

(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 I

H O 4 N 5/91 (2006. 01)

H O 4 N 5/91 N

H O 4 N 5/76 (2006. 01)

H O 4 N 5/76 Z

G O 6 T 7/00 (2006. 01)

G O 6 T 7/00 1 O O B

G O 6 T 7/20 (2006. 01)

G O 6 T 7/00 1 O O D

G O 6 T 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 レッドモンド 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 忠武
 (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】

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、

40

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、

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。

10

【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、

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。

20

【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、

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

30

40

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、

50

を備えた計算システム。

【請求項 1 5】

請求項 1 3 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 1 6】

請求項 1 3 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、を備えた計算システム。

10

【請求項 1 7】

請求項 1 6 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 1 8】

請求項 1 7 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、を備えた計算システム。

20

【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、計算システム上で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

40

【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・ギャラリー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

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、標準的なブル・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。

【 0 0 0 4 】

【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にはなく、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、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の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

【 0 0 0 5 】

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

【 0 0 0 6 】

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く、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

【 0 0 0 7 】

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。

【 0 0 0 8 】

【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

【 0 0 0 9 】

【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。

【 0 0 1 0 】

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キ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。

【 0 0 1 1 】

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 1 0 4 を含むコンピュータ・システムの一例 1 0 2 を示す。メディア分析エージェント 1 0 4 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 1 0 4 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 1 0 4 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、リップ・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 1 0 4 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 のコンテキストでは、分析エージェント 1 0 4 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 1 0 2 は代わりにエージェント 1 0 4 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 1 0 4 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 1 0 2 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。

【 0 0 1 2 】

図示のように、コンピュータ 1 0 2 は、1 つ以上のプロセッサ即ち演算装置 1 3 2、システム・メモリ 1 3 4、およびシステム・メモリ 1 3 4 からプロセッサ 1 3 2 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 1 3 6 を含む。

【 0 0 1 3 】

バス 1 3 6 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) 1 3 8、およびランダム・アクセス・メモリ (RAM) 1 4 0 を含む。基本入出力システム 1 4 2 (BIOS) は、起動中のように、コンピュータ 1 0 2 内のエレメント間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 ROM 1 3 8 内に格納されている。コンピュータ 1 0 2 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 1 4 4、リムーバブル磁気ディスク 1 4 8 に対する読み書きを行なう磁気ディスク・ドライブ 1 4 6、および CD ROM、DVD ROM またはその他のこ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、ROM 138、または RAM 140 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、イントラ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、フロッピー・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、コンピュータ自体も含む。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る。本発明は、このようなサブコンポーネントに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。

【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。

【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と組み合わせる呼び出し、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・ウィンドウ・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

、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。

【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 (ASIC)、現場プログラム可能ゲート・アレイ (FPGA)、プログラマブル論理アレイ (PLA) 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。

【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関検出部 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。

【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。

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

10

20

30

40

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、「H S V」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。H S V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、H S V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、H S Vカラー空間は、カラー量子化部 2 1 8 によって、三次元 (3 D) デカルト座標系で量子化され、それぞれ X および Y に 1 0 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 1 0 個の値を用いることにより、カラー量子化部 2 1 8 は、H S V 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

【 0 0 2 6 】

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 D C ブロックを、カラー量子化部 2 1 8 によって、量子化カラー空間 (例えば、H S V カラー空間) に投影する。3 D 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 3 D カラー・ヒストグラムを形成する。3 D 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 2 1 8 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 2 1 2 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 2 1 2 は、上位 2 0 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

【 0 0 2 7 】

次に、カラー・オブジェクト分析部 2 1 2 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 3 D 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。

【 0 0 2 8 】

【 数 1 】

$$hist_d(k,x,y,z)$$

(1)

30

ここで、k はフレーム番号を示し、x , y , z はカラー・ピンを示す。尚、前述のことから、カラー・オブジェクト分析部 2 1 2 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てはまる画素を考慮することが認められよう。

【 0 0 2 9 】

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 2 1 2 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を有するカラー・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 2 1 2 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

【 0 0 3 0 】

【 数 2 】

$$hist_d^a(x,y,z)$$

(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

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】

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

ここで、 d_o はショットの持続時間、 d_l 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。続いて、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) $0 \leq cor(a, b) \leq 1$, $cor(a, a) = 1$

2) $cor(a, b) = cor(b, a)$

ショット集合化のための拡張ウィンドウ方式

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 (Tc) 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 の拡張ウィンドウ 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】

10

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

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

20

【0039】

【数6】

$$v > \text{mean-var} \quad (6)$$

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。

30

【0040】

【数7】

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

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、

40

$\text{left}(i) = \max[\text{cor}(i, i-1), \text{cor}(i, i-2), \text{cor}(i, i-3)]$

$\text{left}(i+1) = \max[\text{cor}(i+1, i-1), \text{cor}(i+1, i-2)]$

$\text{right}(i) = \max[\text{cor}(i, i+1), \text{cor}(i, i+2), \text{cor}(i, i+3)]$

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

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

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

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。一実施形態によれば、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

30

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。一実施形態によれば、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。

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】

一実施形態によれば、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

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、関連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どうか判定を行い、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する。

【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、アプリケーション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。一実施形態によれば、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・アクセラ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形態は、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キ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。

【 0 0 5 2 】

図示のように、図 7 の方法は、開始すると、ブロック 7 0 2 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 1 0 4 のコントローラ 2 0 2 は、ローカル・アプリケーション（例えば、2 1 0）または外部ソースから、即ち、I / O インターフェース 2 0 8 を介して、指示を受ける。

【 0 0 5 3 】

応答して、メディア分析エージェント 1 0 4 は、ブロック 7 0 4 において、メディア分析エンジン 2 0 4 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 2 0 4 は、選択的にカラー・オブジェクト分析部 2 1 2 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 2 1 4 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、相関検出部 2 1 6 を呼び出し、統計的に意味的に繋がりのあるショットを識別する。

【 0 0 5 4 】

ブロック 7 0 6 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関検出部 2 1 6 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（2 1 8）にこのショットを追加する。拡張ウィンドウ 2 1 8 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 1 0 4 を解放する。

【 0 0 5 5 】

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 8 0 2 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 2 1 8 によって量子化する。

【 0 0 5 6 】

ブロック 8 0 4 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 2 0 2 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。

【 0 0 5 7 】

ブロック 8 0 6 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関検出部 2 1 6 に送る。相関検出部 2 1 6 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 8 0 8 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。

【 0 0 5 8 】

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 2 0 4 は、選択的に、カラー・オブジェクト分析部 2 1 2 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 2 1 4 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 2 1 2 とは異なり、時間スライス分析機能 2 1 4 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。

【 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 9 0 2 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 2 2 0 は、ブロック 9 0 4 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 9 0 6 において、コントローラ 2 0 2 は、時間分析関数 2 2 2 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 2 2 2 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。

【 0 0 6 1 】

ブロック 9 0 8 において、識別したショットのキー・フレームの特徴を相関検出部 2 1 6 に供給する。相関検出部 2 1 6 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キ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 2 0 2 は、拡張ウィンドウ 2 1 8 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。

代替実施形態

図 1 0 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 1 0 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 1 0 4 を実現する部分集合を含む、記憶媒体 / 素子 1 0 0 0 を示す。

【 0 0 6 2 】

ここで用いる場合、記憶媒体 1 0 0 0 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 1 0 0 0 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 1 0 0 0 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 1 0 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。

【 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の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 1 0 4 のアプリケーション 2 1 0 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 1 0 4 は、カラー・オブジェクト分析部 2 1 2、カラー量子化部 2 1 8 および相関検出部 2 1 6 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

オーディオ・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】

- 102 コンピュータ・システム
- 104 革新的メディア分析エージェント
- 132 演算装置
- 134 システム・メモリ
- 136 バス
- 138 リード・オンリ・メモリ (ROM)
- 140 ランダム・アクセス・メモリ (RAM)
- 142 基本入出力システム (BIOS)
- 144 ハード・ディスク・ドライブ
- 146 磁気ディスク・ドライブ
- 148 リムーバブル磁気ディスク
- 150 光ディスク・ドライブ
- 154 SCSI インターフェース
- 156 入出力 (I/O) インターフェース
- 158 オペレーティング・システム
- 160 アプリケーション・プログラム
- 162 プログラム・モジュール
- 164 プログラム・データ
- 166 キーボード
- 168 ポインティング・デバイス
- 170 インターフェース
- 176 リモート・コンピュータ
- 178 メモリ記憶素子
- 180 ローカル・エリア・ネットワーク (LAN)
- 182 ワイド・エリア・ネットワーク (WAN)

30

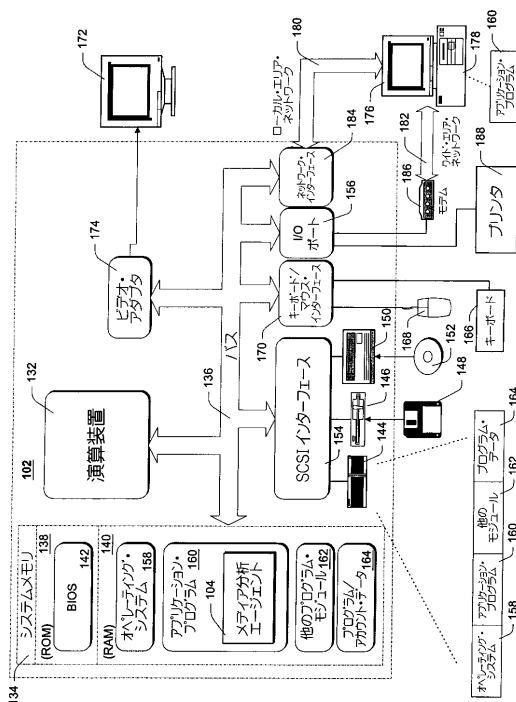
40

50

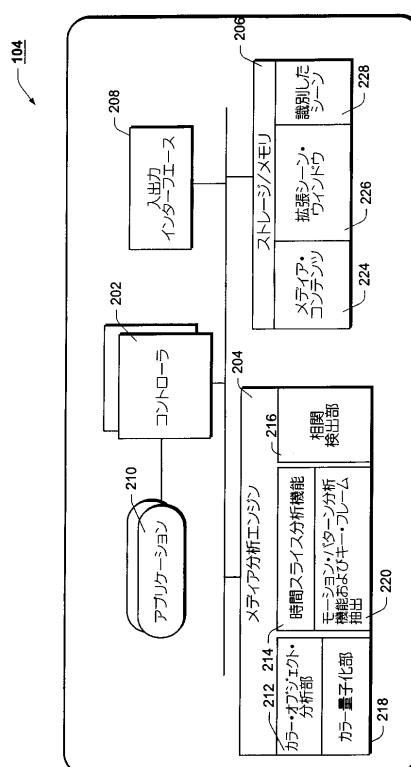
184	アダプタ	
186	モデム	
202	コントローラ	
204	メディア分析エンジン	
206	メモリ/記憶装置	
212	カラー・オブジェクト分析部	
214	時間スライス分析機能	
216	相関検出部	
218	カラー空間量子化部	
220	モーション・パターン分析およびキー・フレーム抽出機能	
224	メディア・コンテンツ	
226	拡張シーン・ウィンドウ・データ構造	
228	シーン・データ構造	
302, 304	フレーム	
306A...N, 308A...N	優勢カラー・オブジェクト	
1000	記憶媒体/素子	

10

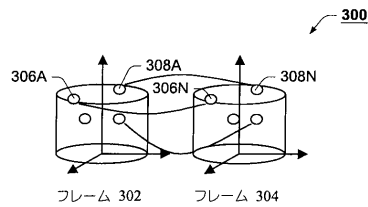
【図1】



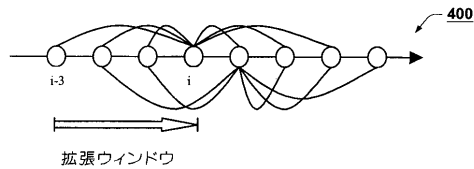
【図2】



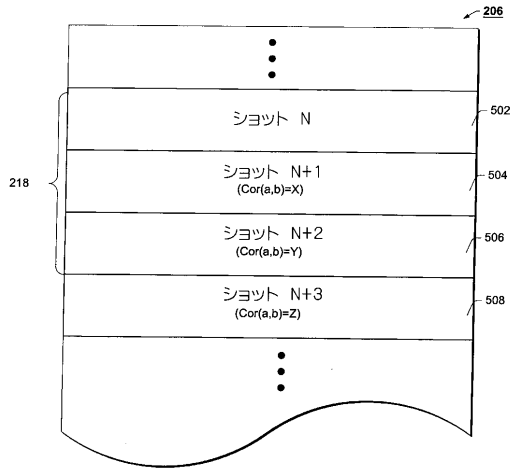
【図 3】



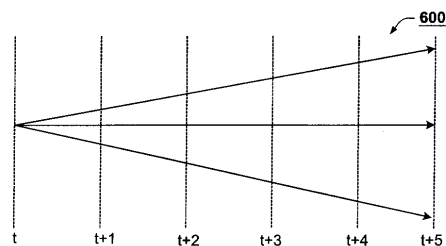
【図 4】



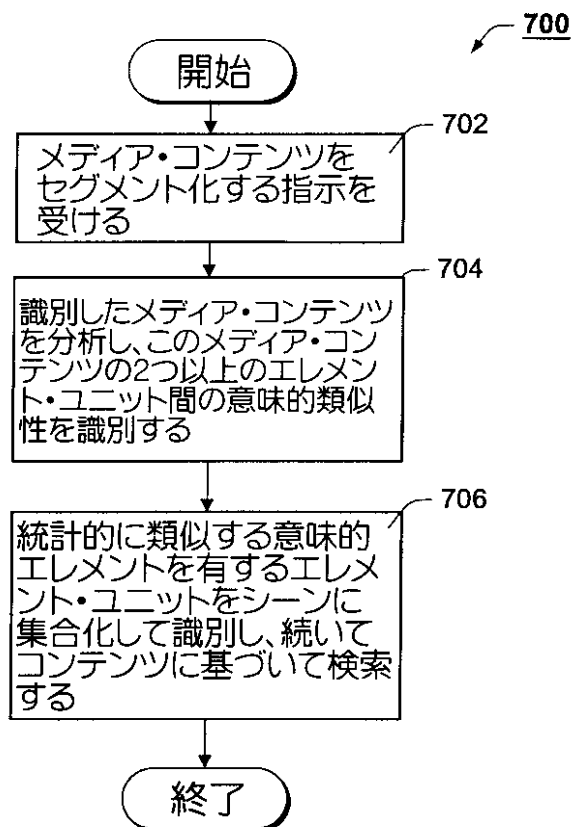
【図 5】



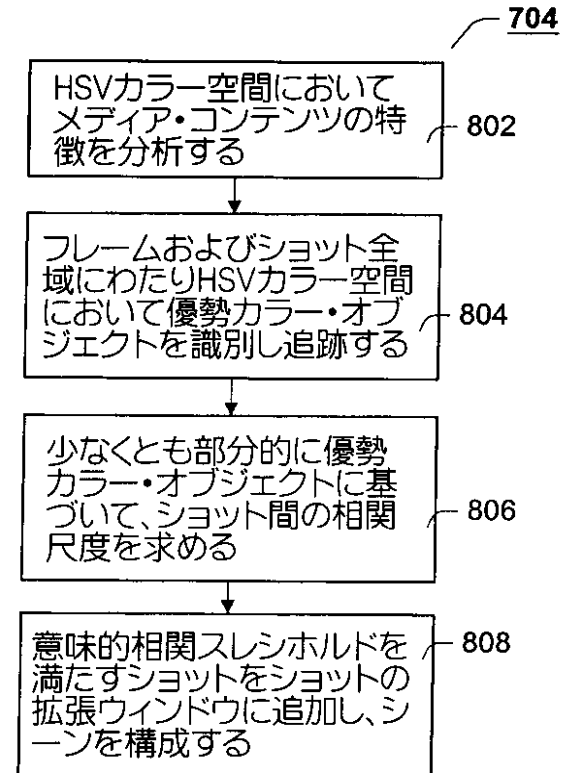
【図 6】



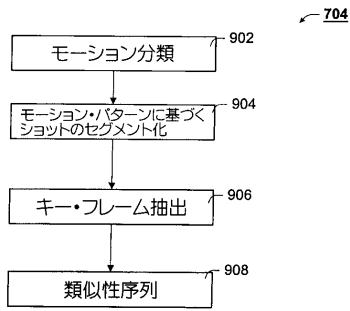
【図 7】



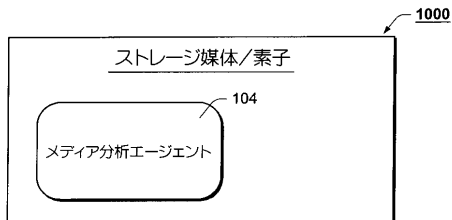
【図 8】



【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 1 0 0 8 7 1 ペキン, ペキン・ユニバーシティ, ビルディング 3 9 , ルーム
3 0 8

(72)発明者 ホン・ジャン・チャン

中華人民共和国 1 0 1 3 0 0 ペキン, シュン・イ・ディストリクト, キャピタル・パラダイス
7 0 1

審査官 田中 絢子

(56)参考文献 特開 2 0 0 0 - 0 6 9 4 2 0 (J P , A)

特開 2 0 0 0 - 1 8 7 7 3 1 (J P , A)

(58)調査した分野(Int.Cl. , D B 名)

H04N 5/76-5/956

G06T 7/00

G06T 7/20