

(19) 日本国特許庁(JP)

## (12) 特許公報(B2)

(11) 特許番号

特許第5510012号  
(P5510012)

(45) 発行日 平成26年6月4日(2014.6.4)

(24) 登録日 平成26年4月4日(2014.4.4)

(51) Int.Cl.

G06F 17/30 (2006.01)

F 1

G06F 17/30 220A  
G06F 17/30 170D

請求項の数 7 (全 31 頁)

(21) 出願番号 特願2010-90608 (P2010-90608)  
 (22) 出願日 平成22年4月9日 (2010.4.9)  
 (65) 公開番号 特開2011-221806 (P2011-221806A)  
 (43) 公開日 平成23年11月4日 (2011.11.4)  
 審査請求日 平成25年3月28日 (2013.3.28)

(73) 特許権者 000002185  
 ソニー株式会社  
 東京都港区港南1丁目7番1号  
 (74) 代理人 100082131  
 弁理士 稲本 義雄  
 (74) 代理人 100121131  
 弁理士 西川 孝  
 (72) 発明者 田中 健司  
 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 川口 邦雄  
 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】 画像処理装置および方法、並びにプログラム

## (57) 【特許請求の範囲】

## 【請求項 1】

動画像コンテンツを構成する画像のうち、第1の頻度で抽出される画像より第1のサイズの第1階層要約データを生成し、データベースに登録する第1階層要約データ生成手段と、

前記動画像コンテンツを構成する画像のうち、前記第1の頻度よりも高い第2の頻度で抽出される画像より、前記第1のサイズよりも小さな第2のサイズの第2階層要約データを生成する第2階層要約データ生成手段と、

前記データベースより、前記第1階層要約データ生成手段により生成された前記第1階層要約データに基づいて、対応する動画像コンテンツを検索する第1の検索手段と、

前記第1の検索手段により検索された動画像コンテンツのうち、前記第2階層要約データ生成手段により生成された前記第2階層要約データに基づいて、対応する動画像コンテンツを検索する第2の検索手段と

を含む画像処理装置。

## 【請求項 2】

前記第1階層要約データは、前記動画像コンテンツを構成する画像のうち、前記第1の頻度で抽出される画像が複数の分割領域に分割された、前記分割領域毎の画素値、輝度値、アクティビティ、音声ボリューム、および所定の音声周波数帯域の振幅の平均値のいずれか、または組み合わせから構成され、前記第1のサイズであり、

前記第2階層要約データは、前記動画像コンテンツを構成する画像のうち、前記第1の

10

20

頻度よりも高い第2の頻度で抽出される画像が複数の分割領域に分割された、前記分割領域毎の画素値、輝度値、アクティビティ、音声ボリューム、または所定の音声周波数帯域の振幅の平均値の少なくともいずれかまたは組み合わせから構成され、前記第2のサイズである

請求項1に記載の画像処理装置。

【請求項3】

前記第1の頻度、および前記第2の頻度は、前記動画像コンテンツのうち、周期的、または非周期的に抽出される画像の頻度である

請求項1に記載の画像処理装置。

【請求項4】

前記周期的とは、所定のフレーム数毎であり、前記非周期的とは、前記動画像コンテンツにおけるシーンチェンジにおけるタイミング毎、または音声データの無音部分から無音部分でなくなるタイミング毎である

請求項3に記載の画像処理装置。

【請求項5】

前記第2の検索手段により検索された動画像コンテンツと、前記第1階層要約データ生成手段により前記第1階層要約データが生成された動画コンテンツとの差分領域を接続することにより、両者のいずれかを削除して、動画コンテンツのデータを圧縮する圧縮手段をさらに含む

請求項1に記載の画像処理装置。

10

【請求項6】

動画像コンテンツを構成する画像のうち、第1の頻度で抽出される画像より第1のサイズの第1階層要約データを生成し、データベースに登録する第1階層要約データ生成手段と、

前記動画像コンテンツを構成する画像のうち、前記第1の頻度よりも高い第2の頻度で抽出される画像より、前記第1のサイズよりも小さな第2のサイズの第2階層要約データを生成する第2階層要約データ生成手段と、

前記データベースより、前記第1階層要約データ生成手段により生成された前記第1階層要約データに基づいて、対応する動画像コンテンツを検索する第1の検索手段と、

前記第1の検索手段により検索された動画像コンテンツのうち、前記第2階層要約データ生成手段により生成された前記第2階層要約データに基づいて、対応する動画像コンテンツを検索する第2の検索手段と

を含む画像処理装置の画像処理方法であって、

前記第1階層要約データ生成手段における、動画像コンテンツを構成する画像のうち、前記第1の頻度で抽出される画像より第1のサイズの第1階層要約データを生成し、データベースに登録する第1階層要約データ生成ステップと、

前記第2階層要約データ生成手段における、前記動画像コンテンツを構成する画像のうち、前記第1の頻度よりも高い第2の頻度で抽出される画像より、前記第1のサイズよりも小さな第2のサイズの第2階層要約データを生成する第2階層要約データ生成ステップと、

前記第1の検索手段における、前記データベースより、前記第1階層要約データ生成ステップの処理により生成された前記第1階層要約データに基づいて、対応する動画像コンテンツを検索する第1の検索ステップと、

前記第2の検索手段における、前記第1の検索ステップの処理により検索された動画像コンテンツのうち、前記第2階層要約データ生成ステップの処理により生成された前記第2階層要約データに基づいて、対応する動画像コンテンツを検索する第2の検索ステップと

を含む画像処理方法。

【請求項7】

動画像コンテンツを構成する画像のうち、第1の頻度で抽出される画像より第1のサイ

20

30

40

50

ズの第1階層要約データを生成し、データベースに登録する第1階層要約データ生成手段と、

前記動画像コンテンツを構成する画像のうち、前記第1の頻度よりも高い第2の頻度で抽出される画像より、前記第1のサイズよりも小さな第2のサイズの第2階層要約データを生成する第2階層要約データ生成手段と、

前記データベースより、前記第1階層要約データ生成手段により生成された前記第1階層要約データに基づいて、対応する動画像コンテンツを検索する第1の検索手段と、

前記第1の検索手段により検索された動画像コンテンツのうち、前記第2階層要約データ生成手段により生成された前記第2階層要約データに基づいて、対応する動画像コンテンツを検索する第2の検索手段と

10

を含む画像処理装置を制御するコンピュータに、

前記第1階層要約データ生成手段における、動画像コンテンツを構成する画像のうち、前記第1の頻度で抽出される画像より第1のサイズの第1階層要約データを生成し、データベースに登録する第1階層要約データ生成ステップと、

前記第2階層要約データ生成手段における、前記動画像コンテンツを構成する画像のうち、前記第1の頻度よりも高い第2の頻度で抽出される画像より、前記第1のサイズよりも小さな第2のサイズの第2階層要約データを生成する第2階層要約データ生成ステップと、

前記第1の検索手段における、前記データベースより、前記第1階層要約データ生成ステップの処理により生成された前記第1階層要約データに基づいて、対応する動画像コンテンツを検索する第1の検索ステップと、

20

前記第2の検索手段における、前記第1の検索ステップの処理により検索された動画像コンテンツのうち、前記第2階層要約データ生成ステップの処理により生成された前記第2階層要約データに基づいて、対応する動画像コンテンツを検索する第2の検索ステップと

を実行させるプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、画像処理装置および方法、並びにプログラムに関し、特に、多くの動画像コンテンツを容易に検索し、管理できるようにする画像処理装置および方法、並びにプログラムに関する。

30

【背景技術】

【0002】

多くの動画像コンテンツを管理する技術が一般に普及しつつある。

【0003】

特に放送局などでは、毎日放送される放送番組が記録され続けることになるため、膨大な量の動画像コンテンツが増え続けることになる。このため、多くの動画像コンテンツから特定の動画像コンテンツを検索できるようにするには、困難を極める。

【0004】

40

これらの問題に対応するため、検索しようとする動画像コンテンツに含まれるいずれかの例示画像のメタデータと、予め登録されている蓄積された動画像コンテンツのメタデータとを比較して、類似する動画像コンテンツを検索する技術が提案されている（特許文献1参照）。

【先行技術文献】

【特許文献】

【0005】

【特許文献1】特開2001-134589号公報

【発明の概要】

【発明が解決しようとする課題】

50

**【 0 0 0 6 】**

しかしながら、特許文献1に記載の技術では、同一のメタデータからなる画像を含む動画像コンテンツが重複して存在している場合、検索結果から、さらに別の例示動画像により検索する必要がある。また、類似した画像を多く含む動画像コンテンツが膨大な場合、検索したい動画像コンテンツのうち、余程特徴的な例示動画像を用いて検索しない限り、管理する動画像コンテンツが増えるほど絞り込むことが困難になる。結果として、欲しい動画像コンテンツを検索するためには、さらなる検索に係る手間や時間が必要となる恐れがある。

**【 0 0 0 7 】**

本発明はこのような状況に鑑みてなされたものであり、特に、膨大な量の動画像コンテンツの管理を容易なものとすると共に、管理している膨大な量の動画像コンテンツの中から検索したい動画像コンテンツを容易に検索できるようにするものである。 10

**【課題を解決するための手段】****【 0 0 0 8 】**

本発明の一側面の画像処理装置は、動画像コンテンツを構成する画像のうち、第1の頻度で抽出される画像より第1のサイズの第1階層要約データを生成し、データベースに登録する第1階層要約データ生成手段と、前記動画像コンテンツを構成する画像のうち、前記第1の頻度よりも高い第2の頻度で抽出される画像より、前記第1のサイズよりも小さな第2のサイズの第2階層要約データを生成する第2階層要約データ生成手段と、前記データベースより、前記第1階層要約データ生成手段により生成された前記第1階層要約データに基づいて、対応する動画像コンテンツを検索する第1の検索手段と、前記第1の検索手段により検索された動画像コンテンツのうち、前記第2階層要約データ生成手段により生成された前記第2階層要約データに基づいて、対応する動画像コンテンツを検索する第2の検索手段とを含む。 20

**【 0 0 0 9 】**

前記第1階層要約データは、前記動画像コンテンツを構成する画像のうち、前記第1の頻度で抽出される画像が複数の分割領域に分割された、前記分割領域毎の画素値、輝度値、アクティビティ、音声ボリューム、および所定の音声周波数帯域の振幅の平均値のいずれか、または組み合わせから構成されるものとし、前記第1のサイズとすることができ、前記第2階層要約データは、前記動画像コンテンツを構成する画像のうち、前記第1の頻度よりも高い第2の頻度で抽出される画像が複数の分割領域に分割された、前記分割領域毎の画素値、輝度値、アクティビティ、音声ボリューム、または所定の音声周波数帯域の振幅の平均値の少なくともいずれかまたは組み合わせから構成されるものとし、前記第2のサイズとすることができる。 30

**【 0 0 1 0 】**

前記第1の頻度、および前記第2の頻度は、前記動画像コンテンツのうち、周期的、または非周期的に抽出される画像の頻度とすることができます。

**【 0 0 1 1 】**

前記周期的とは、所定のフレーム数毎であるものとし、前記非周期的とは、前記動画像コンテンツにおけるシーンチェンジにおけるタイミング毎、または音声データの無音部分から無音部分でなくなるタイミング毎であるものとすることができます。 40

**【 0 0 1 2 】**

前記第2の検索手段により検索された動画像コンテンツと、前記第1階層要約データ生成手段により前記第1階層要約データが生成された動画コンテンツとの差分領域を接続することにより、両者のいずれかを削除して、動画コンテンツのデータを圧縮する圧縮手段をさらに含ませるようにすることができます。

**【 0 0 1 3 】**

本発明の一側面の画像処理方法は、動画像コンテンツを構成する画像のうち、第1の頻度で抽出される画像より第1のサイズの第1階層要約データを生成し、データベースに登録する第1階層要約データ生成手段と、前記動画像コンテンツを構成する画像のうち、前 50

記第1の頻度よりも高い第2の頻度で抽出される画像より、前記第1のサイズよりも小さな第2のサイズの第2階層要約データを生成する第2階層要約データ生成手段と、前記データベースより、前記第1階層要約データ生成手段により生成された前記第1階層要約データに基づいて、対応する動画像コンテンツを検索する第1の検索手段と、前記第1の検索手段により検索された動画像コンテンツのうち、前記第2階層要約データ生成手段により生成された前記第2階層要約データに基づいて、対応する動画像コンテンツを検索する第2の検索手段とを含む画像処理装置の画像処理方法であって、前記第1階層要約データ生成手段における、動画像コンテンツを構成する画像のうち、前記第1の頻度で抽出される画像より第1のサイズの第1階層要約データを生成し、データベースに登録する第1階層要約データ生成ステップと、前記第2階層要約データ生成手段における、前記動画像コンテンツを構成する画像のうち、前記第1の頻度よりも高い第2の頻度で抽出される画像より、前記第1のサイズよりも小さな第2のサイズの第2階層要約データを生成する第2階層要約データ生成ステップと、前記第1の検索手段における、前記データベースより、前記第1階層要約データ生成ステップの処理により生成された前記第1階層要約データに基づいて、対応する動画像コンテンツを検索する第1の検索ステップと、前記第2の検索手段における、前記第1の検索ステップの処理により検索された動画像コンテンツのうち、前記第2階層要約データ生成ステップの処理により生成された前記第2階層要約データに基づいて、対応する動画像コンテンツを検索する第2の検索ステップとを含む。10

#### 【0014】

本発明の一側面のプログラムは、動画像コンテンツを構成する画像のうち、第1の頻度で抽出される画像より第1のサイズの第1階層要約データを生成し、データベースに登録する第1階層要約データ生成手段と、前記動画像コンテンツを構成する画像のうち、前記第1の頻度よりも高い第2の頻度で抽出される画像より、前記第1のサイズよりも小さな第2のサイズの第2階層要約データを生成する第2階層要約データ生成手段と、前記データベースより、前記第1階層要約データ生成手段により生成された前記第1階層要約データに基づいて、対応する動画像コンテンツを検索する第1の検索手段と、前記第1の検索手段により検索された動画像コンテンツのうち、前記第2階層要約データ生成手段により生成された前記第2階層要約データに基づいて、対応する動画像コンテンツを検索する第2の検索手段とを含む画像処理装置を制御するコンピュータに、前記第1階層要約データ生成手段における、動画像コンテンツを構成する画像のうち、前記第1の頻度で抽出される画像より第1のサイズの第1階層要約データを生成し、データベースに登録する第1階層要約データ生成ステップと、前記第2階層要約データ生成手段における、前記動画像コンテンツを構成する画像のうち、前記第1の頻度よりも高い第2の頻度で抽出される画像より、前記第1のサイズよりも小さな第2のサイズの第2階層要約データを生成する第2階層要約データ生成ステップと、前記第1の検索手段における、前記データベースより、前記第1階層要約データ生成ステップの処理により生成された前記第1階層要約データに基づいて、対応する動画像コンテンツを検索する第1の検索ステップと、前記第2の検索手段における、前記第1の検索ステップの処理により検索された動画像コンテンツのうち、前記第2階層要約データ生成ステップの処理により生成された前記第2階層要約データに基づいて、対応する動画像コンテンツを検索する第2の検索ステップとを実行させる。30

#### 【0015】

本発明の一側面においては、動画像コンテンツを構成する画像のうち、第1の頻度で抽出される画像より第1のサイズの第1階層要約データが生成されてデータベースに登録され、前記動画像コンテンツを構成する画像のうち、前記第1の頻度よりも高い第2の頻度で抽出される画像より、前記第1のサイズよりも小さな第2のサイズの第2階層要約データが生成され、前記データベースより、生成された前記第1階層要約データに基づいて、対応する動画像コンテンツが検索され、検索された動画像コンテンツのうち、生成された前記第2階層要約データに基づいて、対応する動画像コンテンツが検索される。40

#### 【0016】

本発明の画像処理装置は、独立した装置であっても良いし、画像処理を行うブロックで50

あっても良い。

【発明の効果】

【0017】

本発明の一側面によれば、多くの動画像コンテンツを容易に検索し、管理することが可能となる。

【図面の簡単な説明】

【0018】

【図1】本発明を適用した画像処理装置の一実施の形態の構成例を示すブロック図である。

【図2】画像蓄積処理を説明するフローチャートである。

10

【図3】第1階層要約データ生成処理を説明するフローチャートである。

【図4】第1階層要約データ生成処理を説明する図である。

【図5】クラスタリング処理を説明するフローチャートである。

【図6】第2階層要約データ生成処理を説明する図である。

【図7】初期クラスタリング処理を説明するフローチャートである。

【図8】初期クラスタリング処理を説明する図である。

【図9】初期クラスタリング処理を説明する図である。

【図10】検索抽出処理を説明するフローチャートである。

【図11】検索抽出処理を説明する図である。

【図12】検索抽出処理を説明する図である。

20

【図13】圧縮処理を説明するフローチャートである。

【図14】圧縮処理を説明する図である。

【図15】汎用のパーソナルコンピュータの構成例を説明する図である。

【発明を実施するための最良の形態】

【0019】

[画像処理装置の構成例]

図1は、本発明を適用した画像処理装置のハードウェアの一実施の形態の構成例を示している。図1の画像処理装置11は、動画像コンテンツを蓄積すると共に、必要に応じて検索して抽出するものである。

【0020】

30

画像処理装置11は、画像取得部21、バッファ22、動画像コンテンツデータ登録部23、記憶部24、検索抽出部25、コンテンツファイル圧縮部26、および表示部27を備えている。

【0021】

画像取得部21は、登録しようとする動画像コンテンツ、または検索しようとする動画像コンテンツの一部となるサンプルデータを取得し、バッファ22に一時的に記憶させる。サンプルデータは、例えば、インターネット上でダウンロード可能な動画像コンテンツの部分的なものや、関連する画像、およびサンプル動画像などである。

【0022】

動画像コンテンツデータ登録部23は、バッファ22に記憶されている、登録しようとする動画像コンテンツについて、第1階層要約データ、および第2階層要約データを生成する。そして、動画像コンテンツデータ登録部23は、第1階層要約データを記憶部24の第1階層要約データDB(データベース)101に登録する。また、動画像コンテンツデータ登録部23は、第2階層要約データからなる第2階層要約データファイル112と動画像コンテンツからなる動画像コンテンツファイル111とをコンテンツデータペア102として記憶部24に記憶させる。この際、記憶部24に記憶されている動画像コンテンツがクラスタリング(クラスに分類)されている場合、動画像コンテンツデータ登録部23は、第1階層要約データに基づいて、動画像コンテンツをいずれかのクラスに分類して、記憶させる。

【0023】

50

尚、図1の記憶部24においては、第1階層要約データDB101、および複数のクラスA乃至Dが含まれている。各クラスA乃至Dには、それぞれコンテンツデータペア102-1乃至102-a, 102-b乃至102-c, 102-d乃至102-e, 102-f乃至102-gが含まれている。さらに、コンテンツデータペア102-1乃至102-aには、それぞれ動画像コンテンツファイル111-1乃至111-a、および第2階層要約データファイル112-1乃至112-aが含まれている。また、コンテンツデータペア102-b乃至102-cには、それぞれ動画像コンテンツファイル111-b乃至111-c、および第2階層要約データファイル112-b乃至112-cが含まれている。さらに、コンテンツデータペア102-d乃至102-eには、それぞれ動画像コンテンツファイル111-d乃至111-e、および第2階層要約データファイル112-d乃至112-eが含まれている。また、コンテンツデータペア102-f乃至102-gには、それぞれ動画像コンテンツファイル111-f乃至111-g、および第2階層要約データファイル112-f乃至112-gが含まれている。

#### 【0024】

尚、これらを特に区別する必要がない場合、単に、コンテンツデータペア102、動画像コンテンツファイル111、および第2階層要約データファイル112と称するものとし、他の構成についても同様に称するものとする。また、これらのクラスA乃至Dの構成については、第1階層要約データDB101に管理される第1階層要約データに対して所属するクラスの情報が対応して記録されているのみであり、記憶部24内において、クラスA乃至Dといった個別のフォルダなどにより管理されているわけではない。図1の記憶部24内の記載は、あくまでもクラス構造を模式的に示したものであって、実際には、記憶部24内において、コンテンツデータペア102が個別に分類されたクラス単位で統一された領域に記録されているわけではない。また、コンテンツデータペア102における動画像コンテンツファイル111、および第2階層要約データファイル112については、対応付けて記録されるが、個別に存在するファイルであり、個別に管理可能なものである。

#### 【0025】

動画像コンテンツデータ登録部23は、アクティビティ計算部41、画像スライド部42、第1階層要約データ生成部43、第2階層要約データ生成部44、およびクラスタリング部45を含む。アクティビティ計算部41は、動画像コンテンツに含まれる画像の隣接画素間の画素値の差分値を画素値とした画像をアクティビティ画像として求める。画像スライド部42は、アクティビティ計算部41により計算されたアクティビティ画像に基づいて、アクティビティの最も高い位置を画像の中心としてスライドさせる。すなわち、画像の中でアクティビティの最も高い位置は、人間の視覚において意識を引き付ける位置であるので、その位置を画像の中心とすることにより、抽出される第1階層要約データ、および第2階層要約データが正規化されるようとする。

#### 【0026】

第1階層要約データ生成部43は、画像分割部61、RGB画素値平均値計算部62、およびシーンチェンジ検出部63を備えている。第1階層要約データ生成部43は、入力される動画像コンテンツの所定間隔毎のフレーム、または、シーンチェンジ検出部63により検出されたシーンチェンジのタイミングとなるフレームを注目フレームとして、その注目フレームより第1階層要約データを生成する。すなわち、第1階層要約データ生成部43は、画像分割部61を制御して、注目フレームを所定数の分割領域に分割させる。次に、第1階層要約データ生成部43は、RGB画素値平均値計算部62を制御して、分割領域毎にRGB(赤色画素、緑色画素、青色画素)のそれぞれの画素値の平均値を求める。そして、第1階層要約データ生成部43は、このようにして分割領域毎に求められたRGBのそれぞれの平均値を要素とし、100バイト程度の情報量からなるフレーム単位のベクトルを第1階層要約データとして求め、第1階層要約データDB101に登録する。

#### 【0027】

10

20

30

40

50

第2階層要約データ生成部44は、画像分割部71、およびRGB画素値平均値計算部72を備えており、動画像コンテンツの全フレームより第2階層要約データを生成する。尚、画像分割部71、およびRGB画素値平均値計算部72については、それぞれ第1階層要約データ生成部43の画像分割部61、およびRGB画素値平均値計算部62と同様のものである。

#### 【0028】

すなわち、第2階層要約データ生成部44は、画像分割部71を制御して、注目フレームを所定数の分割領域に分割させる。次に、第2階層要約データ生成部44は、RGB画素値平均値計算部72を制御して、分割領域毎にRGB(赤色画素、緑色画素、青色画素)のそれぞれの画素値の平均値を求めさせる。そして、第2階層要約データ生成部44は、このようにして分割領域毎に求められたRGBのそれぞれの平均値を要素とし、30バイト程度の情報量からなるフレーム単位のベクトルを第2階層要約データとして求める。さらに、第2階層要約データ生成部44は、このようにして求められた第2階層要約データを時系列配置したファイルを第2階層要約データファイル112と、動画像コンテンツファイル111とを、コンテンツデータペア102として記憶部24に記憶させる。

10

#### 【0029】

尚、第1階層要約データ、および第2階層要約データは、いずれにおいても、いずれの動画像コンテンツファイルに所属するものであるかを示す情報が含まれている。また、第1階層要約データは、30フレーム毎に100バイト程度の情報量の情報であり、第2階層要約データは、1フレーム毎に30バイト程度の情報量の情報であるものとしているが、第1階層要約データの方が、第2階層要約データよりもフレーム数の頻度が低く、情報量が多いものであれば、他のフレーム数毎であって、異なる情報量であってもよい。すなわち、第1階層要約データについては、第1階層要約データDB101において検索に利用されるため、単独で利用されるため、多少大きめのデータでもよい。しかしながら、第2階層要約データは、時系列に管理される複数のものを含む第2階層要約データファイル112の単位で利用されることになるため、個々には小さなデータサイズとする必要があるためである。

20

#### 【0030】

クラスタリング部45は、重心計算部81、距離計算部82、および初期登録部83を備えている。クラスタリング部45は、記憶部24に記憶されている複数のコンテンツデータペア102を第1階層要約データDB101に登録されている第1階層要約データに基づいて、類似したもの同士を同一クラスに分類する(クラスタリングする)。

30

#### 【0031】

クラスタリング部45は、既に、クラスタリングが完了している記憶部24に新たに動画像コンテンツファイル111を登録するとき、その新たに登録しようとする動画像コンテンツの第1階層要約データに基づいてクラスを分類する。すなわち、クラスタリング部45は、重心計算部81を制御して、各クラスに属するベクトルからなる第1階層要約データの重心ベクトルを計算させる。次に、クラスタリング部45は、距離計算部82を制御して各クラスの重心ベクトルと、新たに登録しようとする動画像コンテンツの第1階層要約データを構成するベクトルとの距離を計算させる。そして、クラスタリング部45は、新たに登録しようとする動画像コンテンツの第1階層要約データからなるベクトルとの距離が最小となる重心ベクトルのクラスに、新たに登録しようとする動画像コンテンツを分類する。

40

#### 【0032】

また、記憶部24に登録されているコンテンツデータペア102がクラスに分類されていない場合、クラスタリング部45は、各コンテンツデータペア102をクラスに分類する。すなわち、クラスタリング部45は、距離計算部82を制御して、各動画像コンテンツの第1階層要約データDB101に登録されている第1階層要約データを構成するベクトル間の距離を計算し、所定の距離の範囲内に存在する動画像コンテンツデータを同一のクラスに分類する。そして、クラスタリング部45は、初期登録部83を制御して、各動画

50

像コンテンツが分類された第1階層要約データのクラスを登録する。

#### 【0033】

検索抽出部25は、アクティビティ計算部151、画像スライド部152、第1階層予約データ検索部153、第2階層要約データ検索抽出部154、クラス検索部155、および検索結果表示画像生成部156を備えている。検索抽出部25は、画像取得部21により取得されバッファ22に記憶された検索しようとする動画像コンテンツに関連するサンプルデータに基づいて、動画像コンテンツを検索し、検索結果を表示部27に表示する。

#### 【0034】

より詳細には、アクティビティ計算部151、および画像スライド部152は、動画像コンテンツデータ登録部23におけるアクティビティ計算部41、および画像スライド部42と同様のものである。すなわち、アクティビティ計算部151は、検索しようとする動画像コンテンツに関連する画像(フレーム)の隣接画素間の差分値からなるアクティビティを計算し、アクティビティを画素値とするアクティビティ画像を生成する。画像スライド部152は、計算されたアクティビティ画像のうち、最もアクティビティの大きな位置を画像の中心位置にスライドさせる。

#### 【0035】

第1階層要約データ検索部153は、画像分割部171、RGBC画素値平均値計算部172、および比較部173を備えている。第1階層要約データ検索部153は、検索しようとする動画像コンテンツに関連するサンプルデータの代表画像より第1階層要約データを生成する。すなわち、第1階層要約データ検索部153は、画像分割部171を制御して、代表画像を所定数の分割領域に分割させる。次に、第1階層要約データ検索部153は、RGBC画素値平均値計算部172を制御して、分割領域毎にRGBC(赤色画素、緑色画素、青色画素)のそれぞれの画素値の平均値を求めさせる。そして、第1階層要約データ検索部153は、このようにして分割領域毎に求められたRGBCのそれぞれの平均値を要素とし、100バイト程度の情報量からなるフレーム単位のベクトルを第1階層要約データとして求める。尚、サンプルデータの代表画像は、サンプルデータを構成する動画像コンテンツの先頭画像、またはシーンチェンジのタイミングにおける画像である。

#### 【0036】

第1階層要約データ検索部153は、比較部173を制御して、サンプルデータより求められた代表画像の第1階層要約データと、第1階層要約データDB101に登録された第1階層要約データとをそれぞれ構成するベクトルとを比較し、一致、または類似しているか否かを判定させる。そして、第1階層要約データ検索部153は、サンプルデータの第1階層要約データに対応する検索対象となる動画像コンテンツの候補を検索する。この際、クラス検索部155により代表画像の第1階層要約データのベクトルと、各クラスに分類される第1階層要約データの重心ベクトルとの一致、または類似しているか否かの比較結果に基づいて、一致するクラスの重心ベクトルが求められていた場合、第1階層要約データ検索部153は、対象となるクラスに属する第1階層要約データと一致、または類似するか否かの比較結果により、サンプルデータの第1階層要約データに対応する動画像コンテンツを検索する。すなわち、各クラスは、第1階層要約データを構成するベクトル間の距離が近いもの同士が集められて構成される。このため、第1階層要約データ検索部153は、第1階層要約データDB101に登録されている第1階層要約データのうち、代表画像の第1階層要約データと一致する重心ベクトルを持つクラスに属する第1階層要約データのみを比較すればよいので検索速度を向上させることができる。

#### 【0037】

第2階層要約データ検索抽出部154は、検索しようとするサンプルデータの動画像コンテンツの動画像情報に基づいて、第2検索要約データを生成する。そして、第2階層要約データ検索抽出部154は、サンプルデータの第2検索要約データと、第1階層要約データ検索部153により検索された第1階層要約データの動画像コンテンツの第2階層要約データファイル112とが一致するものを検索する。すなわち、第2階層要約データ検

10

20

30

40

50

索抽出部 154 は、画像分割部 181 を制御して、検索しようとするサンプルデータの動画像コンテンツの動画像情報の各フレームを所定数の分割領域に分割させる。次に、第2階層要約データ検索抽出部 154 は、RGB 画素値平均値計算部 182 を制御して、分割領域毎に RGB (赤色画素、緑色画素、青色画素) のそれぞれの画素値の平均値を求める。そして、第2階層要約データ検索抽出部 154 は、このようにして分割領域毎に求められた RGB のそれぞれの平均値を要素とし、30 バイト程度の情報量からなるフレーム単位のベクトルをサンプルデータの第2階層要約データとして求める。すなわち、第2階層要約データファイル 112 は、毎フレームごとに求められた第2階層要約データが時系列に配置されたものである。従って、サンプルデータの第2階層要約データが時系列に配置されたものが、登録された第2階層要約データファイル 112 の時系列に配置された第2階層要約データ上で、1フレームずつスライドされるとき、対応するフレーム同士の第2階層要約データ間の差分が順次求められるとき、その差分の最小値が所定の閾値よりも小さければ、検索対象となる動画像コンテンツと一致、または類似しているものとして検索される。また、検索対象となる動画像コンテンツと一致、または類似していれば、その差分が最小となるときのスライド位置が、そのサンプルデータの、検索しようとする動画像コンテンツ上の再生位置として特定される。

## 【0038】

クラス検索部 155 は、サンプルデータの代表画像より求められた第1階層要約データを構成するベクトルと、記憶部 24 に記憶されている第1階層要約データ DB 101 に登録された各クラスの重心ベクトルとの距離を求め、所定の範囲内のクラスを検索する。すなわち、クラス検索部 155 は、重心計算部 191 を制御して、クラス毎にそのクラスに属する第1階層要約データを構成するベクトルの重心ベクトルを求める。そして、クラス検索部 155 は、代表画像の第1階層要約データからなるベクトルと、重心ベクトルとの距離を求めて所定の距離内の重心ベクトルのクラスを検索する。

## 【0039】

検索結果表示画像生成部 156 は、第2階層要約データ検索抽出部 154 により検索された動画像コンテンツの第1階層要約データに対応する画像を、LCD (Liquid Crystal Display) などからなる表示部 27 に表示する。

## 【0040】

コンテンツファイル圧縮部 26 は、同一部分検索部 211、不一致部分検索部 212、編集部 213、および更新部 214 を備えている。コンテンツファイル圧縮部 26 は、第2階層要約データ検索抽出部 154 により検索された第2階層要約データファイル 112 における同一部分、および不一致部分を検索すると共に、相互に同一部分を1つのみ残して消去し、その他の不一致部分と接合するなどの編集を行い、動画像コンテンツファイル 111 のデータ量を圧縮する。すなわち、コンテンツファイル圧縮部 26 は、同一部分検索部 211 を制御して、検索された第2階層要約データファイル 112 のうち、同一部分となる範囲を検索させる。コンテンツファイル圧縮部 26 は、不一致部分検索部 212 を制御して、検索されたもののうち、不一致部分となる範囲を検索させる。次に、コンテンツファイル圧縮部 26 は、編集部 213 を制御して、同一部分となる範囲を1箇所残すのみとし、その他を削除させ、必要に応じて不一致部分を接合するように編集させる。そして、コンテンツファイル圧縮部 26 は、新たに生成された動画像コンテンツファイル 111 により動画像コンテンツファイル 111 を更新する。

## 【0041】

## [動画像コンテンツ蓄積処理]

次に、図 1 の画像処理装置 11 における動画像コンテンツ蓄積処理について説明する。

## 【0042】

ステップ S11 において、画像取得部 21 は、記憶部 24 に蓄積すべき動画像コンテンツを取得し、バッファ 22 に格納する。

## 【0043】

ステップ S12 において、動画像コンテンツデータ登録部 23 は、バッファ 22 より動

10

20

30

40

50

画像コンテンツを読み出して、第1階層要約データ生成処理を実行し、第1階層要約データを生成して記憶部24の第1階層要約データDB101に登録する。

【0044】

[第1階層要約データ生成処理]

ここで、図3のフローチャートを参照して、第1階層要約データ生成処理について説明する。

【0045】

ステップS31において、動画像コンテンツデータ登録部23は、フレーム数をカウントするためのカウンタ $f$ を1に初期化する。

【0046】

ステップS32において、動画像コンテンツデータ登録部23は、順次動画像コンテンツを読み出し、未処理の次のフレーム画像が存在するか否かを判定する。ステップS32において、未処理の次のフレーム画像が存在しない場合、処理は終了する。一方、ステップS32において、次の動画像コンテンツのフレーム画像が存在する場合、処理は、ステップS33に進む。

10

【0047】

ステップS33において、第1階層要約データ生成部43は、シーンチェンジ検出部63を制御して、供給されてくるフレーム画像が直前の画像との比較によりシーンチェンジしているフレーム画像であるか否かを判定する。すなわち、シーンチェンジ検出部63は、例えば、処理対象であるフレーム画像と直前画像との各画素の画素値の差分値の総和を求めて、所定の閾値よりも大きくなる場合、シーンチェンジが発生したものとみなす。すなわち、連続する動画像の場合、直前のフレーム画像との変化は小さいと考えられるため、画素値の差分値の総和は、所定の閾値よりも小さいと考えられる。しかしながら、シーンチェンジが発生する場合、直前のフレーム画像とは全く異なる画像となることが多いため、各画素間の画素値の差分値の総和は大きく変化するものと考えられる。そこで、シーンチェンジ検出部63は、各画素間の画素値の差分値の総和と、所定の閾値との比較によりシーンチェンジが発生しているか否かを判定する。

20

【0048】

ステップS33において、例えば、直前の画像との各画素の画素値の差分値の総和が所定の閾値よりも小さく、シーンチェンジが発生していないものとみなされた場合、処理は、ステップS34に進む。

30

【0049】

ステップS34において、動画像コンテンツデータ登録部23は、カウンタ $f$ が30であるか否かを判定し、30ではない場合、処理は、ステップS35に進み、カウンタ $f$ が1インクリメントされて、処理は、ステップS32に戻る。すなわち、シーンチェンジがない限り、30フレームが経過するまで、ステップS32乃至S35の処理が繰り返される。

【0050】

そして、ステップS33において、直前の画像との各画素の画素値の差分値の総和が所定の閾値よりも大きく、シーンチェンジが発生しているとみなされた場合、または、ステップS34において、カウンタ $f$ が30になった場合、処理は、ステップS36に進む。

40

【0051】

ステップS36において、動画像コンテンツデータ登録部23は、バッファ22より直前のフレームから30フレーム目にあたる注目フレームの画像をバッファ22より抽出する。

【0052】

ステップS37において、動画像コンテンツデータ登録部23は、アクティビティ計算部41を制御して、抽出した注目フレームの画像におけるアクティビティを計算させる。より詳細には、アクティビティ計算部41は、各画素について水平垂直に隣接する画素間の差分の合計を画素のアクティビティとして計算し、全画素について同様に計算すること

50

で、全画素の画素値がアクティビティからなるアクティビティ画像を計算により求める。

【0053】

ステップS38において、動画像コンテンツデータ登録部23は、画像スライド部42を制御して、求められたアクティビティ画像のうち、アクティビティが最も高い画素の位置を、注目フレームの画像における中心位置に移動させる。すなわち、アクティビティが高い画素位置は、人間の視覚において、最も視線が向けられ易い位置であるため、注目フレームにおけるその最も視線が向けられ易い位置が画像の中心位置となるようにスライドされて移動されることで、以降の処理が正規化される。尚、複数に最も高いアクティビティを取る画素が存在する場合、その最もアクティビティが高い画素により構成される被写体に視線が向けられ易いことになるため、その場合、それらの画素の重心位置を注目フレームの画像の中心位置とするようにしてもよい。

10

【0054】

ステップS39において、第1階層要約データ生成部43は、画像分割部61を制御して、中心位置がスライドされた第Fフレームを、N個の分割領域P(n)に分割する。すなわち、例えば、N=32である場合、第Fフレームの画像は、図4で示されるように分割領域P(1)乃至P(32)に分割される。

【0055】

ステップS40において、第1階層要約データ生成部43は、分割領域P(n)をカウントするためのカウンタnを1に初期化する。

20

【0056】

ステップS41において、第1階層要約データ生成部43は、RGB画素値平均値計算部62を制御して、分割領域P(n)のRGBの各画素について、画素値の平均値をそれぞれ求める。

20

【0057】

ステップS42において、第1階層要約データ生成部43は、カウンタnが所定の分割数Nであるか否かを判定する。すなわち、図4の場合、第1階層要約データ生成部43は、カウンタnが所定の分割数N=32であるか否かを判定する。ステップS42において、カウンタnが所定の分割数Nではない場合、ステップS43において、第1階層要約データ生成部43は、カウンタnを1インクリメントして、処理は、ステップS41に戻る。すなわち、全ての分割領域P(n)のRGBの各画素について、画素値の平均値がそれぞれ求められる。

30

【0058】

そして、ステップS42において、カウンタnが分割数Nに達した場合、処理は、ステップS43に進む。

【0059】

ステップS44において、第1階層要約データ生成部43は、分割領域P(1)乃至P(N)のそれぞれのRGBの平均値を要素とするベクトルからなり、100バイト程度のデータ量の第1階層要約データC1を第1階層要約データDB101に登録して記憶部24に記憶させ、処理は、ステップS31に戻る。

40

【0060】

以上の処理により動画像コンテンツのうち、先頭から順次30フレーム毎に、または、シーンチェンジがあるとき、そのときのフレームが注目フレームとされて、順次、所定数の分割領域P(n)に分割されて、それぞれのRGBの平均値が求められる。そして、求められた各分割領域P(n)のRGBの平均値を要素とするベクトルであって、100バイト程度のデータ量の第1階層要約データC1が、第1階層要約データDB101に順次登録される。すなわち、例えば、図4で示されるように、注目フレームが32分割され、RGBのそれぞれの階調が16であって、1バイトで表現されるとき、第1階層要約データC1は、各要素が1バイトで表現され、96バイトからなる96次元のベクトルにより構成されるものとなる。このため、第1階層要約データDB101に登録されている情報は、時系列の情報を含まない第1階層要約データを構成するベクトルの集合からなる特徴量であ

50

ると考えることができる。

【0061】

ここで、図2のフローチャートの説明に戻る。

【0062】

ステップS12において、登録しようとする動画像データの第1階層要約データが記憶部24の第1階層要約データDB101に登録されると、ステップS13において、クラスタリング部45は、クラスタリング処理を実行し、登録しようとする動画像コンテンツの第1階層要約データのクラスを分類する。

【0063】

[クラスタリング処理]

ここで、図5のフローチャートを参照し、クラスタリング処理について説明する。

10

【0064】

ステップS51において、クラスタリング部45は、第1階層要約データDB101に登録されている第1階層要約データのうち、登録しようとする動画像コンテンツを除いたものであって、クラスに分類されているものを抽出する。さらに、クラスタリング部45は、重心計算部81を制御して、同一クラスに分類されている第1階層要約データからなるベクトルの重心ベクトルを計算させる。すなわち、例えば、図1のクラスAにおいては、クラスタリング部45は、重心計算部81を制御して、クラスAのコンテンツデータペア102-1乃至102-aのそれぞれのに対応する第1階層要約データからなるベクトルの重心ベクトルを求めさせる。同様に、クラスB乃至Dについても、それぞれの第1階層要約データからなるベクトルの重心ベクトルが求められる。

20

【0065】

ステップS52において、クラスタリング部45は、登録しようとする動画像コンテンツの第1階層要約データのうち、未処理のものを処理対象の第1階層要約データC1に設定する。

【0066】

ステップS53において、クラスタリング部45は、クラスをカウントするカウンタmを1に初期化する。

【0067】

ステップS54において、クラスタリング部45は、距離計算部82を制御して、処理対象の第1階層要約データC1を構成するベクトルと、クラスmの重心ベクトルとの距離dmを計算させる。より詳細には、距離計算部82は、例えば、自乗ノルムを用いて、それぞれのベクトル間距離dmを求める。

30

【0068】

ステップS55において、クラスタリング部45は、求められたベクトル間の距離dmが所定の閾値th1よりも小さいか否かを判定し、例えば、ベクトル間の距離dmが所定の閾値th1よりも小さい場合、処理は、ステップS56に進む。

【0069】

ステップS56において、クラスタリング部45は、処理対象とする第1階層要約データC1をクラスmに仮登録する。

40

【0070】

ステップS57において、クラスタリング部45は、カウンタmが分類されているクラス数Mに達したか否かを判定し、達していない場合、処理は、ステップS58に進む。そして、ステップS58において、クラスタリング部45は、カウンタmを1インクリメントして、処理は、ステップS54に戻る。すなわち、全てのクラスの重心ベクトルとのベクトル間距離が所定の閾値th1よりも小さいか否かが判定されるまで、ステップS54乃至S58の処理が繰り返される。

【0071】

そして、ステップS57において、カウンタmがクラス数Mであると判定された場合、ステップS59において、クラスタリング部45は、処理対象となっている第1階層要約

50

データ C 1 が、いずれかのクラスに分類されて仮登録されているか否かを判定する。そして、ステップ S 5 9 において、いずれかのクラスに分類されて仮登録されている場合、処理は、ステップ S 6 0 に進む。

【 0 0 7 2 】

ステップ S 6 0 において、クラスタリング部 4 5 は、処理対象とする第 1 階層要約データ C 1 が仮登録されているクラスのうち、距離  $d_m$  が最小となるクラス  $m$  を処理対象とする第 1 階層要約データ C 1 のクラス  $m$  に分類する。すなわち、処理対象とする第 1 階層要約データ C 1 が、複数のクラスに分類されている場合、そのいずれかに分類する必要があるため、第 1 階層要約データ C 1 のベクトルが、重心ベクトルと最も近い（最も類似している）クラスに分類される。

10

【 0 0 7 3 】

ステップ S 6 1 において、クラスタリング部 4 5 は、分類されたクラスに、同一の動画像コンテンツの第 1 階層要約データが既に登録されているか否かを判定する。ステップ S 6 1 において、例えば、分類されたクラスに、同一の動画像コンテンツの第 1 階層要約データが登録されていない場合、ステップ S 6 2 において、クラスタリング部 4 5 は、分類されたクラスに処理対象となる第 1 階層要約データを分類し、第 1 階層要約データ DB 1 0 1 に登録する。

【 0 0 7 4 】

一方、ステップ S 5 9 において、いずれのクラスにも分類されていない場合、または、ステップ S 6 1 において、既に、同一の動画像コンテンツの第 1 階層要約データが登録されている場合、処理は、ステップ S 6 3 に進む。

20

【 0 0 7 5 】

ステップ S 6 3 において、クラスタリング部 4 5 は、登録しようとする動画像コンテンツについて、未処理の第 1 階層要約データが存在するか否かを判定し、未処理の第 1 階層要約データが存在する場合、処理は、ステップ S 5 2 に戻る。すなわち、サンプルデータの動画像コンテンツについて、未処理の第 1 階層要約データが存在しない状態となるまで、ステップ S 5 2 乃至 S 6 3 の処理が繰り返される。

【 0 0 7 6 】

そして、ステップ S 6 3 において、登録しようとする動画像コンテンツについて、未処理の第 1 階層要約データが存在しないと判定された場合、処理は、終了する。

30

【 0 0 7 7 】

以上の処理により、登録しようとする動画像コンテンツの第 1 階層要約データについて、各クラスに最大 1 つの第 1 階層要約データが分類される。

【 0 0 7 8 】

ここで、図 2 のフローチャートの説明に戻る。

【 0 0 7 9 】

ステップ S 1 3 において、登録しようとする動画像コンテンツの第 1 階層要約データについて、各クラスに最大 1 つの第 1 階層要約データが分類されると、処理は、ステップ S 1 4 に進む。

【 0 0 8 0 】

ステップ S 1 4 において、第 2 階層要約データ生成部 4 4 は、第 2 階層要約データ生成処理を実行し、第 2 階層要約データからなる第 2 階層要約データファイル 1 1 2 を生成し、動画像コンテンツファイル 1 1 1 と共にコンテンツデータペア 1 0 2 として記憶部 2 4 に記憶させる。

40

【 0 0 8 1 】

[ 第 2 階層要約データ生成処理 ]

ここで、図 6 のフローチャートを参照して、第 2 階層要約データ生成処理について説明する。

【 0 0 8 2 】

ステップ S 8 1 において、動画像コンテンツデータ登録部 2 3 は、フレーム数をカウン

50

トするためのカウンタ F を 1 に初期化する。

【0083】

ステップ S 8 2 において、動画像コンテンツデータ登録部 2 3 は、バッファ 2 2 に蓄積されている、登録しようとする動画像コンテンツに次のフレーム画像が存在するか否かを判定し、次のフレーム画像が存在しない場合、処理は終了する。また、ステップ S 8 2 において、次のフレーム画像が存在する場合、処理は、ステップ S 8 3 に進む。

【0084】

ステップ S 8 3 において、動画像コンテンツデータ登録部 2 3 は、バッファ 2 2 よりフレーム ( F ) を注目フレームの画像としてバッファ 2 2 より抽出する。

【0085】

ステップ S 8 4 において、動画像コンテンツデータ登録部 2 3 は、アクティビティ計算部 4 1 を制御して、抽出した注目フレームの画像におけるアクティビティ画像を計算させる。

【0086】

ステップ S 8 5 において、動画像コンテンツデータ登録部 2 3 は、画像スライド部 4 2 を制御して、求められたアクティビティ画像のうち、アクティビティが最も高い画素の位置を、注目フレームの画像における中心位置に移動させる。

【0087】

ステップ S 8 6 において、第 2 階層要約データ生成部 4 4 は、画像分割部 7 1 を制御して、中心位置がスライドされた第 F フレームを、N 個の分割領域 P ( n ) に分割する。

10

【0088】

ステップ S 8 7 において、第 2 階層要約データ生成部 4 4 は、分割領域 P ( n ) をカウントするためのカウンタ n を 1 に初期化する。

【0089】

ステップ S 8 8 において、第 2 階層要約データ生成部 4 4 は、R G B 画素値平均値計算部 7 2 を制御して、分割領域 P ( n ) の R G B の各画素について、画素値の平均値をそれぞれ求める。

【0090】

ステップ S 8 9 において、第 2 階層要約データ生成部 4 4 は、カウンタ n が所定の分割数 N であるか否かを判定する。ステップ S 8 9 において、カウンタ n が所定の分割数 N ではない場合、ステップ S 9 0 において、第 2 階層要約データ生成部 4 4 は、カウンタ n を 1 インクリメントして、処理は、ステップ S 8 8 に戻る。すなわち、全ての分割領域 P ( n ) の R G B の各画素について、画素値の平均値がそれぞれ求められる。

30

【0091】

そして、ステップ S 9 0 において、カウンタ n が分割数 N に達した場合、処理は、ステップ S 9 1 に進む。

【0092】

ステップ S 9 1 において、第 2 階層要約データ生成部 4 4 は、分割領域 P ( 1 ) 乃至 P ( N ) のそれぞれの R G B の平均値を要素とするベクトルからなり、30 バイト程度のデータ量の第 2 階層要約データ C 2 を新規の第 2 階層要約データファイル 112 に登録する。この際、第 2 階層要約データ生成部 4 4 は、処理されたフレーム順に第 2 階層要約データ C 2 を新規の第 2 階層要約データファイル 112 に登録して記憶部 24 に記憶させる。

40

【0093】

そして、ステップ S 9 2 において、カウンタ F が 1 インクリメントされると、処理は、ステップ S 8 2 に戻る。

【0094】

以上の処理により動画像コンテンツの各フレームが順次注目フレームとされて、所定数の分割領域 P ( n ) に分割されて、それぞれの R G B の平均値が求められる。そして、求められた各分割領域 P ( n ) の R G B の平均値を要素とするベクトルであって、30 バイト程度のデータ量の第 2 階層要約データ C 2 が、順次、処理されたフレーム順に第 2 階層

50

要約データファイル112に登録される。すなわち、例えば、図4で示されるように、注目フレームが32分割され、RGBのそれぞれの階調が4であって、2ビットで表現されるとき、1つの第2階層要約データC2は、各要素が2ビットで表現され、24バイトからなる96次元のベクトルにより構成されるものとなる。さらに、第2階層要約データファイル112には、この第2階層要約データC2が時系列にフレーム順に記憶される構造となる。このため、第2階層要約データファイル112に記憶されている第2階層要約データC2をフレーム順に配置したデータは、時系列の情報を含む第2階層要約データC2のベクトルからなる特徴量であるといえる。

【0095】

ここで、図2のフローチャートの説明に戻る。

10

【0096】

ステップS14において、第2階層要約データ生成処理により第2階層要約データファイル112が生成されると、処理は、ステップS15に進む。

【0097】

ステップS15において、動画像コンテンツデータ登録部23は、登録しようとする動画像コンテンツからなる動画像コンテンツファイル111と対応する第2階層要約データファイル112とからなるコンテンツデータペア102を記憶部24に登録する。

【0098】

以上の処理により動画像コンテンツを登録する際、所定間隔毎のフレーム単位、またはシーンチェンジのあったフレーム単位で抽出される第1階層要約データが第1階層要約データDB101に登録される。また、毎フレーム単位で抽出される第2階層要約データがフレーム順に配置された状態で第2階層要約データファイル112として登録され、動画像コンテンツファイル111とのコンテンツデータペア102として記憶部24に登録される。

20

【0099】

尚、第1階層要約データDB101の情報は、毎フレーム単位ではないが、所定間隔のフレーム単位、またはシーンチェンジのあるフレーム単位で抽出される第1階層要約データC1そのものがデータベースとして登録されているのみである。このため、第1階層要約データDB101の情報は、時系列の情報を持たない代わりに、データ検索に優れたデータベースとして登録されている。このため、第1階層要約データDB101の情報は、後述する検索しようとするサンプルデータの動画像コンテンツのデータとの照合により、検索しようとするサンプルデータの動画像コンテンツを特定するのに優れている特徴量であるといえる。

30

【0100】

また、第1階層要約データについては、第1階層要約データDB101において、クラスタリングされて登録されているため、第1階層要約データを検索する際、クラス毎の重心ベクトル等を用いて、いずれのクラスに所属しているのかを特定した後、さらに、特定されたクラス内の第1階層要約データを検索するようにすることで、検索速度を向上させることが可能となる。

40

【0101】

[初期クラスタリング処理]

さらに、以上においては、既に記憶部24に登録されている動画像コンテンツファイル111は、いずれもクラスタリングされていることが前提となる処理であった。しかしながら、クラスタリングされることなく動画像コンテンツファイル111が記憶部24に登録されている場合、初期クラスタリング処理が必要となる。そこで、次に、図7のフローチャートを参照して、初期クラスタリング処理について説明する。尚、第1階層要約データ生成処理により、全ての動画像コンテンツについての第1階層要約データが、第1階層要約データDB101に登録済みであることを前提とする。このため、図2のフローチャートにおけるステップS13の処理を除くステップS11, S12, S14, S15の処理によりクラスタリングされていない状態で記憶部24に動画像コンテンツが記憶されてい

50

るものとする。

【0102】

ステップS111において、クラスタリング部45は、第1階層要約データDB101のうち、注目する第1階層要約データC1(q)と、比較対象となる第1階層要約データC1(r)とを識別するカウンタq,rをq=1,r=2に初期化する。

【0103】

ステップS112において、クラスタリング部45は、距離計算部82を制御して、第1階層要約データC1(q)と第1階層要約データC1(r)との距離dmを計算させる。

【0104】

ステップS113において、クラスタリング部45は、計算された距離dmが、所定の閾値th1よりも小さいか否かを判定し、例えば、距離dmが、所定の閾値th1よりも小さい場合、すなわち、類似していると判定された場合、処理は、ステップS114に進む。

【0105】

ステップS114において、クラスタリング部45は、初期登録部83を制御して、第1階層要約データC1(r)の属する動画像コンテンツの他の第1階層要約データC1が、注目する第1階層要約データC1(q)と同一のクラスに既に分類されて登録されているか否かを判定する。

【0106】

ステップS114において、第1階層要約データC1(r)の属する動画像コンテンツの他の第1階層要約データC1が、注目する第1階層要約データC1(q)と同一のクラスに既に分類されて登録されていない場合、処理は、ステップS115に進む。

【0107】

ステップS115において、初期登録部83は、第1階層要約データC1(r)を注目する第1階層要約データC1(q)と同一のクラスに分類して登録する。

【0108】

すなわち、注目する第1階層要約データC1(q)と同一のクラスとみなされたその他の比較対象となる第1階層要約データC1(r)が既に登録されていると、同一のクラスに同一の動画像コンテンツから異なる第1階層要約データが複数登録されることになる。クラスとは、動画像コンテンツ同士で類似している、または同一であるフレーム画像をまとめたものである。従って、同一の動画像コンテンツの第1階層要約データが、複数も同一のクラスに登録する必要がない。このため、既に登録されたものがない場合にのみ、比較対象となる第1階層要約データC1(r)が注目する第1階層要約データC1(q)と同一のクラスに分類されて登録される。

【0109】

ステップS116において、クラスタリング部45は、カウンタrが、第1階層要約データC1の総数Qとなったか否かを判定し、カウンタrが総数Qではない場合、ステップS117において、カウンタrを1インクリメントして処理は、ステップS112に戻る。

【0110】

一方、ステップS113において、距離dmが所定の閾値th1より大きく、また、ステップS114において、注目する第1階層要約データC1(q)と同一のクラスとみなされたその他の比較対象となる第1階層要約データC1(r)が既に登録されている場合、ステップS115の処理はスキップされて、処理は、ステップS116に進む。

【0111】

すなわち、注目する第1階層要約データC1(q)と、それ以外の全ての比較対象となる第1階層要約データC1(r)との比較処理が全て完了するまで、ステップS112乃至S116の処理が繰り返される。

【0112】

10

20

30

40

50

そして、ステップ S 116において、カウンタ  $r$  が総数  $Q$  になった場合、ステップ S 118において、クラスタリング部 45 は、カウンタ  $q$  が総数  $Q$  になったか否かを判定する。ステップ S 116において、カウンタ  $q$  が総数  $Q$  ではない場合、ステップ S 119において、クラスタリング部 45 は、カウンタ  $q$  を 1 インクリメントし、さらに、ステップ S 120において、カウンタ  $r$  をカウンタ  $q$  より 1 大きな値にする。そして、処理は、ステップ S 112 に戻る。

#### 【0113】

すなわち、第 1 階層要約データ DB101 の第 1 階層要約データ間の比較をするため、重複した比較をする必要がないので、注目する第 1 階層要約データ C1( $q$ ) に対して、比較対象となる第 1 階層要約データ C1( $r$ ) は、カウンタ  $q$  より上位のものであればよい。そこで、カウンタ  $r$  は、 $q + 1$  から総数  $Q$  までの処理となる。そして、これらの比較が、全ての第 1 階層要約データ間でなされるまで、ステップ S 112 乃至 S 120 の処理が繰り返される。

#### 【0114】

そして、ステップ S 118において、カウンタ  $q$  が総数  $Q$  であると判定された場合、処理は終了する。

#### 【0115】

以上の処理により、第 1 階層要約データ間で類似しているもの、または同一のものについて、それぞれが同一のクラスに分類され、さらに、各クラスには同一の動画像コンテンツに属する第 1 階層要約データは、最大 1 個となるように分類される。この結果、例えば、異なる動画像コンテンツから 6 種類の第 1 階層要約データが集められて 1 クラスが設けられるような場合、図 8 の最上段の 6 枚のフレーム画像で示されるように、同一の人物が存在するような画像が抽出される。または、図 8 の上から 2 段目の 6 枚のフレーム画像で示されるように、同一の灯台と海が存在するような画像や、図 8 の上から 3 段目の 6 枚のフレーム画像で示されるように、2 機の航空機が存在するような画像や、図 8 の上から 4 段目の 6 枚のフレーム画像で示されるように、同一のロケット発射台が存在するような画像が抽出され、それぞれの 6 枚が同一のクラスとして分類される。すなわち、図 8 における各段の 6 枚のフレーム画像の第 1 階層要約データは、相互にベクトル間距離が所定の距離内の画像である。

#### 【0116】

尚、同一のクラスに分類される第 1 階層要約データ数毎にクラス数を求めた場合、例えば、図 9 で示されるような傾向が見られることが知られている。すなわち、図 9 においては、縦軸が、同一のクラスに分類される第 1 階層要約データ数（同一のクラスに分類されるフレーム画像数）を示し、横軸が、クラス数を示している。この結果、同一のクラスに分類される第 1 階層要約データ数が少数（例えば、3 乃至 5 個）のクラス数は非常に多く、同一のクラスに分類される第 1 階層要約データ数が多数（例えば、16 個以上）であると極端に少ないことがわかる。このことから、第 1 階層要約データ数の多いクラスについては、クラスを特定すれば、対応する動画像コンテンツも特定されることがわかる。従って、動画像コンテンツを検索するときには、クラスを特定し、特定したクラスに属する第 1 階層要約データとの比較をすることで、検索処理負荷を低減し、検索速度を向上させることが可能になることがわかる。また、図 9 で示されるような、クラスの分布をそのまま用いて、データマイニングすることも可能となり、例えば、毎日同一の番組を動画像コンテンツとして蓄積するような場合、月間の利用が多かったシーンを統計的に分析することも可能である。

#### 【0117】

##### [検索抽出処理]

次に、図 10 のフローチャートを参照して、検索抽出処理について説明する。

#### 【0118】

ステップ S 141において、画像取得部 21 は、検索しようとする動画像コンテンツのサンプルデータを構成する動画像コンテンツを取得し、バッファ 22 に格納する。

10

20

30

40

50

## 【0119】

ステップS142において、検索抽出部25は、バッファ22に格納されているサンプルデータの動画像より代表画像となるフレームの画像を抽出する。そして、検索抽出部25は、アクティビティ計算部151を制御して、検索しようとする動画像コンテンツの代表画像となるフレームにおけるアクティビティ画像を計算させる。

## 【0120】

ステップS143において、検索抽出部25は、画像スライド部152を制御して、求められたアクティビティ画像のうち、アクティビティが最も高い画素の位置を、代表画像となるフレームの画像における中心位置に移動させる。

## 【0121】

ステップS144において、第1階層要約データ検索部153は、画像分割部171を制御して、中心位置がスライドされた代表画像となるフレームを、N個の分割領域P(n)に分割する。

10

## 【0122】

ステップS145において、第1階層要約データ検索部153は、分割領域P(n)をカウントするためのカウンタnを1に初期化する。

## 【0123】

ステップS146において、第1階層要約データ検索部153は、RGB画素値平均値計算部172を制御して、分割領域P(n)のRGBの各画素について、画素値の平均値をそれぞれ求める。

20

## 【0124】

ステップS147において、第1階層要約データ検索部153は、カウンタnが所定の分割数Nであるか否かを判定する。ステップS147において、カウンタnが所定の分割数Nではない場合、ステップS148において、第1階層要約データ検索部153は、カウンタnを1インクリメントして、処理は、ステップS146に戻る。すなわち、全ての分割領域P(n)のRGBの各画素について、画素値の平均値がそれぞれ求められる。

## 【0125】

そして、ステップS147において、カウンタnが分割数Nに達した場合、処理は、ステップS149に進む。

## 【0126】

30

ステップS149において、第1階層要約データ検索部153は、分割領域P(1)乃至P(N)のそれぞれのRGBの平均値を要素とするベクトルからなり、100バイト程度のデータ量の、検索しようとする動画像コンテンツのサンプルデータの代表画像の第1階層要約データC1tを記憶し、処理は、ステップS150に進む。

## 【0127】

ステップS150において、クラス検索部155は、第1階層要約データDB101に登録されている第1階層要約データのうち、クラスに分類されているものを抽出する。さらに、クラス検索部155は、重心計算部191を制御して、同一クラスに分類されている第1階層要約データからなるベクトルの重心ベクトルを計算させる。

## 【0128】

40

ステップS151において、クラス検索部155は、クラスを識別するカウンタmを1に初期化する。

## 【0129】

ステップS152において、クラス検索部155は、比較部192を制御して、代表画像のフレームの第1階層要約データC1tのベクトルと、クラスmの重心ベクトルとを比較し、一致(または類似)するか否かを判定する。このとき、例えば、3ビットで0乃至7の数値を用いて表現した(3, 4, 0, 4)および(2, 4, 1, 4)は、例えば、2ビットで0乃至3の数値を用いて表現した(1, 2, 0, 2)および(1, 2, 0, 2)に写像される。このため、このようなビット変換により一致する場合については、一致する、または類似するものと判定するようにしてもよい。また、例えば3ビットで、(3,

50

$(2, 3, 0)$  および  $(4, 2, 4, 0)$  の 2 つのベクトルの場合、そのまま 2 ビットにすると、 $(1, 1, 1, 0)$  および  $(2, 1, 2, 0)$  となり、一致しないが、量子化ステップの半分、すなわち 1 を加算すると、 $(4, 3, 4, 1)$  および  $(5, 3, 5, 1)$  となる。そこで、これを 2 ビットにすると  $(2, 1, 2, 0)$  および  $(2, 1, 2, 0)$  となって、一致させることができる。すなわち、量子化を行う際、予め量子化ステップの半分の値を加算しておき、その後、2 ビットに変換することで一致するような場合も、一致する、または類似するものとみなすようにしてもよい。このようにすることで、数値上の比較は完全一致とすることで、極端に検索対象が狭くなることによる、不一致の可能性を低減させることが可能となる。

## 【0130】

10

ステップ S 152 において、第 1 階層要約データ  $C_{1t}$  のベクトルと、クラス  $m$  の重心ベクトルとが一致する場合、ステップ S 153 において、クラス検索部 155 は、クラス  $m$  を登録する。また、ステップ S 152 において、第 1 階層要約データ  $C_{1t}$  のベクトルと、クラス  $m$  の重心ベクトルとが一致しない場合、ステップ S 153 の処理は、スキップされる。

## 【0131】

ステップ S 154 において、クラス検索部 155 は、カウンタ  $m$  がクラス数  $M$  に達したか否かを判定し、クラス数  $M$  に達していない場合、ステップ S 155 において、カウンタ  $m$  を 1 インクリメントして、処理は、ステップ S 152 に進む。

## 【0132】

20

そして、全てのクラス  $m$  の重心ベクトルとの比較がなされると、ステップ S 154 において、カウンタ  $m$  がクラス数  $M$  に達したものとみなされ、処理は、ステップ S 156 に進む。

## 【0133】

ステップ S 156 において、第 1 階層要約データ検索部 153 は、上述したステップ S 153 の処理により登録された、クラス  $m$  の集合を処理対象のクラス  $x$  とする。尚、クラス  $x$  は複数のクラス  $m$  とするようにしてもよい。

## 【0134】

ステップ S 157 において、第 1 階層要約データ検索部 153 は、クラス  $x$  に属する第 1 階層要約データ  $C_{1(s)}$  を識別するカウンタ  $s$  を 1 に初期化する。

30

## 【0135】

ステップ S 158 において、第 1 階層要約データ検索部 153 は、比較部 173 を制御して、代表画像のフレームの第 1 階層要約データ  $C_{1t}$  のベクトルと、クラス  $x$  に属する第 1 階層要約データ  $C_{1(s)}$  のベクトルとが一致するか否かを比較し、一致するか否かを判定する。尚、ベクトル間の一致については、ステップ S 152 における処理で説明したように、ビット数変換や、量子化ステップ幅の半分を加算した上でビット数を変化させるといった処理により一致しても、一致したものとして扱うものとする。

## 【0136】

ステップ S 158 において、第 1 階層要約データ  $C_{1t}$  のベクトルと、クラス  $x$  に属する第 1 階層要約データ  $C_{1(s)}$  のベクトルとが一致する場合、ステップ S 159 において、第 1 階層要約データ検索部 153 は、クラス  $x$  に属する第 1 階層要約データ  $C_{1(s)}$  のベクトルを登録する。また、ステップ S 158 において、第 1 階層要約データ  $C_{1t}$  のベクトルと、クラス  $x$  に属する第 1 階層要約データ  $C_{1(s)}$  のベクトルとが一致しない場合、ステップ S 159 の処理は、スキップされる。

40

## 【0137】

ステップ S 160 において、第 1 階層要約データ検索部 153 は、カウンタ  $s$  がクラス  $x$  に属する第 1 階層要約データ  $C_{1(s)}$  の総数  $S$  に達したか否かを判定し、総数  $S$  に達していない場合、ステップ S 161 において、カウンタ  $s$  を 1 インクリメントして、処理は、ステップ S 158 に戻る。

## 【0138】

50

そして、全てのクラス $\times$ に属する第1階層要約データC1(s)のベクトルとの比較がなされると、ステップS160において、カウンタ $s$ が総数 $S$ に達したものとみなされ、処理は、ステップS162に進む。

【0139】

ステップS162において、クラス $\times$ に属する第1階層要約データC1(s)のベクトルのうち、検索しようとする動画像コンテンツのサンプルデータより抽出された第1階層要約データC1tのベクトルと一致した第1階層要約データC1の動画像コンテンツに対応するコンテンツデータペア102を読み出す。

【0140】

ステップS163において、第2階層要約データ検索抽出部154は、画像分割部181、および、RGB画素値平均値計算部182を、第2階層要約データ生成部44が画像分割部71、およびRGB画素値平均値計算部72を制御するのと同様に制御して、第2階層要約データ生成処理を実行して、検索しようとする動画像コンテンツのサンプルデータの第2階層要約データファイル112を生成する。尚、第2階層要約データ生成処理については、図6のフローチャートを参照して説明した処理と同様であるので、その説明は省略する。

10

【0141】

ステップS164において、第2階層要約データ検索抽出部154は、スライドマッチング検索部183を制御して、検索しようとする動画像コンテンツのサンプルデータの第2階層要約データファイル112と、ステップS162の処理により読み出されたコンテンツデータペア102に含まれる第2階層要約データファイル112とをスライドさせながら、フレーム間の第2階層要約データ間の差分絶対値和からなる類似度を求める。

20

【0142】

そして、ステップS165において、第2階層要約データ検索抽出部154は、類似度に基づいて、検索しようとする動画像コンテンツと、動画像コンテンツにおけるサンプルデータの再生位置を特定する。より詳細には、第2階層要約データ検索抽出部154は、ずらしたフレーム数に対応して得られる類似度のうち、最小となるものが、所定の閾値よりも小さい場合、サンプルデータの動画像コンテンツと、ステップS162において読み出されたコンテンツデータペア102の動画像コンテンツとが一致するものとみなし、検索しようとする動画像コンテンツとして特定する。このとき、類似度が最小となるスライド位置が、検索しようとする動画像コンテンツにおける、サンプルデータの再生位置であることも特定されることになる。

30

【0143】

すなわち、スライドマッチング検索部183は、図11の最上段で示されるサンプルデータの第2階層要約データファイル112Aと、2段目に示されるステップS162の処理で読み出された第2階層要約データファイル112Bとの対応する第2階層要約データ間の差分絶対値和を類似度として求める。尚、図11においては、サンプルデータの第2階層要約データファイル112Aは、フレームF201乃至F203より構成され、それぞれの第2階層要約データは、1次元のベクトルとして表現されており、フレーム順に10, 11, 7と表記されている。また、第2階層要約データファイル112Bは、フレームF111乃至F115より構成され、それぞれ第2階層要約データは、1次元のベクトルとして表現されており、フレーム順に8, 10, 11, 7, 6と表記されている。

40

【0144】

最初の処理では、最上段の第2階層要約データファイル112Aと中段の第2階層要約データファイル112Bとが対応するフレーム順にその差分絶対値和が類似度として求められる。この場合、類似度は、 $7 (= |10 - 8| + |11 - 10| + |7 - 11|)$ として求められる。次の処理では、第2階層要約データファイル112Aが、図11の下段で示されるように、図中右方向に1フレーム分スライドされることにより、類似度は、 $0 (= |10 - 10| + |11 - 11| + |7 - 7|)$ となる。すなわち、この場合、第2階層要約データファイル112Bの動画像コンテンツが、第2階層要約データファイル1

50

12Aが求められた動画像より検索しようとしていた動画像コンテンツであることが特定される。また、この際、検索しようとした動画像コンテンツと一致するタイミングが、第2階層要約データファイル112BにおけるフレームF112乃至F113であることが特定されるため、一致する動画像の再生位置を特定することができる。尚、図11の例においては、類似度が0となる場合に第2階層要約データファイルが一致したものとしているが、所定の閾値よりも小さい値となれば、一致するものと特定するようにしてもよい。

#### 【0145】

例えは、図12で示されるように、異なる動画像コンテンツによるフレーム毎の第2階層要約データの値の変化が類似している、または一致しているもの同士は同一の動画像コンテンツであるものとみなすことができる。

10

#### 【0146】

尚、図12においては、s10t0乃至9は動画像コンテンツを識別する識別子であり、横軸がフレーム番号を、縦軸が各フレームにおける1次元のベクトル要素からなる第2階層要約データの変化を示している。また、s10t0乃至9は、スライドマッチングにより類似度が最小（最も類似している状態）となるようにフレームのスライド位置の調整が済んだ波形である。

#### 【0147】

すなわち、図12においては、例えは、s10t1で示される波形が、サンプルデータの動画像コンテンツより求められた第2階層要約データファイルのものであるとすれば、s10t2, 3の動画像コンテンツの第2階層要約データファイルの波形と、十分に類似しているということが言える。この結果、s10t1の動画像コンテンツは、s10t2, 3の動画像コンテンツと一致しているとみなすことができ、検索しようとしていた動画像コンテンツが、s10t2, 3の波形で示されるものであったということができる。

20

#### 【0148】

ステップS166において、検索結果表示画像生成部156は、類似度が所定の値より低い第2階層要約データファイル112と同一のコンテンツデータペア102の動画像コンテンツファイル111を読み出し、再生位置と共に検索結果として表示部27に表示する。

#### 【0149】

以上の処理により、サンプルデータの動画像コンテンツを入力するだけで、容易で、かつ、迅速に目的とする記憶部24に登録された動画像コンテンツを検索することが可能となる。

30

#### 【0150】

また、第1階層要約データ、および第2階層要約データといった階層化された要約データを用いることにより、上位階層の時系列の情報を持たない特徴量である第1階層要約データによる絞込みを行った後に、下位階層となる時系列の情報を持った第2階層要約データファイルによる検索を行うようにしたので、検索負荷を低減し、検索速度の高速化を実現することが可能となる。

#### 【0151】

さらに、第1階層要約データは、データベースとして管理することにより、検索を容易で高速なものとすることが可能となる。また、第1階層要約データは、クラスタリング処理されてクラス単位で管理されているため、クラス毎の重心ベクトルを利用することにより、第1階層要約データを絞り込むことができ、絞り込んだ第1階層要約データのみを詳細に比較するだけでよくなるので、第1階層要約データのみの検索処理を高速に実現することが可能となる。

40

#### 【0152】

##### [圧縮処理]

以上においてはサンプルデータである1部の動画像コンテンツを入力するだけで、容易で、かつ、迅速に目的とする動画像コンテンツを検索される例について説明してきたが、同一のシーンを含んだ、個別に異なる動画像コンテンツが大量に検索されてしまう場合が

50

ある。この場合、重複するシーンを含む動画像コンテンツが大量に記憶部24に蓄積されていることとなる。そこで、重複しているシーンについては削除した上で、必要なシーンだけ繋ぎ合わせて必要最小限の動画像コンテンツファイルとすることで、動画像コンテンツファイルを圧縮し、記憶部24の記憶容量を節約するようにしてよい。

【0153】

ここで、図13を参照して、上述した圧縮処理について説明する。

【0154】

ステップS181において、検索抽出部25は、検索抽出処理を実行して、サンプルデータとなる1部の動画像コンテンツに基づいて、同一、または類似する動画像コンテンツを検索して、再生位置と共に特定する。尚、検索抽出処理については、図10のフローチャートを参照して、説明した処理と同様であるので、その説明は省略するものとする。 10

【0155】

ステップS182において、コンテンツファイル圧縮部26は、同一部分検索部211を制御して、検索された複数の動画像コンテンツファイルより、同一部分を検索させる。すなわち、図11を参照して説明したようにフレーム単位でスライドマッチングにより第2階層要約データファイル112が比較されることにより、サンプルデータの動画像コンテンツのどのフレームと、同一、または類似していると見なされた動画像コンテンツのどのフレームとが一致するかについては検索されることになる。

【0156】

この結果、複数の動画像コンテンツが検索された場合、サンプルの動画像コンテンツの一致部分を揃える事で、検索結果となる複数の動画像コンテンツ間のフレーム間の対応は認識することができる。そこで、同一部分検索部211は、動画像コンテンツファイルの一致するフレームを揃えて比較し、同一部分を検索する。例えば、図14の上段で示される動画像コンテンツファイルCont1の第2階層要約データファイル112と、図14の2段目で示される動画像コンテンツファイルCont2の第2階層要約データファイル112とは、いずれもフレームF1の位置で揃えられている。尚、図14においては、帯状で示される範囲のうち、動画像コンテンツCont1の第2階層要約データファイル112が、右上がりの斜線部で示されたフレームFs1乃至Fe1の範囲であり、動画像コンテンツCont2の第2階層要約データファイル112が、左上がりの斜線部で示されたフレームFs2乃至Fe2の範囲である。したがって、同一部分検索部211は、動画コンテンツCont1(または2)におけるフレームFs1乃至フレームFe2の範囲が同一部分であることを検索する。 20 30

【0157】

ステップS183において、コンテンツファイル圧縮部26は、不一致部分検索部212を制御して、検索された複数の動画像コンテンツファイルの第2階層要約データファイル112より、不一致部分を検索させる。すなわち、図14の場合、動画像コンテンツCont2におけるフレームFs2乃至フレームFs1の範囲、および動画像コンテンツCont1におけるフレームFe2乃至コンテンツフレームFe1の範囲が不一致部分となる。

【0158】

ステップS184において、コンテンツファイル圧縮部26は、編集部213を制御して、同一部分、および不一致部分を編集させて繋ぎ合せる。すなわち、編集部213は、第2階層要約データファイル112に対応する動画コンテンツCont1(または2)の動画像コンテンツファイル111におけるフレームFs1乃至フレームFe2、並びに、動画像コンテンツCont2の動画像コンテンツファイル111におけるフレームFs2乃至フレームFs1の範囲、および動画像コンテンツCont1の動画像コンテンツファイル111におけるフレームFe2乃至コンテンツフレームFe1の範囲を繋ぎ合せる。この場合、動画像コンテンツCont2を基準にすると、図14の3段目で示されるように、動画像コンテンツCont1におけるフレームFe2乃至コンテンツフレームFe1の範囲を編集により繋ぎ合せるのみでよいので、格子状の範囲で示されるように繋ぎ合わ 40 50

せて、図14の3段目で示されるような新たな動画像コンテンツを生成する。尚、図14で示される各帶は、第2階層要約データファイル112の構成であり、編集部213は、第2階層要約データファイル112の構成に基づいて、動画像コンテンツファイル111を編集する。

【0159】

ステップS185において、更新部214は、新たに生成した動画像コンテンツに対応するように、第1階層要約データDB101、並びに、動画像コンテンツファイル111、および第2階層要約データファイル112からなるコンテンツデータペア102を更新する。この際、更新部214は、不要となった動画像コンテンツContent1,2のコンテンツデータペア102は削除する。

10

【0160】

以上の処理により動画像コンテンツファイルを実質的に圧縮することが可能となり、動画像コンテンツにかかる記憶部24の記憶容量を節約することが可能となる。また、このような圧縮処理により、検索対象となる動画像コンテンツそのものが減少することになるので、管理にかかる負担を低減させると共に、サンプルデータによる動画像コンテンツの検索の負荷を低減し、検索速度を向上させることができるとなる。

【0161】

尚、以上においては、第1階層要約データおよび第2階層要約データとして、画像を複数領域に分割し、分割領域毎のRGBの画素値の平均値を用いる例について説明してきたが、フレーム単位の特徴量を生成できればよいのでこれ限るものではない。例えば、分割領域毎のRGBの画素値の平均値に代えて、分割領域毎の輝度値の平均値、分割領域毎のアクティビティの値の平均値とするようにしてもよい。また、分割領域毎の輝度ヒストグラムを、データ化したものなどでもよく、例えば、輝度方向の分解能を3乃至5ビット程度としたものとし、例えば、4ビットの場合、16次元のベクトルとするようにしてもよい。また、輝度値ではなくRGBの画素値を用いたものでもよいし、RGBで表現される色立体を、各軸を、例えば、2乃至4ビットの分解能で分割し、それぞれに対して頻度を付したものとしてもよく、3ビットとして、 $8^3 = 512$ 次元のベクトルとするようにしてもよい。さらに、画像のみならず、音声の情報に基づいてフレーム単位で第1階層要約データおよび第2階層要約データを生成するようにしてもよく、例えば、音声ボリューム、または所定の音声周波数帯域毎の振幅の平均値や、これらを画像の情報と組み合わせたものとしても良い。

20

【0162】

また、画像に対して、色クラスタリング（RGBの3次元で表現される色の分布に対し、ベクトル量子化を試みる一般的な画像処理手法）を行い、最も頻繁に用いられている色を代表色として、そのRGBの値を3次元ベクトルで表現したものとしてもよい。

30

【0163】

さらに、画像を分割する際、画面下部、上部などにテロップが頻繁に表示される画像の場合、テロップが頻繁に登場する領域は、より大雑把に分割するといった工夫をすることによりテロップによる影響を緩和するようにしても良い。

【0164】

40

また、以上においては、第1階層要約データについては、30フレーム毎、またはシーンチェンジのタイミングにおけるフレームにおいて、求められるものとしてきたが、その他の周期的に求められるものとしたり、または、動画像コンテンツに変化がおきる非周期的なものであってもよい。従って、周期的なものとしては、例えば、他のフレーム数間隔おきのフレームから第1階層要約データを求めるようにしてもよいし、非周期的なものとしては、音声の無音部分が終了するタイミングのフレームから第1階層要約データを求めるようにしてもよい。さらに、第2階層要約データについては、毎フレーム求める例について説明してきたが、第1階層要約データよりも頻度が低いフレーム数で求められれば良いものであるので、第1階層要約データよりも頻度が低いという条件を満たす限り、第1階層要約データのように、数フレーム間隔といった周期的なものでもよいし、シーン

50

チェンジなどの所定の条件を満たすような非周期的なものであってもよい。

【0165】

また、以上においては、第1階層要約データ、および第2階層要約データの抽出に当たっては、事前にアクティビティ画像を求めてアクティビティの高い位置を中心位置とするようにして、画像を正規化するようにする例について説明してきたが、例えば、画像の自己相関関数が所定量（例えば0.9倍）だけ変化する距離で正規化を行うことにより、スケーリングの自由度を許容して同一画像を検出することができるようにしてよい。

【0166】

さらに、ベクトルの比較により一致するか否かを判定する処理において、距離が所定の値以下におさまることを条件にするが、この所定の値を用途に応じて変化させてもよい。例えば、完全に一致する動画像コンテンツを検索したい場合、所定値を小さくするようにしてもよい。また、同じ画像でもテロップやカラーコレクションの有無などを許容し一致判定を行いたい場合には、所定値を大きくするようにしてもよい。

10

【0167】

また、以上の検出抽出処理においては、第1階層要約データからなるベクトルを比較する際、ベクトルが一致するか否かに基づいて判定する例について説明してきたが、クラスタリング処理や初期クラスタリング処理において用いたように、ベクトル間の距離を求めて、所定の距離より小さく十分に相互が類似しているか否かに基づいて判定するようにしてもよい。逆に、クラスタリング処理や初期クラスタリング処理におけるベクトル間の距離を求めて所定の距離よりも小さく十分に類似しているか否かを判定する処理に代えて、検出抽出処理で用いたベクトルが一致するか否かに基づいて判定するようにしてもよい。

20

【0168】

さらに、以上においては、第1階層要約データ、および第2階層要約データの2階層構造の要約データを用いる例について説明してきたが、さらに、多階層の要約データを構成するようにしても良く、さらに、各階層にクラスタリング処理を施し、クラス分類するようにしてもよい。

【0169】

以上の如く、本発明によれば、膨大な量の動画像コンテンツを容易に管理することが可能になると共に、管理している膨大な量の動画像コンテンツの中から検索したい動画像コンテンツを容易に検索することが可能となる。

30

【0170】

ところで、上述した一連の処理は、ハードウェアにより実行させることもできるが、ソフトウェアにより実行させることもできる。一連の処理をソフトウェアにより実行させる場合には、そのソフトウェアを構成するプログラムが、専用のハードウェアに組み込まれているコンピュータ、または、各種のプログラムをインストールすることで、各種の機能を実行することができる、例えば汎用のパーソナルコンピュータなどに、記録媒体からインストールされる。

【0171】

図15は、汎用のパーソナルコンピュータの構成例を示している。このパーソナルコンピュータは、CPU(Central Processing Unit)1001を内蔵している。CPU1001にはバス1004を介して、入出力インタ-フェイス1005が接続されている。バス1004には、ROM(Read Only Memory)1002およびRAM(Random Access Memory)1003が接続されている。

40

【0172】

入出力インタ-フェイス1005には、ユーザが操作コマンドを入力するキーボード、マウスなどの入力デバイスよりなる入力部1006、処理操作画面や処理結果の画像を表示デバイスに出力する出力部1007、プログラムや各種データを格納するハードディスクドライブなどよりなる記憶部1008、LAN(Local Area Network)アダプタなどによりなり、インターネットに代表されるネットワークを介した通信処理を実行する通信部1009が接続されている。また、磁気ディスク（フレキシブルディスクを含む）、光ディス

50

ク (CD-ROM(Compact Disc-Read Only Memory)、DVD(Digital Versatile Disc)を含む)、光磁気ディスク (MD(Mini Disc)を含む)、もしくは半導体メモリなどのリムーバブルメディア 1011 に対してデータを読み書きするドライブ 1010 が接続されている。

【 0 1 7 3 】

CPU 1001 は、ROM 1002 に記憶されているプログラム、または磁気ディスク、光ディスク、光磁気ディスク、もしくは半導体メモリ等のリムーバブルメディア 1011 から読み出されて記憶部 1008 にインストールされ、記憶部 1008 から RAM 1003 にコードされたプログラムに従って各種の処理を実行する。RAM 1003 にはまた、CPU 1001 が各種の処理を実行する上において必要なデータなども適宜記憶される。

【 0 1 7 4 】

尚、本明細書において、記録媒体に記録されるプログラムを記述するステップは、記載された順序に沿って時系列的に行われる処理は、もちろん、必ずしも時系列的に処理されなくとも、並列的あるいは個別に実行される処理を含むものである。

## 【符号の説明】

【 0 1 7 5 】

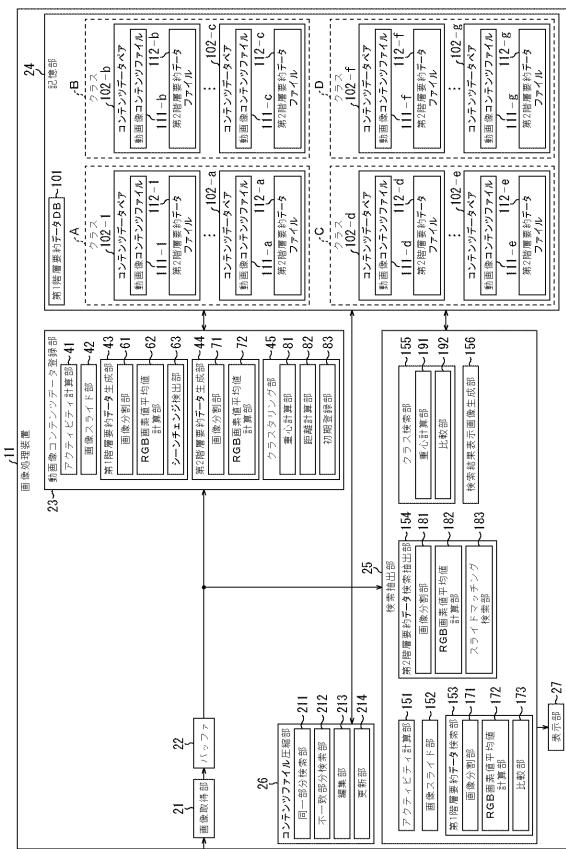
1 1 画像処理装置， 2 1 画像取得部， 2 2 バッファ， 2 3 コンテンツデータ登録部， 2 4 記憶部， 2 5 検索抽出部， 2 6 コンテンツファイル圧縮部， 2 7 表示部， 4 1 アクティビティ計算部， 4 2 画像スライド部， 4 3 第1階層要約データ生成部， 4 4 第2階層要約データ生成部， 4 5 クラスタリング部， 1 5 1 アクティビティ計算部， 1 5 2 画像スライド部， 1 5 3 第1階層要約データ検索部， 1 5 4 第2階層要約データ検索抽出部， 4 5 クラス検索部

10

20

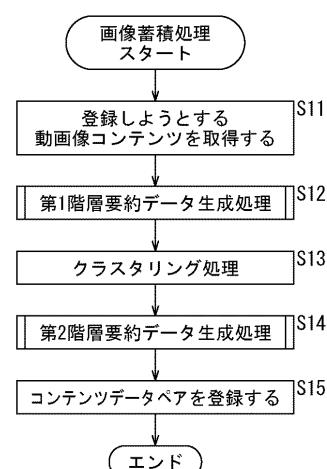
〔 図 1 〕

図 1



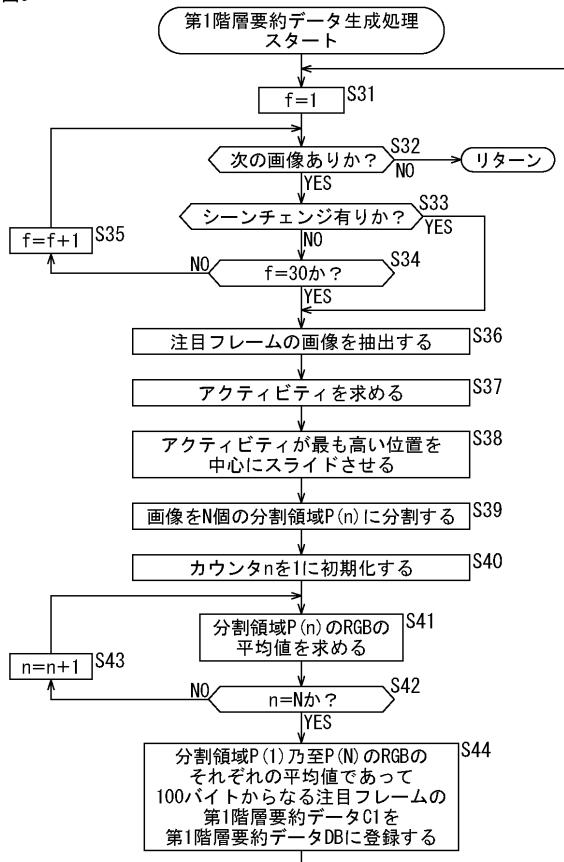
〔 図 2 〕

图2



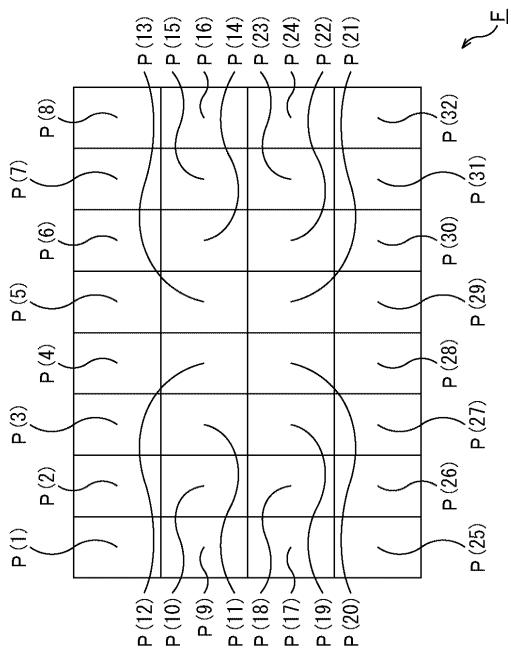
【図3】

図3



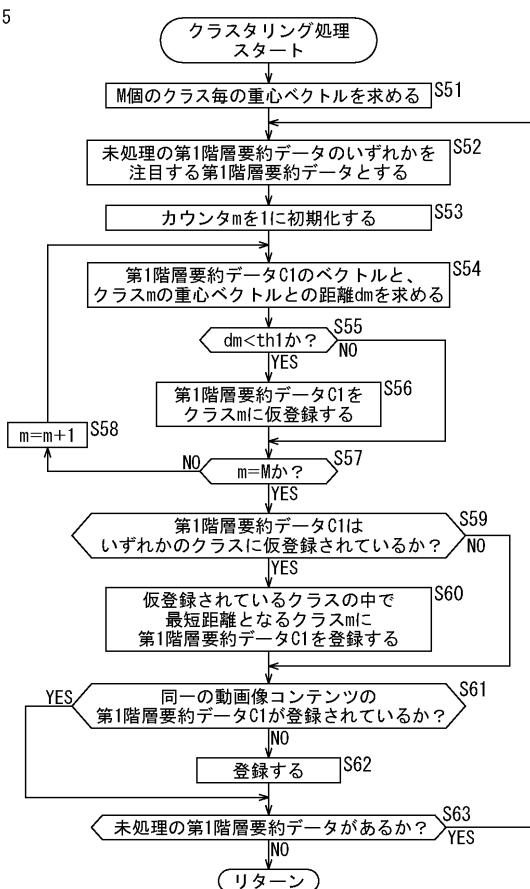
【図4】

図4



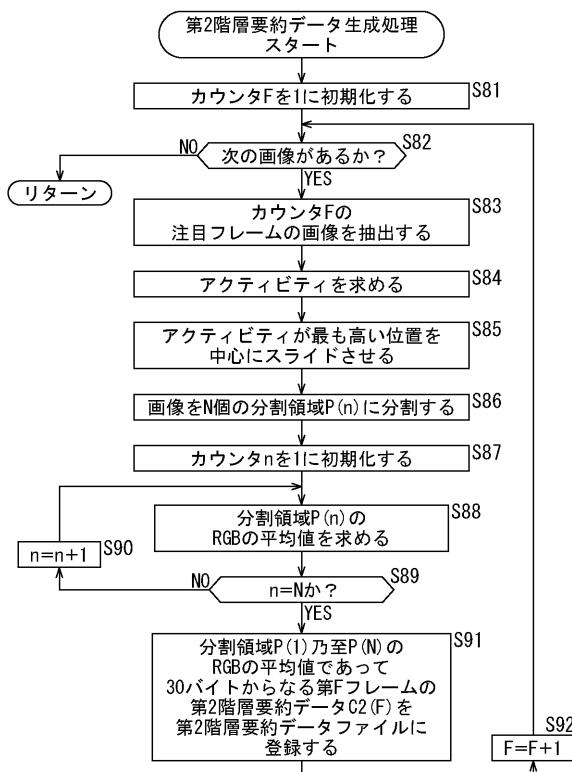
【図5】

図5



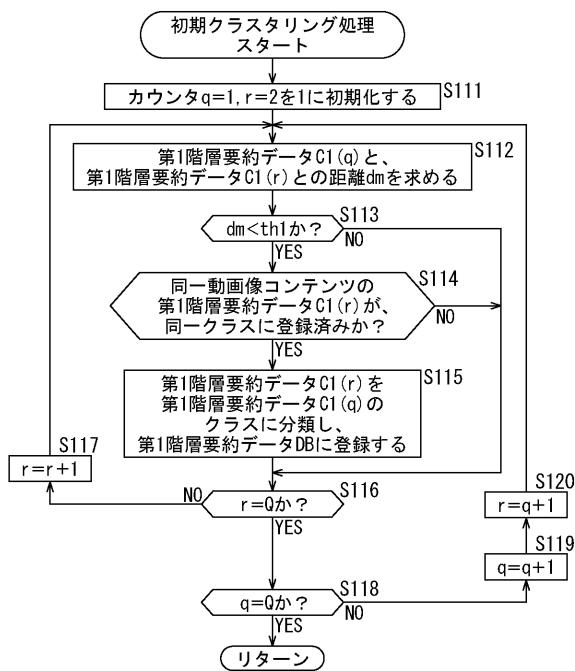
【図6】

図6



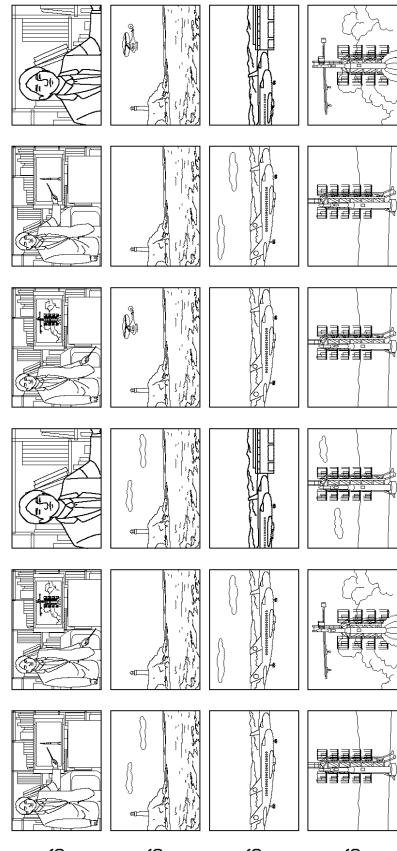
【図7】

図7



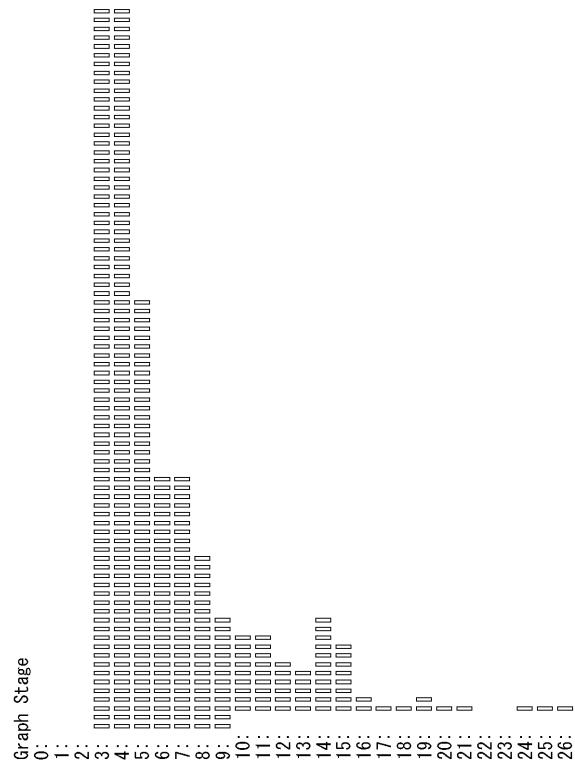
【図8】

図8



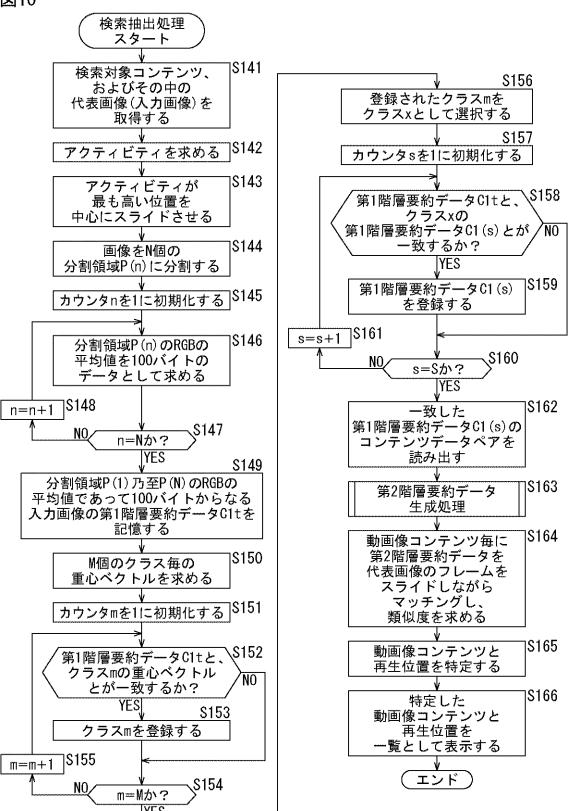
【図9】

図9

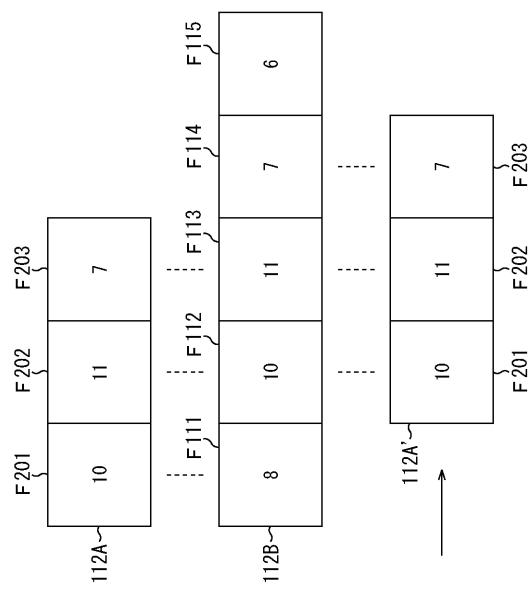


【図10】

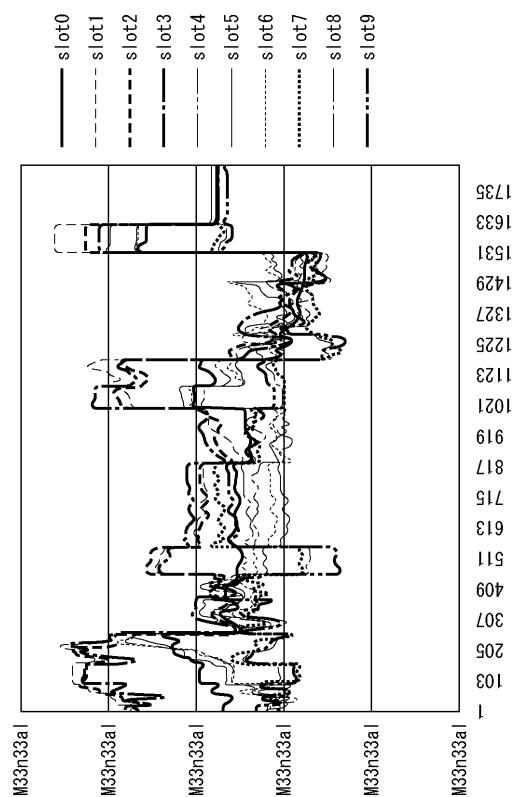
図10



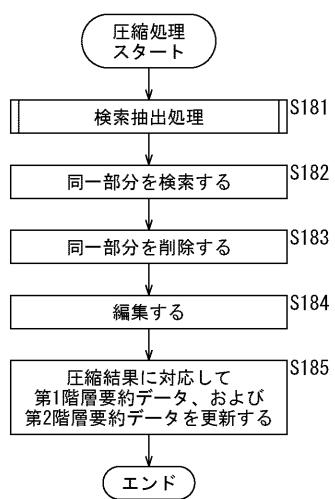
【図 1 1】  
図11



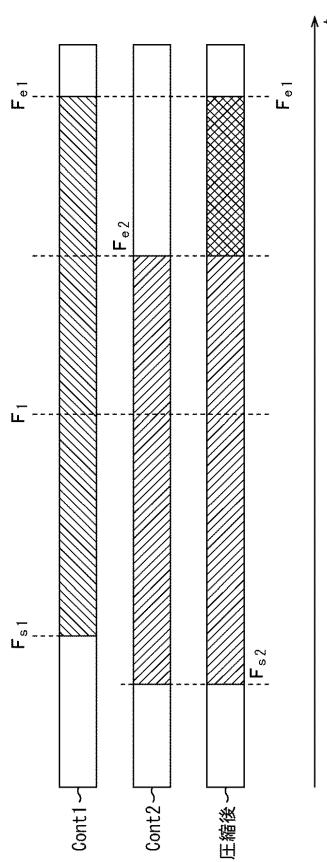
【図 1 2】  
図12



【図 1 3】  
図13

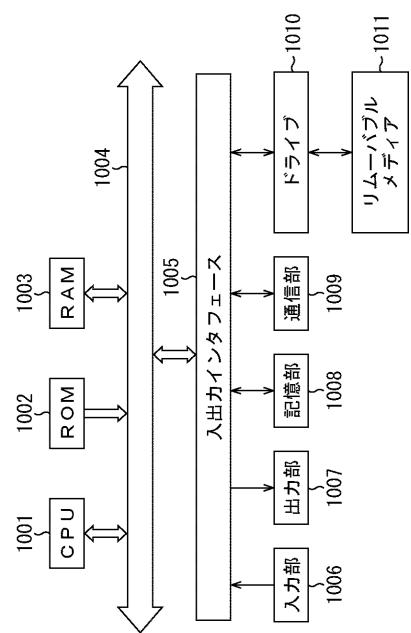


【図 1 4】  
図14



【図15】

図15



---

フロントページの続き

(72)発明者 高橋 義博  
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 田中 和政  
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 菊地 大介  
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 山田 直秀  
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 森崎 裕之  
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 伊知地 和之

(56)参考文献 特開2000-148796 (JP, A)  
特開2004-045565 (JP, A)  
特開2001-216332 (JP, A)  
特開2002-044610 (JP, A)  
特開2010-004434 (JP, A)  
特表2001-527304 (JP, A)  
特開平10-257436 (JP, A)  
特開2006-268866 (JP, A)

(58)調査した分野(Int.Cl., DB名)

G 06 F 17 / 30  
H 04 N 5 / 76  
H 04 N 5 / 91  
C S D B (日本国特許庁)