

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4506832号
(P4506832)

(45) 発行日 平成22年7月21日(2010.7.21)

(24) 登録日 平成22年5月14日(2010.5.14)

(51) Int.Cl.		F I			
HO4N	5/91	(2006.01)	HO4N	5/91	Z
HO4N	5/76	(2006.01)	HO4N	5/76	B
HO4N	5/225	(2006.01)	HO4N	5/225	F
G11B	27/00	(2006.01)	G11B	27/00	A

請求項の数 6 (全 25 頁)

(21) 出願番号	特願2007-341028 (P2007-341028)	(73) 特許権者	000002185
(22) 出願日	平成19年12月28日 (2007.12.28)		ソニー株式会社
(65) 公開番号	特開2009-164828 (P2009-164828A)		東京都港区港南1丁目7番1号
(43) 公開日	平成21年7月23日 (2009.7.23)	(74) 代理人	100112955
審査請求日	平成21年1月26日 (2009.1.26)		弁理士 丸島 敏一
		(72) 発明者	磯部 幸雄
			東京都港区港南1丁目7番1号 ソニー株式会社内
		(72) 発明者	前 篤
			東京都港区港南1丁目7番1号 ソニー株式会社内
		(72) 発明者	長尾 研一郎
			東京都港区港南1丁目7番1号 ソニー株式会社内

最終頁に続く

(54) 【発明の名称】 画像処理装置、インデックス生成装置、撮像装置、インデックス生成方法およびプログラム

(57) 【特許請求の範囲】

【請求項1】

動画データに含まれる複数の画像について属性情報を検出する属性情報検出手段と、
前記属性情報と前記複数の画像のうち当該属性情報に対応する画像の前記動画データにおける位置とを関連付けたインデックスデータを生成するインデックスデータ生成手段と

前記動画データおよび前記インデックスデータを含むストリームデータを生成するストリームデータ生成手段と

を具備し、

前記属性情報は、当該属性情報に対応する画像の特徴量であり、

前記属性情報検出手段は、前記特徴量として前記複数の画像に含まれる顔画像の数を検出し、

前記ストリームデータ生成手段は、ビデオオブジェクトユニットを単位として前記ストリームデータを生成し、前記インデックスデータを最終のビデオオブジェクトユニットに格納する

画像処理装置。

【請求項2】

動画データに含まれる複数の画像の位置と当該画像に関する属性情報とを関連付けたインデックスデータを前記動画データとともに格納するストリームデータからインデックス表示画面を生成するインデックス生成装置であって、

前記属性情報を解析してインデックス表示の対象となる画像を特定する属性情報解析手段と、

前記特定された画像の前記動画データにおける位置を生成する位置生成手段と、

前記生成された位置に基づいて前記動画データの代表画像を生成する代表画像生成手段と、

前記代表画像に基づいて前記インデックス表示画面を生成するインデックス表示画面生成手段と

を具備し、

前記属性情報は、当該属性情報に対応する画像の特徴量であり、

前記特徴量は、前記複数の画像に含まれる顔画像の数を示し、

前記属性情報解析手段は、前記複数の画像に含まれる顔画像の数が変化する画像を前記インデックス表示の対象となる画像として特定する

インデックス生成装置。

【請求項3】

被写体を撮像して動画データを生成する撮像手段と、

前記動画データに含まれる複数の画像について属性情報を検出する属性情報検出手段と

、

前記属性情報と前記複数の画像のうち当該属性情報に対応する画像の前記動画データにおける位置とを関連付けたインデックスデータを生成するインデックスデータ生成手段と

、

前記動画データおよび前記インデックスデータを含むストリームデータを生成するストリームデータ生成手段と

を具備し、

前記属性情報は、当該属性情報に対応する画像の特徴量であり、

前記属性情報検出手段は、前記特徴量として前記複数の画像に含まれる顔画像の数を検出し、

前記ストリームデータ生成手段は、ビデオオブジェクトユニットを単位として前記ストリームデータを生成し、前記インデックスデータを最終のビデオオブジェクトユニットに

格納する

撮像装置。

【請求項4】

被写体を撮像して動画データを生成する撮像手段と、

前記動画データに含まれる複数の画像について属性情報を検出する属性情報検出手段と

、

前記属性情報と前記複数の画像のうち当該属性情報に対応する画像の前記動画データにおける位置とを関連付けたインデックスデータを生成するインデックスデータ生成手段と

、

前記動画データおよび前記インデックスデータを含むストリームデータを生成するストリームデータ生成手段と、

前記ストリームデータに含まれる属性情報を解析してインデックス表示の対象となる画像を特定する属性情報解析手段と、

前記特定された画像の前記動画データにおける位置を生成する位置生成手段と、

前記生成された位置に基づいて前記動画データの代表画像を生成する代表画像生成手段と、

前記代表画像に基づいて前記インデックス表示画面を生成するインデックス表示画面生成手段と

を具備し、

前記属性情報は、当該属性情報に対応する画像の特徴量であり、

前記特徴量は、前記複数の画像に含まれる顔画像の数を示し、

前記属性情報解析手段は、前記複数の画像に含まれる顔画像の数が変化する画像を前記

10

20

30

40

50

インデックス表示の対象となる画像として特定する
撮像装置。

【請求項 5】

動画データに含まれる複数の画像について属性情報を検出する属性情報検出手順と、
前記属性情報と前記複数の画像のうち当該属性情報に対応する画像の前記動画データに
おける位置とを関連付けたインデックスデータを生成するインデックスデータ生成手順と

、
前記動画データおよび前記インデックスデータを含むストリームデータを生成するスト
リームデータ生成手順と、

前記ストリームデータに含まれる属性情報を解析してインデックス表示の対象となる画
像を特定する属性情報解析手順と、

前記特定された画像の前記動画データにおける位置を生成する位置生成手順と、

前記生成された位置に基づいて前記動画データの代表画像を生成する代表画像生成手順
と、

前記代表画像に基づいて前記インデックス表示画面を生成するインデックス表示画面生
成手順と

を具備し、

前記属性情報は、当該属性情報に対応する画像の特徴量であり、

前記特徴量は、前記複数の画像に含まれる顔画像の数を示し、

前記属性情報解析手順は、前記複数の画像に含まれる顔画像の数が変化する画像を前記
インデックス表示の対象となる画像として特定する

インデックス生成方法。

【請求項 6】

動画データに含まれる複数の画像について属性情報を検出する属性情報検出手順と、

前記属性情報と前記複数の画像のうち当該属性情報に対応する画像の前記動画データに
おける位置とを関連付けたインデックスデータを生成するインデックスデータ生成手順と

、
前記動画データおよび前記インデックスデータを含むストリームデータを生成するスト
リームデータ生成手順と、

前記ストリームデータに含まれる属性情報を解析してインデックス表示の対象となる画
像を特定する属性情報解析手順と、

前記特定された画像の前記動画データにおける位置を生成する位置生成手順と、

前記生成された位置に基づいて前記動画データの代表画像を生成する代表画像生成手順
と、

前記代表画像に基づいて前記インデックス表示画面を生成するインデックス表示画面生
成手順と

をコンピュータに実行させるプログラムであって、

前記属性情報は、当該属性情報に対応する画像の特徴量であり、

前記特徴量は、前記複数の画像に含まれる顔画像の数を示し、

前記属性情報解析手順は、前記複数の画像に含まれる顔画像の数が変化する画像を前記
インデックス表示の対象となる画像として特定する

プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、画像処理装置に関し、特に動画データに含まれる複数の画像について属性情
報を検出し、または、属性情報に基づいてインデックス表示画面を生成する画像処理装置
、インデックス生成装置、撮像装置、および、これらにおける処理方法ならびに当該方法
をコンピュータに実行させるプログラムに関する。

【背景技術】

10

20

30

40

50

【0002】

近年、動画を記録するためのデジタルビデオカメラが普及している。デジタルビデオカメラでは、撮像された動画データは符号化されて、DVD (Digital Versatile Disk)、ハードディスク、フラッシュメモリ等の記録媒体に記録される。記録媒体に記録された動画データは、デジタルビデオカメラまたは他の再生機器において再生表示される。

【0003】

動画データを再生表示する際、先頭から全てを再生するのではなく、途中の特定のシーンから再生することをユーザが望む場合がある。その場合、撮像された動画データ全体から特定のシーンを検索するためには、検索に要する時間が必要であるとともに検索作業が煩雑であることが多い。

10

【0004】

そこで、動画データを再生する場合において、この動画データの各部におけるサムネイル画像を生成して、この生成されたサムネイル画像を用いて特定の記録位置から動画を再生させる技術が提案されている。例えば、記録媒体に記録されている動画データの複数のシーンから各シーンを示す画面のサムネイル画像を複数形成し、ユーザによって選択されたサムネイル画像に対応するシーンの先頭から動画の再生を開始する再生装置が提案されている(例えば、特許文献1参照。)

【特許文献1】特開平11-289517号公報(図7)

【発明の開示】

【発明が解決しようとする課題】

20

【0005】

上述の従来技術によれば、各シーンを示す画面のサムネイル画像に基づいて動画データの特定のシーンから再生表示することが可能になる。この各シーンに関する情報は、予め何らかの管理情報ファイルとして保存しておくことが望ましいが、独自のファイルとして保存した場合には動画データが装置間で転送された際にその情報を喪失してしまうおそれがある。

【0006】

本発明はこのような状況に鑑みてなされたものであり、動画データの転送後においてもその動画データの各シーンに関する情報を維持させることを目的とする。

【課題を解決するための手段】

30

【0007】

本発明は、上記課題を解決するためになされたものであり、その第1の側面は、動画データに含まれる複数の画像について属性情報を検出する属性情報検出手段と、上記属性情報と上記複数の画像のうち当該属性情報に対応する画像の上記動画データにおける位置とを関連付けたインデックスデータを生成するインデックスデータ生成手段と、上記動画データおよび上記インデックスデータを含むストリームデータを生成するストリームデータ生成手段とを具備する画像処理装置およびその方法ならびにプログラムである。これにより、属性情報と動画データ中の属性情報に対応する位置とを関連付けたインデックスデータをストリームデータに格納させるという作用をもたらす。

【0008】

40

また、この第1の側面において、上記属性情報は、当該属性情報に対応する画像の特徴量であってもよい。この場合において、上記属性情報検出手段は、上記特徴量として上記複数の画像に含まれる顔画像の数を検出してもよい。また、この場合において、上記ストリームデータ生成手段は、ビデオオブジェクトユニットを単位として上記ストリームデータを生成し、上記インデックスデータを最終のビデオオブジェクトユニットに格納するようにしてもよい。

【0009】

また、本発明の第2の側面は、被写体を撮像して動画データを生成する撮像手段と、上記動画データに含まれる複数の画像について属性情報を検出する属性情報検出手段と、上記属性情報と上記複数の画像のうち当該属性情報に対応する画像の上記動画データにお

50

る位置とを関連付けたインデックスデータを生成するインデックスデータ生成手段と、上記動画データおよび上記インデックスデータを含むストリームデータを生成するストリームデータ生成手段とを具備する撮像装置およびその方法ならびにプログラムである。これにより、属性情報と撮像された動画データ中の属性情報に対応する位置とを関連付けたインデックスデータをストリームデータに格納させるという作用をもたらす。

【0010】

また、本発明の第3の側面は、動画データに含まれる複数の画像の位置と当該画像に関する属性情報とを関連付けたインデックスデータを上記動画データとともに格納するストリームデータからインデックス表示画面を生成するインデックス生成装置であって、上記属性情報を解析してインデックス表示の対象となる画像を特定する属性情報解析手段と、
10
上記特定された画像の上記動画データにおける位置を生成する位置生成手段と、上記生成された位置に基づいて上記動画データの代表画像を生成する代表画像生成手段と、上記代表画像に基づいて上記インデックス表示画面を生成するインデックス表示画面生成手段とを具備することを特徴とするインデックス生成装置およびその方法ならびにプログラムである。これにより、ストリームデータに含まれるインデックスデータによって特定される画像に基づいてインデックス表示画面を生成させるという作用をもたらす。

【0011】

また、この第3の側面において、上記属性情報は、当該属性情報に対応する画像の特徴量であってよい。この場合において、上記特徴量は、上記複数の画像に含まれる顔画像の数を示し、上記属性情報解析手段は、上記複数の画像に含まれる顔画像の数が変化する画像
20
を上記インデックス表示の対象となる画像として特定するようにしてもよい。これにより、顔画像の数が変化する位置を区切りとしてインデックス表示画面を生成させるという作用をもたらす。

【0012】

また、本発明の第4の側面は、被写体を撮像して動画データを生成する撮像手段と、上記動画データに含まれる複数の画像について属性情報を検出する属性情報検出手段と、上記属性情報と上記複数の画像のうち当該属性情報に対応する画像の上記動画データにおける位置とを関連付けたインデックスデータを生成するインデックスデータ生成手段と、上記動画データおよび上記インデックスデータを含むストリームデータを生成するストリームデータ生成手段と、上記ストリームデータに含まれる属性情報を解析してインデックス
30
表示の対象となる画像を特定する属性情報解析手段と、上記特定された画像の上記動画データにおける位置を生成する位置生成手段と、上記生成された位置に基づいて上記動画データの代表画像を生成する代表画像生成手段と、上記代表画像に基づいて上記インデックス表示画面を生成するインデックス表示画面生成手段とを具備する撮像装置およびその方法ならびにプログラムである。これにより、撮像された動画データとともに属性情報と撮像された動画データ中の属性情報に対応する位置とを関連付けたインデックスデータをストリームデータに格納しておいて、インデックスデータから特定される画像に基づいてインデックス表示画面を生成させるという作用をもたらす。

【発明の効果】

【0013】

本発明によれば、動画データの転送後においてもその動画データの各シーンに関する情報を維持することができるという優れた効果を奏し得る。

【発明を実施するための最良の形態】

【0014】

次に本発明の実施の形態について図面を参照して詳細に説明する。ここでは、動画データの記録される記録媒体としてDVD-Video規格によるディスクを想定するが、他の記録媒体においても同様に適用可能である。

【0015】

図1は、DVD-Video規格によるディスクのデータ構造を示す図である。ディスクは円盤状の形状を有しており、内周側にリードイン領域711が設けられ、外周側にリ
50

ードアウト領域 7 1 5 が設けられる。そして、内周側から外周側に向かって順番に、ファイルシステム領域 7 1 2 と、管理情報領域 7 1 3 と、データ記録領域 7 1 4 とが設けられる。ファイルシステム領域 7 1 2 は、ファイルシステムの管理領域であり、ISO 9660 および UDF (Universal Disk Format) の規格に準拠するように設けられている。管理情報領域 7 1 3 は、ディスク上に記録されているビデオコンテンツ全体の管理情報を保持する領域である。データ記録領域 7 1 4 は、各ビデオタイトルセットの内容およびその制御情報を保持する領域である。なお、これら領域は、2048 バイトの論理セクタに区分される。

【0016】

管理情報領域 7 1 3 には、VMG (Video ManaGer) 7 3 0 が保持される。この VMG 7 3 0 は、VMGI (VMG Information) 7 3 1 と、VMGM_VOBS (Video Object Set for VMG Menu) 7 3 2 と、VMGI__BUP (VMGI for Back UP) 7 3 4 とを備えている。VMGI 7 3 1 は、後述するように、ディスク上に記録されているビデオコンテンツのタイトル毎の管理情報やトップメニューの制御情報などを保持する。VMGM_VOBS 7 3 2 は、トップメニューのデータを保持する。また、VMGI__BUP 7 3 4 は、VMGI 7 3 1 のバックアップ用コピーである。

【0017】

データ記録領域 7 1 4 には、VTS (Video Titles Set) 7 4 0 が少なくとも 1 つ保持される。各 VTS 7 4 0 は、それぞれ VTSI (VTS Information) 7 4 1 と、VTSM_VOBS (Video Object Set for VTS Menu) 7 4 2 と、VTS TT_VOBS (Video Object Set for Titles in a VTS) 7 4 3 と、VTSI__BUP (VTSI for Back UP) 7 4 4 とを備えている。VTSI 7 4 1 は、後述するように、そのビデオタイトルセットに含まれるタイトルのチャプター毎の管理情報や制御情報およびチャプターメニューの制御情報などを保持する。VTSM_VOBS 7 4 2 は、チャプターメニューのデータを保持する。VTS TT_VOBS 7 4 3 は、そのビデオタイトルセットに含まれるタイトルのデータを保持する。また、VTSI__BUP 7 4 4 は、VTSI 7 4 1 のバックアップ用コピーである。

【0018】

図 2 は、ビデオタイトルセット (VTS) におけるタイトルとチャプターとの関係を示す図である。上述のように、ビデオコンテンツの実データは VTS TT_VOBS 7 4 3 に記録される。各 VTS 7 4 0 には、最大 99 タイトルまで格納できる。但し、ディスク全体としても最大 99 タイトルまでしか格納できないため、この最大タイトル数は他の VTS の使用状況に影響を受けることになる。

【0019】

VTS 7 4 0 における各タイトルは、少なくとも 1 つのチャプターに区分される。各タイトルの最大チャプター数は 99 である。カムコードにおいては、記録開始から記録終了までの 1 回の記録単位がチャプターとして記録され、例えば、ディスクがイジェクトされた場合、タイトル内で 99 チャプターに達した場合、タイトル内で 99 セルに達した場合、動画記録から静止画記録に移行した場合などの事象が生じるまで同一タイトル内にチャプターが生成されていく。従って、カムコードにおいて繰り返し記録を行っている際にタイトル内のチャプター数が 99 に達すると、そのタイトルが閉じられ、次のチャプターは新たなタイトル内に生成される。

【0020】

各タイトルにおけるチャプターの開始位置を保持するのは VTSI 7 4 1 である。後述するように、VTSI 7 4 1 は、タイトル毎の管理情報 (TTU) として各チャプターの開始位置を示すポインタ (PTT_SRP) を保持している。

【0021】

図 3 は、ビデオオブジェクトセット (VOBS) の構成要素を示す図である。図 1 の VMGM_VOBS 7 3 2、VTSM_VOBS 7 4 2、および、VTS TT_VOBS 7 4 3 は、それぞれのデータをビデオオブジェクトセット (VOBS) として保持するもの

10

20

30

40

50

であり、共通の形式を有する。このVOBS760は、1個以上のビデオオブジェクト（VOB）761の集合である。このVOB761にはVOB_ID番号が付され、識別のために用いられる。このVOB761は、1個以上のセル（Cell）762により構成される。このセル762は、リアルタイム再生単位であり、VOB761と同様にセルID番号が付される。

【0022】

セル762は、1個以上のビデオオブジェクトユニット（VOBU）763を含む。このVOBU763は、ナビゲーションパック（NV_PCK）764を先頭とするパック列である。パックとは、2048バイトの固定データであり、先頭のNV_PCK764の他、映像データを保持する映像パック（V_PCK）765および音声データを保持する音声パック（A_PCK）766を備える。

10

【0023】

NV_PCK764は、パックヘッダ771に続いて、再生制御情報（PCI: Presentation Control Information）772およびデータ検索情報（DSI: Data Search Information）773を保持する。PCI772は、再生表示に関する制御情報であり、ノンシームレス再生のためのアングル情報や、サブピクチャのハイライト表示のための情報などを保持する。DSI773は、ディスクにアクセスするための制御情報であり、シームレス再生のためのアングル情報や、再生時間（0.5秒×n）を単位としたVOBUの検索情報などを保持する。

【0024】

20

図4は、ビデオオブジェクトユニット（VOBU）と符号化画像との関係を示す図である。VOBS760内の画像データは、MPEG-2（Moving Picture Coding Experts Group -2）符号化方式のプログラムストリーム（PS: Program Stream）により符号化されている。このMPEG-2符号化方式では、GOP（Group Of Pictures）と呼ばれる15枚の画像により0.5秒程度の動画像を表している。図4（a）では、1枚のIピクチャと、4枚のPピクチャと、10枚のBピクチャとにより計15枚の画像511によりGOP510を構成する様子が示されている。ここで、Iピクチャは、フレーム内符号化画像であり、他のPピクチャやBピクチャとは独立して符号化される完結画像である。一方、Pピクチャは、時間的に先行するIピクチャまたはPピクチャとの差分を利用して符号化される。また、Bピクチャは、時間的に前後両方向のIピクチャまたはPピクチャとの差分を利用して符号化される。従って、GOPにおいて独立して復号化できるのはIピクチャだけである。

30

【0025】

GOP内の15枚の画像は、ディスク上では図4（b）のように一部の順序が入れ替えられて記録される。これは、上述の符号化方式の特徴に基づくものであり、復号の際に時間的に後続の画像を待つことを回避するためである。例えば、Bピクチャ（B5）を復号化するためにはIピクチャ（I3）とPピクチャ（P6）を参照する必要があるが、図4（b）のような入れ替えを行うことにより、Bピクチャ（B5）を復号化する時点で必要な画像データ（I3およびP6）が揃っていることになる。なお、このような画像521の順序関係を維持するために、PTS（Presentation Time Stamp）522およびDTS（Decoding Time Stamp）523というタイムスタンプが付与される。PTS522は再生出力の時刻管理情報であり、そのタイムスタンプが付与された単位画像をいつ再生出力するかを示す。一方、DTS523は復号の時刻管理情報であり、そのタイムスタンプが付与された単位画像をいつ復号するかを示す。

40

【0026】

符号化された各画像は、図4（c）のように、それぞれ1つ以上のパックに収められる。例えば、Iピクチャ（I3）はV_PCK_I3（532）として保持され、Bピクチャ（B1）はV_PCK_B1（533）として保持される。そして、NV_PCK531などと共にVOBUを構成する。

【0027】

50

図5は、プログラムチェーン(PGC)におけるプログラムチェーン情報(PGCI)とビデオオブジェクトセット(VOSS)との関係を示す図である。プログラムチェーン(PGC)600は、再生単位であるセル762の再生順序を示すプログラムチェーン情報(PGCI)610と、そのセル762を保持するビデオオブジェクトセット(VOSS)620とから構成される。

【0028】

上述のとおり、VOSS620を構成するVOBやセルにはそれぞれID番号が付されているが、これは必ずしも再生順序を表すわけではない。この再生順序を示すのがPGCI610におけるセル番号(CN)の順序に従って、VOSS620におけるセルが、VOB#1・Cell#1、VOB#1・Cell#2、VOB#2・Cell#1、VOB#3・Cell#1、VOB#3・Cell#2、VOB#4・Cell#1、VOB#4・Cell#2、VOB#4・Cell#3、VOB#5・Cell#1の順で再生される。

【0029】

PGC600において、一つ以上の連続するセル番号を有するセルをまとめたものを一つのプログラム(PG)として定義することができる。このプログラムにはプログラム番号が付される。そして、一つ以上の連続するプログラム番号を有するプログラムをまとめたものをチャプター(PTT)として定義することができる。例えば、図5の例では、PG#1から#3がPTT#1として定義され、PG#4および#5がPTT#2として定義されている。従って、VOB#1・Cell#1、VOB#1・Cell#2、VOB#2・Cell#1、VOB#3・Cell#1、VOB#3・Cell#2が1つ目のチャプターとなり、VOB#4・Cell#1、VOB#4・Cell#2、VOB#4・Cell#3、VOB#5・Cell#1が2つ目のチャプターとなる。

【0030】

図6は、ビデオタイトルセット情報(VTSI)のデータ構造を示す図である。このVTSI741は、ビデオタイトルセット情報管理テーブル(VTSI__MAT:VTSI Management Table)、ビデオタイトルセットPTT検索ポインタテーブル(VTS__PTT__SRPT:VTS PTT Search Pointer Table)、ビデオタイトルセットPGCIテーブル(VTS__PGCI:T:VTS PGCI Table)、ビデオタイトルセットメニューPGCIユニットテーブル(VTSM__PGCI__UT:VTS Menu PGCI Unit Table)、ビデオタイトルセット時刻マップテーブル(VTS__TMAPT:VTS Time MAP Table)、ビデオタイトルセットメニュー・セルアドレステーブル(VTSM__C__ADT:VTS Menu Cell Address Table)、ビデオタイトルセットメニュー・VOBUアドレスマップ(VTSM__VOBU__ADMAP:VTS Menu VOBU Address MAP)、ビデオタイトルセット・セルアドレステーブル(VTS__C__ADT:VTS Cell Address Table)、および、ビデオタイトルセット・VOBUアドレスマップ(VTS__VOBU__ADMAP:VTS VOBU Address MAP)というデータ構造を備えることができる。

【0031】

ここで、VTS__PTT__SRPTは、各タイトルにおけるチャプターのポインタへのアクセス情報である。図2により説明したとおり、PTT__SRPTは各チャプターの開始位置を示すポインタであり、具体的には、各チャプターの属するPGCのPGC番号およびそのPGCにおけるPG番号によりチャプターを特定する。このPTT__SRPTをタイトル毎に保持するのがTTUであり、さらにTTUへのポインタ(TTU__SRPT)を保持するのがVTS__PTT__SRPTである。従って、このVTS__PTT__SRPTに含まれる情報を利用することにより所望のチャプターのPTT__SRPTを読み出すことができ、図2のようにVTS__TT__VOSSにおけるチャプターの開始位置を知ることができる。

【0032】

また、VTSIにおけるVTS__PGCI:Tは、ビデオタイトルセットのプログラムチェーンを再生制御するためのPGCIへのアクセス情報である。このビデオタイトルセッ

10

20

30

40

50

トのPGCIは必要に応じて1つ以上設けることができる。このVTS_PGCIへのポインタ(VTS_PGCI_SRP)がVTS_PGCI_Tに保持される。

【0033】

ここで、PGCIは、PGC全般情報(PGC_GI:PGC General Information)、PGCコマンドテーブル(PGC_CMDT:PGC CoMmanD Table)、PGCプログラムマップ(PGC_PGMAP:PGC ProGram Map)、セルプレイバック情報テーブル(C_PBIT:Cell PlayBack Information Table)、および、セル位置情報テーブル(C_POSIT:Cell POSition Information Table)というデータ構造を備えることができる。PGC_CMDTは、セルの再生前後または再生中に実行すべきコマンドを保持する。PGC_PGMAPは、各プログラムの開始セル番号を保持する。C_PBITは、各セルのカテゴリや各セル内の先頭VOBUの開始アドレスなどを保持する。C_POSITは、各セルのVOB_ID番号およびセルID番号を保持する。従って、PGC番号とPG番号が既知であればPGC_PGMAPからプログラムにおける開始セルのセル番号を取得でき、さらにC_POSITによりセル番号からVOB_ID番号およびセルID番号を取得することができる。これにより、図5のようにVOBSにおける具体的な位置情報を知ることができる。すなわち、PTT_SRPのPGC番号およびPG番号からVTS_TT_VOBSにおけるVOB_ID番号およびセルID番号を取得することができるわけである。

10

【0034】

図7は、本発明の実施の形態における顔情報パック769の内部構成例を示す図である。本発明の実施の形態では、一例として、動画データの各画像に含まれる顔画像の数を検出して、顔画像の数が変化する点をシーンの切れ目として動画データの代表画像を生成することを想定する。

20

【0035】

顔情報パック769は、上述の映像パック765や音声パック766と同様に、ビデオオブジェクトユニット(VOBU)763内に格納される。但し、全てのVOBUに格納される必要はなく、特定のVOBUにのみ格納されるようにしておくことが望ましい。撮像時に顔に関する情報を解析して、撮像終了時にまとめて顔に関する情報を格納することを想定すると、最終のVOBUまたは最終から所定数遡ったVOBUに格納しておくようにするのが便利である。

30

【0036】

この顔情報パック769は、パックヘッダおよびそれに続く顔情報パケットを備える。顔情報パケットは、パケットヘッダとそれに続くサブストリーム識別子と顔情報データを備える。顔情報パケットのパケットヘッダには、ストリームの種別を示すストリーム識別子が含まれる。ここでは、ストリーム識別子として、他のストリームと同期が必要なプライベートストリーム(private_stream_1)を意味する「10111101」(2進数)が保持される。また、パケットヘッダに続くサブストリーム識別子として、プロバイダ定義ストリーム(provider defined stream)を意味する「11111111」(2進数)が保持される。

40

【0037】

図8は、本発明の実施の形態における顔情報データの構成例を示す図である。

【0038】

顔情報データは、顔情報識別子(CAO_ID)と、顔情報プロパティ(CAO_FACE_INFO)と、顔情報の本体とから構成される。顔情報の本体は、顔検出情報(FACE_DETECT_INFO)と顔ブロック情報(FACE_BLOCK)とを保持する。

【0039】

顔情報識別子は、顔情報データであることを示す顔情報データ識別子(CAO_DAT_ID)と、顔情報のバージョンを示す顔情報バージョン(CAO_DAT_VER)とを保持する。顔情報バージョンが異なると、顔情報の形式が異なる場合がある。

【0040】

50

顔情報プロパティは、顔情報サイズ (CAO_FINFO_SIZE) と、顔情報データバージョン (FACE_DAT_VER) と、顔ブロック情報フラグ (FBLK_STRUCT_FLG) と、タイムスケール (FBLK_TIMESCALE) と、顔ブロック情報データ数 (FBLK_Ns) と、顔検出情報サイズ (FDETECT_INF_SIZE) と、顔ブロック情報データサイズ (FBLK_SIZE) とを保持する。

【 0 0 4 1 】

顔情報サイズは、顔情報の全体のサイズを示すフィールドである。顔情報データバージョンは、顔情報データのバージョンを示すフィールドである。顔ブロック情報フラグは、顔情報の本体に何が記載されているかを示す有効フラグ群である。この各フラグについては顔情報の説明において後述する。

【 0 0 4 2 】

タイムスケールは、顔ブロック情報における時間の単位を示すフィールドである。顔ブロック情報データ数は、顔ブロック情報のデータ数を示すフィールドである。顔検出情報サイズは、顔検出情報のサイズを示すフィールドである。顔ブロック情報データサイズは、顔ブロック情報の1ブロック分のサイズを示すフィールドである。

【 0 0 4 3 】

顔検出情報は、顔検出セル番号 (CELL_ID) と、顔検出位置 (OFFSET) とを保持する。顔検出セル番号は、顔が検出されたセルの番号を示すフィールドである。顔検出位置は、顔が検出された位置を示すフィールドである。具体的には、顔検出位置は、顔が検出されたビデオオブジェクトユニット (VOBU) 内のナビゲーションパックの位置を、ビデオオブジェクトセット (VOS) の先頭からパック単位で示したものである。

【 0 0 4 4 】

顔ブロック情報は、顔検出時間オフセット (FACE_DETECT_TIME_OFFSET) と、顔基本情報 (FACE_BASIC_INFO) と、顔スコア情報 (FACE_SCORE) と、顔表情情報 (FACE_EXPRESSION_INFO_1) とを保持する。顔検出時間オフセットは、顔が検出された時間のチャプター先頭からのオフセット時間を示すフィールドである。顔基本情報は、検出された顔の x 座標 (X)、y 座標 (Y)、幅 (WIDTH) および高さ (HEIGHT) をそれぞれ示すフィールドである。顔スコア情報は、検出された顔の顔らしさのスコアを示すフィールドである。顔表情情報は、検出された顔の表情を示すフィールドであり、例えば、笑顔に属するか否か等を示す。これら顔ブロック情報の顔検出時間オフセット、顔基本情報、顔スコア情報および顔表情情報に有効なデータが保持されているか否かは、上述の顔ブロック情報フラグにおける有効フラグ群に示される。

【 0 0 4 5 】

図 9 は、本発明の実施の形態における顔情報データにより代表画像が特定される様子を示す図である。

【 0 0 4 6 】

ここでは、ビデオオブジェクトユニット (VOBU) を構成する VOBU 763 のうち、最終の VOBU # p に顔情報パック 769 が格納されていることを想定している。顔情報パック 769 における顔検出情報の顔検出位置 (OFFSET) に基づいて VOBU # i が特定され、VOBU # i の映像パック 765 に含まれる I ピクチャから顔サムネイル画像が生成される。

【 0 0 4 7 】

なお、上述の顔検出時間オフセットを併用することにより、チャプター先頭からの経過時間を画面上に表示することもできる。

【 0 0 4 8 】

図 10 は、本発明の実施の形態における映像記録装置としての撮像装置 100 の構成を示す図である。この撮像装置 100 は、カメラ部 10 と、記録再生処理部 20 と、制御部 30 とを備えている。

【 0 0 4 9 】

カメラ部 10 は、光学ブロック 11 と、カメラ制御部 12 と、信号変換器 13 と、撮像信号処理部 14 と、音声入力部 15 と、音声信号処理部 16 とを備える。光学ブロック 1

10

20

30

40

50

1 は、内部に、被写体を撮像するためのレンズ群、絞り調整機構、フォーカス調整機構、ズーム機構、シャッター機構、フラッシュ機構、および、手ぶれ補正機構などを備える。カメラ制御部 1 2 は、制御部 3 0 から制御信号を受けて、光学ブロック 1 1 に供給する制御信号を生成する。そして、生成した制御信号を光学ブロック 1 1 に供給して、ズーム制御、シャッター制御、および、露出制御などの制御を行う。

【 0 0 5 0 】

信号変換器 1 3 は、例えば C C D (Charge Coupled Device) などの撮像素子により構成され、その結像面に、光学ブロック 1 1 を通じた像が結像される。この信号変換器 1 3 は、シャッター操作に応じて制御部 3 0 から供給される画像取り込みタイミング信号を受けて、結像面に結像されている被写体像を撮像信号に変換し、撮像信号処理部 1 4 に供給する。

10

【 0 0 5 1 】

撮像信号処理部 1 4 は、制御部 3 0 からの制御信号に基づいて、撮像信号についてのガンマ補正や A G C (Auto Gain Control) などの処理を行うとともに、撮像信号をデジタル信号としての画像信号に変換する処理も行う。音声入力部 1 5 は、撮影時の被写体周辺の音声を収集する。この音声入力部 1 5 からの音声信号は音声信号処理部 1 6 に供給される。音声信号処理部 1 6 は、制御部 3 0 からの制御信号に基づいて、音声信号についての補正や A G C などの処理を行うとともに、音声信号をデジタル信号に変換する処理も行う。

【 0 0 5 2 】

記録再生処理部 2 0 は、符号化 / 復号回路 2 1 と、ディスクインターフェース 2 3 と、出力処理部 2 4 と、バッファメモリ 2 5 とを備える。

20

【 0 0 5 3 】

符号化 / 復号回路 2 1 は、カメラ部 1 0 からの画像信号および音声信号や追加記録情報を M P E G 方式等に符号化し多重化して圧縮データに変換する符号化機能を有する。一方、符号化 / 復号回路 2 1 は、圧縮データから画像信号および音声信号や追加記録情報を分離して復号する復号機能を有する。また、符号化 / 復号回路 2 1 は、制御部 3 0 からの制御信号に基づいて、撮像信号処理部 1 4 からの画像信号に対して、自動ホワイトバランス制御、露出補正制御、デジタルズーム倍率に応じた拡大制御などをさらに行う。

【 0 0 5 4 】

ディスクインターフェース 2 3 は、符号化 / 復号回路 2 1 から圧縮データを受けてディスク 4 9 に書き込む。また、ディスクインターフェース 2 3 は、ディスク 4 9 から圧縮データを読み出して符号化 / 復号回路 2 1 に供給する。出力処理部 2 4 は、制御部 3 0 からの制御により、符号化 / 復号回路 2 1 からの圧縮データを制御部 3 0 や出力端子 2 7 乃至 2 9 に供給する。バッファメモリ 2 5 は、例えば S D R A M などにより構成され、符号化 / 復号回路 2 1 における符号化または復号のための作業領域として利用される。

30

【 0 0 5 5 】

制御部 3 0 は、処理装置 3 1 と、R O M (Read Only Memory) 3 3 と、R A M (Random Access Memory) 3 4 と、操作入力部 4 1 を接続するための操作入力インターフェース 3 5 と、表示部 4 2 を接続するための表示制御部 3 6 と、メモ리카ード 4 3 を装填するためのメモ리카ードインターフェース 3 7 と、手ぶれ補正のために角速度を検出する角速度検出器 3 8 と、撮影時刻を記録するための時計回路 3 9 と、G P S 受信部 4 4 とがシステムバス 3 2 を介して接続されることにより構成される。

40

【 0 0 5 6 】

処理装置 3 1 は制御部 3 0 全体の処理を司るものであり、作業領域として R A M 3 4 を使用する。R O M 3 3 には、カメラ部 1 0 を制御するためのプログラムや、画像信号や音声信号の記録制御および再生制御などを実行するためのプログラムが書き込まれている。

【 0 0 5 7 】

操作入力インターフェース 3 5 に接続される操作入力部 4 1 には、撮影モードと再生モードなどの他のモードとを切り換えるモード切り換えキー、ズーム調整キー、露出調整の

50

ためのキー、シャッターキー、動画撮影用キー、表示部 4 2 における表示調整キーなどの複数のキーが設けられている。操作入力インターフェース 3 5 は、操作入力部 4 1 からの操作信号を処理装置 3 1 に伝える。処理装置 3 1 は、操作入力部 4 1 においていずれのキーが操作されたかを判別し、その判別結果に応じた制御処理を行う。

【 0 0 5 8 】

表示制御部 3 6 に接続される表示部 4 2 は、例えば L C D (Liquid Crystal Display) などにより構成され、処理装置 3 1 の制御の下に、カメラ部 1 0 からの画像信号や、ディスク 4 9 から読み出された画像信号を表示する。

【 0 0 5 9 】

メモリカードインターフェース 3 7 は、符号化 / 復号回路 2 1 からの圧縮データをメモリカード 4 3 に書き込む。また、メモリカードインターフェース 3 7 は、メモリカード 4 3 から圧縮データを読み出して符号化 / 復号回路 2 1 に供給する。

【 0 0 6 0 】

時計回路 3 9 は、年、月、日、時間、分、秒などを表わす時間情報を生成する。角速度検出器 3 8 は、撮像装置 1 0 0 に対して外部から加わる角速度を検出するジャイロスコープである。この角速度検出器 3 8 からの角速度情報 $[\omega = (\text{rad} / \text{sec})]$ は、所定間隔毎に処理装置 3 1 に報告される。GPS 受信部 4 4 は、GPS (Global Positioning System) 衛星からの電波を受信して、撮像装置 1 0 0 の現在位置を取得するものである。

【 0 0 6 1 】

図 1 1 は、本発明の実施の形態における撮像装置 1 0 0 の顔情報生成のための機能構成例を示す図である。

【 0 0 6 2 】

操作受付部 1 1 1 は、ユーザによる操作入力を受け付けるものであり、操作入力部 4 1 により実現される。この操作受付部 1 1 1 により受け付けられた操作入力に基づいて、撮像部 1 1 2 に指示が与えられる。

【 0 0 6 3 】

撮像部 1 1 2 は、被写体を撮像して動画データを生成するものであり、例えばカメラ部 1 0 により実現される。この撮像部 1 1 2 により撮像された動画データは記録制御部 1 1 3 および顔検出部 1 1 4 に供給される。

【 0 0 6 4 】

記録制御部 1 1 3 は、撮像された動画データおよびその動画データに関する顔情報をストリーム記録部 1 9 1 に記録するよう制御するものであり、例えば記録再生処理部 2 0 により実現される。

【 0 0 6 5 】

顔検出部 1 1 4 は、撮像された動画データを構成する各画像について、その各画像に含まれる顔画像を検出して、その顔画像に関する情報を取得するものである。この顔検出部 1 1 4 により得られた情報は、顔情報生成部 1 1 5 に供給される。

【 0 0 6 6 】

顔情報生成部 1 1 5 は、顔検出部 1 1 4 によって検出された顔画像について、顔情報データを生成するものである。顔情報データのフォーマットは、例えば図 8 に示したとおりである。この顔情報生成部 1 1 5 により生成された顔情報データは、顔情報保持部 1 1 6 に保持される。

【 0 0 6 7 】

顔情報保持部 1 1 6 は、顔情報生成部 1 1 5 により生成された顔情報データを保持するものである。この顔情報保持部 1 1 6 には、顔検出部 1 1 4 において顔画像が検出されるのに従って、顔情報データが蓄積されていくことになる。

【 0 0 6 8 】

ストリーム記録部 1 9 1 は、撮像された動画データおよび顔情報データをストリームデータとして記録する記録媒体である。このストリーム記録部 1 9 1 は、例えば DVD - Video 規格によるディスクであり、図 3 のように V O B U 内に映像パック 7 6 5 や音声

10

20

30

40

50

パック766を記録するのと同様に、最終のVOBUまたは最終から所定数遡ったVOBUに顔情報データを顔情報パック769として格納する。

【0069】

図12は、本発明の実施の形態における撮像装置100の顔インデックス表示のための機能構成例を示す図である。

【0070】

操作受付部111およびストリーム記録部191は、図11により説明したものと同様である。この例では、操作受付部111により受け付けられた操作入力に基づいて、ストリーム記録部191に記録された顔情報データがインデックスデータとして読み出される。

10

【0071】

顔情報解析部122は、顔情報データを解析してインデックス表示の対象となる画像を特定するものである。具体的には、この顔情報解析部122は、顔情報パック769の顔検出情報における顔検出位置(図8のOFFSET)により、インデックス表示の対象となる画像を特定する。すなわち、ビデオオブジェクトセット(VOS)の先頭からのパック単位の位置により、インデックス表示の対象となる画像が特定される。

【0072】

アドレス生成部123は、顔情報解析部122から供給された顔検出位置に基づいてインデックス表示の対象となる画像にアクセスするためのストリーム記録部191におけるアドレスを生成するものである。上述の顔検出位置は2048バイトのパック単位であるため、パック単位の位置に2048を乗じたものがバイトアドレスとなる。

20

【0073】

顔サムネイル生成部124は、アドレス生成部123によって生成されたアドレスによりストリーム記録部191から画像を取得して、その画像から顔サムネイル画像を生成するものである。具体的には、この顔サムネイル生成部124は、アドレス生成部123によって生成されたアドレスが示すVOBUに含まれるIピクチャを取得し、そのIピクチャを縮小することにより顔サムネイル画像を生成する。

【0074】

顔インデックス生成部125は、顔サムネイル生成部124により生成された顔サムネイル画像を利用して顔インデックス表示画面を生成するものである。この顔インデックス表示画面については図14を参照して後述する。

30

【0075】

表示部126は、顔インデックス生成部125により生成された顔インデックス表示画面を表示するものである。

【0076】

図13は、本発明の実施の形態によるチャプター表示画面の一例を示す図である。このチャプター表示画面では、ディスクに記録されている各チャプターの代表画像がサムネイル811として表示されている。また、このサムネイル811の下部にはそのチャプターのタイトル番号およびチャプター番号812が表示されている。

【0077】

この図13の例では1枚の画面に6つのチャプターが表示されており、それ以外のチャプターを表示させたい場合には、画面ページを変更する必要がある。画面ページを変更するために使用されるのが前ページボタン818または次ページボタン819である。また、選択されたチャプターの顔インデックス表示画面を表示するためには、ユーザは顔インデックス表示ボタン817を押下する。

40

【0078】

ユーザは、操作キーのメニュー・キーによってチャプター表示画面を呼び出し、四方向キーでハイライト表示を移動する。四方向キー真中の決定キーによって、選択したチャプターの再生開始やページの切替えを行うことができる。

【0079】

50

なお、このチャプター表示画面のそれぞれのページを、画像データとしてVMGM__V
OBS732に記録してもよい。サムネイル811の画像自体もメニューページの一部と
して貼り込まれている。一方、ユーザがリモコンなどにおける操作キーにより画面上のサ
ムネイルを選択する際の選択枠815は、サブピクチャとして定義され、NV__PCK7
64のPCI772(図3)におけるハイライト情報(HLI:HighLight Information
)によりハイライト表示の制御が行われる。このようにメニューページとしてメディアに
記録しておくことで、市販されているDVD再生装置でメディアを再生したときでも、顔
インデックス表示画面を表示することができる。

【0080】

図14は、本発明の実施の形態による顔インデックス表示画面の一例を示す図である。
チャプター表示画面において顔インデックス表示ボタン817が押下されると、選択され
たチャプター821の顔インデックス表示画面が表示される。

10

【0081】

この顔インデックス表示画面には、チャプター821において検出された顔の数に変化
する点をシーンの区切りとして、各シーンの代表画像824が表示される。この顔インデ
ックス表示画面においても、チャプター表示画面と同様に、リモコンなどにおける操作キ
ーにより画面上の代表画像824を選択する際の選択枠825が表示される。

【0082】

この図14の例では1枚の画面に6つの代表画像が表示されており、それ以外の代表画
像を表示させたい場合には、画面ページを変更する必要がある。画面ページを変更するた
めに使用されるのが前ページボタン828または次ページボタン829である。また、こ
の例では1枚の画面に3つのチャプターが表示されており、それ以外のチャプターを表示
させたい場合には、表示チャプターを変更する必要がある。表示チャプターを変更するた
めに使用されるのが前チャプターボタン822または次チャプターボタン823である。

20

【0083】

この顔インデックス表示画面においても、ユーザは四方向キーでハイライト表示を移動
する。四方向キー真中の決定キーによって、選択された代表画像に対応する位置からの再
生開始やページの切替えを行うことができる。

【0084】

図15は、本発明の実施の形態における撮像装置100の動画データ変換のための機能
構成例を示す図である。図11の例では撮像の際に顔検出を行って顔情報をストリームデ
ータに格納する実施態様を示したが、ここでは動画データとは独立した動画情報が存在す
る場合に動画情報に基づいてストリームデータに変換する実施態様について説明する。例
えば、顔情報が動画データ以外のメタファイルとして保存されている場合において、動画
データおよび顔情報を1つのストリームデータとして記録することを想定する。

30

【0085】

操作受付部111およびストリーム記録部191は、図11により説明したものと同様
である。この例では、操作受付部111により受け付けられたデータ変換指示入力に基づ
いて、データ変換が行われてストリームデータがストリーム記録部191に記録される。

【0086】

40

動画データ保持部181は動画データを保持するものであり、動画情報保持部182は
その動画データに関する動画情報を保持するものである。すなわち、ここでは、動画デー
タに関する動画情報が動画データ保持部181とは別のファイルとして保持されているこ
とを想定している。

【0087】

データ変換部132は、操作受付部111からの指示に基づいて動画データ保持部18
1に保持される動画データをストリームデータの形式に変換するものであり、例えば処理
装置31により実現される。変換されたストリームデータは、記録制御部133に供給さ
れる。

【0088】

50

記録制御部 133 は、データ変換部 132 によって変換されたストリームデータをストリーム記録部 191 に記録するよう制御するものであり、例えば記録再生処理部 20 により実現される。

【0089】

動画情報解析部 134 は、データ変換部 132 において変換対象となっている動画データに関する動画情報を動画情報保持部 182 から読み出して、その動画情報を解析してその顔画像に関する情報を取得するものであり、例えば処理装置 31 により実現される。

【0090】

顔情報生成部 135 は、動画情報解析部 134 によって取得された顔画像について、顔情報データを生成するものである。顔情報データのフォーマットは、例えば図 8 に示したとおりである。この顔情報生成部 135 により生成された顔情報データは、顔情報保持部 136 に保持される。

10

【0091】

顔情報保持部 136 は、顔情報生成部 135 により生成された顔情報データを保持するものである。この顔情報保持部 136 には、動画情報解析部 134 において顔画像が取得されるのに従って、顔情報データが蓄積されていくことになる。この顔情報データは、記録制御部 133 によってストリームデータの一部として格納される。すなわち、図 3 の V O B U 内の映像パック 765 や音声パック 766 と同様に、最終の V O B U または最終から所定数遡った V O B U に顔情報データが顔情報パック 769 としてストリーム記録部 191 に格納される。

20

【0092】

図 16 は、本発明の実施の形態における撮像装置 100 の動画データ編集のための機能構成例を示す図である。ここでは、ストリーム記録部 191 に記録されているストリームデータが分割編集された際に、顔情報データが更新される実施態様について説明する。

【0093】

操作受付部 111 およびストリーム記録部 191 は、図 11 により説明したものと同様である。この例では、操作受付部 111 により受け付けられたデータ編集指示入力に基づいて、ストリーム記録部 191 に記録されているストリームデータの編集が行われる。

【0094】

ストリーム編集部 142 は、ストリーム記録部 191 に記録されているストリームデータを、操作受付部 111 により受け付けられたデータ編集指示入力に基づいて分割編集するものである。

30

【0095】

顔情報更新部 143 は、ストリーム編集部 142 によって分割編集されたストリームデータについて顔情報データを更新するものである。すなわち、分割点以降のストリームデータについては、元の顔情報データから分割点以前の顔情報データを削除する。一方、分割点以前のストリームデータについては、元の顔情報データから分割点以前の顔情報データを抽出して、新たな顔情報データとして追加する。これら分割点以前の顔情報データおよび分割点以降のストリームデータは、それぞれ別個のストリームデータとして、ストリーム編集部 142 を介してストリーム記録部 191 に記録される。

40

【0096】

次に本発明の実施の形態における撮像装置 100 の動作について図面を参照して説明する。

【0097】

図 17 は、本発明の実施の形態における撮像時の顔情報生成方法の処理手順例を示す図である。

【0098】

まず、操作受付部 111 により撮像の指示がなされると、撮像部 112 において撮像された動画データを構成する各画像について、顔検出部 114 がその各画像に含まれる顔画像を検出して、その顔画像に関する情報を取得する（ステップ S911）。そして、顔情

50

報生成部 115 は、ステップ S911 において検出された顔画像について、顔情報データを生成する（ステップ S912）。これらの処理が、チャプターの終了に至るまで繰り返される（ステップ S913）。

【0099】

そして、1チャプター分の撮像の終了時に、撮像された動画データに関する顔情報が記録制御部 113 によってストリーム記録部 191 に記録される（ステップ S914）。具体的には、最終の VOB U または最終から所定数遡った VOB U に顔情報パック 769 が格納される。

【0100】

図 18 は、本発明の実施の形態における顔インデックス表示方法の処理手順例を示す図である。 10

【0101】

まず、ストリーム記録部 191 に記録された顔情報データが顔情報解析部 122 によって解析され、インデックス表示の対象となる画像が特定される（ステップ S921）。その特定されたインデックス表示の対象となる画像にアクセスするためのストリーム記録部 191 におけるアドレスがアドレス生成部 123 によって生成される（ステップ S922）。この生成されたアドレスの示す VOB U に含まれる I ピクチャが顔サムネイル生成部 124 によって取得され（ステップ S923）、その I ピクチャを縮小することにより顔サムネイル画像が生成される（ステップ S924）。

【0102】

このようにして生成された顔サムネイル画像を利用して、顔インデックス生成部 125 によって顔インデックス表示画面が生成され（ステップ S925）、表示部 126 に顔インデックス表示画面が表示される（ステップ S926）。 20

【0103】

図 19 は、本発明の実施の形態における動画データ変換時の顔情報生成方法の処理手順例を示す図である。

【0104】

まず、データ変換部 132 において変換対象となっている動画データに関する動画情報が動画情報保持部 182 から読み出され、動画情報解析部 134 がその動画情報を解析してその顔画像に関する情報を取得する（ステップ S931）。顔情報生成部 135 は、ステップ S931 において取得された顔画像について、顔情報データを生成する（ステップ S932）。そして、動画データの変換終了時に、変換された動画データに関する顔情報が記録制御部 113 によってストリーム記録部 191 に記録される（ステップ S934）。具体的には、最終の VOB U または最終から所定数遡った VOB U に顔情報パック 769 が格納される。 30

【0105】

図 20 は、本発明の実施の形態における動画データ編集時の顔情報データ更新方法の処理手順例を示す図である。

【0106】

まず、操作受付部 111 により受け付けられたデータ編集指示入力に基づいて、ストリーム編集部 142 はストリーム記録部 191 に記録されているストリームデータを分割編集する（ステップ S941）。そして、分割点以前のストリームデータについては、元のストリームデータの顔情報データから分割点以前の顔情報データが抽出され、新たな顔情報データとして追加される（ステップ S942）。一方、分割点以降のストリームデータについては、元のストリームデータの顔情報データから分割点以前の顔情報データが削除される（ステップ S943）。 40

【0107】

このように、本発明の実施の形態によれば、顔情報生成部 115 によって生成された顔情報パック 769 を最終の VOB U または最終から所定数遡った VOB U に格納しておくことによって、動画データの転送後においてもその動画データの各シーンに関する顔情報 50

を維持することができるとともに、この顔情報に基づいて顔サムネイル生成部 1 2 4 および顔インデックス生成部 1 2 5 によって顔インデックス表示画面を生成して、表示部 1 2 6 に表示することができる。また、顔情報の生成は、撮像時の他、異なるフォーマット間のデータ変換時にも行うことができる。また、ストリームデータの分割編集の際には、分割された両ストリームデータの最終の V O B U または最終から所定数遡った V O B U に顔情報パックを格納することができる。

【 0 1 0 8 】

なお、本発明の実施の形態では、顔画像の数が変化する点をシーンの切れ目として動画データの代表画像を生成しているが、これ以外をシーンの切れ目とすることもできる。一般に、動画データに含まれる複数の画像に関する情報を属性情報とすると、この属性情報としては、画像自体の特徴の他、画像に対応する音声の状態や、その画像が撮像された場所や時刻などの情報が含まれる。例えば、時計回路 3 9、角速度検出器 3 8 および G P S 受信部 4 4 により取得された情報を属性情報として、これら属性情報の変化点をシーンの切れ目とすることができる。

10

【 0 1 0 9 】

また、本発明の実施の形態では、アドレス生成部 1 2 3 において生成されたアドレスが示す V O B U に含まれる I ピクチャから、顔サムネイル生成部 1 2 4 が顔サムネイル画像を生成することを想定したが、顔サムネイル画像を予め生成しておいて顔情報データに格納しておいてもよい。この場合、顔サムネイル生成部 1 2 4 は、顔情報データから顔サムネイル画像を読み出して、顔インデックス生成部 1 2 5 に供給する。

20

【 0 1 1 0 】

なお、本発明の実施の形態は本発明を具現化するための一例を示したものであり、以下に示すように特許請求の範囲における発明特定事項とそれぞれ対応関係を有するが、これに限定されるものではなく本発明の要旨を逸脱しない範囲において種々の変形を施すことができる。

【 0 1 1 6 】

なお、本発明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。

【 図面の簡単な説明 】

30

【 0 1 1 7 】

【 図 1 】 D V D - V i d e o 規格によるディスクのデータ構造を示す図である。

【 図 2 】 ビデオタイトルセット (V T S) におけるタイトルとチャプターとの関係を示す図である。

【 図 3 】 ビデオオブジェクトセット (V O B S) の構成要素を示す図である。

【 図 4 】 ビデオオブジェクトユニット (V O B U) と符号化画像との関係を示す図である。

【 図 5 】 プログラムチェーン (P G C) におけるプログラムチェーン情報 (P G C I) とビデオオブジェクトセット (V O B S) との関係を示す図である。

【 図 6 】 ビデオタイトルセット情報 (V T S I) のデータ構造を示す図である。

40

【 図 7 】 本発明の実施の形態における顔情報パック 7 6 9 の内部構成例を示す図である。

【 図 8 】 本発明の実施の形態における顔情報データの構成例を示す図である。

【 図 9 】 本発明の実施の形態における顔情報データにより代表画像が特定される様子を示す図である。

【 図 1 0 】 本発明の実施の形態における映像記録装置としての撮像装置 1 0 0 の構成を示す図である。

【 図 1 1 】 本発明の実施の形態における撮像装置 1 0 0 の顔情報生成のための機能構成例を示す図である。

【 図 1 2 】 本発明の実施の形態における撮像装置 1 0 0 の顔インデックス表示のための機能構成例を示す図である。

50

【図 1 3】本発明の実施の形態によるチャプター表示画面の一例を示す図である。

【図 1 4】本発明の実施の形態による顔インデックス表示画面の一例を示す図である。

【図 1 5】本発明の実施の形態における撮像装置 1 0 0 の動画データ変換のための機能構成例を示す図である。

【図 1 6】本発明の実施の形態における撮像装置 1 0 0 の動画データ編集のための機能構成例を示す図である。

【図 1 7】本発明の実施の形態における撮像時の顔情報生成方法の処理手順例を示す図である。

【図 1 8】本発明の実施の形態における顔インデックス表示方法の処理手順例を示す図である。

【図 1 9】本発明の実施の形態における動画データ変換時の顔情報生成方法の処理手順例を示す図である。

【図 2 0】本発明の実施の形態における動画データ編集時の顔情報データ更新方法の処理手順例を示す図である。

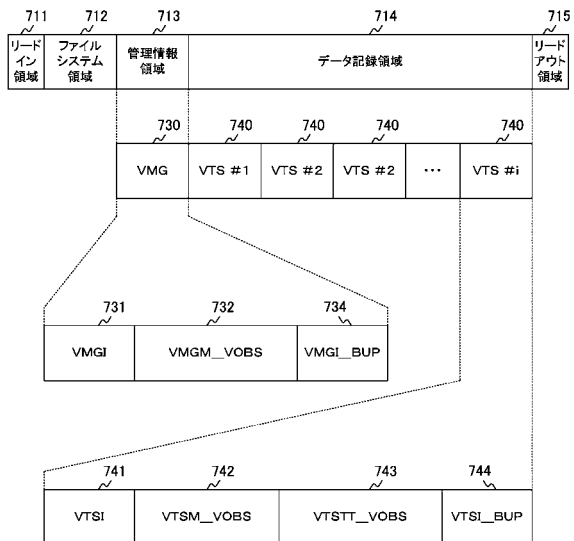
【符号の説明】

【 0 1 1 8 】

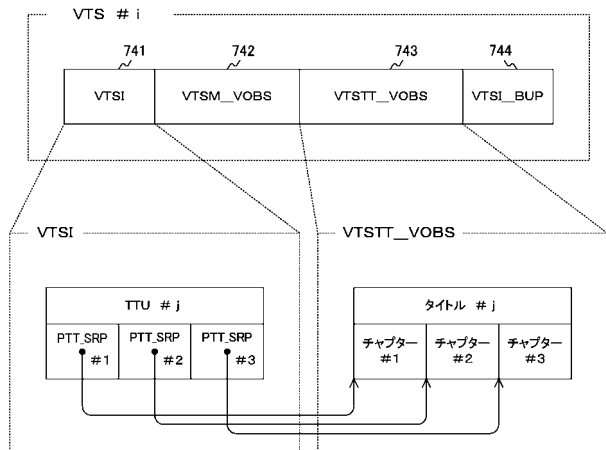
1 0	カメラ部	
1 1	光学ブロック	
1 2	カメラ制御部	
1 3	信号変換器	20
1 4	撮像信号処理部	
1 5	音声入力部	
1 6	音声信号処理部	
2 0	記録再生処理部	
2 1	符号化／復号回路	
2 3	ディスクインターフェース	
2 4	出力処理部	
2 5	バッファメモリ	
2 7 ~ 2 9	出力端子	
3 0	制御部	30
3 1	処理装置	
3 2	システムバス	
3 3	R O M	
3 4	R A M	
3 5	操作入力インターフェース	
3 6	表示制御部	
3 7	メモリカードインターフェース	
3 8	角速度検出器	
3 9	時計回路	
4 1	操作入力部	40
4 2	表示部	
4 3	メモリカード	
4 4	G P S 受信部	
1 0 0	撮像装置	
1 1 1	操作受付部	
1 1 2	撮像部	
1 1 3、1 3 3	記録制御部	
1 1 4	顔検出部	
1 1 5、1 3 5	顔情報生成部	
1 1 6、1 3 6	顔情報保持部	50

- 1 2 2 顔情報解析部
- 1 2 3 アドレス生成部
- 1 2 4 顔サムネイル生成部
- 1 2 5 顔インデックス生成部
- 1 2 6 表示部
- 1 3 2 データ変換部
- 1 3 4 動画情報解析部
- 1 4 2 ストリーム編集部
- 1 4 3 顔情報更新部
- 1 8 1 動画データ保持部
- 1 8 2 動画情報保持部
- 1 9 1 ストリーム記録部

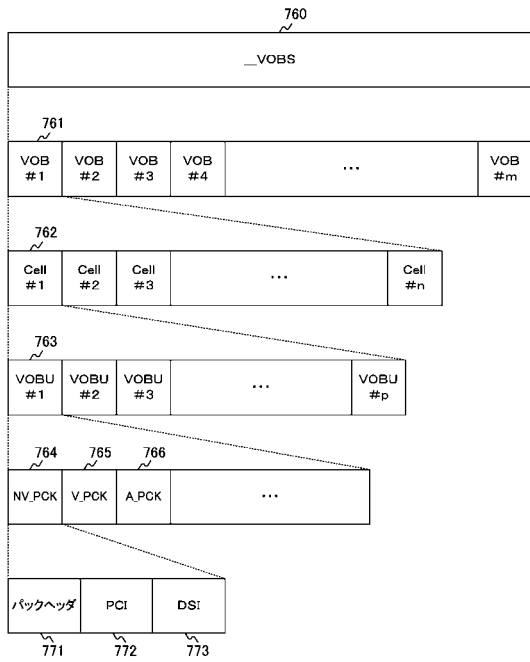
【図1】



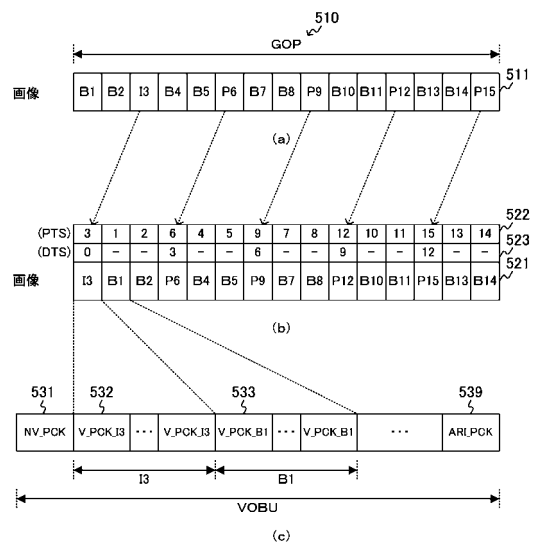
【図2】



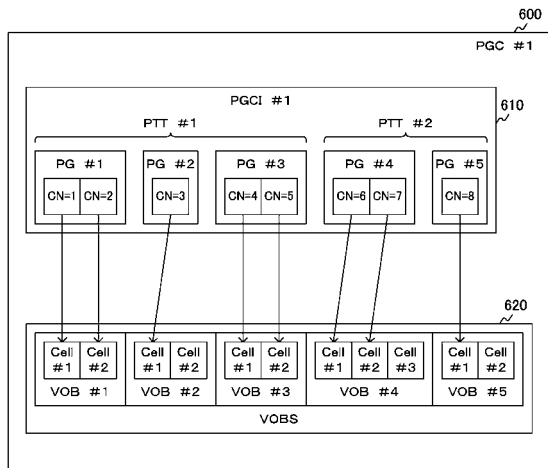
【図3】



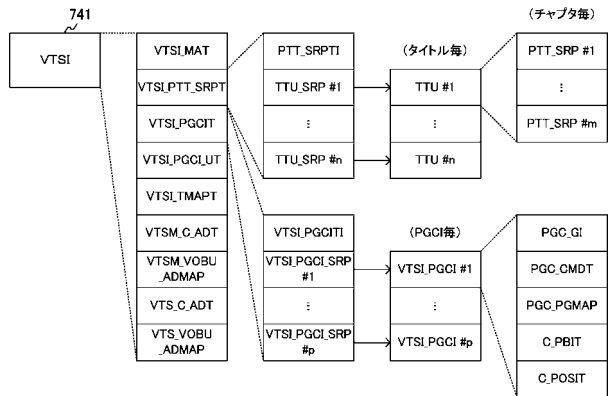
【図4】



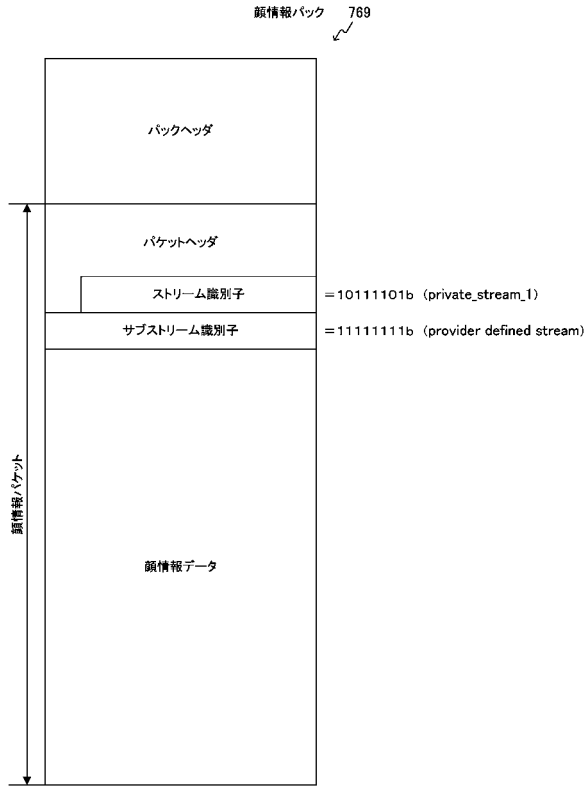
【図5】



【図6】



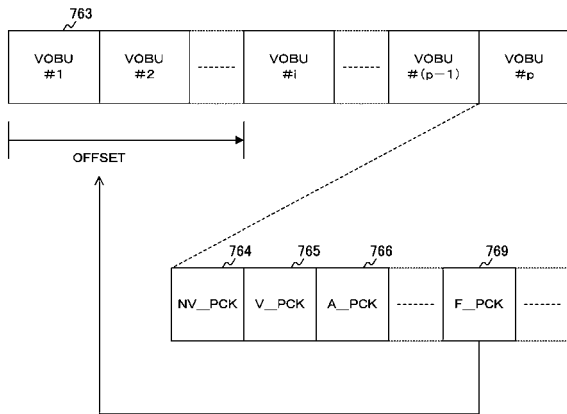
【図7】



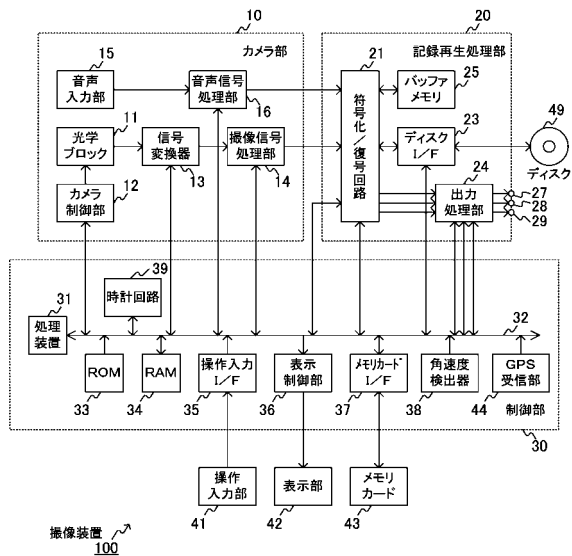
【図8】

CAO_ID	顔情報識別子
CAO_DAT_ID	顔情報データ識別子
CAO_DAT_VER	顔情報バージョン
CAO_FACE_INFO	顔情報プロパティ
CAO_FINFO_SIZE	顔情報サイズ
FACE_DAT_VER	顔情報データバージョン
FBLK_STRUGT_FLG	顔ブロック情報フラグ
FBLK_TIMESCALE	タイムスケール
FBLK_Ns	顔ブロック情報データ数
FDTECT_INF_SIZE	顔検出情報サイズ
FBLK_SIZE	顔ブロック情報データサイズ
for (i=0; i<FBLK_Ns; i++)	顔情報
FACE_DETECT_INFO	顔検出情報
CELL_ID	顔検出セル番号
OFFSET	顔検出位置
FACE_BLOCK	顔ブロック情報
FACE_DETECT_TIME_OFFSET	顔検出時間オフセット
FACE_BASIC_INFO	顔基本情報
X	x座標
Y	y座標
WIDTH	幅
HEIGHT	高さ
FACE_SCORE	顔スコア情報
FACE_EXPRESSION_INFO_1	顔表情情報

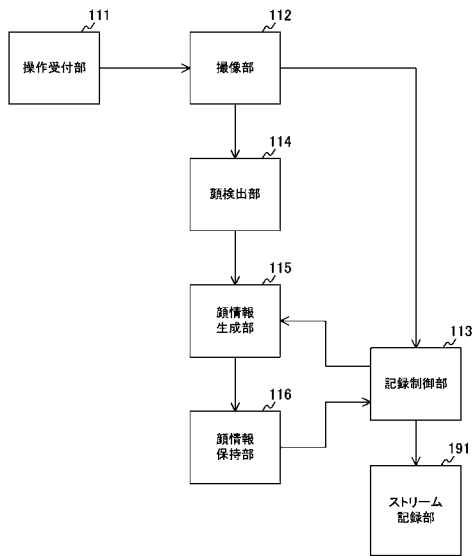
【図9】



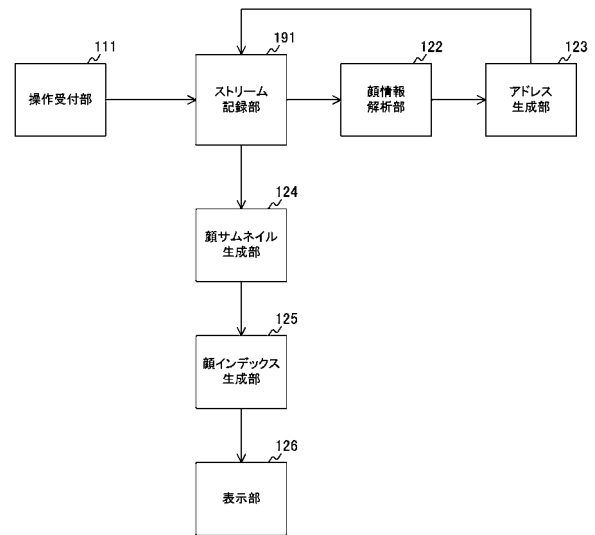
【図10】



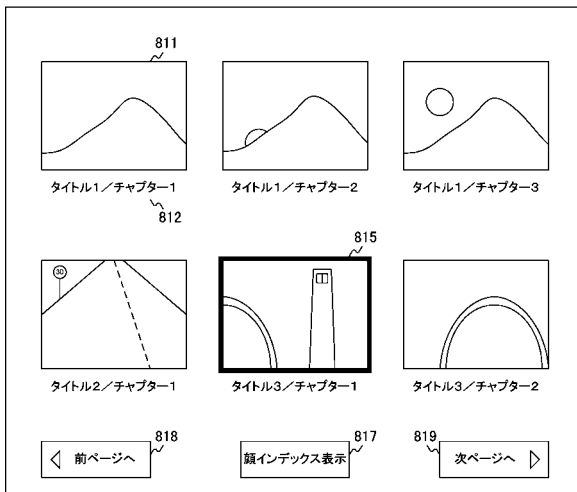
【図 1 1】



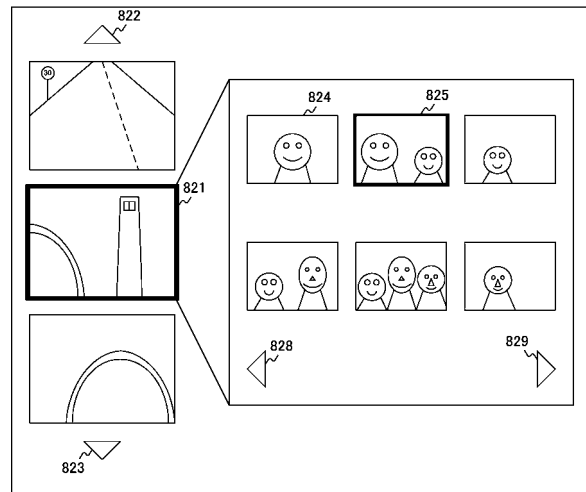
【図 1 2】



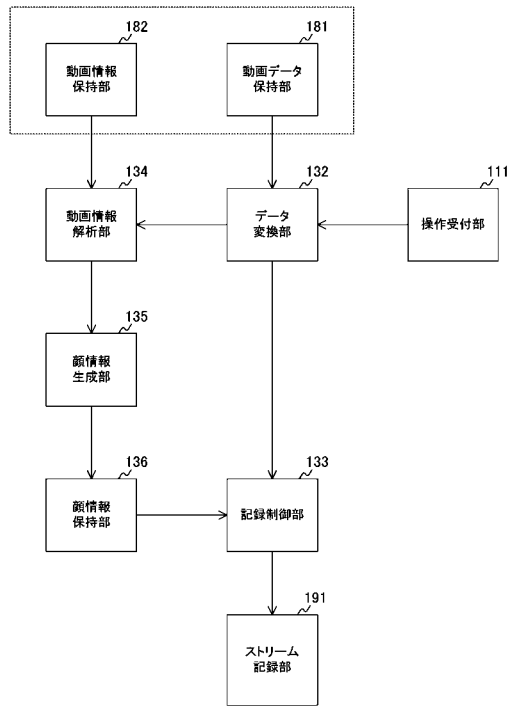
【図 1 3】



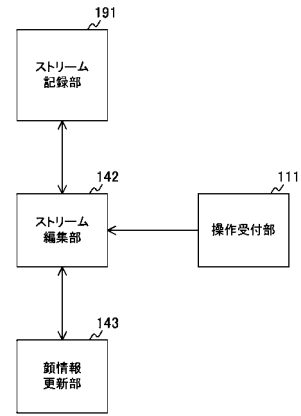
【図 1 4】



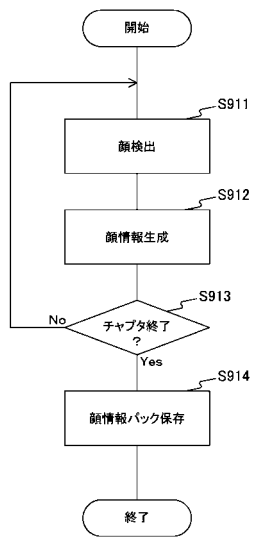
【図15】



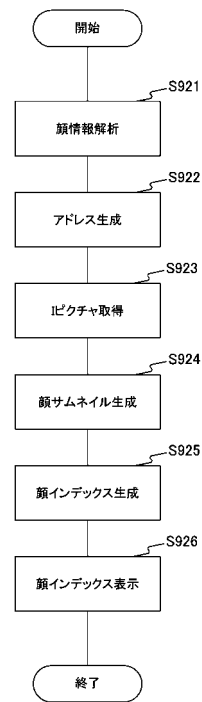
【図16】



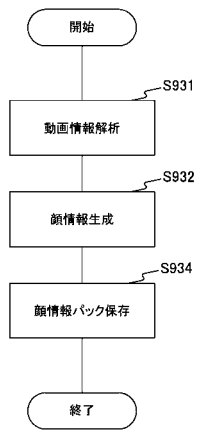
【図17】



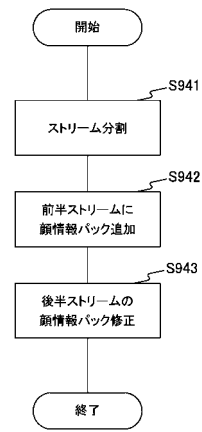
【図18】



【図19】



【図20】



フロントページの続き

(72)発明者 増田 徹
東京都港区港南1丁目7番1号 ソニー株式会社内

審査官 梅岡 信幸

(56)参考文献 特開2007-280325(JP,A)
特開2007-226606(JP,A)
特開2003-204541(JP,A)

(58)調査した分野(Int.Cl., DB名)

H04N	5/76	-	5/956
G11B	20/10	-	20/16
G11B	27/00	-	27/34
H04N	5/222	-	5/257