量設定回路309と、チャンク構成設定回路310からの記録に関するパラメータであるチャンク当たりの符号量VL及びチャンク構成を示すパラメータKが入力される。記録制御回路123は入力されたチャンク当たりの符号量VLとチャンク構成を示すパラメータKを符号量設定回路122に供給する。符号量設定回路122はチャンク当たりのGOP数gnを算出し、GOPがgn個で符号量VLとなるように符号量制御回路117に設定すると共に、ストリームパラメータ付加回路119にも供給する。

#### 【0056】

記録が開始されると、不図示の撮像素子で撮像された画像データが入力端子101から入力され、カメラ信号処理回路102でカメラ信号処理され、輝度信号／色差信号として並べ替え回路103に出力される。

#### 【0057】

並べ替え回路103では、図5(b)に示す時間軸tに対してカメラ信号処理回路102から各フレーム画像F<sub>-2</sub>、F<sub>-1</sub>、F<sub>0</sub>、F<sub>1</sub>、F<sub>2</sub>、F<sub>3</sub>、F<sub>4</sub>、F<sub>5</sub>、F<sub>6</sub>、…が出力される。そして、並べ替え回路103は、これらの画像フレームをF<sub>0</sub>、F<sub>-2</sub>、F<sub>-1</sub>、F<sub>3</sub>、F<sub>1</sub>、F<sub>2</sub>、F<sub>6</sub>、F<sub>4</sub>、F<sub>5</sub>、…の順に並べ替える。また、並べ替え回路103は、フレーム内の画像データを符号化する単位の小ブロックに分割する処理を行い、加算回路104、スイッチ回路105に出力する。

#### 【0058】

記録制御回路123はスイッチ回路105の端子を記録開始時、まず端子aを選択するよう指示し、スイッチ回路105は記録制御回路123により端子aを選択し、フレーム画像F<sub>0</sub>をDCT回路106に送出する。DCT回路106はフレーム画像F<sub>0</sub>に対して離散コサイン変換(Discrete Cosine Transform、以下、「DCT」)を施し、量子化回路107に送出する。

#### 【0059】

量子化回路107では符号量制御回路117から設定された量子化処理で、DCT回路106から供給されるDCT処理されたデータを量子化し、可変長符号化回路114と逆量子化回路108に送出する。

#### 【0060】

可変長符号化回路114は量子化回路107から供給される量子化された画像データをハフマン符号等を用いて可変長符号化する。そして可変長符号化回路114は、図5(b)に示すように、フレーム画像F<sub>0</sub>をフレーム内符号化データI<sub>0</sub>としてストリーム生成回路115に送出する。

#### 【0061】

逆量子化回路108は量子化回路107から供給された量子化された画像データを逆量子化し、逆DCT回路109に送出する。逆DCT回路109は逆量子化回路108で逆量子化された画像データを逆離散コサイン変換(逆DCT)し、加算回路110に送出する。

10

20

30

40

50

**【0062】**

加算回路110は逆DCT回路109から供給される逆DCTされた画像データにスイッチ回路113から供給されるデータを加算してメモリ111に供給する。ここで、スイッチ回路113は記録制御回路123の指示により、端子aが選択され、「0」データを加算回路110に供給する。従って加算回路110は逆DCT回路109から供給されるデータと同じ値をメモリ111に供給し、メモリ111は供給された画像データを記憶する。

**【0063】**

フレーム画像F<sub>0</sub>の処理が終わると、引き続きフレーム画像F<sub>-2</sub>の画像データが並べ替え回路103から加算回路104、動き補償回路112に入力される。この時、スイッチ回路105は記録制御回路123の指示により、端子bが選択される。更に記録制御回路123は動き補償回路112を制御する。動き補償回路112は、記録開始時であるので図5に示すようにフレーム画像F<sub>0</sub>のみを参照して予測誤差が最も小さくなるフレーム画像F<sub>0</sub>のブロックをメモリ111から読み出し加算回路104に供給する。10

**【0064】**

加算回路104は並べ替え回路103から供給されるフレーム画像F<sub>-2</sub>の画像データと動き補償回路112から供給されるデータを引き算し、予測誤差データとしてスイッチ回路105に供給する。

**【0065】**

スイッチ回路105は加算回路104から供給された引き算された予測誤差データをDCT回路106に送出する。DCT回路106は加算回路104から供給される予測誤差データをDCT処理して量子化回路107に供給する。20

**【0066】**

量子化回路107では符号量制御回路117から設定された量子化処理で、DCT回路106から供給されるDCT処理された予測誤差データを量子化し、可変長符号化回路114に送出する。

**【0067】**

可変長符号化回路114は量子化回路107から供給される量子化された予測誤差データを可変長符号化する。そして、可変長符号化回路114は、図5に示すように、フレーム画像F<sub>-2</sub>をフレーム間符号化データB<sub>-2</sub>としてストリーム生成回路115に送出する。30

**【0068】**

次のフレーム画像F<sub>-1</sub>の処理もフレーム画像F<sub>-2</sub>と同様に処理される。処理されたフレーム画像は可変長符号化回路114からフレーム間符号化（双方向予測符号化）データB<sub>-1</sub>としてストリーム生成回路115に送出される。

**【0069】**

次にフレーム画像F<sub>3</sub>が並べ替え回路103から加算回路104、動き補償回路112に入力される。この時、スイッチ回路105は記録制御回路123の指示により、端子bが選択される。そして動き補償回路112はフレーム画像F<sub>0</sub>を参照して予測誤差が最も小さくなるフレーム画像F<sub>0</sub>のブロックをメモリ111から読み出し加算回路104に供給する。40

**【0070】**

加算回路104は並べ替え回路103から供給されるフレーム画像F<sub>3</sub>の画像データと動き補償回路112から供給される動き補償されたフレーム画像F<sub>0</sub>のデータを引き算する。また、加算回路104は、予測誤差データとしてスイッチ回路105に供給する。スイッチ回路105は加算回路104から供給された引き算された予測誤差データをDCT回路106に送出する。DCT回路106は加算回路104から供給される予測誤差データをDCT処理して量子化回路107に供給する。

**【0071】**

量子化回路107では符号量制御回路117から設定された量子化処理で、DCT回路106から供給されるDCT処理された予測誤差データを量子化し、可変長符号化回路150

14と逆量子化回路108に送出する。

**【0072】**

可変長符号化回路114は量子化回路107から供給される量子化された予測誤差データを可変長符号化する。そして、可変長符号化回路114は、図5(b)に示すように、フレーム画像F3をフレーム間符号化(片方向予測符号化)データP<sub>3</sub>としてストリーム生成回路115に送出する。

**【0073】**

逆量子化回路108は量子化回路107から供給された量子化された予測誤差データを逆量子化し、逆DCT回路109に送出する。逆DCT回路109は逆量子化回路108で逆量子化された予測誤差データを逆DCT処理し、加算回路110に送出する。

10

**【0074】**

加算回路110は逆DCT回路109から供給される逆DCT処理された予測誤差データにスイッチ回路113から供給されるデータを加算してメモリ111に供給する。ここで、スイッチ回路113は記録制御回路123の指示により、端子bが選択され、動き補償回路112から供給される動き補償されたフレーム画像F<sub>0</sub>のデータを加算回路110に供給する。従って加算回路110は逆DCT回路109から供給される予測誤差データに動き補償回路112から供給される動き補償されたデータを加算してメモリ111に供給し、メモリ111は供給された画像データを記憶する。

**【0075】**

次にフレーム画像F<sub>1</sub>、F<sub>2</sub>が処理される。即ち、動き補償回路112で、局所復号されたフレーム画像F<sub>0</sub>、F<sub>3</sub>の双方向から予測される以外は上記F<sub>-2</sub>、F<sub>-1</sub>の処理と同じであり、可変長符号化回路114で可変長符号化される。そして、図5(b)に示すように、フレーム画像F<sub>1</sub>、F<sub>2</sub>はフレーム間符号化データB<sub>1</sub>、B<sub>2</sub>としてストリーム生成回路115に送出される。

20

**【0076】**

以下、フレーム画像F<sub>6</sub>、F<sub>4</sub>、F<sub>5</sub>、F<sub>9</sub>、F<sub>7</sub>、F<sub>8</sub>、F<sub>12</sub>、F<sub>10</sub>、F<sub>11</sub>と処理し、図5(b)に示すようにP<sub>6</sub>、B<sub>4</sub>、B<sub>5</sub>、P<sub>9</sub>、B<sub>7</sub>、B<sub>8</sub>、P<sub>12</sub>、B<sub>10</sub>、B<sub>11</sub>を得、ストリーム生成回路115に供給していく。

**【0077】**

ストリーム生成回路115では、図5(c)に示すように、シーケンスヘッダ、GOPヘッダ、ピクチャヘッダ等を付加してストリームを生成し、バッファ116に供給する。そして、ストリーム生成回路115は、図5(b)に示すように、符号化データとして、I<sub>0</sub>、B<sub>-2</sub>、B<sub>-1</sub>、P<sub>3</sub>、B<sub>1</sub>、…、P<sub>12</sub>、B<sub>10</sub>、B<sub>11</sub>で1GOP(GOP<sub>(1)</sub>)を形成する。

30

**【0078】**

ストリームパラメータ付加回路119は、図5(c)に示すように、ストリーム生成回路115に、画像データの符号量VL、フレーム内符号化データI<sub>0</sub>の符号化サイズI<sub>0</sub>(size)をユーザーデータとして付加する。また、ストリームパラメータ付加回路119は、図5(c)に示すように、ストリーム生成回路115に、次のGOPのIフレームまでのオフセット値I<sub>0</sub>(offset)をユーザーデータとして付加する。

40

**【0079】**

バッファ116は記録制御回路123の制御を受け、出力端子118から図3の多重化回路303に画像のストリームデータを送出する。

**【0080】**

符号量制御回路117は符号量設定回路122から設定されたチャンク当たりのGOP数g<sub>n</sub>とチャンク当たりの符号量VLに従い、GOPがg<sub>n</sub>個で符号量がVLとなるようバッファ116の符号量を監視し、量子化回路107での量子化処理を制御する。

**【0081】**

ヘッダ情報生成回路120は、ストリーム生成回路115からMP4ファイルのヘッダに必要とされるチャンクへのオフセット値、チャンク当たりのサンプル数、サンプルのサ

50

イズ等の情報を生成する。そして、ヘッダ情報生成回路 120 は、生成された情報を出力端子 121 から出力し図 3 のヘッダ付加回路 307 に供給する。

#### 【0082】

次に、図 2 を参照してオーディオの符号化について説明する。

#### 【0083】

図 2において、201 はオーディオデータの入力端子である。202 は M D C T 回路である。203 は T N S 回路である。204 はインテンシティステレオ回路である。205 は M / S ステレオ回路である。206 はスケールファクタ計算回路である。207 は量子化回路である。208 は可変長符号化回路である。209 はストリーム生成回路である。210 は出力端子である。211 はヘッダ情報生成回路である。212 は出力端子である。  
10 213 は聴覚心理モデル回路である。214 はビット割り当て回路である。215 は記録制御回路である。216 は入力端子である。

#### 【0084】

入力端子 216 から符号量設定回路 309、チャンク構成設定回路 310 からの記録に関するパラメータであるチャンク当たりの符号量 A L とチャンク構成を示すパラメータ M が記録制御回路 215 に入力される。

#### 【0085】

記録が開始されると、不図示の A / D 変換回路等から入力端子 201 にオーディオデータが入力され、M D C T 回路 202、聴覚心理モデル回路 213 に供給される。

#### 【0086】

M D C T 回路 202 では入力端子 201 から入力されるオーディオデータを 1024 サンプルで 1 フレームとして変形離散コサイン変換（以下、「M D C T」）を施し、T N S 回路 203 に出力する。尚、変形離散コサイン変換は、Modified Discrete Cosine Transform の略称である。T N S 回路 203 では T N S (Temporal Noise Shaping) 処理を施し、インテンシティステレオ回路 204 に出力する。T N S は量子化雑音を信号波形の振幅値に応じて整形することにより音質の向上を図る処理である。ここでは M D C T 係数の一部を時系列信号とみなして線形予測する。これにより、量子化雑音は信号波形の振幅が大きいところに集中する。

#### 【0087】

インテンシティステレオ回路 204 では、高域では左右チャネルの信号パワー差により音源の位置を感じるという聴覚の特性を利用し、左右チャネルの和信号とパワー比を本来の 2 チャネルデータの代わりに用いる。インテンシティステレオ回路 204 から出力されたオーディオデータは M / S ステレオ回路 205 に送出される。

#### 【0088】

M / S ステレオ回路 205 では、左右のチャネルの和信号と差信号を本来の 2 チャネルデータの代わりに用いる。主に低域の信号処理に用いられる。M / S ステレオ回路 205 から出力されたオーディオデータはスケールファクタ計算回路 206 に供給される。

#### 【0089】

スケールファクタ計算回路 206 では聴覚特性に合わせて複数の帯域（スケールファクタバンド）に分割された各帯域ごとの代表値（スケールファクタ）を算出し、帯域毎にスケールファクタを用いて正規化処理を行う。そして、スケールファクタ計算回路 206 は正規化されたデータを量子化回路 207 に送出する。

#### 【0090】

量子化回路 207 はスケールファクタ計算回路 206 から供給されたデータに対して、ビット割り当て回路 214 からの制御に従って量子化を行い、可変長符号化回路 208 に送出する。

#### 【0091】

ここで、入力端子 201 から入力されたオーディオデータは聴覚心理モデル回路 213 にも入力されている。聴覚心理モデル回路 213 では入力されたオーディオデータを F F T 処理し、スケールファクタバンド毎に絶対可聴閾値とマスキング効果を加味して、補正

10

20

30

40

50

された可聴閾値を算出し、ビット割り当て回路 214 に送出する。

【0092】

ビット割り当て回路 214 は、記録制御回路 215 により、聴覚心理モデル回路 213 から供給されたスケールファクタバンド毎の補正された可聴閾値から閾値を越える成分について M フレーム（1 チャンク）で A L ビットとなるようにビットを割り当てていく。

【0093】

可変長符号化回路 208 は量子化回路 207 から量子化されたデータを供給され、ハフマン符号化を施し、符号化したデータをストリーム生成回路 209、ビット割り当て回路 214 に送出する。

【0094】

ストリーム生成回路 209 は、可変長符号化回路 208 から供給される符号化データと、回路 202 ~ 207 からの処理帯域の範囲やスケールファクタ等の符号化パラメータとを用いてオーディオのストリームを生成する。そして、ストリーム生成回路 209 は、生成されたオーディオストリームを出力端子 210 から出力し図 3 の多重化回路 303 に供給する。

10

【0095】

また、ヘッダ情報生成回路 211 は、ストリーム生成回路 209 から MP4 ファイルのヘッダに必要とされるチャンクへのオフセット値、チャンク当たりのサンプル数、サンプルのサイズ等の情報を生成する。そして、ヘッダ情報生成回路 211 は、生成された情報を出力端子 212 から出力し図 3 のヘッダ付加回路 307 に供給する。

20

【0096】

多重化回路 303 は画像符号化回路 302、オーディオ符号化回路 306 から供給される画像とオーディオの符号化データを図 5 (a) に示すようにビデオ、オーディオ 1 チャンク毎にインターリープして多重化する。そして、多重化回路 303 は、ビデオ 1 チャンクとオーディオ 1 チャンクで符号量  $V_L + A_L = L_{fix}$  となるようにして mdat box のデータを生成し、記録媒体 311 に記録していく。

【0097】

更にヘッダ付加回路 307 は画像符号化回路 302 及びオーディオ符号化回路 306 からチャンクへのオフセット値、チャンク当たりのサンプル数、サンプルのサイズ等の情報を供給される。そして、ヘッダ付加回路 307 は、これらの情報から、video track 及び audio track の stco、stsc、stsz box のデータを作成し、多重化回路 303 に供給する。

30

【0098】

また、ヘッダ付加回路 307 は、ビデオチャンク、オーディオチャンクの符号量  $V_L$ 、 $A_L$ 、ビデオチャンク、オーディオチャンクの合計符号量  $L_{fix}$  をユーザーデータとして記録する free box を生成し、多重化回路 303 に供給する。また、ヘッダ付加回路 307 は、ビデオチャンクの構成を示す 1 チャンク当たりのフレーム数  $K$  をユーザーデータとして記録する free box を生成し、多重化回路 303 に供給する。更に、ヘッダ付加回路 307 は、GOP 数  $g_n$ 、オーディオチャンクの構成を示す 1 チャンク当たりのフレーム数  $M$ 、記録される画像データのフレームレート等をユーザーデータとして記録する free box を生成し、多重化回路 303 に供給する。

40

【0099】

多重化回路 303 は、図 5 (d) に示すように、moov box 内に free box F を生成し、moov box を記録媒体 311 に記録する。またユーザーデータには、データが上記形式であることを示す識別情報を入れておくようとする。

【0100】

尚、図 5 (d) では moov box 内に free box F を設けたが、free box F は moov box 内ではなく、moov box、mdat box と同等の位置に設けてもよい。

【0101】

[再生装置の説明]

次に、図 4 を参照して上記のように生成された MP4 ファイルを再生する装置について

50

説明する。

**【0102】**

図4において、401は記録媒体である。402は記録媒体401に記録されたデータを再生する再生回路である。403はバッファである。404は分離回路である。405は可変長復号化回路である。406は逆量子化回路である。407は逆DCT回路である。408は加算回路である。409はメモリである。410は動き補償回路である。411はスイッチ回路である。412は並べ替え回路である。413は出力端子である。414は可変長復号化回路である。415は逆量子化回路である。416はスケールファクタ回路である。417はM/Sステレオ回路である。418はインテンシティステレオ回路である。419はTNS回路である。420は逆MDCT回路である。421は出力端子である。422はヘッダ情報解析回路である。423はストリームユーザーデータ解析回路である。424は再生制御回路である。425は入力端子である。10

**【0103】**

図6は再生処理を示すフローチャートである。

**【0104】**

先ず、入力端子425から再生指示を受けると再生制御回路424は指定されたファイルを再生するために、指定されたファイルが上記形式で記録されたファイルであるかどうかを調べる。

**【0105】**

再生制御回路424は再生回路402を制御して、図6に示すように記録媒体401に記録されたファイルをオープンし(S602)、再生回路402を通してファイルデータを再生し、バッファ403に供給する。20

**【0106】**

ヘッダ情報解析回路422は、再生されてバッファ403に記憶される再生データを解析し、moov boxを検出する。moov boxが検出されると、記録されている画像の大きさ、オーディオの符号化方式等の符号化の基本パラメータを取得する(S604)。

**【0107】**

次に、free boxから、上記形式で記録したユーザー定義データがあることを示す識別情報があるかどうかを検出する(S605)。

**【0108】**

free box Fが検出されると、ビデオチャンクの構成を示す1チャンク当たりのフレーム数K、GOP数gn、オーディオチャンクの構成を示す1チャンク当たりのフレーム数Mを取得する(S606)。更にビデオチャンク、オーディオチャンクの符号量VL、AL、ビデオチャンク、オーディオチャンクの合計符号量Lfixを取得する(S607)。そしてstsc、stsz、stco boxは読み込みまず、ストリームデータのあるmdat boxを検出する(S608)。

**【0109】**

また、上記S605でfree box Fがない場合、再生制御回路424は、通常のファイルとして、stsc、stsz、stco boxを検出する(S610)。また、チャンクの構成パラメータのチャンクオフセット値、サンプル数、サンプルサイズを再生回路402を介してバッファ403に可能な限り読み出す(S611)。40

**【0110】**

本実施形態においては、ヘッダ情報解析回路422がチャンクの構成及びサイズを示すユーザデータを検出すると、再生制御回路424はstsc box、stsz box、stco boxのパラメータを使わずにmdat box内のストリームデータにアクセスする。

**【0111】**

即ち、ユーザデータが検出された場合、そのMP4ファイルは、1チャンク毎に固定長化されて記録されているため、オフセット値を使用せず、所望のチャンクを検出することができる。

**【0112】**

10

20

30

40

50

再生時において、各ビデオチャンクの先頭からのオフセット位置は( $L_{fix} \times n$ )バイト( $n = 0, 1, 2, \dots$ )で求められる。また、オーディオチャンクの先頭からのオフセット位置は( $L_{fix} \times n + VL$ )バイト( $n = 0, 1, 2, \dots$ )で求められる。

#### 【0113】

再生回路402は再生制御回路424の制御により、記録媒体401に記録されたファイルのmdat boxのストリームデータを再生し、バッファ403に供給する。バッファ403に記憶されたストリームデータはバッファの占有状態等をみて読み出しを始め、分離回路404に供給される。

#### 【0114】

分離回路404ではビデオのチャンクデータとオーディオのチャンクデータを分離し、ビデオのチャンクデータは可変長復号化回路405、ストリームユーユーザーデータ解析回路423に供給し、オーディオチャンクのデータは可変長復号化回路414に供給する。可変長復号化回路405は分離回路404から供給される再生されたストリームデータを可変長復号し、逆量子化回路406に供給する。

#### 【0115】

逆量子化回路406は可変長復号化回路405から供給される可変長復号化されたデータを逆量子化し、逆DCT回路407に送出する。逆DCT回路407は逆量子化回路406から供給される逆量子化されたデータに逆DCT処理を施し、加算回路408に送出する。加算回路408は逆DCT回路407から供給される逆DCT処理されたデータとスイッチ回路411から供給されるデータを加算する。

#### 【0116】

ここで、記録媒体401から再生されるストリームデータは、先ずGOP<sub>(1)</sub>のフレーム内符号化されたI<sub>0</sub>が再生される。再生制御回路424はスイッチ回路411の端子aを選択するよう制御し、スイッチ回路411は加算回路408にデータ‘0’を供給する。加算回路408はスイッチ回路411から供給される‘0’データと逆DCT回路407から供給される逆DCT処理されたデータとを加算して、再生されたフレームF<sub>0</sub>としてメモリ409と並べ替え回路412に供給する。メモリ409は加算回路408から供給される加算データを記憶する。

#### 【0117】

GOP<sub>(1)</sub>のフレーム内符号化データI<sub>0</sub>の次に再生されるのは双方向予測符号化されたピクチャデータB<sub>-2</sub>、B<sub>-1</sub>である。また、逆DCT回路407までの再生手続きは上記フレーム内符号化データI<sub>0</sub>で説明した再生処理と同様であるので省略する。

#### 【0118】

逆DCT回路407から逆DCT処理された双方向予測符号化されたピクチャデータが加算回路408に供給される。再生制御回路424はこの時、スイッチ回路411の端子bを選択するようスイッチ回路411を制御し、スイッチ回路411は端子bを選択し、加算回路408に動き補償回路410からのデータを供給する。

#### 【0119】

動き補償回路410は、再生されてくるストリームから、符号化時に生成されストリーム内に記録された動きベクトルを検出する。そして、動き補償回路410は、参照ブロックのデータ(この場合は記録開始時であるため、再生されたフレーム内符号化データF<sub>0</sub>からのデータのみ)をメモリ409から読み出してスイッチ回路411の端子bに供給する。

#### 【0120】

加算回路408は逆DCT回路407から供給される逆DCT処理されたデータとスイッチ回路411から供給される動き補償されたデータを加算する。また、加算回路408は、再生されたフレームF<sub>-2</sub>、F<sub>-1</sub>として並べ替え回路412に供給する。

#### 【0121】

次に、片方向方向予測符号化されたピクチャデータP<sub>3</sub>が再生され、逆DCT回路40

10

20

30

40

50

7までの再生処理は上記フレーム内符号化データ $I_0$ で説明した再生処理と同様であるので省略する。

#### 【0122】

逆DCT回路407から逆DCT処理された片方向予測符号化されたピクチャデータが加算回路408に供給される。再生制御回路424は、スイッチ回路411の端子bを選択するようスイッチ回路411を制御し、スイッチ回路411は端子bを選択し、加算回路408に動き補償回路410からのデータを供給する。

#### 【0123】

動き補償回路410は再生されてくるストリームから、符号化時に生成され、ストリーム内に記録された動きベクトルを検出する。そして、動き補償回路410は、参照ブロックのデータ（再生されたフレーム内符号化データ $F_0$ からのデータ）をメモリ409から読み出してスイッチ回路411の端子bに供給する。10

#### 【0124】

加算回路408は逆DCT回路407から供給される逆DCT処理されたデータとスイッチ回路411から供給される動き補償されたデータを加算し、再生されたフレーム $F_3$ としてメモリ409と並べ替え回路412に供給する。メモリ409は加算回路408から供給される加算データを記憶する。

#### 【0125】

次にピクチャ $B_1$ 、 $B_2$ が再生される。 $B_{-2}$ 、 $B_{-1}$ と違う点は、記録開始時のフレームではないため、双方向予測としてフレーム $F_0$ と $F_3$ から再生されることである。それ以外は上記 $B_{-2}$ 、 $B_{-1}$ と同様の処理によって再生される。20

#### 【0126】

以上説明したように、 $P_6$ 、 $B_4$ 、 $B_5$ 、 $\dots$ が順次再生されていく。

#### 【0127】

並べ替え回路412は順次再生されてくるフレーム $F_0$ 、 $F_{-2}$ 、 $F_{-1}$ 、 $F_3$ 、 $F_1$ 、 $F_2$ 、 $F_6$ 、 $F_4$ 、 $F_5$ 、 $\dots$ を並べ換える。即ち、並べ替え回路412は、再生されるフレームを $F_{-2}$ 、 $F_{-1}$ 、 $F_0$ 、 $F_1$ 、 $F_2$ 、 $F_3$ 、 $F_4$ 、 $F_5$ 、 $F_6$ 、 $\dots$ に並べ換えて出力端子413に出力する。

#### 【0128】

可変長復号化回路414に供給されたオーディオチャンクの再生データは、可変長復号され、逆量子化回路415に供給される。逆量子化回路415では復号された再生データを逆量子化し、スケールファクタ回路416に送出する。30

#### 【0129】

スケールファクタ回路416では正規化された成分をスケールファクタ値から元に戻し、M/Sステレオ回路417に送出する。M/Sステレオ回路417では和信号、差信号から左右チャネル信号を再生し、インテンシティステレオ回路418に送出する。

#### 【0130】

インテンシティステレオ回路418は右チャネルの和信号とパワー比から左右チャネル信号を再生し、TNS回路419に送出する。

#### 【0131】

TNS回路419では符号化時に用いた予測フィルタと逆特性のフィルタを用いてデータを再生し、逆MDCT回路420に送出する。40

#### 【0132】

逆MDCT回路420では再生されたデータに対して逆MDCTを施して再生オーディオデータを得、出力端子421から出力する。

#### 【0133】

<特殊再生>

以上、通常再生について説明したが、次に特殊再生について説明する。

#### 【0134】

再生制御回路424はmdat boxのストリームのファイル先頭からのオフセット値 $S_t(0)$ 50

`ffset`)をファイルの再生開始時に取得し保持している。

【0135】

ここで、ストリーム内の画像フレームの  $n$  番目を再生中に図4の入力端子425から順方向への特殊再生の指示信号が入力されるとする。例えば  $s$  秒先の時間を指定されたとする。

【0136】

$s$  秒間のフレーム数は  $s \times \text{frame\_rate} = u \dots (1)$

`frame_rate`は再生開始時、MP4ファイルのfree box Fのユーザーデータとして記録されているパラメータから取得する。

【0137】

再生されるフレームは  $n + u$  番目のフレームであるが、MPEGで圧縮されたストリームデータを特殊再生する場合、フレーム内符号化されたデータIから再生しなければならない。従って  $n + u$  番目のフレームの直前のフレーム内符号化された画像データ(イントラフレーム)を再生する。

【0138】

まず、 $n + u$  番目のフレームが含まれているチャンクまでのチャンク数  $r$  が計算される。

$$r = (n + u) / K \dots (2)$$

但し、 $r$  の小数点以下は切り捨てて小さいほうの整数にまるめる。

【0139】

次に再生されるチャンクまでのオフセット値  $P_r$  が

$$P_r = S_t(\text{offset}) + L_{fix} \times r \dots (3)$$

で算出される。

【0140】

次に再生されるイントラフレームのあるGOPのチャンク内の番号  $q$  を算出する。

$$q = (n + u - r) / (K / g_n) \dots (4)$$

但し、 $q$  の小数点以下は切り捨てて小さいほうの整数にまるめる。

【0141】

再生制御回路424は上記式(1)~(4)により各値を算出した後、バッファ403、可変長符号化回路405等の復号回路を初期化し、オーディオの出力端子421への出力をミュートする。

【0142】

再生回路402を制御して、 $P_r$  バイトにあるGOP( $Pr$ )まで記録媒体401をシークし、GOP( $Pr$ )からデータを再生し、バッファ403に供給する。バッファ403に記憶された再生データは直ちに分離回路404を介してストリームユーザーデータ解析回路423に供給される。

【0143】

ストリームユーザーデータ解析回路423ではGOP( $Pr$ )のユーザーデータを解析し、次のイントラフレームまでのオフセット値  $I(\text{offset})$  を取得し、再生制御回路424に供給する。

【0144】

再生制御回路424は次のイントラフレームまでのオフセット値  $I(\text{offset})$  が取得されるとすぐに再生回路402を介して記録媒体401のファイルをオフセット値  $I(\text{offset})$  分シークさせ、つぎのGOP( $Pr+1$ )を再生させる。そしてGOP( $Pr+1$ )のユーザーデータから次のイントラフレームへのオフセット値を取得する。

【0145】

上記処理を繰り返して特殊再生の目的であるGOP( $Pr+q$ )に辿り着く。再生制御回路424が再生回路402を介してGOP( $Pr+q$ )を再生し、バッファ403に記憶していく。分離回路404はバッファ403に記憶されたGOP( $Pr+q$ )の再生データをストリームユーザーデータ解析回路423に供給する。ストリームユーザーデータ解析回路423はユ

10

20

30

40

50

ユーザーデータからイントラフレームのサイズ I (size) を取得し、再生制御回路 424 にそのサイズ (I (size)) を伝える。再生制御回路 424 はサイズ情報 I (size) をもとに、イントラフレームのみを可変長復号化回路 405 以下に送出し、特殊再生の再生画像を得ることができる。そして、次の特殊再生画像を式 (1) ~ (4) を用いて計算を始める。

#### 【0146】

以上説明したように、記録時に、画像データとオーディオデータのチャンク構成を決め、画像データのチャンクの符号量 VL、オーディオデータのチャンクの符号量 AL としたときに、両チャンクの符号量 L fix = VL + AL が一定となるように符号化を行う。

#### 【0147】

更に画像データのストリームには、GOP 毎に画像データの符号量 VL とフレーム内符号化データ (イントラフレーム) の符号量 I (size)、次の GOP のイントラフレームまでのオフセット値 I (offset) をユーザーデータとして記録する。10

#### 【0148】

更に 1 チャンク当たりのフレーム数 K, M、符号化単位数 gn 等の各チャンクの構成を示す情報等をユーザーデータとして free box に格納して記録するようとする。また、符号化のフレームレート、各チャンクの符号量 VL, AL、及び合計の符号量 L fix、上記構成であることを示す識別情報等をユーザーデータとして free box に格納して記録するようとする。

#### 【0149】

再生時には free box から上記ストリームであることを認識すると、free box とストリームに記録されているユーザーデータの各情報からビデオ及びオーディオの各チャンクのストリームデータにアクセスする。これにより、通常再生も特殊再生も行うようにし、stco、stsc、stsz box 等のヘッダ情報を読み込まなくとも再生可能としたので、大きな容量を必要とするヘッダを記憶するメモリが不要になる。20

#### 【0150】

また、ヘッダにある符号化に関する上記 stco、stsc、stsz box 等のパラメータを読み込んで解析する必要がないのでストリームデータへのアクセスが早く行えるようになり、再生の待ち時間を減少でき、シームレスの再生が可能になる。

#### 【0151】

従来では、上記 stco、stsc、stsz box 等のパラメータの情報量が多すぎてメモリに 1 回で読み込めない場合に、再生しながら記録媒体からパラメータを読み直してメモリのパラメータ情報を更新することが必要であった。しかし、その必要はなくなり、機器本体での処理負荷も低減できる。30

#### 【0152】

また、ファイルは通常の MP4 ファイルであるので、他の機器での再生も可能である。

#### 【0153】

更に、1 つのファイルで実現できるので、記録媒体間でコピーされたデータを本装置で再生したとしても、上記再生処理と同様に再生できる。

#### 【0154】

また、上記例ではオーディオの符号化方式を MPEG - AAC として説明したが、MP EG I, II, III 等の他の他の符号化方式でも同様に記録再生でき、同様の効果が得られる。40

#### 【0155】

##### [他の実施形態]

以上、本発明に係る実施形態について具体例を用いて詳述したが、本発明は、例えば、システム、装置、方法、プログラム若しくは記憶媒体（記録媒体）等としての実施態様をとることが可能である。具体的には、複数の機器から構成されるシステムに適用しても良いし、また、一つの機器からなる装置に適用しても良い。

#### 【0156】

また、本発明の目的は、図示の機能ブロック及び動作において、いずれの部分をハード

50

ウェア回路により実現し、或いはコンピュータを用いたソフトウェア処理によって実現しても達成されることは言うまでもない。

【0157】

尚、本発明は、前述した実施形態の機能を実現するソフトウェアのプログラムを、システムあるいは装置に直接あるいは遠隔から供給することによって達成される場合も含む。その場合、システム等のコンピュータが該プログラムコードを読み出して実行することになる。

【0158】

従って、本発明の機能処理をコンピュータで実現するために、該コンピュータにインストールされるプログラムコード自体も本発明を実現するものである。つまり、本発明は、本発明の機能処理を実現するためのコンピュータプログラム自体も含まれる。10

【0159】

その場合、プログラムの機能を有していれば、オブジェクトコード、インタプリタにより実行されるプログラム、OSに供給するスクリプトデータ等の形態であっても良い。

【0160】

プログラムを供給するための記録媒体（記憶媒体）としては、例えば、フレキシブルディスク、ハードディスク、光ディスク、光磁気ディスク等がある。その他にも、MO、CD-ROM、CD-R、CD-RW、磁気テープ、不揮発性のメモリカード、ROM、DVD(DVD-ROM、DVD-R)等がある。

【0161】

その他、プログラムの供給方法としては、クライアントコンピュータのブラウザを用いてインターネットのホームページに接続し、該ホームページから本発明のコンピュータプログラムそのものをダウンロードすることもできる。また圧縮され自動インストール機能を含むファイルをハードディスク等の記録媒体にダウンロードすることによっても供給できる。また、本発明のプログラムを構成するプログラムコードを複数のファイルに分割し、それぞれのファイルを異なるホームページからダウンロードすることによっても実現可能である。つまり、本発明の機能処理をコンピュータで実現するためのプログラムファイルを複数のユーザに対してダウンロードさせるWWWサーバも、本発明に含まれるものである。20

【0162】

また、本発明のプログラムを暗号化してCD-ROM等の記憶媒体に格納してユーザに配布し、所定の条件をクリアしたユーザが、インターネットを介してホームページから暗号化を解く鍵情報をダウンロードすることもできる。この場合、ダウンロードした鍵情報を使用することにより暗号化されたプログラムを実行してコンピュータにインストールさせて実現する。30

【0163】

また、コンピュータが、読み出したプログラムを実行することによって、前述した実施形態の機能が実現される他、そのプログラムの指示に基づき、コンピュータ上で稼動しているOS等が、実際の処理の一部又は全部を行うことによっても実現され得る。

【0164】

更に、記録媒体から読み出されたプログラムが、コンピュータに挿入された機能拡張ボードやコンピュータに接続された機能拡張ユニットのメモリに書き込まれた後、該ボード等のCPU等が実際の処理の一部又は全部を行うことによっても実現される。40

【図面の簡単な説明】

【0165】

【図1】図3の画像符号化回路302の詳細構成を示すブロック図である。

【図2】図3のオーディオ符号化回路305の詳細構成を示すブロック図である。

【図3】本発明に係る実施形態の記録装置を示すブロック図である。

【図4】本発明に係る実施形態の再生装置を示すブロック図である。

【図5】(a)は画像ストリーム及びオーディオストリームとチャンクとの関係を説明す50

る図、(b)は符号化のフレーム順序を説明する図、(c)は画像ストリームのデータ構成を説明する図、(d)は本実施形態のfree boxの構成を説明する図である。

【図6】本実施形態の再生処理を示すフローチャートである。

【図7】MP4ファイルに記録される画像データ及びオーディオデータの構成を説明する図である。

【図8】MP4ファイルのmoov box内のstsc、stsz、stco boxを説明する図である。

【図9】Quicktime(登録商標)フォーマットの構成を説明する図である。

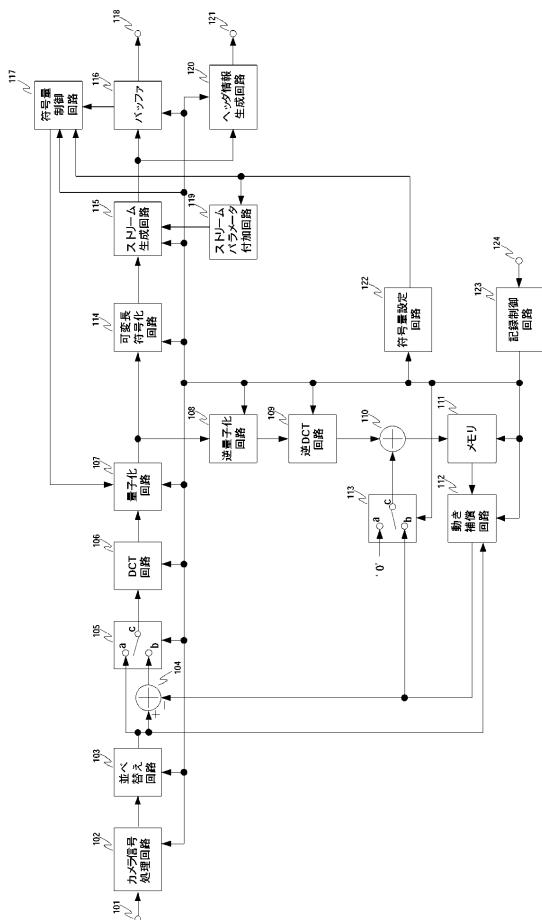
【符号の説明】

【0166】

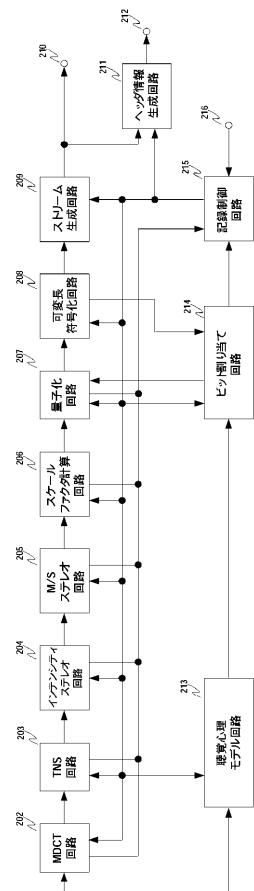
101	画像データ入力端子	10
102	カメラ信号処理回路	
103	並べ替え回路	
104	加算回路	
105	スイッチ回路	
106	DCT回路	
107	量子化回路	
108	逆量子化回路	
109	逆DCT回路	
110	加算回路	
111	メモリ	20
112	動き補償回路	
113	スイッチ回路	
114	可変長符号化回路	
115	ストリーム生成回路	
116	バッファ	
117	符号量制御回路	
118	ストリームデータ出力端子	
119	ストリームパラメータ付加回路	
120	ヘッダ情報生成回路	
121	ヘッダ情報出力端子	30
122	符号量設定回路	
123	記録制御回路	
124	制御データ入力端子	
201	オーディオデータ入力端子	
202	MDC T回路	
203	TNS回路	
204	インテンシティステレオ回路	
205	M/Sステレオ回路	
206	スケールファクタ計算回路	
207	量子化回路	40
208	可変長符号化回路	
209	ストリーム生成回路	
210	出力端子	
211	ヘッダ情報生成回路	
212	出力端子	
213	聴覚心理モデル回路	
214	ピット割り当て回路	
215	記録制御回路	
216	入力端子	
301	画像データ入力端子	50

3 0 2	画像符号化回路	
3 0 3	多重化回路	
3 0 4	記録回路	
3 0 5	オーディオデータ入力端子	
3 0 6	オーディオ符号化回路	
3 0 7	ヘッダ付加回路	
3 0 8	パラメータ入力端子	
3 0 9	符号量設定回路	
3 1 0	チャンク構成設定回路	
3 1 1	記録媒体	10
4 0 1	記録媒体	
4 0 2	再生回路	
4 0 3	バッファ	
4 0 4	分離回路	
4 0 5	可変長復号化回路	
4 0 6	逆量子化回路	
4 0 7	逆 D C T 回路	
4 0 8	加算回路	
4 0 9	メモリ	
4 1 0	動き補償回路	20
4 1 1	スイッチ回路	
4 1 2	並べ替え回路	
4 1 3	出力端子	
4 1 4	可変長復号化回路	
4 1 5	逆量子化回路	
4 1 6	スケールファクタ回路	
4 1 7	M / S ステレオ回路	
4 1 8	インテンシティステレオ回路	
4 1 9	T N S 回路	
4 2 0	逆 M D C T 回路	30
4 2 1	出力端子	
4 2 2	ヘッダ情報解析回路	
4 2 3	ストリームユーザーデータ解析回路	
4 2 4	再生制御回路	
4 2 5	入力端子	

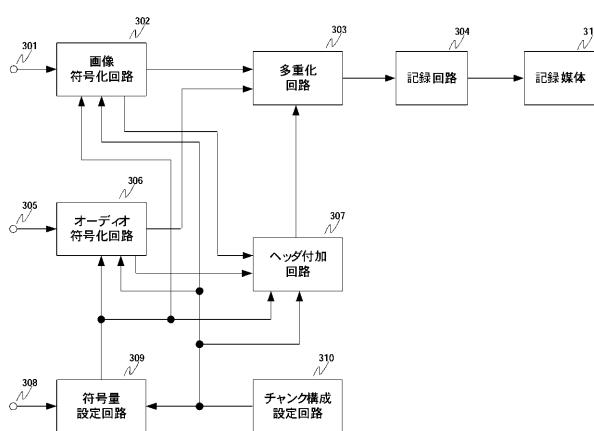
【図1】



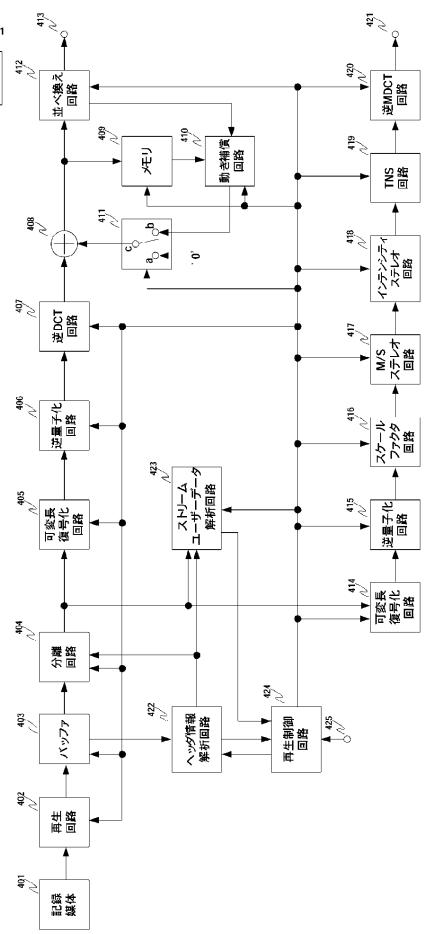
【図2】



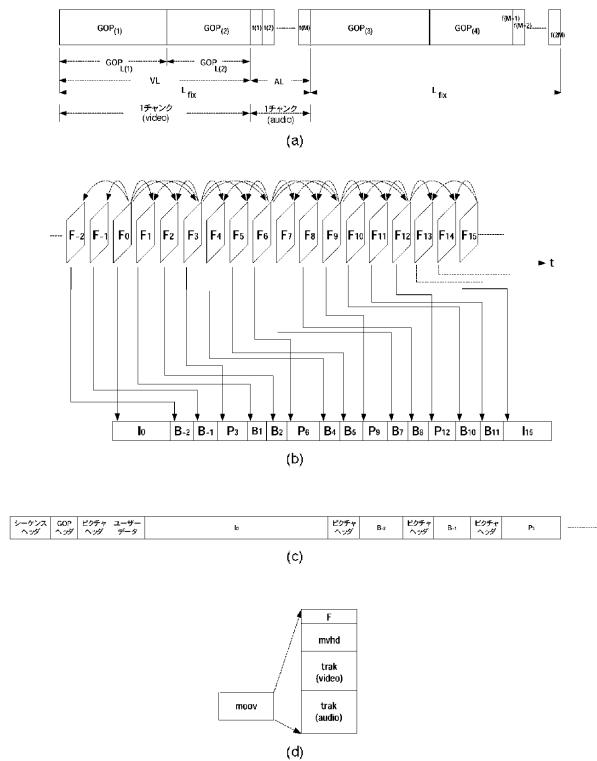
【図3】



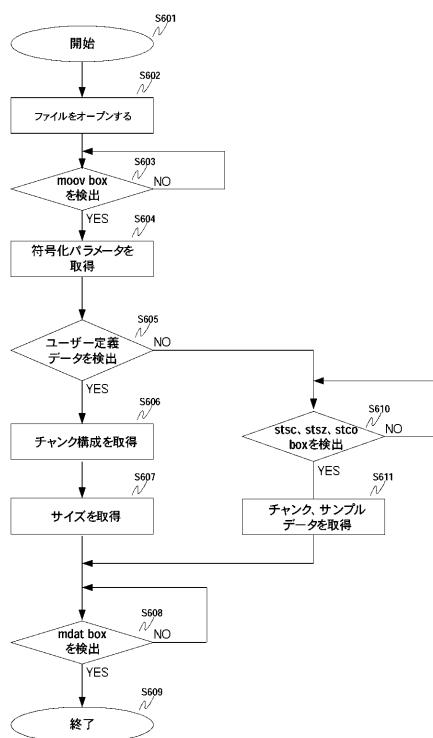
【図4】



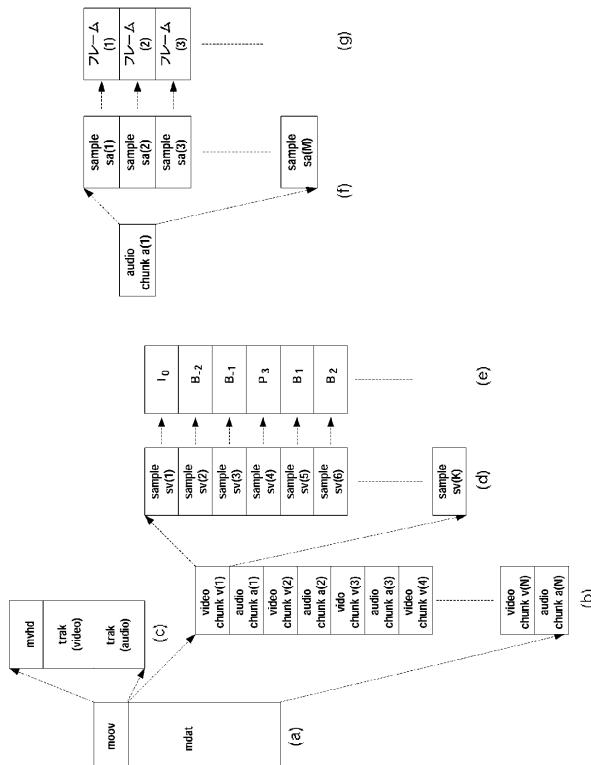
【図5】



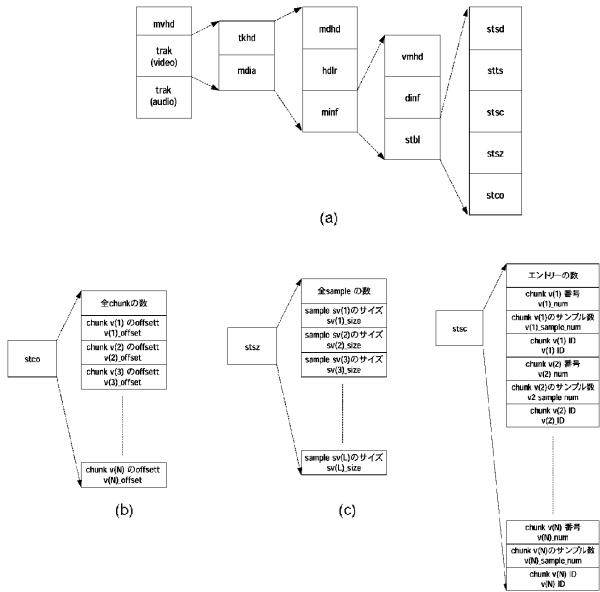
【図6】



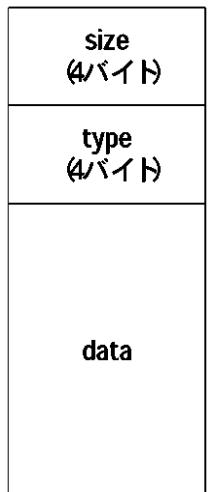
【図7】



【図8】



【図9】



---

フロントページの続き

審査官 梅本 章子

(56)参考文献 特開2004-282703(JP,A)  
特開2005-038516(JP,A)  
特開2005-223442(JP,A)

(58)調査した分野(Int.Cl., DB名)

H 04 N	5 / 7 6	-	5 / 9 5 6
H 04 N	7 / 2 4	-	7 / 6 8
G 11 B	2 0 / 1 0	-	2 0 / 1 2
G 11 B	2 7 / 0 0	-	2 7 / 3 4