

(19) 日本国特許庁(JP)

## (12) 公表特許公報(A)

(11) 特許出願公表番号

特表2006-510114  
(P2006-510114A)

(43) 公表日 平成18年3月23日(2006.3.23)

(51) Int.C1.

G06F 17/30

(2006.01)

F 1

G06F 17/30

G06F 17/30

210A

220C

テーマコード(参考)

5B075

		審査請求 未請求 予備審査請求 有 (全 16 頁)
(21) 出願番号	特願2004-560808 (P2004-560808)	(71) 出願人 390009531
(86) (22) 出願日	平成15年12月12日 (2003.12.12)	インターナショナル・ビジネス・マシーンズ・コーポレーション
(85) 翻訳文提出日	平成17年8月10日 (2005.8.10)	INTERNATIONAL BUSINESS MACHINES CORPORATION
(86) 國際出願番号	PCT/US2003/039548	アメリカ合衆国10504 ニューヨーク
(87) 國際公開番号	W02004/055639	州 アーモンク ニューオーチャード
(87) 國際公開日	平成16年7月1日 (2004.7.1)	ロード
(31) 優先権主張番号	10/319,295	(74) 代理人 100086243
(32) 優先日	平成14年12月13日 (2002.12.13)	弁理士 坂口 博
(33) 優先権主張国	米国(US)	(74) 代理人 100091568
		弁理士 市位 嘉宏
		(74) 代理人 100108501
		弁理士 上野 剛史

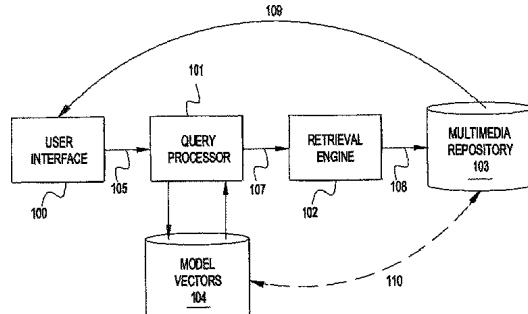
最終頁に続く

(54) 【発明の名称】概念モデル空間におけるコンテンツの表現及びこれを検索するための方法及び装置

## (57) 【要約】

【課題】 モデル・ベクトルを用いてマルチメディア文書を索引付けすること。

【解決手段】 マルチメディア文書からモデル・ベクトル表現を抽出するための方法及び装置。モデル・ベクトルは、マルチメディア文書が一組のカテゴリに属すること、又は、一組の意味概念が当該文書に関連することに対する確信性の多次元的表現を与える。モデル・ベクトルはマルチメディア文書と関連して、そのコンテンツの索引又はカテゴリ化を与えることができ、当該マルチメディア文書を比較、検索、分類、又はクラスタ化するために用いることができる。モデル・ベクトルは、情報探索、マルチメディア・コンテンツの個人化、及びマルチメディア情報貯蔵庫照会、の目的のために用いることができる。



**【特許請求の範囲】****【請求項 1】**

マルチメディア文書を表わすための少なくとも1つのモデル・ベクトルを生成する方法であって、

複数の概念検出器を前記マルチメディア文書に適用するステップと、  
前記マルチメディア文書を各々の検出器に関して採点するステップと、  
前記採点を多次元空間にマッピングして、少なくとも1つのベクトル表現を形成するステップと、  
を有する方法。

**【請求項 2】**

前記複数の概念検出器が、カテゴリ、対象、事象、シーン、及び人々についての固定の語彙目録に対応する、請求項1に記載の方法。

**【請求項 3】**

前記文書が、オーディオ、ビジュアル、テキスト、及びスピーチといった多数の様式を有し、前記概念検出器が前記マルチメディア文書に存在する単一の又は多数の様式について動作する、請求項1に記載の方法。

**【請求項 4】**

前記概念検出器が、前記マルチメディア文書から抽出された文書形態のコンテンツベースの記述子について動作する、請求項1に記載の方法。

**【請求項 5】**

前記概念検出器が、前記マルチメディア文書に関連するメタデータについて動作する、請求項1に記載の方法。

**【請求項 6】**

前記概念検出器が、前記マルチメディア文書に関連する概念情報について動作する、請求項1に記載の方法。

**【請求項 7】**

前記概念検出器が知識ベースで動作する、請求項1に記載の方法。

**【請求項 8】**

前記検出器が、訓練されたモデル統計分類器に対応する、請求項1に記載の方法。

**【請求項 9】**

前記採点が、前記マルチメディア文書において概念を検出することの確信性、前記マルチメディア文書に対する概念の関連性、及び前記概念に関する前記検出器の信頼性のうちの少なくとも1つに基づく、請求項1に記載の方法。

**【請求項 10】**

前記マッピングが、前記採点を連結するステップ、線形変換を実行するステップ、非線形変換を実行するステップ、量子化を実行するステップ、及びしきい値化によって次元数を削減するステップのうちの少なくとも1つに対応する、請求項1に記載の方法。

**【請求項 11】**

前記モデル・ベクトルの次元に対する前記採点のマッピングが、一対一、一対多、多対一、又は多対多とすることができる、請求項10に記載の方法。

**【請求項 12】**

前記複数の検出器が、マルチメディア文書の前記索引付けのために準備される、請求項1に記載の方法。

**【請求項 13】**

前記採点及びマッピング方法が、マルチメディア文書の前記索引付けのために準備される、請求項1に記載の方法。

**【請求項 14】**

モデル・ベクトルを用いてマルチメディア文書を索引付けするための方法であって、  
1つ又はそれ以上のモデル・ベクトルを各マルチメディア文書に対して生成するステップと、

10

20

30

40

50

前記モデル・ベクトルをこれと対応するマルチメディア文書と関連付けるステップと、前記関連付けられたモデル・ベクトルの値に基づいて、前記マルチメディア文書にアクセスするための索引を作るステップと、  
を有する方法。

【請求項 15】

マルチメディア文書を表わすために少なくとも1つのモデル・ベクトルを生成する前記ステップが、

複数の概念検出器を前記マルチメディア文書に適用するステップと、  
前記マルチメディア文書を各々の検出器に関して採点するステップと、  
前記採点を多次元空間にマッピングして、少なくとも1つのベクトル表現を形成するステップと、  
を有する請求項14に記載の方法。

【請求項 16】

前記複数の検出器が、マルチメディア文書の前記索引付けのために準備される、請求項15に記載の方法。

【請求項 17】

前記採点方法及びマッピング方法が、マルチメディア文書の前記索引付けのために準備される、請求項15に記載の方法。

【請求項 18】

多数のモデル・ベクトルが、多数の様式、形態、記述子、又はモデル、の各々に基づいて、各マルチメディア文書について生成される、請求項14に記載の方法。

【請求項 19】

前記関連付けるステップが、データベースのキー値、メディア・ロケータ、又は他の種類の識別子、に基づく、請求項14に記載の方法。

【請求項 20】

前記索引が、前記モデル・ベクトル値に基づいて、類似性検索、隣接アクセス、又は範囲検索、を可能にする、請求項14に記載の方法。

【請求項 21】

アプリケーションにおいてモデル・ベクトルをマルチメディア文書と併せて用いるための方法であつて、

各マルチメディア文書を表わすために、少なくとも1つのモデル・ベクトルを、  
複数の概念検出器を前記マルチメディア文書に適用するステップと、  
前記マルチメディア文書を各検出器に関して採点するステップと、  
前記採点を多次元空間にマッピングして、少なくとも1つのベクトル表現を形成するステップと、  
によって生成するステップと、

前記少なくとも1つのベクトル表現の値に基づいて、前記マルチメディア文書について少なくとも1つの動作を実行するステップと、  
を有する方法。

【請求項 22】

前記少なくとも1つの動作が、マルチメディア情報貯蔵庫から文書を検索して取り出すために、前記モデル・ベクトルを用いるステップを含む、請求項21に記載の方法。

【請求項 23】

前記少なくとも1つの動作が、マルチメディア情報をフィルタ処理するステップ、要約するステップ、及び個人化するステップのうちの少なくとも1つを有する請求項21に記載の方法。

【請求項 24】

前記少なくとも1つの動作が、データ発掘するステップを有する、請求項21に記載の方法。

【請求項 25】

10

30

40

50

前記少なくとも1つの動作が、前記文書をクラスタ化するステップを有する請求項21に記載の方法。

【請求項26】

前記少なくとも1つの動作が、前記文書を分類化するステップを有する請求項21に記載の方法。

【請求項27】

マルチメディア文書を表わすために少なくとも1つのモデル・ベクトルを生成する方法を実行する機械により実行可能な命令プログラムを実体的に具体化する機械読取可能なプログラム格納装置であって、前記方法が、

複数の概念検出器を前記マルチメディア文書に適用するステップと、

前記マルチメディア文書を各々の検出器に関して採点するステップと、

前記採点を多次元空間にマッピングして、少なくとも1つのベクトル表現を形成するステップと、

を有する、プログラム格納装置。

【請求項28】

アプリケーションにおいてモデル・ベクトルをマルチメディア文書と併せて用いるための方法を実行する機械により実行可能な命令プログラムを実体的に具体化する機械読取可能なプログラム格納装置であって、前記方法が、

各マルチメディア文書を表わすために、少なくとも1つのモデル・ベクトルを、

複数の概念検出器を前記マルチメディア文書に適用するステップと、

前記マルチメディア文書を各検出器に関して採点するステップと、

前記採点を多次元空間にマッピングして、少なくとも1つのベクトル表現を形成するステップと、

によって生成するステップと、

前記少なくとも1つのベクトル表現の値に基づいて、前記マルチメディア文書について少なくとも1つの動作を実行するステップと、

を有する、プログラム格納装置。

【請求項29】

アプリケーションにおいてモデル・ベクトルをマルチメディア文書と併せて用いるシステムであって、

各マルチメディア文書を表わすために少なくとも1つのモデル・ベクトルを生成する少なくとも1つのモデル・ベクトル生成コンポーネントと、

前記少なくとも1つのベクトル表現の値に基づいて、前記マルチメディア文書について少なくとも1つの動作を実行するための少なくとも1つの文書処理コンポーネントと、を有するシステム。

【請求項30】

前記少なくとも1つのモデル・ベクトル生成コンポーネントが、

複数の概念検出器を前記マルチメディア文書に適用するための少なくとも1つの概念検出器適用コンポーネントと、

前記マルチメディア文書を各検出器に関して採点するための採点コンポーネントと、

前記採点を多次元空間にマッピングして、少なくとも1つのベクトル表現を形成するためのマッピング・コンポーネントと、

を有する、請求項29に記載のシステム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、マルチメディア文書を索引付けするためにモデル・ベクトルを用いることに関し、より詳細には、モデル・ベクトル表現を形成して、モデル・ベクトルをマルチメディア文書と関連付けて索引を与え、モデル・ベクトルを用いてマルチメディア文書を検索し、分類し、クラスタ化する方法及び装置に関する。本発明は、さらに、情報の探索、マ

10

20

30

40

50

ルチメディア・コンテンツの個人化、及びマルチメディア情報貯蔵庫照会、の目的のためにモデル・ベクトルを用いることに関する。

【背景技術】

【0002】

ますます増加しているビデオ、イメージ、テキストその他のマルチメディア文書形態のデジタル情報の量は、該情報を索引付けし、検索し、カテゴリ化し、かつ、組織化するためのより効率的な方法に対する必要性を駆り立てている。コンテンツ分析、形態抽出、及び分類における最近の進歩は、マルチメディア文書を効率的に検索し、フィルタ処理するための機能を改善しつつある。しかし、色、テクスチャ、形状、動きのようなマルチメディア・コンテンツから自動的に抽出することができる低レベルの形態記述と、マルチメディア・システムのユーザにとって意味のある対象、事象、シーン、人々のような意味記述との間で顕著なギャップが残る。

【0003】

マルチメディアの索引付けの問題は、人為的で、半自動的な、又は完全に自動的な処理を要求する多数の手法によって対処することができる。1つの手法は、人が人為的に、ラベル、カテゴリ、又は記述をマルチメディア文書に割り当てる可能にする注釈ツール又はカタログ化ツールを用いるものである。例えば、著者、M.Naphade、C.Y.Lin、J.R.Smith、B.Tseng及びS.Basuは、(非特許文献1)という表題の論文において、ラベルをビデオ・ショットに割り当てる可能にするビデオ注釈ツールについて述べている。著者らは、さらに、能動的学習に基づいてラベルを割り当てる半自動的な方法を教示する。完全に自動的な手法もまた可能である。例えば、著者M.Naphade、S.Basu、及びJ.R.Smithは、(非特許文献2)において、低レベルのビジュアル形態の統計的なモデル化に基づいて、ラベルを自動的にビデオ・システムに割り当てる方法について述べている。自動ラベル付け技術は、自動的に割り当たられたラベルに基づいて、ビデオの検索を可能にすることに対して有益であるが、索引付けは、少ないボキャブラリの照合値に限定されるため、ユーザがラベル用語の1つと照合しない検索語を入力した場合には、検索によって、どのような目標マルチメディア文書も見出されることはない。

【0004】

自動システムが、ラベル、カテゴリ、及び記述をマルチメディア文書に割り当てるための改善しつつある機能であるとすると、これらの記述を用いて、これらの文書を索引付けし、検索し、分類し、及びクラスタ化するためのより意味のある方法を与るために、これらの記述を活用する新規な技術が必要になる。さらに、システムは、自動システムの不確実性又は信頼性、並びに、マルチメディア文書に割り当たされたあらゆるラベル、カテゴリ、又は記述の関連性を考慮して、効率的な索引を与るべきである。

【0005】

【非特許文献1】「Learning to Annotate VideoDatabases」IS&T/SPIE Symposium on Electronic Imaging: Science and Technology-Storage & Retrieval for Image and VideoDatabases X、カリフォルニア州サンノゼ、2002年1月

【非特許文献2】「Statistical Modeling Approach to Content-based Video Retrieval」IEEE International Conference on Acoustics, Speech and Signal Processing(ICA-2002)、2002年5月

【発明の開示】

【発明が解決しようとする課題】

【0006】

したがって、本発明の目的は、あらゆる自動的なラベル付け、及び、確信性、信頼性及び関連性といったその対応する採点の結果を取り込むモデル・ベクトル表現を用いて、マルチメディア文書を索引付けするための方法及び装置を提供することである。

【0007】

本発明の別の目的は、情報の探索、マルチメディア・コンテンツの個人化、及びマルチメディア情報貯蔵庫照会、の用途において、モデル・ベクトル表現を用いることである。

10

20

30

40

50

**【課題を解決するための手段】****【0008】**

上述その他の目的は、マルチメディア文書の分類又はラベル付けの結果、及びあらゆる対応する不確実性、信頼性又は関連性採点を、該マルチメディア文書の検索、分類、及びクラスタ化に用いることができる多次元ベクトルに包み込むモデル・ベクトル表現を用いて、マルチメディア文書を索引付けするための装置及び方法を提供する本発明により実現される。モデル・ベクトル表現は、字句エンティティを多次元ベクトル空間の次元にマッピングすることを含み、これにより、文書がその多次元空間において表わされ、索引付けされることが可能になる。

**【0009】**

モデル・ベクトル表現の利点は、これが語彙目録全体にわたる広範囲のラベル付けを取り込むことである。これはまた、ラベル又は分類結果の不確実性を取り込む小型の表現を与える。モデル・ベクトル表現は、さらに、その実値多次元性質が距離空間における効率的な索引付けを可能にして、該モデル・ベクトル表現の距離又は類似性において簡単な計算を可能にすることにより、索引付けにおいても利点を有する。このことは、モデル・ベクトルを、マルチメディア文書の類似性検索、関連性フィードバックベースの検索、分類、クラスタ化、フィルタ処理などのために用いる効率的な方法を可能にする。

**【発明を実施するための最良の形態】****【0010】**

本発明は、以下で、添付図面を特に参照して、より詳細に述べられる。

**【0011】**

図1は、本発明の特徴を有するマルチメディア情報検索システムの一例を示す。図示されるように、ユーザは、ユーザ・インターフェース(100)を通して、ステップ(105)において、照会をマルチメディア情報検索システムに向けて発行する。この照会は、照会プロセッサ(101)により取り扱われる。照会プロセッサは、ステップ(106)において、格納された索引値(104)の組を検索して、ユーザの照会に対する照合を見出す。索引値がモデル・ベクトル形態である場合には、この値は、各々の索引付けされたマルチメディア文書の意味次元に関する多次元ベクトルに対応する。この照合は、ステップ(107)において検索エンジンに渡されて、照合するマルチメディア文書が、ステップ(108)においてマルチメディア貯蔵庫(103)から取り出される。対応するマルチメディア文書は、モデル・ベクトルと、貯蔵庫内の特定のマルチメディア文書との関連(110)に基づいて求められる。マルチメディア文書は、次いで、ステップ(109)においてユーザに戻されて、ユーザ・インターフェース(100)で表示される。モデル・ベクトル表現は、ユーザの照会に対する照合を見出すのに用いることができる一組の格納された索引値(103)を表わすための1つの方法を与える。

**【0012】**

モデル・ベクトル表現は、一連の検出器又は分類器をマルチメディア文書に適用した結果を包み込む。例えば、{「車」、「ボート」、「列車」}の概念がマルチメディア文書に示されているかどうかを検出することにより、以下の語彙目録{「車」、「ボート」、「列車」}からの字句エンティティを割り当てる一組の分類器を考慮されたい。検出問題は、各概念が存在することに対する確実性を反映する採点を割り当てるこにより、各概念の有無を検出する一組のバイナリ分類器として見ることができる。例えば、このシステムは、「車」に対して採点0.75を与えることができ、これは、「車」ラベルが割り当てられる確信性は75%であることを意味すると解釈することができる。一方、システムは、「列車」に対しては採点0.25を与えることができ、これは、「列車」ラベルが割り当てられる確信性は25%であることを意味すると解釈することができる。全体的には、システムは、これらの多数の検出器に対する採点をもたらし、モデル・ベクトルはこれらの採点を单一の表現で取り込み、これは、次いで、マルチメディア文書に対する索引として用いられることになる。

**【0013】**

10

20

30

40

50

図2は、マルチメディア文書又は照会に対するモデル・ベクトルを生成する工程を示す。マルチメディア文書(200)は、最初に、複数の検出器(201)により操作されて、検出器の各々の基になる概念に関連して採点される。検出器自体は、固定の語彙目録(204)、又は固定のカテゴリ、対象、事象、シーン、又は人々の組に対応することができる。例えば、米国議会図書館グラフィック材料同意語辞典は、写真その他の種類のグラフィック文書をカタログ化するための一組のカテゴリを与える。検出器は、各検出器がTGMカテゴリの1つに対応するように造り、用いることができる。語彙目録(204)の概念はまた、一般的、特定的、又は抽象的なものとすることができます。例えば、或る概念は、「橋を示すシーン」のような一般的なエンティティに対応することができます。或いは、或る概念は、「ゴールデンゲートブリッジを示すシーン」のような特定のエンティティに対応することができます。最後に、或る概念は、「現代文明」のような抽象的なエンティティに対応することができます。検出器(201)は、訓練されたモデルその他の種類の統計分類器に対応することができます。訓練(205)の場合においては、ラベル付けされたマルチメディア文書の例は、学習工程において、検出器(201)モデル及びそれらのパラメータを定義するように用いることができる。検出器(202)の出力は、次いで、マッピング工程(202)において変換されて、モデル・ベクトル(203)が生成される。モデル・ベクトルは、語彙目録(204)の概念に関連するマルチメディア文書(200)の集合的な採点形態を与える。さらに、モデル・ベクトル(203)は、語彙目録に対するその採点を考慮することにより、マルチメディア文書(200)についての推論を可能にする。

10

20

30

40

#### 【0014】

図3は、一組の検出器を用いて、マルチメディア文書が分析され採点される、マルチメディア文書に対するモデル・ベクトルを生成する1つの実施形態を示す。各マルチメディア文書(300)は、N個の検出器(301ないし303)を適用することにより分類される。検出器は、サポートベクトルマシン、ガウス混合モデル、隠れマルコフモデル、ニューラルネット、ベイズネット、線形判別分析などを含む任意の数の形態を取ることができます。各検出器は、特定の意味概念を表わすことができる。例えば、語彙目録{「車」、「ボート」、「列車」}を考えると、検出器は、検出器1=「車」、検出器2=「ボート」、検出器3=「列車」の概念を表わすことがある。すなわち、検出器1は、「車」の概念がマルチメディア文書に関連があるかどうかを判断し、他の検出器も同様に動作する。検出器は、以前に、供給された地上較正によりラベル付けされた例に基づいて、モデルを学習するか又は作成する技術を用いることによって、それぞれの概念を検出するように訓練されていた可能性がある。

#### 【0015】

検出器(301ないし303)は、各々の検出を実行するために、マルチメディア文書(300)に関する様々な情報を用いることができる。例えば、検出器(301ないし303)は、マルチメディア文書(300)構成する情報の1つ又はそれ以上の様式(視覚、オーディオ、スピーチ、テキスト)を用いることができる。検出器(301ないし303)は、さらに、マルチメディア文書(300)からの情報についての異なる様式から抽出された色、テクスチャ、形状、動き、音の周波数、空間的又は時間的配置のようなコンテンツベースの記述子を用いることができる。例示的な記述子は、色ヒストグラム、縁ヒストグラム、動きベクトル、形状境界記述子などを含む。検出器(301ないし303)は、さらに、マルチメディア文書(300)に関するメタデータを用いることができる。例えば、題名、著者、作成日、ジャンルなどのような情報を用いることができる。さらに、マルチメディア文書(300)と他の文書との関連のような他の概念情報を用いることができる。検出器(301ないし303)は、さらに、語彙目録又はマルチメディア情報貯蔵庫に関する情報及び知識の組織に基づいた推測及び推論を可能にする知識ベース又は意味ネットを用いることができる。

#### 【0016】

各検出器においては、採点(305)が各マルチメディア文書に対して生成される。採

50

点は、マルチメディア文書（300）に関連する検出器により、それぞれの概念のモデル化に対する情報を与える。採点は、確信性又は不確実性（総合的に「確信性」と呼ばれる）のような多くのことを反映することができ、これによって検出器は、文書における概念、文書に対する概念の関連性、又は概念を検出する際の検出器の信頼性を検出する。例えば、上述の検出器1を考慮すると、採点は、該検出器がマルチメディア文書において「車」の表現を検出することができる確信性を示すことができる。確信性は、識別境界又はしきい値に対する近接性に関することがある。例えば、マルチメディア文書が、「車」を検出するための識別境界から遠い場合には、高い確信性があると判断される。しかし、マルチメディア文書が識別境界から近い場合には、低い確信性があると判断される。関連性採点は、概念がマルチメディア文書に対してどれだけ関連があるかを示す。例えば、「車」が部分的に示されるだけであったり、又は、マルチメディア文書の大部分を構成しない場合には、関連性採点は低いと判断することができる。或いは、信頼性採点は、検出器がそれぞれの概念を検出することに対してどれだけ信頼できるかを示すことができる。例えば、検出器1が、「車」の数個の例のみを用いて訓練された場合には、低い信頼性採点が求められることになる。しかし、これが多くの例を用いて訓練された場合には、高い信頼性採点が求められることになる。採点自体は、一次元値を生成するといったこれらの属性の1つだけを反映することができる。しかし、採点は、さらに、情報を多数の属性に与えることにより、他次元とすることができる。

#### 【0017】

検出器の各々に対して採点が生成されると、これらは、マップされて（304）、モデル・ベクトル（306）を生成する。各検出器（301ないし303）がその分類を実行するために、多数の様式（例えば、イメージ、ビデオ、オーディオ、テキスト、スピーチ）を用いる場合のような幾つかの場合においては、単一のモデル・ベクトル（306）が各マルチメディア文書（300）に対して生成される。或いは、各検出器が1つの様式のみを用いる場合においては、多数のモデル・ベクトル（306）を各マルチメディア文書に対して生成することができる。この場合には、多数のモデル・ベクトルを各マルチメディア文書に対して生成して、例えば、一方をオーディオ様式に対して、他方をイメージ様式に対してなどのように、多数の採点を反映することができる。

#### 【0018】

モデル・ベクトル又は複数のモデル・ベクトルを生成するためのマッピング（304）は、検出器から生成された採点の組み合わせ又は集合を与える。幾つかの場合においては、マッピングは、N次元のベクトルを生成するためにN個の採点を連結させるという単純な動作を与える。例えば、一次元の確信性採点が各検出器により生成される（すなわち、分類器1が採点C1、分類器2が採点C2、及び分類器3が採点C3を生成する）上述の3つの要素の語彙目録{「車」、「ボート」、「列車」}を考慮すると、連結工程は、三次元のモデル・ベクトルM=[C1, C2, C3]を生成する。或いは、マッピング（304）は、確信性採点の線形重み付け又は変換を生成することができる。

#### 【0019】

確信性採点は、検出器の信頼性又は分類結果の関連性により重み付けすることができる。3つの検出器の各々に対する信頼性採点R1、R2、R3のそれぞれを考慮されたい。重み付けにより、マッピング（304）は、信頼性採点R<sub>i</sub>を確信性採点C<sub>i</sub>で乗算することにより、三次元のモデル・ベクトルM=[R1 \* C1, R2 \* C2, R3 \* C3]を生成することができる。或いは、3つの検出器の各々に対する関連性採点L1、L2、L3を考慮すると、マッピング（304）は、関連性採点L<sub>i</sub>を確信性採点C<sub>i</sub>で乗算することにより、三次元のモデル・ベクトルM=[L1 \* C1, L2 \* C2, L3 \* C3]を生成することができる。主要構成要素分析、特異値分解、ウェーブレット変換、離散コサイン変換といった場合におけるような他のマッピング（304）は、線形変換及び/又は次元数の削減を与えることができる。或いは、サポートベクトルマシン、ニューラルネットといった場合におけるようなマッピング（304）は、非線形変換を与えることができる。マッピング（304）は、さらに、離散空間又はバイナリ値空間に対する量子化を含

10

20

30

40

50

むことができる。例えば、マッピング段階（304）において、検出器からの確信性採点（305）をしきい値化することにより、各概念がマルチメディア文書（300）にあるかどうかを示すバイナリモデル・ベクトルを生成することができる。

#### 【0020】

全体的に、マッピング（304）は、個々の概念又は検出器（301ないし303）からモデル・ベクトル（306）の個々の次元への様々な特定のマッピングをもたらすことができる。採点（305）を連結するマッピング（304）におけるような幾つかの場合には、モデル・ベクトルの次元に対する概念の一対一のマッピングが生成される。しかし、他の場合においては、例えば、モデル・ベクトル（306）の次元数を元の概念空間に関連して減少させるように、多対一のマッピングを生成することが望ましいとすることができる。他の場合においては、マッピング（304）は、例えば、ある程度の冗長性をモデル・ベクトル（306）に可能にするように一対多又は多対多とすることができる。

#### 【0021】

図4は、検出器の採点に基づいて生成されたモデル・ベクトルの例を示す。固定の語彙目録 = {「都市風景」、「顔」、「室内」、「景色」、「独白」、「野外」、「人々」、「テキスト\_オーバーレイ」}に対する検出器の結果（400）の採点された組が与えられると、検出器の採点（400）の各々を多次元のモデル・ベクトル（401）の独特な次元にマッピングすることにより、モデル・ベクトル（401）が生成される。本例においては、0.35である「都市風景」の採点がモデル・ベクトルの第1次元にマップされる。0.87である「顔」の採点が第2次元にマップされ、以下同様に続く。異なるマルチメディア文書に対応するモデル・ベクトル間の照合を助けるために、モデル・ベクトルの次元に対する検出器の採点の一貫したマッピングを用いることができる。同様に、より大きい語彙目録に対する例示的な検出器の採点（402）は、モデル・ベクトルの次元（403）にマップすることができる。この場合においては、「動物」に対する採点がモデル・ベクトルの第1次元にマップされる。「海辺」に対する採点が第2次元にマップされ、以下同様に続く。

#### 【0022】

図5は、モデル・ベクトルを用いたマルチメディア文書の索引付け工程を示す。第1に、K個のマルチメディア文書の集合（500）がモデル・ベクトル生成工程（501）において分析されて、一組のM個のモデル・ベクトル（502）を生成する。モデル・ベクトル生成工程（501）は、マルチメディア文書（500）のすべてにわたり、固定の語彙目録（505）及び対応する検出器の組を用いて、モデル・ベクトル（502）を生成する際の一貫性を可能にすることができる。さらに、モデル・ベクトル生成工程（501）はまた、同じ理由のために、マルチメディア文書（500）のすべてにわたり、採点（305）及びマッピング（304）に対して固定のパラメータの組を用いることができる。モデル・ベクトル（502）が生成されると、これらは対応するマルチメディア文書（500）と関連付けられる。例えば、この関連は、各モデル・ベクトル（502）と各マルチメディア文書（500）との間の一次キー・外部キー関係を述べるデータベースのキー値を用いて表わすことができる。或いは、この関連は、各モデル・ベクトル（502）を、対応するマルチメディア文書（500）のアドレスを与えるメディア・ロケータと関連付けることにより表わすことができる。或いは、各マルチメディア文書（500）を独自に識別する識別子を用いて、モデル・ベクトル（502）の各々の関連が表わされるようになることが可能である。さらに、各マルチメディア文書（500）のヘッダ又はメタデータ・フィールドにおけるモデル・ベクトルの値を表わすことによって、又は、ウォーターマーク付け又は何らかの他の持続的な関連付け方法によりモデル・ベクトル（502）の値を各マルチメディア文書（500）と持続的に関連付けることによって、モデル・ベクトル（502）を、直接、各マルチメディア文書（500）と関連付けることも可能である。

#### 【0023】

モデル・ベクトル（502）が生成され、これらのマルチメディア文書（500）との

10

20

30

40

50

関連が表わされると、該モデル・ベクトル(502)の値に基づいて、該マルチメディア文書(500)に対するアクセスを可能にする索引が作成される。この索引は、例えば、類似性検索又は隣接検索を可能にするような近接性をベースにしたアクセスを可能にすることができる。これらの場合においては、アクセスは、照会モデル・ベクトルを供給することにより達成することができ、同様なモデル・ベクトル又は最も近い目標モデル・ベクトルの固定サイズの組が索引から見出される。索引は、さらに、範囲をベースとしたアクセスをサポートすることができ、この場合には、照会モデル・ベクトルが供給され、該照会モデル・ベクトルから一定の距離の中にあるすべての目標モデル・ベクトルが該索引から見出される。

#### 【0024】

図6は、モデル・ベクトルを用いた照会工程を示す。モデル・ベクトルの値は、マルチメディア文書を取り出すために、ユーザの照会と照合される。ユーザ(600)は、照会(601)をマルチメディア文書検索システム(609)に対して発行する。照会は、ユーザにより与えられた例示的なモデル・ベクトル形態とすることができます。検索インターフェースは、任意的に、例えば、ユーザが照会に関連する意味概念を識別し、モデル・ベクトル表現を構築するために採点を割り当てることを可能にするインターフェースを与えることにより、ユーザが、照会モデル・ベクトルを開発することを可能にすることができます。或いは、検索インターフェースは、任意的に、マルチメディア文書をユーザに表示して、ユーザが、どのマルチメディア文書が照会に関連するかを選択することを可能にすることができます。システムは、次いで、事前計算された関連のモデル・ベクトルを照会のために用いるか、又は、照会時にモデル・ベクトルを生成する。或いは、照会は、ユーザにより与えられたマルチメディア文書の例の形態とすることができます、この場合には、照会マルチメディア文書は、照会モデル・ベクトルを作成するためのモデル・ベクトル生成工程を用いて分析し、処理することができる。

#### 【0025】

照会モデル・ベクトルが利用可能になると、これらはステップ(602)において、格納されたモデル・ベクトル値(606)と照合される。照合工程は、索引構造体を用いて目標モデル・ベクトルの照合を識別することを含むことができる。この照合は、上述のように、類似性検索、隣接検索、又は範囲の照合を含むことができる。照合工程(602)は、照会モデル・ベクトルに照合する格納されたモデル・ベクトル(606)を識別する照合リスト(603)を生成する。照合リスト(603)は、次いで、ステップ(604)において、任意的に採点される。照合の採点は、モデル・ベクトル値を用いて、距離空間計算に基づいて求めることができる。例えば、単一の照会モデル・ベクトルを考慮すると、照合の採点は、ユークリッド距離又はマンハッタン距離のような距離関数を用いて、多次元モデルのベクトル空間において測定された近接性に基づくものとすることができます。或いは、照合工程は、モデル・ベクトルの次元の幾つかだけを使用することが可能である。例えば、(400ないし401)におけるモデル・ベクトルを考慮すると、ユーザが「人間関連の」概念にのみ関心がある場合には、選択的に、次元2(「顔」)及び次元7(「人々」)を用いることができる。多数の照会モデル・ベクトルが与えられる場合においては、照合の採点は、個々のモデル・ベクトルからの距離の採点を組み合わせることにより取得することができる。照会モデル・ベクトルの中心を計算し、中心モデル・ベクトルを照会として用いるような他の選択肢も可能である。

#### 【0026】

採点された照合リストは、次いで、例えば、最良の照合をリストの最上位に移動するために、任意的に、ステップ(605)においてランク付けされる。次いで、ランク付けされたリストは、例えば、10個の最良の照合をリスト上に保持するために、任意的に、ステップ(606)において切り捨てられる。次いで、結果(607)がユーザに戻される。任意的に、検索システムは、これらのマルチメディア文書を結果リスト(607)におけるモデル・ベクトルに関連するマルチメディア貯蔵庫(610)から取り出して、該文書をユーザに表示する。

10

20

30

40

50

## 【0027】

結果がユーザに表示されると、ユーザは、例えば、関連性フィードバック技術を用いて、肯定的な例又は否定的な例を結果リスト(607)から識別することにより、検索を改善することができる。検索システム(609)は、この情報を照会処理と併せて用いて、モデル・ベクトルの採点に基づいて照合を取り出すことができる。

## 【0028】

モデル・ベクトルは、マルチメディア文書を取り出すために用いることができるが、これらはまた、マルチメディア文書のクラスタ化及び分類のために用いることができる。例えば、モデル・ベクトルを多次元距離空間において分析して、集積的クラスタ処理のような様々な技術を用いて、クラスタを識別することができる。モデル・ベクトルは、さらに、差別的又は生成的モデル化に基づくもののような様々な監視された学習方法を用いて分類することができる。例示的な分類器は、サポートベクトルマシン及びガウス混合モデルを含む。能動的学習及びブースティングのような他の技術もまた分類目的のためにモデル・ベクトル値に適用することができる。

## 【0029】

モデル・ベクトルはまた、情報の探索及びマルチメディア貯蔵庫の発掘のためにも用いることができる。例えば、モデル・ベクトル集合の次元の相関を調査して、概念がマルチメディア文書に現れるときの共起性についての情報を明らかにすることができる。

## 【0030】

図7は、モデル・ベクトルを用いたマルチメディア文書の適合を示す。このモデル・ベクトルは、マルチメディア貯蔵庫からのマルチメディア文書又は情報をフィルタ処理し、要約し、又は個人化するために用いることができる。ユーザ(700)は、要求(708)をマルチメディア情報システムに向けて発行する。この要求はステップ(701)で取り扱われる。要求は、(601)におけるように特定のユーザの照会を含むことができ、この場合、ユーザは例示的なモデル・ベクトル又はマルチメディア文書を供給するか、又は意味概念を識別する。或いは、要求はログイン形態であってもよく、ここでは、ユーザのプロファイルがプリファレンス情報と併せて、すでに格納されている。この場合においては、ユーザのプリファレンス情報は、任意的に、ステップ(702)でチェックすることができる。プリファレンス情報はまた、例示的なモデル・ベクトル、マルチメディア文書、又は識別された意味概念の形態で格納されてもよい。ユーザの照会及びユーザのプリファレンス情報は、次いで、集約され処理されて、照会モデル・ベクトルを生成し、これらは次いで、格納されたモデル・ベクトル(704)と照合され、これを取り出すように用いられ、次いで、ステップ(706)において、マルチメディア貯蔵庫(705)から選択的にマルチメディア文書を取り出すための索引として用いられる。これにより、モデル・ベクトル値に基づいたマルチメディア文書のフィルタ処理が与えられる。

## 【0031】

任意的に、ステップ(707)において、マルチメディア文書に関連するモデル・ベクトルを照会モデル・ベクトルと組み合わせて用いて、該マルチメディア文書のコンテンツに適合させることができる。この適合は、特定の照会に対するユーザのプリファレンスにより、マルチメディア・コンテンツを個人化することができる。例えば、ユーザのプリファレンスは、「スポーツ」の概念が重要であることを示す場合がある。この場合には、取り出されたマルチメディア文書、例えば、「ニュース」ビデオを処理して、「スポーツ」セグメントだけを抽出することができる。或いは、この適合は、例えば、「非スポーツ」セグメントを圧縮し、「スポーツ」セグメントからのハイライトを抽出することにより、コンテンツを要約することができる。

## 【0032】

本発明は、好ましい実施形態を参照して述べられた。当業者であれば、特許請求の範囲に述べられる本発明の精神及び範囲から離れることなく修正を実行することができることが明らかであろう。

## 【図面の簡単な説明】

10

20

30

40

50

## 【 0 0 3 3 】

【図1】照会プロセッサが検索のためにモデル・ベクトルの索引を用いるマルチメディア情報検索システムを示す。

【図2】語彙目録及び訓練された検出器を与えられた場合のマルチメディア文書に対するモデル・ベクトルの生成を示す。

【図3】マルチメディア文書に対するモデル・ベクトルを生成するための検出、採点、及びマッピング工程を示す。

【図4】検出器の採点に基づいて生成されたモデル・ベクトルの例を示す。

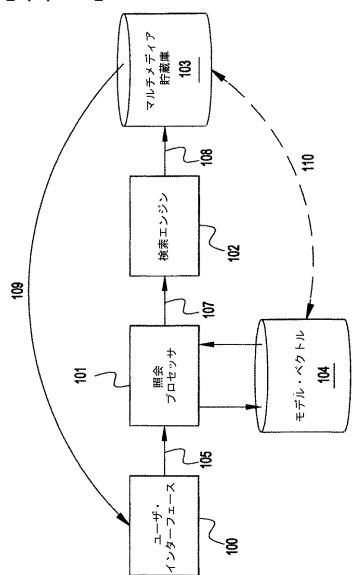
【図5】モデル・ベクトルを用いたマルチメディア文書の索引付け工程を示す。

【図6】モデル・ベクトルを用いた照会工程を示す。

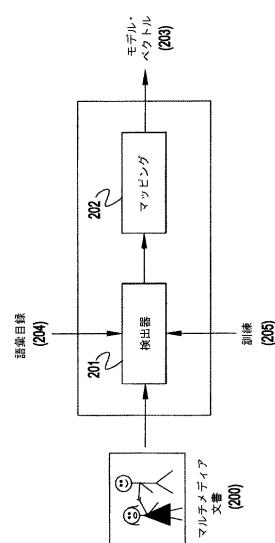
【図7】モデル・ベクトルを用いたマルチメディア文書の適合を示す。

10

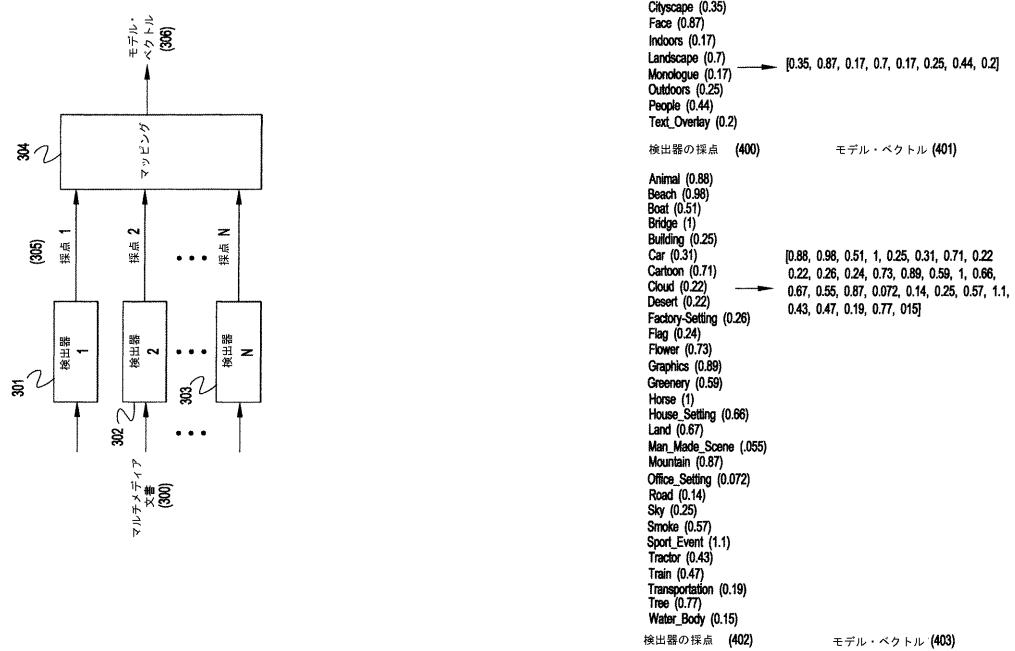
【図1】



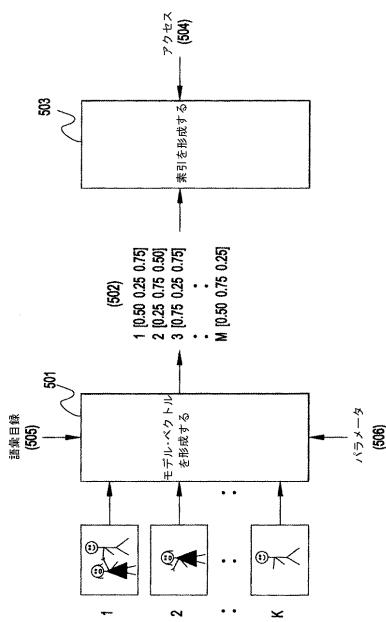
【図2】



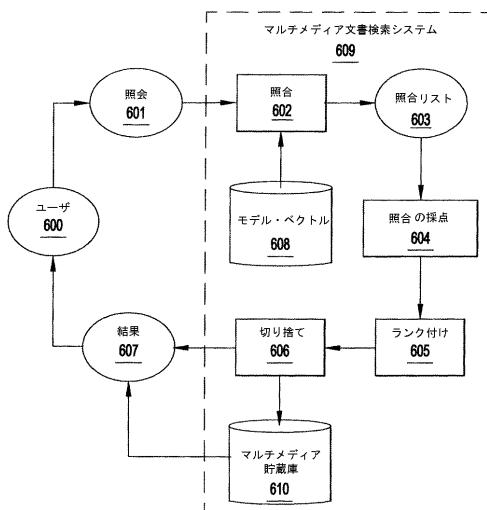
【 図 3 】



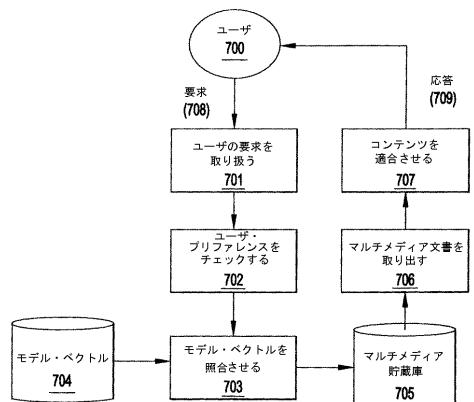
【図5】



【図6】



【図7】



## 【国際調査報告】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US03/39548															
<b>A. CLASSIFICATION OF SUBJECT MATTER</b> IPC(7) : G06F 17/30 US CL : 707/5, 104.1 345/589, 839 According to International Patent Classification (IPC) or to both national classification and IPC																	
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) U.S. : 707/5, 104.1 345/589, 839																	
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched																	
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)																	
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b> <table border="1"> <thead> <tr> <th>Category *</th> <th>Citation of document, with indication, where appropriate, of the relevant passages</th> <th>Relevant to claim No.</th> </tr> </thead> <tbody> <tr> <td>Y</td> <td>US 5,794,178 A(CAID et al) 11 August 1998, col. 1, lines 41-45, col. 2, lines 32-41, col. 6, lines 10-13.</td> <td>1-30.</td> </tr> <tr> <td>Y,E</td> <td>U.S Patent Application No. 20040073543 A1(KIM et al) 15 April 2004, Figs 5-7, paragraphs 0055-0058.</td> <td>1-30.</td> </tr> <tr> <td>A</td> <td>US 6,084,595 A(BACH et al) 4 July 2000, see entire references.</td> <td>1-30</td> </tr> <tr> <td>A</td> <td>US 6,405,166 B1(HUANG et al) 11 June 2002, see entire references.</td> <td>1-30</td> </tr> </tbody> </table>			Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	Y	US 5,794,178 A(CAID et al) 11 August 1998, col. 1, lines 41-45, col. 2, lines 32-41, col. 6, lines 10-13.	1-30.	Y,E	U.S Patent Application No. 20040073543 A1(KIM et al) 15 April 2004, Figs 5-7, paragraphs 0055-0058.	1-30.	A	US 6,084,595 A(BACH et al) 4 July 2000, see entire references.	1-30	A	US 6,405,166 B1(HUANG et al) 11 June 2002, see entire references.	1-30
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.															
Y	US 5,794,178 A(CAID et al) 11 August 1998, col. 1, lines 41-45, col. 2, lines 32-41, col. 6, lines 10-13.	1-30.															
Y,E	U.S Patent Application No. 20040073543 A1(KIM et al) 15 April 2004, Figs 5-7, paragraphs 0055-0058.	1-30.															
A	US 6,084,595 A(BACH et al) 4 July 2000, see entire references.	1-30															
A	US 6,405,166 B1(HUANG et al) 11 June 2002, see entire references.	1-30															
<input type="checkbox"/> Further documents are listed in the continuation of Box C.		<input type="checkbox"/> See patent family annex.															
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed																	
Date of the actual completion of the international search 03 September 2004 (03.09.2004)		Date of mailing of the international search report 15 SEP 2004															
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No. (703)305-3230		Authorized officer JOHN BREENE <i>Peggy Harrod</i> Telephone No. 703-305-3900															

---

フロントページの続き

(81)指定国 AP(BW,GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,HU,IE,IT,LU,MC,NL,PT,RO,SE,SI,SK,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NI,NO,NZ,OM,PH,PL,PT,RO,RU,SC,SD,SE,SG,SK,SL,TJ,TM,TN,TR,TT,TZ,UA,UG,UZ,VC,VN,YU,ZA,ZM,ZW

(72)発明者 スミス、ジョン

アメリカ合衆国11040ニューヨーク州、ニュー イリデ パーク、ファレル ストリート40

(72)発明者 ナファード、ミリンド

アメリカ合衆国61802イリノイ州アーバナ、オークランド アベニュー803、104号

(72)発明者 ナットセブ、アポスター

アメリカ合衆国10606ニューヨーク州、ホワイトプレーンズ、マーティン アベニュー25、904号

F ターム(参考) 5B075 ND16 NK06 NK31 NR12 NS00 PR06 QP01 QP05 QT00