

(19)日本国特許庁(JP)

(12)公表特許公報(A)

(11)公表番号
特表2024-535035
(P2024-535035A)

(43)公表日 令和6年9月26日(2024.9.26)

(51)国際特許分類
G 0 6 F 16/535(2019.01)
G 0 6 F 16/532(2019.01)

F I
G 0 6 F 16/535
G 0 6 F 16/532

テーマコード (参考)
5 B 1 7 5

審査請求 未請求 予備審査請求 未請求 (全29頁)			
(21)出願番号	特願2024-516634(P2024-516634)	(71)出願人	390009531
(86)(22)出願日	令和4年9月15日(2022.9.15)		インターナショナル・ビジネス・マシー
(85)翻訳文提出日	令和6年3月14日(2024.3.14)		ンズ・コーポレーション
(86)国際出願番号	PCT/EP2022/075650		INTERNATIONAL BUSI
(87)国際公開番号	WO2023/041648		NESS MACHINES CORPO
(87)国際公開日	令和5年3月23日(2023.3.23)		RATION
(31)優先権主張番号	17/479,172		アメリカ合衆国10504 ニューヨー
(32)優先日	令和3年9月20日(2021.9.20)		ク州 アーモンク ニュー オーチャード
(33)優先権主張国・地域又は機関	米国(US)		ロード
(81)指定国・地域	AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA ,RW,SD,SL,ST,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,RU,TJ,TM),EP(AL,A T,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR ,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC, 最終頁に続く	(74)代理人	100112690 弁理士 太佐 種一
		最終頁に続く	

(54)【発明の名称】 複合画像のパーソナライズされた画像構成要素に基づく 認知画像検索

(57)【要約】

本発明の実施形態は、電子検索を実行するコンピュータ実装方法を対象とする。コンピュータ実装方法は、プロセッサを使用して、複数の電子的に識別可能なオブジェクトを含む複合電子画像を受信することを含み、複合電子画像がユーザに関連付けられている。プロセッサを使用して、複数の電子的に識別可能なオブジェクトのそれぞれについて少なくとも1つのサブ画像を提供することによって、複合電子画像をサブ画像にセグメント化する。サブ画像のそれぞれに対して、プロセッサを使用して、パーソナライズされたサブ画像検索動作を実行する。パーソナライズされたサブ画像検索動作は、サブ画像の中から検索対象サブ画像を選択することと、検索対象サブ画像をユーザのパーソナライズされたメタデータと関連付けることと、ユーザのパーソナライズされたメタデータに少なくとも部分的に基づいて、データベースを検索して検索画像のセットを返すことと、を含む。

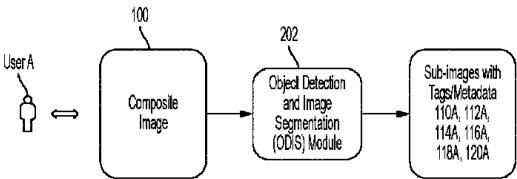


FIG. 2

【特許請求の範囲】

【請求項 1】

電子検索を実行するコンピュータ実装方法であって

プロセッサを使用して、複数の電子的に識別可能なオブジェクトを含む複合電子画像を受信することであり、前記複合電子画像がユーザに関連付けられている、前記受信することと、

前記プロセッサを使用して、前記複数の電子的に識別可能なオブジェクトのそれぞれについて少なくとも1つのサブ画像を提供することによって、前記複合電子画像を前記サブ画像にセグメント化することと、

前記サブ画像のそれぞれについて、前記プロセッサを使用して、パーソナライズされたサブ画像検索動作を実行することであり、前記パーソナライズされたサブ画像検索動作が、

前記サブ画像の中から検索対象サブ画像を選択すること、

前記検索対象サブ画像を前記ユーザのパーソナライズされたメタデータに関連付けること、および

前記ユーザの前記パーソナライズされたメタデータに少なくとも部分的に基づいて、データベースを検索して検索画像のセットを返すことを含む、前記実行することと、

を含むコンピュータ実装方法。

【請求項 2】

前記ユーザの前記パーソナライズされたメタデータが、

前記検索対象サブ画像の前記ユーザに対する関連性の記述と、

前記検索対象サブ画像の前記複合画像に対する相対サイズと、

前記複合画像内の前記検索対象サブ画像の相対位置と、

を含む、請求項 1 に記載のコンピュータ実装方法。

【請求項 3】

前記ユーザの前記パーソナライズされたメタデータが、前記ユーザに関する情報のコーパスから導出される、請求項 2 に記載のコンピュータ実装方法。

【請求項 4】

前記パーソナライズされたサブ画像検索動作が、

前記検索対象サブ画像に関連性レベルを割り当てることであって、

前記関連性レベルが、前記ユーザに対する前記検索対象サブ画像の関連性を含む、前記割り当てることと、

前記関連性レベルに少なくとも部分的に基づいて、前記検索対象サブ画像を前記サブ画像に対してランク付けすることと、

をさらに含む、請求項 1 から 3 のいずれかに記載のコンピュータ実装方法。

【請求項 5】

前記パーソナライズされたサブ画像検索動作が、

前記パーソナライズされたサブ画像検索動作の結果を前記ユーザに提示することと、

前記パーソナライズされたサブ画像検索動作の前記結果に関する前記ユーザからのフィードバックを受信することと、

をさらに含む、請求項 1 から 4 のいずれかに記載のコンピュータ実装方法。

【請求項 6】

前記フィードバックに基づいて、前記検索対象サブ画像について前記パーソナライズされたサブ画像検索動作を繰り返すことをさらに含む、請求項 5 に記載のコンピュータ実装方法。

【請求項 7】

前記プロセッサが、

前記ユーザに関する情報の前記コーパスに基づいて、前記ユーザの前記パーソナライズされたメタデータを生成し、

10

20

30

40

50

前記ユーザに関する情報の前記コーパスに少なくとも部分的に基づいて、前記ユーザに対する前記検索対象サブ画像の前記関連性を決定する、
ように訓練された分類器を備える、請求項 3 に記載のコンピュータ実装方法。

【請求項 8】

電子検索を実行するためのコンピュータ・システムであって、前記コンピュータ・システムがプロセッサに通信可能に結合されたメモリを備え、前記プロセッサが、

複数の電子的に識別可能なオブジェクトを含む複合電子画像を受信することであり、前記複合電子画像がユーザに関連付けられている、前記受信することと、

前記複数の電子的に識別可能なオブジェクトのそれぞれについて少なくとも 1 つのサブ画像を提供することによって、前記複合電子画像を前記サブ画像にセグメント化することと、

前記サブ画像のそれぞれについて、パーソナライズされたサブ画像検索動作を実行することであり、

前記サブ画像の中から検索対象サブ画像を選択すること、

前記検索対象サブ画像を前記ユーザのパーソナライズされたメタデータに関連付けること、および

前記ユーザの前記パーソナライズされたメタデータに少なくとも部分的に基づいて、データベースを検索して検索画像のセットを返すこと、

を含む、前記パーソナライズされたサブ画像検索動作を実行することと、

を含むプロセッサ動作を実行するように構成されている、

コンピュータ・システム。

【請求項 9】

前記ユーザの前記パーソナライズされたメタデータが、前記ユーザに対する前記検索対象サブ画像の関連性の記述を含む、請求項 8 に記載のコンピュータ・システム。

【請求項 10】

前記ユーザの前記パーソナライズされたメタデータが、前記ユーザに関する情報のコーパスから導出される、請求項 9 に記載のコンピュータ・システム。

【請求項 11】

前記パーソナライズされたサブ画像検索動作が、

前記検索対象サブ画像に関連性レベルを割り当てることであって、

前記関連性レベルが、前記ユーザに対する前記検索対象サブ画像の関連性を含む、前記割り当てることと、

前記関連性レベルに少なくとも部分的に基づいて、前記検索対象サブ画像を前記サブ画像に対してランク付けすることと、

をさらに含む、請求項 8 から 10 のいずれかに記載のコンピュータ・システム。

【請求項 12】

前記パーソナライズされたサブ画像検索動作が、

前記パーソナライズされたサブ画像検索動作の結果を前記ユーザに提示することと、

前記パーソナライズされたサブ画像検索動作の前記結果に関する前記ユーザからのフィードバックを受信することと、

をさらに含む、請求項 8 から 11 のいずれかに記載のコンピュータ・システム。

【請求項 13】

前記プロセッサ動作が、前記フィードバックに基づいて、前記検索対象サブ画像について前記パーソナライズされたサブ画像検索動作を繰り返すことをさらに含む、請求項 12 に記載のコンピュータ・システム。

【請求項 14】

前記プロセッサが、

前記ユーザに関する情報の前記コーパスに基づいて、前記ユーザの前記パーソナライズされたメタデータを生成し、

前記ユーザに関する情報の前記コーパスに少なくとも部分的に基づいて、前記ユーザに

10

20

30

40

50

対する前記検索対象サブ画像の前記関連性を決定する、
ように訓練された分類器を備える、請求項 10 に記載のコンピュータ・システム。

【請求項 15】

電子検索を実行するためのコンピュータ・プログラム製品であって、前記コンピュータ・プログラム製品がコンピュータ可読記憶媒体に記憶されたコンピュータ可読プログラムを備え、前記コンピュータ可読プログラムがプロセッサ上で実行されると、前記プロセッサに、

複数の電子的に識別可能なオブジェクトを含む複合電子画像を受信することであり、前記複合電子画像がユーザに関連付けられている、前記受信することと、

前記複数の電子的に識別可能なオブジェクトのそれぞれについて少なくとも 1 つのサブ画像を提供することによって、前記複合電子画像を前記サブ画像にセグメント化することと、

前記サブ画像のそれぞれについて、パーソナライズされたサブ画像検索動作を実行することであり、

前記サブ画像の中から検索対象サブ画像を選択すること、

前記検索対象サブ画像を前記ユーザのパーソナライズされたメタデータに関連付けること、および

前記ユーザの前記パーソナライズされたメタデータに少なくとも部分的に基づいて、データベースを検索して検索画像のセットを返すこと、

を含む、パーソナライズされたサブ画像検索動作を実行することと、

を含む方法を実行させる、

コンピュータ・プログラム製品。

【請求項 16】

前記ユーザの前記パーソナライズされたメタデータが、前記ユーザに対する前記検索対象サブ画像の関連性の記述を含む、請求項 15 に記載のコンピュータ・プログラム製品。

【請求項 17】

前記ユーザの前記パーソナライズされたメタデータが、前記ユーザに関する情報のコーパスから導出される、請求項 16 に記載のコンピュータ・プログラム製品。

【請求項 18】

前記パーソナライズされたサブ画像検索動作が、

前記検索対象サブ画像に関連性レベルを割り当てることであって、

前記関連性レベルが、前記ユーザに対する前記検索対象サブ画像の関連性を含む、前記割り当てることと、

前記関連性レベルに少なくとも部分的に基づいて、前記検索対象サブ画像を前記サブ画像に対してランク付けすることと、

をさらに含む、請求項 15 から 17 のいずれかに記載のコンピュータ・プログラム製品。

【請求項 19】

前記パーソナライズされたサブ画像検索動作が、

前記パーソナライズされたサブ画像検索動作の結果を前記ユーザに提示することと、

前記パーソナライズされたサブ画像検索動作の前記結果に関する前記ユーザからのフィードバックを受信することと、

をさらに含む、

前記プロセッサ動作が、前記フィードバックに基づいて、前記検索対象サブ画像について前記パーソナライズされたサブ画像検索動作を繰り返すことをさらに含む、

請求項 15 から 18 のいずれかに記載のコンピュータ・プログラム製品。

【請求項 20】

前記プロセッサが、

前記ユーザに関する情報の前記コーパスに基づいて、前記ユーザの前記パーソナライズされたメタデータを生成し、

前記ユーザに関する情報の前記コーパスに少なくとも部分的に基づいて、前記ユーザに

10

20

30

40

50

対する前記検索対象サブ画像の前記関連性を決定する、
ように訓練された分類器を備える、請求項 17 に記載のコンピュータ・プログラム製品
【請求項 21】

コンピュータ・プログラムであって、前記プログラムがコンピュータ上で実行されると、
請求項 1 から 7 のいずれかに記載の方法を実行するように適合されたプログラム・コード
手段を含む、コンピュータ・プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、一般にプログラマブル・コンピュータに関する。より詳細には、本発明は、
複合画像のパーソナライズされた画像構成要素またはサブ画像に基づいて画像検索を認知
的に実行するコンピューティング・システム、コンピュータ実装方法、およびコンピュ
ータ・プログラム製品に関する。

【背景技術】

【0002】

オンライン検索エンジンは、ユーザが検索クエリではなく主に画像に基づいていわゆる
画像検索を実行することを可能にする検索機能を含む。「逆画像検索」として知られる技
術は、コンテンツ・ベースの画像検索 (content-based image re
trieval、CBIR) クエリ技術であり、事実上、画像ベースの検索クエリとして
使用されるサンプル画像を CBIR システムに提供することを含む。逆画像検索は、検索
用語がないことが特徴であるため、ユーザは、正しい結果を返したり、返さなかったりす
るキーワードまたは用語を推測する必要がなくなる。逆画像検索により、ユーザは、特定
のサンプル画像に関連するコンテンツ、すなわち、画像の人気度、操作されたバージョン
、派生作品などを見つけることができる。

【0003】

複合画像とは、複数の異なる識別可能なオブジェクトを含む画像である。例えば、単一
の複合画像が、建物、建物の前を通過する車、建物に歩いて入る 2 人の人物、建物の隣に
ある木などを含み得る。オブジェクト検出は、コンピュータ・ビジョンおよび画像処理に
関連するコンピュータ技術であり、デジタル画像およびビデオ内の特定のクラスの意味オ
ブジェクト (例えば、人間、建物、または車) のインスタンスを検出することを扱う。オ
ブジェクト検出は、画像アノテーション、車両の計数、およびアクティビティ認識などの
コンピュータ・ビジョン・タスクにおいて広く使用されている。自動画像アノテーション
は、コンピュータ・システムがデジタル画像にキャプションまたはキーワードの形態でメ
タデータを自動的に割り当てるプロセスであり、これにより、自動画像アノテーションを
画像検索システムで使用して、データベースから関心のある画像を検索し、整理し、特定
することが可能になる。

【発明の概要】

【0004】

本発明の実施形態は、電子検索を実行するコンピュータ実装方法を対象とする。コンピ
ュータ実装方法は、プロセッサを使用して、複数の電子的に識別可能なオブジェクトを含
む複合電子画像を受信することを含み、複合電子画像がユーザに関連付けられている。プ
ロセッサを使用して、複数の電子的に識別可能なオブジェクトのそれぞれについて少なく
とも 1 つのサブ画像を提供することによって、複合電子画像をサブ画像にセグメント化す
る。サブ画像のそれぞれに対して、プロセッサを使用して、パーソナライズされたサブ画
像検索動作を実行する。パーソナライズされたサブ画像検索動作は、サブ画像の中から検
索対象サブ画像を選択することと、検索対象サブ画像をユーザのパーソナライズされたメ
タデータと関連付けることと、ユーザのパーソナライズされたメタデータに少なくとも部
分的に基づいて、データベースを検索して検索画像のセットを返すことと、を含む。

【0005】

本発明の実施形態は、上述のコンピュータ実装方法と実質的に同じ特徴を有するコンピ

10

20

30

40

50

ュータ・システムおよびコンピュータ・プログラム製品も対象とする。

一態様によると、電子検索を実行するためのコンピュータ・システムが提供され、本コンピュータ・システムは、プロセッサに通信可能に結合されたメモリを備え、プロセッサは、複数の電子的に識別可能なオブジェクトを含む複合電子画像を受信することであって、複合電子画像がユーザに関連付けられている、受信することと、複数の電子的に識別可能なオブジェクトのそれぞれについて少なくとも1つのサブ画像を提供することによって、複合電子画像をサブ画像にセグメント化することと、サブ画像のそれぞれについて、パーソナライズされたサブ画像検索動作を実行することであって、サブ画像の中から検索対象サブ画像を選択すること、検索対象サブ画像をユーザのパーソナライズされたメタデータと関連付けること、およびユーザのパーソナライズされたメタデータに少なくとも部分的に基づいて、データベースを検索して検索画像のセットを返すことを含む、実行することと、を含むプロセッサ動作を実行するように構成されている。

10

【0006】

別の態様によると、電子検索を実行するためのコンピュータ・プログラム製品が提供され、コンピュータ・プログラム製品は、コンピュータ可読記憶媒体に記憶されたコンピュータ可読プログラムを含み、コンピュータ可読プログラムは、プロセッサ上で実行されると、プロセッサに、複数の電子的に識別可能なオブジェクトを含む複合電子画像を受信することであって、複合電子画像がユーザに関連付けられている、受信することと、複数の電子的に識別可能なオブジェクトのそれぞれについて少なくとも1つのサブ画像を提供することによって、複合電子画像をサブ画像にセグメント化することと、サブ画像のそれぞれについて、パーソナライズされたサブ画像検索動作を実行することであって、サブ画像の中から検索対象サブ画像を選択すること、検索対象サブ画像をユーザのパーソナライズされたメタデータと関連付けること、およびユーザのパーソナライズされたメタデータに少なくとも部分的に基づいて、データベースを検索して検索画像のセットを返すこと、を含む、パーソナライズされたサブ画像検索動作を実行することと、を含む方法を実行させる。

20

【0007】

追加の特徴および利点は、本明細書に記載された技術によって実現される。他の実施形態および態様は、本明細書に詳細に記載されている。より良く理解するために、説明および図面を参照されたい。

30

【0008】

次に、本発明の好ましい実施形態を、単なる例として、以下の図面を参照して説明する。

【図面の簡単な説明】

【0009】

【図1】本発明の実施形態によるパーソナライズされたサブ画像検索システムに入力することができる複合画像を示す図である。

【図2】本発明の実施形態によるパーソナライズされたサブ画像検索システムにおいて使用することができるオブジェクト検出および画像セグメント化モジュールを示す図である。

40

【図3】本発明の実施形態によるパーソナライズされたサブ画像検索システムで使用することができる画像構成要素認知検索モジュールを示す図である。

【図4A】本発明の実施形態によるパーソナライズされたサブ画像検索システムにおいて使用することができる画像構成要素認知検索モジュールを示す図である。

【図4B】本発明の実施形態に従って生成されたパーソナライズされたタグ／メタデータを有するサブ画像の例を示す図である。

【図5】本発明の実施形態による方法を示す流れ図である。

【図6A】本発明の実施形態によるパーソナライズされたサブ画像検索システムを示す、ブロック図と流れ図を組み合わせた図である。

【図6B】本発明の実施形態による、図6Aに示されたシステムおよび流れ図によって利

50

用される式を示す図である。

【図 7】本発明の実施形態を実施するために利用することができる機械学習システムを示す図である。

【図 8】本発明の実施形態による、図 7 に示された機械学習システムによって実施され得る学習段階を示す図である。

【図 9】本発明の様々な実施形態を実施することができる例示的なコンピューティング・システムの詳細を示す図である。

【発明を実施するための形態】

【0010】

添付の図および開示された実施形態の以下の詳細な説明において、図に示された様々な要素には、3桁の参照番号が付されている。場合によっては、各参照番号の左端の数字は、その要素が最初に示されている図に対応する。

【0011】

簡潔にするために、本発明の態様を作成および使用することに関連する従来技術は、本明細書で詳細に説明する場合もあれば、説明しない場合もある。特に、本明細書に記載される様々な技術的特徴を実施するためのコンピューティング・システムおよび特定のコンピュータ・プログラムの様々な態様は、よく知られている。したがって、簡潔にするために、多くの従来実施態様の詳細は、本明細書では簡単に言及されるだけであり、あるいはよく知られているシステムもしくはプロセスまたはその両方の詳細を提供することなく完全に省略される。

【0012】

本明細書に記載されるシステムの機能ユニットの多くは、モジュールとしてラベル付けされている。本発明の実施形態は、多種多様なモジュールの実装に適用される。例えば、モジュールは、カスタム VLSI 回路もしくはゲート・アレイ、ロジック・チップなどの既製の半導体、トランジスタ、または他の個別部品を含むハードウェア回路として実装することができる。モジュールは、フィールド・プログラマブル・ゲート・アレイ、プログラマブル・アレイ・ロジック、プログラマブル・ロジック・デバイスなどのプログラマブル・ハードウェア・デバイスで実装することもできる。モジュールは、様々なタイプのプロセッサによる実行のためにソフトウェアで実装することもできる。実行可能コードの識別されたモジュールは、例えば、オブジェクト、プロシージャ、または関数として編成することができる。それにもかかわらず、識別されたモジュールの実行可能ファイルは、物理的に一緒に配置されている必要はなく、論理的に結合されると、モジュールとして機能し、モジュールの規定された目的を達成する、異なる場所に格納された異なる命令を含むことができる。

【0013】

本明細書に示されるシステムの様々な構成要素、モジュール、サブ機能などは、図示および説明を容易にするために別々に示されている。本発明の実施形態では、様々な構成要素、モジュール、サブ機能などによって実行される動作は、特に別段の記載がない限り、本明細書に記載される本発明の様々な実施形態の範囲から逸脱することなく、図示されたものとは異なるように分散させることができる。

【0014】

便宜上、本明細書に記載される技術的動作の一部は、非公式な表現を用いて伝えられる。例えば、キャッシュ・メモリにデータが記憶されたプロセッサは、そのデータを「知っている」プロセッサとして説明され得る。同様に、ユーザがロード・データ・コマンドをプロセッサに送信することは、ユーザがプロセッサにデータをロードするように「伝える」こととして説明され得る。この詳細な説明における任意のそのような非公式な表現は、非公式な表現がより公式的で技術的な機能および動作に対応していることをカバーしていると読まれるべきであり、当業者は、そのような非公式の表現が、非公式な表現がより公式的で技術的な機能および動作に対応していることをカバーしていることを理解するであ

10

20

30

40

50

ることが理解される。

【0015】

本明細書で提供される説明は、「画像」を参照する。「画像」という用語の使用は、特にそうではないと記載されない限り、コンピュータによって分析され、メモリに記憶され、電子的に送信され、コンピュータ・ディスプレイに表示され得る、画像の電子的またはデジタル表現を指すことを理解されたい。

【0016】

ここで、本発明の実施形態に関連する技術の概要に目を向けると、本明細書で前述したように、複数の異なるオブジェクトが多くの場合所与の画像内に表示される。例えば、単一の画像は、建物、建物の前を通過する車、建物に歩いて入る2人の人物、建物の隣の木などを含むことがある。複数の識別可能なオブジェクトが描写された画像は、複合画像と呼ばれる。複合画像の場合、既知のオンライン検索エンジンは、複合画像全体に類似した画像を探すことによって画像検索を行う。しかしながら、ユーザが複合画像内の1つまたは複数のオブジェクトのみを検索することに関心がある場合、既知の画像検索技術では、ユーザは複数の編集ステップを行って、ユーザにとって関心のあるオブジェクトが主要オブジェクトである新しい画像を作成し、その後編集された画像を検索する必要がある。コンピュータと対話する際の「ユーザ・エクスペリエンス」(UX)は、システムに特定のタスクを実行させるためにユーザがシステムに入力する必要がある情報量によって影響を受けることが知られている。

【0017】

ここで本発明の態様の概要に目を向けると、本発明の実施形態は、ユーザが複合画像内の特定のオブジェクトに焦点を当てた画像検索を行いたい状況におけるUXを改善する。本発明の実施形態は、複合画像のコンピュータ生成されパーソナライズされた画像構成要素または複合画像サブ画像に基づいて画像検索を認知的に実行するコンピューティング・システム、コンピュータ実装方法、およびコンピュータ・プログラム製品を提供する。本発明の実施形態によると、コンピューティング・システムは、画像検索要求を提出する際に、ユーザが関心のあるオブジェクトを特定するためのアクションを取ることを要求することなく、ユーザにとって関心のある複合画像内のオブジェクトを認知的に決定するように構成される。例えば、本発明の実施形態では、複合画像内の1つまたは複数のオブジェクトに焦点を当てた画像検索を実行したいユーザが、ユーザにとって関心のあるオブジェクトが主要オブジェクトである新しい画像を作成するために、複数の編集ステップを行う必要がない。

【0018】

本発明の実施形態では、ユーザからの複合画像および画像検索要求の受信に応答して、コンピュータ・システムは、複合画像に対してオブジェクト検出および画像セグメント化プロセスを自動的に実行して、複合画像内の識別可能なオブジェクトを検出し、複合画像をサブ画像にセグメント化し、各サブ画像が識別可能なオブジェクトのうちの少なくとも1つに対応する。任意選択で、自動画像アノテーションをサブ画像に適用して、記述的メタデータのサブ画像への初期割り当てを生成することができる。認知プロセッサは、サブ画像と、任意選択で、メタデータの初期割り当てとを受信する。本発明の実施形態によると、認知プロセッサは、画像処理機能および表現ベースの自然言語処理機能を備える。自然言語処理機能は、IBM Watson (R)などのロバストな表現ベースの認知データ分析技術を使用して実装することができる。IBM Watson (R)は、自然言語を理解し、証拠に基づいて仮説を生成し、その過程で学習することによって、コンピュータよりも人間のように情報を処理する、表現ベースの認知データ分析技術である。加えて、表現ベースの認知コンピュータ分析は、「キーワード」よりも「単語パターン」の方がより柔軟な検索能力があること、および表現ベースの認知データ分析によって処理することができるデータ量が非常に多いことなどを含む、いくつかの理由で、キーワードベースのコンピュータ分析よりも優れた計算能力を提供する。

【0019】

本発明の実施形態による認知プロセッサは、サブ画像、任意選択の初期メタデータ、およびユーザのコーパスを分析して、ユーザに対するサブ画像の関連性のレベルを決定し、その関連性のレベルを自然言語で捕捉し、関連性のレベルをサブ画像のメタデータに組み込んで、サブ画像のパーソナライズされたメタデータを作成するという第1の認知分析タスクを実行する。本発明の一部の実施形態では、初期メタデータは、ユーザに対するサブ画像の関連性のレベルを決定するタスクを実行する際に、認知プロセッサを補強または支援するために使用することができる。

【0020】

画像ベースの検索エンジンは、各サブ画像について画像検索結果のセットが生成されるように、各サブ画像およびそれに関連付けられたパーソナライズされたメタデータに対して画像検索を実行する。認知プロセッサは、各サブ画像、各サブ画像に関連付けられたパーソナライズされたメタデータ、および任意選択でユーザのコーパスを分析して、各サブ画像をユーザに対する関連性レベル（または重要度レベル）に基づいてランク付けするという第2の認知タスクを実行する。本発明の実施形態では、各サブ画像の関連性スコアは、複合画像に対するサブ画像の相対的なサイズ、および複合画像内のサブ画像の相対的な位置の関数とすることができる。ランク付けされた各サブ画像およびそれに関連付けられた検索結果のセットは、例えば、コンピュータ・ディスプレイを使用して、レビューのためにユーザに提示することができる。本発明の一部の実施形態では、認知プロセッサは、しきい値を上回るランキング・レベル（または重要度レベル）を有するサブ画像のみを表示するように構成され得る。

10

20

【0021】

本発明の一部の実施形態では、ユーザは、検索結果に関するユーザ・フィードバックを認知プロセッサに提供することができ、ユーザ・フィードバックは、記憶され、第1および第2の認知プロセッサ・タスクの将来の実行を増強または改善するために使用することができる。本発明の一部の実施形態では、ユーザ・フィードバックは、ユーザが表示された検索結果とどのように対話するかから導出することができる。例えば、ユーザが、他のサブ画像の検索結果をクリックすることなく、4番目にランク付けされたサブ画像およびそれに関連付けられた検索結果を即座にクリックした場合、認知プロセッサは、4番目にランク付けされたサブ画像のランクが低すぎたと判断することができる。ユーザが、他のサブ画像の検索結果をクリックすることなく、最上位にランク付けされたサブ画像およびそれに関連付けられた検索結果を即座にクリックした場合、認知プロセッサは、最上位にランク付けされたサブ画像が適切にランク付けされたと判断することができる。本発明の一部の実施形態では、認知プロセッサは、ディスプレイを介してランキングに関する質問をユーザに提示することによって、ユーザ・フィードバックを直接求めることができる。例えば、認知プロセッサは、認知プロセッサによってランク付けされた上位4つのサブ画像のユーザのランク付けをディスプレイで入力するようにユーザに求めることができる。

30

【0022】

本発明の一部の実施形態では、認知プロセッサは、ユーザ・フィードバックを評価して、ユーザ・フィードバックが現在の画像検索の品質を改善するかどうかを判定することができる。認知プロセッサが、ユーザ・フィードバックによって現在の画像検索を改善できると判断した場合、認知プロセッサは、ユーザ・フィードバックに基づいて、その第1および第2の認知タスクを更新し、次いで、画像検索を繰り返すことができる。本発明の一部の実施形態では、上述した画像検索の繰り返しは、オプションとしてユーザに提供され、ユーザがユーザ承認を入力した場合にのみ実行され得る。

40

【0023】

本発明の一部の実施形態では、第1および第2の認知タスクは、サブ画像が検索される前にランク付けされるように、画像検索に先立って行うことができる。本発明の一部の実施形態では、第1および第2の認知タスクは、ユーザが、検索画像とともに、ユーザにとって関心のある複合画像内のオブジェクトの自然言語による識別を入力することによって、さらに補強することができる。例えば、ユーザは、複合画像と、「左下隅の花」という

50

自然言語テキストとを含む画像検索要求を提出することができる。認知プロセッサには自然言語処理機能が含まれているため、関心のあるオブジェクトの自然言語による識別に特定のフォーマットを必要とする必要性がないため、UXへの影響は最小限に抑えられる。認知プロセッサは、その自然言語処理機能を使用して、テキスト入力の意味を解釈し、その意味を使用して、複合画像の左下隅の花が、オブジェクト検出プロセスによって識別されたサブ画像の中に確実に含まれるようにする。認知プロセッサはまた、テキスト入力の意味を使用して、花を示すサブ画像に適切なランキングを適用する。

【0024】

本発明の一部の実施形態では、認知プロセッサは、画像処理アルゴリズム、機械学習アルゴリズム、および自然言語処理アルゴリズムを有する訓練された分類器を使用して、そのタスクおよび他の認知または評価動作を行うことができる。

10

【0025】

本発明の一部の実施形態では、認知プロセッサの自然言語処理機能は、自然言語の質問に対する回答を提供する既知のタイプのQ & Aシステムの修正版であるパーソナライズされたQ & A機能を含むことができる。非限定的な例として、認知プロセッサは、IBM (R) によって開発されたDeepQA技術の特徴および機能のすべてを含むことができる。DeepQAは、データ・リポジトリに問い合わせ、自然言語処理、機械学習、情報検索、仮説生成、仮説スコアリング、最終ランキング、および回答マージの要素を適用して結論に達することで、自然言語の質問に答えるQ & Aシステムである。このようなQ & Aシステムは、教育環境の自然な質疑応答パラダイムなど、ある種の意味論的なクエリおよび検索動作で人間を支援することができる。IBM (R) のDeepQA技術などのQ & Aシステムでは、多くの場合、非構造化情報管理アーキテクチャ(unstructured information management architecture、UIMA) が使用され、このUIMAは、非構造化情報の分析およびIBM (R) によって開発された検索技術との統合のためのマルチ・モーダル・アナリティクスの開発、発見、構成、および展開のためのコンポーネント・ソフトウェア・アーキテクチャである。認知プロセッサ・タスクに適用される場合、Q & A機能は、ユーザに対する所与のサブ画像の関連性は何か、またはユーザに対する各サブ画像の関連性に基づくサブ画像の適切なランク付けは何かなどの問い合わせに答えるために使用することができる。

20

【0026】

次に、本発明の実施形態のより詳細な説明に目を向けると、図1は、本発明の態様による画像構成要素認知検索モジュール302 (図3に示す) によって実行される分析および画像検索の対象となり得る複合画像100を示す。複合画像は、図示するように構成および配置された、飛行機112、マンション112、複数の植木鉢114、2人の人物116、看板(売り物件、貸し物件など) 118、および木120を含む複数のオブジェクトを含む。

30

【0027】

図2は、オブジェクト検出および画像セグメント化(object detection and image segmentation、ODIS) モジュール202を示す。本発明の一部の実施形態では、ODISモジュール202は、認知検索モジュール302 (図3に示す) 内に組み込むことができ、複合画像100に対してオブジェクト検出および画像セグメント化動作を行うように構成されている。ODISモジュール202は、ユーザAから複合画像100を受信し、複合画像100内の電子的に識別可能なオブジェクトを検出し、複合画像100をサブ画像112A、114A、115A、118A、120A (図3に示す) にセグメント化し、各サブ画像は電子的に識別可能なオブジェクトの少なくとも1つに対応する。一般に、オブジェクトは、選択された粒度レベルでオブジェクトを電子的に認識し分類することができる場合、電子的に識別可能である。例えば、単一の葉が電子的に識別可能であっても、ODISモジュール202の粒度は、木はオブジェクトとして識別されるが、木の各個々の葉は識別されないように設定することができる。任意選択で、ODISモジュール202は、タグもしくは記述的メタデータまたは

40

50

その両方の初期割り当てをサブ画像 112 A、114 A、116 A、118 A、120 A に適用するために使用することができる自動画像アノテーション機能を含むことができる。

【0028】

図 3 は、認知検索モジュール 302 と、サブ画像 112 A、114 A、116 A、118 A、120 A、ユーザ A コーパス 320、および他のユーザ A コンテキスト & 調整 (other User A context & adjustments、OUCA) 330 を含む、認知画像検索モジュールへの入力とを示す。ユーザ A コーパス 320 は、ユーザ・プロフィール 322 およびユーザ・アクティビティ 324 を含む。ユーザ・プロフィール 322 は、ユーザ A によって作成され、ユーザ A に関連付けられた設定および情報の集合である。ユーザ・プロフィール 322 には、ユーザ A の名前、年齢、写真、および知識または専門知識などの個人のの特徴など、ユーザ A を識別するために使用される重要な情報が含まれる。ユーザ・プロフィール 322 は、ユーザ A のソーシャル・メディア・サイトでユーザ A が使用しているプロフィールからダウンロードすることができる。本発明の一部の実施形態では、ユーザ・プロフィール 322 は、特に、職業、趣味、関心、音楽の好み、好きな著者、読んだ本などのプロフィール情報を含む、パーソナライズされたタグ 432、442 およびパーソナライズされたメタデータ 434、444 (すべて図 4 B に示されている) を構築するのを支援するプロフィール情報をユーザ A から引き出すように構築することができる。ユーザ・プロフィール 322 の情報は、ユーザ A によって自動的に提出される。OUCA 330 は、焦点、複合画像 100 のサイズ、および複合画像 100 内のサブ画像の目立ち具合 (prominence) などの「入力画像プロパティ」を含むことができる。OUCA 330 は、オブジェクトが複合画像 100 の前面であるか背景であるかをさらに含むことができる。OUCA 330 は、現在のパーソナライズされたサブ画像検索結果 312 に関するユーザ A からのフィードバックをさらに含むことができる。OUCA 330 は、認知プロセッサ 202 によって行われた過去の複合画像検索、ならびに現在の複合画像検索と他の過去の複合画像検索との間の任意の重複 (例えば、共通のサブ画像) をさらに含むことができる。本発明の実施形態によると、認知検索モジュール 302 は、様々な入力 (110 A、112 A、114 A、116 A、118 A、120 A、320、330) を分析して、パーソナライズされたサブ画像検索結果 312 を生成する。方法 500 (図 5 に示す) は、本発明の実施形態による、パーソナライズされたサブ画像検索結果 312 を生成するために認知検索モジュール 302 によって実行される動作を示す。方法 500 は、図 5 の説明に関連して本明細書において引き続きさらに詳細に説明される。方法 600 (図 6 A に示す) は、本発明の実施形態による、パーソナライズされたサブ画像検索結果 312 を生成するために認知検索モジュール 302 によって実行される動作を示す。方法 600 は、図 6 A の説明に関連して本明細書において引き続きさらに詳細に説明される。

【0029】

図 4 A は、本発明の実施形態による認知検索モジュール 302 A を示す。認知検索モジュール 302 A は、認知検索モジュール 302 (図 3 に示す) によって実行される動作のすべてを実行することができるが、認知検索モジュール 302 A が本発明の実施形態に従ってどのように実装され得るかの追加的な詳細を提供する。認知検索モジュール 302 A は、図示されるように構成および配置された、ODIS モジュール 202 と、画像処理モジュール 402 と、認知プロセッサ 404 と、検索エンジンとを含む。モジュール 202、402、404、405 のすべては、本発明の実施形態に従ってそのモジュールの機能を実行するために、必要に応じて使用することができる表現ベースの自然言語処理機能を含む。自然言語処理機能は、IBM Watson (R) などのロバストな表現ベースの認知データ分析技術を使用して実装することができる。IBM Watson (R) は、自然言語を理解し、証拠に基づいて仮説を生成し、その過程で学習することによって、コンピュータよりも人間のように情報を処理する、表現ベースの認知データ分析技術である。加えて、表現ベースの認知コンピュータ分析は、「キーワード」よりも「単語パターン

」の方がより柔軟な検索能力があること、および表現ベースの認知データ分析によって処理することができるデータ量が非常に多いことなどを含む、いくつかの理由で、キーワードベースのコンピュータ分析よりも優れた計算能力を提供する。

【0030】

図4Aに示すODIS202は、図2に示すODIS202と同じ特徴および機能を含む。本発明の一部の実施形態では、ODIS202は、認知検索モジュール302Aの外部にあってもよく、または認知検索モジュール302A内に統合されてもよい。画像処理モジュール402は、サブ画像110A、112A、114A、116A、118A、120Aが生成された後に、認知プロセッサ404によって実行される分析のための画像処理を提供する。認知プロセッサ404は、パーソナライズされたタグ／メタデータを有するサブ画像430（図4に示す）と、パーソナライズされたタグ／メタデータおよびユーザ検索ガイダンスを有するサブ画像440（図4に示す）とを構築するために使用される一次認知分析を、パーソナライズされたタグ／メタデータを有するサブ画像430と、パーソナライズされたタグ／メタデータおよびユーザ検索ガイダンスを有するサブ画像440とをランク付けするために使用される一次認知分析とともに実行する。検索エンジン406は、パーソナライズされたタグ／メタデータを有するサブ画像430もしくはパーソナライズされたタグ／メタデータおよびユーザ検索ガイダンスを有するサブ画像440またはその両方に基づいて、画像検索を実行する。検索エンジン406は、検索エンジン406がネットワーク410（例えば、ローカル・ネットワーク、ワイド・エリア・ネットワーク、インターネットなど）にアクセスして、パーソナライズされたタグ／メタデータを有するサブ画像430もしくはパーソナライズされたタグ／メタデータおよびユーザ検索ガイダンスを有するサブ画像440またはその両方に一致するデータを、ブログ、フォーラム、ニュース・サイト、レビュー・サイト、データ・リポジトリなどの様々なロケーション・タイプを表す様々なウェブ・サーバ420から引き出すことを可能にするブラウザ機能を含む。

【0031】

図4Bは、例として示されている、パーソナライズされたタグ／メタデータを有するサブ画像430と、パーソナライズされたタグ／メタデータおよびユーザ検索ガイダンスを有するサブ画像440との詳細を示す。複合画像100の他のサブ画像についても同様の例を生成することができる。パーソナライズされたタグ／メタデータを有するサブ画像430によって示される例では、サブ画像112Aは、ODSモジュール202、画像処理モジュール402、および認知プロセッサ404によって処理されており、現在、画像検索を行うために検索エンジン406によって使用される準備ができています。同様に、パーソナライズされたタグ／メタデータおよびユーザ検索ガイダンスを有するサブ画像440によって示される例では、サブ画像114Aは、ODSモジュール202、画像処理モジュール402、および認知プロセッサ404によって処理されており、現在、画像検索を行うために検索エンジン406によって使用される準備ができています。次いで、パーソナライズされたタグ／メタデータを有するサブ画像430、ならびにパーソナライズされたタグ／メタデータおよびユーザ検索ガイダンスを有するサブ画像440は、認知プロセッサ404によってランク付けされ、認知検索モジュール302Aによって、パーソナライズされたサブ画像検索結果312として出力される。

【0032】

図5は、本発明の実施形態によるコンピュータ実装方法500を示す。方法500は、認知検索モジュール302、302A（図3および4Aに示す）によって実行することができる。適切な場合には、方法500の説明は、モジュール302、302Aの対応する要素を参照する。本発明の実施形態によると、方法500は、「開始」ブロック202で始まり、次いで、ブロック204に進み、ODISモジュール202は、複合画像100をサブ画像112A、114A、116A、118A、120Aにセグメント化し、各サブ画像112A、114A、116A、118A、120Aは複合画像100内の電子的に識別可能なオブジェクトを含む。任意選択で、ODISモジュール202は、サブ画像

112A、114A、116A、118A、120Aに自動画像アノテーションを適用して、サブ画像112A、114A、116A、118A、120Aへの記述的メタデータの初期割り当てを生成することができる。方法500は、次いで、ブロック506に進み、認知プロセッサ404は、サブ画像112A、114A、116A、118A、120Aと、任意選択で、メタデータの初期割り当てを受信する。本発明の実施形態によると、認知プロセッサ404は、画像処理および表現ベースの自然言語処理機能を使用して、サブ画像、任意選択の初期メタデータ、およびユーザAコーパス320を分析し、ユーザAに対するサブ画像の関連性のレベルを判定し、その関連性のレベルを自然言語で捕捉し、関連性のレベルをサブ画像のメタデータに組み込んで、サブ画像のパーソナライズされたメタデータを作成するという第1の認知分析タスク（TASK-1）を実行する。本発明の一部の実施形態では、初期メタデータは、ユーザAに対するサブ画像の関連性のレベルを決定するタスクを実行する際に、認知プロセッサ404を補強または支援するために使用することができる。

10

【0033】

ブロック508において、検索エンジン406は、各サブ画像について画像検索結果のセットが生成されるように、各サブ画像およびそれに関連付けられたパーソナライズされたメタデータに対して画像検索を実行する。ブロック510において、認知プロセッサ510は、各サブ画像、各サブ画像に関連付けられたパーソナライズされたメタデータ、および任意選択でユーザAコーパスを分析して、ユーザAに対する関連性レベル（または重要度レベル）に基づいて各サブ画像をランク付けするという第2の認知タスク（TASK-2）を実行する。ブロック512において、認知プロセッサは、例えば、コンピュータ・ディスプレイを使用してレビューするために、各ランク付けされたサブ画像およびそれに関連付けられた検索結果のセットをユーザに表示する。本発明の一部の実施形態では、認知プロセッサは、しきい値を上回るランキング・レベル（または重要度レベル）を有するサブ画像のみを表示するように構成され得る。

20

【0034】

判定ブロック514において、方法500は、ユーザAが検索結果に関するフィードバックを認知プロセッサ404に提供したか否かを判定する。判定ブロック514における問い合わせに対する答えが「はい」である場合、ブロック516において、ユーザ・フィードバックが記憶され、認知プロセッサ404によって実行される分析を増強または改善するために使用される。例えば、認知プロセッサ404が分類器710（図7に示す）を使用して実装される本発明の実施形態では、ユーザ・フィードバックは、分類器710の追加の訓練データとして使用される。本発明の一部の実施形態では、ユーザ・フィードバックは、ユーザAが、表示された検索結果とどのように対話したかから導出することができる。本発明の一部の実施形態では、認知プロセッサ404は、ディスプレイを介してランキングに関する質問をユーザAに提示することによって、ユーザ・フィードバックを直接求めることができる。

30

【0035】

次いで、方法500は、判定ブロック518に進み、判定ブロック514に戻って追加のユーザ・フィードバックをチェックするか否か、またはブロック504に戻って現在の複合画像100の分析を繰り返すか否かを判定する。本発明の一部の実施形態には、認知プロセッサ404は、判定ブロック518において、ユーザ・フィードバックを評価して、ユーザ・フィードバックが現在の画像検索の品質を改善するか否かを判定することができる。認知プロセッサ404が、判定ブロック518において、ユーザ・フィードバックによって現在の画像検索を改善することができる判断した場合、認知プロセッサ404は、ユーザ・フィードバックに基づいて、その第1および第2の認知タスクを更新し、次いで、ブロック504に戻ることによって、画像検索を繰り返すことができる。本発明の一部の実施形態では、上述した画像検索の繰り返しは、オプションとしてユーザAに提供され、ユーザAがユーザ承認を入力した場合にのみ実行され得る。ブロック518における問い合わせに対する答えが「いいえ」である場合、方法500は、判定ブロック514

40

50

に戻り、ユーザ・フィードバックのチェックを継続する。判定ブロック 514 において追加のユーザ・フィードバックが受信されない場合、方法 500 は、判定ブロック 520 に進み、検索のために提出される複合画像 100 がさらにあるかどうかを評価する。判定ブロック 520 における問い合わせに対する答えが「いいえ」である場合、方法 500 は、ブロック 522 に進み、待機し、その後判定ブロック 520 に戻る。判定ブロック 520 における問い合わせに対する答えが「はい」である場合、方法 500 はブロック 502 に戻る。

【0036】

本発明の一部の実施形態では、第 1 および第 2 の認知タスクは、サブ画像が検索される前にランク付けされるように、画像検索に先立って行うことができる。本発明の一部の実施形態では、第 1 および第 2 の認知タスクは、ユーザ A が、検索画像とともに、（例えば、図 4 B に示すパーソナライズされたタグ／メタデータおよびユーザ検索ガイダンスを有するサブ画像 440 によって示されるように）ユーザにとって関心のある複合画像内のオブジェクトの自然言語による識別を入力することによって、さらに補強することができる。

10

【0037】

図 6 A は、本発明の実施形態によるコンピュータ実装方法 500 を示し、図 6 B は、方法 600 において利用することができる式 A～D を示す。方法 600 は、認知検索モジュール 302、302 A（図 3 および図 4 A に示す）によって実行され得る。適切な場合には、方法 600 の説明は、モジュール 302、302 A の対応する要素を参照する。本発明の実施形態に従って、ユーザ A は、複合画像 100 を認知検索モジュール 302、302 A に入力する。ブロック 602 において、方法 600 は、画像認識技術を使用して、複合画像 100 内の異なる別々のサブ画像 112 A、114 A、116 A、118 A、120 A を識別する。ブロック 604 において、方法 600 は、OUC A 330 およびユーザ A コーパス 320 に基づいて、パーソナライズされたタグおよびメタデータを作成する。ブロック 606 において、方法 600 は、ブロック 606 に列挙されたような多くのコンテキスト要因、ならびに OUC A 330 もしくはユーザ A コーパス 320 またはその両方のいずれかに基づいて、複合画像 100 内の各サブ画像 112 A、114 A、116 A、118 A、120 A の相対的な重要度を識別し、割り当てる。ブロック 606 では、任意選択で、ユーザ A が、サブ画像 112 A、114 A、116 A、118 A、120 A のうちの 1 つもしくは 2 つ、またはサブ画像 112 A、114 A、116 A、118 A、120 A の一部を丸で囲むことによって、サブ画像 112 A、114 A、116 A、118 A、120 A の相対的な重要度を修正することができる。

20

30

【0038】

ブロック 608 において、検索エンジン 406 は、パーソナライズされたタグ／メタデータを有するサブ画像に基づいて画像検索を行う。ブロック 610 では、ブロック 608 で検索エンジン 406 によって出力された検索結果が組み合わされて、相対的な重要度に基づいてソートされ、優先順位が付けられる。次いで、ブロック 610 で生成された検索結果がユーザ A に提示される。ユーザ A は、リンクを開く、特定の部分をズームする、任意のダウンロード、後続のページで検索を修正する、または後続の検索という形で、検索結果に関するフィードバックを提供する。

40

【0039】

本発明の実施形態では、OUC A 330 は、ユーザ・プロフィールおよび関心に限定されることなく、検索された画像の履歴、および画像に関連しないアクション（文書、ブラウジングなど）の履歴を含むいくつかの入力から構築される。本発明の一部の実施形態では、コンテキストはタグのセットとして構築される。このセットは、方法 600 がユーザのアクティビティから「学習する」につれ、新しい情報で絶えず更新される。

【0040】

本発明の実施形態では、各サブ画像 112 A、114 A、116 A、118 A、120 A は、ユーザ・コンテキストに対してスコア付けされ得る。各サブ画像には、複合画像 1

50

00の撮影者によって割り当てられたキュレートされた(curated)タグ、ソーシャル・ネットワーク上のユーザAの1人または複数の「友人」によって割り当てられたクラウド・ソーシングされたタグ、画像認識アルゴリズムによって割り当てられた自動生成されたタグ、および類似の複合画像または類似のサブ画像のためにユーザAによって使用されたユーザAの履歴からのタグなど、1つまたは複数のタグを割り当てることができる。

【0041】

タグの割り当てが完了すると、既知のタグ・ステミング技術を使用して、関連するタグを追加する。すべてのタグに和集合関数(union function)が適用される。関連性スコアは、「I」(サブ画像タグ)と「C」(ユーザ・コンテキスト・タグ)との間で計算される。関連性スコアは、図6Bに示す式A~Cによって示されるジャカード指数を使用して計算される。各サブ画像の最終的な関連性スコアは、類似度スコアと、複合画像に対するサブ画像の相対的なサイズと、複合画像内のサブ画像の相対的な位置との関数である。最終的な関連性スコアは、図6Bの式Dに示す線形重み付け関数を使用して計算することができる。

10

【0042】

ここで、本明細書に開示される本発明の好ましい実施形態の態様に使用することができる機械学習技術のさらなる詳細を提供する。本明細書に記載されるプロセッサの様々なタイプのコンピュータ制御機能は、機械学習もしくは自然言語処理技術またはその両方を使用して実装することができる。一般に、機械学習技術は、機械学習アルゴリズムもしくは自然言語処理アルゴリズムまたはその両方のセットを実行するように構成されたプログラマブル・コンピュータとして実装することができる、いわゆる「ニューラル・ネットワーク」上で実行される。ニューラル・ネットワークは、神経生理学、認知科学／心理学、物理学(統計力学)、制御理論、コンピュータ・サイエンス、人工知能、統計学／数学、パターン認識、コンピュータ・ビジョン、並列処理、およびハードウェア(例えば、デジタル／アナログ／VLSI／光学)を含む、様々な専門分野からの知識を組み込む。

20

【0043】

ニューラル・ネットワークおよびそれらの機械学習アルゴリズムの基本的な機能は、ある種の機械知覚を通じて非構造化センサ・データを解釈することによってパターンを認識することである。ネイティブの形式の非構造化実世界データ(例えば、画像、音声、テキスト、または時系列データ)は、コンピュータが理解および操作することができる数値形式(例えば、大きさおよび方向を有するベクトル)に変換される。機械学習アルゴリズムは、実世界のデータ・ベクトルに含まれるパターン(または関係)が明らかにされ、学習されるまで、実世界のデータ・ベクトルに対して学習ベースの分析を複数回繰り返し実行する。学習されたパターン／関係は、例えば、実世界データの分類(またはラベリング)および実世界データのクラスタリングを含む、様々なタスクを実行するために使用することができる予測モデルとして機能する。分類タスクは、多くの場合、ラベルとデータとの相関関係を認識するためにニューラル・ネットワーク(すなわち、モデル)を訓練するためのラベル付きデータセットの使用に依存する。これは教師あり学習として知られている。分類タスクの例としては、画像内のオブジェクト(例えば、停止標識、歩行者、車線マーカなど)を識別すること、ビデオ内のジェスチャを認識すること、声を検出すること、オーディオ内の声を検出すること、特定の話者を識別すること、スピーチをテキストに書き換えることなどが含まれる。クラスタリング・タスクは、オブジェクト間の類似性を識別し、それらの共通する特性に従ってグループ化し、それらを他のグループのオブジェクトと区別する。これらのグループは「クラスタ」として知られている。

30

40

【0044】

好ましい実施形態に従って、本発明の態様を実施するために使用することができる機械学習技術の一例について、図7および図8を参照して説明する。本発明の実施形態に従って構成および配置された機械学習モデルについて、図7を参照して説明する。本明細書に記載される本発明の実施形態のうちの1つまたは複数を実施することが可能な例示的なコ

50

ンピューティング・システムおよびネットワーク・アーキテクチャの詳細な説明は、図 9 を参照して提供される。

【0045】

図 7 は、本明細書に記載される本発明の好ましい実施形態の様々な態様を実施することができる分類器システム 700 を示すブロック図を示す。より具体的には、システム 700 の機能は、本発明の実施形態においてコンピュータ機能を実装するために使用することができる、様々なモデルおよびサブモデルを生成するために本発明の実施形態において使用される。システム 700 は、ネットワーク 704 を介して分類器 710 と通信する複数のデータ・ソース 702 を含む。本発明の一部の実施形態では、データ・ソース 702 は、ネットワーク 704 をバイパスし、分類器 710 に直接供給することができる。データ・ソース 702 は、本発明の実施形態に従って分類器 710 によって評価されるデータ／情報入力を提供する。データ・ソース 702 は、分類器 710 によって作成されたモデル 716 を訓練もしくは更新またはその両方を行うために分類器 710 によって使用されるデータ／情報入力も提供する。データ・ソース 702 は、リアル・タイム・データを収集するように構成されたセンサ、データ・リポジトリ（訓練データ・リポジトリを含む）、および他の分類器からの出力を含むが、これらに限定されない多種多様なデータ・ソースとして実装することができる。ネットワーク 704 は、ローカル・ネットワーク、ワイド・エリア・ネットワーク、プライベート・ネットワーク、インターネットなどを含むが、これらに限定されない任意のタイプの通信ネットワークとすることができる。

10

20

【0046】

分類器 710 は、処理システム 900（図 9 に示す）などのプログラマブル・コンピュータによって実行されるアルゴリズムとして実装することができる。図 7 に示すように、分類器 710 は、一連の機械学習（ML）アルゴリズム 712 と、自然言語処理（NLP）アルゴリズム 714 と、ML アルゴリズム 712 によって生成された（または学習された）関係（または予測）アルゴリズムであるモデル 716 と、を含む。分類器 710 のアルゴリズム 712、714、716 は、図示および説明を容易にするために別々に示されている。本発明の実施形態では、分類器 710 の様々なアルゴリズム 712、714、716 によって実行される機能は、図示したものとは異なるように分散させることができる。例えば、分類器 710 がサブタスクを有するタスク全体を実行するように構成されている場合、ML アルゴリズム 712 の一部が各サブタスクを実行し、ML アルゴリズム 712 の一部がタスク全体を実行するように、一連の ML アルゴリズム 712 をセグメント化することができる。さらに、本発明の一部の実施形態では、NLP アルゴリズム 714 を ML アルゴリズム 712 内に統合することができる。

30

【0047】

NLP アルゴリズム 714 は、分類器 710、より具体的には ML アルゴリズム 712 が、自然言語データ（テキストおよび音声）を受信し、言語処理、情報検索、および機械学習の要素を適用して、自然言語入力から意味を導出し、導出された意味に基づいて潜在的にアクションを起こすことを可能にする音声認識機能を含む。本発明の実施形態に従って使用される NLP アルゴリズム 714 は、分類器 710 が結果 720 を自然言語（テキストおよび音声）に変換して、結果 720 の態様を自然言語コミュニケーションとして伝えることを可能にする音声合成機能を含むこともできる。

40

【0048】

NLP アルゴリズム 714 および ML アルゴリズム 712 は、データ・ソース 702 から入力データ（すなわち、訓練データおよび分析中のデータ）を受信し、評価する。ML アルゴリズム 712 は、入力データのフォーマットを解釈して利用するために必要な機能を含む。例えば、データ・ソース 702 が画像データを含む場合、ML アルゴリズム 712 は、画像データを解釈するように構成された視覚認識ソフトウェアを含むことができる。ML アルゴリズム 712 は、分類器 710 が完了するように設計されているタスク全体およびサブタスクをモデル化する 1 つまたは複数のモデル 716 を時間の経過とともに作成／訓練／更新するために、受信した訓練データ（例えば、データ・ソース 702 のうち

50

の1つまたは複数から受信したデータ)に機械学習技術を適用する。

【0049】

ここで図7および図8をまとめて参照すると、図8は、上述したモデル716を生成するためにMLアルゴリズム712によって実行される学習段階800の例を示す。学習段階800において、分類器710は、訓練データから特徴を抽出し、その特徴を、MLアルゴリズム712によって認識および分析され得るベクトル表現に変換する。特徴ベクトルは、MLアルゴリズム712によって分析され、ターゲット・モデル(またはモデルのタスク)に対して訓練データを「分類」し、分類された訓練データ間の関係を明らかにする。MLアルゴリズム712の適切な実施態様の例は、ニューラル・ネットワーク、サポート・ベクター・マシン(SVM)、ロジスティック回帰、決定木、隠れマルコフ・モデル(HMM)などを含むが、これらに限定されない。MLアルゴリズム712によって実行される学習または訓練は、教師あり、教師なし、または教師ありおよび教師なし学習の態様を含むハイブリッドとすることができる。教師あり学習は、訓練データが既に利用可能であり、分類／ラベル付けされている場合である。教師なし学習は、訓練データが分類／ラベル付けされていない場合であり、したがって、分類器710およびMLアルゴリズム712の反復によって開発されなければならない。教師なし学習は、例えば、クラスタリング、異常検出、ニューラル・ネットワーク、深層学習などを含む追加の学習／訓練方法を利用することができる。

10

【0050】

モデル716がMLアルゴリズム712によって十分に訓練されると、「実世界」データを生成するデータ・ソース702にアクセスし、「実世界」データをモデル716に適用して、使用可能な結果720のバージョンを生成する。本発明の一部の実施形態では、結果720は、分類器710にフィードバックされ、モデル716を更新もしくは改良またはその両方を行うための追加の訓練データとして、MLアルゴリズム712によって使用され得る。

20

【0051】

本発明の実施形態では、MLアルゴリズム712およびモデル716は、特定の結果／決定の全体的な精度を改善するために、それらの結果／決定(結果720を含む)のうちの様々なものに信頼水準(confidence level、CL)を適用するように構成され得る。MLアルゴリズム712もしくはモデル716またはその両方が、CLの値が所定のしきい値(TH)を下回る(すなわち、 $CL < TH$)という判定を行うか、または結果を生成する場合、この結果／判定は、その判定／結果が有効ではないという結論を正当化するのに十分に低い「信頼性」を有するものとして分類され得て、この結論は、その判定／結果が下流の処理においていつ処理されるか、どのように処理されるか、もしくは処理されるかどうか、またはその組み合わせを決定するために使用することができる。CL > THの場合、その判定／結果は、有効であると見なすことができ、この結論は、その判定／結果が下流の処理においていつ処理されるか、どのように処理されるか、もしくは処理されるかどうか、またはその組み合わせを決定するために使用することができる。多くの異なる所定のTHレベルを提供することができる。CL > THの判定／結果は、判定／結果が下流の処理においていつ処理されるか、どのように処理されるか、もしくは処理されるか、またはその組み合わせを優先順位付けするために、最も高いCL > THから最も低いCL > THまでランク付けすることができる。

30

40

【0052】

本発明の実施形態では、分類器710は、結果720に信頼水準(CL)を適用するように構成することができる。分類器710が、結果720のCLが所定のしきい値(TH)を下回る(すなわち、 $CL < TH$)と判定した場合、結果720は、結果720の「信頼性なし」の分類を正当化するのに十分に低いものとして分類され得る。CL > THの場合、結果720は、結果720が有効であるという判定を正当化するのに十分に高いものとして分類され得る。CL > THの結果720を最も高いCL > THから最も低いCL > THまでランク付けすることができるように、多くの異なる所定のTHレベルを提供する

50

ことができる。

【0053】

分類器710によって、より具体的にはMLアルゴリズム712によって実行される機能は、重み付き有向グラフとして編成することができ、ノードは（例えば、人間の脳のニューロンに倣ってモデル化された）人工ニューロンであり、重み付き有向エッジがノードを接続する。分類器710の有向グラフは、特定のノードが入力層ノードを形成し、特定のノードが隠れ層ノードを形成し、特定のノードが出力層ノードを形成するように編成することができる。入力層ノードは隠れ層ノードに結合し、出力ノードは出力層ノードに結合する。各ノードは、それぞれが接続強度を有する方向矢印として表すことができる接続経路によって、隣接する層のすべてのノードに接続される。複数の入力層、複数の隠れ層、および複数の出力層を設けることができる。複数の隠れ層が設けられた場合、分類器710は、分類器710の割り当てられたタスクを実行するための教師なし深層学習を実行することができる。

10

【0054】

人間の脳の機能と同様に、各入力層ノードは、接続強度調整もノード加算もなしに入力を受け取る。各隠れ層ノードは、その入力を、関連する接続経路に関連付けられた接続強度に従って、すべての入力層ノードから受け取る。同様の結合強度乗算およびノード加算が、隠れ層ノードおよび出力層ノードに対して実行される。

【0055】

分類器710の重み付き有向グラフは、データ・レコード（例えば、データ・ソース702からの出力）を一度に1つずつ処理し、レコードの当初の任意の分類をレコードの既知の実際の分類と比較することによって「学習」する。「逆伝搬（back-propagation）」（すなわち、「誤差の逆方向伝搬（backward propagation of errors）」）として知られる訓練方法を用いて、最初のレコードの初期分類からの誤差が分類器710の重み付き有向グラフにフィードバックされ、2回目に重み付き有向グラフの重み付き接続を修正するために使用され、このフィードバックプロセスが何度も反復して続く。分類器710の重み付き有向グラフの訓練段階では、各レコードの正しい分類がわかっているため、出力ノードに「正しい」値を割り当てることができる。例えば、正しいクラスに対応するノードのノード値は「1」（または0.9）であり、その他のノードのノード値は「0」（または0.1）である。したがって、出力ノードの重み付き有向グラフの計算値をこれらの「正しい」値と比較し、各ノードの誤差項を計算することが可能である（すなわち、「デルタ」ルール）。次いで、これらの誤差項は、次の反復において出力値が「正しい」値に近づくように、隠れ層の重みを調整するために使用される。

20

30

【0056】

図9は、本発明の好ましい実施形態の態様に従って1つまたは複数のコンピュータ処理動作を実施するために使用することができるコンピュータ・システム900の高レベル・ブロック図を示す。1つの例示的なコンピュータ・システム900が示されているが、コンピュータ・システム900は、コンピュータ・システム900を追加のシステム（図示せず）に接続する通信経路925を含み、インターネット、イントラネット、もしくはワイヤレス通信ネットワーク、またはその組み合わせなどの、1つまたは複数の、ワイド・エリア・ネットワーク（WAN）もしくはローカル・エリア・ネットワーク（LAN）またはその両方を含むことができる。コンピュータ・システム900および追加のシステムは、例えば、それらの間でデータを通信するために、通信経路925を介して通信する。本発明の一部の実施形態では、追加のシステムは、1つまたは複数のクラウド・コンピューティング・システム50として実装することができる。クラウド・コンピューティング・システム50は、コンピュータ・システム900を使用して実装することができる本詳細な説明に記載されたあらゆるコンピューティング・システムを含む、コンピュータ・システム900の機能の一部またはすべてを（任意の組み合わせで）補完、サポート、または置き換えることができる。加えて、本詳細な説明に記載された様々なコンピューティン

40

50

グ・システムの機能の一部またはすべてを、クラウド・コンピューティング・システム 50 のノードとして実装することができる。

【0057】

コンピュータ・システム 900 は、プロセッサ 902 などの 1 つまたは複数のプロセッサを含む。プロセッサ 902 は、通信インフラ 904（例えば、通信バス、クロスオーバー・バー、またはネットワーク）に接続されている。コンピュータ・システム 900 は、ディスプレイ・ユニット 908 に表示するために通信インフラ 904 から（または図示されていないフレーム・バッファから）グラフィックス、テキスト、および他のデータを転送するディスプレイ・インターフェース 906 を含むことができる。コンピュータ・システム 900 は、メイン・メモリ 910、好ましくはランダム・アクセス・メモリ（RAM）も含み、二次メモリ 912 を含むこともできる。二次メモリ 912 は、例えば、ハード・ディスク・ドライブ 914 もしくはリムーバブル・ストレージ・ドライブ 916（例えば、フレキシブル・ディスク・ドライブ、磁気ドライブ、または光ディスク・ドライブを表す）またはその両方を含むことができる。リムーバブル・ストレージ・ドライブ 916 は、当業者によく知られている方法で、リムーバブル・ストレージ・ユニット 918 から読み取り、もしくはそれに書き込み、またはその両方を行う。リムーバブル・ストレージ・ユニット 918 は、例えば、フレキシブル・ディスク、コンパクト・ディスク、磁気テープ、または光学ディスク、フラッシュ・ドライブ、ソリッド・ステート・メモリなどを表し、これらは、リムーバブル・ストレージ・ドライブ 916 によって読み取られ、書き込まれる。理解されるように、リムーバブル・ストレージ・ユニット 918 は、コンピュータ・ソフトウェアもしくはデータまたはその両方が記憶されたコンピュータ可読媒体を含む。

10

20

【0058】

本発明の代替の実施形態では、二次メモリ 912 は、コンピュータ・プログラムまたは他の命令をコンピュータ・システムにロードできるようにするための他の同様の手段を含むことができる。このような手段は、例えば、リムーバブル・ストレージ・ユニット 920 およびインターフェース 922 を含むことができる。このような手段の例としては、（ビデオ・ゲーム機器に見られるような）プログラム・パッケージおよびパッケージ・インターフェース、（EPROM または PROM などの）リムーバブル・メモリ・チップおよび関連付けられたソケット、ならびにソフトウェアおよびデータがリムーバブル・ストレージ・ユニット 920 からコンピュータ・システム 900 への転送を可能にする他のリムーバブル・ストレージ・ユニット 920 およびインターフェース 922 を挙げることができる。

30

【0059】

コンピュータ・システム 900 は、通信インターフェース 924 を含むこともできる。通信インターフェース 924 により、コンピュータ・システムと外部デバイスとの間でソフトウェアおよびデータを転送することが可能になる。通信インターフェース 924 の例としては、モデム、（イーサネット・カードなどの）ネットワーク・インターフェース、通信ポート、または PCMCIA スロットおよびカードなどを挙げることができる。通信インターフェース 924 を介して転送されるソフトウェアおよびデータは、例えば、通信インターフェース 924 によって受信可能な電子信号、電磁信号、光信号、またはその他の信号であり得る信号の形態である。これらの信号は、通信経路（すなわち、チャンネル）925 を介して通信インターフェース 924 に供給される。通信経路 925 は、信号を搬送し、ワイヤすなわちケーブル、光ファイバ、電話線、携帯電話リンク、RF リンク、もしくは他の通信チャンネル、またはその組み合わせを使用して実装することができる。

40

【0060】

本発明の様々な実施形態は、関連する図面を参照して本明細書に記載されている。本発明の範囲から逸脱することなく、本発明の代替の実施形態を考案することができる。様々な接続および位置関係（例えば、上、下、隣接など）が、以下の説明および図面における要素間に記載されている。これらの接続もしくは位置関係またはその両方は、別段の指定

50

がない限り、直接的または間接的であってもよく、本発明は、この点において限定することは意図されていない。したがって、エンティティの結合は、直接的または間接的な結合のいずれかを指すことができ、エンティティ間の位置関係は、直接的または間接的な位置関係のいずれかを指す。さらに、本明細書に記載される様々なタスクおよびプロセス・ステップは、本明細書では詳細に説明されない追加のステップまたは機能を有する、より包括的な手順またはプロセスに組み込むことができる。

【0061】

以下の定義および略語は、特許請求の範囲および本明細書の解釈のために使用される。本明細書で使用される場合、用語「備える (comprise)」、「備えている (comprising)」、「含む (include)」、「含んでいる (including)」、「有する (has)」、「有している (having)」、「含有する (contains)」もしくは「含有している (containing)」またはそれらの任意の他の変形は、非排他的な包含をカバーすることが意図されている。例えば、要素のリストを含む組成物、混合物、プロセス、方法、物品、または装置は、必ずしもそれらの要素のみに限定されず、明示的に列挙されていないか、またはそのような組成物、混合物、プロセス、方法、物品、もしくは装置に固有の他の要素を含むことができる。

【0062】

本明細書で使用される用語は、本発明の特定の実施形態を説明することのみを目的としており、本発明を限定することは意図されていない。本明細書で使用される場合、単数形「a」、「an」および「the」は、文脈が明らかにそうでないことを示さない限り、複数形も同様に含むことが意図されている。用語「備える」もしくは「備えている」またはその両方は、本明細書で使用される場合、述べられた特徴、整数、ステップ、動作、要素、もしくは構成要素またはその組み合わせの存在を指定するが、1つもしくは複数の他の特徴、整数、ステップ、動作、要素、構成要素、もしくはそれらのグループまたはその組み合わせの存在あるいは追加を排除するものではないことをさらに理解されよう。

【0063】

さらに、「例示的」という用語およびその変形は、本明細書では、「例、事例、または例示として役立つこと」を意味するために使用される。「例示的」として本明細書に記載される任意の実施形態または設計は、必ずしも、他の実施形態もしくは設計よりも好ましいまたは有利であると解釈されるべきではない。「少なくとも1つの」、「1つまたは複数の」という用語、およびそれらの変形は、1以上の任意の整数、すなわち、1、2、3、4などを含むことができる。「複数 (a plurality)」という用語およびその変形は、2以上の任意の整数、すなわち、2、3、4、5などを含むことができる。「接続」という用語およびその変形は、間接的な「接続」と直接的な「接続」の両方を含むことができる。

【0064】

用語「約」、「実質的に」、「およそ」およびそれらの変形は、本出願の出願時に利用可能な機器に基づく特定の量の測定に関連付けられた誤差の程度を含むことが意図されている。例えば、「約」は、所与の値の±8%または5%または2%の範囲を含むことができる。

【0065】

本明細書で使用される場合、機械学習アルゴリズムの文脈では、用語「入力データ」およびその変形は、訓練、学習、もしくは分類動作、またはその組み合わせを実行するために機械学習アルゴリズムにおいて受信され、機械学習アルゴリズムによって使用される、あらゆるタイプのデータまたは他の情報を網羅することが意図されている。

【0066】

本明細書で使用される場合、機械学習アルゴリズムの文脈では、用語「訓練データ」およびその変形は、訓練もしくは学習動作またはその両方を行うために機械学習アルゴリズムにおいて受信され、機械学習アルゴリズムによって使用される、あらゆるタイプのデータまたは他の情報を網羅することが意図されている。

【0067】

本明細書で使用される場合、機械学習アルゴリズムの文脈では、用語「アプリケーション・データ」、「実世界データ」、「実データ」、およびそれらの変形は、分類動作を行うために機械学習アルゴリズムにおいて受信され、機械学習アルゴリズムによって使用される、あらゆるタイプのデータまたは他の情報を網羅することが意図されている。

【0068】

「信号通信中」、「通信中」、「通信可能に結合された」という語句、およびそれらの変形は、本明細書では互換的に使用することができ、あらゆるシステム、ハードウェア、ソフトウェア、プロトコル、またはフォーマットを使用して、情報またはデータを交換するために電気信号を使用するあらゆる結合、接続、または相互作用を指すことができ、交換が無線で行われるか有線接続を介して行われるかは無関係である。

【0069】

本発明は、システム、方法、もしくはコンピュータ・プログラム製品、またはその組み合わせであってもよい。コンピュータ・プログラム製品は、プロセッサに本発明の態様を実行させるためのコンピュータ可読プログラム命令を有するコンピュータ可読記憶媒体を含むことができる。

【0070】

コンピュータ可読記憶媒体は、命令実行デバイスが使用するための命令を保持および記憶することができる有形のデバイスとすることができる。コンピュータ可読記憶媒体は、例えば、電子記憶デバイス、磁気記憶デバイス、光記憶デバイス、電磁記憶デバイス、半導体記憶デバイス、または前述の任意の適切な組み合わせであってもよいが、これらに限定されない。コンピュータ可読記憶媒体のより具体的な例の非網羅的なリストには、ポータブル・コンピュータ・ディスクレット、ハード・ディスク、ランダム・アクセス・メモリ（RAM）、読み取り専用メモリ（ROM）、消去可能プログラム可能読み取り専用メモリ（EPROMまたはフラッシュ・メモリ）、スタティック・ランダム・アクセス・メモリ（SRAM）、ポータブル・コンパクト・ディスク読み取り専用メモリ（CD-ROM）、デジタル・バーサタイル・ディスク（DVD）、メモリ・スティック、フレキシブル・ディスク、パンチ・カードまたは命令が記録された溝内の隆起構造などの機械的に符号化されたデバイス、および前述の任意の適切な組み合わせが含まれる。コンピュータ可読記憶媒体は、本明細書で使用される場合、電波もしくは他の自由に伝播する電磁波、導波路もしくは他の伝送媒体を通して伝播する電磁波（例えば、光ファイバ・ケーブルを通過する光パルス）、またはワイヤを通して送信される電気信号などの、それ自体一過性の信号であると解釈されるべきではない。

【0071】

本明細書に記載されたコンピュータ可読プログラム命令は、コンピュータ可読記憶媒体からそれぞれのコンピューティング／処理デバイスに、あるいはネットワーク、例えばインターネット、ローカル・エリア・ネットワーク、ワイド・エリア・ネットワークもしくはワイヤレス・ネットワークまたはその組み合わせを介して外部コンピュータもしくは外部記憶デバイスにダウンロードすることができる。ネットワークは、銅線伝送ケーブル、光伝送ファイバ、無線伝送、ルータ、ファイアウォール、スイッチ、ゲートウェイ・コンピュータもしくはエッジ・サーバまたはその組み合わせを含むことができる。各コンピューティング／処理装置のネットワーク・アダプタ・カードまたはネットワーク・インターフェースは、ネットワークからコンピュータ可読プログラム命令を受信し、コンピュータ可読プログラム命令をそれぞれのコンピューティング／処理装置内のコンピュータ可読記憶媒体に記憶するために転送する。

【0072】

本発明の動作を実行するためのコンピュータ可読プログラム命令は、アセンブラ命令、命令セット・アーキテクチャ（ISA）命令、機械命令、機械依存命令、マイクロコード、ファームウェア命令、状態設定データ、あるいはSmalltalk（R）、C++などのオブジェクト指向プログラミング言語、および「C」プログラミング言語または同

10

20

30

40

50

様のプログラミング言語などの従来の手続き型プログラミング言語を含む、1つまたは複数のプログラミング言語の任意の組み合わせで記述されたソース・コードまたはオブジェクト・コードのいずれかであってもよい。コンピュータ可読プログラム命令は、完全にユーザのコンピュータ上で、一部はユーザのコンピュータ上でスタンドアロン・ソフトウェア・パッケージとして、一部はユーザのコンピュータ上で、一部はリモート・コンピュータ上で、または完全にリモート・コンピュータもしくはサーバ上で実行されてもよい。後者のシナリオでは、リモート・コンピュータは、ローカル・エリア・ネットワーク（LAN）もしくはワイド・エリア・ネットワーク（WAN）を含む任意のタイプのネットワークを介してユーザのコンピュータに接続されてもよく、または接続は、（例えば、インターネット・サービス・プロバイダを使用してインターネットを介して）外部コンピュータに対して行われてもよい。一部の実施形態では、例えば、プログラマブル・ロジック回路、フィールド・プログラマブル・ゲート・アレイ（FPGA）、またはプログラマブル・ロジック・アレイ（PLA）を含む電子回路は、本発明の態様を行うために、コンピュータ可読プログラム命令の状態情報を利用して電子回路をパーソナライズすることによって、コンピュータ可読プログラム命令を実行することができる。

10

20

30

40

50

【0073】

本発明の態様は、本発明の実施形態による方法、装置（システム）、およびコンピュータ・プログラム製品の流れ図もしくはブロック図またはその両方を参照して本明細書に記載されている。流れ図もしくはブロック図またはその両方の各ブロック、ならびに流れ図もしくはブロック図またはその両方のブロックの組み合わせは、コンピュータ可読プログラム命令によって実施することができることを理解されよう。

【0074】

これらのコンピュータ可読プログラム命令は、汎用コンピュータ、専用コンピュータ、または他のプログラム可能なデータ処理装置のプロセッサに提供されて、コンピュータまたは他のプログラム可能なデータ処理装置のプロセッサを介して実行される命令が、流れ図もしくはブロック図またはその両方のブロックで指定された機能／行為を実施するための手段を作成するように、機械を生成することができる。これらのコンピュータ可読プログラム命令はまた、内部に命令が記憶されたコンピュータ可読記憶媒体が、流れ図もしくはブロック図またはその両方のブロックで指定された機能／行為の態様を実施する命令を含む製品を構成するように、コンピュータ、プログラム可能なデータ処理装置、もしくは他の装置、またはその組み合わせに特定のやり方で機能するように指示することができるコンピュータ可読記憶媒体に記憶されてもよい。

【0075】

コンピュータ可読プログラム命令はまた、コンピュータ、他のプログラム可能なデータ処理装置、または他のデバイスにロードされ、コンピュータ、他のプログラム可能な装置、または他のデバイス上で実行される命令が、流れ図もしくはブロック図またはその両方のブロックで指定された機能／行為を実施するように、一連の動作ステップをコンピュータ、他のプログラム可能な装置、または他のデバイス上で実行させて、コンピュータ実施プロセスを生成することができる。

【0076】

図中の流れ図およびブロック図は、本発明の様々な実施形態によるシステム、方法、およびコンピュータ・プログラム製品の可能な実施態様のアーキテクチャ、機能、および動作を示す。これに関して、流れ図またはブロック図の各ブロックは、指定された論理機能を実施するための1つもしくは複数の実行可能命令を構成する、モジュール、セグメント、または命令の一部を表すことができる。一部の代替の実施形態では、ブロックに示されている機能は、図に示されている順序とは異なる順序で行われてもよい。例えば、連続して示されている2つのブロックは、実際には、実質的に同時に実行することができ、またはブロックは、関与する機能性に応じて、時には逆の順番で実行することができる。ブロック図もしくは流れ図またはその両方の各ブロック、およびブロック図もしくは流れ図またはその両方のブロックの組み合わせは、指定された機能もしくは行為を実行する、また

は専用のハードウェアおよびコンピュータ命令の組み合わせを実行する専用のハードウェア・ベースのシステムによって実施することができることにも留意されたい。

【0077】

本明細書で使用される用語は、特定の実施形態を説明することのみを目的としており、本発明を限定することは意図されていない。本明細書で使用される場合、単数形「a」、「an」および「the」は、文脈が明らかにそうでないことを示さない限り、複数形も同様に含むことが意図されている。用語「備える」もしくは「備えている」またはその両方は、本明細書で使用される場合、述べられた特徴、整数、ステップ、動作、要素、もしくは構成要素またはその組み合わせの存在を指定するが、1つもしくは複数の他の特徴、整数、ステップ、動作、要素、構成要素、もしくはそれらのグループまたはその組み合わせの存在あるいは追加を排除するものではないことをさらに理解されよう。

10

【0078】

以下の特許請求の範囲におけるすべての手段またはステップに加えて機能要素の対応する構造、材料、行為、および均等物は、具体的に特許請求される他の特許請求される要素と組み合わせて機能を実行するための任意の構造、材料、または行為を含むことが意図されている。本発明の記載は、例示および説明を目的として提示されたが、開示された形態において網羅的であること、または本発明を限定することは意図されていない。本発明の範囲および思想から逸脱することなく、多くの変更形態および変形形態が当業者には明らかであろう。本実施形態は、本発明の原理および実際の用途を最もよく説明するために、ならびに他の当業者が、企図される特定の使用に適するような様々な修正を伴う様々な実施形態について本発明を理解することを可能にするために、選択され、説明された。

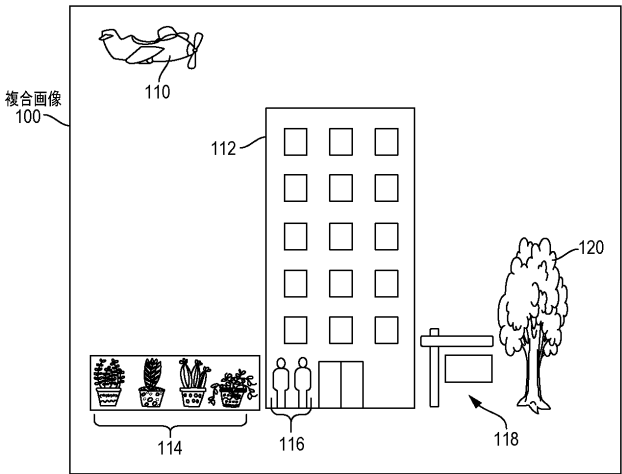
20

【0079】

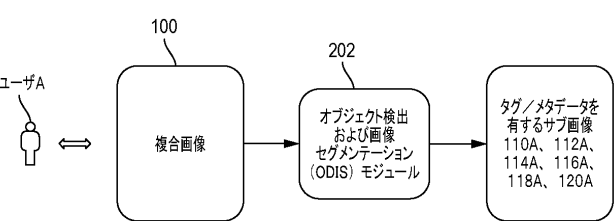
当業者は、現在および将来の両方において、以下の特許請求の範囲に含まれる様々な改善および強化を行い得ることが理解されるであろう。

【図面】

【図1】



【図2】

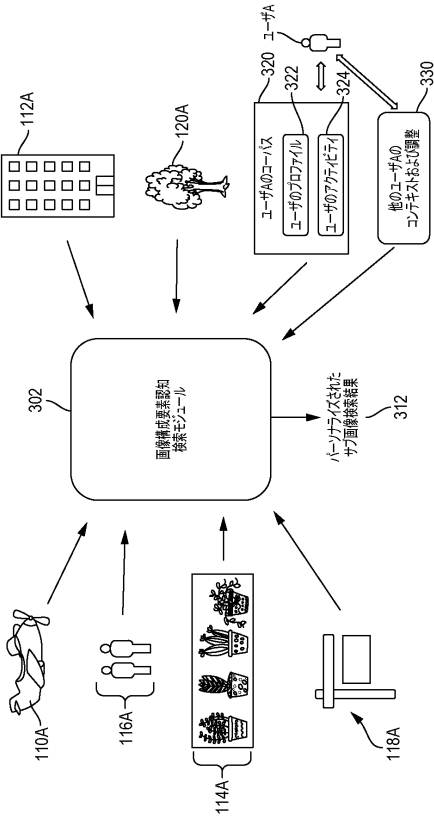


30

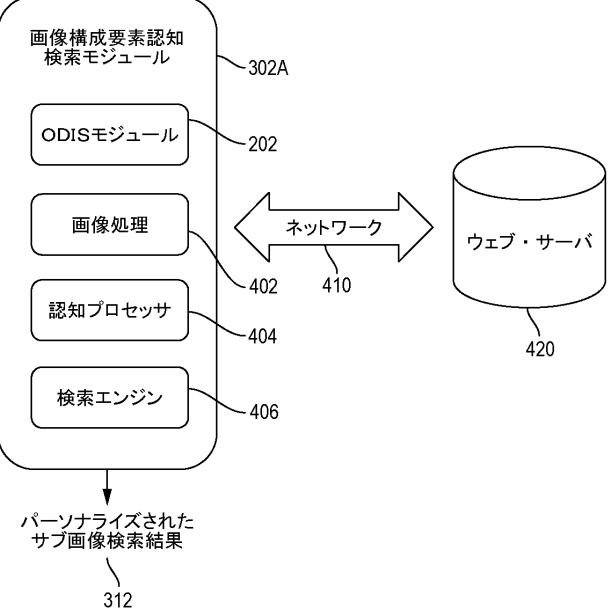
40

50

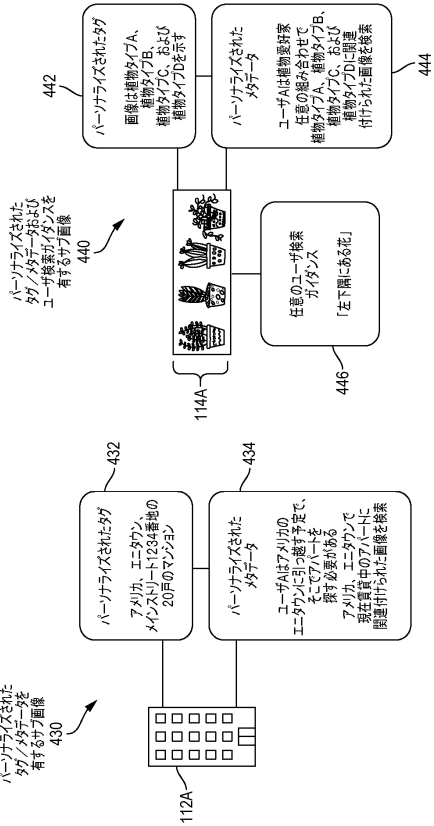
【図 3】



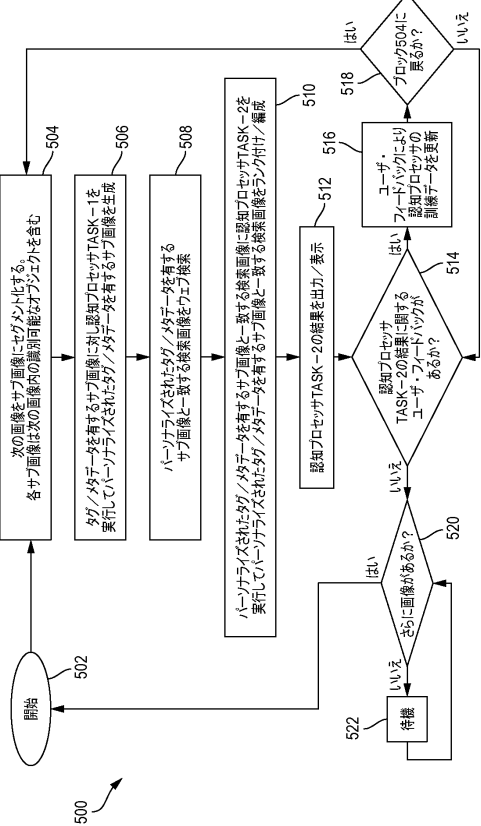
【図 4 A】



【図 4 B】



【図 5】



10

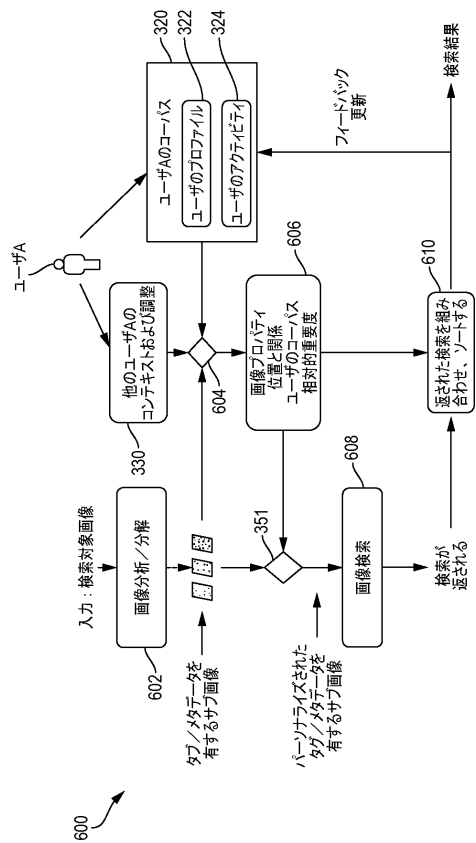
20

30

40

50

【図 6 A】



【図 6 B】

式 (Eq) A $I = \{ \text{キュレートされたタグ} \} \cup \{ \text{クラウド・ソーシングされたタグ} \} \cup \{ \text{自動タグ} \} \cup \{ \text{履歴タグ} \}$

式 (Eq) B $S = \{ \text{ユーザ} \cdot \text{コーパス} \cdot \text{タグ} \}$

式 (Eq) C $\text{類似度スコア} r = |I \cap S| / |I \cup S|$

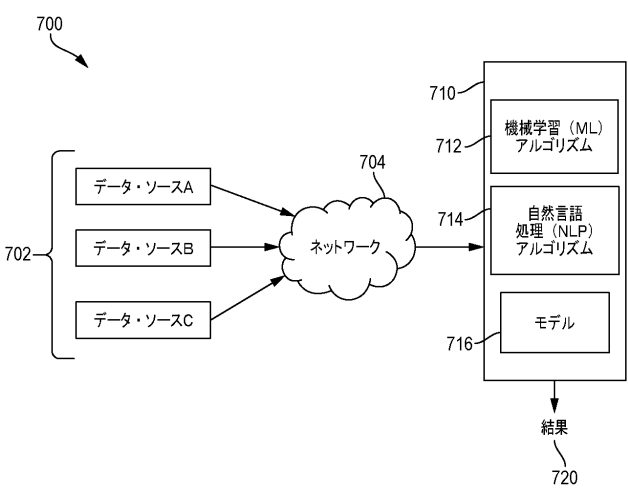
式 (Eq) D $r_s = W1 * (r) + W2 * (\text{サブ画像の面積} / \text{画像の面積}) + W3 * (\text{相対位置スコア})$

ここで
* 相対位置スコアはサブ画像の中心から複合画像の中心までの距離を測定する逆正規化関数を用いて計算され、
* W1、W2、W3は各スコアに関連付けられた重み付けを表す

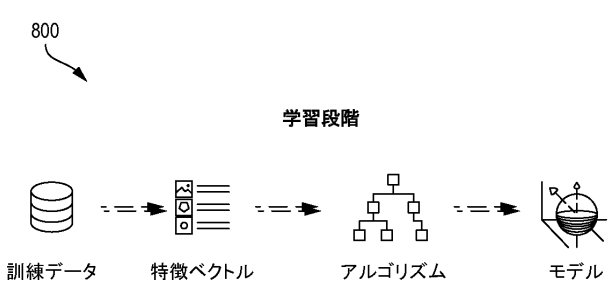
10

20

【図 7】



【図 8】

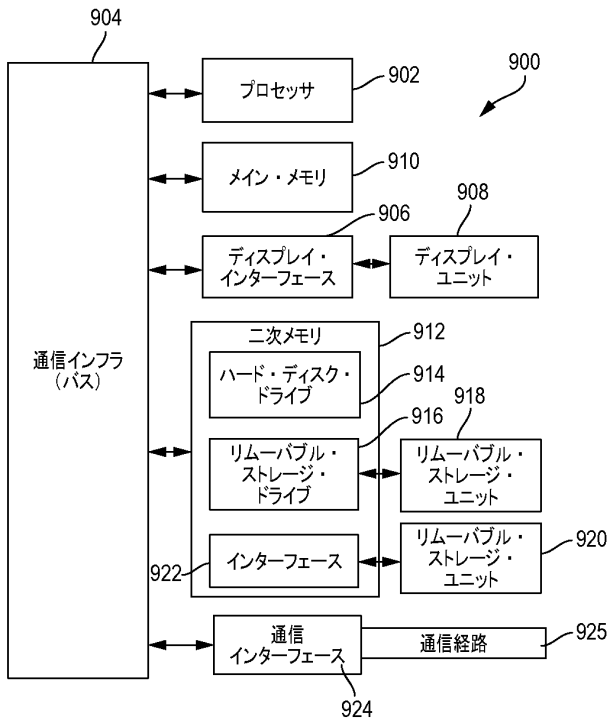


30

40

50

【図 9】



10

20

30

40

50

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2022/075650

A. CLASSIFICATION OF SUBJECT MATTER INV. G06F16/535 G06F16/583 G06F16/9535 ADD.		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) G06F		
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) EPO-Internal		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Iker Gondra: "Personalized Content-Based Image Retrieval" In: "Artificial Intelligence for Maximizing Content Based Image Retrieval", 1 January 2009 (2009-01-01), Information science reference, USA, XP055354046, ISBN: 978-1-60566-174-2 pages 261-287, DOI: 10.4018/978-1-60566-174-2.ch012, abstract page 263 page 266 - page 268; figure 5 page 274 - page 275 ----- -/--	1-21
☒ Further documents are listed in the continuation of Box C. ☐ 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 but 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 "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 16 December 2022		Date of mailing of the international search report 23/12/2022
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Authorized officer Bercan, Radu

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2022/075650

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	SMITH J R ET AL: "VISUALSEEK: A FULLY AUTOMATED CONTENT-BASED IMAGE QUERY SYSTEM", PROCEEDINGS OF ACM MULTIMEDIA 96. BOSTON, NOV. 18 - 22, 1996; [PROCEEDINGS OF ACM MULTIMEDIA], NEW YORK, ACM, US, 18 November 1996 (1996-11-18), pages 87-98, XP000734712, DOI: 10.1145/244130.244151 ISBN: 978-0-89791-871-8 Section 5 - "Multiple Regions Query"; figures 1,2,10 -----	1-21

フロントページの続き

MK,MT,NL,NO,PL,PT,RO,RS,SE,SI,SK,SM,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,KM,ML,MR,N
E,SN,TD,TG),AE,AG,AL,AM,AO,AT,AU,AZ,BA,BB,BG,BH,BN,BR,BW,BY,BZ,CA,CH,CL,CN,CO,CR,CU,
CV,CZ,DE,DJ,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IQ,IR,IS,IT,J
M,JO,JP,KE,KG,KH,KN,KP,KR,KW,KZ,LA,LC,LK,LR,LS,LU,LY,MA,MD,ME,MG,MK,MN,MW,MX,MY
,MZ,NA,NG,NI,NO,NZ,OM,PA,PE,PG,PH,PL,PT,QA,RO,RS,RU,RW,SA,SC,SD,SE,SG,SK,SL,ST,SV,SY,T
H,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,WS,ZA,ZM,ZW

(74)代理人 100120710

弁理士 片岡 忠彦

(72)発明者 バラスブラマニアン、スワミナサン

アメリカ合衆国 4 8 2 4 3 - 1 6 7 5 ミシガン州 デトロイト ルネサンス センター 4 0 0 ス
イート 2 7 1 0

(72)発明者 デ、ラダ

インド 7 0 0 1 5 6 コルカタ ウェスト ベンガル、5 9 . 0 エム アーテリアル ロード オブ
ニュータウン (ラジャルハット)、プロット エヌオー. 0 8 - エムエーアール、ディーエルエフ
アイティー パーク

(72)発明者 ジャミル、マムヌーン

アメリカ合衆国 1 0 5 9 8 ニューヨーク州 ヨークタウン・ハイツ ルート 1 3 4 キッチャワン
・ロード 1 1 0 1

(72)発明者 ウー、ジェン

アメリカ合衆国 7 8 7 5 8 - 3 4 0 0 テキサス州 オースティン バーネット ロード 1 1 5 0 1

(72)発明者 ヴァスデヴァン、シェラネロア

アメリカ合衆国 7 8 7 5 8 - 3 4 0 0 テキサス州 オースティン バーネット ロード 1 1 5 0 1

F ターム (参考) 5B175 DA02 GB05 HA01