

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第3867342号

(P3867342)

(45) 発行日 平成19年1月10日(2007.1.10)

(24) 登録日 平成18年10月20日(2006.10.20)

(51) Int. Cl.	F I	
G 1 1 B 20/12 (2006.01)	G 1 1 B 20/12	1 0 2
G 1 1 B 20/10 (2006.01)	G 1 1 B 20/10	3 0 1 Z
H O 4 J 3/00 (2006.01)	H O 4 J 3/00	M
H O 4 N 5/92 (2006.01)	H O 4 N 5/92	H
H O 4 N 7/26 (2006.01)	H O 4 N 7/13	Z

請求項の数 13 (全 16 頁)

(21) 出願番号	特願平9-111599	(73) 特許権者	000002185
(22) 出願日	平成9年4月28日(1997.4.28)		ソニー株式会社
(65) 公開番号	特開平10-83632		東京都品川区北品川6丁目7番35号
(43) 公開日	平成10年3月31日(1998.3.31)	(74) 代理人	100067736
審査請求日	平成15年10月16日(2003.10.16)		弁理士 小池 晃
(31) 優先権主張番号	特願平8-111682	(74) 代理人	100086335
(32) 優先日	平成8年5月2日(1996.5.2)		弁理士 田村 榮一
(33) 優先権主張国	日本国(JP)	(74) 代理人	100096677
			弁理士 伊賀 誠司
		(72) 発明者	安田 幹太
			東京都品川区北品川6丁目7番35号 ソニー株式会社内
		(72) 発明者	田原 勝己
			東京都品川区北品川6丁目7番35号 ソニー株式会社内

最終頁に続く

(54) 【発明の名称】 符号化装置および方法、伝送方法、並びに信号記録媒体

(57) 【特許請求の範囲】

【請求項1】

アクセスユニット単位で選択的に切り替えられる複数のビデオデータを可変長符号化して得られる複数のビデオストリームから、上記アクセスユニット毎のデコード時刻と同一のデコード時刻に対するアクセスユニットの大きさをそれぞれ検出するアクセスユニット検出手段と、

上記アクセスユニット検出手段により検出された複数のアクセスユニットの大きさをデコード時刻毎に比較して、アクセスユニットの最大値を仮想アクセスユニットとする最大値検出手段と、

上記仮想アクセスユニットから構成される仮想ビデオストリームを構成する仮想ビデオストリーム構成手段と、

上記ビデオストリームのバッファ占有量が上記仮想ビデオストリームのバッファ占有量以下となるように、上記ビデオストリームのアクセスユニットの大きさと上記仮想アクセスユニットの大きさとの差分を調整しながら、上記複数のビデオストリームをパケット化するパケット化手段と

を有することを特徴とする符号化装置。

【請求項2】

上記パケット化手段は、上記差分に対してパディングパケットを用いてパケット化することを特徴とする請求項1記載の符号化装置。

【請求項3】

10

20

上記パケット化手段は、上記パディングパケット自体でパックを構成している場合には、パケット化しないことを特徴とする請求項 2 記載の符号化装置。

【請求項 4】

上記仮想ビデオストリーム構成手段は、上記仮想ビデオストリームを用いて、デコーダへの供給時刻及びパケットの大きさを決定し、

上記パケット化手段は、上記デコーダへの供給時刻及びパケットの大きさの情報を使用して、上記ビデオストリームをパケット化する

ことを特徴とする請求項 1 記載の符号化装置。

【請求項 5】

上記ビデオストリームは M P E G 規格のエレメンタリストリームであることを特徴とする請求項 1 記載の符号化装置。

10

【請求項 6】

上記パケット化手段によりパケット化されて生成されたストリームは、M P E G 規格のプログラムストリームであることを特徴とする請求項 1 記載の符号化装置。

【請求項 7】

上記複数のビデオデータは、異なるカメラアングルから撮像されたデータであることを特徴とする請求項 1 記載の符号化装置。

【請求項 8】

アクセスユニット単位で選択的に切り替えられる複数のビデオデータを可変長符号化して得られる複数のビデオストリームから、上記アクセスユニット毎のデコード時刻と同一のデコード時刻に対するアクセスユニットの大きさをそれぞれ検出するアクセスユニット検出工程と、

20

上記アクセスユニット検出工程により検出された複数のアクセスユニットの大きさをデコード時刻毎に比較して、アクセスユニットの最大値を仮想アクセスユニットとする最大値検出工程と、

上記仮想アクセスユニットから構成される仮想ビデオストリームを構成する仮想ビデオストリーム構成工程と、

上記ビデオストリームのバッファ占有量が上記仮想ビデオストリームのバッファ占有量以下となるように、上記ビデオストリームのアクセスユニットの大きさと上記仮想アクセスユニットの大きさととの差分を調整しながら、上記複数のビデオストリームをパケット化するパケット化工程と

30

を有することを特徴とする符号化方法。

【請求項 9】

上記パケット化工程では、上記差分に対してパディングパケットを用いてパケット化することを特徴とする請求項 8 記載の符号化方法。

【請求項 10】

アクセスユニット単位で選択的に切り替えられる複数のビデオデータを可変長符号化して得られる複数のビデオストリームから、上記アクセスユニット毎のデコード時刻と同一のデコード時刻に対するアクセスユニットの大きさをそれぞれ検出するアクセスユニット検出工程と、

40

上記アクセスユニット検出工程により検出された複数のアクセスユニットの大きさをデコード時刻毎に比較して、アクセスユニットの最大値を仮想アクセスユニットとする最大値検出工程と、

上記仮想アクセスユニットから構成される仮想ビデオストリームを構成する仮想ビデオストリーム構成工程と、

上記ビデオストリームのバッファ占有量が上記仮想ビデオストリームのバッファ占有量以下となるように、上記ビデオストリームのアクセスユニットの大きさと上記仮想アクセスユニットの大きさととの差分を調整しながら、上記複数のビデオストリームをパケット化するパケット化工程と、

上記夫々パケット化された複数のストリームを 1 つのストリームとして伝送する工程と

50

を有することを特徴とする伝送方法。

【請求項 1 1】

上記パケット化工程では、上記差分に対してパディングパケットを用いてパケット化することを特徴とする請求項 1 0 記載の伝送方法。

【請求項 1 2】

記録信号が記録される信号記録媒体において、

上記記録信号は、

アクセスユニット単位で選択的に切り替えられる複数のビデオデータを可変長符号化して得られる複数のビデオストリームから、上記アクセスユニット毎のデコード時刻と同一のデコード時刻に対するアクセスユニットの大きさをそれぞれ検出するアクセスユニット

10

検出工程と、
上記アクセスユニット検出工程により検出された複数のアクセスユニットの大きさをデコード時刻毎に比較して、アクセスユニットの最大値を仮想アクセスユニットとする最大値検出工程と、

上記仮想アクセスユニットから構成される仮想ビデオストリームを構成する仮想ビデオストリーム構成工程と、

上記ビデオストリームのバッファ占有量が上記仮想ビデオストリームのバッファ占有量以下となるように、上記ビデオストリームのアクセスユニットの大きさと上記仮想アクセスユニットの大きさととの差分を調整しながら、上記複数のビデオストリームをパケット化するパケット化工程と

20

により得られるものであることを特徴とする信号記録媒体。

【請求項 1 3】

上記パケット化工程では、上記差分に対してパディングパケットを用いてパケット化することを特徴とする請求項 1 2 記載の信号記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、動画像信号及び音響信号などを、例えば光磁気ディスクや磁気テープなどの記録媒体に記録し、これを再生してディスプレイなどに表示したり、テレビ会議システム、テレビ電話システム、放送用機器など、動画像信号及び音響信号などを伝送路を介して

30

送信側から受信側に伝送し、受信側において、これを受信し、表示する場合などに用いて好適な、符号化装置および方法、伝送方法、並びに信号記録媒体に関する。

【0002】

【従来の技術】

映像信号や音声信号などを、光磁気ディスクや磁気テープなどの記録媒体に記録し、それらを再生してディスプレイなどに表示したり、テレビ会議システム、テレビ電話システムなどにおいて、送信側が、所定の伝送路を介して、それらの信号を伝送し、受信側が、それらの信号を受信し、表示する場合などにおいて、最近、これらの信号は、A/D変換した後、いわゆるMPEG(Moving Picture Experts Group)方式で符号化して取り扱われることが多くなりつつある。

40

【0003】

ここで、上記MPEGとは、ISO/IEC JTC1/SC29(International Organization for Standardization / International Electrotechnical Commission, Joint Technical Committee 1 / Sub Committee 29: 国際標準化機構 / 国際電気標準会議、合同技術委員会 1 / 専門部会 29) の蓄積用動画像符号化の検討組織の略称であり、MPEG1 標準としてISO11172が、MPEG2 標準としてISO13818がある。これらの国際標準において、マルチメディア多重化の項目でISO11172-1及びISO13818-1が、映像の項目でISO11172-2及びISO13818-2が、また音声の項目でISO11172-3及びISO13818-3がそれぞれ標準化されている。

【0004】

50

通常、映像と音声は同時に扱うので、一般的なシステムにおいては、映像信号、音声信号および、関連するデータといった複数のデータをまとめて（多重化して）記録・伝送する。そして、再生するときに、多重化されたデータを、映像信号や音声信号といったデータの種別別に分離した後復号して、それらのデータを同期させて再生する。

【 0 0 0 5 】

データを多重化する場合、所定の数の映像信号および音声信号を個別に符号化し、各信号に対する符号化ストリームを生成した後、それらの符号化ストリームが多重化される。

【 0 0 0 6 】

M P E Gシステム（ISO/IEC13818-1 あるいは ISO/IEC11172-1）では、この多重化ストリームを規定している。以下にこのM P E Gシステムにおけるデコーダモデルおよび多重化ストリームの構造について説明する。ここでは説明を簡単にするためM P E G 2（ISO/IEC13818-1）プログラムストリームおよびM P E G 1システム（ISO/IEC11172-1）ストリームについて述べるが、M P E G 2システムのトランスポートストリーム（ISO/IEC13818-1）もM P E G 2プログラムストリームと同様の原理によりデコードされる。

10

【 0 0 0 7 】

M P E Gシステムでは、仮想的なデコーダのモデル（S T D：システムターゲットデコーダ）が規定されていて、多重化システムストリームは、このS T Dにおいて正しく、すなわちバッファの破綻をきたさないように、復号されるものとして定義されている。

【 0 0 0 8 】

ここでこのS T D（システムターゲットデコーダ）の動作を説明する。図6はS T D（システムターゲットデコーダ）の一例の概略構成を表し、図7（A）及び（B）は、M P E G 2システムのプログラムストリームの構造及びM P E G 2トランスポートストリームの構造をそれぞれ示す。

20

【 0 0 0 9 】

S T Dは内部にシステムタイムクロック（S T C：System Time Clock）16と呼ばれる基準時計を持っている。このS T C 16はある時間間隔で増加している。一方M P E G 2システムのプログラムストリームは複数のアクセスユニットで構成されており、そのストリームには、図7に示すようにシステムクロックリファレンス（S C R：System Clock Reference）と呼ばれる時間情報がパックヘッダとよばれる領域にエンコードされている。デコーダはS T CがS C Rに等しくなるとその該当するパック、すなわちプログラムストリームの構成単位を、あるレート、すなわちパックヘッダの mux_rate field にエンコードされている値で読み出す。

30

【 0 0 1 0 】

読み出されたパックは直ちにその構成単位であるパケットの種類に応じて、分離器11で各エレメンタリーストリーム、すなわちビデオストリームやオーディオストリームなどに分離され、各エレメンタリーストリームのデコーダバッファ、すなわちビデオバッファ12やオーディオバッファ14に入力される。

【 0 0 1 1 】

パケットヘッダには、図7に示すように、デコーディングタイムスタンプ（D T S：Decoding Time Stamp）、プレゼンテーションタイムスタンプ（P T S：Presentation Time Stamp）と呼ばれる時間情報を記述するフィールドが用意されていて、それぞれエレメンタリーストリームのデコード単位（アクセスユニット）のデコードされるべき時刻と表示されるべき時刻を表している。特に、P T Sは、アクセスユニットが表示される時刻を表しており、D T Sは、アクセスユニットが復号される時刻を表している。ただし、D T S = P T Sとなるアクセスユニットについては、P T Sの値のみエンコードされる。デコーダバッファ12に入力されたアクセスユニットは、このS T Cの値がD T Sの値に等しくなったときにバッファから引き抜かれて各デコーダ、すなわちビデオデコーダ13やオーディオデコーダ15に入力されデコードされる。

40

【 0 0 1 2 】

このようにS T D（システムターゲットデコーダ）においては、同じ基準時計S T C 16

50

に対するデコード時刻の情報が、各エレメンタリーストリームのパケットにエンコードされているため、ビデオ、オーディオ、その他のデータを同期させて再生することが可能となっている。

【0013】

また、多重化においては、このSTDの各エレメンタリーストリームのデコーダバッファがオーバーフローおよびアンダーフローしないように、STDへのパックの供給時刻SCR（システムクロックリファレンス）が決定され、アクセスユニットをパケット化することが要求される。ここで、上記オーバーフローとは、バッファに供給されるデータがバッファ容量を越えてしまうことを意味し、アンダーフローは、アクセスユニットがデコードされるべき時刻にバッファにまだ到達していない状態を示す。このようなデコード時のバッファの破綻なく、上記オーバーフローやアンダーフローを生じさせないように、複数のビットストリームを多重化する技術を、本件出願人は、先に特願平7-341951号の明細書および図面において提案している。

10

【0014】

以上は、図7の(A)のMPEG2プログラムストリームについての説明であるが、図7の(B)のMPEG2トランスポートストリームについても同様な構造を持つ。図7の(B)のトランスポートストリームヘッダは、上記ISO/IEC13818-1において規定される、同期バイト(sync_byte)から巡回カウンタ(continuity_counter)までの4バイトのことである。クロックリファレンス及びデコード時刻は、図7の(A)のMPEGプログラムストリームの場合と同様である。

20

【0015】

MPEGビデオにはGOP(Group of Pictures : グループオブピクチャ)という構造があり、この単位で独立して、すなわち該当するGOPをデコードするのに前のGOPに属するピクチャを必要としないように、エンコーディングを行うことができる。よって複数のビデオストリームがある時、GOPを切り替えの単位として複数のビデオストリームを切り替えることができる。

【0016】

いま、上記の条件、すなわちビデオストリームがGOPで閉じている条件でエンコードされている2種類の異なるプログラムストリームがあって、それぞれを独立に多重化を行う場合を考える。ただし、プログラムストリームの切り替えを可能とするために、GOPの境界が同一のビデオパケット中に含まれることがないように制約を与える。

30

【0017】

図8は、このような条件で2本のプログラムストリームのそれぞれを独立に多重化する場合の例及びその2つのプログラムストリームを選択的に切り換えて出力する場合の例を示している。図8の(a)に示すように、プログラムストリームPS0のパックPK0とパックPK1にビデオストリームV0のGOP0のデータが、プログラムストリームPS0のパックPK2とパックPK3にビデオV0のGOP1のデータが多重化されている。また、図8の(b)に示すように、プログラムストリームPS1のパックPK0、PK1、PK2にビデオV1のGOP0のデータが、プログラムストリームPS1のパックPK3にビデオV1のGOP1のデータが多重化されている。

40

【0018】

これらの図8の(a)、(b)に示すような別々に多重化された2本のプログラムストリームが1つの記録媒体に記録されていて、例えば図6の読取装置10が選択的にプログラムストリームをパック単位で切り替えて出力することができるシステムを想定するとき、上述したGOP(グループオブピクチャ)の独立性により、プログラムストリームを切り替え点において切り替えたとき、ビデオ再生を途切れさせることなく連続的に再生することが可能となる。

【0019】

例えば、図8の(c)に示すように、プログラムストリームPS0のパックPK0、PK1を読み出した後、続けてプログラムストリームPS1のパックPK3を読み出せば、図

50

6のビデオバッファ12には、ビデオV0のGOP0のデータが入力された後、ビデオV1のGOP1のデータが入力されるので、ビデオV0からビデオV1に映像が切り替わっても連続的に再生することが可能になる。この例では2本のプログラムストリームが記録媒体に記録されているとしたが、2本以上でも同様である。以下これらのGOPの切替え点のパックをエントリーポイントと呼ぶ。

【0020】

【発明が解決しようとする課題】

ところで、記録媒体に複数のプログラムストリームが記録されていて、読取装置がエントリーポイントにおいて読み出すプログラムストリームを選択的に切り替える機能を持っているとき、媒体に記録される複数のプログラムストリームの多重化を、各々独立に通常の方法で行うと、デコーダで正しくデコードできない場合が生じる。これは次の2つの理由による。

10

【0021】

理由1．SCR（システムクロックリファレンス）の不整合

パックヘッダにエンコードされるSCRは、デコーダへの読み出し開始時間を示すので、読み出されてデコーダに入力される2つの隣接するパックについて、

（後のパックにエンコードされているSCR）

（前のパックにエンコードされているSCR）

+（前パックの転送時間）

すなわち、

20

（後のパックにエンコードされているSCR）

（前のパックにエンコードされているSCR）

+（前パックの大きさ）/（読み出しレート）

の条件を満足しなければならない。従って、プログラムストリームPS0をパックPK0、PK1、PK2、PK3、...と順に読み出す場合には上記の条件が成り立っても（個々のプログラムストリームは上記条件が成り立つように多重化されている）、図8の(c)に示すように、プログラムストリームPS0をパックPK0、PK1と読み出した後に、エントリーポイントでプログラムストリームを切り替えて、プログラムストリームPS1のパックPK3を読み出してデコーダに入力しようとする、プログラムストリームPS0とプログラムストリームPS1の多重化はそれぞれ独立に行われているので、上記の条件を満足できなくなる、すなわち前のパックを読み終わったときに、STC（システムタイムクロック）が後のパックにエンコードされているSCR（システムクロックリファレンス）の値より大きくなり、後のパックを読み出すことができなくなることがある。

30

【0022】

理由2．バッファの破綻

読取装置が読み出すプログラムストリームの切り替えを行うと、結果としてデコーダバッファの破綻（オーバーフロー、アンダーフロー）が生じる可能性がある。

【0023】

この理由2を、図9を参照しながら説明する。図9は、ビデオデコーダバッファにおけるデータの占有量の遷移を表している。ここで図9の(a)は、例えば図8の(a)に示すようなプログラムストリームPS0をパックPK0、PK1、PK2、PK3、...と順に読み出した時のバッファの状態であり、(ア)はビデオV0のGOP0のデータ、(イ)はビデオV0のGOP1のデータである。図9の(b)は、例えば図8の(b)に示すようなプログラムストリームPS1をパックPK0、PK1、PK2、PK3、...と順に読み出した時のバッファの状態であり、(ウ)はビデオV1のGOP0のデータ、(エ)はビデオV1のGOP1のデータである。図9の(a)、(b)ともに、それぞれ連続したプログラムストリームであるから、当然バッファは破綻することなく多重化されているはずである。ところが、この例のように多重化されたプログラムストリームを、例えば図8の(c)に示すように、読取装置がプログラムストリームPS0のパックPK0、PK1と読んだ後、ビデオを切り替えるために、プログラムストリームPS1のパックPK3を

40

50

読んだとすると、バッファにはビデオV0のGOP0のデータが入力された後ビデオV1のGOP1のデータが供給されるので、バッファの占有量の遷移は、図9の(c)のようになる。ここで(オ)はビデオV0のGOP0のデータであり、(カ)はビデオV1のGOP1のデータである。

【0024】

ビデオV1のGOP1のデータをデコードするときに、読み出しはSCR(システムクロックリファレンス)によって、バッファからの引き抜きはDTS(デコーディングタイムスタンプ)によって決まり、データの入力および引き抜きタイミングは(カ)と同様になるので、図9の(c)のようにバッファのオーバーフローが発生してしまう。

【0025】

本発明は、上述した実情に鑑みてなされたものであり、アクセスユニットの読みだしタイミングとそのデコードされた後のバッファの状態を複数のプログラムストリームについて同じにすることにより、複数のプログラムストリームをエン트리ポイントにおいて切り替えて読み出したときに、SCRの不整合が発生することなく、またバッファの破綻が起こらないようなプログラムストリームを生成し得るような符号化装置および方法、伝送方法、並びにこれらの方法や装置により得られたデジタル信号が記録されたデジタル信号記録媒体の提供を目的とする。

【0026】

【課題を解決するための手段】

本発明は、上述したような課題を解決するために、アクセスユニット単位で選択的に切り替えられる複数のビデオデータを可変長符号化して得られる複数のビデオストリームから、アクセスユニット毎のデコード時刻と同一のデコード時刻に対するアクセスユニットの大きさをそれぞれ検出し、検出された複数のアクセスユニットの大きさをデコード時刻毎に比較して、アクセスユニットの最大値を仮想アクセスユニットとし、仮想アクセスユニットから構成される仮想ビデオストリームを構成し、ビデオストリームのバッファ占有量が仮想ビデオストリームのバッファ占有量以下となるように、ビデオストリームのアクセスユニットの大きさと仮想アクセスユニットの大きさとを差分を調整しながら、複数のビデオストリームをパケット化することを特徴とする。

【0027】

すなわち、本発明は、複数のデジタル信号のビットストリームの符号化単位であるアクセスユニットの大きさとデコード時刻とを検出し、デコード時刻毎に、得られる複数のアクセスユニットの大きさを比較して、それらの最大値を選択し、デコード時刻においてアクセスユニットの大きさが上記選択された最大値に等しくなるような仮想ストリームを構成し、各々のデジタル信号のビットストリームを符号化する際に、アクセスユニットが上記仮想ストリームのアクセスユニットの大きさに満たないときに、その差分と大きさの等しいパディングパケットを用いてパケット化することを特徴とする。この場合、上記パケット化する際、アクセスユニットが仮想的なビデオストリームのアクセスユニットの大きさに満たない場合には、その差分と大きさの等しいパディングパケットをパケット化するが、あるいは差分の大きさがパックの大きさよりも大きいときは何もパケット化しないことが好ましい。

【0028】

また本発明は、さらに、パケット化して得られるストリームを1つのチャンネルとして複数のチャンネルを持つトランスポートストリームに符号化して伝送することを特徴とする。

【0029】

具体的には、複数のいわゆるMPEGビデオストリームのアクセスユニットの大きさとデコード(表示)時刻を検出する手段と、デコード時刻毎に各々のビデオストリームに属するアクセスユニットの最大値を選択する手段と、デコード時刻においてアクセスユニットの大きさが選ばれた最大値に等しくなるような1つの仮想的なビデオストリームを構成する手段と、この仮想的なビデオストリームをあたかも実在するビデオストリームのように扱いデコーダへの供給時刻(クロックリフレンス)およびパケットの大きさ決定する手段

10

20

30

40

50

と、実際に各々のビデオストリームをパケット化する際、アクセスユニットが仮想的なビデオストリームのアクセスユニットの大きさに満たない場合には、その差分と大きさの等しいパディングパケットをパケット化する手段と、差分の大きさがパックの大きさよりも大きいときはなにもパケット化しない手段をもつ。

【0030】

また、本発明の他の具体的な構成としては、複数のMPEGビデオストリームのアクセスユニットの大きさとデコード（表示）時刻を検出する手段と、デコード時刻毎に、得られる複数のアクセスユニットの大きさを比較して、それらの最大値を選択する手段と、デコード時刻においてアクセスユニットの大きさが選ばれた最大値に等しくなるような1つの仮想的なビデオストリームを構成する手段と、ビデオストリームとしてこの仮想的なビデオストリームをあたかも実在するビデオストリームのように扱い、オーディオその他のストリームと共に、それらのデコーダへの供給時刻（クロックリファレンス）およびパケットの大きさ決定する手段と、実際に各々のビデオストリームをパケット化する際、アクセスユニットが仮想的なビデオストリームのアクセスユニットの大きさに満たない場合には、その差分と大きさの等しいパディングパケットをパケット化する手段と、パケット化して得られるストリームを1つのチャンネルとして、複数のチャンネル持つMPEGトランスポートストリームに符号化して伝送する手段をもつ。

10

【0031】

【発明の実施の形態】

以下、本発明に係る実施の形態について、図面を参照しながら説明する。

20

【0032】

図1は、本発明に係るデジタル信号符号化装置の実施の形態の概略構成を示している。この第1の実施の形態においては、同一シーンを異なるカメラアングルから撮影した複数のビデオ、例えば3種類のビデオストリーム $V S_0 \sim V S_2$ についての符号化を考える。

【0033】

同じシーンを撮影したものであるから、当然オーディオや字幕データなどのような、上記ビデオストリーム $V S_0 \sim V S_2$ と一緒に多重化すべきデータ（オーディオストリームASや、その他のストリームTS等）は、3つのビデオストリーム $V S_0 \sim V S_2$ に対して全く同一のものを使うことができるため、それぞれ1つのストリームが3つのビデオストリーム $V S_0 \sim V S_2$ と多重化される。最終的に得られるプログラムストリーム（ $P S_0 \sim P S_2$ ）を復号側においてGOP（グループオブピクチャ）毎に切り替えて再生するとき、切替え点において連続的にビデオが再生可能であるためには、切替え点の時刻とフィールドパリティ（トップフィールドで始まるかボトムフィールドで始まるか）がビデオストリーム毎に同一である必要がある。この条件を満たすために、ここではビデオのピクチャタイプ、トップフィールドファーストフラグ、リピートファーストフィールドフラグはカメラアングルの異なるビデオで同じようにコーディングを行うが、実際には必ずしも同一にする必要はない。

30

【0034】

なお、トップフィールドファーストフラグ及びリピートファーストフラグは、MPEG2において定義されているフラグであり、トップフィールドファーストフラグはインターレースフレームを画面表示する場合、トップフィールドとボトムフィールドのどちらを最初に出力するかを示すフラグである。また、リピートファーストフラグは、例えば映画等のフィルムソース（24コマ）をインターレースビデオ信号（30フレーム）に変換する場合、3:2プルダウン手法を用いて生成された信号を符号化する際に除去される冗長フィールドを示すフラグのことである。

40

【0035】

本実施の形態のシステムでは、異なるカメラアングルに対するビデオストリームをデコードするとき、どのビデオストリームをデコードしてもバッファからアクセスユニットを引き抜いた（デコード）後のビデオバッファの状態が同じになるように多重化することが可能である。これにより、GOP単位でプログラムストリームを切り替えたときにバッファ

50

の状態を同じにすることができ、その結果、バッファの破綻無くシームレスなビデオ再生が可能となっている。以下、本実施の形態のシステムを説明する。

【0036】

図1において、アクセスユニット検出器50, 51, 52a~52cは、各エレメンタリーストリームのアクセスユニットの大きさとそのアクセスユニットのデコード時刻DTS(表示時刻とデコード時刻が異なる場合は表示時刻PTSも)を検出する。エレメンタリーストリームのエンコーダと多重化器が一体のシステムでは、これらの情報はエンコーダが出力できる可能性が大きく、その時はエンコーダの出力値を用いることができる。図1の例では、オーディオストリームASがアクセスユニット検出器50に送られ、異なるカメラアングルから撮影されエンコードされた例えば3種類のビデオストリームVS₀, VS₁, VS₂がそれぞれアクセスユニット検出器52a, 52b, 52cに送られ、その他のストリームTSがアクセスユニット検出器51に送られる。そして、各アクセスユニット検出器50, 51, 52a~52cにおいて、それぞれのエレメンタリーストリームのアクセスユニットの大きさとデコード時刻DTS(必要に応じてPTS)等が検出される。

10

【0037】

最大値検出器53は、アクセスユニット検出器52a, 52b, 52cからの、3種類のビデオストリームVS₀, VS₁, VS₂より各々得られるアクセスユニットの大きさを比較し、アクセスユニットの最大値をデコード時刻毎に選択する。

【0038】

図2にこの選択の様子を示す。図2ではわかりやすいように各々のビデオストリームVS₀, VS₁, VS₂に対するアクセスユニットを時間的にずらして書いているが、これらは同一のデコード時刻t_n(n=1, 2, ...)に対するアクセスユニットの大きさを表している。デコード時刻t_N毎にアクセスユニットの最大値が仮想的なビデオストリームPVSのアクセスユニットとして選ばれている。仮想的なビデオストリームPVSは、このデコード間隔とアクセスユニットの大きさをもつものとする。

20

【0039】

スケジューラ54は、この仮想的なビデオストリームPVSとその他の多重化すべきエレメンタリーストリーム、すなわちオーディオストリームASと、字幕データなどの他のストリームTSについての、アクセスユニットの大きさとデコード時刻の情報を受け取って、パケット化すべきエレメンタリーストリームの種類、パケットの大きさ、パックに付けるSCR(システムクロックリファレンス)などの制御情報を出力する。このスケジューラ54は通常の多重化方式において用いるものと同じでよい。例えば、このスケジューリングの技術として、本件出願人が先に特願平7-341951号の明細書および図面に開示した技術等を用いることができる。また、他のスケジューリングの技術を用いてもよい。

30

【0040】

パケット化器55a, 55b, 55cは、スケジューラ54によって出力された制御情報を用いて、エレメンタリーストリームのパケット化を行う。ただしスケジューラ54は仮想的なビデオストリーム(仮想ビデオ)PVSに対してスケジューリングを行っているので、得られる制御情報をそのまま用いて実際のビデオストリーム(実ビデオ)をパケット化しても、アクセスユニットの引き抜き時におけるバッファの占有量は必ずしも同じにならない。ところが、

40

(実ビデオのアクセスユニットの大きさ)

(仮想ビデオのアクセスユニットの大きさ)

が常に成り立つので、パケット化を行う際、ビデオのアクセスユニットの大きさが仮想ビデオのアクセスユニットの大きさに満たない場合にパディングパケットをパケット化すると、バッファから引き抜かれた(デコードされた)後のバッファの状態は、異なるアングルのビデオによらず一定になる。

【0041】

図3を用いてこれらのパケット化器55a~55cの動作を説明する。

50

【 0 0 4 2 】

図 3 の (a) は仮想的なビデオのアクセスユニット PVAU の大きさを表し、図 3 の (b) は実際に多重化すべきビデオのアクセスユニット VAU の大きさを表す。図 3 の (c) の実線は、多重化スケジューラによって、仮想ビデオがパケットにどのように分割されるかを示している。このスケジューラの出力情報を用いて実際にビデオパケット VP をパケット化するときには、仮想ビデオと実際のビデオストリームの各アクセスユニット PVAU と VAU との大きさに差があるため、図 3 の斜線 (ハッチング) が施された部分を調整しなければならない。パケット化器 55 a ~ 55 c は、図 3 の (c) で斜線が施された部分において、図 3 の (d) に示すように、パディングパケット PP を出力する機能を持つ。ビデオストリームが図 3 の (d) のようにパケット化されることにより、他のエレメンタリーストリームを多重化して生成されるプログラムストリームは図 3 の (e) のようになる。この例では、各パケット化器においてビデオ、オーディオ、字幕データそれぞれ 1 ストリームについての多重化を行った例を示している。

10

【 0 0 4 3 】

また、このように仮想ビデオと実際のビデオストリームのアクセスユニットの差分を調整するとき、パディングパケット PP それ自体でパックを構成している場合には、このパディングパケット PP を転送しなくても、各エレメンタリーストリームのバッファの遷移には全く影響を与えない。多重化におけるオーバーヘッドを減らすため、すなわち無駄なデータの蓄積を防ぐため、このような場合には、パケット化器 55 a ~ 55 c はパディングパケット PP をパケット化しない機能をもつ。本例において、図 3 の (d) の x で示す部分のパディングパケット PP は、それ自体がパックを構成しているので、図 3 の (e) のプログラムストリーム中にはパケット化されていない。ただし、パケット化器 55 a ~ 55 c のハードウェアを簡略化したいようなときには、パディングパケット PP を全てパケット化しても差し支えないが、多重化のオーバーヘッドは増える。

20

【 0 0 4 4 】

本例の多重化器によって多重化を行ったときのバッファの挙動は、例えば図 4 のようになる。この図 4 において、点線は仮想的なビデオに対してスケジューリングを行ったときのビデオバッファの状態の遷移を表す。実線は、仮想的なビデオストリームに対してスケジューリングを行って得られる情報をもとに実際のビデオストリームを多重化したときのビデオバッファの状態を表している。また、時刻 t_1, t_2, t_3, \dots は、デコード時刻を表している。

30

【 0 0 4 5 】

この図 4 中の x 印までは、仮想ビデオと同じタイミングでビデオバッファへのロードが行われる。一方、図 4 中の x 印から 印までの区間では、上記したように仮想ビデオと実ビデオ (実際のビデオストリーム) のアクセスユニットとの差分がパディングパケットで置き換えられるので、ビデオバッファへのロードは行われず、次のアクセスユニットのロードの開始点 (図中 印のところ)、あるいはデコード時刻 t_n ($n = 1, 2, \dots$) になるまでバッファのデータ量は変わらない。

【 0 0 4 6 】

この結果、全ての時刻において、
 (実際のビデオのバッファ占有量)
 (仮想的なビデオのバッファ占有量)
 が常に成り立つ。

40

【 0 0 4 7 】

したがって、異なるカメラアングルから撮影されたビデオデータを符号化し、本方法により多重化を行って生成される複数のプログラムストリームは、全ての時刻によって上記の関係式を満足する。よって、仮想ビデオがバッファの破綻無く多重化されているならば、これらのプログラムストリームをエンターポイントで切り替えてデコードしても、バッファの破綻は起きない。

【 0 0 4 8 】

50

そして、これらのプログラムストリーム PS_0 、 PS_1 及び PS_2 が、例えばディスク等の記録媒体 56 に記録される。

【0049】

次に、2つの異なるカメラアングルから撮影されエンコードされたビデオを多重化する様子を図5に表す。図5の(a)はビデオV0のアクセスユニットの大きさを表し、(b)はビデオV1のビデオアクセスユニットの大きさを表す。この例では、GOP0、GOP1両方とも4枚のアクセスユニットから構成されている。GOPは、MPEG2に定義されているグループオブピクチャの意味である。図5の(c)は、アクセスユニットの大きさのどちらか大きな方を選択することによって得られる、仮想ビデオのアクセスユニットの大きさを表す。図5の(d)は、仮想ビデオに対してスケジューリングを行った結果、仮想ビデオがどのようにパッケージサイズされるかを示している。実際のビデオを多重化するときは、図5の(e)、(f)に示すように、仮想ビデオとの差分が生じるときはパディングを行う。ただし、仮想ビデオの packets が3つ以上の packets に分割されるときには、図5の(e)、(f)のようにパディング packets を後詰めして2つの packets にしてしまってもよい。スケジューリングの結果として、仮想ビデオに対するプログラムストリームが図5の(g)のように多重化されるとすると、本実施の形態を用いて実ビデオを多重化してできるプログラムストリームは図5の(h)、(i)のようになる。

10

【0050】

こうしてできるプログラムストリームは、図5の(h)、(i)のどの矢印に従って切り替えてデコーダに入力しても、上に示したようにバッファの破綻を引き起こすことなく、映像・音声のシームレスな再生が可能になる。なお、この例では2つのビデオを考えたが3つ以上のビデオに対しても同様である。

20

【0051】

以上説明した例においては、プログラムストリームに対して本発明の実施の形態の方法を適用したが、伝送用途に使われるトランスポートストリームに対しても、本方法は有効である。トランスポートストリームでは、一本のストリームが複数のチャンネルで構成され、各チャンネルが一つの独立したストリームに相当している。また各チャンネルは、各々独立してタイムベースをもつことができる。したがって、上記の複数プログラムストリームに対する方法をそのまま各チャンネルに対して行い、複数チャンネルのストリームの一本のトランスポートストリームに多重化し、その多重化されたトランスポートストリームを伝送する。それにより、複数のチャンネルを切り替えたときに、異なるカメラアングルから撮影された映像にシームレスに切り替えることが可能になる。

30

【0052】

また、上述した実施の形態では、異なるカメラアングルから撮影された映像を多重化することを考えたが、全く内容の関係の無い画像に関しても本発明は適用可能である。また、本発明は、ビデオにとどまらずオーディオやその他のデータについても適用可能である。

【0053】

【発明の効果】

本発明によれば、複数のデジタル信号のビットストリームの符号化単位であるアクセスユニットの大きさとデコード時刻とを検出し、デコード時刻毎に、得られる複数のアクセスユニットの大きさを比較して、それらの最大値を選択し、デコード時刻においてアクセスユニットの大きさが上記選択された最大値に等しくなるような仮想ストリームを構成し、各々のデジタル信号のビットストリームをパッケージ化する際に、アクセスユニットが上記仮想ストリームのアクセスユニットの大きさに満たないときに、その差分と大きさの等しいパディング packets を用いてパッケージ化しているため、複数のデジタル信号のビットストリームを切り替えて復号しても、復号バッファを破綻させずに連続的に再生することができる。

40

【0054】

すなわち、本発明によれば、複数の多重化ストリームを切り替えて再生する機能をもつシステム用に多重化を行う際、エレメンタリーストリームからアクセスユニットの大きさと

50

デコード時刻を求め、各デコード時刻毎にアクセスユニットの最大値を選び、その最大値をアクセスユニットの大きさとする仮想的なビデオストリームを考えて多重化のスケジューリングを行い、実際のビデオを多重化するときには仮想ビデオと実ビデオの差分が生じる場合はパディングを行うことにより、複数の多重化ストリームを切り替えて再生するときには、バッファの破綻を起こさず、かつ映像や音声を途切れさせることが無いようにできる。ここで、この複数の多重化ストリームは蓄積媒体の各部に記録されていても良いし、複数ストリームをさらに一本のトランスポートストリームにまとめられて放送に用いられても良い。

【図面の簡単な説明】

【図 1】本発明の実施の形態となるデジタル信号符号化装置の概略構成を示すブロック図である。 10

【図 2】仮想ビデオのアクセスユニットの決定を説明するための図である。

【図 3】上記実施の形態によって生成されるプログラムストリームの一例を説明するための図である。

【図 4】上記実施の形態によって生成されるプログラムストリームをデコードしたときのバッファの挙動を説明するための図である。

【図 5】上記実施の形態を 2 つの異なるカメラアングルに対するビデオデータに対して適用したときに得られるプログラムストリームを説明するための図である。

【図 6】いわゆる M P E G 規格におけるシステムターゲットデコーダ (S T D) の概略構成を示すブロック図である。 20

【図 7】いわゆる M P E G 規格におけるプログラムストリームおよびトランスポートストリームの構造の一例を示す図である。

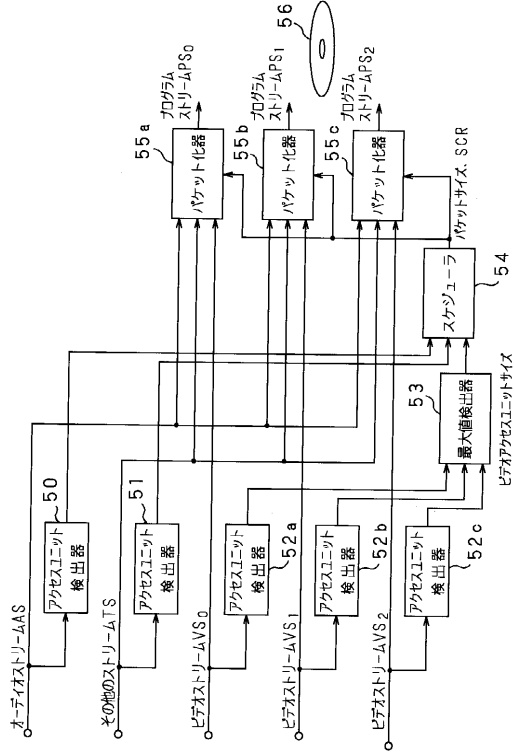
【図 8】2 本のプログラムストリームをそれぞれ独立に多重化したときに得られるストリーム及び 2 本のプログラムストリームを切り替えたときに得られるストリームの一例を説明するための図である。

【図 9】プログラムストリーム切り替え時におけるバッファの挙動を説明するための図である。

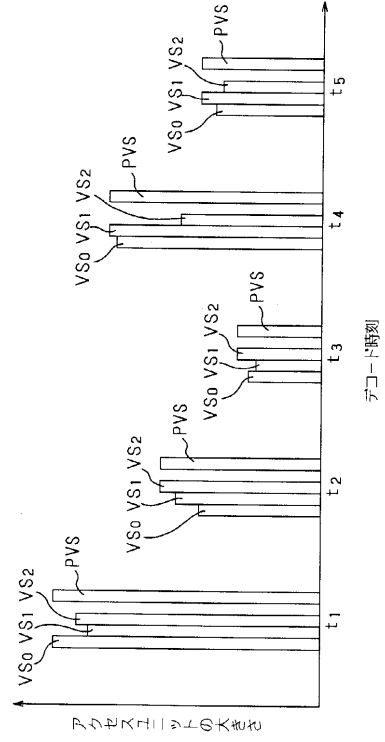
【符号の説明】

5 0 , 5 1 , 5 2 a ~ 5 2 c アクセスユニット検出器、 5 3 最大値検出器、 5 4 スケジューラ、 5 5 a ~ 5 5 c パケット化器 30

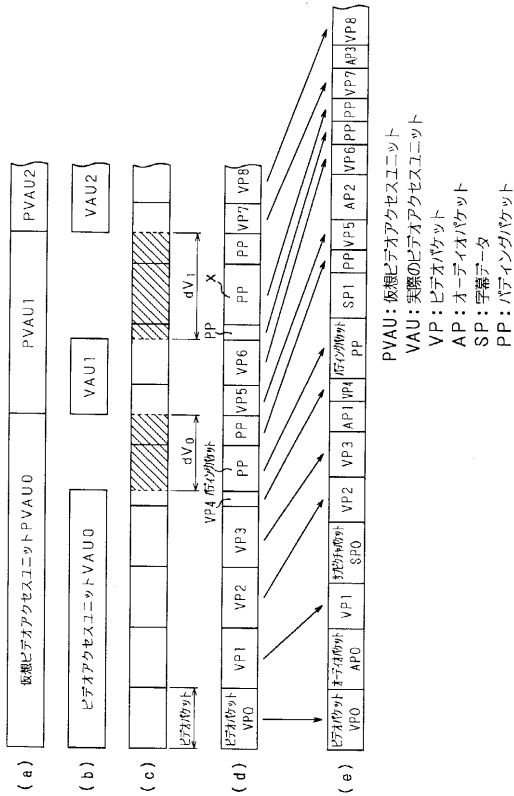
【 図 1 】



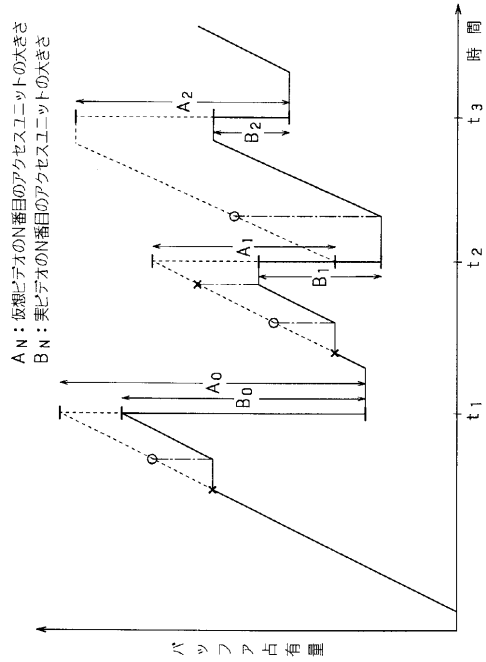
【 図 2 】



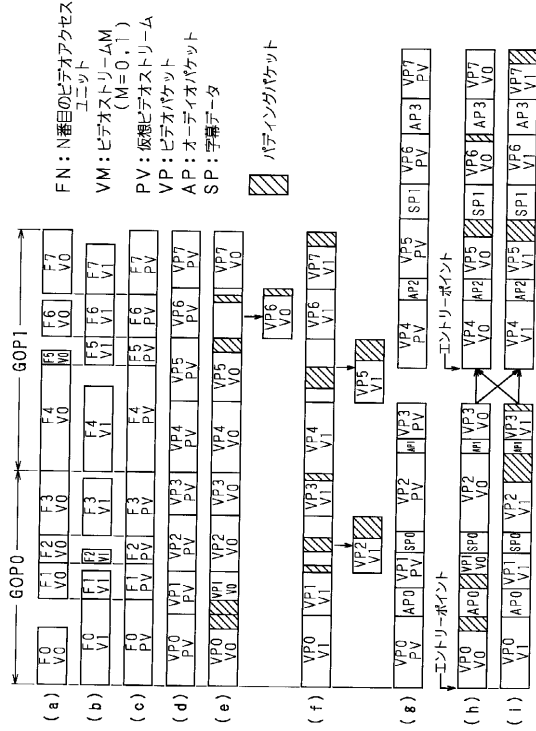
【 図 3 】



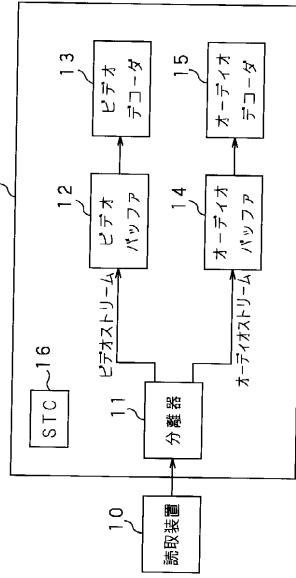
【 図 4 】



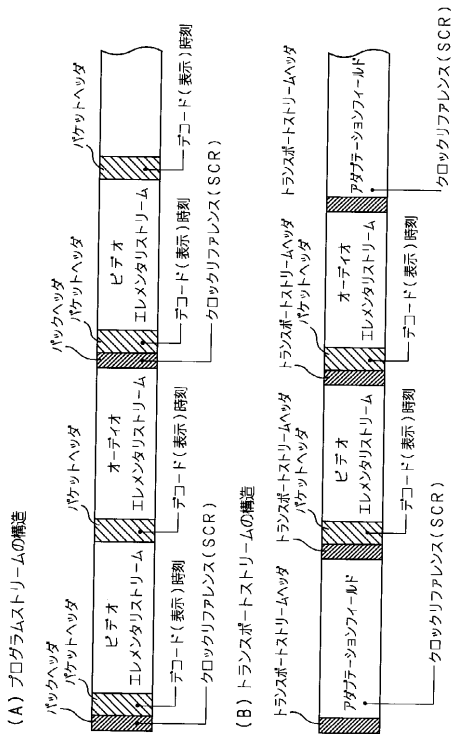
【 図 5 】



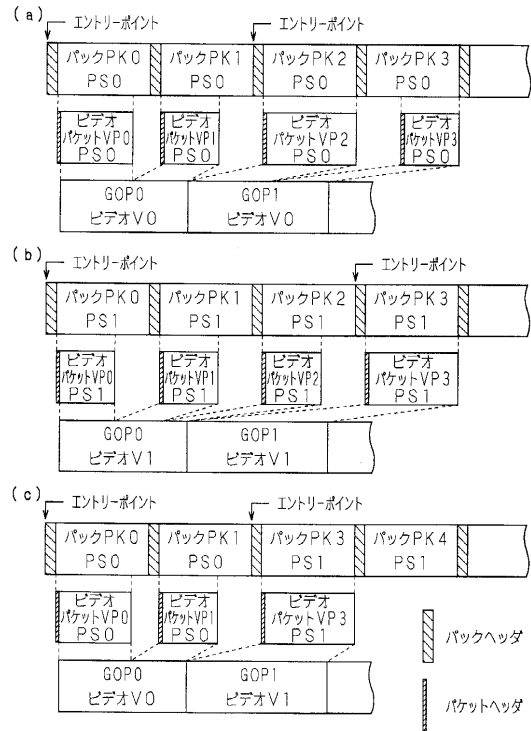
【 図 6 】



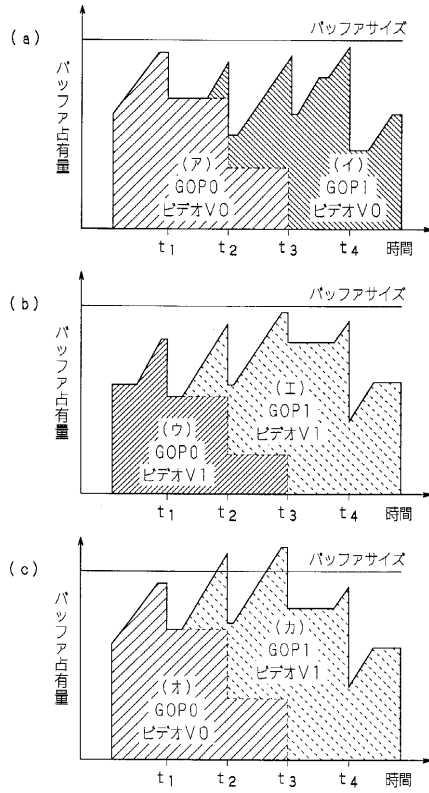
【 図 7 】



【 図 8 】



【 図 9 】



フロントページの続き

- (72)発明者 大石 義明
東京都品川区北品川6丁目7番35号 ソニー株式会社内
- (72)発明者 根岸 慎治
東京都品川区北品川6丁目7番35号 ソニー株式会社内

審査官 齋藤 哲

- (56)参考文献 特開平10-027461(JP,A)
特開平07-284096(JP,A)
特開平08-098160(JP,A)
国際公開第97/013367(WO,A1)
国際公開第95/012197(WO,A1)

(58)調査した分野(Int.Cl., DB名)

G11B 20/10
H04J 3/00
H04N 5/92
H04N 7/26