

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4981221号
(P4981221)

(45) 発行日 平成24年7月18日(2012.7.18)

(24) 登録日 平成24年4月27日(2012.4.27)

(51) Int.Cl.

F 1

HO4N	5/91	(2006.01)	HO4N	5/91	N
HO4N	5/76	(2006.01)	HO4N	5/76	Z
G06T	7/00	(2006.01)	G06T	7/00	100B
G06T	7/20	(2006.01)	G06T	7/00	100D
			G06T	7/20	A

請求項の数 18 外国語出願 (全 21 頁)

(21) 出願番号

特願2001-227960 (P2001-227960)

(22) 出願日

平成13年7月27日 (2001.7.27)

(65) 公開番号

特開2002-125178 (P2002-125178A)

(43) 公開日

平成14年4月26日 (2002.4.26)

審査請求日

平成20年7月24日 (2008.7.24)

(31) 優先権主張番号

09/627762

(32) 優先日

平成12年7月28日 (2000.7.28)

(33) 優先権主張国

米国 (US)

(73) 特許権者

500046438

マイクロソフト コーポレーション
アメリカ合衆国 ワシントン州 9805
2-6399 レッドmond ワン マイ
クロソフト ウェイ

(74) 代理人

100140109

弁理士 小野 新次郎

(74) 代理人

100089705

弁理士 社本 一夫

(74) 代理人

100071124

弁理士 今井 庄亮

(74) 代理人

100076691

弁理士 増井 忠式

(74) 代理人

100075270

弁理士 小林 泰

最終頁に続く

(54) 【発明の名称】メディア・セグメント化システムおよび関連する方法

(57) 【特許請求の範囲】

【請求項 1】

カラー空間において、受けたメディア・コンテンツに含まれる各フレームについて、カラー情報を分析して、画素情報が最も多い1つ以上の優勢カラー・オブジェクトを識別するステップと、

前記受けたメディア・コンテンツ全体にわたって、前記カラー空間において、フレームを跨いで前記識別された1つ以上の優勢カラー・オブジェクトを追跡して、各ショットについて、最も長い持続時間有する優勢カラー・オブジェクトを識別するステップと、

該識別された優勢カラー・オブジェクトから各ショットについて優勢カラー・オブジェクト・ヒストグラムを生成するステップと、

前記優勢カラー・オブジェクト・ヒストグラム間のヒストグラム交点を計算してショット間の相関を定量化し、前記計算されたヒストグラム交点が最小相関スレシホールドを満たすショットを集合化することにより、シーンを識別するステップと、

を備えた、コンピュータにより実行する方法。

【請求項 2】

請求項1記載の方法において、前記カラー情報を分析するステップは、

各フレームの画素、またはイントラ・エンコード・フレームのDCブロックを、量子化カラー空間に投影するステップと、

前記画素および/またはブロックの正規分布を生成して、前記メディア・コンテンツのカラー・ヒストグラムを形成するステップと、

10

20

を含む、コンピュータにより実行する方法。

【請求項 3】

請求項 2 記載の方法であって、更に、

前記カラー・ヒストグラムにおいて局部最大点を識別するステップと、

前記識別した局部最大値の各々を取り囲むN個の量子化ユニットの球として、カラー・オブジェクトを定義するステップと、

を含む、コンピュータにより実行する方法。

【請求項 4】

実行されると、請求項 1 記載のコンピュータにより実行する方法をコンピュータに実施させる複数の実行可能命令を備えた記憶媒体。

10

【請求項 5】

計算システムであって、

複数の実行可能命令を含む記憶媒体と、

前記記憶媒体に結合され、少なくとも前記複数の実行可能命令の部分集合にアクセスして実行して、請求項 1 記載のコンピュータにより実行する方法を実施する、実行ユニットと、

を備えた計算システム。

【請求項 6】

受けたメディアのショットに関連した各フレームについて、カラー空間内において、画素情報が最も多い1つ以上の優勢カラー・オブジェクトを識別するステップと、

20

前記カラー空間において、前記識別された1つ以上の優勢カラー・オブジェクトをフレームを跨いで追跡して、各ショットについて、最も長い持続時間を有する優勢カラー・オブジェクトを識別するステップと、

該識別された優勢カラー・オブジェクトから各ショットについて優勢カラー・オブジェクト・ヒストグラムを生成し、前記優勢カラー・オブジェクト・ヒストグラム間のヒストグラム交点を計算することにより、ショット間の相関スコアを生成するステップと、

前記相関スコアが最小相関スレシホールドを満たすショットをシーンにセグメント化するステップと、

を備えた、コンピュータにより実行する方法。

【請求項 7】

30

請求項 6 記載の方法であって、更に、

受けたメディアのフレームから1つ以上のスライスを抽出し、前記受けたメディアのフレームの1つ以上のモーション・パターンを分析するステップと、

前記フレームの前記モーション・パターンに基づいて、フレーム間の相関スコアを生成するステップと、

フレーム間の前記相関スコアに基づいて、ショット内においてセグメント境界を選択するステップと、

を含む、コンピュータにより実行する方法。

【請求項 8】

40

請求項 6 記載の方法において、前記セグメント化するステップは、

動的拡張ウィンドウに、相関スコアが所定のスレシホールドを超過したショットを収容するステップ、

を含む、コンピュータにより実行する方法。

【請求項 9】

請求項 8 記載の方法において、前記相関スコアを生成するステップは、

前記フレームのモーション・パターンに基づいて、ショットの各セグメントに対して1つ以上のキー・フレームを選択するステップと、

前記ショットのキー・フレームの視覚的特徴に基づいて、ショット間の前記相関スコアを生成するステップと、

を更に含む、方法。

50

【請求項 10】

複数の命令を備えた記憶媒体であって、前記複数の命令は、実行されると、コンピュータに、

メディア分析エージェントを実現して、受けたメディアのショットに関連した各フレームについて、カラー空間内において、画素情報が最も多い1つ以上の優勢カラー・オブジェクトを識別するステップと、

前記カラー空間において、前記識別された1つ以上の優勢カラー・オブジェクトをフレームを跨いで追跡して、各ショットについて、最も長い持続時間を有する優勢カラー・オブジェクトを識別するステップと、

該識別された優勢カラー・オブジェクトから各ショットについて優勢カラー・オブジェクト・ヒストグラムを生成し、前記優勢カラー・オブジェクト・ヒストグラム間のヒストグラム交点を計算することにより、ショット間の相関スコアを生成するステップと、 10

前記相関スコアが最小相関スレシホールドを満たすショットをシーンにセグメント化するステップと、

を含む方法を実行させる、記憶媒体。

【請求項 11】

請求項10記載の記憶媒体であって、前記方法は、更に、

受けたメディアのフレームから1つ以上のスライスを抽出して、前記受けたメディアのフレームの1つ以上のモーション・パターンを分析するステップと、

前記フレームの前記モーション・パターンに基づいてフレーム間の相関スコアを生成するステップと、 20

フレーム間の前記相関スコアに基づいてショット境界を選択するステップと、
を含む、記憶媒体。

【請求項 12】

請求項10記載の記憶媒体であって、前記方法は、更に、

動的拡張ウィンドウに、相関スコアが所定のスレシホールドを超過したショットを収容するステップ

を含む、記憶媒体。

【請求項 13】

計算システムであって、

メディア・コンテンツを受け取り、供給するメモリ・デバイスと、

前記メモリ・デバイスに結合されたメディア分析エージェントであって、

前記メディア・コンテンツ内のショットに関連したカラー・ヒストグラムを生成し、

前記カラー・ヒストグラムにおける局部最大値からカラー・オブジェクトを識別し、

各ショットに関連した各フレームについて、画素情報が最も多い1つ以上の優勢カラー・オブジェクトを識別し、

前記識別された1つ以上の優勢カラー・オブジェクトをフレームを跨いで追跡して、各ショットについて、最も長い持続時間を有する優勢カラー・オブジェクトを識別し、

該識別された優勢カラー・オブジェクトから各ショットについて優勢カラー・オブジェクト・ヒストグラムを生成し、前記優勢カラー・オブジェクト・ヒストグラム間のヒストグラム交点を計算することにより、ショット間の相関スコアを生成し、 40

前記相関スコアが最小相関スレシホールドを満たすショットをシーンにセグメント化する

ように構成された、メディア分析エージェントと、
を備えた計算システム。

【請求項 14】

請求項13記載の計算システムにおいて、前記メディア分析エージェントは、

メディア・フレームの画素および／またはイントラ・エンコード・フレームのDCブロックを量子化カラー空間に投影し、前記フレームのカラー・ヒストグラムを生成するカラー・オブジェクト分析部、

を備えた計算システム。

【請求項 15】

請求項13記載の計算システムにおいて、前記メディア分析エージェントは、更に、前記カラー・オブジェクト分析部からの複数のショットに関連する1つ以上の属性を受け取り、前記ショットの2つ以上の間の相関スコアを計算する相関検出部、を備えた計算システム。

【請求項 16】

請求項13記載の計算システムにおいて、前記メディア分析エージェントは、1つ以上のフレームから一次元スライスを抽出し、前記スライスの1つ以上のモーション・パターンを分析してショットの境界を検出する時間スライス分析部、
10
を備えた計算システム。

【請求項 17】

請求項16記載の計算システムにおいて、前記メディア分析エージェントは、更に、前記時間スライス分析部からの複数のショットに関連する1つ以上の属性を受け取り、前記ショットの2つ以上の間の相関スコアを計算する相関検出部、を備えた計算システム。

【請求項 18】

請求項17記載の計算システムにおいて、前記メディア分析エージェントは、更に、前記相関検出部に結合され、シーンに含ませるために全てのショットを統計的に分析するまで、前記シーンを規定する意味的に相関のあるショットを保持する、動的サイズ可変拡張ウィンドウ、
20
を備えた計算システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、一般的に、画像処理に関し、更に特定すれば、メディア・セグメント化システムおよび関連する方法に関する。

【0002】

【従来の技術】

近年における処理、記憶およびネットワーク技術の改良により、多くのパーソナル計算システムがマルチメディア・オブジェクト（例えば、オーディオ、グラフィックおよびビデオ・コンテンツ）を受信し、処理し、レンダリングする能力を有するようになった。このようなマルチメディア・レンダリングに応用した計算パワーの一例は、例えば、ネットワークを通じてリモート・サーバから、適切に構成した計算システムにビデオ・コンテンツを「流入」させ、計算システム上のレンダリングが可能となったことである。レンダリング・システムの多くは、典型的なビデオ・カセット・プレーヤ/レコーダ（VCR）に類似した機能性を備えている。しかしながら、計算パワーの増大により、更に一層進んだ機能に対する消費者の期待が大きくなっている。正にこののような期待の主要な一例として、関連する（即ち、ユーザに特別な関心がある）メディア・コンテンツに迅速にアクセスできることがあげられる。従来技術のシステムは、この期待に沿うことができなかった。
30

【0003】

大量のメディアに対処しアクセスするために、種々の画像データベースおよびビジュアル情報システムが、最近では入手可能となっている。このようなシステムは、医療画像管理、CAD/CAMシステム、犯罪識別システム、クリップアート・ギャラリ等を含む多種多様のアプリケーションにおいて用いられている。従来技術のシステムは、多数の探索技法のいずれかを用いて、関連情報にアクセスし検索することができる。概して、このような従来技術のシステムは、テキストに基づくキーワード手法を利用し、このようなメディア・コンテンツをインデックス化し検索する。このような手法によれば、各フレーム、ショットまたはシーン（各々1つ以上の前者によって構成される）を、データベース・オブジェクトとして格納し、データベース内の各画像（例えば、フレーム、ショット、シーン
40

40

50

)には、当該オブジェクトのテキスト記述が手作業で生成され、関連付けられている。これらのキーワード記述子は、標準的なプール・クエリによって探索することができ、クエリ・テキストの正確なまたは確率論的な照合に基づいて、検索が行われる。

【0004】

【発明が解決しようとする課題】

このような従来技術のシステムは、このような技術に対する欲求を刺激する働きをするものの、真のコンテンツに基づくメディア探索に役立つシステムは従来技術ではなく、したがって特定のメディア・コンテンツに精度高くアクセスし検索するという要望に完全に対応することはできない。テキストのみに基づくシステムには、固有の問題がいくつかある。分類階層を構築するために必要な記述キーワードの自動生成または意味情報の抽出は、現在の視覚および情報技術の能力を超えている。その結果、このような画像のテキスト記述は、手作業で生成せざるを得ない。キーワード記述子の手作業による入力は、煩わしく時間のかかるプロセスであり、不正確さや記述的限界の可能性があることを認めなければならない。更に、テキストおよびパターンのようなあるビジュアル・プロパティは、多くの場合、僅かなテキスト記述子で適切にまたは高精度に記述することは、特に汎用インデックス化および検索アプリケーションでは、不可能ではないにしても困難である。

10

【0005】

コンテンツに基づいてメディアを定性的にセグメント化しようとする他の手法も論じられてきたが、全ては計算に費用がかかり、結果として、ほぼリアル・タイムの消費者用アプリケーションには不適切である。これら従来技術の手法は、典型的に、フレーム間の類似したマテリアルを識別し、ショットの境界を検出する。ショットの境界は、多くの場合、編集点、例えば、カメラ・フェードを示すのであって、意味的な境界ではないことを当業者は認めよう。更に、計算上の複雑さが伴うために、このようなショットは、多くの場合スタティックとして、または編集点以前または以降の固定数のフレームとして定義される（例えば、直前の3フレーム、および後続の3フレーム）。この点について、このような従来技術のシステムは、典型的に、固定のフレーム・ウインドウを利用してショットを定義する。

20

【0006】

対照的に、シーンは意味的に同様なショットから成り、したがって多数のショット境界を含む可能性がある。これに応じて、2つのショット間におけるフレームの視覚的類似性に基づく従来技術の手法は、良い結果が得られないことが多い、シーンを識別しセグメント化するためには、ショット間における意味的相関の定量的尺度が必要となっている。

30

【0007】

したがって、従来技術のシステムに共通して付随する固有の欠点によって煩わされることのないメディア・セグメント化システムおよび関連する方法を提案する。

【0008】

【課題を解決するための手段】

本発明は、意味的レベルにおいて、メディア・コンテンツの迅速なアクセスおよび検索に役立つ、メディア・セグメント化システムおよび関連する方法に関する。本発明の実施態様の一例によれば、メディア・コンテンツを受信し、受信したメディアの連続ショットの1つ以上の属性を分析することから成る方法を提案する。少なくとも部分的に1つ以上の属性に基づいて、連続ショットの各々について相関スコアを生成する。シーンのセグメント化を実行し、意味的に繋がりのあるショットを集合化する。

40

【0009】

【発明の実施の形態】

図面全体を通じて、同じ参照番号を用いて同様のコンポーネントおよび構造を引用することとする。

【0010】

本発明は、コンテンツに基づくシーン・セグメント化システムおよび関連する方法に関する。この点に関して、本発明は、主にテキスト・キーワードを拠り所としていた、従来技

50

術の画像記憶および検索システムに共通して付随していた多数の欠点を克服する。本発明の発明的特質は、ビジュアル・メディア・コンテンツのコンテクストにおいて展開する。しかしながら、本発明はそのように限定される訳ではなく、以下に紹介する革新的なメディア分析エージェントは、ここに記載する発明的概念を利用して、例えば、オーディオ・コンテンツ、グラフィック・コンテンツ等を含む多種多様のマルチメディア・コンテンツのいずれに対しても、コンテンツに基づくメディア・セグメント化を実行することは認められよう。この点について、以下に提示する実施形態例は、本発明の範囲および精神を単に例示するに過ぎない。

【0011】

本発明を説明するにあたり、前述の図面を参照しながら、ネットワーク・アーキテクチャおよび関連する方法の例について説明する。しかしながら、ここに記載するアーキテクチャおよび方法に対する変更も、本発明から逸脱することなく、当然可能であることを注記しておく。実際には、このような代替実施形態は、本発明の範囲および精神の中で予想されることである。

計算システムの一例

図1は、革新的メディア分析エージェント104を含むコンピュータ・システムの一例102を示す。メディア分析エージェント104は、メディア・コンテンツを分析し、ショットの各フレームにおいて1つ以上のオブジェクトを識別し、同様のオブジェクトを含むショットをシーンにセグメント化し、格納して後にコンテンツに基づいてアクセスし検索しようというものである。先に紹介したように、そして以下に続く説明に基づいて認められるように、分析エージェント104は、本発明の精神および範囲から逸脱することなく、コンテンツに基づく探索の目的で、他のタイプのメディアを識別しセグメント化するためにも用いることができる。尚、図1では、分析エージェント104を別個の単体アプリケーションとして図示するが、例えば、メディア・プレーヤ、メディア・ライブラリ、リッパ・アプリケーション(ripper application)等のアプリケーションの関数としても実現可能であることは認められよう。以下の論述から、コンピュータは、あるクラスの汎用または特殊目的計算プラットフォームのいずれをも表わすことを意図しており、革新的分析エージェント104を備えると、先に紹介した第1の実施態様例にしたがって本発明の教示を実現することは明白であろう。尚、図1のコンテクストでは、分析エージェント104をソフトウェア・アプリケーションとして図示しているが、コンピュータ・システム102は代わりにエージェント104のハードウェア実施態様も同様に対応可能であることは認められよう。この点について、分析エージェント104の説明を別にすれば、能力が高いコンピュータあるいは低いコンピュータであっても、本発明の精神および範囲から逸脱することなく、置換することも可能であるので、コンピュータ・システム102についての以下の説明は、単に例示であることを意図している。

【0012】

図示のように、コンピュータ102は、1つ以上のプロセッサ即ち演算装置132、システム・メモリ134、およびシステム・メモリ134からプロセッサ132までを含む種々のシステム・コンポーネントを結合するバス136を含む。

【0013】

バス136は、数種類のバス構造の内1つ以上を表わし、メモリ・バスまたはメモリ・コントローラ、周辺バス、加速グラフィックス・ポート、および種々のバス・アーキテクチャのいずれかを用いるプロセッサまたはローカル・バスを含む。システム・メモリは、リード・オンリ・メモリ(ROM)138、およびランダム・アクセス・メモリ(RAM)140を含む。基本入出力システム142(BIOS)は、起動中のように、コンピュータ102内のエレメント間ににおけるデータ転送を補助する基本的なルーティンを含み、典型的にROM138内に格納されている。コンピュータ102は、更に、図示しない、ハード・ディスクに対する読み書きを行なうハード・ディスク・ドライブ144、リムーバブル磁気ディスク148に対する読み書きを行なう磁気ディスク・ドライブ146、およびCD-ROM、DVD-ROMまたはその他のこののような光媒体に対して読み書きを行

10

20

30

40

50

なう光ディスク・ドライブ150を含む。ハード・ディスク・ドライブ144、磁気ディスク・ドライブ146、および光ディスク・ドライブ150は、SCSIインターフェース154またはその他の何らかの適当なバス・インターフェースによって、バス136に接続されている。ドライブおよびこれらに伴うコンピュータ読み取り可能媒体は、コンピュータ読み取り可能命令、データ構造、プログラム・モジュール、およびコンピュータ102のその他のデータの不揮発性格納を行なう。

【0014】

ここに記載する動作環境の一例は、ハード・ディスク144、リムーバブル磁気ディスク148、およびリムーバブル光ディスク152を用いているが、他のタイプのコンピュータ読み取り可能媒体で、コンピュータによるアクセスが可能なデータを格納可能な、磁気カセット、フラッシュ・メモリ・カード、ディジタル・ビデオ・ディスク、ランダム・アクセス・メモリ(RAM)、リード・オンリ・メモリ(ROM)等も、動作環境の一例において使用可能であることは、当業者には認められよう。

【0015】

ハード・ディスク144、磁気ディスク148、光ディスク152、ROM138、またはRAM140上には、オペレーティング・システム158、例えば、本発明の教示を組み込んだ分析エージェント104を含む1つ以上のアプリケーション・プログラム160、その他のプログラム・モジュール162、およびプログラム・データ164(例えば、得られた言語モデル・データ構造等)を含む多数のプログラム・モジュールを格納することができる。ユーザは、キーボード166およびポインティング・デバイス168のような入力デバイスによって、コマンドおよび情報をコンピュータ102に入力することができる。他の入力デバイス(図示せず)には、マイクロフォン、ジョイスティック、ゲーム・パッド、衛星ディッシュ、スキャナ等を含むことができる。これらおよびその他の入力デバイスは、バス136に結合されているインターフェース170を介して、演算装置132に接続されている。モニタ172またはその他の形式の表示装置も、ビデオ・アダプタ174のようなインターフェースを介して、バス136に接続されている。モニタ172に加えて、パーソナル・コンピュータは多くの場合、スピーカやプリンタのような、その他の出力周辺装置(図示せず)も含む。

【0016】

図示のように、コンピュータ102は、リモート・コンピュータ176のような1つ以上のリモート・コンピュータへの論理接続を用いて、ネットワーク環境において動作することも可能である。リモート・コンピュータ176は、別のパーソナル・コンピュータ、パーソナル・ディジタル・アシスタント、サーバ、ルータまたはその他のネットワーク・デバイス、ネットワーク「シン・クライアント」PC(network thin-client PC)、ピア・デバイス、あるいは共通のネットワーク・ノードとすることもでき、典型的に、コンピュータ102に関して先に述べたエレメントの多くまたは全てを含むが、図1にはメモリ記憶素子178のみを図示している。この点に関して、革新的分析エージェント104は、例えば、計算システム176のようなリモート計算システムによって呼び出し、利用することも可能である。

【0017】

図示のように、図1に示す論理接続は、ローカル・エリア・ネットワーク(LAN)180およびワイド・エリア・ネットワーク(WAN)182を含む。このようなネットワーク環境は、事務所、企業規模のコンピュータ・ネットワーク、インターネットおよびインターネットにおいては、一般的である。一実施形態では、リモート・コンピュータ176は、ワシントン州、RedmondのMicrosoft Corporationが製造し販売する「Internet Explorer」ウェブ・ブラウザのような、インターネット・ウェブ・ブラウザ・プログラムを実行する。

【0018】

LANネットワーク環境で用いる場合、コンピュータ102は、ネットワーク・インターフェースまたはアダプタ184を介してローカル・ネットワーク180に接続する。WA

10

20

30

40

50

Nネットワーク環境で用いる場合、コンピュータ102は、典型的に、モデム186、またはインターネットのようなWAN182を通じて通信を確立するその他の手段を含む。モデム186は、内蔵でも外付けでもよく、入出力(I/O)インターフェース156を介してバス136に接続することができる。ネットワークの接続性に加えて、パーソナル・コンピュータ102またはその一部に関係して図示したプログラム・モジュールを、リモート・メモリ記憶装置に格納することもできる。尚、図示のネットワーク接続は、一例であり、コンピュータ間に通信リンクを確率するその他の手段も使用可能であることは認められよう。

【0019】

一般に、コンピュータ102のデータ・プロセッサは、異なる時点においてコンピュータの種々のコンピュータ読み取り可能記憶媒体に格納されている命令によってプログラムされている。プログラムおよびオペレーティング・システムは、典型的に、例えば、フロッピ・ディスクまたはCD-ROM上で配付される。そこから、これらをコンピュータの二次メモリにロードする。実行時に、これらを少なくとも部分的にコンピュータの一次電子メモリにロードする。ここに記載する発明は、これらおよびその他の種々のタイプのコンピュータ読み取り可能記憶媒体が、マイクロプロセッサまたはその他のデータ・プロセッサと共に以下で説明する革新的なステップを実施する命令またはプログラムを含む場合、このようなメディアを含む。また、本発明は、以下に説明する方法および技術にしたがってプログラムされる場合、コンピュータ自身も含む。更に、コンピュータのあるサブコンポーネントは、以下に説明する機能およびステップを実行するようにプログラムすることができます。本発明は、このようなサブコンポーネントを記載のようにプログラムする場合、これらを含む。加えて、ここに記載する発明は、種々のタイプのメモリ媒体上に具体化する、データ構造を含む。これについては以下で説明する。

10

20

【0020】

例示の目的上、オペレーティング・システムのようなプログラムおよびその他の実行可能プログラム・コンポーネントは、ここでは離散ブロックとして図示するが、このようなプログラムおよびコンポーネントは、種々の時点にコンピュータの異なる記憶コンポーネント内に位置し、コンピュータのデータ・プロセッサが実行することは認められよう。

メディア分析エージェントの一例

図2は、本発明の一実施態様にしたがって、本発明の教示を組み込んだメディア分析エージェントの一例104のブロック図を示す。図2に示す実施形態の一例によれば、メディア分析エージェント104は、1つ以上のコントローラ202、メディア分析エンジン204、メモリ/記憶装置206、入出力インターフェース、およびオプションとして、図示のように各々通信状態に結合した1つ以上のアプリケーション210を含むように図示されている。先に紹介したように、メディア分析エージェント104は、受信したメディア・フレームのコンテンツを分析し、少なくとも部分的に、受信したメディアのコンテンツの1つ以上の属性に基づいて、メディアを異種のシーンにセグメント化する。

30

【0021】

本発明の実施形態の一例によれば、メディア分析エージェント104は、選択的に、1つ以上のカラー・オブジェクト分析部212、時間スライス分析機能214、および相関検出部216を呼び出し、メディア・シーンの識別およびセグメント化を行なう。一実施形態によれば、カラー・オブジェクト分析部212または時間スライス分析機能のいずれかを、相関検出部216と組み合わせて呼び出し、ショット間の意味的類似性を識別し、シーンの検出およびセグメント化に役立てる。図示のように、カラー・オブジェクト分析部は、カラー空間量子化部218を含む。時間スライス分析機能214は、モーション・パターン分析およびキー・フレーム抽出機能220を備えるように示されている。これについては以下で詳細に説明する。図示のように、ストレージ/メモリ206は、1つ以上の受信したメディア・コンテンツ224、拡張シーン・ウインドウ・データ構造226および/または識別したシーン・データ構造228のためにメモリを含む。以下で更に詳細に論ずるが、メディア分析エンジン204は、受信したメディアの1つ以上の属性(例えば

40

50

、カラー、テクスチャ、空間・時間情報等)を分析し、ショット間の意味的類似性を識別する。少なくとも部分的に意味的類似性の分析に基づいて、メディア分析エンジン204は、受信したメディアをシーンにセグメント化し、後にコンテンツに基づいてアクセスおよび検索が行なえるようにする。

【0022】

尚、多数の異種機能ブロックとして図示したが、1つ以上のエレメント202～228を1つ以上のブロックに組み込んでもよいことは認められよう。同様に、分析エージェント104は、図よりも少ない機能ブロック、即ち、オブジェクト識別機能212または時間スライス分析機能214の一方のみを用いても実施可能であり、本発明の精神および範囲から逸脱する訳ではない。

10

【0023】

一実施態様によれば、コントローラ202は、例えば、ローカル・メモリ・ストレージ(206)、リモート・メディア・プロバイダおよび/またはコンテンツ・ストレージ・ソースを含み、ネットワークを通じてメディア分析エージェント104に結合されている多数のソースのいずれからも、メディア・コンテンツを受信する(例えば、図7参照)。一実施態様によれば、メディア・コンテンツは、コントローラ202によってリモート・ソースから受信し、分析のためにストレージ/メモリ224に置かれる。一実施態様によれば、メディアはホスト・コンピュータ102によって圧縮形態で受信され、メディア分析エージェント104に提示する前に伸長する。代替実施態様では、コントローラ202は選択的に、メディア分析エージェント104内に常駐するデコーダ・アプリケーション(例えば、210)を呼び出し、圧縮形態で受信したメディアをデコードしてから、メディア分析エージェント204を選択的に呼び出す。先に述べた本発明の革新的な面を別として、コントローラ202は、例えば、プロセッサ、特定用途集積回路(A S I C)、現場プログラム可能ゲート・アレイ(F P G A)、プログラマブル論理アレイ(P L A)等のような、当技術分野では公知の多種多様の制御ロジックのいずれをも代表することを意図している。更に、コントローラ202は、(例えば、演算装置132によって)実行すると、前述の制御機能を実現する複数の実行可能命令として、ソフトウェアで実現してもよいことは認められて当然である。

20

【0024】

メディア分析エンジン204は、コントローラ202によって選択的に呼び出され、受信したメディアのショット間における意味的相関を識別し、精度高くメディアをシーンにセグメント化する。せいぜいフレーム間の視覚的類似性を識別しショット境界を識別する程度の従来技術のシステムとは異なり、メディア分析エンジン204は、選択的に1つ以上の意味、コンテキスト分析部212および/または214を呼び出し、連続するショットにおけるフレームの1つ以上の定性的属性を定量化する。以下で更に詳しく論ずるが、これらの属性の定量化を相關検出部216に提出し、ショットが意味的に相関付けられているか、即ち、類似しているか否か判定を行なう。類似性のスレシホールドを満たす場合、ショットを少なくともシーンの部分集合として集合化する。

30

【0025】

図2に示す例によれば、メディア分析エンジン204は、2つの意味コンテキスト・メディア分析部、カラー・オブジェクト分析部212、および時間スライス分析機能214を備えるように示されている。以下の説明から認められるように、1つ以上のメディア分析部を呼び出してシーンのセグメント化を行なうことができる。したがって、意味的分析能力が高いまたは低いメディア分析エンジンを、図示の実施形態と置換してもよく、本発明の精神および範囲から逸脱する訳ではない。

40

カラー・オブジェクト分析部

図示のように、カラー・オブジェクト分析部212は、カラー量子化部218を含む。視覚的類似性を測定する従来技術の手法とは異なり、カラー・オブジェクト分析部212は、カラー量子化部を呼び出して、連続するショット内における1つ以上のフレームの1つ以上の優勢なオブジェクトおよび/または環境オブジェクトについて、カラー空間におけ

50

るカラー・ヒストグラムを計算する。一実施態様によれば、色相、彩度、および値、即ち、「HSV」カラー空間をカラー量子化に用い、ヒストグラムを計算する。HSVカラー空間は、他のカラー空間よりもこの分析に対して多数の利点をもたらす。なぜなら、これは自然で近似的に知覚が均一であり、HSVカラー空間の量子化によって、コンパクトであるが完全なカラーの集合体を生成するからである。一実施態様によれば、HSVカラー空間は、カラー量子化部218によって、三次元(3D)デカルト座標系で量子化され、それぞれXおよびYに10個の値、Z(明度)に5個の値が与えられる。当業者は、カラー(即ち、XおよびY座標)を示すために10個の値を用いることにより、カラー量子化部218は、HSV空間において僅かに異なるカラー間でさえも区別し、照明条件が変化しても、より多くのオブジェクトを判別することが可能となることを認めよう。

10

【0026】

ビデオ・ショットの優勢なカラー・オブジェクトを判定するために、各フレームの画素、および/またはショットのイントラ・エンコード(I)フレームにおけるDCブロックを、カラー量子化部218によって、量子化カラー空間(例えば、HSVカラー空間)に投影する。3Dカラー空間におけるこれらの画素の正規分布は、フレームの3Dカラー・ヒストグラムを形成する。3Dカラー・ヒストグラムにおける全ての優勢局部最大点を識別し、カラー空間における狭い近傍(例えば、3つ量子化単位の直径を有する)内にて、各局部最大点周囲の球をカラー・オブジェクトとして、カラー量子化部218によって定義する(例えば、図3参照)。一旦カラー・オブジェクトを識別したなら、カラー・オブジェクト分析部212は、画素情報が最も多い1つ以上のオブジェクトを識別する。これらのオブジェクトは、フレームに対して最も有意なカラー情報を含むものとして識別され、ノイズに対して立ち直りが早い(resilient)。一実施態様によれば、カラー・オブジェクト分析部212は、上位20個のオブジェクトを優勢オブジェクトとして選択する。

20

【0027】

次に、カラー・オブジェクト分析部212は、優勢カラー・オブジェクト内に含まれる画素のみをカウントすることによって、各フレーム毎に3D優勢カラー・ヒストグラムを生成する。一実施態様によれば、優勢カラー・ヒストグラムは、次のように示される。

【0028】

【数1】

$$hist_d(k, x, y, z) \quad (1)$$

30

ここで、kはフレーム番号を示し、x, y, zはカラー・ビンを示す。尚、前述のことから、カラー・オブジェクト分析部212は空間ドメインではオブジェクトのセグメント化を行なわないが、オブジェクトのカラー空間における優勢領域に当たる画素を考慮することが認められよう。

【0029】

一旦優勢オブジェクトを識別したなら、カラー・オブジェクト分析部212は、フレームを跨がってカラー空間内のオブジェクトを追跡し、ショットの優勢オブジェクトを識別する。2つの連続するフレームにおいて2つのカラー・オブジェクトの中心が十分に近い場合、これら2つのカラー・オブジェクトを同じカラー・オブジェクトとして認識する。このようなカラー追跡プロセスは、ショット内のフレーム全てを追跡するまで続ける。追跡の後、ショットにおいて最も長い持続時間有するカラー・オブジェクトのみを優勢オブジェクトとして保持する。一実施形態によれば、カラー・オブジェクト分析部212は、各ショット毎に、総合的優勢カラー・ヒストグラムを形成する。これは、以下のように表わされる。

40

【0030】

【数2】

$$hist_d^a(x, y, z) \quad (2)$$

ここで、aはショットを示す。総合的優勢カラー・ヒストグラムは、ショット全体における

50

るフレームに共通な優勢カラー・オブジェクトから成る。一実施態様によれば、カラー・オブジェクト分析部 212 は、ショットにおいて持続時間が長い方のカラー・オブジェクトに重み値を適用する。数学的には、次のように表わされる。

【0031】

【数3】

$$hist_d^A(x, y, z,) = hist_d^a(x, y, z) \times \frac{d_1}{d_o} \quad (3)$$

ここで、 d_o はショットの持続時間、 d_1 はカラー (x, y, z) を有する優勢カラー・オブジェクトの持続時間である。一実施態様によれば、カラー・オブジェクト分析部 212 は、更にヒストグラム (3) を詳細化するにあたり、ショット内の各優勢カラー・オブジェクトの平均サイズを正規化する。したがって、ショットの優勢カラー・ヒストグラムは、フレーム内における構造的コンテンツ、およびショット内における時間的コンテンツ双方を表わす。また、これらの優勢カラー・オブジェクトは、多くの場合ショットにおける優勢オブジェクトまたは背景を表わし、2つのショットにおけるこれらカラー・オブジェクト間の相関は、2つのショット間の相関の優れた指標となる。

【0032】

一旦図3に移り、本発明の一態様にしたがって、識別したオブジェクトを収容した、一例のHSVカラー空間の模式図を提示する。図示のように、2つのカラー空間円筒が図示されており、各々、例えば、フレーム (N) 302 およびフレーム ($N + 1$) 304 を表わす。HSVカラー・ヒストグラム 302 および 304 は、関連するフレーム (302, 304) 内における優勢カラー・オブジェクト (例えば、306A...N, 308A...N) を識別し、ショット内における優勢カラー・オブジェクトを識別するために用いられる。続けて、このようなHSVカラー・ヒストグラムを利用して、優勢カラー・オブジェクトを識別し、シーンのセグメント化を行なう。

【0033】

引き続き図2を参照すると、一旦カラー・オブジェクト分析部 212 によって優勢カラー・オブジェクト・ヒストグラムを生成したなら、コントローラ 202 は選択的に相関検出部 216 のインスタンスを呼び出し、2つのショット a および b 間の相関スコアを求める。相関検出部 216 は、多数の統計技法のいずれも用いても、ショット間の相関を識別することができる。例示の一実施態様によれば、相関検出部 216 は、ショット a および B 間の相関を計算する際に、2つのショットの2つの優勢カラー・ヒストグラム間のヒストグラムの交点を計算する。数学的に表わすと次の通りであり、

【0034】

【数4】

$$Cor(a, b) = \sum x \sum y \sum z \min[hist_d^A(x, y, z), hist_d^B(x, y, z)] \quad (4)$$

次の特性を有する。

【0035】

- 1) 0 cor (a, b) 1, cor (a, a) = 1
- 2) cor (a, b) = cor (b, a)

ショット集合化のための拡張ウインドウ方式

少なくとも部分的に、相関検出部 261 によって行なった相関分析に基づいて、コントローラ 202 は、ショットを集合化し、メディアをシーンにセグメント化する。例示の一実施態様によれば、コントローラ 202 は、最小相関スレシホールド (T_c) を満たすショットを集合化する。

【0036】

本発明の一態様によれば、コントローラ 202 は、メモリ 206 の拡張ウインドウ 218 を利用して、少なくとも部分的に相関検出部 216 によって求めた相関スコアに基づいて

10

20

30

40

50

、相関付けられた連続ショットを1つのシーンに集合化する。以下の論述に基づいて、拡張ウィンドウ技術によって、多くのショット対を比較したり、複雑なショット・リンクを構成する必要性をなくすことによって、実施態様の計算上の複雑さを低下させることは認められよう。

【0037】

むしろ、時間的制約、即ち、時間的に接近しているショットは同じシーンに属する可能性が高いことを考慮すると、2つのショット間の相関スコアを、次の時間近接性（または牽引）係数で重み付けする。

【0038】

【数5】

$$w=1/(1+d/C) \quad (5)$$

ここで、dは2つのショット間の最短距離、Cは少なくとも部分的に平均ショット長によって決定した定数である。一実施態様によれば、コントローラ202は、あらゆるシーンが少なくとも3つのショットを含むと仮定している。最初に、最初の3つのショットが新たなシーンを形成し、拡張ウィンドウのサイズを3にセットする。新たなショットを分析する毎に、その相関スコアをウィンドウ内の最後の3つのショットと比較し、3つの相関スコア間の最大値を決定する。次に、計算した最大値が平均最大相関スコアから、拡張ウィンドウ内に含まれるショットのあらゆる変動を減じた値よりも大きい場合、このショットを拡張ウィンドウ内の現シーンに吸収する。数学的には、コントローラ202によって行われる比較は次のように表わされる。

【0039】

【数6】

$$v > mean-var \quad (6)$$

最大値(v)がこのスレシホールドを満たさない場合、続くショットをもう少し多く分析し、現ショットが新たなシーンの開始を表わすことの信頼性を高める。前述のシーンに含まれるという要件を満たさない1ショットがシーンの中にある場合が多いことが経験的にわかっている。しかしながら、続くショットの分析を追加することによって、現ショットが意味的シーンの終端を表わすのではないことを確認することができる。コントローラ202が、1つ以上の続くショットが基準(6)を満たすと判定した場合、拡張ウィンドウ218内で展開したシーン内に、以前のシーンをいずれでも含ませることができる。数学的には、コントローラ202は、現ショットiの牽引比(attraction ratio)を、以下のように、新たなシーンに向けて形成する。

【0040】

【数7】

$$R(i) = (right(i) + right(i+1)) / (left(i) + left(i+1)) \quad (7)$$

ここで、 $R(i) > T$ および $R(i) > R(i-1)$ および $R(i) > R(i+1)$ の場合、

$left(i) = max[cor(i, i-1), cor(i, i-2), cor(i, i-3)]$

$left(i+1) = max[cor(i+1, i-1), cor(i+1, i-2)]$

$right(i) = max[cor(i, i+1), cor(i, i+2), cor(i, i+3)]$

$right(i+1) = max[cor(i+1, i+2), cor(i+1, i+3), cor(i+1, i+4)]$

Tはスレシホールドである。

【0041】

一実施態様によれば、コントローラ202はスレシホールドTを1.5にセットする。したがって、右側からのショットiへの牽引が左側からよりも大きい場合、現ショットiは新たなシーンを開始すると判断する。それ以外の場合、コントローラ202は、ショットiを拡張ウィンドウの現シーンに置く。拡張ウィンドウについては、図4および図5を参照

10

20

30

40

50

しながら、以下で図に基づいて説明する。

【0042】

図4に移り、少なくとも部分的に前述の相関測定に基づく拡張シーン・ウィンドウにショットを追加する図式表現を簡潔に示す。図示のように、コントローラ202は、ショットの相関値を以前のショットおよび連続するショットと比較し、ショットの各々をどのシーンに置くべきか特定する。

【0043】

図5は、本発明の一態様による、拡張シーン・ウィンドウ218を含むストレージ/メモリ206を図で示す。先に紹介したように、拡張ウィンドウ218は、相関測定による規定にしたがって、同様の意味的内容を有するショットを集合化するために用いられる。

10

空間・時間分析機能

カラー・オブジェクト分析部212に加えて、あるいはこれの代わりに、コントローラ202は、空間・時間スライス分析機能214のインスタンスを選択的に呼び出して、離散シーンを識別したショット間の意味的類似性を識別することもできる。図示のように、時間スライス分析機能214は、モーション分析機能220および空間分析機能222を含む。以下で更に詳しく説明するが、時間スライス分析機能214は、連続ショットにおけるフレームから、1つ以上の水平および垂直一次元(1D)スライスを切り出し、スライスのモーション・パターンを定量化し、キー・フレームを選択して、ショットの各モーション・セグメントを表わす。少なくとも部分的に1Dショットからの定量化した特徴に基づいて、コントローラ202は、相関検出部216のインスタンスを呼び出し、隣接するショットからの定量化した特徴間の類似性を測定し、シーンの境界を識別する。この場合も、従来技術とは異なり、時間スライス分析機能は、ショット間の意味的繋がり(cohesion)を識別し、メディア・シーンを識別する。

20

【0044】

一実施形態によれば、時間スライス分析機能214は、受信したメディアの垂直および水平面から1Dスライスを切り出し、モーション分析機能220のインスタンスを選択的に呼び出す。モーション分析機能220は、各ショットのスライスを、更に小さなセグメントに区分する。各セグメントは、一致したモーション・パターンを有する。二次元空間・時間スライスにおいて、時間テクスチャは、モーションの軌跡を示す情報を含む。一実施態様によれば、従来のテクスチャ分析方法は、例えば、同時発生マトリクス計算のように用いられ、ショットにおけるモーション・パターンを特徴化する。一実施態様によれば、15個の同時発生マトリクスを計算し、3つの異なる方向で5つのシーンに跨がる軌跡をモデル化しつつ、各マトリクスの平滑性および明暗を表わす30個の特徴を抽出する。

30

15個の同時発生マトリクスの一例を、図6を参照しながら示す。

【0045】

一実施態様によれば、モーション分析機能220は、以下の4つのタイプの1つにしたがって、少なくとも部分的に基礎モーションに基づいて、各セグメントのモーションを特徴付ける。1)無モーションまたは僅かなモーション、2)ズーム・インまたはズーム・アウト、3)パン、および4)ティルト。少なくとも部分的に、一致するモーション・パターンの各セグメントに対するモーション・パターンに基づいて、以下の規則にしたがって1つ以上のキー・フレームを選択する。

40

【0046】

1)無モーションまたは僅かなモーション：任意のフレームを選択しインデックス化する。

2)ズーム・インまたはズーム・アウト：最初および最後のフレームを選択しインデックス化する。

【0047】

3)パン：対応する垂直スライスを選択しインデックス化する。

4)ティルト：対応する水平スライスを選択しインデックス化する。

一旦インデックス化すると、各ショットは、少なくとも部分的に前述のモーション分析に

50

基づいて抽出した1つ以上のキー・フレーム集合の特徴によって表わされる。キー・フレームの特徴は、カラー・ヒストグラム、またはその他の画像の特徴とすることができる。少なくとも部分的にショットのキー・フレームの特徴に基づいて、相関検出部はショットの類似性尺度を計算し、ショットが意味的に関係するか否か判定を行い、関係がある場合、コントローラ202はショットをシーンにセグメント化し、少なくとも一時的にメモリ228に格納する。一実施形態では、コントローラ202は、ショット間の類似性を計算する際、2つのショットのキー・フレーム間のヒストグラム交点を識別する。また、先に紹介したように、一実施形態によれば、メディア分析エージェント104は、拡張ウィンドウを利用して、意味的類似性のスレシホールドを満たすショットから、動的にシーンを生成する。

10

【0048】

ここで用いる場合、ストレージ/メモリ206および入出力インターフェース208は、各々、これらが当技術分野において周知の通りのエレメントを表わすことを意図している。ストレージ/メモリ206は、メディア分析エージェント104が、少なくとも一時的に、メディア・コンテンツ224、拡張シーン・ウィンドウ226および/または識別したシーン228を維持するために利用する。I/Oインターフェース208は、メディア分析エージェント104が、外部エレメントおよびシステムと通信を行うことによって、容易に分散アーキテクチャおよびリモート動作を行なえるようになる。

20

【0049】

アプリケーション210は、シーンのセグメント化のために意味的に繋がりのあるショットを自動的に識別するためにメディア分析エンジン204を用いることができる、あるいはメディア分析エンジン204によって用いることができる多種多様のアプリケーションを含むことを意図している。この点について、アプリエーション210は、グラフィカル・ユーザ・インターフェース(GUI)、メディア・プレーヤ、メディア生成器、メディア・データベース・コントローラ等も含むことができる。

20

【0050】

前述のことから、メディア分析エージェントは、多数の代替実施形態においても実現できることが認められよう。一実施態様によれば、メディア分析エージェント104は、単体アプリケーションとして、例えば、メディア・デコーダ・アプリケーション、メディア・レンダリング・アプリケーション、ブラウザ・アプリケーション、メディア・プレーヤ・アプリケーション等のような上位マルチメディア・アプリケーションの部分集合としてソフトウェアで実現する。あるいは、メディア分析エージェント104は、例えば、用途特定集積回路(ASIC)、コントローラ、プログラマブル・ロジック・デバイス(PLD)、マルチメディア・アクセレレータ周辺機器等において、ハードウェアで実現することも可能である。このような代替実施態様は、本発明の範囲および精神に該当するものとする。

30

動作および実施態様の一例

以上、図1ないし図6を参照しながら、メディア分析エージェント104の動作環境および機能エレメントについて紹介したが、これより図7ないし図10を参照しながら、システムの動作について以下に更に詳しく展開していく。以下では、メディア分析エージェント104の動作は、意味的にビデオ・メディアをセグメント化するというコンテクストで進めていくが、これは図示を容易にするためであり、限定のためではない。しかしながら、メディア分析エージェント104は、例えば、オーディオ・コンテンツのような他のタイプのメディアでも意味的にセグメント化するように拡張可能であることを、当業者は認めよう。

40

【0051】

図7は、本発明の一実施形態にしたがって、動的にメディアを意味的に類似するユニットにセグメント化する方法のフロー・チャートおよび一例を示す。更に具体的には、図示する実施態様の一例では、図7は、少なくとも部分的にシーンを構成するショット間の意味的類似性に基づいて、ビデオ・コンテンツを動的にシーンにセグメント化する方法の一例

50

を提示する。

【0052】

図示のように、図7の方法は、開始すると、ブロック702において、メディア・コンテンツをセグメント化する指示を受ける。更に具体的には、メディア分析エージェント104のコントローラ202は、ローカル・アプリケーション（例えば、210）または外部ソースから、即ち、I/Oインターフェース208を介して、指示を受ける。

【0053】

応答して、メディア分析エージェント104は、ブロック704において、メディア分析エンジン204のインスタンスを呼び出し、識別したメディア・コンテンツを分析し、メディアを構成するショット間の意味的類似性を識別する。先に紹介したように、メディア分析エンジン204は、選択的にカラー・オブジェクト分析部212を呼び出してカラー・オブジェクトのセグメント化を行なうか、時間スライス分析機能214を呼び出してメディア・コンテンツの時間スライス分析を行なう。少なくとも部分的にこのような分析に基づいて、相關検出部216を呼び出し、統計的に意味的に繋がりがあるショットを識別する。
10

【0054】

ブロック706において、統計的に意味的に関係があると判断されたショットを共に集合化し、意味的に関係するメディア・コンテンツのシーンを形成する。

先に紹介したように、あるショットが以前および／または後続のショットに意味的に関係があると一旦相關検出部216が判断したなら、シーンを定義するショットの拡張ウィンドウ（218）にこのショットを追加する。拡張ウィンドウ218の利用によって、従来技術では共通して付随していた厄介な計算上の複雑性から、メディア分析エージェント104を解放する。
20

【0055】

図8は、本発明の一態様によるカラー・オブジェクト・セグメント化方法の一例のフロー・チャートを示す。図示した実施形態の一例において、この方法はブロック802から開始し、HSVカラー空間においてメディア・コンテンツを分析する。即ち、フレームからのコンテンツを、HSVカラー空間においてカラー量子化部218によって量子化する。

【0056】

ブロック804において、HSVカラー空間において、優勢オブジェクトを識別し、フレームおよびショット全域にわたって追跡する。更に具体的には、先に紹介したように、コントローラ202がHSVカラー空間内のオブジェクトを識別し、フレームの境界を超えてこのようなオブジェクトを追跡する。フレーム間のオブジェクトの位置が殆ど動かない場合、これは類似意味構造の指標である。
30

【0057】

ブロック806において、HSVカラー空間内における優勢カラー・オブジェクトに関する情報を相關検出部216に送る。相關検出部216は、少なくとも部分的に連続ショット内の優勢カラー・オブジェクトに基づいて、意味的類似性の尺度を求める。ブロック808において、統計的に他のショットと意味的に類似するショットを、ショットの拡張ウィンドウに集合化する。一旦意味的に類似するショットを全て識別したなら（したがって、拡張ウィンドウ内に保持したなら）、ショットをシーンとして定義し、後のアクセスおよび検索のために格納する。
40

【0058】

図9は、本発明の位置態様による時間スライス分析方法の一例のフロー・チャートを示す。先に紹介したように、メディア分析エンジン204は、選択的に、カラー・オブジェクト分析部212の代わりとして、またはこれに加えて、時間スライス分析機能214を呼び出し、意味的に類似するショットを識別し、シーンとしてセグメント化することができる。カラー・オブジェクト分析部212とは異なり、時間スライス分析機能214は、受信したメディア・コンテンツのモーションおよび空間・時間テクスチャ属性を分析し、シーンをセグメント化する。
50

【 0 0 5 9 】

このように、図9に示す実施形態の一例によれば、この方法を開始すると、ブロック902において、1つ以上の連続ショットの1つ以上のフレームから一次元水平および垂直スライスを抽出する。モーション分析機能220は、ブロック904において、少なくとも部分的にセグメントのモーション属性に基づいてスライスを更に小さなセグメントに繰り返し区分する。

【 0 0 6 0 】

ブロック906において、コントローラ202は、時間分析関数222を選択的に呼び出し、モーション・パターン分析に基づいてショットのキー・フレームを抽出し、これらキー・フレームの特徴を抽出して、ショットのビジュアル・コンテンツを表わす。図示した実施形態の一例によれば、先に紹介したように、時間分析機能222はキー・フレームの1つ以上のモーション、カラーおよび/または時間テクスチャ属性を抽出し、ショットのビジュアル・コンテンツを表わす。

10

【 0 0 6 1 】

ブロック908において、識別したショットのキー・フレームの特徴を相關検出部216に供給する。相關検出部216は、少なくとも部分的にこれらの特徴に基づいて、ショット間の意味的類似性の統計的尺度を求める。前述のように、統計的に類似した意味のコンテクストを有するショットを共に集合化し、シーンを形成する。前述のように、コントローラ202は、拡張ウィンドウ218を用いて、シーン・セグメント化においてショットを集合化することも可能である。

20

代替実施形態

図10は、本発明の更に別の実施形態にしたがって、本発明の教示を実現する命令を含む複数の命令を格納した記憶媒体のブロック図である。概略的に、図10は、複数の実行可能命令を格納し、少なくともその一部が、実行すると、本発明のメディア分析エージェント104を実現する部分集合を含む、記憶媒体/素子1000を示す。

【 0 0 6 2 】

ここで用いる場合、記憶媒体1000は、例えば、揮発性メモリ素子、不揮発性メモリ素子、磁気記憶媒体、光記憶媒体等のように、当業者には公知の多数の記憶素子および/または記憶媒体のいずれをも表わすことを意図している。同様に、実行可能命令は、例えば、C++、ビジュアル・ベーシック、ハイパーテキスト・マークアップ言語(HTML)、Java(登録商標)、extensible Markup Language(XML)等のように、当技術分野では公知の多数のソフトウェア言語のいずれをも反映することを意図している。更に、記憶媒体/素子1000は、いずれのホスト・システムと一緒に位置する必要はないことも認められよう。即ち、記憶媒体/素子1000は、実行システムに通信状態に結合されこれによるアクセスが可能なりモート・サーバ内に常駐してもよい。したがって、本発明の精神および範囲内において代替記憶媒体およびソフトウェアの実施形態が考えられるので、図10のソフトウェア実施態様は、例示として見なすこととする。

30

【 0 0 6 3 】

以上、構造的特徴および/または方法論的ステップに特定的な言語で本発明を説明したが、添付した特許請求の範囲に規定する本発明は、必ずしも記載した具体的な特徴またはステップに限定される訳ではないことは理解されよう。例えば、ここに提示した発明的概念は、複数のオーディオ・コンテンツを収容した記憶媒体(例えば、音楽CD)上で、別個のオーディオ・コンテンツ(例えば、歌)を識別するためにも用いることができる。この代替実施態様によれば、メディア分析エージェント104のアプリケーション210は、記憶媒体上のオーディオ・コンテンツのカラー表現を生成する。例えば、スペクトル分析等のように、多数の教示のいずれでも、このオーディオ・ビジュアル変換を実行するためには用いることができる。一旦オーディオ・ビジュアル変換が完了したなら、メディア分析エージェント104は、カラー・オブジェクト分析部212、カラー量子化部218および相関検出部216を選択的に呼び出し、前述した本発明の教示にしたがって、複数のオ

40

50

ーディオ・コンテンツから意味的に別個のオーディオ・コンテンツを識別する。したがって、前述の具体的な特徴およびステップは、ここに紹介した広義の発明的概念の実施態様の一例として開示したに過ぎないことは認められよう。

【図面の簡単な説明】

【図1】本発明の教示を組み込んだ計算システムの一例のブロック図である。

【図2】本発明の実施形態の一例にしたがって、コンテンツに基づくシーン・セグメント化を実行するメディア分析エージェントの一例のブロック図である。

【図3】本発明の一態様による、カラー・オブジェクトのセグメント化および追跡を示す模式図である。

【図4】本発明の一態様による、拡張ウィンドウ・ショット集合化技法を示す模式図である。10

【図5】本発明の一態様による拡張シーン・ウィンドウを備えたデータ構造の模式図である。

【図6】本発明の一態様による時間的スライス分析に用いる同時発生マトリクスの模式図である。

【図7】本発明の実施形態の一例による、コンテンツに基づくシーン・セグメント化方法の一例のフロー・チャートである。

【図8】本発明の一態様にしたがって、メディア・コンテンツのショット間における意味的類似性を識別する、カラー・オブジェクト・セグメント化方法の一例のフロー・チャートである。20

【図9】本発明の一態様にしたがって、メディア・コンテンツのショット間における意味的類似性を識別する、時間スライス分析方法の一例のフロー・チャートである。

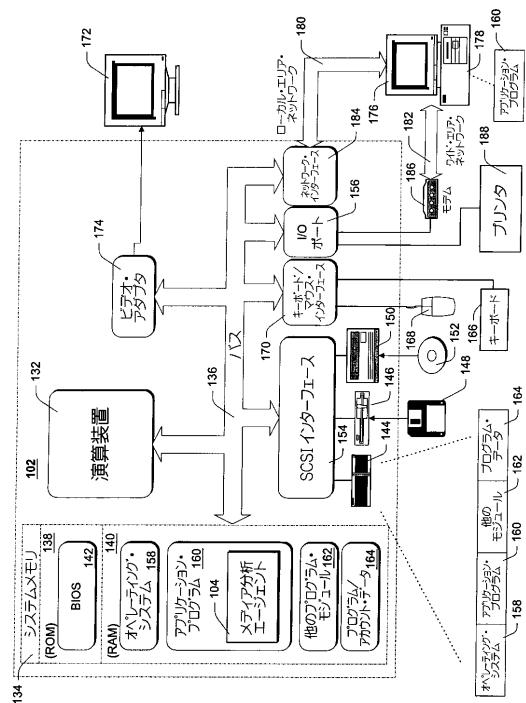
【図10】複数の実行可能命令を格納し、少なくともその部分集合が、本発明の教示を組み込んだメディア分析エージェントを実現する、記憶媒体の一例のブロック図である。

【符号の説明】

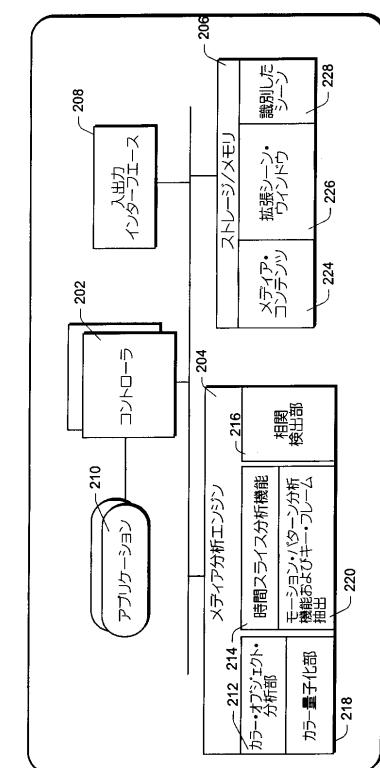
1 0 2	コンピュータ・システム	
1 0 4	革新的メディア分析エージェント	
1 3 2	演算装置	
1 3 4	システム・メモリ	
1 3 6	バス	30
1 3 8	リード・オンリ・メモリ(ROM)	
1 4 0	ランダム・アクセス・メモリ(RAM)	
1 4 2	基本入出力システム(BIOS)	
1 4 4	ハード・ディスク・ドライブ	
1 4 6	磁気ディスク・ドライブ	
1 4 8	リムーバブル磁気ディスク	
1 5 0	光ディスク・ドライブ	
1 5 4	SCSIインターフェース	
1 5 6	入出力(I/O)インターフェース	
1 5 8	オペレーティング・システム	40
1 6 0	アプリケーション・プログラム	
1 6 2	プログラム・モジュール	
1 6 4	プログラム・データ	
1 6 6	キーボード	
1 6 8	ポインティング・デバイス	
1 7 0	インターフェース	
1 7 6	リモート・コンピュータ	
1 7 8	メモリ記憶素子	
1 8 0	ローカル・エリア・ネットワーク(LAN)	
1 8 2	ワイド・エリア・ネットワーク(WAN)	50

1 8 4	アダプタ	
1 8 6	モデム	
2 0 2	コントローラ	
2 0 4	メディア分析エンジン	
2 0 6	メモリ / 記憶装置	
2 1 2	カラー・オブジェクト分析部	
2 1 4	時間スライス分析機能	
2 1 6	相関検出部	
2 1 8	カラー空間量子化部	
2 2 0	モーション・パターン分析およびキー・フレーム抽出機能	10
2 2 4	メディア・コンテンツ	
2 2 6	拡張シーン・ウィンドウ・データ構造	
2 2 8	シーン・データ構造	
3 0 2 , 3 0 4	フレーム	
3 0 6 A ... N , 3 0 8 A ... N	優勢カラー・オブジェクト	
1 0 0 0	記憶媒体 / 素子	

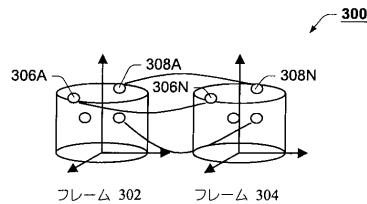
【図1】



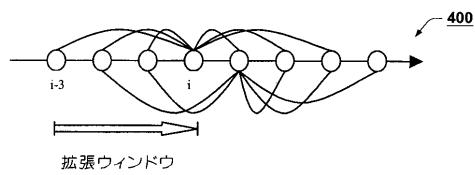
【図2】



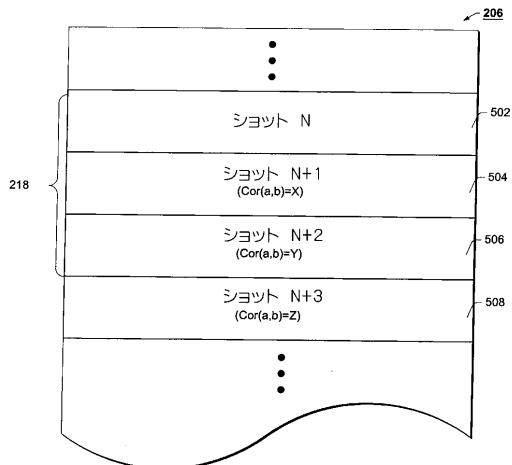
【図3】



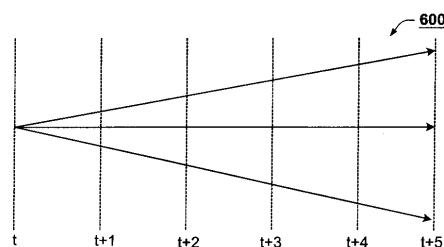
【図4】



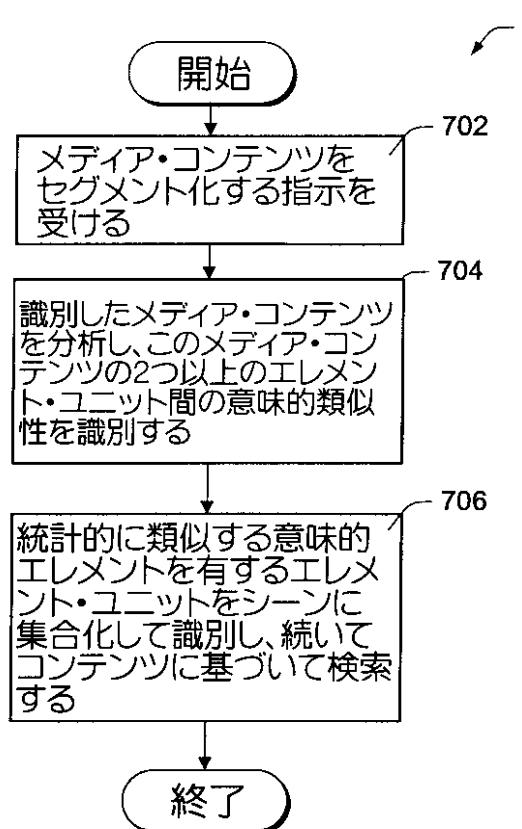
【図5】



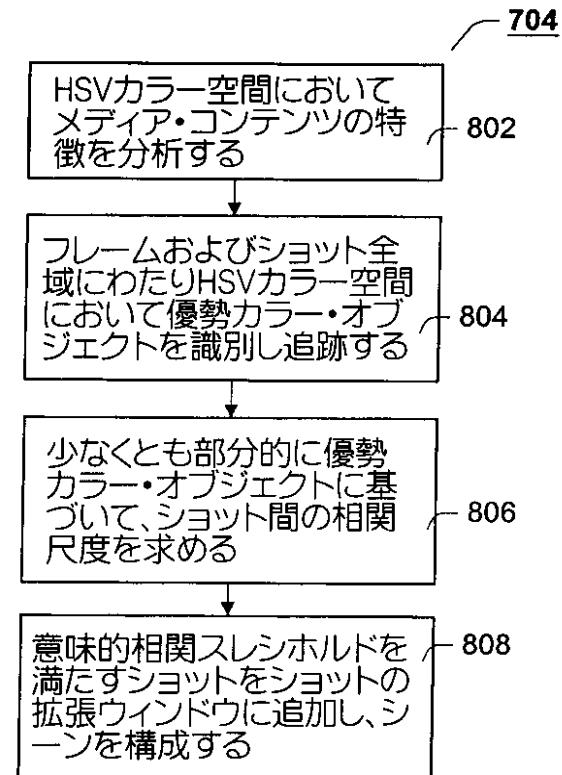
【図6】



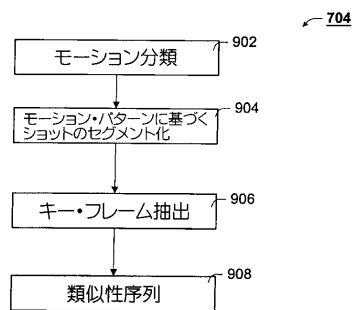
【図7】



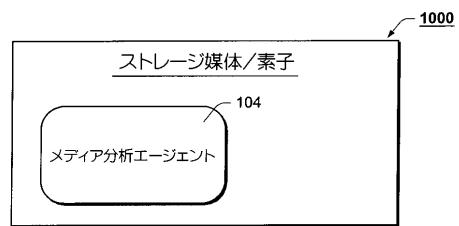
【図8】



【図9】



【図10】



フロントページの続き

(74)代理人 100096013

弁理士 富田 博行

(72)発明者 トン・リン

中華人民共和国 100871 ペキン, ペキン・ユニバーシティ, ビルディング 39, ルーム
308

(72)発明者 ホン・ジャン・チャン

中華人民共和国 101300 ペキン, シュン・イ・ディストリクト, キャピタル・パラダイス
701

審査官 田中 紗子

(56)参考文献 特開2000-069420 (JP, A)

特開2000-187731 (JP, A)

(58)調査した分野(Int.Cl., DB名)

H04N 5/76-5/956

G06T 7/00

G06T 7/20