

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4027662号
(P4027662)

(45) 発行日 平成19年12月26日(2007.12.26)

(24) 登録日 平成19年10月19日(2007.10.19)

(51) Int. Cl.		F I		
G06F	17/30	(2006.01)	G O 6 F	17/30 1 7 O D
G06T	7/20	(2006.01)	G O 6 T	7/20 C
H04N	5/76	(2006.01)	H O 4 N	5/76 B

請求項の数 6 (全 24 頁)

(21) 出願番号	特願2001-508725 (P2001-508725)	(73) 特許権者	000005049
(86) (22) 出願日	平成12年6月29日(2000.6.29)		シャープ株式会社
(86) 国際出願番号	PCT/JP2000/004299		大阪府大阪市阿倍野区長池町22番22号
(87) 国際公開番号	W02001/003005	(74) 代理人	100064746
(87) 国際公開日	平成13年1月11日(2001.1.11)		弁理士 深見 久郎
審査請求日	平成14年7月19日(2002.7.19)	(74) 代理人	100085132
審査番号	不服2004-18022 (P2004-18022/J1)		弁理士 森田 俊雄
審査請求日	平成16年9月1日(2004.9.1)	(74) 代理人	100083703
(31) 優先権主張番号	特願平11-184606		弁理士 仲村 義平
(32) 優先日	平成11年6月30日(1999.6.30)	(74) 代理人	100096781
(33) 優先権主張国	日本国(JP)		弁理士 堀井 豊
(31) 優先権主張番号	特願平11-339019	(74) 代理人	100098316
(32) 優先日	平成11年11月30日(1999.11.30)		弁理士 野田 久登
(33) 優先権主張国	日本国(JP)	(74) 代理人	100109162
			弁理士 酒井 将行

最終頁に続く

(54) 【発明の名称】 動画像検索装置、および動画像検索情報管理装置

(57) 【特許請求の範囲】

【請求項1】

動画像データを構成する1以上のシーンの各々に対応した検索情報を利用して、所望の画像を検索するための動画像検索装置であって、

前記動画像データには前記検索情報が対応付けられており、

前記検索情報は、前記シーンの各々について、当該シーンを空間的に分割することにより得られる、時間方向には当該シーンの長さを持った複数の空間分割ブロックを単位とする要約情報を含み、

前記空間分割ブロック単位の要約情報は、当該空間分割ブロック内の動画像データに基づいて算出される画像特性を示した数値情報であり、かつ当該シーン内での相対値で与えられ、

前記検索情報を読み出して管理するための情報管理手段と、

前記情報管理手段に接続され、外部から与えられる第1の検索要求に応答して、前記情報管理手段から前記検索情報に含まれる前記空間分割ブロック単位の要約情報を受取り、前記空間分割ブロック単位の要約情報を用いて前記第1の検索要求に合致するシーンを検索するための第1のシーン検索手段を含む、動画像検索装置。

【請求項2】

前記検索情報はさらに、前記シーンの各々について、当該シーンを時間的に分割することにより得られる、空間方向には当該シーンの画像サイズを持った複数の時間分割ブロックを単位とする要約情報を含み、

10

20

前記時間分割ブロック単位の要約情報は、当該時間分割ブロック内の動画像データに基づいて算出される画像特性を示した数値情報であり、

前記情報管理手段に接続され、外部から与えられる第2の検索要求に応答して、前記情報管理手段から前記検索情報に含まれる前記時間分割ブロック単位の要約情報を受取り、前記時間分割ブロック単位の要約情報を用いて前記第2の検索要求に合致するシーンを検索するための第2のシーン検索手段を含む、請求項1に記載の動画像検索装置。

【請求項3】

前記検索情報はさらに、前記シーンの各々を単位とする要約情報を含み、

前記シーン単位の要約情報は、当該シーン内の動画像データに基づいて算出される画像特性を示した数値情報であり、

前記情報管理手段に接続され、外部から与えられる第3の検索要求に応答して、前記情報管理手段から前記検索情報に含まれる前記シーン単位の要約情報を受取り、前記シーン単位の要約情報を用いて前記第3の検索要求に合致するシーンを検索するための第3のシーン検索手段をさらに含む、請求項1または2に記載の動画像検索装置。

【請求項4】

動画像データを構成する1以上のシーンの各々に対応した検索情報を管理するための動画像検索情報管理装置であって、

前記検索情報は、前記シーンの各々について、当該シーンを空間的に分割することにより得られる、時間方向には当該シーンの長さを持った複数の空間分割ブロックを単位とする要約情報を含み、

前記空間分割ブロック単位の要約情報は、当該空間分割ブロック内の動画像データに基づいて算出される画像特性を示した数値情報であり、かつ当該シーン内での相対値で与えられ、

前記検索情報を読み込み前記検索情報に含まれる個々の要約情報を保持する保持手段と、外部から要求を受け、前記外部からの要求が前記空間分割ブロック単位の要約情報を要求する第1の要求であるか否かを判断する判断手段と、前記外部からの要求が前記第1の要求であると判断された場合に、前記検索情報に含まれる前記空間分割ブロック単位の要約情報を選択して出力する出力手段とを備えることを特徴とする、動画像検索情報管理装置。

【請求項5】

前記検索情報はさらに、前記シーンの各々について、当該シーンを時間的に分割することにより得られる、空間方向には当該シーンの画像サイズを持った複数の時間分割ブロックを単位とする要約情報を含み、

前記時間分割ブロック単位の要約情報は、当該時間分割ブロック内の動画像データに基づいて算出される画像特性を示した数値情報であり、

前記判断手段はさらに、前記外部からの要求が前記時間分割ブロック単位の要約情報を要求する第2の要求であるか否かを判断し、前記出力手段はさらに、前記外部からの要求が前記第2の要求であると判断された場合に、前記検索情報に含まれる前記時間分割ブロック単位の要約情報を選択して出力することを特徴とする、請求項4に記載の動画像検索情報管理装置。

【請求項6】

前記検索情報はさらに、前記シーンの各々を単位とする要約情報を含み、

前記シーン単位の要約情報は、当該シーン内の動画像データに基づいて算出される画像特性を示した数値情報であり、

前記判断手段はさらに、前記外部からの要求が前記シーン単位の要約情報を要求する第3の要求であるか否かを判断し、前記出力手段はさらに、前記外部からの要求が前記第3の要求であると判断された場合に、前記検索情報に含まれる前記シーン単位の要約情報を選択して出力することを特徴とする、請求項4または5に記載の動画像検索情報管理装置。

【発明の詳細な説明】

【 0 0 0 1 】

【 発明の属する技術分野 】

本発明は、動画像を検索するための情報を管理する装置（動画像検索情報管理装置）、および動画像を検索するための装置（動画像検索装置）に関し、より詳細には、動画像を構成する複数のシーンの中から、所望のシーンを効率的に検索して取出すことを可能とする、動画像検索装置、および動画像検索情報管理装置に関する。

【 0 0 0 2 】

【 従来の技術 】

コンピュータの性能の向上、記憶装置の大容量化、通信のためのインフラストラクチャの整備などにより、大量の画像をデータベース化しておき必要なときに必要な画像を配信するサービスなどが普及しつつある。画像をデジタル化する技術も普及しているため、個人で動画像のデータベースを構築することも可能である。

【 0 0 0 3 】

このように動画像をデータベース化すると、その中からいかにして所望の動画像を検索するかが問題となる。本発明は、そのような動画像の検索の技術の改良に関する。

【 0 0 0 4 】

本明細書において使用する「動画像」という用語は、RGB（Red, Green, Blue）信号からなる、加工されていない動画像データ、およびMPEG（Moving Picture Experts Group）などのように、デジタル化され、所定の動画圧縮方式によって符号化された動画像データの両者を含むものとする。

【 0 0 0 5 】

「シーン」という用語は、連続する複数の画像フレームからなる、動画像の構成単位であって、例えば録画の開始から停止までの1回の撮影で得られた動画像、または編集済みの動画像の中で、二つの編集点で両端を区切られた一連の動画像のことをいう。編集箇所がフェード、ディゾルブなどの特殊な効果を含んでいる場合も含む。一つのシーンは複数のシーンを含んでもよい。

【 0 0 0 6 】

「シーンチェンジ点」とは、二つのシーンが切替わる時間軸上の位置を指し、シーンが切替わった直後のシーンの先頭フレームを「シーンチェンジフレーム」という。

【 0 0 0 7 】

「キーフレーム」とは、検索などで利用される、各シーンまたは動画像全体を代表するような特徴的なフレームをいう。

【 0 0 0 8 】

動画像データは、時間軸に沿って変化する画像情報からなるコンテンツである。動画像データは、その全てが検索の対象となる。したがって、所望の動画像を検索するための最も単純な方法は、動画像の全体を見て所望の動画像を選ぶことである。しかし、動画像の全体を見るためにはある時間を要する。したがって、データベースに登録されている動画像の数が増えると、所望の動画像を取出すために要する時間が極めて大きくなる。したがってこの方法は現実的ではない。

【 0 0 0 9 】

効率的に動画像を検索するための一般的な方法では、予め各動画像の内容を代表する種々の検索のための情報（検索情報）を各動画像について用意し、これに基づいて所望の動画像を検索する。

【 0 0 1 0 】

検索情報として用いられる情報は、画像に付加された情報（付加情報）、画像の特性を表わす情報（特性情報）、および画像の概要を表わす情報（概要情報）を含む。付加情報は、動画像に附随して用意されるタイトルまたは説明文を含む。特性情報は、画像信号から導出される輝度値の頻度もしくはフレーム間での動き情報など、画像の特性を表わす情報を含む。概要情報は、画像の特性情報または外部からの入力によって定められるシーンチェンジ点またはキーフレームなどを含む。

【 0 0 1 1 】

上記した情報の中で、特性情報は特に便利でかつ重要である。たとえば動画像を代表するキーフレームは特性情報を利用して選択できる。特性情報は数値化することが容易で、所望の動画像の特性を特性情報の形で数値化すれば、要求された特性情報と一致する特性を有する動画像を抽出することも容易にできる。以下の説明では「検索情報」とはこの特性情報のことをいうものとする。

【 0 0 1 2 】

一般に動画像は複数個のシーンを含む。検索者が動画像を検索するのは、多くの場合、検索の対象となる1または複数の動画像から、動画像中に含まれている所望のフレームまたは所望のシーンを発見するためである。検索情報を用いて動画像の中から所望のフレームまたはシーンを発見するための基本的な方法は、予め各フレームに関する検索情報または各シーンに関する検索情報を求めて記録しておき、これらに基づいてフレームまたはシーンの選択を行なうことである。

10

【 0 0 1 3 】

この点に関し特許文献1には、シーンチェンジ点を検出するためのフレームの特徴量として、フレームごとの動きベクトルの総和、ならびに、フレームを構成する複数個の小領域の、フレーム内およびフレーム間の連続度を用いる装置および方法が開示されている。

【 0 0 1 4 】

また特許文献2には、各シーン内の動きベクトルの大きさの平均値および画像の色ヒストグラムをシーンの特徴量として用いて、各シーンの代表画像を画像インデックスとして表示する際の制御方法が開示されている。

20

【 0 0 1 5 】

こうした従来技術を考慮して、動画像を検索するための情報を記録するための装置（動画像検索情報記録装置）および動画像を検索するための装置（動画像検索装置）を作成したとすれば、以下に示すようなものとなると考えられる。

【 0 0 1 6 】

図1を参照して、想定された従来の動画像検索情報記録装置は、入力される動画像データを解析してシーンに分割し、シーンの分割位置（シーンチェンジ）などの動画像の構造を表わす情報（動画像構造情報）を出力するための解析部601と、解析部601の出力と入力された動画像データとを受けるとに接続され、解析部601から出力された動画像構造情報および動画像データに基づいて、シーン内の各フレームに対応する動きベクトルの総和などの、画像の要約情報を検索情報として生成し、出力するフレーム単位の検索情報生成部602と、フレーム単位の検索情報生成部602の出力を受けるとに接続され、フレーム単位の検索情報生成部602から出力されるフレーム単位の要約情報に基づいて、あるシーンの全体に対する動きベクトルの大きさの平均値などの要約情報を検索情報として生成し出力するためのシーン単位の検索情報生成部603と、解析部601から出力される動画像構造情報、フレーム単位の検索情報生成部602から出力される各フレーム単位の検索情報、およびシーン単位の検索情報生成部603から出力されるシーン全体に対する検索情報を、所定の形式に配置することにより、動画像に対応する検索情報を生成し出力するための動画像の検索情報生成部604と、もとの動画像データと、動画像の検索情報生成部604から出力される、動画像に対応する検索情報とを対応付けて記録媒体606に記憶するための記録部605とを含む。

30

40

【 0 0 1 7 】

なお、1フレームは、動画像を時間的に分割した単位であると考えられることができる。したがってフレームは時間分割ブロックの一例と呼ぶことができる。

【 0 0 1 8 】

図2を参照して、想定される従来の画像検索装置は、記録媒体606から動画像検索情報および動画像データを読み出すための読出部701と、読出部701から供給される各動画像検索情報を保持し、検索情報に対する要求を受けて、シーン単位の検索情報またはフレーム単位の検索情報を動画像の構造情報とともに出力するための検索情報管理部702

50

と、シーン単位の検索の要求を受けて、検索情報管理部 702 にシーン単位の検索情報と動画像構造情報とを要求し、受取ったシーン単位の検索情報と動画像構造情報とに基づいて、検索要求に合致したシーンを検出して、検出されたシーンに関する情報を出力するためのシーン単位の検索実行部 703 と、フレーム単位の検索の要求を受けて、検索情報管理部 702 に対してシーン内のフレーム単位の検索情報および動画像構造情報を要求し、受取ったシーン内のフレーム単位の検索情報および動画像情報に基づいて、検索要求に合致したシーンまたはフレームを検出して、検出されたシーンまたはフレームに関する情報を出力するためのフレーム単位の検索実行部 704 と、シーン単位の検索実行部 703 およびフレーム単位の検索実行部 704 から出力される、検出されたシーンに関する情報または検出されたフレームに関する情報に基づいて、それら検出されたシーンまたはフレームに対応する画像データを読み出す部 701 を介して取出し、出力するためのデータ管理部 705 と、オペレータから与えられる検索指示に応答し、検索要求をシーン単位の検索実行部 703 またはフレーム単位の検索実行部 704 に与え、データ管理部 705 から与えられる動画像データを検索結果として表示する処理を繰返し、検索終了時には検索結果を出力するための検索制御部 706 とを含む。

10

【0019】

図 1 に示される装置は以下のように動作することが想定される。動画像データが入力されると、解析部 601 がその動画像データをシーン単位に分割する。解析部 601 はまた、動画像構造情報を出力する。

【0020】

20

フレーム単位の検索情報生成部 602 は、解析部 601 から出力された動画像構造情報および動画像データに基づいて、シーン内のフレームに対応する要約情報を生成する。フレーム単位の検索情報生成部 602 はさらに、こうして生成された要約情報をシーン単位の検索情報生成部 603 および動画像の検索情報生成部 604 に検索情報として出力する。

【0021】

シーン単位の検索情報生成部 603 は、フレーム単位の検索情報生成部 602 から与えられたフレーム単位の要約情報に基づき、シーン全体に対する要約情報を生成する。シーン単位の検索情報生成部 603 はさらに、こうして生成された要約情報をシーン単位の検索情報生成部 603 に検索情報として与える。

30

【0022】

動画像の検索情報生成部 604 は、解析部 601 から動画像構造情報を、フレーム単位の検索情報生成部 602 からフレーム単位の検索情報を、シーン単位の検索情報生成部 603 からシーン全体のための検索情報をそれぞれ受け、所定の形式で配置して動画像に対応する検索情報を生成する。動画像の検索情報生成部 604 は、こうして生成された、動画像に対応する検索情報を記録部 605 に与える。

【0023】

記録部 605 は、もとの動画像データと、動画像の検索情報生成部 604 から与えられた検索情報とを記録媒体 606 に記録する。このとき記録部 605 は、記録される動画像データと検索情報とが対応していることを示す情報を動画像データまたは検索情報のいずれか一方または双方に付加するか、または動画像データおよび検索情報とは別に記録する。

40

【0024】

図 2 を参照して、従来の検索装置は以下のように動作することが想定される。検索の開始時に、オペレータは検索制御部 706 に対して検索指示を与える。検索制御部 706 は、この検索指示に応答してシーン単位の検索、またはフレーム単位の検索のいずれかを選択する。

【0025】

たとえばシーン単位の検索が選択されたものとする。すると、検索制御部 706 はシーン単位の検索実行部 703 に対して検索の要求を与える。シーン単位の検索実行部 703

50

はこの要求に応答して、検索情報管理部 702 に対してシーン単位の検索情報と動画像構造情報とを要求し、検索情報管理部 702 から出力されたシーン単位の検索情報と動画像情報とを受取る。

【0026】

シーン単位の検索実行部 703 は、このシーン単位の検索情報と動画像構造情報とに基づいて、検索制御部 706 から与えられた検索要求に合致したシーンを検出する。そしてシーン単位の検索実行部 703 は、検出されたシーンに関する情報をデータ管理部 705 に対して出力する。

【0027】

データ管理部 705 は、シーン単位の検索実行部 703 から与えられたシーンに関する情報に基づいて、対応する画像データを読出部 701 を介して記録媒体 606 から読出し、検索制御部 706 に与える。

10

【0028】

検索制御部 706 は、この画像データを表示する。オペレータは、表示された画像データを見て、次の検索方針を決め、次の検索指示を検索制御部 706 に与える。以下、同様の処理が繰返される。

【0029】

こうして、オペレータが取出したいと考えていた画像が取出されると検索は終了する。検索結果は検索制御部 706 から出力される。

【0030】

20

従来の技術では、シーンに関する検索情報は、そのシーンに含まれるフレーム単位の検索情報に基づいて生成されている。これは、シーンを構成する最小の単位はフレームであるとな一般的に認識されていることによる。

【0031】

【特許文献 1】

特開平 9 - 284702 号公報

【0032】

【特許文献 2】

特開平 7 - 38842 号公報

【0033】

30

【発明が解決しようとする課題】

このように個々のフレーム単位の検索情報に基づいてシーンの検索情報を生成すると、各フレームを小領域に分割して得られる、各フレーム内の空間的な特徴（画面分割ブロック単位の検索情報）を検索情報として利用することができる。しかし、これらの従来の装置では時間軸上である長さを有するシーン全体にわたる、画像の空間的な特徴を扱うことができないという問題がある。この明細書では、このようにシーン全体にわたる、画像の空間的な特徴を「空間分割ブロック単位の検索情報」と呼ぶ。

【0034】

たとえば、画面の中央部に激しく動く物体が映っている、という特徴を有するシーンを取出そうとする場合を考える。従来の技術では、このようなシーンをシーン単位の検索情報から検索することができず、各シーンに含まれる各フレームに付加された検索情報まで降りて検索する必要がある。そのため従来の技術では、そのような検索は効率が極めて低くなる。

40

【0035】

本発明は、従来技術の、上記したような問題点に鑑みてなされたものであり、シーンの特徴を的確に表わす新たな指標を用いて動画像検索を効率的に行なうことを可能とする動画像の検索情報を記録する装置および動画像を検索するための装置を提供することを目的とする。

【0036】

本発明の他の目的は、シーンを時間的なブロックに分割してその特徴を抽出するだけで

50

なく、シーンを、時間軸上では分割せず、空間的にのみ分割してシーンの別の特徴を抽出することにより、動画像の検索を効率的に行なうことを可能とする、動画像検索装置、および動画像検索情報管理装置を提供することである。

【0037】

本発明のさらに他の目的は、シーンの特徴を的確に表わす新たな指標を用いて動画像の検索を効率的に行なうことが可能で、かつ検索のための情報の情報量が少なく済む、動画像検索装置、および動画像検索情報管理装置を提供することである。

【0038】

【課題を解決するための手段】

上記目的を達成するために、本発明にかかる動画像検索情報記録装置は、動画像データと、動画像データにより表わされる動画像を時間軸上で1以上の時間軸上の部分動画像に分割する動画像構造情報とに基づいて、1以上の部分動画像の各々に対応して検索情報を生成する検索情報生成部と、検索情報と動画像構造情報とを、対応する動画像データとともに記録媒体に記録する記録部とを含む。検索情報生成部は、1以上の時間軸上の部分動画像の各々を空間的に分割することにより得られる1以上の空間的部分動画像の要約情報を生成するための第1の要約情報生成部を含む。

【0039】

時間軸上の部分動画像の各々を空間的に分割することにより得られる空間的部分動画像から要約情報が生成される。この要約情報は、本発明で新たに導入されたものであり、時間軸上の部分動画像の中の、時間的な一部から得られるのではない。この要約情報は、空間的には画像の一部のみの、しかし時間軸上では部分動画像の全体にわたる空間的部分動画像から得られる。したがって、この要約情報は、ある部分動画像の時間的な全体を通じた、その部分動画像の空間的な特徴を表わす。この新たな指標を用いて、部分動画像の、時間よりも空間的な特徴に重点をおいた検索を効率的に行なうことが可能となる。

【0040】

この発明の他の局面にかかる動画像検索装置は、動画像を構成する1以上の部分動画像の各々に対応した検索情報を利用して、所望の画像を検索するための動画像検索装置である。この動画像を表わす動画像データには、検索情報が対応付けられており、検索情報は、1以上の時間軸上の部分動画像の各々を空間的に分割することにより得られる1以上の空間的部分動画像の要約情報を含む。この発明にかかる動画像検索装置は、検索情報を読み出して管理するための情報抽出部と、この情報管理部に接続され、外部から与えられる、時間軸上の部分動画像を検索の単位とする第1の検索要求に応答して、検索情報に含まれる空間的部分動画像の要約情報を用いて、第1の検索要求に合致する時間軸上の部分動画像を検索するための第1の部分動画像検索部を含む。

【0041】

時間軸上の部分動画像の各々を空間的に分割することにより得られる空間的部分動画像から得られた要約情報は、本発明で新たに導入されたものであり、空間的には画像の一部のみの、しかし時間軸上では部分動画像の全体にわたる空間的部分動画像から得られる。したがって、この要約情報は、ある部分動画像の時間的な全体を通じた、その部分動画像の空間的な特徴を表わす。したがってこの発明にかかる動画像検索装置によれば、この新たな指標を用いて、部分動画像の、時間よりも空間的な特徴に重点をおいた検索を効率的に行なうことができる。

【0042】

【発明の実施の形態】

図3に、本実施の形態にかかる動画像検索情報記録装置40のブロック図を示す。図3を参照して、この動画像検索情報記録装置40は、動画像データを受け、動画像の構造を解析してシーンに分割し、シーンチェンジ点などの動画像構造情報を出力するための解析部101と、動画像データと、解析部101から出力される動画像構造情報とを受け、各シーンに含まれる各フレームの画像データに対する要約情報（フレームという時間分割ブロックを単位とする要約情報）を生成し出力するための第1の検索情報生成部102と、

10

20

30

40

50

解析部 101 から出力された動画像構造情報、および動画像データを受け、各シーンに関する検索情報として、シーン内の全てのフレームの同一分割ブロック位置の画像データから構成される空間分割ブロックの要約情報を生成し出力するための第 2 の検索情報生成部 103 と、第 1 の検索情報生成部 102 から出力されたシーン内の各フレーム（時間分割ブロック）に対する要約情報、および / または、第 2 の検索情報生成部 103 から出力されたシーン内の各空間分割ブロックに対する要約情報を入力とし、第 1 の検索情報生成部 102 からの要約情報、および / または第 2 の検索情報生成部 103 からの要約情報の一方または双方に基づき、シーン全体に対する要約情報を生成し出力するための第 3 の検索情報生成部 104 と、第 1 の検索情報生成部 102 から出力された要約情報、第 2 の検索情報生成部 103 から出力された要約情報、第 3 の検索情報生成部 104 から出力される要約情報、および解析部 101 から出力される動画像構造情報を所定の形式に配置することにより、動画像に対応する検索情報を生成し出力するための第 4 の検索情報生成部 105 と、第 4 の検索情報生成部 105 から出力される、動画像に対する検索情報と、もとの動画像データとを受け、それらの一方または双方に検索情報と動画像データとの対応情報などを付加するかまたは別データとし、検索情報および動画像データとの対応情報などを記録媒体 107 に記録するための記録部 106 とを含む。

10

【0043】

本願発明の装置が扱う動画像データは、RGB 信号などの原画像信号、MPEG (Moving Picture Experts Group) などの動画圧縮方式を用いて符号化された画像データのいずれでも良い。本実施の形態では、説明の簡便のため、動画像データは符号化されたものであると仮定する。

20

【0044】

図 3 に示された各ブロックの機能についてさらに詳細に以下に説明する。解析部 101 は、動画像を時間的な小単位（シーン）に分割する。そのために解析部 101 は、画像データ中のシーンチェンジ点の検出を行なう。解析部 101 はこの例では、符号化画像データに含まれる符号化ブロックの予測モードの、フレーム内での頻度情報を用いてシーンチェンジ点を検出する。

【0045】

たとえば、図 4 を参照して、解析部 101 は、動画像 120 の中の第 1 のシーンチェンジフレーム 130、第 2 のシーンチェンジフレーム 132、第 3 のシーンチェンジフレーム 134、第 4 のシーンチェンジフレーム 136 を検出する。それによって解析部 101 は、第 1 のシーンチェンジフレーム 130 と第 2 のシーンチェンジフレーム 132 との間の画像データ（第 1 のシーンチェンジフレーム 130 を含む）からなる第 1 のシーン 140、第 2 のシーンチェンジフレーム 132 と第 3 のシーンチェンジフレーム 134 との間の画像データ（第 2 のシーンチェンジフレーム 132 を含む）からなる第 2 のシーン 142、第 3 のシーンチェンジフレーム 134 と第 4 のシーンチェンジフレーム 136 との間の画像データ（第 3 のシーンチェンジフレーム 134 を含む）からなる第 3 のシーン 144、および第 4 のシーンチェンジフレーム 136 から動画像 120 の最後までまでの画像データ（第 4 のシーンチェンジフレーム 136 を含む）からなる第 4 のシーン 146 とに動画像 120 を分割する。すなわち、解析部 101 は、動画像 120 をこの例では 4 つのシーン 140、142、144、146 に分割する。

30

40

【0046】

解析部 101 は、第 1 のシーンチェンジフレーム 130、第 2 のシーンチェンジフレーム 132、第 3 のシーンチェンジフレーム 134、第 4 のシーンチェンジフレーム 136 などのシーンチェンジフレームの、画像データ中の位置を表わす情報を、動画像構造情報として出力する機能を有する。

【0047】

図 5 を参照して、第 1 の検索情報生成部 102 は、解析部 101 から出力された動画像構造情報に基づいて動画像データを分割することにより得られたあるシーン 160 に対して以下のような処理を行なって、シーン 160 に含まれる各フレームの要約情報を作成す

50

る。なおシーン 160 はフレーム 162 A から 162 N を含むものとする。

【0048】

第 1 の検索情報生成部 102 は、フレーム 162 A ~ 162 N の各々について、動きベクトルのフレーム内の頻度情報 166 A から 166 N を求める。また第 1 の検索情報生成部 102 は、フレーム 162 A ~ 162 N の各々における動きベクトルの平均値 168 A から 168 N を求める。

【0049】

このようにして求められた頻度情報 166 と平均値情報 168 とが、この例では各フレームの要約情報 170 を構成する。要約情報は、フレーム番号の関数（すなわち時間の関数）として与えられる。

10

【0050】

この要約情報は、シーン 160 に含まれる各フレームの空間内に分布する情報を圧縮して得られた検索情報である。各フレームは、画像を時間的に分割した単位であって、時間分割ブロックの一種である。したがって要約情報 170 は、時間分割ブロック単位の検索情報といえることができる。

【0051】

第 2 の検索情報生成部 103 は、解析部 101 から出力された動画像構造情報（シーンチェンジ位置の情報）と、動画像データとを受ける。第 2 の検索情報生成部 103 はこの動画像構造情報と動画像データとを用いて以下のようにして空間分割ブロックに対する要約情報を生成する。

20

【0052】

なおここで空間分割ブロックとは以下のようなものをいう。一つのシーン内の全てのフレームの各々を同じ方法で複数個の分割ブロックに空間的に分割する。それら全てのフレームの、対応する位置の分割ブロックの全体をそのシーンの空間分割ブロックと呼ぶ。したがって、一つのシーンが n 個のフレームを含み、各フレームが m 個の分割ブロックに分割される場合には、そのシーンは m 個の空間分割ブロックに分割され、各空間分割ブロックは n 個の分割ブロックを含む。

【0053】

具体的には、図 6 を参照して、第 2 の検索情報生成部 103 ではシーン 160 に対して以下のような処理が行なわれる。前述のとおり、シーン 160 が複数個のフレーム 162 A ~ 162 N を含むものとする。これらフレーム 162 A ~ 162 N が各々、 M 個の分割ブロックに空間的に分割される。ここで、分割ブロックとは、画像面を複数に分割したときの、各小領域を指す。たとえば図 6 に示されるフレーム 162 A は、複数個の分割ブロック 180 A 1 から 180 A M に分割される。他のフレーム 162 B から 162 N も同様である。

30

【0054】

そして、複数個のフレーム 162 A ~ 162 N の各々の、同じ位置の分割ブロックによって、シーン 160 の一つの空間分割ブロックが形成される。たとえば N 個のフレームの各々の M 番目の分割ブロックの集合（その要素の数は N と等しい。）によって空間分割ブロック 182 - M が形成される。他の空間分割ブロックも同様に形成される。

40

【0055】

こうして得られた M 個の空間分割ブロックの各々に対して、時間軸方向に分布する情報を圧縮して要約情報 188 を生成する。ここでいう要約情報 188 は、各空間分割ブロックに対する要約情報であり、各空間分割ブロックに含まれる分割ブロックの、フレーム内の位置（座標）に対する関数値である。

【0056】

典型的には、要約情報 188 は、頻度情報 184 と平均値情報 186 とを含む。頻度情報 184 としては、各符号化ブロックの予測モードの空間分割ブロック内の頻度情報、または動き補償予測符号化時の動きベクトルの空間分割ブロック内の頻度情報を用いることができる。たとえば図示されていないが空間分割ブロック 182 - 1 からは頻度情報 18

50

4 Aが得られ、空間分割ブロック182 - 2からは頻度情報184 Bが得られ、以下同様にして空間分割ブロック182 - Mからは頻度情報184 Mが得られる。

【0057】

また平均値情報186としては、空間分割ブロック内の動きベクトルの総和もしくは平均値を用いることができる。たとえば図示されていないが空間分割ブロック182 - 1からは平均値情報186 Aが得られ、空間分割ブロック182 - 2からは平均値情報186 Bが得られ、以下同様にして空間分割ブロック182 - Mからは平均値情報186 Mが得られる。

【0058】

これ以外にも要約情報としては、空間分割ブロック内の動きベクトルの標準偏差などを用いることもできる。

【0059】

第3の検索情報生成部104は、第1の検索情報生成部102から出力された要約情報、または、第2の検索情報生成部103から出力された要約情報を受け、これらの一方または双方に基づいて、シーン全体に対する要約情報を生成し出力する。

【0060】

ここで、シーン全体に対する要約情報とは、シーンを検索する際に用いられる検索情報である。シーン全体に対する要約情報としては、動きベクトル、輝度値、または色差値などの、シーン全体における頻度情報、平均値、標準偏差などを用いることができる。

【0061】

第4の検索情報生成部105は、第1の検索情報生成部102から出力された要約情報と、第2の検索情報生成部103から出力された要約情報と、第3の検索情報生成部104から出力された要約情報と、解析部101から出力された動画像構造情報とを所定の形式に配置することにより、動画像に対応する検索情報を生成し出力して記録部106に与える。

【0062】

記録部106は、第4の検索情報生成部105から出力された、動画像に対応する検索情報と、もとの動画像データとを受けて、検索情報のみ、または検索情報と動画像データとを記録媒体107に記録する。このとき記録部106は、動画像データと検索情報とが対応していることを示す情報を動画像データもしくは検索情報のいずれか一方もしくは双方に付加するか、または動画像データおよび検索情報とは別データとして記録する。

【0063】

上に構造と一般的動作とを説明した本実施の形態にかかる動画像検索情報記録装置40において用いられる検索情報としては、動き補償予測符号化された動画像データに含まれる動きベクトルに基づいて得られる動きの大きさの情報をを用いるものとする。

【0064】

より具体的には、時間分割ブロック(フレーム)単位の検索情報は、シーン内の各フレームに対して、各符号化ブロックに付加された動きベクトルから求められた、フレーム全体における動きベクトルの絶対値の総和として求められる。なお、動きベクトルの絶対値の総和にかえて動きベクトルの絶対値の平均としても同じことである。実際、一フレームに含まれる符号化ブロックの数が定まっていれば、動きベクトルの絶対値の総和とその平均値との一方から容易に他方が求められる。

【0065】

空間分割ブロック単位の検索情報としては、各空間分割ブロックにおける動きベクトルの絶対値の総和が使用される。ここでも、動きベクトルの総和に替えてその平均を検索情報として用いることができる。空間分割ブロックを構成する各分割ブロックが符号化ブロックと一致する場合には、各分割ブロックに対応する動きベクトルの数は1となる。各分割ブロックが複数個の符号化ブロックを含む場合には、各分割ブロックに対応する動きベクトルの数はその分割ブロックに含まれる符号化ブロックの数と等しい。

【0066】

10

20

30

40

50

または空間分割ブロック単位の検索情報は、各空間分割ブロック内での動きベクトルの絶対値の総和を、空間分割ブロックを構成する分割ブロックの数で除した値でもよい。この場合の分割ブロックの数は、シーンを構成するフレームの数と等しい。さらにまた、空間分割ブロック単位の検索情報は、各空間分割ブロック内での動きベクトルの絶対値の総和を、空間分割ブロックを構成する分割ブロックに含まれる符号化ブロックの数の総和で除した値でもよい。

【0067】

時間分割ブロックまたは空間分割ブロックの検索情報として、動きベクトルの絶対値の総和を用いたとすると、シーンに含まれる全ての動きベクトルの絶対値の総和を計算することができる。本実施の形態では、こうして得られた総和をシーン単位の検索情報とする

10

【0068】

次に、空間分割ブロック単位の検索情報の生成に用いられる分割ブロックの具体例について説明する。図7に示されるように、シーン160がフレーム162A~162Nからなっているものとする。これらフレーム162A~162Nの動きベクトルの絶対値の総和から、シーン160のシーン単位検索情報Iが求められる。

【0069】

図8に示されるように、フレーム162A~162Nはそれぞれ、 2×2 個の分割ブロックに分割されているものとする。すなわちフレーム162A~162Nはそれぞれ4つの分割ブロックに分割されている。たとえばフレーム162Aは4つの分割ブロック190A1から190A4に分割される。以下のフレームについても図示しないが同様である。たとえばフレーム162Bはフレーム190B1から190B4に分割される、などである。したがって先の空間分割ブロックの説明からわかるように、シーン160は4つの空間分割ブロックに分割され、これら4つの空間分割ブロックに対して、各空間分割ブロック内の動きベクトルの絶対値の総和からそれぞれ空間分割ブロック単位の検索情報I1からI4がいずれも実数値として求められる。

20

【0070】

したがって、シーン160のためのシーン検索情報として、たとえばこれら検索情報I, I₁, I₂, I₃, I₄を順番に配列して得られる{I, I₁, I₂, I₃, I₄}を用いることができる。この検索情報をもとの画像データとともに記録すればよい。

30

【0071】

ただし、上記した検索情報は冗長性を含む。これは、この例ではシーン単位検索情報Iが空間分割ブロック単位の検索情報の和と等しいことによる。たとえばシーン単位検索情報Iは、空間分割ブロック単位の検索情報I₁からI₄より求めることができる。または、空間分割ブロック単位の検索情報I₁からI₄のうちの任意の3つとシーン単位検索情報Iとから、空間分割ブロック単位の検索情報の残りの一つを計算することができる。

【0072】

そこで、本実施の形態の装置では、I₄については検索情報には含ませないようにする。さらに本実施の形態の装置では、各空間分割ブロック単位検索情報のそのままの値ではなく、シーン単位の検索情報に対する各空間分割ブロック単位検索情報の値の百分率で示す。第1から第3の空間分割ブロックの検索情報の、シーン単位の検索情報Iに対する比率をP₁、P₂およびP₃とすれば、このときの空間分割ブロック単位の検索情報は図9に示されるようにして求められる。すなわち、第1の空間分割ブロックとシーン単位検索情報とから第1の空間分割ブロックの検索情報P₁が計算される。第2の空間分割ブロックとシーン単位検索情報とから第2の空間分割ブロックの検索情報P₂が計算される。第3の空間分割ブロックとシーン単位検索情報とから第3の空間分割ブロックの検索情報P₃が計算される。第4の空間分割ブロックについて検索情報を求めると冗長となるので、求めることはしない。つまり、シーン160は4つの空間分割ブロックに分割されるが、そのうち3つの検索情報を求めればよい。一般的にP_kは、 $P_k = 100 \times I_k / I$ によって求められる。

40

50

【 0 0 7 3 】

このように空間分割ブロックの検索情報として、シーン単位検索情報に対する百分率として記録すると、以下のような利点を得られる。第 1 に検索情報を記録するための信号のダイナミックレンジが制限されるため、より効率的に検索情報を記録できる。すなわち記録される情報量がそうでない場合より減少する。第 2 に、検索情報の値が正規化されるため、別々のシーンに付された検索情報を互いに比較することが可能となる。第 3 に、既に述べたとおり、記録すべき変数を一つ省略することができる。

【 0 0 7 4 】

このようにして動画像検索情報が第 4 の検索情報生成部 1 0 5 によって生成され、記録部 1 0 6 を介して記録媒体 1 0 7 に記録される。

10

【 0 0 7 5 】

さて、図 3 に示された動画像検索情報記録装置 4 0 によって記録媒体 1 0 7 に記録された動画像データおよび検索情報を用いて動画像検索をする装置について以下説明する。図 1 0 を参照して、この実施の形態にかかる動画像検索装置 5 0 は、記録媒体 1 0 7 に記録された動画像データとその検索情報とを読み出すための読出部 5 0 1 と、読出部 5 0 1 により読出された動画像の検索情報を保持し、検索要求に応じてシーン単位検索情報、空間分割ブロック単位検索情報、または時間分割ブロック単位検索情報を、動画像構造情報とともに出力するための検索情報管理部 5 0 2 と、シーン単位の検索要求に応答して、検索情報管理部 5 0 2 に対して要求を出し、検索情報管理部 5 0 2 からシーン全体に対する検索情報と、動画像構造情報とを受け、シーン単位の検索要求に合致したシーンを検出して、検出されたシーンに関する情報を出力するための第 1 の検索実行部 5 0 3 と、空間分割ブロック単位の検索要求に応答して、検索情報管理部 5 0 2 に対して要求を出し、検索情報管理部 5 0 2 から受取った、シーン内の各空間分割ブロックに対する要約情報と、動画像構造情報とに基づいて、検索要求に合致したシーンを検出して検出されたシーンに関する情報を出力するための第 2 の検索実行部 5 0 4 と、時間分割ブロック単位の検索要求に応答して検索情報管理部 5 0 2 に対して要求を出し、検索情報管理部 5 0 2 から受取ったシーン内の各フレームに対する要約情報と、動画像構造情報とに基づいて検索要求に合致したシーンまたはフレームを検出し、検出されたシーンまたはフレームに関する情報を出力するための第 3 の検索実行部 5 0 5 と、読出部 5 0 1 によって読出された動画像情報を保持し、第 1 の検索実行部 5 0 3、第 2 の検索実行部 5 0 4 または第 3 の検索実行部 5 0 5 から出力される検索結果に基づいて、検索結果に基づき検出されるシーンまたはフレームの画像データを出力するためのデータ管理部 5 0 6 と、オペレータからの検索指示入力に応答して、シーン単位の検索要求を第 1 の検索実行部 5 0 3 に、空間分割ブロック単位の検索要求を第 2 の検索実行部 5 0 4 に、または時間分割ブロック単位の検索要求を第 3 の検索実行部 5 0 5 に、それぞれ出力し、その結果データ管理部 5 0 6 から出力されるシーンまたはフレームの画像データを表示し、検索終了時には検索結果を出力するための検索制御部 5 0 7 とを含む。

20

30

【 0 0 7 6 】

この動画像検索装置 5 0 は以下のように動作する。検索開始時、オペレータは検索制御部 5 0 7 に対して検索指示を与える。検索制御部 5 0 7 は、この検索指示に応じて、シーン単位の検索、空間分割ブロック単位での検索、または時間分割ブロック単位での検索のいずれかを選択し、第 1 の検索実行部 5 0 3、第 2 の検索実行部 5 0 4、または第 3 の検索実行部 5 0 5 のいずれかに対して検索要求を出す。

40

【 0 0 7 7 】

第 1 の検索実行部 5 0 3、第 2 の検索実行部 5 0 4、第 3 の検索実行部 5 0 5 のうち、検索制御部 5 0 7 から検索要求を受取ったものは、検索情報管理部 5 0 2 から与えられる動画像構造情報と、シーン単位、空間分割ブロック単位、または時間分割ブロック（フレーム）単位の要約情報に基づいて、検索要求に合致したシーンまたはフレームを検出し、検出されたシーンまたはフレームに関する情報をデータ管理部 5 0 6 に与える。

【 0 0 7 8 】

50

データ管理部 506 は、第 1 の検索実行部 503、第 2 の検索実行部 504 または第 3 の検索実行部 505 から受取った、シーンまたはフレームに関する情報に対応するシーンまたはフレームの画像を動画データから抽出し検索制御部 507 に与える。

【0079】

検索制御部 507 は、データ管理部 506 から与えられた画像を表示する。表示された画像により、オペレータはその画像が所望のシーンであるかどうかを判断し、所望のシーンであれば検索結果を出力するように指示を検索制御部 507 に与える。検索制御部 507 はこの指示に応答して、検索結果を出力する。

【0080】

検索制御部 507 により表示された画像が所望のシーンを表わすものではなかった場合、オペレータはさらに次の検索指示を検索制御部 507 に与える。以下検索制御部 507 は上述したと同様の動作を行なって、検索結果のシーンの表示を行なう。以下、検索が終了するまでこの動作を繰り返す。

【0081】

具体的な動画検索の例について説明する。この実施の形態では、検索情報として、動き補償予測符号化された動画データに含まれる動きベクトルに基づいて得られる動きの大きさに関する情報を用いる。

【0082】

所望のシーンまたはシーン内のフレームは、次のようにして検索される。

【0083】

シーン単位の検索情報を利用すると、たとえば全体的に動きの大きなシーンと、全体的に動きの小さなシーンとを選別することができる。

【0084】

空間分割ブロック単位の検索情報を利用すると、たとえば一シーンにわたって動きが空間的に遍在するようなシーンと、動きが空間的に比較的均等に分布するシーンとを選別できる。また、動きの遍在する部分が画像のどの部分かを指定することができる。たとえば、画像の中央部分に激しく動く部分（被写体）があるシーン、または画像の下半分に激しく動く部分があるシーン（空を含む画像）などを選び出すことができる。

【0085】

本実施の形態では、空間分割ブロックという、従来は考慮されていなかったものを単位とする検索情報をシーンの検索に用いる。そのため、次のような効果が得られる。

【0086】

図 11 から図 13 を参照して、図 12 (A) に示すシーン A と、図 13 (A) に示すシーン B とについて、図 11 (A) に示す基準となるシーンとの類似度を計算する場合を考える。こうした「類似度」計算により、基準となるシーンと類似した動き特徴を有するシーンを検出することができる。図 12 から図 13 は、シーン単位の検索情報のみを用いた場合を示す。

【0087】

図 11 から図 13 の (B) に、それぞれの (A) に示す画像に対するシーン単位検索情報を示す。図 11 から図 13 の (A) に示されるように、シーン A およびシーン B は、基準となるシーンに対して、シーン全体として同じような動きの強さを持っている。そのため、図 11 から図 13 の (B) に示されるように、対応のシーン単位検索情報も同様の値となる。すなわち、シーン A、B のいずれも、基準となるシーンと類似していると判断される。

【0088】

続いて、図 14 から図 16 を参照して、空間分割ブロック単位の検索情報を用いた場合の類似度検索の特徴について説明する。図 14 から図 16 の (A) にはそれぞれ基準となるシーン、シーン A、B を示す。図示されるように、各シーンの各フレームは 4 つの分割ブロックを含む。そのうちそれぞれ第 1 から第 3 の分割ブロックからなる空間分割ブロックに対して、空間分割ブロック単位の検索情報 P_1 、 P_2 および P_3 が計算されている。ま

10

20

30

40

50

た図14から図16の(B)に示すように、基準となるシーン、シーンAおよびBに対するシーン単位の検索情報はほぼ同じ値である。

【0089】

しかし、図14から図16の(C)に示されるように、空間分割ブロック単位の検索情報で各シーンを比較すると、次の事実が分る。すなわち、シーンAの方が、基準となるシーンの空間分割ブロック単位検索情報の分布と類似した分布を有する。これは、シーンBでは人物が右の方によっているため、空間的に動きの偏りができるためである。基準となるシーンおよびシーンAはいずれも、人物がほぼ中央にいる。したがって、基準となるシーンと類似したシーンとしてシーンAは検出されるが、シーンBは検出されない。つまり、シーン全体としての動きはシーンA、Bとも基準となるシーンと類似している場合であ

10

っても、動きが遍在しているか否かによってシーンを選別することができる。

【0090】

さらに、フレーム(時間分割ブロック)単位の検索情報を利用すれば、たとえばシーン内の特に動きの激しいフレーム、逆に動きの止まっているようなフレームを選別することができる。

【0091】

上記したシーン単位の検索情報、空間分割ブロック単位での検索情報、および時間分割ブロック単位での検索情報を組み合わせることによって、さらに細かく所望のシーンの検索をすることができる。

【0092】

20

たとえば、全体的に動きの大きなシーンを選び、特に画像の中央部分に激しく動く部分を持つシーンに絞り込み、さらに絞り込まれた各シーン内の、特に動きの激しいフレームをキーフレームとして取出す、という形で動画像検索を実行することができる。

【0093】

このように、本実施の形態の装置によれば、多面的な条件指定により、画像の動きなどの画像特徴に基づいて、所望のシーンの画像を効率的に検索することができる。検索に要する時間も短縮される。

【0094】

上記した実施の形態は、本発明を実施するための一つの例に過ぎない。この他にも種々の変形をこれらの装置に対して行なうことができる。

30

【0095】

たとえば、この実施の形態の装置では解析部101が自動的に動画像データに対する信号解析によりシーンチェンジ点を見出している。しかし本発明はこれには限定されず、たとえばオペレータが動画像を見て、オペレータの主観によって動画像の構造を解析しシーンの分割位置にタグを付すようにしてもよい。この場合解析部101は、動画像データを同様に解析するが、その際、動画像データに付されたタグを見て動画像をシーンに分割する。

【0096】

上記した実施の形態では、フレームの要約情報の頻度情報166として動き補償予測符号化時の動きベクトルのフレーム内頻度情報を用いた。しかしこの発明はこれには限定されず、頻度情報として各符号化ブロックの予測モードのフレーム内頻度情報を用いることもできる。

40

【0097】

また、要約情報としては、上記したほかに、フレーム内の動きベクトルの総和もしくは標準偏差、または原画像信号もしくは復号画像信号を用いて計算されるフレーム内の平均輝度値、もしくは平均色差値などを用いてもよい。

【0098】

検索方法および手順は、ここに示した組合せには限定されない。選択される候補を絞り込む順序を変えることも可能である。また、生成される検索情報によっても、検索のための方法およびその手順は変化する。

50

【 0 0 9 9 】

またたとえば、上記した実施の形態の装置では、空間分割ブロックとして、1フレームを $2 \times 2 = 4$ 分割した分割ブロックの1シーンにわたる集合を用いた。分割の個数は4個には限定されない。たとえば図17に示されるように 4×4 、 8×8 、またはそれ以上の数による分割を行なうことも可能である。 4×4 分割の場合には、百分率で表わした空間分割ブロック単位検索情報は P_1 から P_{15} である。 P_{16} は100%から他の空間分割ブロックの検索情報の和を減ずることにより得られる。 8×8 分割の場合には、百分率で表わした空間分割ブロック単位検索情報は P_1 から P_{63} である。 P_{64} も100%から他の空間分割ブロックの検索情報の和を減ずることにより得られる。

【 0 1 0 0 】

さらに、このように複数通りの分割方法のいずれを用いるかを選択できるようにしてもよい。この場合、選択された個数(「解像度」と呼ぶことができる。)にしたがって分割された分割ブロックによって空間分割ブロックを構成し、その各々に対して検索情報を求める。さらに、検索情報を記録するときには、シーン単位検索情報と、選択された解像度による空間分割ブロック単位の検索情報に加えて、図18にその値と分割数との対応関係が示される分割数指定フラグを検索情報に付加する。分割数指定フラグが「0」であれば $\{P_1, P_2, P_3\}$ 、「1」であれば $\{P_1, P_2, \dots, P_{15}\}$ 、「2」であれば $\{P_1, P_2, \dots, P_{63}\}$ 、「3」であれば $\{P_1, P_2, \dots, P_{256}\}$ の空間分割ブロック単位の検索情報が記録される。

【 0 1 0 1 】

このように空間分割ブロックの解像度を複数通り指定できることにより、以下のような効果が生じる。

【 0 1 0 2 】

まず、画像のうち上下左右のおおまかな画像の動きの分布を知りたいという、比較的簡易な検索のみが要求されるデータベースでは、 2×2 など、小さな分割数を用いる。検索情報の記録量を抑えることができる。

【 0 1 0 3 】

画像の特定の一部または離れた2つの部分で動きが大きいシーンを抽出したい、というような要求が頻繁にあるデータベースでは、 8×8 など、比較的大きな分割数を用いる。高機能な検索を実現することができる。高解像度の検索情報は、低解像度の検索情報を含むので、低解像度の検索が必要なときにも、低解像度の検索情報を重複して持つ必要はない。

【 0 1 0 4 】

以上説明した実施の形態では、 2×2 、 4×4 、および 8×8 のように、縦方向の分割数と横方向の分割数とが等しく、かついずれも2のべき乗である。こうした分割数を選択的に用いると、分割の解像度が異なる空間分割ブロック単位検索情報を記録したデータベース間においても検索情報の間の対応付けが可能となる。また、異なる分割解像度どうしでの検索情報を比較することも可能となる。

【 0 1 0 5 】

さらに、図19に示されるように、複数種類の解像度の空間分割ブロック単位の検索情報を階層的に記録してもよい。たとえば、シーンの検索情報は、シーン単位の検索情報200と、 2×2 分割による空間分割ブロック単位の検索情報202と、 4×4 分割による空間分割ブロック単位の検索情報204と、 8×8 分割による空間分割ブロック単位の検索情報206とを含む。

【 0 1 0 6 】

この階層構造では、ある層の空間分割ブロック単位の検索情報は、一段高解像度の4つの空間ブロック単位の検索情報で表わされる。したがって、各層において、4つの空間分割ブロック単位の検索情報のうちの一つを記録しなくとも、検索に必要な全ての情報を得ることができる。

【 0 1 0 7 】

10

20

30

40

50

図 19 に示した例による検索情報は、図 17 に示した 8×8 にフレームを分割して形成される空間分割ブロックを単位とする検索情報を記録するのと同じ解像度の検索情報を含む。一方で図 19 に示した検索情報によれば、他の解像度の検索情報を容易に得ることができるという利点を有する。

【0108】

上に説明した実施の形態では、シーン内の全フレームを、 2×2 、 4×4 、 8×8 などのように $n \times n$ の配置に分割して、空間分割ブロックを形成している。しかし本発明によれば、フレームの分割方法はこれには限られない。たとえば図 20 から図 24 に示されるような形でフレームを分割して、得られた空間分割ブロックを単位として検索情報を生成するようにしてもよい。図 20 から図 24 に示した例では、シーン内の全フレームを 2 のべき乗 (2^n) 個 (図 20 では $n = 2$ 、図 21 では $n = 3$ 、図 22 では $n = 4$ 、図 23 では $n = 5$ 、図 24 では $n = 6$) に分割している。このような分割方法を用いても、上に述べた実施の形態の装置で得られたものと同様の検索機能を実現するための検索情報を生成することができる。

【0109】

図 20 ~ 図 24 に示したようなフレームの分割方法によっても、上に記載した実施の形態と同様に、異なる解像度の検索情報同士の対応付けが容易である。また、異なる解像度の検索情報を互いに容易に比較することができる。

なお、図 5 に示される例の場合は、たとえばフレーム 162N を一つの時間分割ブロック 164N として要約情報を求めている。しかし本発明はこうした実現例に限定されない。たとえば一つの時間分割ブロックが 2 以上のフレームを含むようにしてもよい。また、各フレームをサブサンプリングしてサブフレームを生成し、一時間分割ブロックが一または複数個のサブフレームを含むようにしてもよい。またシーン内のフレームを適宜間引いて、間引いた各フレームに対する要約情報を用いるとしてもよい。

【0110】

なお、図 6 に示される例または図 8 に示される例の場合は、空間分割ブロックに含まれる全ての分割ブロックを用いて空間分割ブロックの要約情報を求めている。しかし本発明はこうした実現例に限定されない。必ずしも空間分割ブロック内の全ての分割ブロックに対応する画像データを用いる必要はなく、分割ブロックを適宜に間引いて、または分割ブロック内の画像データを適宜に間引いて、空間分割ブロックの要約情報を生成してもよい。

【0111】

なお、図 5 および図 6 を参照して説明した、時間分割ブロック単位もしくは空間分割ブロック単位の検索情報、またはシーン単位の検索情報を生成する際に、動画を撮影したカメラの動きによるバックグラウンドの動きへの影響を除くために、カメラの動きに相当する動きベクトル分を補正して検索情報を生成するようにしてもよい。

【0112】

さらに、上記した実施の形態では、動画画像検索情報記録装置 40 と動画画像検索装置 50 とは分離された別の装置である。そして検索情報は記録媒体 107 を介して動画画像検索情報記録装置 40 から動画画像検索装置 50 に渡されている。これに替えて、例えば、図 3 に示す第 1 の検索情報生成部 102、第 2 の検索情報生成部 103、第 3 の検索情報生成部 104 で生成された各種の検索情報を、解析部 101 から出力される動画画像構造情報とあわせて、直接図 10 の検索情報管理部 502 に渡すようにしてもよい。この場合の検索処理は、いわゆるリアルタイム検索と呼ばれる。

【0113】

上記した実施の形態でフレームを分割した分割ブロックは、符号化ブロックと一致してもよい。符号化ブロックとは別の、適当な大きさのブロックを新たに定義してもよい。上の実施の形態では、分割ブロックは対称形であり、かつ分割ブロックの配置も対称である。しかし分割ブロックを新しく定義する場合には、ブロックが対称形である必要はない。またブロックの配置に関しても対称である必要はない。また、画像面内を適宜に間引くこ

10

20

30

40

50

とも可能である。すなわち、画像内に分割ブロックが含まれない領域があってもよい。また、分割ブロックが部分的に重複してもよい。すなわち、画像内に複数の分割ブロックに含まれる領域があってもよい。

【0114】

上記した実施の形態では、図3に示される第3の検索情報生成部104は、第1の検索情報生成部102から出力される要約情報および第2の検索情報生成部103から出力される要約情報を受けて、シーン全体に関する要約情報を生成している。しかし本発明はこのような実現例には限定されない。たとえば第3の検索情報生成部104は、入力された画像データから直接にシーン全体に関する要約情報を生成するようにしてもよい。

【0115】

また、上記した実施の形態では、各種の要約情報（検索情報）を、物理的に定義された一シーンを単位として求めている。しかし検索情報を、シーンを単位として求める必要は必ずしもない。たとえば1つのシーンを複数に分割したサブシーンを単位として検索情報を求めてもよい。すなわち「シーン」という語は「サブシーン」をも意味する。逆に、複数のシーンからなる複合シーンを単位として検索情報を求めることもできる。すなわち「シーン」という語は、「複数のシーンの集まり」をも意味する。

【0116】

また、上記した実施の形態では、1つの動画画像データに対して検索情報を生成し、記録し、検索を実行しているが、本発明はこうした実現例に限定されない。すなわち、複数の動画画像データを検索の対象とし、検索情報を生成し、記録し、複数の動画画像データ中から所望のシーンまたはフレームを検索することも想定するものである。

【0117】

その一例として、図25に示す、複数のシーンの各々が独立した動画画像データとして蓄積されている画像データベースからの動画画像検索を考える。昨今の、動画画像を撮影することが可能なデジタルカメラなどで動画画像の撮影を行なった場合には、各撮影された動画画像、すなわち1回の録画開始から録画停止までの1シーンを表わす動画画像データが、それぞれ独立したデータファイルとして生成される。したがって、図25に示すような画像データベースも容易に想起できる。

【0118】

このような画像データベースに対して、図3に示される動画画像検索情報記録装置40によって検索情報を生成する。このとき、入力される動画画像データは、データベースに存在する複数の動画画像データである。格動画画像データは既にシーンに分解されている（1つの動画画像データは各々一つのシーンのみを含んでいる）ため、動画画像構造情報は不要である。したがって解析部101では動画画像の構造解析を行なう必要はない。第1、第2、第3の検索情報生成部102、103、104は、動画画像データからそれぞれフレーム（時間分割ブロック）、空間分割ブロック、シーン全体に対する要約情報を生成する。第4の検索情報生成部105は、これらの要約情報を所定の形式に配置して動画画像に対する検索情報を生成する。記録部106は、各動画画像データと各動画画像に対応する検索情報とを受け、各動画画像データと検索情報との対応情報をいずれか一方もしくは双方に付加するか、または別データとし、検索情報と対応情報とを記録媒体107に記録する。記録媒体107を元の画像データベース内としてもよい。また、各検索情報は各々独立して記録しても、または画像データベースに対応する検索情報としてまとめて記録してもよい。

【0119】

また、このような画像データベースならびに生成および記録された検索情報から、図10に示される動画画像検索装置によって検索を実行する。このとき、読出部501で読出される検索情報管理装置502で管理されるのは、データベースに存在する複数の動画画像データに対応した検索情報であり、各動画画像データに対応した動画画像構造情報は不要である。第1、第2、第3の検索実行部503、504、505ではそれぞれ、対応する検索情報を用いて検索要求と合致したシーンまたはフレームを検出し、検出されたシーンまたはフレームに関する情報を出力する。データ管理部506および検索制御部507は前記した

10

20

30

40

50

実施の形態と同様に動作し、検索結果が得られる。

【0120】

今回開示された実施の形態はすべての点で例示であって制限的なものではないと考えられるべきである。本発明の範囲は上記した説明ではなくて特許請求の範囲によって示され、特許請求の範囲と均等の意味および範囲内でのすべての変更が含まれることが意図される。

【図面の簡単な説明】

【図1】 想定された従来の動画像の検索情報記録装置の概略ブロック図である。

【図2】 想定された従来の動画像の検索装置の概略ブロック図である。

【図3】 本発明の一実施の形態にかかる動画像検索情報記録装置のブロック図である。 10

【図4】 本発明の一実施の形態にかかる動画像検索情報記録装置において、動画像をシーンに分解する過程を示す図である。

【図5】 本発明の一実施の形態にかかる動画像検索情報記録装置において、時間分割ブロック単位の検索情報が生成される過程を示す図である。

【図6】 本発明の一実施の形態にかかる動画像検索情報記録装置において、空間分割ブロック単位の検索情報が生成される過程を示す図である。

【図7】 本発明の一実施の形態にかかる動画像検索情報記録装置におけるシーン単位の検索情報の概念を説明するための図である。

【図8】 本発明の一実施の形態にかかる動画像検索情報記録装置における空間分割ブロック単位の検索情報の概念を説明するための図である。 20

【図9】 本発明の一実施の形態にかかる動画像検索情報記録装置における、冗長性を排した空間分割ブロック単位の検索情報の概念を説明するための図である。

【図10】 本発明の一実施の形態にかかる動画像検索装置のブロック図である。

【図11】 シーン単位検索情報による類似度検索における、基準となるシーンを示す図である。

【図12】 シーン単位検索情報による類似度検索における、一つのシーンを示す図である。

【図13】 シーン単位検索情報による類似度検索における、他のシーンを示す図である。

【図14】 空間分割ブロック単位検索情報による類似度検索における、基準となるシーンを示す図である。 30

【図15】 空間分割ブロック単位検索情報による類似度検索における、一つのシーンを示す図である。

【図16】 空間分割ブロック単位検索情報による類似度検索における、他のシーンを示す図である。

【図17】 本発明の一実施の形態にかかる動画像検索情報記録装置において、シーン単位検索情報の記録と、空間分割ブロック単位の検索情報の選択的な記録という概念を示す図である。

【図18】 分割指定数フラグの値の意味を表形式で示す図である。

【図19】 本発明の一実施の形態にかかる動画像検索情報記録装置における、シーン単位の検索情報の記録と、空間分割ブロック単位の検索情報の階層的な記録という概念を表わす図である。 40

【図20】 フレームを 2^2 個に分割するときの分割例を示す図である。

【図21】 フレームを 2^3 個に分割するときの分割例を示す図である。

【図22】 フレームを 2^4 個に分割するときの分割例を示す図である。

【図23】 フレームを 2^5 個に分割するときの分割例を示す図である。

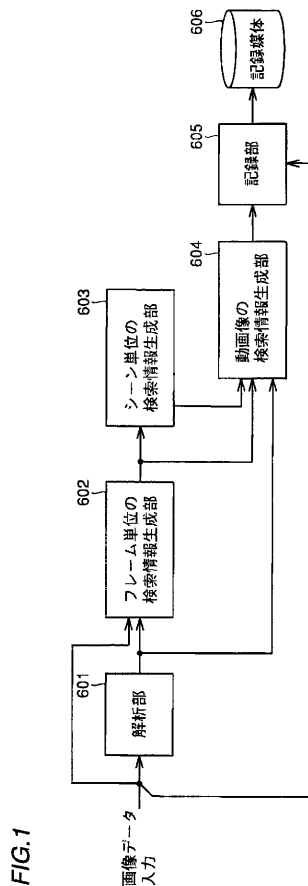
【図24】 フレームを 2^6 個に分割するときの分割例を示す図である。

【図25】 動画像データがシーン単位に分割して蓄積されている動画像データベースを示す図である。

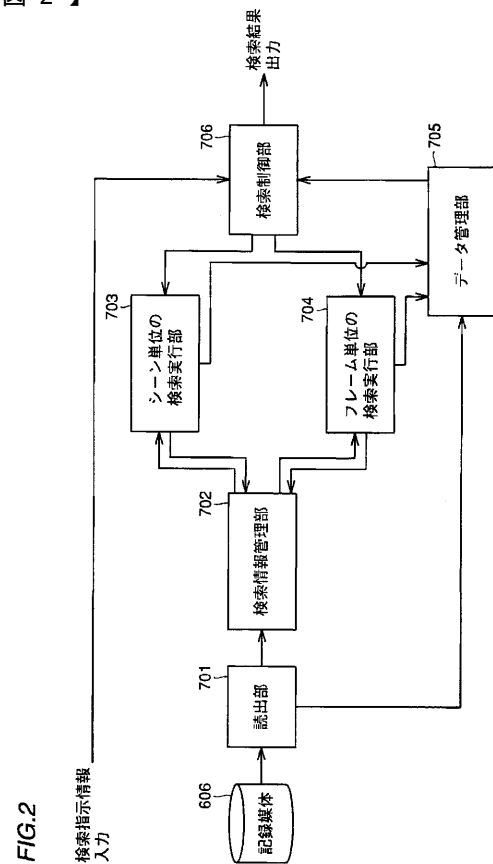
【符号の説明】

101 解析部、102 第1の検索情報生成部、103 第2の検索情報生成部、104 第3の検索情報生成部、105 第4の検索情報生成部、106 記録部、107 記録媒体、501 読出部、502 検索情報管理部、503 第1の検索実行部、504 第2の検索実行部、505 第3の検索実行部、506 データ管理部、507 検索制御部、601 解析部、602 フレーム単位の検索情報生成部、603 シーン単位の検索情報生成部、604 動画像の検索情報生成部、605 記録部、606 記録媒体、701 読出部、702 検索情報管理部、703 シーン単位の検索実行部、704 フレーム単位の検索実行部、705 データ管理部、706 検索制御部。

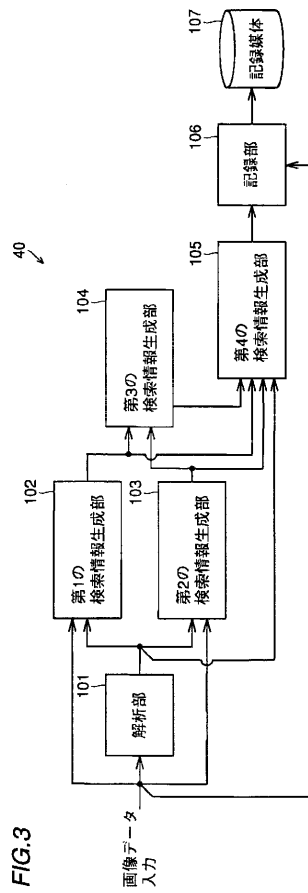
【図1】



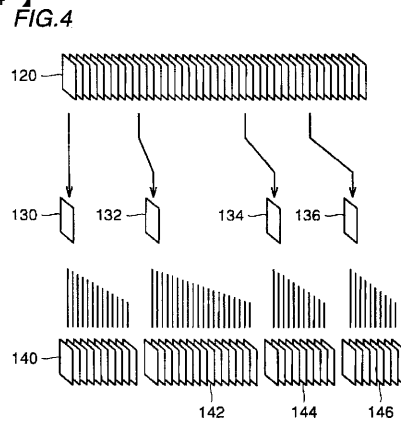
【図2】



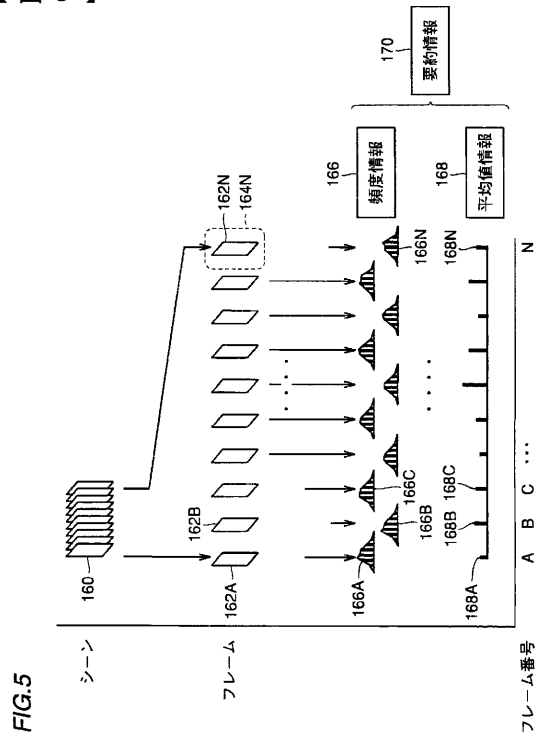
【 図 3 】



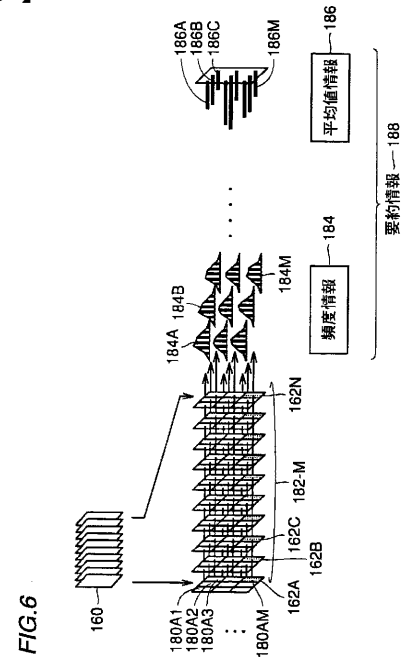
【 図 4 】

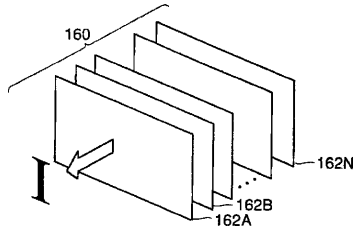
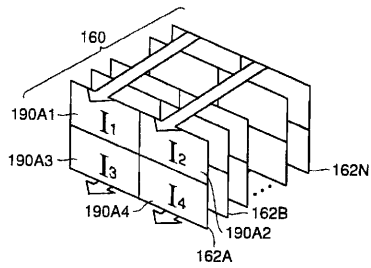
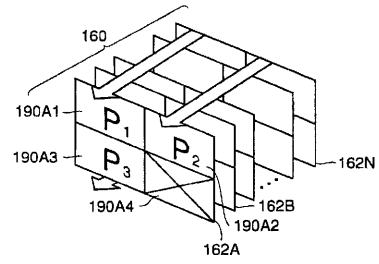


【 図 5 】

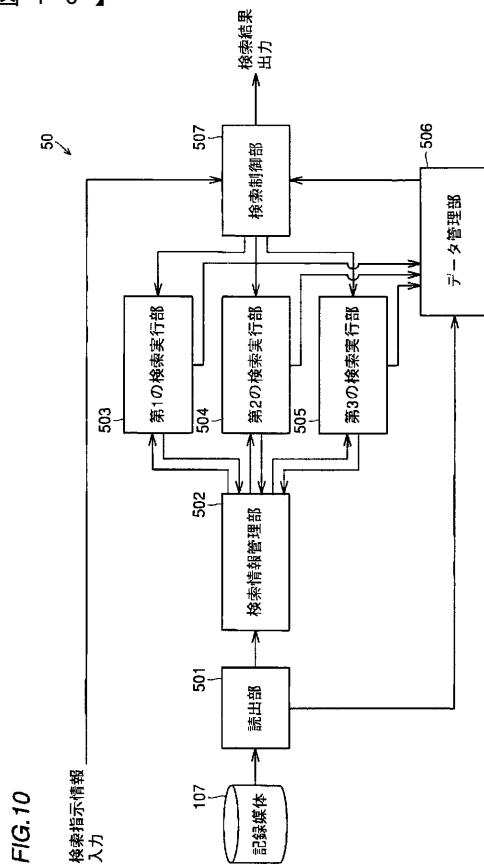


【 図 6 】



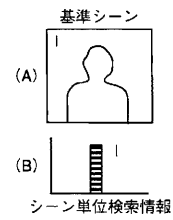
【図7】
FIG.7【図8】
FIG.8【図9】
FIG.9

【図10】



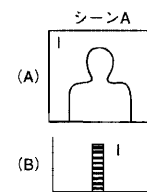
【図11】

FIG.11



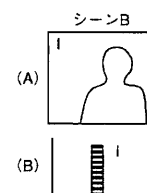
【図12】

FIG.12



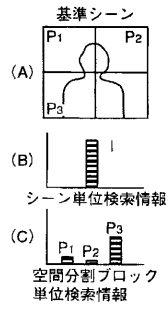
【図13】

FIG.13



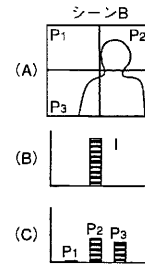
【図 14】

FIG.14



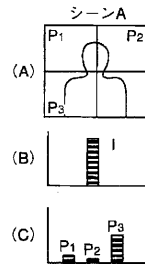
【図 16】

FIG.16



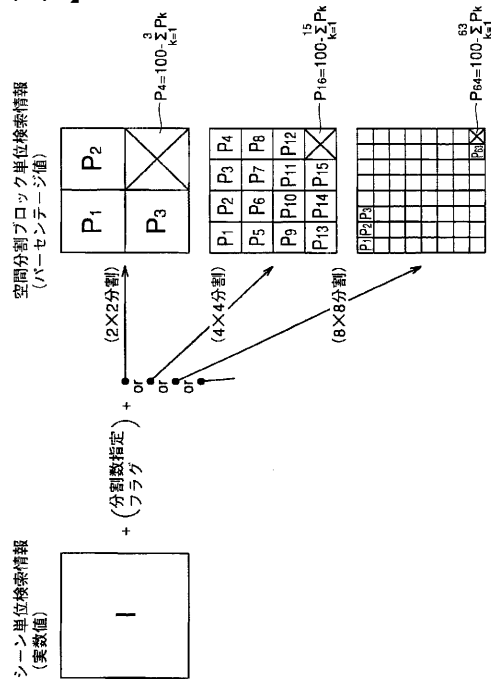
【図 15】

FIG.15



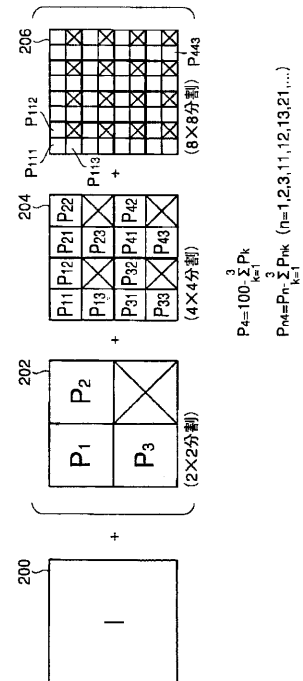
【図 17】

FIG.17



【図 19】

FIG.19



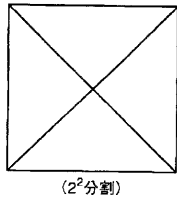
【図 18】

FIG.18

分割数指定フラグ	分割数
0	2×2
1	4×4
2	8×8
3	16×16

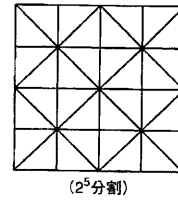
【図 2 0】

FIG.20



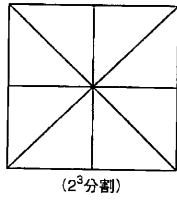
【図 2 3】

FIG.23



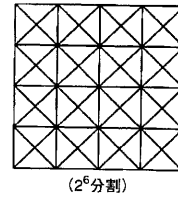
【図 2 1】

FIG.21



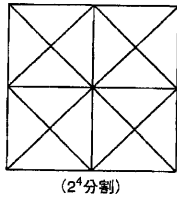
【図 2 4】

FIG.24



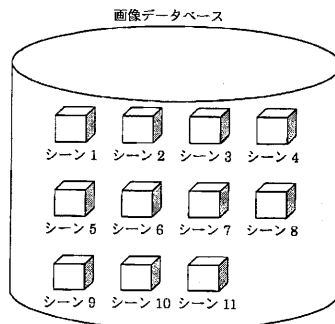
【図 2 2】

FIG.22



【図 2 5】

FIG.25



フロントページの続き

(72)発明者 渡部 秀一

千葉県千葉市緑区誉田町2 - 24 - 7 ラポール誉田A125

合議体

審判長 田口 英雄

審判官 青柳 光代

審判官 長 由紀子

(56)参考文献 宮坂, 吉田, MPEGビットストリーム中の動きベクトルの動画像検索への応用, 1999年電子情報通信学会総合大会講演論文集 情報・システム(2), 日本, 社団法人電子情報通信学会, 1999年3月28日, P. 12, D - 11 - 12

粕谷, 宮森, 富永, 圧縮ビデオ中の動きベクトルを利用した高速検索手法の提案, 1996年電子情報通信学会情報・システムソサイエティ大会講演論文集, 日本, 社団法人電子情報通信学会, 1996年 9月21日, P. 278, D - 276

粕谷, 前田, 宮森, 富永, 圧縮映像中のパラメータを利用した高速照合とその検索方式の提案, 情報処理学会研究報告, 日本, 社団法人情報処理学会, 1997年 6月 6日, VOL. 97, NO. 58, P. 25 - 32 (97 - AVM - 17 - 5)

(58)調査した分野(Int.Cl., DB名)

G06F17/30, H04N5/76