

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関  
国際事務局

(43) 国際公開日  
2012 年 12 月 27 日(27.12.2012)



(10) 国際公開番号  
**WO 2012/176317 A1**

- (51) 国際特許分類:  
*G06T 1/00* (2006.01) *G06T 7/00* (2006.01)  
*G06F 17/30* (2006.01)
- (21) 国際出願番号: PCT/JP2011/064463
- (22) 国際出願日: 2011 年 6 月 23 日(23.06.2011)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (71) 出願人 (米国を除く全ての指定国について): サイバーアイ・エンタテインメント株式会社 (CYBER AI ENTERTAINMENT INC.) [JP/JP]; 〒1580095 東京都世田谷区瀬田 1-5-1 Tokyo (JP).
- (72) 発明者; および
- (75) 発明者/出願人 (米国についてのみ): 久寿良木 健(KUTARAGI Ken) [JP/JP]; 〒1580095 東京都世田谷区瀬田 1-5-1 サイバーアイ・エンタテインメント株式会社内 Tokyo (JP). 薄 隆(USUKI Takashi) [JP/JP]; 〒1580095 東京都世田谷区瀬田 1-5-1 サイバーアイ・エンタテインメント

株式会社内 Tokyo (JP). 横手 靖彦(YOKOTE Yasuhiko) [JP/JP]; 〒1580095 東京都世田谷区瀬田 1-5-1 サイバーアイ・エンタテインメント株式会社内 Tokyo (JP).

(74) 代理人: 辻居 幸一, 外(TSUJII Koichi et al.); 〒1008355 東京都千代田区丸の内 3 丁目 3 番 1 号 新東京ビル 中村合同特許法律事務所 Tokyo (JP).

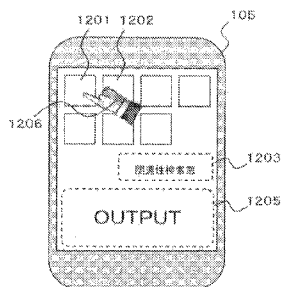
(81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

[続葉有]

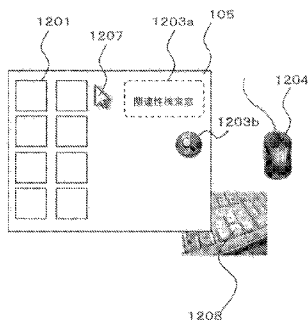
(54) Title: IMAGE RECOGNITION SYSTEM-EQUIPPED INTEREST GRAPH COLLECTION SYSTEM USING RELATIONSHIP SEARCH

(54) 発明の名称: 画像認識システムを組み込んだ関連性検索によるインタレスト・グラフ収集システム

[図12A]  
(A)



(B)



1203 Relationship search window

(57) Abstract: Provided is a system which comprehensively and effectively acquires, as an interest graph, user-specific objects and regions of interest. This system is a search system using, as an input means, image information that involves various objects and subjects, wherein by making an inquiry to an image recognition engine on the server side via a network about selected images, various general objects and the like contained in a designated image are extracted and recognized in real time, and by notifying a relationship search engine on the server side via the image recognition engine of image constituent element groups contained in the recognized input images, the relationship search engine extracts related element groups with respect to the individual image constituent element groups, and the related element groups as a relationship graph having the respective extracted related element groups as nodes are visually displayed together with the depths of the relationships between the nodes.

(57) 要約: ユーザ固有の関心の対象や領域を、インタレスト・グラフとして、網羅的、かつ効果的に獲得するシステムを提供する。本発明に係るシステムは、様々な物体や主題を内包する画像情報を入力手段として用いる検索システムであって、選択した画像をネットワークを経由してサーバ側の画像認識エンジンに問い合わせる事により、指定された画像に含まれる様々な一般物体等をリアルタイムに抽出・認識し、それら認識された入力画像に含まれる画像構成要素群を、画像認識エンジン経由でサーバ側の関連性検索エンジンに通知する事で、関連性検索エンジンが個々の画像構成要素群それぞれに対し関連要素群を抽出し、抽出された関連要素群それぞれをノードとする関連性グラフとして、各ノード間の関連性の深さと共に視覚的に表現する。



(84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI

(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:

— 国際調査報告 (条約第 21 条(3))

## 明 細 書

発明の名称：

画像認識システムを組み込んだ関連性検索によるインタレスト・グラフ収集システム

### 技術分野

[0001] 本発明は、ネットワーク経由でサーバ側に構築されている画像認識エンジンを備えた関連性検索システムを利用する事で、上記画像認識エンジンによって認識可能となった画像構成要素群、及び上記関連性検索システムによって上記各画像構成要素に関連付けられたその他の要素間の繋がりと、繋がりの深さを視覚化してユーザのネットワーク端末上に提示する事で、ユーザによる視覚的な対象物の検索や探索を可能にするシステムを提供すると共に、その過程を通じてユーザの関心の在り所やその関心の推移をサーバ側で分析し、それらをユーザ毎の、あるいは特定のユーザ群毎の、或いはユーザ全体に亘るインタレスト・グラフとしてサーバ側で収集するシステムに関する。

### 背景技術

[0002] 旧来、ユーザ毎の関心を把握する代表的な手段として、書面或いはそれに代わる方法でユーザにアンケートを送付し、いくつかの選択可能な候補群の中からユーザにとって特に関心のあるジャンル、さらには特定の商品・サービスに関する認知度等を段階的な評価尺度と共に求め、それらの回答を集計してマーケティング活動に利用する方法が多用されていた。また近年ではインターネットを活用した各種サービスの登場により、契約開始時のユーザ登録の一環として、ユーザにとって特に関心のあるジャンル等の記入を求める事等により、ユーザ毎の嗜好を反映した関連商品やサービスの提供が可能となっている。

[0003] また、インターネットを活用した物品販売を行ういくつかのサイトでは、ユーザの購入履歴やサイトの閲覧履歴から、お勧めの商品や関連するサービス群を追加的に提示したり、同一商品を購入した他のユーザが他に何を購入

したかの履歴情報を基に、それら同様の商品のリコメンデーションをまだ未購入のユーザの端末上に提示したりするなど、より多様に分化するユーザ群に向けて、より適合精度の高い商品やサービスの提供が可能となっている（特許文献1）。

[0004] さらに、近年のソーシャルネットワーク・サービスの拡大に伴い、それぞれのユーザ・プロフィールに対しユーザ自身で自らの関心領域や好み等を記載したり、他のユーザが投稿した文章や写真或いは動画等に対して「いいね！（ＬＩＫＥ）」釦をクリックする事により、当該ユーザからの好意的な反応をサイトに反映させると共に、これらＬＩＫＥ情報に基づいた新たな試みも既に始まっている（特許文献2）。また、140文字以内というような短いつぶやきをインターネット上で発信するサービスでは、特定の発信者や特定の話題に対し、関心のある多数のユーザがフォローするという特性を活用して、その主たる話題の内容やテーマを分類・解析して、ユーザの関心の在り所を効果的に求めようというアイデアも提案されている。

[0005] ユーザの関心の在り所を求めようとする例として、ユーザの閲覧したファイル間を伝播している語から、ユーザの変わりゆく興味をリアルタイムに推定するアルゴリズムを実装した装置がある（特許文献3）。

[0006] すなわち、特許文献3に開示された装置は、ユーザが閲覧した履歴から複数のファイルに含まれる語をファイル毎にテキストとして入力する手段と、そのテキストから単語単位に分割する手段と、ユーザが閲覧した複数のファイル間でユーザが参照した「伝播する語」を抽出する手段と、一または複数の前記「伝播する語」を記憶する手段と、前記「伝播する語」の全てのファイルに対する出現頻度から所定の「影響度」と、「伝播する語」の特定のファイル内に出現する程度を表す所定のi D F値を求める手段と、「影響度」と前記i D F値との関数である「影響度i D F値」に応じてユーザの興味ある語の集合をユーザ・プロフィール情報として抽出する手段とを備える。

[0007] また、コンテンツ体系をユーザとアイテムの関係からなるグラフ体系で表現し、ユーザが意味情報に依拠して興味の向くコンテンツを容易且つ的確に

探索することを可能にする装置等も開示されている（特許文献４）。

[0008] すなわち、特許文献４に開示された装置等は、各個人毎の興味がクラス階層化された興味オントロジを表す興味オントロジデータの供給を受けて供給された興味オントロジ間の近似度を計測することによりユーザ間の興味の近似度を計測する近似度計測手段と、上記近似度計測手段による計測結果に基づいて興味オントロジ間の近似度が所定の範囲内にあるユーザコミュニティを識別可能なユーザグラフのデータを形成するユーザグラフ形成手段と、上記ユーザグラフ形成手段によって形成されたユーザグラフのデータを管理すると共に当該ユーザグラフのデータに基づくグラフの各ノードを成す複数のユーザ間を結ぶエッジにタクソノミによる意味情報を付与してユーザ間の関係をグラフベースで再構築するユーザグラフ再構築手段とを備えたことを特徴とする。

## 先行技術文献

## 特許文献

- [0009] 特許文献１：特開２００１－２２２５８６号公報  
特許文献２：米国特許第７７８８２６０号明細書  
特許文献３：特開２００７－２６４７１８号公報  
特許文献４：特開２００９－２０５２８９号公報

## 発明の概要

## 発明が解決しようとする課題

- [0010] しかしながらユーザの興味がある領域の取得という点では、旧来のアンケートやユーザ登録時のプロフィール記入による手法では、個々のユーザにとっては大雑把な領域に対する好みや関心の表明に止まり、また、近年のソーシャルネットワーク・サービス中で多用される「ＬＩＫＥ！」釦にしても、自分の関心領域であるというより友人の関心領域を間接的にフォローしているに過ぎないとも言える。そもそも友人リスト自体が必ずしも自分の友人全体を包括的に指し示すものでもなかった。また１４０文字以内のネットワー

ク上の様々なつぶやきに対してフォローが可能なサービスにおいては、そのつぶやきの頻度、リアルタイム性、話題の新規性という点からも、従来のネットワーク・サービスに比べてより新鮮なユーザの関心領域の取得が可能になると期待されているものの、自分の何気ないつぶやきや他の人のつぶやきをフォローする事で、ユーザ自身の持つ関心領域が網羅的かつ的確に取得可能かという点、その対象となる領域自体が極めて限定的なものとなっていた。

[0011] また、特許文献3や特許文献4に開示された技術においても、文書や単語を前提とした処理が行われている。そもそも文字による表意的な表現には、その背景となる文化や慣習の相違が色濃く反映しており、世界中の人々にとって直観的かつ共通のコミュニケーション手段であるとは、およそ言えない。古来、「百聞は一見にしかず」の諺にあるように、数多の言葉よりたった一枚の映像の方が的確に状況を表わしているケースが多い。映像はその中に主題や状況の描写と共に様々なオブジェクトを内包しており、どのような国の人々であってもおよそ認識可能である。人間以外の動物も、かなりの部分の情報を視覚から獲得していて、瞬時に周囲の状況を把握し次の行動判断を行っている。しかるに、従来の文字主体の方法では多言語であるが故の表現上の曖昧さが残ると同時に、刻々と遷り変るユーザの関心の在り所や好奇心の行き先等を効果的かつリアルタイムに獲得する、もしくはインタラクティブに獲得しようとする事は困難であった。

[0012] さらに、膨大な数の人々を対象とし、それら個々の人々の関心のある領域、或いはありそうな領域を効果的かつ網羅的に獲得する事は時間的にも方法論的にも従来は極めて困難で、一定の範囲内における関心の捕捉に止まり、ユーザの関心の在り所をより広範な領域に拡げ効果的に探索し獲得する方法が待ち望まれていた。

[0013] そこで本発明では、文字を介さず主題や周囲状況と共に様々なオブジェクトを内包する画像情報を活用し、ユーザの刻々と遷り変る関心の在り所を、ユーザによる興味のある画像の検索および探索という過程で効果的に捕捉す

る為に、それらの画像に含まれる個々の画像構成要素を画像認識エンジンの助けを借りてリアルタイムに検出し、それら認識された個々の画像構成要素とさらに関連性の高い他の関連要素群を、関連性検索システムの助けを借りて、併せて可視化する事で、ユーザが視覚的かつインタラクティブに興味のある対象を検索したり探索可能とさせる事で、従来に比べ効果的にユーザ固有の関心の対象や領域を、サーバ側でユーザ毎の、或いは特定のユーザ群の、或いはユーザ全体に共通のインタレスト・グラフとして、網羅的、かつ効果的に獲得する事を目的としている。

### 課題を解決するための手段

[0014] 本発明に係るインタレスト・グラフ収集システムは、一形態として、キーワードやメタデータ、或いは文章といったような表意的な文字による入力手段によらず、様々な物体(object)や主題(subject)を内包する画像情報を入力手段として用いる検索システムであって、インターネット上或いは専用ネットワーク上に存在する多数の画像群、或いはネットワーク端末経由でユーザがインターネット上にアップロードした画像の中から、ユーザが関心を持った画像全体、或いは画像の特定の領域をユーザがネットワーク端末上で選択し、それら選択した画像をネットワークを経由してサーバ側の画像認識エンジンに問い合わせる事により、インターネットを介してサーバ側の画像認識エンジンが、選択された画像全体、或いは指定された画像領域に含まれる様々な一般物体・特定物体・人・顔・シーン・文字・記号・イラストレーション、ロゴ、ファビコン(等)をリアルタイムに抽出・認識し、それら認識された入力画像に含まれる画像構成要素群を、画像認識エンジン経由でサーバ側の関連性検索エンジンに通知する事で、関連性検索エンジンが個々の画像構成要素群それぞれに対し直接的及び間接的に一定以上の関連があると判断したその他の関連要素群を、関連性検索エンジン内の関連性知識データベース上に学習可能な状態で収納されている要素間の直接関連性を記述した多次元特徴ベクトルを基に抽出し、上記画像認識エンジンにより認識された画像構成要素群及び関連性検索エンジンにより抽出された関連要素群それぞれをノ

ードとする関連性グラフとして、各ノード間の関連性の深さと共に、ユーザのネットワーク端末上に二次元画像、或いは奥行きを持った三次元画像、或いは上記関連性グラフの観察時間としての時間軸変数を加えた四次元時空間画像として視覚的に表現する事を可能にしたことを特徴とする。

[0015] また、他の形態として、本発明に係るインタレスト・グラフ収集システムは、上記関連性検索操作において、ネットワーク端末上に表示されている関連性グラフ上の任意のノードを、ユーザがタッチスクリーン上でタップ或いはタッチして選択するか、ポインタのカーソルを任意のノード上に移動し選択する、或いは関連性グラフ上の任意の領域に向けてユーザがタッチスクリーン上でフリックするか、ポインタのカーソルを関連性グラフ上の任意の領域に移動し画面全体をドラッグしスクロールする、或いは方向キー等による同様の操作、或いはユーザによるジェスチャーや視線、音声、或いは脳波を用いた同様の効果を及ぼす入力操作を用いる事により、当該選択されたノード、或いは移動後の領域を中心とする新たな関連性グラフを、そこに至る途中経過も含めて関連性検索エンジンがネットワーク端末に向け追加的に送出する事で、ユーザにとって関心のあるノード或いは領域をユーザが関連性グラフ上でシームレスに辿りながら、複数のノード間にまたがる広範な関連性として視覚的に認識する事が可能に構成されたことを特徴とする。

[0016] また、他の形態として、本発明に係るインタレスト・グラフ収集システムは、上記関連性検索操作において、上記画像認識エンジンが提示した複数の画像構成要素群の中から、ユーザが選択した特定の画像構成要素、或いはネットワーク端末上に表示されている関連性グラフ上の特定のノードをユーザがタッチスクリーン上でダブルタップするかピンチアウトする操作、或いはポインタ等の操作による当該ノードの選択と当該ノードを中心とする領域の拡大、或いはユーザによるジェスチャーや視線、音声、或いは脳波を用いた同様の効果を及ぼす入力操作を用いる事により、当該ノードを中心とするより詳細な関連性グラフをユーザのネットワーク端末上に視覚的に表現する事を可能にした上で、それらの一連の操作を当該ノードに対するユーザの一定

の関心の存在であるとし、当該ユーザを中心ノードとする各要素間の直接関連性を記述した多次元特徴ベクトル上で、当該ノードに係るユーザの関心の深さを表わす特徴ベクトル値を適応的に強める事で、当該ユーザを中心ノードとする個々のユーザに対応したインタレスト・グラフとして獲得可能にすると共に、当該インタレスト・グラフを広範なユーザに拡大して獲得する事により、特定のユーザクラス、或いはユーザ全体にまたがる統計的な広義のインタレスト・グラフとして収集可能に構成されたことを特徴とする。

[0017] また、他の形態として、本発明に係るインタレスト・グラフ収集システムは、上記関連性検索操作において、ユーザが着目し選択したノードを関連性グラフ上で辿る事なく、それらノード画像をネットワークを經由してサーバ側の画像認識エンジンに再び問い合わせる事により、画像認識エンジンの助けを借りて当該ノードに関する新たな画像構成要素群を獲得し、それら画像構成要素群を起点とする新たな関連要素群を関連性検索エンジンからネットワーク端末に向け送出する事で、ユーザが当該ノードに対する新たな関連性をそれら相互の関連性の深さと共に関連性グラフ上で視覚的に認識する事を可能にした上で、直前の同様の操作において当該ノードへの起点となっている画像構成要素から当該ノードに至る一連のノード間の関連性の存在をユーザが認識し利用していると関連性検索エンジンが推測し、各要素間の直接関連性を記述した多次元特徴ベクトル上で、それら一連の関連性を構成するそれぞれのノード間の直接的な関係性の深さを表わす特徴ベクトル値を適応的に強める事で、関連性検索エンジン内の関連性知識データベースの追加学習が更に可能になるよう構成されたことを特徴とする。

[0018] また、他の形態として、本発明に係るインタレスト・グラフ収集システムは、上記関連性検索操作において、画像認識エンジンにより認識可能となった画像構成要素群、及びそれら画像構成要素群それぞれに対応する各関連要素群に対し、それらを代表する写真、イラストレーション、文字、記号、ロゴ、ファビコン等から生成した縮小画像サムネイルを関連性検索エンジンが元画像に代えてネットワーク端末に向け送出する事により、関連性グラフ上

のノードとして画像サムネイル単位での表示・選択が更に可能になるよう構成されたことを特徴とする。

[0019] また、他の形態として、本発明に係るインタレスト・グラフ収集システムは、上記関連性検索操作において、複数のノードをサーバ側の画像認識エンジンに問い合わせる事を可能にした上で、画像認識プロセスに備わる入力条件選択機能として、論理的な演算子（AND、OR）を導入し、ANDを選択した場合はそれぞれのノード間で共通かつ直接的に関連するノードを、ORを選択した場合はそれぞれのノードのいずれか一つ以上に直接的に関連するノードを、相互の関連度の深さと共にネットワーク端末上に視覚的に表現する事が更に可能になるよう構成されたことを特徴とする。

[0020] また、他の形態として、本発明に係るインタレスト・グラフ収集システムは、上記関連性検索操作において、複数のノードをサーバ側の画像認識エンジンに問い合わせる事を可能にした上で、画像認識プロセスに備わる入力条件選択機能として、関連性探索演算子（Connection Search）を導入し、一見全く関連性がないと思われるような複数のノード間の関係を、それぞれの入力ノード群に対し直接的及び間接的に関連するその他のノードを経由する一連の関連性として探索する事で、異なるレイヤ（階層）に及ぶノード間の間接的な関係を発見し、それらノード間の最短パスを含む関連性グラフとしてネットワーク端末上に表示する事を可能にすると同時に、上記関連性探索プロセスにおいて、発見された複数のノード間にまたがる上記間接的関係を、関連性検索エンジン内の関連性知識データベース内に学習の上獲得する事で、以降の同様或いは類似の関連性探索要求に備える事が更に可能になるよう構成されたことを特徴とする。

[0021] また、他の形態として、本発明に係るインタレスト・グラフ収集システムは、上記関連性検索操作において、ユーザと間接的な関係にあるノード、或いはユーザとの関連性が凡そないとされているその他のノードに対し、それらノードを当該ユーザとの直接的な関係として結び付ける接続演算子（LIKE）、及び既に結びつけられている当該ノードとユーザ間の直接的な関

係を切断する切断演算子（D I S L I K E）を導入する事で、当該ユーザを中心ノードとする各要素間の直接関連性を記述した多次元特徴ベクトル上で、当該ノードに係るユーザの関心の深さを表わす値を増加、或いは減少、或いは滅消させる事で、当該ユーザを中心ノードとする、個々のユーザに対応したインタレスト・グラフの更新が更に可能になるよう構成されたことを特徴とする。

[0022] また、他の形態として、本発明に係るインタレスト・グラフ収集システムは、上記関連性検索操作において、ユーザ以外のノードを対象にした新たな直接的関係の存在及び非存在の可能性を、それらの複数のノードが直接的に結び付けられるべきだとして提起する参照演算子（R E F E R E N C E）、及び既に直接的に結び付いているもののその直接的な関連の存在が疑わしいとして直接的関係の非存在を提起する非参照演算子（U N R E F E R E N C E）を導入する事で、それらの新たなノード間の直接関連性の存在、或いは非存在の可能性を関連性検索エンジンがユーザに対し喚起する事を可能にした上で、特定の権限を有するスーパーバイザ或いは一定数以上のユーザから関連がある或いはないと判断されたノード間の関連性に係る特徴ベクトルの値を関連性検索エンジンが更新可能にした上で、当該ノード群に係る更新された関連性グラフとしてネットワーク端末上に反映可能にすると共に、それらの新たな直接関連性の存在或いは非存在に係る更新情報を、全てのユーザに通知する事が更に可能になるよう構成されたことを特徴とする。

## 発明の効果

[0023] 本発明に係るシステムによって、文字による検索という多言語対応が必要となる情報検索手段から、文字によらず画像そのものを入力手段とする情報検索処理が可能となる事で、より広範な国と地域のユーザを対象とする言語フリーの検索システムが提供可能となる。また、検索入力と検索結果の双方を従来の文字から画像情報に置き換える事で、人間にとってより直観的な情報の検索・発見が可能となる。併せて、表示画面サイズの比較的小さなネットワーク端末群であっても、入出力操作を文字から画像サムネイルやアイコ

ン等の画像ベースのユーザ・インターフェース（UI）に移行させる事で、指先一つ或いはユーザによる簡単なポインティング操作で高度な検索操作を可能にするなど、従来に比べより軽快な検索環境を提供出来る事から、従来より頻度の高い検索や探索を喚起する事が可能となり、それらをサーバ側で統計的に処理する事で、より鮮度の高いダイナミックなインタレスト・グラフの取得が可能になるという効果を奏する。

### 図面の簡単な説明

[0024] [図1]本発明に係るシステムの一実施形態におけるシステム構成を説明する説明図である。

[図2]本発明に係るシステムの一実施形態におけるサーバ及びM D Bの機能ブロックを説明する説明図である。

[図3]本発明に係るシステムの一実施形態におけるネットワーク端末装置の機能ブロックを説明する説明図である。

[図4]本発明に係るシステムの一実施形態における画像認識フローを説明する説明図である。

[図5]本発明に係るシステムの一実施形態における他の画像認識フローを説明する説明図である。

[図6A]本発明に係るシステムの他の実施形態における画像認識システムの機能ブロックを説明する説明図である。

[図6B]本発明に係るシステムの他の実施形態における一般物体認識部の機能ブロックを説明する説明図である。

[図7]本発明に係るシステムの他の実施形態における局所特徴量抽出部の機能を説明する説明図である。

[図8]本発明に係るシステムの他の実施形態におけるカテゴリ認識部の機能ブロックを説明する説明図である。

[図9]本発明に係るシステムの他の実施形態におけるシーン認識部の機能ブロックを説明する説明図である。

[図10]本発明に係るシステムの他の実施形態におけるシーン分類の一例を説

明する説明図である。

[図11]本発明に係るシステムの他の実施形態における特定物体認識部の機能ブロックを説明する説明図である。

[図12A]本発明に係るシステムの一実施形態におけるユーザインタフェースを説明する説明図である。

[図12B]本発明に係るシステムの他の実施形態におけるユーザインタフェースを説明する説明図である。

[図12C]本発明に係るシステムの他の実施形態におけるユーザインタフェースを説明する説明図である。

[図12D]本発明に係るシステムの他の実施形態におけるユーザインタフェースを説明する説明図である。

[図12E]本発明に係るシステムの他の実施形態におけるユーザインタフェースを説明する説明図である。

[図13]本発明に係るシステムのグラフ演算部の機能ブロックを説明する説明図である。

[図14A]本発明に係るシステムの一実施形態におけるグラフを表現するための基本的データ構造を説明する説明図である。

[図14B]本発明に係るシステムの一実施形態におけるノードデータとリンクデータとを操作する為のオペレーションを説明する説明図である。

[図14C]本発明に係るシステムの一実施形態におけるグラフ構造及びリンク表現を説明する説明図である。

[図15]本発明に係るシステムの一実施形態における関連画像とのリンク構造例及び動作例を説明する説明図である。

[図16]本発明に係るシステムの一実施形態におけるシナリオと対応するグラフ構造等を説明する説明図である。

[図17]本発明に係るシステムの一実施形態におけるシナリオと対応するグラフ構造等を説明する説明図である。

[図18]本発明に係るシステムの一実施形態における関連性導出演算の例を説

明する説明図である。

[図19]本発明に係るシステムの一実施形態におけるインタレスト・グラフ獲得例を説明する説明図である。

[図20A]本発明に係るシステムの一実施形態におけるグラフ構造の表示例を説明する説明図である。

[図20B]本発明に係るシステムの他の実施形態におけるグラフ構造の表示例を説明する説明図である。

[図20C]本発明に係るシステムの他の実施形態におけるグラフ構造の表示例を説明する説明図である。

[図21]本発明に係るシステムの他の実施形態における動作例を説明する説明図である。

[図22]本発明に係るシステムの他の実施形態における動作例を説明する説明図である。

[図23]本発明に係るシステムの他の実施形態における具体的な連想関係を説明する説明図である。

[図24A]本発明に係るシステムの他の実施形態における動作例を説明する説明図である。

[図24B]本発明に係るシステムの他の実施形態における動作例を説明する説明図である。

[図25]本発明に係るシステムの他の実施形態における動作例を説明する説明図である。

[図26]本発明に係るシステムの他の実施形態における動作例を説明する説明図である。

[図27]本発明に係るシステムの一実施形態における統計情報処理部の機能ブロックを説明する説明図である。

[図28]本発明に係るシステムの一実施形態における特定ユーザフィルター処理部の機能ブロックを説明する説明図である。

[図29]本発明に係るシステムの一実施形態における表現例を説明する説明図

である。

[図30]本発明に係るシステムの他の実施形態における動作例を説明する説明図である。

[図31]本発明に係るシステムの一実施形態におけるインタレスト・グラフの構成例を説明する説明図である。

[図32]本発明に係るシステムの他の実施形態におけるインタレスト・グラフの表示例を説明する説明図である。

[図33]本発明に係るシステムの一実施形態におけるソーシャル・グラフ獲得処理を説明する説明図である。

[図34]本発明に係るシステムの一実施形態におけるインタレスト・グラフ収集プロセスを説明する説明図である。

[図35]本発明に係るシステムの他の実施形態におけるユーザ毎のインタレスト・グラフに対応する多次元特徴ベクトルを説明する説明図である。

### 発明を実施するための形態

[0025] 以下、本発明に係るシステムを実施するための形態について、図を参照しながら詳述する。

[0026] 図1に、本発明に係るシステムの一実施形態を示す。システム100は、サーバ101と、グラフデータベース（以下、「GDB」ともいう）102Aと、マザーデータベース（以下、「MDB」ともいう）102Bと、ユーザが使用する複数台のネットワーク端末装置105a～105dとからなり、サーバ101とGDB102A及びMDB102Bとは接続103で接続され、サーバ101及びネットワーク装置105はネットワーク或いはインターネット104に接続されている。

[0027] ここで、サーバとは、クライアントからの要求に応じてデータを処理し、その結果をサービスとして提供する1又は複数のコンピュータプログラムであり、1台のコンピュータシステム上に実装することも、複数のコンピュータからなるシステム群に分散して実装することもできる。また、他のサーバ機能と並列して1又は複数のコンピュータシステム上に実装することもでき

る。さらに、独立した複数の処理機能を有するように構成することもできる。以下、本明細書では、サーバの意義を上記のとおり位置付ける。

[0028] また、ハードウェアとしてのコンピュータシステムは、最も基本的な構成としては、命令バス及びデータバスで接続された算術論理演算ユニットと制御ユニットと記憶装置と入出力装置とを備えた電子計算機である。入出力インタフェースを介して入出力装置から入力された情報（ビットデータ）に基づき算術論理演算ユニットにおいて算術演算や論理演算、比較演算、シフト演算等が実行される。実行されたデータは必要に応じて記憶装置に記憶され、入出力装置から出力される。これら一連の処理は、記憶装置に記憶されたソフトウェアプログラムによって制御される。本発明の実施形態において使用される各サーバマシンも、上記のとりのコンピュータとしての基本機能を最低限備えたハードウェアであり、オペレーティングシステムやデバイスドライバ、ミドルウェア、そしてアプリケーションソフトウェアといったプログラム群によって制御されている。

[0029] 図2に、本発明に係るシステムの一実施形態におけるサーバ101、ならびに、GDB102A及びMDB102Bの機能ブロックを示す。サーバ101は、ソフトウェア機能ブロックとして、領域処理部201と、一般物体認識部202と、特定物体認識部203と、MDB検索部206と、MDB学習部207と、MDBマネージメント部208と、ネットワーク通信制御部204と、データ検索処理部205と、グラフ演算部221と、グラフ記憶部222と、グラフ管理部223と、関連性演算部224と、統計情報処理部209と、特定ユーザフィルター処理部210とからなる。

なお、領域処理部201、一般物体認識部202、特定物体認識部203、MDB検索部206、MDB学習部207、及びMDBマネージメント208は、画像認識エンジン200を構成している。画像認識エンジン200は後述の図6A記載の画像認識システムに置き換えても良い。また、グラフ演算部221、グラフ記憶部222、グラフ管理部223、及び関連性演算部224は、関連性検索エンジン220を構成している。

サーバ１０１の機能ブロックは必ずしもこれらに限定されるものではないが、これら代表的な機能について簡単に説明する。

[0030] 領域処理部２０１は、画像内の領域分割や部分画像の切出しなどを行う。一般物体認識部２０２は、画像中に含まれる物体を一般的な名称（カテゴリ）で認識する。また、物体認識部２０３は、MDBに登録された情報と照合し、物体を同定する。

ネットワーク通信制御部２０４は、画像の入力出力処理や、ネットワーク端末との情報通信制御などを行う。データ検索処理部２０５は、リンク先から情報を収集したり、集合知の問合せや、収集、ならびに検索などを行う。

MDB検索部２０６は、物体の名称等のタグデータ等を検索する。MDB学習部２０７は、新規設計データの追加や詳細情報の追加、時間情報の登録、付帯情報の登録、更新、追加などを行う。MDBマネジメント部２０８は、設計データから特徴点や特徴量の抽出、付帯情報からカテゴリ情報を抽出しカテゴリデータへの登録や、カテゴリデータ内のカテゴリ分類の拡張、分割、更新、統合、修正、新規カテゴリの登録などを行う。

[0031] 関連性検索エンジン２２０は、上述したとおり、少なくともグラフ演算部２２１と、グラフ記憶部２２２と、グラフ管理部２２３と、関連性演算部２２４とを含む。グラフ演算部２２１は、サーバ上で実行される各種グラフ演算を処理し、グラフ記憶部２２２は、グラフデータベースに格納されたノードデータやリンクデータを用いたグラフ構造をメモリ上に展開し、後段の処理が行いやすいようにデータ形式を整え、グラフ管理部２２３は、グラフ演算部２２１で実行される多数のグラフ演算の管理・調停を行う。さらに、関連性演算部２２４では、ノード間の関連性をグラフマイニング手法を用いて計算する。

[0032] 統計情報処理部２０９は、GDB１０２Ａに蓄積されたグラフデータを用いた統計情報処理を行う。特定ユーザフィルター処理部２１０は、ユーザの主観に基づいた検索結果のフィルタリングを行う。例えば、各ノードに付与されたタイプ情報から部分グラフを抽出してグラフマイニング処理をするこ

とで共起確率に基づくユーザの興味を処理することができる。

[0033] また、GDB102Aは、ノードデータ231と、リンクデータ232とからなる。GDB102Aは必ずしもこれらに限定されるものではないが、これら代表的な機能について簡単に説明する。

[0034] ノードデータ231は、ノードに関するデータを格納する。データ構造の一例については、図14A(D)に基づいて後述する。

リンクデータ232は、リンクに関するデータを格納する。リンク構造の一例については、図14A(E)に基づいて後述する。

[0035] また、MDB102Bは、設計データ251と、付帯情報データ252と、特微量データ253と、カテゴリデータ254と、不特定物体データ255とからなる。MDB102Bは必ずしもこれらに限定されるものではないが、これら代表的な機能について簡単に説明する。

[0036] 設計データ251は、物体を製造するためのデータベースから生成された、物体の構造や形状、寸法、部品の接続情報、配置図、可動部、可動範囲、重量、剛性など、物体を構成し、製造するのに必要な基本情報を保持している。

付加情報データ252は、物体の名称、製造者、部品番号、日時、素材、組成、加工情報など物体に関するあらゆる情報を保持している。

特微量データ253は、設計情報に基づいて生成される個々の物体の特徴点や特微量情報を保持している。

カテゴリデータ254は、一般物体認識部において物体のカテゴリ分類を行う際に使用する情報を保持している。

不特定物体データ255は、現時点で特定物体認識不能の物体に関する情報を保持しており、その後に類似の特徴を有する物体が頻繁に検出されれば、新たな特定物体として新規登録処理される。

[0037] 図3に、本発明に係るシステムの一実施形態におけるネットワーク端末装置を示す。ネットワーク端末装置105a～105dは、広くユーザが使用するクライアント端末装置であり、コンピュータ、携帯情報端末(PDAや

パッド)、携帯電話機等が含まれる。つまり、ネットワーク端末装置105a~105dは、多様な種類の電子情報機器が多数インターネットなどのネットワークに接続されている様子をあらわしている。以下、ネットワーク端末装置105と記載した場合にはネットワークに接続されたネットワーク端末装置105a~105dのいずれか1台をさす。ネットワーク端末装置105a~105dはすべて同一機種である必要はない。同等の機能(あるいは実施可能な最低限の機能)を備えた端末装置であればよい。ここではネットワーク端末装置105の代表的な機能ブロックについて説明する。

[0038] 図3のネットワーク端末105の事例として、動画像入力機能と表示機能が一体で存在する場合と、それぞれが別体で存在する場合とがある。前者では携帯電話や最新のスマートフォンのように、ネットワーク端末105には、操作部105-01と、表示部105-02と、音声入出力部105-03と、画像送受信部105-04と、カメラ部105-05と、ネットワーク通信部105-06と、CPU105-07と、記憶部105-08と、電源部105-09と、位置情報取得部105-10と、各種センサ群105-11とが一体として備わっており、後者では動画撮影カメラとTVのように、入出力機能がそれぞれ別体のものとして存在している。

[0039] 本発明においては、これら入出力部が一体となったネットワーク端末を事例にとり説明を行う。操作部105-01は、例えばタッチパッド(ディスプレイに内蔵されるものを含む)、キー入力部、ポインティングデバイス、ジョグダイヤル等の入力デバイスで構成される。表示部105-02は、各出力機器に応じた解像度及びビデオメモリを備えたディスプレイ部である。音声入出力部105-03は、音声認識用のマイク、スピーカ等の入出力デバイスで構成される。画像送受信部105-04は、ネットワーク端末105で撮影した動画像データをサーバに送信し、あるいはサーバから配信されて来る動画像データを受信するために必要なコーデック部、メモリユニット等が含まれる。なお、動画像データには、静止画像も含まれる。カメラ部105-05はCCDやMOSセンサ等の撮像デバイスを含む動画像撮影手段

である。ネットワーク通信部 105-06 はインターネット等のネットワークに接続するためのインタフェースであって、有線／無線のいずれかである事を問わない。CPU 105-07 は中央処理装置であり、記憶部 105-08 はフラッシュメモリ等の一時記憶装置であり、電源部 105-09 はネットワーク端末全体に電力を供給する為の電池等を指す。位置データ検出部 105-10 は GPS 等の位置情報検出デバイスであり、各種センサ群 105-11 には加速度センサ、傾きセンサ、磁気センサなどが含まれる。

[0040] A. 画像認識処理

次に、図 4 に従って、本発明に係るシステムの一実施形態における画像認識システムの全体的なフローを説明する。

[0041] 画像認識処理の開始 (S401) は、一例として、ネットワーク端末装置 105 からアップロードされたり、サーバからクローリングして収集されることによる元画像の入力からはじまる (S402)。元画像は、もともとサーバ上に存在するものを使用してもよい。元画像は、2次元画像であるか3次元画像であるかを問わない。また、元画像の入力に際しては、その元画像におけるどの物体の着目領域の指示がポインティングデバイス等のデバイス (不図示) を通じてなされる場合と、着目点の指示がなく元画像全体が処理対象として入力される場合がある。次に、S404 において一般物体認識処理が行われる。一般物体認識処理には、例えば、BoF (Bag-of-Features) の手法を採用することができる。この一般物体認識処理では、検出された物体のカテゴリ (物体の一般名称) の認識まで行われる。ただし、着目点の指示がされた場合は、カテゴリの認識ができた場合とできなかった場合とで処理が分岐し、その判断は S405 で行われる。カテゴリ認識ができなかった場合には、S406 へ進み、既存のカテゴリの取り扱いについて判断がなされる (S407 或いは S408) が、着目点に指示の有無に関わらず物体のカテゴリ認識が出来た場合には、特定物体認識処理へ進むべく S409 へ進む。

[0042] S405 の判断において特定物体認識処理へ進む場合は、まず、S409

において個別物体画像の切り出し処理が行われる。そして、切り出された個別物体画像について、特定物体認識処理が行われる（S 4 1 0）。特定物体認識処理では、M D B 1 0 2 B に登録された設計データ 2 5 1 から抽出された特徴量データ 2 5 3 に基づく一致度を算出する評価関数によって物体の同定が試みられる。

[0043] 一方で、S 4 0 5 の判断において一般物体認識ができなかった場合には、S 4 0 6 へ進み、着目物体の持つ特徴量と、M D B 1 0 2 B が把握している既存のカテゴリに属する物体の特徴量との情報距離に基づいて、その着目物体を含む新たなカテゴリを登録するか（S 4 0 7）、その着目物体に近接する既存カテゴリの拡張を検討するか（S 4 0 8）の判断が行われる。新たなカテゴリを登録する（S 4 0 7）場合にはS 4 0 4 に復帰し、既存のカテゴリを拡張する（S 4 0 8）場合にはS 4 0 9 へ進む。

[0044] S 4 1 1 においては、特定物体の同定ができたかどうか判断される。特定物体の同定ができた場合には、S 4 1 3 に進み、S 4 0 9 において切り出された個別物体画像にM D B 1 0 2 B に登録されている物体の詳細データよりもさらに精細な情報が含まれるかが判断される。もし、S 4 1 3 においてY e s と判断された場合には、S 4 1 4 に進み、M D B 1 0 2 B におけるその物体の詳細データはM D B 学習部 2 0 7 により更新され、より精細な情報を有するようになる。一方で、S 4 1 3 においてN o と判断された場合には、S 4 1 5 に進み、次の判断がなされる。

[0045] S 4 1 5 は、S 4 0 5 において一般物体認識ができなかったと判断された場合であって、S 4 0 6 の判断でS 4 0 8、S 4 0 9、S 4 1 0 へ進み、特定物体の認識ができた（S 4 1 1 でY e s）場合に判断される。S 4 1 5 において、特定された物体が既存のカテゴリであった場合には、M D B 1 0 2 B に登録されているその既存カテゴリの定義を拡張、あるいは拡張によりカテゴリ内の物体の情報距離が分散する場合には分割、あるいは近接カテゴリとの情報距離が前記カテゴリ内の物体間の情報距離と同程度以下になった場合には統合、あるいは特定された物体の登録により既存の物体の情報の齟齬

を発見した場合には修正を施し、カテゴリデータ 254 をアップデートする (S416)。一方で、S415において、特定された物体が既存のカテゴリでなかった場合には、S407へジャンプし、新規カテゴリとして登録する。

[0046] S411において、特定物体の認識ができなかった場合は、MDB102Bにおいてその物体はいったん「未確認物体」として登録され、将来の処理に備えて認識処理を終了する (S417)。S416において既存カテゴリを拡張してアップデートされた場合も認識処理を終了する (S417)。

[0047] 図5は、図4における特定物体認識処理及び学習処理の一部について、他の実施形態を示すフロー図である。以下、詳細に説明する。

[0048] まず、S501より特定物体認識処理が開始される。ここで入力されるデータは、単一物体の画像のほか、同レイヤの設計データを使用することができる。さらに、画像にリンクされた設計データや、設計データそれ自体（製品全体のみならず、部品であってもよい）を使用することもできる。

次に、S502においてMDBで生成された特徴量データ253を元に、元画像内の特徴点及び特徴量が抽出され、MDBより生成される特徴量データと比較される。ここで、MDBより特徴量データを生成・比較する方法には、以下の2種類がある。

[0049] 1つ目は、物体を構成する最小単位（設計データなどで表されている）ごとの3次元情報を元に、あらゆる角度から2次元平面に写像し、その写像画像から物体の特定に用いる特徴量などを生成する。比較する際には、その特徴量を元に入力画像から特徴量を抽出し、出現部位や頻度などを比較する方法である (S504)。なお、ここでの特徴量は、一例として輪郭抽出法やSURF法やなどに基づいて生成される。

[0050] 2つ目は、物体を構成する最小単位（設計データなど）の集合からなる3次元形状情報を投影角度や拡大率などを変えながら2次元平面に写像するプロセスを評価関数として、物体の特徴点および特徴量との差異を一致度として判定する方法（チューン方式）である (S505)。

- [0051] なお、従来は、サンプルとなる画像そのものをとにかくたくさん集めて特徴量あるいは評価関数による同定処理（S 5 0 2）を行なっていたことに鑑みれば、S 5 0 4 や S 5 0 5 で説明したMDBより生成する方法（S 5 0 3）は、同定の確度を上げるという点で従来技術より有利な効果を奏する。
- [0052] 次に、S 5 0 6 において、物体の同定ができたかが判断される。もし、同定されたと判断された場合には、S 5 1 0 に進み、同定に使用したデータがMDBのデータよりも詳細であるか、最新であるかが判断され、これらの判断に基づいて、物体固有の情報（設計データなど）や時間情報（物体の型式、バージョン情報）がMDBに更新登録され、特定物体認識処理を抜ける。つまり、情報の登録及びMDB更新が、データベース学習処理となる。
- [0053] 一方で、S 5 0 6 において、物体の同定ができなかったと判断されると、画像情報以外の情報（画像中の文字やロゴ）を抽出して物体同定処理を行う。例えば、画像中の物体に非常に有名なメーカのロゴが映し出されてさえいれば、物体の大半がフレームから外れていたとしても同定が容易に行える場合がある。そうして、S 5 0 9 に進み、再び物体の同定ができたかどうか判断される。物体の同定ができていれば（S 5 0 9 においてY e s）、S 5 1 0 に進み、同定に使用したデータがMDBのデータよりも詳細であるか、最新であるかが判断され、これらの判断に基づいて、物体固有の情報（設計データなど）や時間情報（物体の型式、バージョン情報）がMDBに更新登録され、特定物体認識処理を抜ける。
- [0054] 一方で、物体の同定ができなかったならば（S 5 0 9 においてN o）、不特定物体としてMDBに登録され、将来の更新や新カテゴリの生成に備え（S 5 1 1）、特定物体認識処理を終了する（S 5 1 2）。
- [0055] なお、S 5 0 7 で示した、画像情報以外の情報による同定処理とともに、あるいはこれに替えて、集合知を利用して物体の同定に役立てることもできる（S 5 0 8）。このS 5 0 8の処理は、例えば、ネット上の百科事典を検索したり、Q & A 掲示板に自動投稿したりすることで実施される。ネット上の百科事典をシステムが自ら検索するには、一般物体認識で得られたカテゴ

りとともにM D Bより生成された特徴量を用いて検索クエリを作成し検索を実行する。そして返送されてきた内容から新たな特徴量を抽出して物体の同定が可能か再度試みる。また、Q & A 掲示板に自動投稿する場合は、一般物体認識で得られたカテゴリとともに元画像を掲示板にアップロードする。このとき、あらかじめ用意された定型文を自動編集して、「この〇〇について、型式を教えてください」とか、「この△△△について、設計情報を公開されているウェブサイトを教えてください」といったクエリを投稿する。そうして、他のユーザ（人間を含む）から「それは、x x - x x x xです」とか、「その△△△の設計データは、<http://www.aaabbb.com/cad/data.dxf>から入手可能です」といったアドバイスが寄せられる。システムは、これらのアドバイスを分析及び評価し、指定されたU R Lへアクセスし物体の設計データ等のダウンロードを試みる。新たに得られた設計データに基づいて物体の同定に成功すれば、得られた新たなデータがM D Bに追加され、データベースが更新される。

[0056] [画像認識システム]

ここで、図6 Aに、本発明に係るシステムの他の実施形態における画像認識システムの機能ブロックを示す。図6 Aに示す画像認識システム2 0 2は、サーバ1 0 1の一部として運用することも可能であるし、あるいは、サーバ1 0 1とは独立したサーバシステムとしても運用可能である。画像認識システム2 0 2は、サーバ1 0 1における一般物体認識部や特定物体認識部に対応する一般物体認識システム及び特定物体認識システムに加え、シーンを認識するためのシーン認識システムをも備えている。サーバ1 0 1における画像認識機能部の別形態あるいは応用例として、以下に詳述する。

[0057] 画像認識システム2 0 2は、ネットワーク通信制御部2 0 4と、領域処理部2 0 1と、データ検索処理部2 0 5と、一般物体認識システム1 0 6と、シーン認識システム1 0 8と、特定物体認識システム1 1 0と、画像カテゴリデータベース1 0 7と、シーン構成要素データベース1 0 9と、M D B 1 1で構成される。一般物体認識システム1 0 6は、一般物体認識部1 0 6

ー01と、カテゴリ認識部106-02と、カテゴリ学習部106-03と、新規カテゴリ登録部106-04とで構成され、シーン認識システム108は、領域抽出部108-01と、特徴抽出部108-02と、重み学習部108-03と、シーン認識部108-04とで構成され、特定物体認識システム110は、特定物体認識部110-01と、MDB検索部110-02と、MDB学習部110-03と、新規MDB登録部110-04とで構成され、画像カテゴリデータベース107は、カテゴリ分類データベース107-01と、不特定カテゴリデータ107-02で構成され、シーン構成要素データベース109は、シーン要素データベース109-01と、メタデータ辞書109-02とで構成され、MDB111は、詳細設計データ111-01と、付帯情報データ111-02と、特徴量データ111-03と、不特定物体データ111-04とで構成される。画像認識システム202の機能ブロックは必ずしもこれらに限定されるものではないが、これら代表的な機能について簡単に説明する。

[0058] 一般物体認識システム106は、画像中に含まれる物体を一般的な名称、あるいはカテゴリで認識する。ここでいうカテゴリは階層的であり、同じ一般物体として認識されているものでも、さらに細分化されたカテゴリ（同じ椅子でも4本足の「椅子」もあれば、全く足の無い「座椅子」のようなものまで含まれる）や、さらに大枠的なカテゴリ（椅子も机もタンスも含めて「家具」のカテゴリとして大分類される）としても分類及び認識が可能である。カテゴリ認識は、この分類を意味するClassification、即ち既知のクラスに物体を分類するという命題であり、カテゴリはまたクラスとも呼ばれる。

[0059] 一般物体認識プロセスにおいて、入力画像中の物体と参照物体画像との比較照合を行った結果、それらが同一形状であるか類似形状である場合、あるいは極めて類似した特徴を併せ持ち他のカテゴリとも主要な特徴において明らかに類似度が低いと認められる場合に、認識された物体に対し対応する既知のカテゴリ（クラス）を意味する一般名称を付与する。それらの各カテゴリを特徴付ける必須要素を詳細に記述したデータベースがカテゴリ分類デー

データベース 107-01 であり、それらのいずれにも分類する事ができない物体は、不特定カテゴリデータ 107-02 として一旦分類し、将来の新たなカテゴリ登録、あるいは既存カテゴリの定義範囲の拡大に備える。

[0060] 一般物体認識部 106-01 では、入力された画像中の物体の特徴点から局所特徴量を抽出し、それらの局所特徴量が予め学習によって得られた所定の特徴量の記述と似ているか似ていないかを比較して、当該物体が既知の一般物体であるかどうかを判別するプロセスを実行する。

[0061] カテゴリ認識部 106-02 では、一般物体認識可能となった物体がどのカテゴリ（クラス）に属するかを、カテゴリ分類データベース 107-01 との照合において特定あるいは推定し、その結果、特定カテゴリにおいてデータベースに追加あるいは修正を加えるような追加の特徴量が見出された場合には、カテゴリ学習部 106-03 において再学習した上で、カテゴリ分類データベース 107-01 の当該一般物体に関する記述をアップデートする。また一旦、不特定カテゴリデータとされた物体とその特徴量が別に検出された他の不特定物体の特徴量と極めて類似であると判定された場合には、それらは新たに発見された同一の未知のカテゴリ物体である可能性が高いとして、新規カテゴリ登録部 106-04 において、カテゴリ分類データベース 107-01 にそれらの特徴量が新規登録され新たな一般名称が付与される。

[0062] シーン認識システム 108 では、入力画像全体あるいは一部を支配している特徴的な構成要素を、性質の異なる複数の特徴抽出システムを用いて検出し、それらをシーン構成要素データベース 109 に記載されているシーン要素データベース 109-01 と多次元空間上で参照する事で、それぞれの入力要素群が特定シーン内に検出されるパターンを統計処理により求め、画像全体あるいは一部を支配している領域が特定のシーンであるかどうかを認識する。併せて、入力画像に付帯しているメタデータ群と、シーン構成要素データベース 109 に予め登録済みの、メタデータ辞書 109-02 に記載されている構成要素とを照合し、シーン検出の精度を一段と向上させる事も可

能となる。領域抽出部 108-01 では、画像全体を必要に応じて複数の領域に分割して、領域毎にシーン判別を可能にする。例えば、都市空間内のビルの屋上に設置した高解像度の監視カメラからは、交差点や数多くの店舗のエントランス等の複数のシーンを見渡す事ができる。特徴抽出部 108-02 は、指定した画像領域内における検出された複数の特徴点の局所特徴量、色情報や物体の形状等、利用可能な様々な特徴量から得られる認識結果を後段の重み学習部 108-03 に入力し、それぞれの要素が特定シーンにおいて共起する確率を求め、シーン認識部 108-04 に入力して最終的な入力画像に対するシーン判別を行う。

[0063] 特定物体認識システム 110 は、入力された画像から検出された物体の特徴を、予め M D B 内に収納されている特定物体群の特徴と逐次照合し、最終的に物体を同定 (Identification) 処理する。地球上に存在する特定物体の総数は膨大で、それら全ての特定物体との照合を行う事はおよそ現実的ではない。従って、後述するように特定物体認識システムの前段において、予め一定の範囲内に物体のカテゴリや探索範囲を絞り込んでおく必要がある。特定物体認識部 110-01 では、検出された特徴点における局所特徴量と、学習によって得られた M D B 内の特徴パラメータ群とを比較し、当該物体がどの特定物体に当て嵌まるかの判別を統計処理により判別する。M D B には、その時点で入手可能な特定物体に関する詳細なデータが記載されている。一例としてそれらが工業製品であるのならば、設計データ 111-01 として設計図や C A D データ等から抽出された、物体の構造、形状、寸法、配置図、可動部、可動範囲、重量、剛性、仕上げなど、物体を構成し製造するのに必要な基本情報が M D B 内に保持される。付帯情報データ 111-02 には、物体の名称、製造者、部品番号、日時、素材、組成、加工情報など物体に関するあらゆる情報が保持される。特徴量データ 111-03 には、設計情報に基づいて生成される個々の物体の特徴点や特徴量情報が保持される。不特定物体データ 111-04 は、その時点ではどの特定物体にも属していない物体等のデータとして、将来の解析に備え M D B 内に暫定的に収納され

る。MDM検索部110-02は特定物体に対応する詳細データを検索する機能を提供し、MDB学習部110-03は適応的かつ動的な学習プロセスを通して、MDBの記載内容に対し追加・修正を行う。また一旦、不特定物体として不特定物体データ111-04とされた物体も、その後に類似の特徴を有する物体が頻繁に検出されれば、新規MDB登録部110-04により新たな特定物体として新規登録処理される。

[0064] 図6Bに、一般物体認識部106-01のシステム構成例及び機能ブロックの実施例を示す。一般物体認識部106-01の機能ブロックは必ずしもこれらに限定されるものではないが、代表的な特徴抽出手法としてBag-of-Features（以下、BoFと呼称する）を適用した場合の一般物体認識手法について、以下に簡単に説明する。一般物体認識部106-01は、学習部106-01aと、Visual Word辞書（Code Book）106-01eと、ベクトル量子化部106-01fと、ベクトル量子化ヒストグラム部106-01gと、ベクトル量子化ヒストグラム識別部106-01hで構成され、学習部106-01aは、局所特徴量抽出部106-01bと、クラスタリング部106-01cと、Visual Word作成部106-01dで構成され、ベクトル量子化ヒストグラム識別部106-01hは、Support Vector Machine（以下、SVMと呼称する）部106-01iで構成される。

[0065] BoFは、画像中に現れる特徴点を様々な方法で抽出し、その相対位置関係を用いずに多数の局所特徴量（Visual Word）の集合体として表現し、それらを学習によって得られた、様々な物体から抽出されたVisual Word辞書（Code Book）106-01eと比較照合して、それら局所特徴量の出現頻度がどの物体に最も近いかを判別する代表的な物体認識手法として広く知られている。

[0066] 図7（A）に、局所特徴量抽出の代表的な事例として、Scale Invariant Feature Transform（以下、SIFTと呼称する）を用いた事例を示す。SIFTは、画像の大きさ・回転・照明の変化に頑健な特徴点検出及び特徴量抽出アルゴリズムの一つであり、一枚の画像の中から複数の特徴的な輝度勾配

の分布を原画像に対する異なる平滑化画像の差分（一例としてDifference-of-Gaussian、以下D o Gと呼称する）を用いて検出し、その代表点である極値（重心位置）を求め特徴点（キーポイント）として抽出する手法である。D o Gを求める過程において、得られるガウス窓（Gaussian Window）の開口量からそれぞれの各特徴点におけるスケールを求め、その支配範囲における局所特徴量を計算する。キーポイント検出プロセスにおいて、画像中に頻出するエッジ上では開口が極端に小さく、有用な特徴量とはなりにくいことからキーポイントから除外される。同様にD o G出力の小さい点（コントラスト比の小さい領域）も、原画像に含まれるノイズの影響を受ける可能性が高いことからキーポイントから除外される。図7（A）では、それらのプロセスを用いて検出された複数のキーポイントと、そのスケールを白い円で示す。

[0067] 次に、検出されたそれぞれの特徴点に対して、代表するオリエンテーション（主成分の方向）を求める。オリエンテーションは10度刻みに全36方向について輝度勾配強度を求め、それが最大値となるオリエンテーションを、そのキーポイントを代表するオリエンテーションとして採用する。次に、各キーポイント周りの検出されたスケール領域において、主要な輝度勾配の代表点を求め、それを各キーポイントの主たるオリエンテーションとする。その上で、各キーポイントのスケールに基づく周辺領域全体を、上記で求めたオリエンテーションに従い回転させながら4×4の計16領域に分割し、各ブロック内で45度ずつ8方向の勾配方向ヒストグラムを生成して、それらの結果から16ブロック×8方向＝計128次元の特徴ベクトルを求める。これらの操作により、画像の回転及びスケールに頑健なSIFT特徴量を得る事が可能になる。最終的に、それら128次元の特徴ベクトルの大きさを正規化して、照明変化にも頑健な局所特徴量を得る。

[0068] 図6Bにおいて、学習部106-01aを構成する局所特徴量抽出部106-01bによって得られた128次元の特徴ベクトルは、後段のクラスタリング部106-01cによって多次元の特徴ベクトル群にクラスタ分割され、Visual Word作成部106-01dで、それぞれの重心ベクトルを元に特

徴ベクトル毎にVisual Wordが生成される。クラスタリングの手法として、k-means法やmean-shift法が知られている。生成されたVisual Wordは、Visual Word辞書 (C o d e B o o k) 1 0 6 - 0 1 e に収納され、それを元に入力画像から抽出されたVisual Wordを照合し、ベクトル量子化部 1 0 6 - 0 1 f で特徴毎にベクトル量子化を行う。その後、ベクトル量子化ヒストグラム部 1 0 6 - 0 1 g において、各次元に対するヒストグラムを生成する。

[0069] 図 7 (B) に、生成されたVisual Word (C o d e B o o k) 、図 7 (C) に抽出されたベクトル量子化ヒストグラム例を示す。ヒストグラムの各ビンの総数 (次元数) は数千〜数万と多く、入力画像によっては特徴の一致が全くないヒストグラムのビンも数多く存在する一方、特徴の一致が顕著なビンもあり、それらを一括してヒストグラムの全ビンの値の総和が 1 になるように正規化処理が行われる。得られたベクトル量子化ヒストグラムは、後段のベクトル量子化ヒストグラム識別部 1 0 6 - 0 1 h へと入力され、代表的な識別器であるSupport Vector Machine (以下SVMと呼称する) 1 0 6 - 0 1 i において、物体の属するクラス (一般物体) を認識処理する。ここでの認識結果は、Visual Word辞書に対する学習プロセスとしても利用可能である。また、他の手法 (メタデータや集合知) も含めた認識判断も、同様にVisual Word辞書に対する学習フィードバックとして利用可能で、同一クラスの特徴を最も適切に記述し他のクラスとの分離度を良好に保つように、アダプティブな修正・校正を継続する事が可能となる。

[0070] 図 8 に、前記の一般物体認識部 1 0 6 - 0 1 を含む、一般物体認識システム 1 0 6 全体の概略ブロック図を示す。一般物体 (クラス) は様々なカテゴリに属していて、それらは多重的な階層構造を成している。一例を挙げると、人間は「哺乳類」という上位カテゴリに属し、哺乳類は「動物」というさらに上位のカテゴリに属しているといった具合である。人間はまた、髪の色や目の色、大人か子供か?といった別のカテゴリでも認識が可能である。これらの認識判断を行うには、カテゴリ分類データベース 1 0 7 - 0 1 の存在が欠かせない。これは現在までの人類の「知」の集積庫であり、今後の学習

や発見によって、さらに新たな「知」が補充され継続的に進化して行くものでもある。一般物体認識部 106-01 で同定されたクラス（究極的には、人類がこれまでに識別している、全ての名詞の総数）は、様々な多次元的かつ階層的な構造を以ってこのカテゴリ分類データベース 107-01 内に記述されている。学習過程において、認識された一般物体はカテゴリ分類データベース 107-01 と照合され、カテゴリ検出部 106-02 で所属カテゴリが認識される。その後、カテゴリ学習部 106-03 に認識結果が引き渡され、カテゴリ分類データベース 107-01 内の記述との整合性が詳細にチェックされる。通常一般物体認識された物体は、複数の認識結果を内包する場合が多い。例えば「昆虫」であると認識した場合に、目の構造や手足の数、触角の有無、全体の骨格構造や羽の大きさ、胴体の色彩や表面のテクスチャ等でも認識・分類可能で、カテゴリ分類データベース 107-01 内の当該一般物体詳細記述と共に照合される。カテゴリ学習部 106-03 では、これらの照合結果を基に、カテゴリ分類データベース 107-01 への追加・修正が必要に応じて適応的に行われる。その結果、既存カテゴリのいずれにも分類できず、「新種の昆虫」ではないか？と推測される場合には、新規カテゴリ登録部 106-04 が、これらの情報をカテゴリ分類データベース 107-01 内に登録する。一方、その時点では不明な物体は、不特定カテゴリデータ 107-02 として、将来の解析や照合に備え一時的にカテゴリ分類データベース 107-01 内に収納される。

[0071] 図 9 に、入力画像に含まれるシーンを認識判別するシーン認識システム 108 の本発明における代表的な実施例をブロック図で示す。学習画像及び入力画像からは、一般に複数の物体が認識可能となるケースが多い。例えば、「空」「太陽」「地面」などの領域と同時に「木」や「草」そして「動物」等の物体が認識可能となる場合、それらが「動物園」なのか「アフリカ」なのかは全体の景色や、それ以外に発見される物体との共起関係等から類推する事になる。例えば、柵や掲示板等が同時に発見され多くの見物客で賑わっていれば「動物園」である可能性が高まるが、全体のスケールが大きく、遠

くに「キリマンジャロ」のような大きな山を臨み、様々な動物が混在して草原上にいるような場合には、そこが「アフリカ」である可能性が一気に高まる。このような場合、さらに認識可能な物体や状況、共起関係等を知識データベースであるシーン構成要素データベース 109 に照合し、より総合的に判断を下す必要も出てくる。例えば、全画面の 9 割が「アフリカ」を指し示していると推定されても、それらが四角枠内に切り取られ壁に張り付いていれば、実際はポスターや写真である確率が極めて高くなる。

[0072] シーン認識システム 108 は、領域抽出部 108-01 と、特徴抽出部 108-02 と、強識別部 108-03 と、シーン認識部 108-04 と、シーン構成要素データベース 109 から構成され、特徴抽出部 108-02 は、一般物体認識部 108-05 と、色情報抽出部 108-06 と、物体形状抽出部と、コンテキスト抽出部と、弱識別器 108-09~12 とで構成され、シーン認識部 108-04 は、シーン分類部 108-13 と、シーン学習部 108-14 と、新規シーン登録部 108-15 で構成され、シーン構成要素データベース 109 は、シーン要素データベース 109-01 と、メタデータ 109-02 で構成される。

[0073] 領域抽出部 108-01 は、背景や他の物体の影響を受けずに目的とする物体の特徴を効果的に抽出するために、対象画像の領域抽出を行う。領域抽出手法の例として、グラフベースの領域分割法 (Graph-Based Image Segmentation) 等が知られている。抽出された物体画像は、局所特徴量抽出部 108-05、色情報抽出部 108-06、物体形状抽出部 108-07、コンテキスト抽出部 108-08 にそれぞれ入力され、それらの各抽出部から得られた特徴量が弱識別器 108-09~12 において識別処理され、多次元の特徴量群として統合モデリングされる。それらモデリング化された特徴量群を、重み付け学習機能を有する強識別器 108-03 に入力し、最終的な物体画像に対する認識判定結果を得る。前記の弱識別器の例として SVM、強識別器の例として AdaBoost 等が上げられる。

[0074] 一般に入力画像には複数の物体や、それらの上位概念である複数のカテゴ

りが含まれている場合が多く、人間はそこから一目で特定のシーンや状況（コンテキスト）を思い浮かべる事ができる。一方、単独の物体や単一のカテゴリのみを提示された場合、それだけで入力画像がどういうシーンを表わしているのかを、直接的に判断するのは困難である。通常は、それらの物体が存在している状況や相互の位置関係、またそれぞれの物体やカテゴリが同時に出現する確率（共起関係）が、その後のシーン判別に対して重要な意味を持ってくる。前項で画像認識可能となった物体群やカテゴリ群は、シーン要素データベース 109-01 内に記述されている、シーン毎の各要素群の頻出確率を基に照合処理され、後段のシーン認識部 108-04 において、入力画像がいかなるシーンを表現しているのかが統計的手法を用いて決定される。これとは別の判断材料として、画像に付帯しているメタデータ 109-02 も有用な情報となり得る。しかし、時には人間が付したメタデータ自体が、思い込みや明らかな誤り、あるいは比喩として画像を間接的に捉えている場合等もあり、必ずしも入力画像中の物体やカテゴリを正しく表わしているとは限らない場合がある。このような場合にも、画像認識システムで得られる結果や、知識情報システムで共起関係等を基に得られる結果を勘案して、最終的に物体やカテゴリの認識処理が行われる事が望ましい。また、一つの画像からは複数のシーンが得られる場合も多い（「海」であると同時に「海水浴場」であったりもする）。その場合は、複数のシーン名が共に付される。さらに画像に付すべきシーン名として、例えば「海」あるいは「海水浴場」のいずれがより相応しいかは、当該画像のみからでは判断が難しく、前後の文脈（コンテキスト）や全体との相関関係やそれぞれの出現共起関係を基に、知識データベースの助けを借りて最終的に判断する必要も出てくる。

[0075] 図 10 (A) にシーン要素データベース 109-01 の記述例を示す。シーン (A) は複数のカテゴリ、カテゴリ m とカテゴリ n を含み、カテゴリ m の構成要素として一般物体  $\alpha$  と一般物体  $\beta$  が、カテゴリ n の構成要素として一般物体  $\gamma$  と特定物体  $\delta$  と特定物体  $\varepsilon$  が、それぞれの出現確率と共に記述されている。

[0076] 図10(B)にシーン『交差点』の構成要素例を示す。『交差点』を構成するカテゴリとして『道路』を考えた場合、そこには複数車線からなる「幹線道路」や片側一車線の「一般道路」、あるいは「歩道」といったような様々な道路が存在する。『道路』上に「車線分離表示」や「横断歩道表示」「進行方向指示」のような『路面表示』が同時に見つかれば、かなりの確率でそこは『交差点』か『交差点』に近い場所であろうと推測できる。その上で複数の道路が交差し、『信号機』がそれぞれにあり、そこに「進行表示」や「歩行者専用表示」等も見つかれば、間違いなくそこは『交差点』であると断定できるだろう。その上で、信号機に「交差点名表示」があり文字が認識可能な場合には、『交差点』はその場所の特定までされた事になる。道路上には他にも「速度規制表示」や「駐停車規制表示」あるいは「一時停止表示」のような『道路標識』カテゴリがあるだろうし、歩道や横断歩道上に『人』がいれば、それぞれの「体型」や「服装」のみならず、「顔」も特定できる可能性もある。車道には「乗用車」や「大型車」のような様々な形態の『車』があり、それぞれの「ナンバープレート」も認識可能かも知れない。また歩道や車道には「自転車」や「オートバイ」等の『二輪車』が走っている場合が多く、「ナンバープレート」が付されていれば「自動二輪」や「オートバイ」として認識する事も可能となるだろう。歩道には「犬」や「猫」などの『動物』が発見される場合もある。紐で「人」と繋がっている場合は「犬」の確率が高まる。単独でうろうろしている小型動物の場合は「猫」かもしれないし、空を飛んでいれば「鳥」であろうと推察できる。視界内には「店の看板」や「電柱の看板」あるいは車体やガラス窓に張ってある「車の看板」のようなカテゴリ『看板』が見つかる可能性が高い。『交差点』の周囲には「建物」があり、時には「歩道橋」や「橋」等の『建造物』を見る事ができるだろうし、道路上には「電柱」や「街路灯」「街路樹」のような一般物体だけではなく、「道路工事」中の臨時のガードレールや看板のような『路上構築物』も見つかるかも知れない。そういった中で、例えば物体が「不明瞭」であるとかオクルージョン（隠蔽）等により「不完全」であるよう

な場合には、『不明』カテゴリとして将来の認識処理に委ねられる事になる。

[0077] 図11に、特定物体認識システム110のシステム全体の構成例、及び機能ブロックを示す。特定物体認識システム110は、一般物体認識システム106と、シーン認識システム108と、MDB111と、特定物体認識部110-01と、MDB検索部110-02と、MDB学習部110-03と、新規MDB登録部110-04とで構成され、特定物体認識部110-01は、二次元写像部110-05と、個別画像切り出し部110-06と、局所特微量抽出部110-07と、クラスタリング部110-08と、Visual Word作成部110-09と、Visual Word辞書(Code Book)110-10と、ベクトル量子化110-11と、ベクトル量子化ヒストグラム部110-12と、ベクトル量子化ヒストグラム識別部110-13と、SVM部110-14と、形状特微量抽出部110-15と、形状比較部110-16と、色情報抽出部110-17と、色彩比較部110-18とで構成される。

[0078] 一般物体認識システム106により、対象物体の属するクラス(カテゴリ)が認識可能になった時点で、物体がさらに特定の物体としても認識可能か？という絞り込みのプロセスに移る事が出来る。クラスがある程度特定されないと、無数の特定物体群からの検索を余儀なくされ、時間的にもコスト的にもとても実用的とは言えない。これらの絞り込みプロセスには、一般物体認識システム106によるクラスの絞り込み以外にも、シーン認識システム108の認識結果から対象の絞り込みを進める事も可能となる。また特定物体認識システムから得られる有用な特微量を用いて、さらなる絞り込みが可能になるだけでなく、物体の一部にユニークな識別情報(商品名とか、特定の商標やロゴ)等が認識可能な場合、あるいは有用なメタデータ等が付されているケースでは、さらなるピンポイントの絞り込みも可能となる。

[0079] それら絞り込まれたいくつかの可能性の中から、複数の物体候補の詳細データや設計データをMDB検索部110-02がMDB111内から順次引

き出し、それらを基に入力画像とのマッチングプロセスに移る。物体が人工物体でない場合や、詳細な設計データ自体が存在していない場合においても、写真等があればそれぞれの特徴を詳細に突き合わせる事で、或る程度の特定物体認識も可能となる。しかし、入力画像と比較画像の見え方がほとんど同じというケースは稀で、それぞれを違う物体として認識してしまうケースがある。反面、物体が人工物体でありCAD等の詳細なデータベースも存在している場合には、二次元写像部110-05が入力画像の見え方に応じMDB内の三次元データを可視化（レンダリング）する事により、極めて高精度の特徴量のマッチングを行う事が可能になる。この場合、二次元写像部110-05における全方向の詳細レンダリングを行う事は計算時間とコストの不要な増大を招く事から、入力画像の見え方に応じた絞り込みが必要となる。一方、MDBを用いた高精度のレンダリング画像から得られる物体の様々な特徴量群は、予め学習プロセスにおいて十分時間をかけて求めておく事も可能となり、実用システムを構築する上でより効果的となる。

[0080] 特定物体認識部110-01では、物体の局所特徴量を局所特徴量抽出部110-07で検出し、クラスタリング部110-08でそれぞれの特徴量を複数の類似特徴群に分離した後、Visual Word作成部110-09で多次元の特徴量セットに変換して、それらをVisual Word辞書110-10に登録する。これらは、多数の学習画像に対し十分な認識精度が得られるまで継続して行われる。学習画像が写真である場合は、画像の解像度不足やノイズの影響、オクルージョンの影響、対象物体画像以外の物体から受ける影響等が避けられないが、MDBを基にしている場合は、対象画像の抜き出しも理想的に行う事が可能で、従来手法に比べて大幅に分解能を高めた特定物体認識システムを構成する事が可能となる。入力画像は、個別画像切り出し部110-06で目的とする特定物体の凡その領域が切り出された後に、局所特徴量抽出部110-07で特徴点及び特徴量を求め、予め学習により用意されたVisual Word辞書110-10を用い、個々の特徴量毎にベクトル量子化された後にベクトル量子化ヒストグラム部110-12にて多次元の特徴量に

展開され、ベクトル量子化ヒストグラム識別部 110-13 にて物体が参照物体と同一か判断される。識別器の例として、SVM (Support Vector Machine) 110-14 が知られているが、他にも判断の重みを学習の上で可能にするAdaBoost等も有効な識別器として利用するケースが多い。これらの識別結果は、MDB学習部 110-03 を通じて、MDB自体への修正と項目追加というフィードバックループにも利用可能で、依然として未確認物体である場合には新規MDB登録部 110-04 に保留され、将来の類似物体の出現や、新たなMDBの登録に備える。

[0081] また、局所特徴量のみならず、検出精度をさらに向上させる目的で物体の形状特徴を利用する事も可能となる。入力画像から切り出された物体は、形状特徴量抽出部 110-15 を経由して形状比較部 110-16 に入力され、物体の形状的な特徴を用いた識別が行われる。その結果はMDB検索部 110-02 にフィードバックされ、可能性のある特定物体に対応したMDBへの絞り込みが行われる。形状特徴量抽出手段の例として、HOG (Histograms of Oriented Gradients) が知られている。形状特徴量は、またMDBを用いた二次元写像を得るための不要なレンダリング処理を減らす目的でも有用となる。

[0082] また、物体の色彩的な特徴や表面処理（テクスチャ）も、画像認識精度を上げる目的で有用である。切り出された入力画像は、色情報抽出部 110-17 に入力され、色彩比較部 110-18 で物体の色情報、あるいはテクスチャ等の抽出が行われ、その結果をMDB検索部 110-02 にフィードバックする事で、比較対象となるMDBのさらなる絞り込みを行う事が可能となる。これら、一連のプロセスを通じて、特定物体認識が効果的に行われる。

#### [0083] B. インタレスト・グラフ収集処理

次に、図12～図36に基づいて、本発明に係るシステムの一実施形態におけるインタレスト・グラフ収集処理について説明する。

[0084] 図12A乃至図12Eは、本発明に係るシステムの一実施形態におけるユ

ーザインタフェースを説明している。図 12A (A) において、ネットワーク端末装置 105 上のディスプレイには、画像 1201、1202 のほか、何枚かの画像と、関連性検索窓 1203 と、出力窓 (OUTPUT) 1205 とが表示されている。

[0085] 画像 1201、1202 は、上記画像認識エンジン 200 により認識可能となった画像構成要素群、及びそれら画像構成要素群それぞれに対応する関連要素群に対し、それらを代表する写真、イラストレーション、文字、記号、ロゴ、ファビコン等から生成した縮小画像サムネイルを関連性検索エンジンが元画像に代えてネットワーク端末に向け送出した画像タイルであり、ユーザの指 (1206) などの操作により、画面上の任意の位置にドラッグ可能となっている。

[0086] また、関連性検索窓 1203 は、ネットワーク端末装置 105 に係るホーム画面、或いはその上で動作する特定のアプリケーションが管理する画面など、任意の画面上に配置させることができる。典型的な実施形態としては、ネットワーク端末装置 105 の起動後のホーム画面上に常駐させて、いつでもユーザが関連性検索窓 1203 に対して検索対象となる画像全体、或いは画像の特定領域を選択した後にドラッグ&ドロップして画像認識及びそれに引き続く関連性検索プロセスを開始する事が出来るように構成する事が可能となる。一例として、図 12A (A) においては、ユーザは画像 1201 を関連性検索窓 1203 にドラッグ&ドロップする操作を示している。

[0087] なお、関連性検索窓 1203 を特に用意せず、ユーザが関心を持った画像全体、或いは画像の特定の領域をユーザがネットワーク端末 105 上で選択し、それら選択した画像をネットワーク経由でサーバ側の画像認識エンジン 200 に問い合わせることが出来るように構成されたインタフェースであれば、どのようなインタフェースを採用してもよい。

例えば、関連性検索窓への検索対象画像の投入という操作に代えて、検索の対象となる画像全体、或いは特定の画像領域を明示的にネットワーク端末 105 の表示スクリーン上でダブルタップする等の操作により、サーバ側の

画像認識エンジン 200 に当該選択画像の認識処理を問い合わせる事も可能である。

[0088] また PC 等では、図 12A (B) のように、タッチパネルでの入力操作に代わりマウス等のポインティングデバイス 1204 を用い、カーソル 1207 を画像 1201 上に移動させ、対象画像 1201 を関連性検索窓 1203 a (あるいは、関連性検索に結び付けられたアイコン 1203 b) に直接ドラッグ&ドロップしたり、マウスのカーソルを画像 1201 上に合わせダブルクリックしたりする事により、サーバ側の画像認識エンジン 200 に当該選択画像の認識処理を問い合わせる事も可能である。

[0089] 図 12B (A) は、上記選択画像 1201 に対する関連性検索結果としてサーバ 101 からネットワーク端末 105 に向けて送出される画像構成要素群、及びそれらとさらに関連性の高いその他の関連要素群とをノードとした関連性グラフをネットワーク端末 105 のスクリーン全体に表示し、それらの各ノードをユーザがタッチスクリーン上でフリック (1210) することで、関連性グラフ上の各ノードを左から右にシームレスに辿っている様子を示す。また、ノードとなっている画像を選択的にユーザがタップ或いはタッチする事により、当該ノードを中心とした関連性グラフの表示となるようにネットワーク端末 105 側で全体の表示を自動的にスクロールさせる事も可能である。図 12B (A) の領域 1209 は、関連性グラフの一例を示し、その一部が切り取られてネットワーク端末 105 上に描かれている様子を図示している。実際の関連性グラフの大きさは当例示よりも遥かに大きくなるケースが多く、ネットワーク端末 105 上に表示しきれない領域 1209 に属するノード、及びそれら相互の関連性であるリンク情報を関連性検索エンジン 220 がネットワーク端末 105 に向けユーザのスクロール操作に伴い追加的に送出することで、ユーザにとって関心のあるノード或いは領域を関連性グラフ上でシームレスに辿りながら、複数のノード間にまたがる広範な関連性としてユーザに視覚的に提示する事が可能となる。

図 12B (A) では、フリックした結果、グレープジュース 1220 の関

連要素としてのオレンジジュース 1 2 2 1 及びグレープ 1 2 2 2 が表示され、さらに、グレープ 1 2 2 の関連要素としてのフルーツ群 1 2 2 3 ~ 1 2 2 6 が表示されている。

[0090] また、図 1 2 B (B) においては、図 1 2 B (A) において、グレープジュース 1 2 2 0 を明示的に選択 (2 回以上のタップやタッチ等) し、これをネットワーク経由してサーバ側の画像認識エンジン 2 0 0 に問い合わせる事により、画像認識エンジン 2 0 0 において認識された画像構成要素群であるボトルキャップ 1 2 3 1、ボトル 1 2 3 2、及び製造元会社のロゴ 1 2 3 3 が表示されている。

[0091] なお、上記スクロール操作はユーザによるジェスチャーや視線、音声、或いは脳波等を用いた同様の効果を及ぼす入力操作を用いても良い (図には描かれていないが、ピンチイン／ピンチアウトを含むジェスチャーの検知や、視線或いは脳波の検知等は、既に用いられている多くのセンシング技術を導入することができる)。また、二次元的なスクロール操作に限らず、関連性グラフを三次元空間やさらに多次元の空間内に配置する事も可能となる。

[0092] 図 1 2 C に事例として示すように、関連性グラフの表現を、複数のノード群とそれら相互の関連性と関連性の強さを視覚的に表わした幾何学的なグラフとしてだけではなく、限られた画像表示サイズとならざるを得ない携帯端末等に向けて、タイル状に並べた等しいサイズの画像タイルの集合として表現する方法も有用となる。この場合は、(1) 最初の入力画像 (1 5 0 1)、(2) 画像認識エンジンにより検出・認識された複数の画像構成要素候補群 (1 2 5 1)、(3) それら個々の画像構成群それぞれと関連する他の関連要素群 (1 2 5 2 或いは 1 2 5 3) を、それぞれ異なる要素群としてネットワーク端末 1 0 5 の表示画面上の割り当てられたエリアに並べて表示する方法が有用となる。(3) (1 2 5 2 或いは 1 2 5 3) の関連要素群は、それぞれ一次繋がり、二次繋がり、三次繋がり…という関連性の度合いに応じたレイヤ毎の表示とする事で、ユーザによるタッチパネル上のフリック操作で、高速に画面をスクロール (1 2 5 4) 可能にして全体の関連性を効果的に閲覧可能さ

せる事が出来る。この様なタイル状のノードをマトリックス状に配列した場合、それぞれのノード間の関連性の強さは、各ノード近傍に数値や記号等のデータとして附記する事が可能である。

[0093] これらタイル状に表現された関連性グラフの任意のノード画像を、再びネットワークを經由してサーバ側の画像認識エンジン200に問い合わせ、入力された画像から新たな画像構成要素群を得る操作により、それらを起点とする新たな関連性グラフを関連性検索エンジン220経由で取得する事が可能となる。このユーザインタフェースの一実装例として、任意のノード画像をユーザが明示的にダブルタップする事により、当該画像に対する新たな画像構成要素の検出及び画像認識要請をネットワーク端末105からサーバ101に対して行ない、画像認識エンジン200が新たな画像構成要素群をサーバ側で検出・認識し、結果をネットワーク端末105に返す事で、ネットワーク端末105側の表示画面に当該画像認識要素群として再提示する事が可能となる(図12B(B))。これら新たな画像構成要素群は、従来の関連性グラフに代えて表示しても良いし、半透明表示機能等を用いて従来の関連性グラフ上に重畳して表示しても良い。これら画像構成要素候補群にユーザが関心を示さない場合は、元の関連性グラフに復帰させる。或いは、ユーザが新たに検出された画像構成要素に関心を示した場合は、当該画像構成要素を表わすノードをユーザがダブルタップする事により、当該画像構成要素を中心とする新たな関連性グラフに係る新たな関連ノード群がサーバ側からネットワーク端末105に向けて送出され、ユーザは新たな関連性グラフを獲得する事が可能になる。

[0094] インタレスト・グラフの収集には、画像認識エンジン200、関連性検索エンジン220、統計情報処理部209、及び特定ユーザフィルター処理部210が関係する。これらは、すべてサーバ101の一部として運用することも可能であるし、あるいは、それぞれサーバ101とは独立したサーバシステムとしても運用可能である。

[0095] 図13に、関連性検索エンジン220におけるグラフ演算部221の詳細

機能ブロックの一構成例を示す。グラフ演算部 2 2 1 は、部分グラフ生成部 1 3 0 1 と、多次元特徴ベクトル生成部 1 3 0 2 と、関連要素ノード抽出部 1 3 0 3 とからなり、必要に応じて、グラフデータベース 1 0 2 A 及び関連性知識データベース 1 3 1 0 からデータの授受を行う。

[0096] 部分グラフ構成部 1 3 0 1 は、画像認識エンジン 2 0 0 により抽出された画像構成要素に対応するノードを入力とし、GDB 1 0 2 A にアクセスしながら当該ノードの部分グラフを生成する。

[0097] 多次元特徴ベクトル生成部 1 3 0 2 は、部分グラフから関連性演算部 2 2 4 での計算（後述の図 1 8）によって多次元特徴ベクトルを生成する。

[0098] 関連要素ノード抽出部 1 3 0 3 は、得られた複数の多次元特徴ベクトルの距離を、例えば、ユークリッド距離を計ったり、マハラノビス距離を計ったりして求め、関連要素ノードを抽出する。

[0099] 図 1 4 A に、本発明に係るシステムの一実施形態におけるグラフを表現するための基本的データ構造を示す。

[0100] [基本的なグラフ演算]

図 1 4 A (A) に示す通り、生成時刻及び `value` (1 4 0 2) に `hash` 演算 1 4 0 4 を施すことで `key` (1 4 0 1) を得る。例えば、`hash` 演算 1 4 0 4 がハッシュアルゴリズム SHA-1 を用いた場合には、`key` は 1 6 0 ビット長になる。`key` (1 4 0 1) は `locate` 演算 1 4 0 3 によって `value` (1 4 0 2) を得る。例えば `locate` 演算 1 4 0 3 には分散ハッシュテーブル (Distributed Hash Table) 方が利用できる。本実施例では、この `key` と `value` の関係を ”(key, {value})” で表現し (図 1 4 A (B))、ノードデータ及びリンクデータとして GDB 1 0 2 A に格納する単位とする。

例えば、図 1 4 A (C) の 2 つのノードがリンクされている場合、ノード `n1` (1 4 1 0) は、”(`n1`, {ノード `n1`})” で、ノード `n2` (1 4 1 1) は、”(`n2`, {ノード `n2`})” で表現する。`n1` や `n2` はそれぞれノード `n1` (1 4 1 0)、ノード `n2` (1 4 1 1) の `key` であり、ノード実体ノード `n1` (1 4 1 0)、ノード

ード $n_2$  (1 4 1 1) をそれぞれ `hash` 演算しそれぞれの `key` を得る。また、リンク  $l_1$  (1 4 1 2) は、ノードと同様に “( $l_1$ , { $n_1$ ,  $n_2$ })” で表現し、{ $n_1$ ,  $n_2$ } を `hash` 演算することでその `key` ( $l_1$ ) 1 4 1 2 を得る。

[0101] 図 1 4 A (D) にノードが保持するデータの構造を示す。タイプ欄には、ノードの保持するデータのタイプが格納される。タイプには、“`USER`”、“`OBJECT`”、“`META`”、“`URI`”、“`EXT`” の 4 種類を定義している。“`USER`” はノードがユーザを表現していることを、“`OBJECT`” はノードがオブジェクトを表現していることを示している。“`META`” はノードがユーザ或いはオブジェクトのメタデータを、“`SUBJECT`” (Subjectivity) はノードがユーザの主観を、“`URI`” はノードがユーザ或いはオブジェクトへの `URI` を示している。“`EXT`” はタイプの拡張の為に用意されており、拡張されたデータはデータ欄に格納される。

[0102] 図 1 4 A (E) に、リンクが保持するデータの構造を示す。タイプ欄には、リンクのタイプが格納される。タイプには、“`UNDIRECTED`”、“`DIRECTED`”、の 2 種類が定義されている。“`UNDIRECTED`” は当該リンクが無向リンクであることを、“`DIRECTED`” は当該リンクが有向リンクであることを示している。また、データ欄には、左ノードの `key`、右ノードの `key`、重み ( $w$ ) 及び関数 ( $f$ ) が格納される。重みにはリンクの太さを表現する値を使っても良いし、後述する多次元特徴ベクトルを圧縮した値を使っても良い。

[0103] これらノードやリンクの “(`key`, {`value`})” で表現されるデータは `immutable` (データ不変) という性質を有しており、すなわち `write-once-read-many` のセマンティクス (書き込みは一回のみだが読み出しは複数回可能) を有するが、そのセマンティクスには限定されない。例えば、`write-many-read-many` (書き込み読み出し共に複数回可能) のセマンティクスでも良い。その場合には、ノード及びリンク共に修正時刻の欄が追加される。

[0104] 図 1 4 A に示したノードデータとリンクデータは `GDB 1 0 2 A` に格納される。図 1 4 B にそれらのデータを操作する為の `GDB 1 0 2 A` のオペレー

ションを示す。図中、典型的なオペレーションセットとして、“CREATE”、“CONNECT”、“NODE”、“LINK”、“SUBGRAPH”の5つのオペレーションを定義するが、拡張しても良い。例えば、データのセマンティクスがwrite-many-read-manyの場合には、“DESTROY”や“UPDATE”のオペレーションがあっても良い。

[0105] “CREATE”は指定したタイプのノードを生成する。“CONNECT”は、指定した2個のノードを指定したタイプでつなげるリンクを生成する。“NODE”はkeyに対応するノードデータを取得する。“LINK”はkeyに対応するリンクデータを取得する。“SUBGRAPH”は指定したノードの部分グラフを取得する。

[0106] 図14Cに、本発明に係るシステムの一実施形態におけるグラフ構造及びリンク表現を示す。

[0107] まず、図14C(A)に簡単なグラフ構造を示す。4つのノード1401、1402、1403、1404とそれらのリンクl1(1405)、l2(1406)、l4(1407)、l3(1408)がある。それらの表現を図14C(B)に示す。リンクは特に示さない限り無向リンクである。“(l1, {n1, n2})”は、ノードn1(1501)とノードn2(1502)との間のリンクである。ノードn1(1501)からノードn2(1502)への有向リンクを表す場合には、“(l1, {n1, n2}’)”と表現する。

[0108] また、図14C(C)には、リンクが静的ではなく、動的な関数で表現される場合が示されている。“(n1, {f(n1, n2)})”は、ノードn1(1401)とノードn2(1402)との間のリンクが、関数“f(n1, n2)”で計算される。例えば、ノードn1(1401)がある画像の多次元特徴量ベクトルであり、ノードn2(1402)が別の画像の多次元特徴量ベクトルである時には、“f(n1, n2)”はその情報距離を比較する演算になる。また、ノード間リンクの確率であってもよい。

[0109] 図15に、本発明に係るシステムの一実施形態における検索関連画像と画像構成要素、及び関連要素画像群の視覚的リンク構造例、および動作例を図

示する。

[0110] [視覚的グラフ表現]

図15は、画像1501を関連性検索窓1203にドラッグ&ドロップした後の関連性グラフを示している。画像1501は画像認識エンジン200、または、画像認識システム202で処理され、3個の画像構成要素が抽出されている。すなわち、画像構成要素Wine(1502)、Wine Glass(1503)、Wine Bottle(1504)の3個である。図中では、それらにWine, Wine Glass, Wine Bottleというラベルが付されているが、これらは画面に出力されることはなく、あくまで本発明における説明を容易にするためのものである。これらの画像構成要素は関連検索エンジン220で処理され、それぞれ1505から1518の関連要素群が抽出されている。例えば、画像1502は5個の関連要素群の画像と関連がある。それらは、Olive(1505)、Cheese(1506)、Bread(1507)、Fine Dish(1508)、Wine Glass(1508)の5個である。Wine Glass(1503)はDecanter(1509)を関連要素とし関連がある。Wine Bottle(1504)は8個の関連要素群の画像と関連がある。それらは、Wine(1502)、Wine Glass(1503)、Decanter(1509)、Cork(1511)、Grape(1513)、DRC(1515)、Wine Cellar(1516)、Oak Barrel(1517)の8個である。

[0111] また、画像間のリンク線の太さにも意味を持たせている。太いリンク線は細いリンク線よりも強い関連度を表す。例えば、Wine(1502)は、Olive(1505)とCheese(1506)とにリンクがあるが、Cheese(1506)とのリンクの方がOlive(1505)のリンクよりも本事例では太い。つまり、Wine(1502)とCheese(1506)との関係の方が強いという関係を表している。

[0112] このような関連性は、図15に描かれた範囲を超えて大きく広がっているが、図15では例示的にその一部を表示している。これら広範な関連性グラフは、ノード間の関連度を基準にして、一定以上の関連性を有するもの同士を選択的に表示することもできる。

[0113] 次に、例えば、Decanter(1509)を関連性検索窓に再投入する。この場

合には、画像認識エンジン200が画像Decanter(1509)を処理して、新たな画像構成要素群を抽出し、それに関連した新たな関連要素群を関連性検索エンジン220から抽出して表示する事で、図15とは異なる関連性グラフが展開される。

[0114] 図16(A)に、図15のシナリオに対応するグラフ構造との関係を示す。ここでは、画像1501とその画像構成要素1502から1504の画像に対応するグラフ構造を示している。4個のノード1601、1602、1603、1604は、それぞれの画像構成要素に対応している。

これらの関係を表現するために、GDB102Aにはデータセット1605が格納されている。

[0115] 画像構成要素には様々のノードがリンクされているが、一例として、図16(B)にそれぞれの画像構成要素に対するメタデータの一部を示す。ノード1602にはノード1610(red)とノード1611(white)の2個のメタデータが、ノード1603にはノード1612(crystal)、ノード1613(company name)、ノード1614(creator name)の3個のメタデータが、ノード1604にはノード1615(name)、ノード1616(vintage)、ノード1617(winery)の3個のメタデータがある。さらにこれらのメタデータはさらなる関連ノードとリンクしている(不図示)。

[0116] 図17(A)に、図15のシナリオに対応する関連要素に関するグラフ構造を示す。ここでは、画像1502とその関連要素1501、1503、1504及び1505から1508の画像に対するグラフ構造を示す。7個のノード1601~1604、及び1701~1704がそれぞれ関連要素に対応している。

これらの関係を表現するために、GDB102Aにはデータセット1705が格納されている。

[0117] 図17(B)に関連要素それぞれに対するグラフ構造を示す。図ではスペースの関係上一部のグラフ構造を示している。例えば、ノード1604にはメタデータに対応したノード群1710とさらなる関連リンク群1711と

のリンクがある。同様なリンクは他の関連要素のノード 1 6 0 1 から 1 6 0 3 にも存在している。

[0118] [関連性導出演算]

図 1 8 は本発明に係わる関連性導出演算の一例で、関連性検索エンジン 2 2 0 における関連性演算部 2 2 4 での処理を示したものである。図 1 7 で見たように、画像構成要素の画像と関連要素を構成しているノードとの間には複雑なグラフ構造が存在している。例えば、図 1 8 (A) のグラフが与えられているとする。これは、2つのノード間のグラフ構造から抽出された部分グラフである。ここで、それぞれのノード間のリンクの  $f$  を計算する(図 1 4 (E) の関数  $f$  を計算する)。関数  $f$  は確率であったりベクトルであったりノードとリンクのタイプによって変わる。例えば、リンク 1 8 0 1 の  $f$  を計算し得られた値を行の一要素とし、これをすべてのリンクに繰り返すと、図 1 8 (B) の行列 ( $v1$ ) を得る。図 1 8 (C) は行列 ( $v1$ ) の各行をビンに対応させたヒストグラム図として描いた。この行列 ( $v1$ ) を多次元特徴ベクトルとしてノード間の関連性の計算に用いる。つまり、この多次元特徴ベクトルはノード間の直接的な関係性の強さを表している。図 1 8 ではノード 1 8 0 1 ( $n1$ ) とノード 1 8 0 9 ( $n2$ ) の間の関連性を当該多次元特徴ベクトルで表し、関連性知識データベース 1 3 1 0 に記録する。また、G D B 1 0 2 A にはノード  $n1$  とノード  $n2$  との間にリンクが生成されたことになるので、すなわち、" $(l_{n1-n2}, \{f(v1)\})$ " (ここで  $f(v1)$  は関連性知識データベースへのアクセス関数／メソッドである) なるリンクデータを G D B 1 0 2 A に格納する。この様にして、ノード間の関連性を学習していく。

[0119] なお、 $f(v1)$  による多次元特徴ベクトルを次元圧縮した値をリンク線の太さに割り当ててもよい。この場合、多次元特徴ベクトルの次元圧縮値が大きくなるほど、リンク線は太いものとしてグラフ上に表現することができる。次元圧縮には、公知の演算方法を用いることができる。

[0120] [基本的なインタレスト・グラフの獲得]

図 1 9 に、本発明に係わるインタレスト・グラフ獲得の一例を示す。図 1

9 (A) は関連性グラフをユーザに対応するノード 1903 (当該ノードのタイプが” USER” である) を中心に簡略化して描いた。ノード 1903 は 3 つのオブジェクトに対応したノード 1904、1905、1906 (それらのノードはタイプが” OBJECT” である) と繋がっている。図 18 に記載の手順でノード 1903 とノード 1904、ノード 1905、及びノード 1906 の間のそれぞれの多次元特徴ベクトルを計算し合計したものが図 19 (A) 中の多次元特徴ベクトル 1901 である。

ここで、2 つのオブジェクトがノード 1903 に追加されたとする (図 19 (B))。ノード 1913 とノード 1914 である。同様にノード 1903 とノード 1913、ノード 1914 の間のそれぞれの多次元特徴ベクトルを計算し合計した結果、図 19 (B) 中の多次元特徴ベクトル 1911 が得られる。点線円 1902 と点線円 1912 における特徴ベクトルの違いに注意されたい。この様に多次元特徴ベクトルを適応的に強めることで当該ユーザ 1903 を中心ノードとするインタレスト・グラフを獲得する。

[0121] 前記操作を GDB 102A に格納されたタイプが” USER” のノードそれぞれに適応する事で、個々のユーザに対応したインタレスト・グラフを獲得することができる。

また、関連性差分部 224 による計算を特定のユーザの集まりに対して適応すれば、当該グループのユーザ (いわゆるユーザクラスタ) に関する特徴を表したものになるし、ユーザ全体の当該計算を適応すれば、ユーザ全体に関する特徴を表したものになる。そして、詳細は後述するが統計情報処理部 209 により当該ユーザを中心とした多次元特徴ベクトル群は統計的なインタレスト・グラフを表現する。

[0122] [グラフ構造の表示例]

図 20A 乃至図 20C に、本発明に係るシステムの一実施形態におけるグラフ構造の表示例を示す。

[0123] 図 20A において、グラフ構造におけるノードに対応する画像には、画像 2001 (例えば、イラストレーション) を使用することもできるし、ロゴ

2002や画像サムネイル2003を使用することもできる。さらには、MDB102Bから公式の画像(2004)を使用することもできる。

ここで、会社のロゴの場合には、一例として複数の意味が含まれることに留意されたい。具体的には、会社そのものを指し示し、かつ、その会社の商品も指し示すことがあり得る。

[0124] 図20Bに、関連性グラフを観察時間としての時間軸変数と併せて視覚的に表示する例を示す。

[0125] 図20B(A)は、関連性グラフを横軸として時間軸を、左が過去で右が未来になるように表示した例である。ユーザは表示面をフリップしてもよい(1210)。その場合、時間軸の表示は一昨日、3日前、等々の過去、或いは明後日、3日後、等々の未来の時間軸になり対応するグラフ表示も変化する。図20B(B)は、時間軸を変化させるスクロールバー2011を用意し、ある時刻での関連性グラフを表示し(2012)、スクロールバーを移動させることで別の時間軸での関連性グラフを表示した(2013)例である。

[0126] なお、図20Bの時間軸に加えて、位置情報を元に地図上、或いは地球儀上に紐付けられた関連性グラフを表示しても良い。

[0127] 図20Cは、あるノードを中心とする、より詳細な関連性グラフを表示する例である。表示2021において、ノード2022をダブルタップする(2023)か、ピンチアウトする(図では非表示)事により(図20C(A))、当該ノード2022を中心とした関連性グラフがより詳細化(2031)されて表示される。

例えば、図20C(B)では、ノード2024のさらなるつながりのあるノード2032や、新しいノード2033が追加表示されている。

[0128] [インタレスト・グラフ成長のための演算子]

図21に、本発明に係るシステムの他の実施形態における動作例を示す。図21(A)において、関連性検索窓1203には、入力検索条件として論理演算子(AND2101及びOR2102)が導入される。ここで、AN

D (2101) を指定した場合、画像2104と画像2105とから始まるノード間で共通かつ直接的に関連するノード群が選択される。すなわち、ノード2104とノード2105からの直接リンクのあるノード群が選ばれる。一方、OR (2102) の場合、それぞれのノードのいずれか1つ以上に直接的に関連するノード群が選択される。すなわち、ノード2104から直接リンクのあるノードとノード2105から直接リンクのあるノードの両方を選ばれる。

[0129] 図21 (B) は、AND演算子2101の動作例を示している。グラフデータベース102A上の関連性グラフを探索することで、画像2104に対応するノード2106と画像2105に対応するノード2107から共通かつ直接的に関連するノード2108とノード2109とが選択される。図21 (B) では、ノード2108はイタリアのトスカーナ地方に関するノードを表しており、ノード2109は、ワイナリーに関するノードが表示されている。

[0130] 図22に、本発明に係るシステムの他の実施形態における動作例を示す。

[0131] 図22 (A) は、関連性検索窓の検索条件としてCONNECTION SEARCH (2103) を選んだときの動作を示している。図では2個の画像 (2201と2203) を関連性検索窓にドラッグ&ドロップしている状態を表している。図22 (B) において、画像2201に対応するノード2202からノード2206 (something 1) は到達可能な状態を表し、画像2203に対応するノード2204からのど2209 (something 2) も到達可能な状態を表している。ここで、GDB102Aを検索することで、ノード2206とノード2209のグラフ構造中のリンクを検索し、両ノード間に直接的、或いは間接的リンクが存在する場合には、当該ノードが表示されることになる。

[0132] 直接的なリンクが存在する場合には、GDB102Aから該当する複数のリンクを取り出し、画像へのURLを保持するノードに到達するたびに当該画像を表示していく。

また間接的リンクが存在する場合には、後述の統計情報処理部209を利

用して、ノード2202をルートとする部分グラフをGDB102Aから抽出し多次元特徴ベクトル生成部1302で生成された多次元特徴ベクトルを対象として、例えば、当該多次元ベクトルの共起確率よりも大きな確率を持つ多次元特徴ベクトルをもつノード群を選び出し、ノード2201とノード2203を間接的につないでいく。この方法の場合、当該ノード間をつなぐパスは複数存在する場合があるが、その時には、当該パス上のノードの数が最小になるパス、あるいは、当該パス上のノード間の重みが最小になるパスを最短パスとして、それを含む関連性グラフを表示しても良い。

[0133] CONNECTION SEARCH (2103) の変形としては、1つの画像、例えば画像2201のみを関連性検索窓にドラッグ&ドロップして前記手法により選出されたリンクをつないでいっても良い。

なお、この後には、ノード2202とノード2204との間に直接のリンク(2210)を生成しても良い。

[0134] 前記間接的リンクが存在する場合には、CONNECTION SEACH演算子2103により図23に示した具体的な連想関係を導く事が可能になる。すなわち、画像2301のワインボトルは、ワインという材料を介して画像2302のワイングラスとの関連が抽出され、さらにグラスという種類の物品を介して高級ワイングラス社製のワイングラスの画像2303との関連が抽出され、さらに同じ材質の椅子の画像2304との関連が抽出され、さらに同じ椅子という種類の物品を介して画像2305との関連が抽出され、さらに同じ作者であるという事実を介して画像2306との関連が抽出され、さらに同じ材質であることを介して画像2307との関連が抽出され、さらに同じ形状であるということを介して画像2308との関連が抽出され、さらに、同じ色であることを介して画像2309との関連が抽出され、さらに同じアパレル業者であるという事実を介して画像2310との関連が抽出されている。

このような非直接的なパスが複数発見される場合、前記したように、最も中継ノード数の少ない、或いは、当該パス上のノード間の重みが最小となるような間接関連性を抽出する事が可能である。

さらには、これら複数の非直接的なパスを辿る事により、思いもよらないノード間の関連を発見することができる。

[0135] 図24Aに、本発明に係るシステムの他の実施形態における動作例を示す。

[0136] 図24Aにおいて、ユーザ2401に対応したノード2402には、2つのオブジェクト2403と2404とが関連付けられているとする。例えば、図22に示した操作によって新しいオブジェクト2410が見つかった場合、それに対して接続演算子LIKE(2420)を適用すると(図24A(A))、ユーザのノード2402とオブジェクトのノード2410との間にリンク2411が生成され、直接的な関連付けが行われる(図24A(B))。結果、”(2411, {user A, object C})”なるリンクデータが、新たにGDB102Aに登録される。

[0137] また、上記更新が行われると、この更新による新たな直接関連性の存在(リンク2411の存在)をユーザ2402とユーザ2402に直接的に関連するユーザ2405とに通知するよう構成することもできる。すなわち、図24A(C)に示すように、ユーザ2402のネットワーク端末(不図示)及びユーザ2405のネットワーク端末105上に新たな直接関連性の存在(リンク2411の存在)が通知されている。

[0138] また、図24B(A)においてオブジェクト2410に対して切断演算子DISLIKE(2421)を適用すると、リンク2411が切断され、リンクデータ”(2412, {object C, user A})”という有向リンクを生成しGDB102Aに登録する(図24B(B))。

[0139] これらの接続演算子LIKEや切断演算子DISLIKEによって、ユーザに対応したノードと他のノードとの間の直接的な関連性が変化することで、相当するグラフ構造も変化する。ここで、あるノードを中心にリンクのあるユーザのノードとの間のリンクを関連性演算部224で図18及び図19と同様の手段で計算することにより得られた多次元特徴ベクトルは、当該ノードに対するユーザの関心の深さを表すことになる。

[0140] 図25に、本発明に係るシステムの他の実施形態における動作例を示す。

[0141] 図25において、ユーザ2501に対応したノード2502には、2つのオブジェクト2503と2504とが関連付けられているとする（図25（A））。一方で、ユーザ2511に対応したノード2512には、3つのオブジェクト2513、2514、2515が関連付けられているとする（図25（B））。

ここでまずオブジェクト2504とオブジェクト2515との間には直接的リンク2510は存在しないとする。しかし本発明における関連性検索エンジン220は図23にみてきたような間接的な関連性を見つけ出す可能性がある。そこで、ユーザ2501のネットワーク端末上にオブジェクト2515の存在の可能性がユーザ2501に対して喚起された場合（図25（C））、当該ユーザは、それらを直接に結びつける演算子Reference（2506）を実行することができる。これにより、オブジェクト2504とオブジェクト2515の間のリンクが提起され、関連性差分部224の処理により新たな多次元特徴ベクトルが生成される。このリンク生成の要請が複数発生し所定の閾値を超えた場合、或いは特定の権限を有するスーパーバイザにより、オブジェクト2504とオブジェクト2515を直接関連付けるリンクを図14Bの”CONNECT”オペレーションによって生成する。

[0142] また、ユーザには特定の権限を付与しても良い。その場合、演算子Referenceによるリンク生成の要請は即実行され、オブジェクト2504とオブジェクト2515を直接関連付けるリンクを図14Bの”CONNECT”オペレーションによって生成する。

[0143] 上記と同様に、Olive Tree(1519)とGrape(1513)との間に点線の暫定的なリンク(1520)が描かれている（図15）。当初の関連性グラフの観点からは両者の距離は離れているが、演算子Referenceによって、ユーザが直接関連付けることを提案することができる。その際に、当該関連性の有無についての提案をきっかけにしたユーザ間のコミュニケーション

を誘発してもよい。その結果、当該提案が正当である場合には、リンク 1 5 2 0 は実線（直接的関係として確立された）のリンクとして更新することができる。

[0144] 図 2 6 に、本発明に係るシステムの他の実施形態における動作例を示す。

[0145] 図 2 6 では、図 2 4 と同様に、ユーザ 2 3 0 1 に対応したノード 2 3 0 2 には、2 つオブジェクト 2 3 0 3 と 2 3 0 4 とが関連付けられている（図 2 6（A））。一方で、ユーザ 2 4 0 1 に対応したノード 2 4 0 2 には、3 つのオブジェクト 2 4 0 3、2 4 0 4、2 4 0 5 が関連付けられ（図 2 6（B））、オブジェクト 2 3 0 4 とオブジェクト 2 4 0 5 との間には直接の関連付けリンク 2 5 0 1 が存在している状態を表している。ユーザ 2 3 0 1 は、それらの関連付けが疑わしいとして演算子 `Unreference`（2 4 0 6）を実行する。この直接的関係の非存在を提起する要請が複数発生し所定の閾値を超えた場合、例えば、一定数を超えたユーザが `Unreference` 演算を実行した場合には、オブジェクト 2 3 0 4 とオブジェクト 2 4 0 5 との間の直接的関連は誤認として切断処理される。或いはスーパーバイザの権限によって、この要請が確定した場合には、オブジェクト 2 3 0 4 とオブジェクト 2 4 0 5 との間の直接的関連は同様に切断され得る。

[0146] [統計情報処理部]

図 2 7 を用いて、統計情報処理部 2 0 9 の一実施形態における機能ブロック構成を説明する。

統計情報処理部 2 0 9 は 3 つの要素から構成される。グラフ・ベクトル構成部 2 7 0 1、推論エンジン部 2 7 0 2、およびグラフマイニング処理部 2 7 0 3 の 3 つであり、推論エンジン部 2 7 0 3 はさらに決定木処理部 2 7 1 0、ベイジアンネットワーク処理部 2 7 1 1 から、グラフマイニング処理部 2 7 0 3 はパターンマイニング処理部 2 7 1 2、RWR(Random Walk with Restarts)処理部 2 7 1 3 から構成されている。なお、グラフマイニングの処理手順はこれらに限らない。

[0147] 図 2 7 のグラフ・ベクトル構成部 2 7 0 1 で、GDB 1 0 2 A からのデー

タ、及び／乃至統計情報データベース２７０４からのデータを用いて、入力ノードに関係した部分グラフを抽出し、関連性演算部２２４での処理から多次元特徴ベクトルを得て、推論エンジン部２７０２の入力とする。推論エンジン部２７０２では、知識データベース２７０５の情報を参考に、確率的推論モデルの１つである決定木法による処理を実行する決定木処理部２７１０で処理されたり、あるいは、ベイジアンネットワーク構成部２７１１で処理されたりして、頻出主要部分抽出のためのグラフマイニング処理部２７０３に送られる。グラフマイニング処理部２７０３では、パターンマイニング手法やRWR手法をもちいたグラフマイニング処理を行った結果の部分グラフを（図２８（Ｃ））を生成する。

[0148] [ユーザの主観によるフィルター処理]

図２８に本発明に係るシステムの一実施形態における特定ユーザフィルター処理部２１０の構成を示す。当該処理部は３個の要素から構成される。多次元ベクトル構成部２８０１、主観フィルター構成部２８０２、及び多次元ベクトル処理部２８０３である。GDB１０２Ａから取り出され、例えば統計情報処理部２０９で処理された結果の部分グラフは、多次元ベクトル構成部２８０１にて多次元ベクトルとして再構成される。ユーザ情報は主観フィルター処理部２８０２において、ユーザデータベース２８０４の情報をを用いてユーザの主観的評価フィルターを多次元特徴ベクトル（図２８（Ｂ））として生成する。これら２個の構成部（２８０１と２８０２）からの出力の多次元特徴ベクトルが、多次元特徴ベクトル処理部２８０３で処理され、図２８（Ｃ）のユーザの主観的評価を反映したノード間のつながりの深さを表した多次元特徴ベクトルとして再構成される。

[0149] ユーザデータベース２８０４からはユーザとリンクのあるノードの評価が数値化されている。当該数値化は、学習によるプロセス、或いはユーザにより直接的に階指定しても良いし、ユーザと当該ノードとの間のリンク数を用いて求めても良い。当該多次元特徴ベクトルの該当ビンの値と当該評価を演算することで、それぞれのユーザの好みに即した重み付け処理が可能である

。

- [0150] また、前記プロセスをより広範なユーザに対応するプレファレンスとして一般化する一手法として、広範なユーザの共通する主観として、次のような要素として表現することも可能である。

|       |                           |
|-------|---------------------------|
| 外見・外観 | ハンサム、ぼろぼろ、清潔、美しい、愛らしい、・・・ |
| 評価表現  | 満足、すばらしい、・・・              |
| ラベリング | 偉い、尊敬、傲慢、・・・              |
| 嗜好    | 好き、美味しい、まろやか、甘味、・・・       |
| 心理状態  | 危機感、怒り、寂しい、悲しい、気力、・・・     |

これらはユーザデータベース 2804 に登録可能であると同時に、これら主観が” SUBJECT” として汎用的に適用可能となる。前記主観フィルター構成部 2802 は上記主観要素を構成する部分グラフから多次元特徴ベクトル（図 28（B））を生成する。当該多次元特徴ベクトルの各ビンの値は、上記主観要素を反映した数値として利用可能である。

- [0151] 図 29 に、前記ユーザ毎に異なる主観性を視覚的に関連性グラフとして表現する事例を示す。図 29（A）ではユーザ 2901 はオブジェクト 2902 はじめとする 6 個のオブジェクトが直接的・間接的に繋がっている状態を示している。それぞれの関連性の深さは 2 次まで表示されている。それぞれのリンク線の太さは関連性の強さを表している。前記に関わる処理を経由することでユーザ 2901 がオブジェクト 2906 に特別な関心を持っていることが判れば、当該オブジェクトを強調表示（2908）しても良い。

- [0152] 一方で、図 29（B）ではユーザ 2911 はユーザ 2901 と類似の関連性グラフを共有しているが、その興味の対象がオブジェクト 2902 である場合、当該オブジェクトをハイライトや特別な装飾効果を含む視覚効果（2912）を施して表示しても良い。

- [0153] 図 28（A）における主観フィルター構成部 2802 の入力に、代わりにユーザが検索しようとしている時間帯や、日時、季節、時代、場所といった時間軸や位置情報を反映した環境フィルターを構成する多次元ベクトルを用いても良い。

[0154] 図30に、本発明に係るシステムの他の実施形態における動作例を示す。

[0155] 図30(A)に示す画像3001に対応したノード3002とグラフ構造3003のメタデータが既に存在していた場合、図30(B)に示すように画像認識エンジン200が上記メタデータの正当性を優先的に比較処理することで図30(A)に示すメタデータのグラフ構造と、図30(B)に示すメタデータのグラフ構造との一致を確認する事が可能となる。これにより、大幅な処理時間の短縮が実現できる。

[0156] 図31(A)にインタレスト・グラフの一例を示す。図では簡略化の為にノードとしてユーザとオブジェクト(モノ)のみを描いているが、実際には画像からは画像認識システム202によってコンテキストやシーンのようなモノ以外の情報も抽出されているのでインタレスト・グラフの要素になる。

図には、3人のユーザ3101から3103と6個のオブジェクト3110から3115の関係が描かれている。ユーザ3101はオブジェクト3110、3111、3112に興味があり、ユーザ3102はオブジェクト3111、3113、3114に興味があり、ユーザ3103はオブジェクト3110、3111、3113、3115に興味があることが描かれている。

このインタレスト・グラフはGDB102Aのデータからユーザと関連のあるノードであり、グラフ演算部221が抽出したものであり、関連性検索エンジン220におけるグラフ記憶部222に存在している。

上記接続演算子LIKE、切断演算子DISLIKE、参照演算子Reference、非参照演算子UnreferenceによってGDB102Aの情報は時々刻々と変化しているので、この図31(A)のインタレスト・グラフも動的なインタレスト・グラフとして獲得できる。

[0157] 図31(B)では、ユーザ3102が例えば接続演算子LIKEによってオブジェクト3112と新しい関係(リンク3201)を結んでいる。このような操作が多くのユーザによって行われ、サーバにおけるオブジェクト3112へのリンク数が増加(図31(B)では1から2に変化した)する。そ

の際に、リンク数に所定の閾値を設定し、それを超えることを、当該ノードに対する着目点の大幅な変化としてとらえ、当該ノード（３１１２）に係るノード（３１０４）にその変化を通知する。例えば、図ではノード３１０４は広告主を例示しており、広告主に対してオブジェクト３１１２へのリンク数が閾値を超えて変化したとして通知する事が可能となる。

[0158] また、上記通知はオブジェクト３１１２と直接関係のあるユーザ３１０１や３１０２に対して行っても良い。また、上記通知は、広告主による、対象オブジェクトに関する購買意欲を喚起する広告やリコメンデーションの提示を可能にするものであっても良い。例えば、図３１（Ｂ）では、ノード３１０４が広告主の場合を例示しており、オブジェクト３１１２（対応画像は３１２０）に関しての広告をユーザ３１０１やユーザ３１０２に通知する事が可能となる。

[0159] 図３２では、ユーザ自身を中心ノードとするインタレスト・グラフを表示する際に、プライバシー保護の観点から当該ユーザのみに限定してインタレスト・グラフを表示する例を示している。図３２では、ユーザ３２０１を中心にインタレスト・グラフが表示されているが、ユーザ３２０２を中心にしたインタレスト・グラフ（灰色の囲み３２１０）はユーザ３２０２のプライバシーに係わることであり、ユーザ３２０１の観点からは表示されないことが要件となる。これらは、ノードのタイプ”USER”をサーバ側で区別することで表示・非表示を制御できる。

[0160] 図３３に、ソーシャル・グラフ獲得の一実施例を示す。本発明における画像認識エンジン２００或いは画像認識システム２０２を組み込んだ一連の関連性検索エンジン２２０を利用した視覚的な関連性検索の過程を通じて獲得した包括的なインタレスト・グラフから、特定のユーザにノードのタイプが”USER”である人を表すノードを抜き出して平面３３０１に写像する事で、人と人との関連性を含むソーシャル・グラフを獲得することができる。

[0161] 図３４にインタレスト・グラフ収集に関するプロセス図の概略を示す。プロセス全体の系は、リアルタイム系とバックグラウンド系とに分かれ、グラ

フ記憶部 2 2 2 が両者を繋いでいる。バックグラウンド系には G D B 1 0 2 A、関連性演算部 2 2 4、統計情報処理部 2 0 9 を配する。リアルタイム系には、画像認識システム 2 0 2（或いは画像認識エンジン 2 0 0（不図示））、グラフ演算部 2 2 1、ネットワーク通信制御部 2 0 4 を配する。ネットワーク通信英魚部 2 0 4 はインターネットを含むネットワークを介してネットワーク端末 1 0 5 に繋がっている。

[0162] インタレスト・グラフは、ユーザと一次繋がりノード群の中から、ユーザとの関連度の大きい順に所定数の要素群を多次元ベクトルを選択抽出し、ユーザ固有の有限長の多次元特徴ベクトルとして求める。

[0163] 図 3 5 はユーザ毎のインタレスト・グラフに対応する多次元特徴ベクトルのイメージである。全てのノードに対応するインタレスト・グラフの総候補次元数は、G D B 1 0 2 A に登録されている網羅的なノード数のオーダーに達する事から、当該ユーザとの関連性の大きい順に一定数を候補から抽出し、図 3 5 で示すように、有限長の多次元特徴ベクトルとしてユーザデータベース 2 8 0 4 に収納される。図 3 5（A）では、Aさんの、図 3 5（B）ではBさんの多次元特徴ベクトルを例示している。この2つの多次元特徴ベクトルを相互に比較することで、AさんとBさんが共通に関心を示す対象を検出することができる。例えば、点線円 3 6 0 1 と点線円 3 6 0 2 には同じ対象物に対する関心のベクトル値が表現されており、これらが共通に大きいことが判明すれば、双方の当該対象物に対するインタレストが似通っていることを知ることができる。

## 符号の説明

- [0164] 1 0 0 インタレスト・グラフ収集システム  
1 0 1 サーバ  
1 0 2 A グラフデータベース（G D B）  
1 0 2 B マザーデータベース（M D B）  
1 0 3 接続  
1 0 4 ネットワーク（あるいはインターネット）

- 105 a ~ 105 d ネットワーク端末装置
- 106 一般物体認識システム
- 107 画像カテゴリデータベース
- 108 シーン認識システム
- 109 シーン構成要素データベース
- 110 特定物体認識システム
- 200 画像認識エンジン
- 209 統計情報処理部
- 210 特定ユーザフィルター処理部
- 220 関連性検索エンジン
- 221 グラフ演算部
- 222 グラフ記憶部
- 223 グラフ管理部
- 224 関連性演算部

## 請求の範囲

### [請求項1]

キーワードやメタデータ、或いは文章といったような表意的な文字による入力手段によらず、様々な物体(object)や主題(subject)を内包する画像情報を入力手段として用いる検索システムであって、インターネット上或いは専用ネットワーク上に存在する多数の画像群、或いはネットワーク端末経由でユーザがインターネット上にアップロードした画像の中から、ユーザが関心を持った画像全体、或いは画像の特定の領域をユーザがネットワーク端末上で選択し、それら選択した画像をネットワークを経由してサーバ側の画像認識エンジンに問い合わせる事により、インターネットを介してサーバ側の画像認識エンジンが、選択された画像全体、或いは指定された画像領域に含まれる様々な一般物体・特定物体・人・顔・シーン・文字・記号・イラストレーション、ロゴ、ファビコン(等)をリアルタイムに抽出・認識し、それら認識された入力画像に含まれる画像構成要素群を、画像認識エンジン経由でサーバ側の関連性検索エンジンに通知する事で、関連性検索エンジンが個々の画像構成要素群それぞれに対し直接的及び間接的に一定以上の関連があると判断したその他の関連要素群を、関連性検索エンジン内の関連性知識データベース上に学習可能な状態で収納されている要素間の直接関連性を記述した多次元特徴ベクトルを基に抽出し、上記画像認識エンジンにより認識された画像構成要素群及び関連性検索エンジンにより抽出された関連要素群それぞれをノードとする関連性グラフとして、各ノード間の関連性の深さと共に、ユーザのネットワーク端末上に二次元画像、或いは奥行きを持った三次元画像、或いは上記関連性グラフの観察時間としての時間軸変数を加えた四次元時空間画像として視覚的に表現する事を可能にしたことを特徴とするインタレスト・グラフ収集システム。

### [請求項2]

上記関連性検索操作において、ネットワーク端末上に表示されている関連性グラフ上の任意のノードを、ユーザがタッチスクリーン上で

タップ或いはタッチして選択するか、ポインタのカーソルを任意のノード上に移動し選択する、或いは関連性グラフ上の任意の領域に向けてユーザがタッチスクリーン上でフリックするか、ポインタのカーソルを関連性グラフ上の任意の領域に移動し画面全体をドラッグしスクロールする、或いは方向キー等による同様の操作、或いはユーザによるジェスチャーや視線、音声、或いは脳波を用いた同様の効果を及ぼす入力操作を用いる事により、当該選択されたノード、或いは移動後の領域を中心とする新たな関連性グラフを、そこに至る途中経過も含めて関連性検索エンジンがネットワーク端末に向け追加的に送出する事で、ユーザにとって関心のあるノード或いは領域をユーザが関連性グラフ上でシームレスに辿りながら、複数のノード間にまたがる広範な関連性として視覚的に認識する事が可能な、請求項 1 に記載のインタレスト・グラフ収集システム。

[請求項3]

上記関連性検索操作において、上記画像認識エンジンが提示した複数の画像構成要素群の中から、ユーザが選択した特定の画像構成要素、或いはネットワーク端末上に表示されている関連性グラフ上の特定のノードをユーザがタッチスクリーン上でダブルタップするかピンチアウトする操作、或いはポインタ等の操作による当該ノードの選択と当該ノードを中心とする領域の拡大、或いはユーザによるジェスチャーや視線、音声、或いは脳波を用いた同様の効果を及ぼす入力操作を用いる事により、当該ノードを中心とするより詳細な関連性グラフをユーザのネットワーク端末上に視覚的に表現する事を可能にした上で、それらの一連の操作を当該ノードに対するユーザの一定の関心の存在であるとし、当該ユーザを中心ノードとする各要素間の直接関連性を記述した多次元特徴ベクトル上で、当該ノードに係るユーザの関心の深さを表わす特徴ベクトル値を適応的に強める事で、当該ユーザを中心ノードとする個々のユーザに対応したインタレスト・グラフとして獲得可能にすると共に、当該インタレスト・グラフを広範なユーザ

に拡大して獲得する事により、特定のユーザクラス、或いはユーザ全体にまたがる統計的な広義のインタレスト・グラフとして収集可能な、請求項1に記載のインタレスト・グラフ収集システム。

[請求項4]       上記関連性検索操作において、ユーザが着目し選択したノードを関連性グラフ上で辿る事なく、それらノード画像をネットワークを経由してサーバ側の画像認識エンジンに再び問い合わせる事により、画像認識エンジンの助けを借りて当該ノードに関する新たな画像構成要素群を獲得し、それら画像構成要素群を起点とする新たな関連要素群を関連性検索エンジンからネットワーク端末に向け送出する事で、ユーザが当該ノードに対する新たな関連性をそれら相互の関連性の深さと共に関連性グラフ上で視覚的に認識する事を可能にした上で、直前の同様の操作において当該ノードへの起点となっている画像構成要素から当該ノードに至る一連のノード間の関連性の存在をユーザが認識し利用していると関連性検索エンジンが推測し、各要素間の直接関連性を記述した多次元特徴ベクトル上で、それら一連の関連性を構成するそれぞれのノード間の直接的な関係性の深さを表わす特徴ベクトル値を適応的に強める事で、関連性検索エンジン内の関連性知識データベースの追加学習を更に可能にする、請求項1に記載のインタレスト・グラフ収集システム。

[請求項5]       上記関連性検索操作において、画像認識エンジンにより認識可能となった画像構成要素群、及びそれら画像構成要素群それぞれに対応する各関連要素群に対し、それらを代表する写真、イラストレーション、文字、記号、ロゴ、ファビコン等から生成した縮小画像サムネイルを関連性検索エンジンが元画像に代えてネットワーク端末に向け送出する事により、関連性グラフ上のノードとして画像サムネイル単位での表示・選択を更に可能にする、請求項1に記載のインタレスト・グラフ収集システム。

[請求項6]       上記関連性検索操作において、複数のノードをサーバ側の画像認識

エンジンに問い合わせる事を可能にした上で、画像認識プロセスに備わる入力条件選択機能として、論理的な演算子（AND、OR）を導入し、ANDを選択した場合はそれぞれのノード間で共通かつ直接的に関連するノードを、ORを選択した場合はそれぞれのノードのいずれか一つ以上に直接的に関連するノードを、相互の関連度の深さと共にネットワーク端末上に視覚的に表現する事を更に可能にする、請求項1に記載のインタレスト・グラフ収集システム。

[請求項7]      上記関連性検索操作において、複数のノードをサーバ側の画像認識エンジンに問い合わせる事を可能にした上で、画像認識プロセスに備わる入力条件選択機能として、関連性探索演算子（Connection Search）を導入し、一見全く関連性がないと思われるような複数のノード間の関係を、それぞれの入力ノード群に対し直接的及び間接的に関連するその他のノードを経由する一連の関連性として探索する事で、異なるレイヤ（階層）に及ぶノード間の間接的な関係を発見し、それらノード間の最短パスを含む関連性グラフとしてネットワーク端末上に表示する事を可能にすると同時に、上記関連性探索プロセスにおいて、発見された複数のノード間にまたがる上記間接的な関係を、関連性検索エンジン内の関連性知識データベース内に学習の上獲得する事で、以降の同様或いは類似の関連性探索要求に備える事が更に可能な、請求項1に記載のインタレスト・グラフ収集システム。

[請求項8]      上記関連性検索操作において、ユーザと間接的な関係にあるノード、或いはユーザとの関連性が凡そないとされているその他のノードに対し、それらノードを当該ユーザとの直接的な関係として結び付ける接続演算子（LIKE）、及び既に結びつけられている当該ノードとユーザ間の直接的な関係を切断する切断演算子（DISLIKE）を導入する事で、当該ユーザを中心ノードとする各要素間の直接関連性を記述した多次元特徴ベクトル上で、当該ノードに係るユーザの関心

の深さを表わす値を増加、或いは減少、或いは滅消させる事で、当該ユーザを中心ノードとする、個々のユーザに対応したインタレスト・グラフの更新が更に可能な、請求項1に記載のインタレスト・グラフ収集システム。

[請求項9]       上記関連性検索操作において、ユーザ以外のノードを対象にした新たな直接的関係の存在及び非存在の可能性を、それらの複数のノードが直接的に結び付けられるべきだとして提起する参照演算子（REFERENCE）、及び既に直接的に結び付いているもののその直接的な関連の存在が疑わしいとして直接的関係の非存在を提起する非参照演算子（UNREFERENCE）を導入する事で、それらの新たなノード間の直接関連性の存在、或いは非存在の可能性を関連性検索エンジンがユーザに対し喚起する事を可能にした上で、特定の権限を有するスーパーバイザ或いは一定数以上のユーザから関連がある或いはないと判断されたノード間の関連性に係る特徴ベクトルの値を関連性検索エンジンが更新可能にした上で、当該ノード群に係る更新された関連性グラフとしてネットワーク端末上に反映可能にすると共に、それらの新たな直接関連性の存在或いは非存在に係る更新情報を、全てのユーザに通知する事が更に可能な、請求項1に記載のインタレスト・グラフ収集システム。

[請求項10]       上記関連性検索操作において、関連性検索エンジン内に関連性知識データベースとして収納されている要素間の直接関連性を記述した多次元特徴ベクトルに対し、ユーザそれぞれの主観的な評価が反映可能になるような重み付け操作を可能にし、それら修飾された多次元特徴ベクトルを基に、各ノード間相互の関連性及び関連性の深さを、ユーザのネットワーク端末上に、ユーザがそれぞれ持つ感じ方の違いという要因を反映した関連性グラフとして視覚的に表現する事を更に可能にする、請求項1に記載のインタレスト・グラフ収集システム。

[請求項11]       上記関連性検索操作において、関連性検索エンジン内に関連性知識

データベースとして収納されている要素間の直接関連性を記述した多次元特徴ベクトルに対し、ユーザが検索しようとする時間帯や日時、季節、時代、場所といった環境フィルターを適用可能にし、それら修飾された多次元特徴ベクトルを基に、各ノード間相互の関連性及び関連性の深さを、ユーザのネットワーク端末上に、観察時間や地域（Location）特性といった時空間要因を反映した関連性グラフとして表現する事を更に可能にする、請求項1に記載のインタレスト・グラフ収集システム。

[請求項12]      上記関連性検索操作において、ユーザが選択したノードに対応した画像自体に予めメタデータが付帯している場合、当該メタデータと当該画像との適合性、或いはメタデータ自体の正当性を検証する目的で、上記画像認識エンジンが当該画像から抽出可能な様々な特徴群及びそれらの特徴量と、メタデータで表現される物体(object)や主題(subject)が有する固有の特徴群及びそれらの特徴量とを優先的に照合する事で、画像認識エンジンの画像認識に係る処理時間を短縮可能とすると共に、当該画像には無関係か或いは一定以上の関係が認められないメタデータを排除する事を更に可能にする、請求項1に記載のインタレスト・グラフ収集システム。

[請求項13]      上記ユーザによる一連の関連性検索の過程を通じて、人以外のノードに対する特定のユーザクラスタ或いはユーザ全体集団に及ぶ視覚的な関心の在り所や関心の遷移が一定以上変化した時点で、当該ノードに対する着目度の大幅な変化を統計情報化し、関連性グラフ上でそれらノードに直接的に関わるユーザ、或いはユーザクラスタ、或いはユーザ全体、或いは特定の第三者に速やかに通知する事を更に可能にする、請求項1に記載のインタレスト・グラフ収集システム。

[請求項14]      上記関連性グラフの表示に関して、ユーザ自身を中心ノードとするインタレスト・グラフの表示は、プライバシー保護の観点から当該ユーザのみに限定する事を更に可能とする、請求項1に記載のインタレ

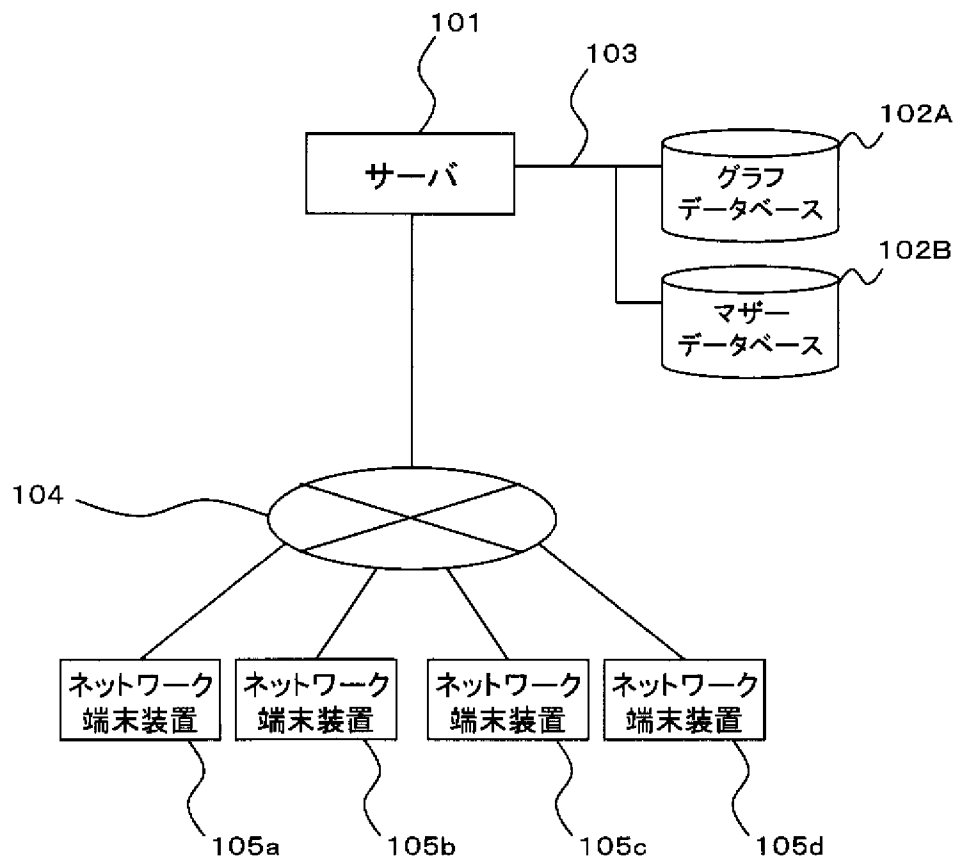
スト・グラフ収集システム。

[請求項15]      上記インタレスト・グラフを活用して、特定のノードが表わす商品やサービスに対し一定以上の関心を示した特定のユーザ、或いは特定のユーザ群に対し、それら商品或いはサービスへの購買意欲を喚起する広告やリコメンデーションの直接的な提示や、それらを提供する第三者への繋ぎ込みを可能にすると同時に、ユーザの属性やユーザの属する時空間要因を加味した上記関連性グラフを活用して、さらにそれらと直接的或いは間接的に関連する他の商品やサービスに対する広告やリコメンデーションの直接的な提示や、それらを提供する第三者への繋ぎ込みを更に可能にする、請求項1に記載のインタレスト・グラフ収集システム。

[請求項16]      上記インタレスト・グラフを活用する事により提示可能になる広告やサービス、或いはリコメンデーションを表わす視覚情報及び／又はリンク情報を、対象となるユーザのネットワーク端末上に表示されている関連性グラフ上に、表示／非表示を選択可能な状態で提示することを更に可能にする、請求項1に記載のインタレスト・グラフ収集システム。

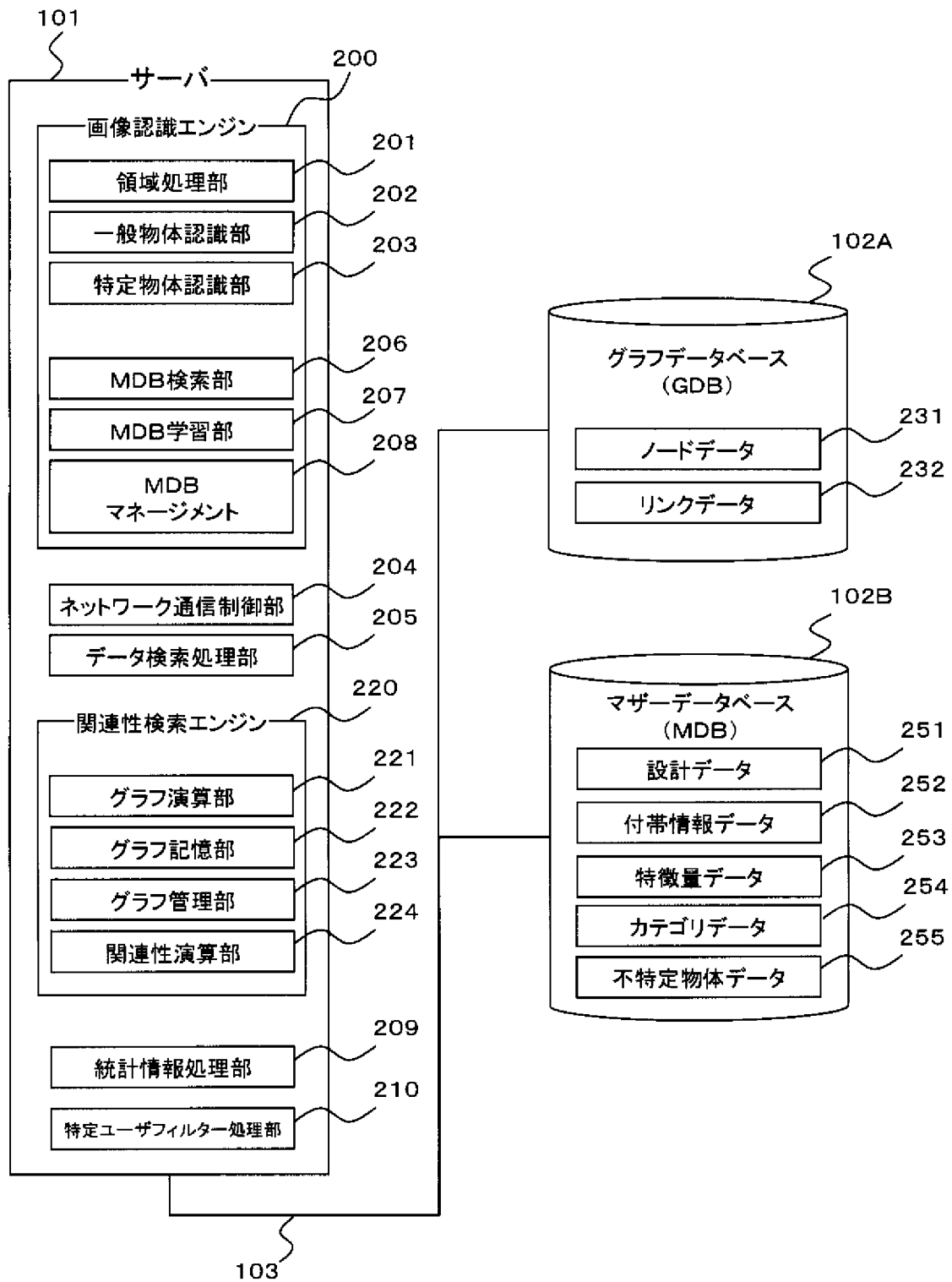
[請求項17]      上記画像認識エンジンを組み込んだ一連の視覚的な関連性検索の過程を通じて獲得したインタレスト・グラフを活用して、同様の関心を持つ広範なユーザ間のコミュニケーションを、上記関連性検索を通じて喚起し、インタレスト・グラフに加えて人と人、人と人以外の関係性を包含する、広義の動的なソーシャル・グラフを、ネットワークを介して獲得する事を更に可能にする、請求項1に記載のインタレスト・グラフ収集システム。

[図1]

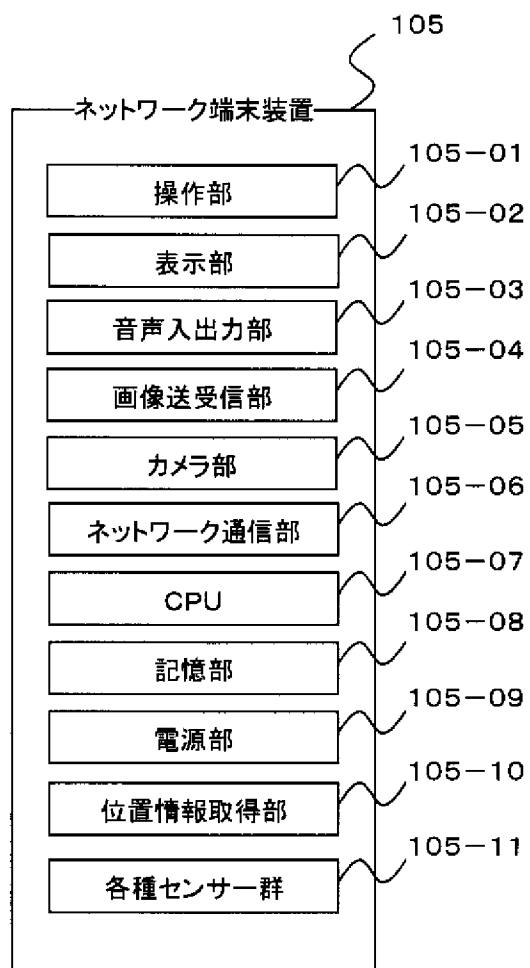
100

[図2]

## 関連性検索システム

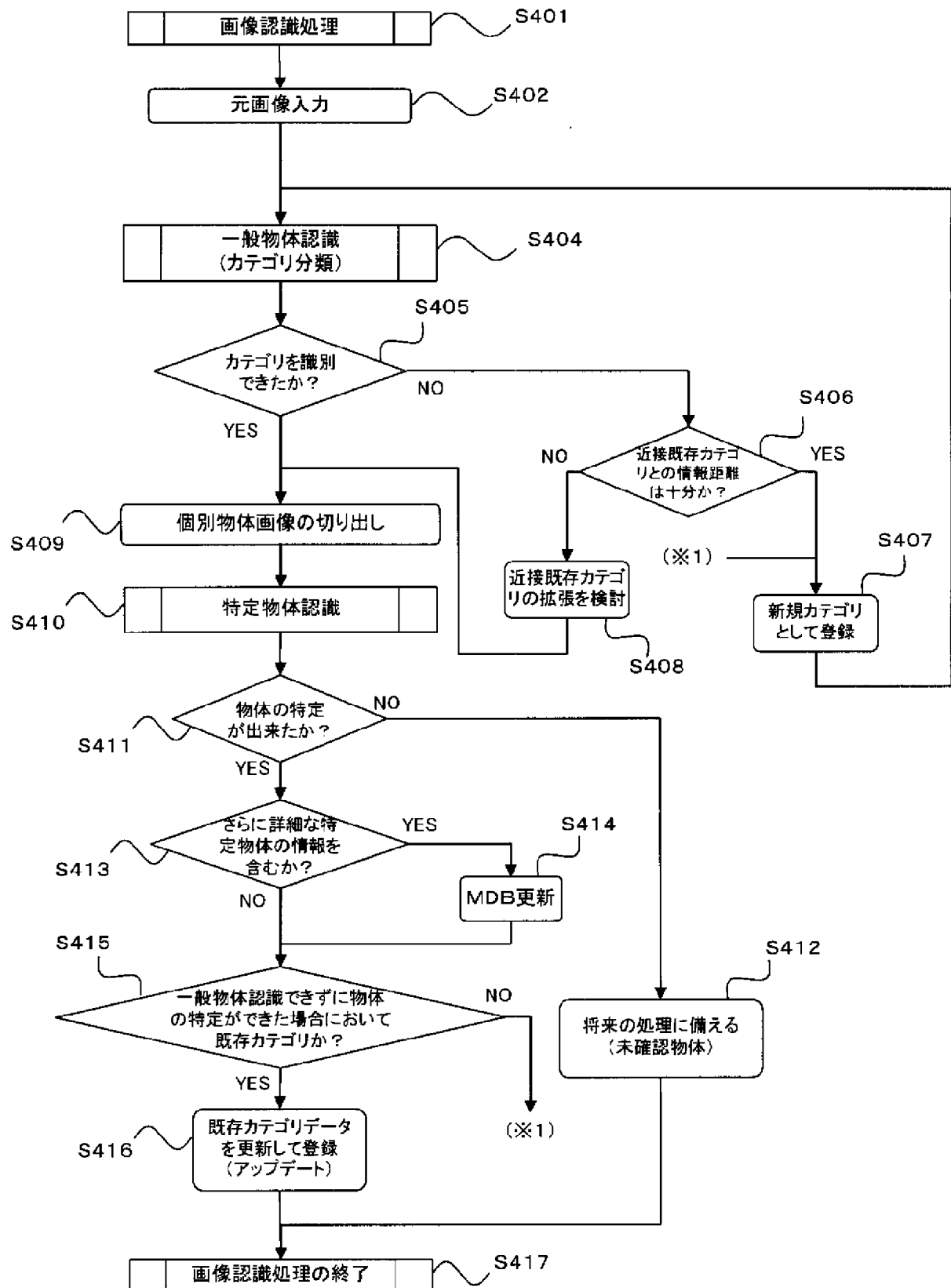


[図3]

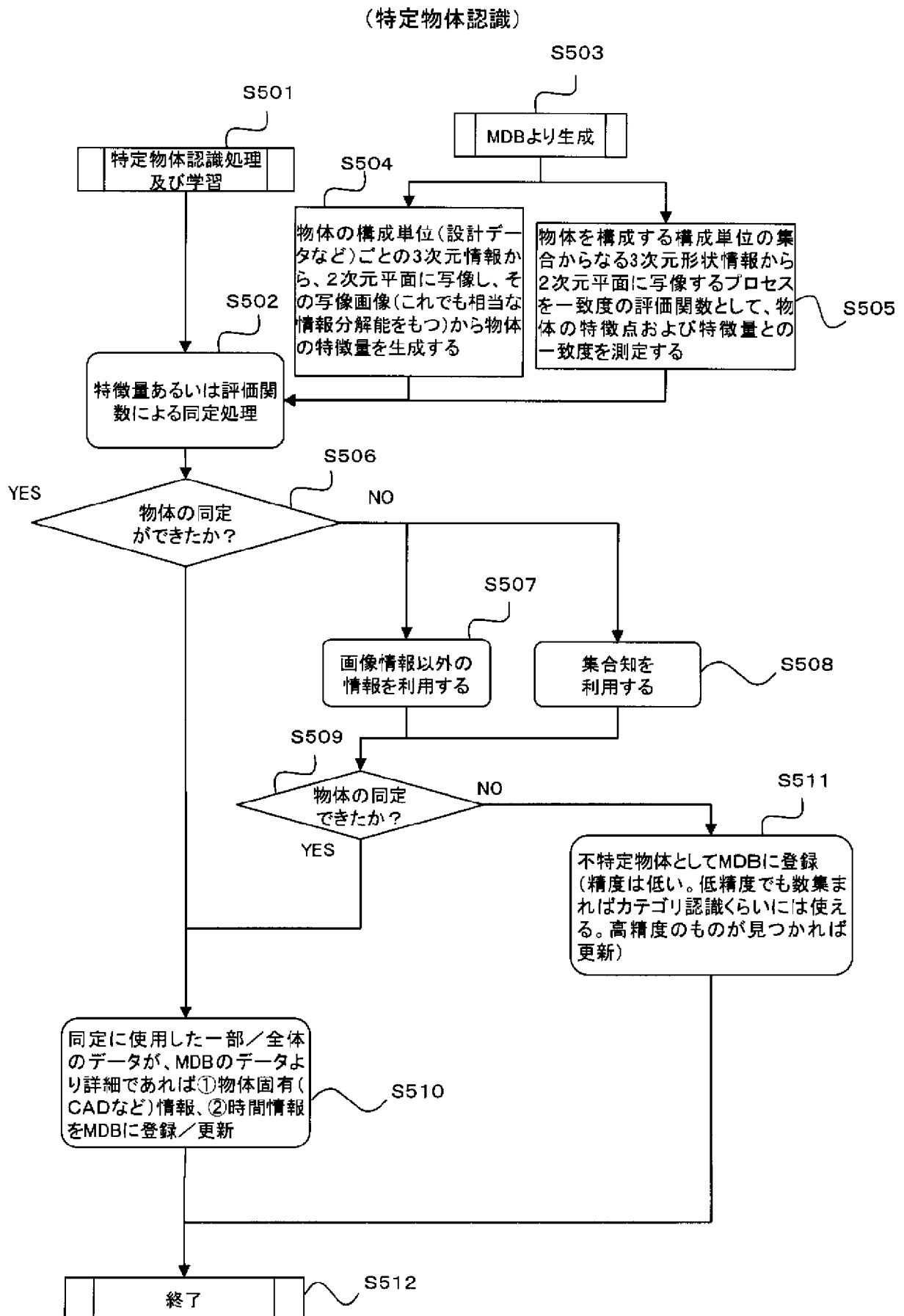


[図4]

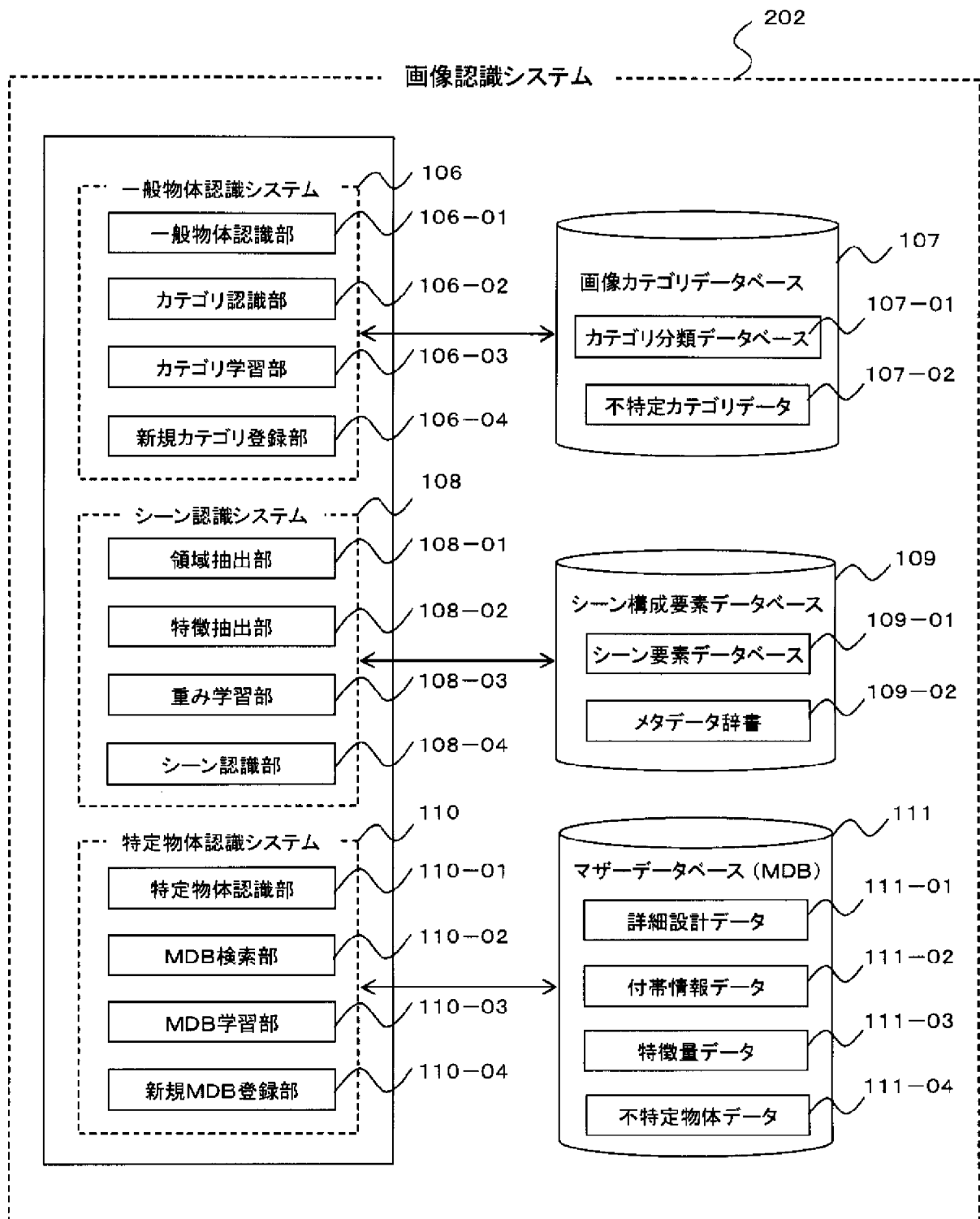
(認識フロー全般)



[図5]



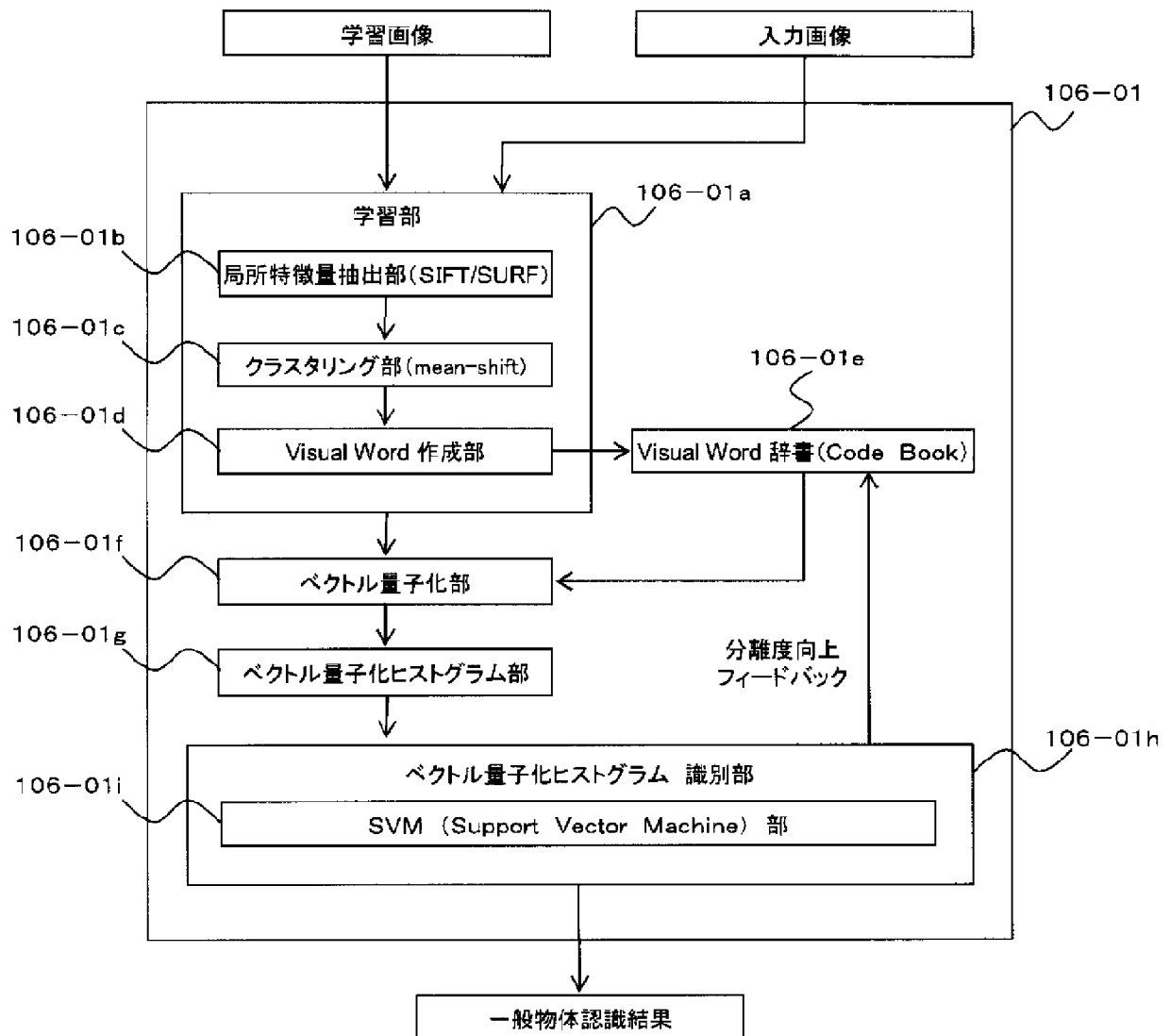
[図6A]



[図6B]

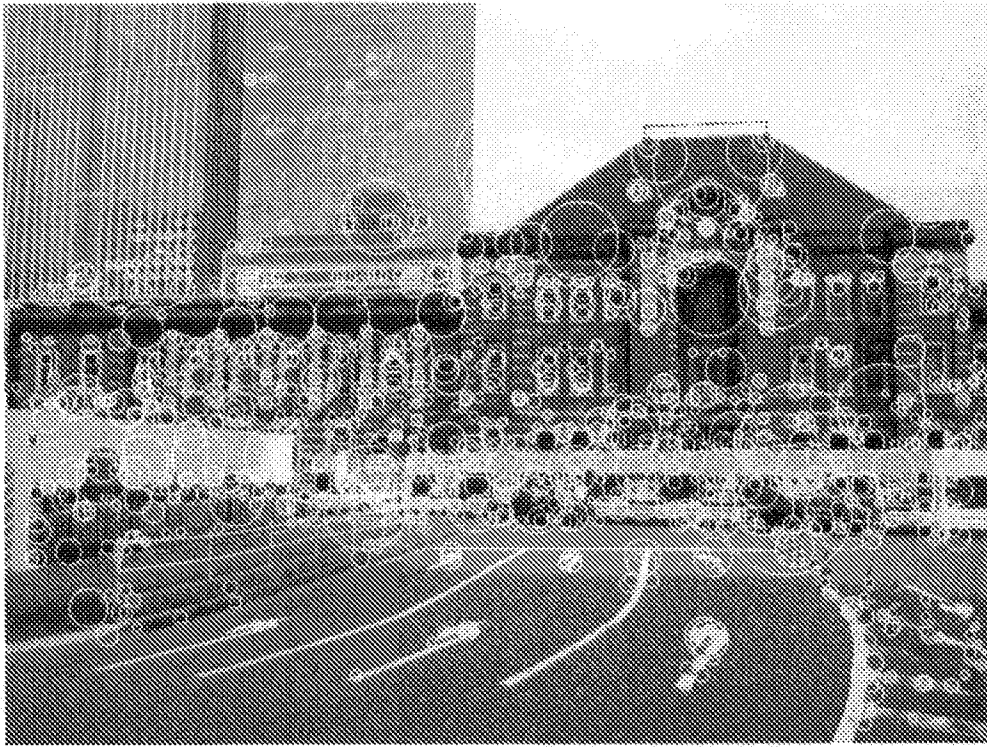
## 一般物体認識部

(Bag-of-Feature)



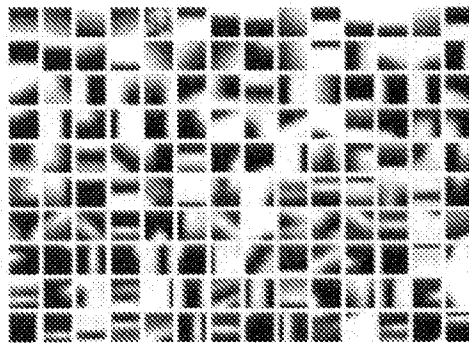
[図7]

(A)



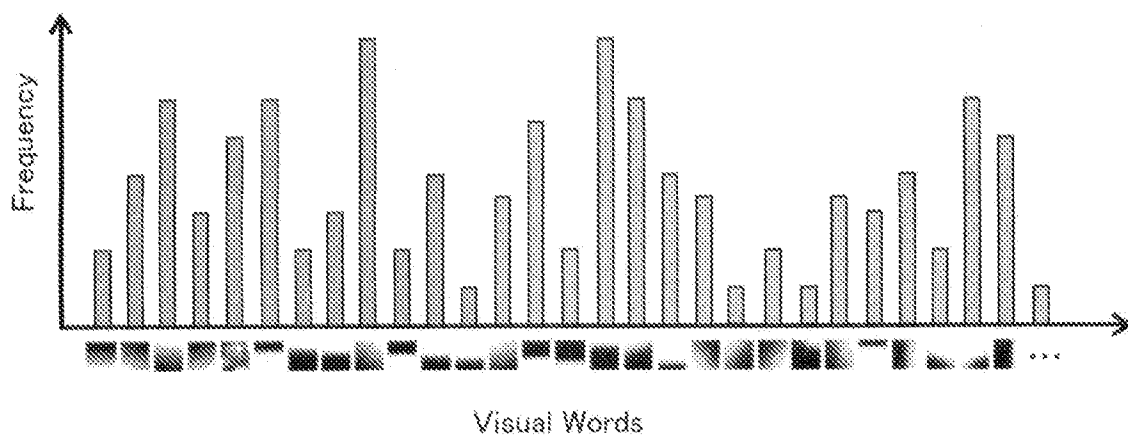
SIFT を用いた特徴点／特徴量抽出

(B)



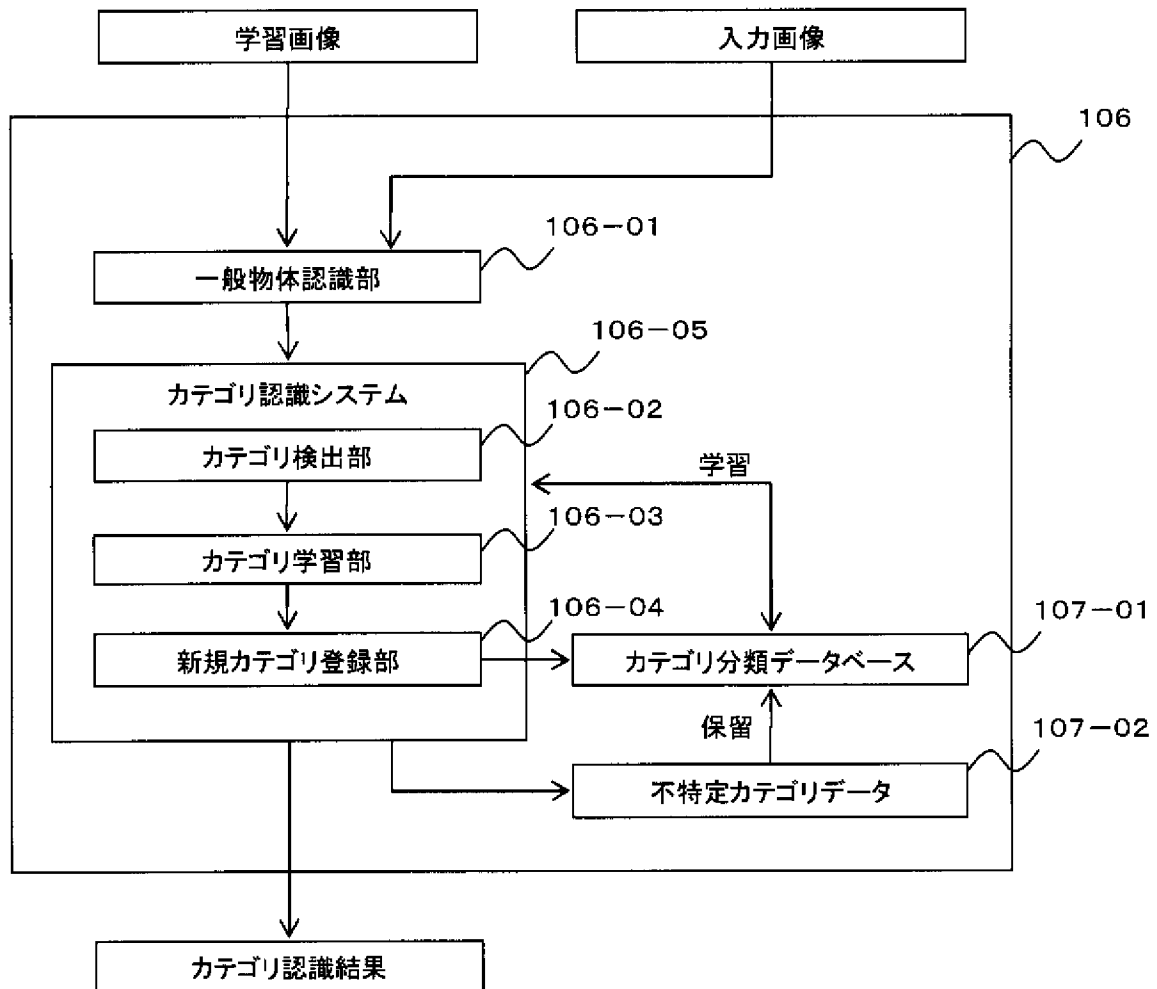
Codebook

(C)



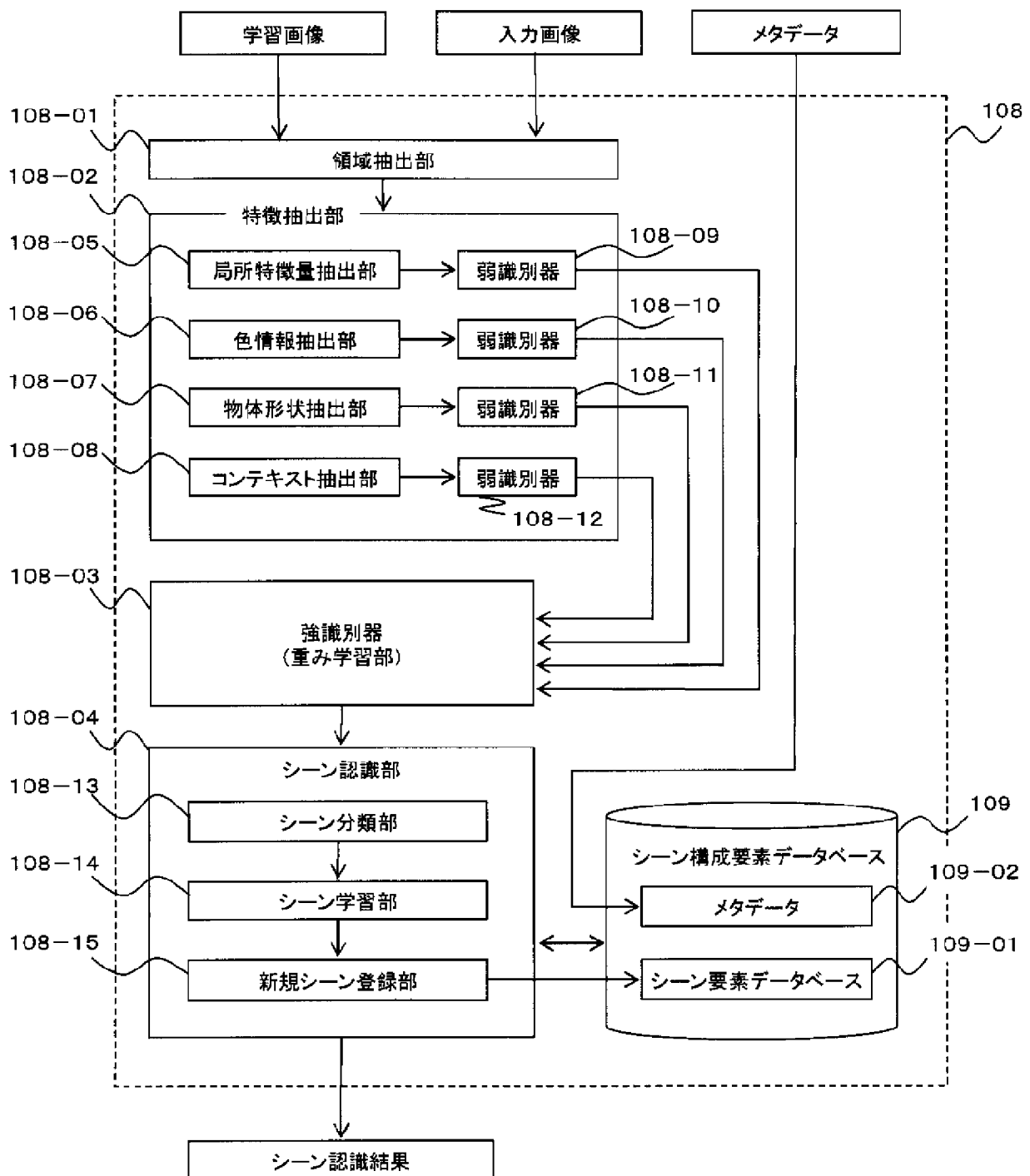
[図8]

## 一般物体認識システム



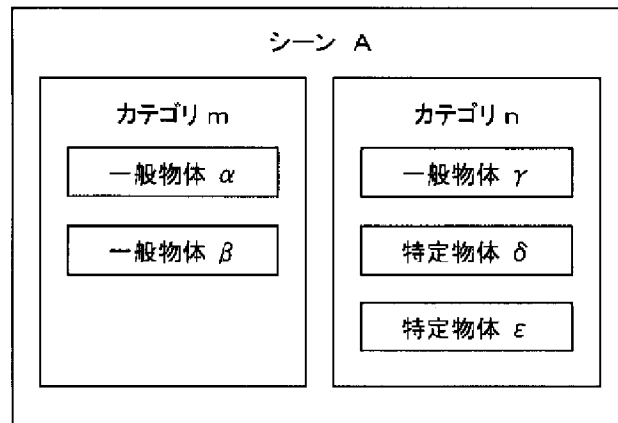
[図9]

## シーン認識システム



[図10]

(A)

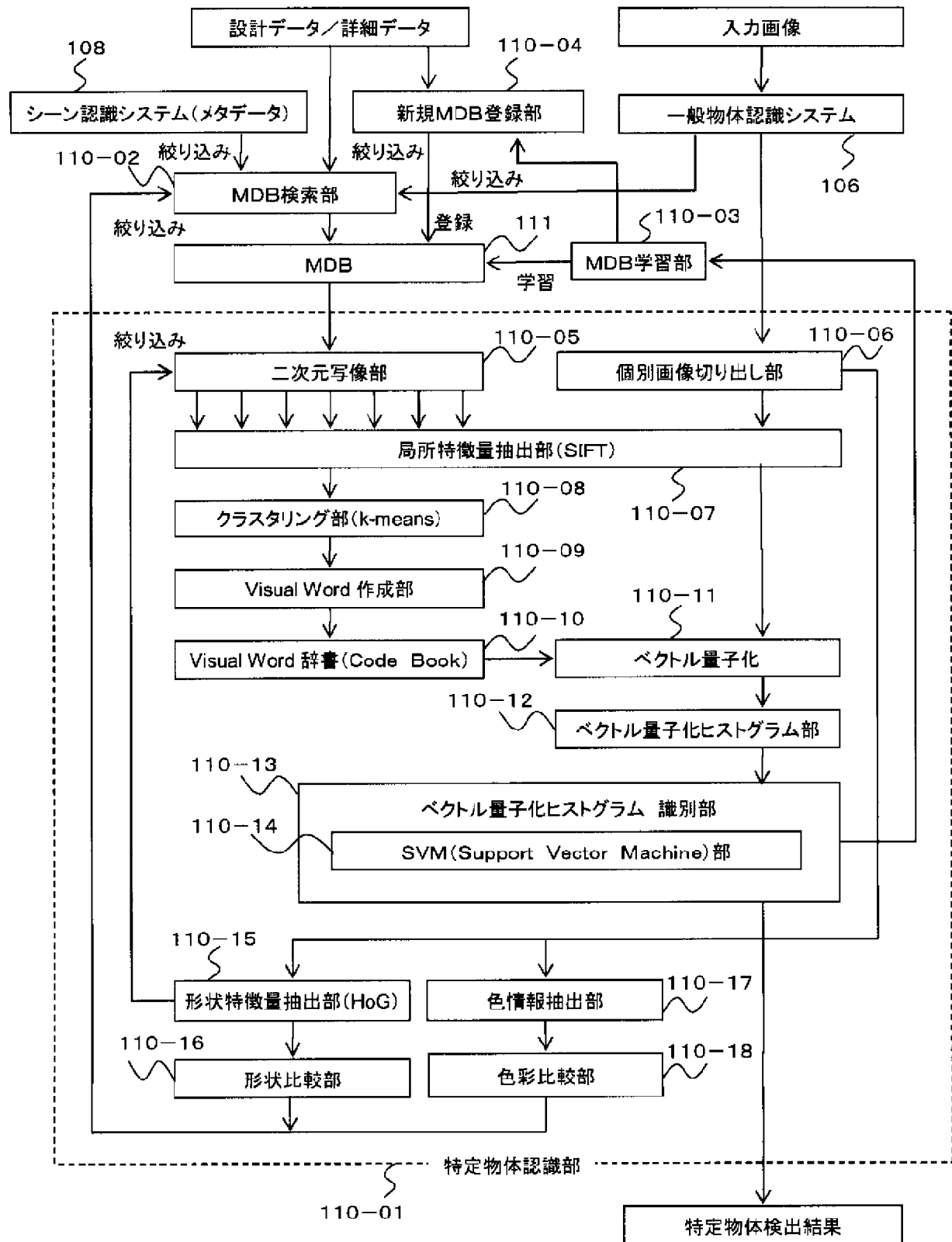


(B)



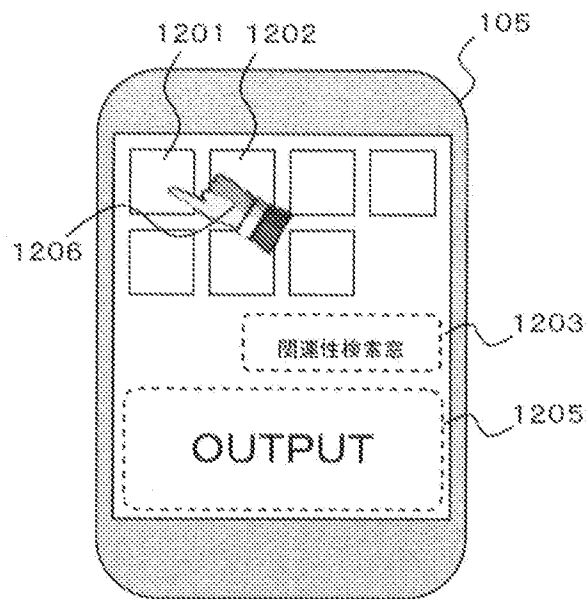
[図11]

## 特定物体認識システム

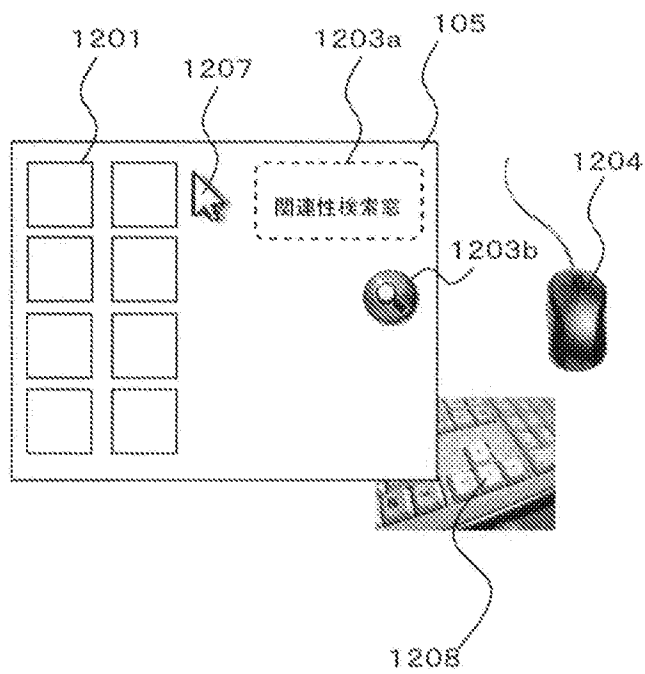


[図12A]

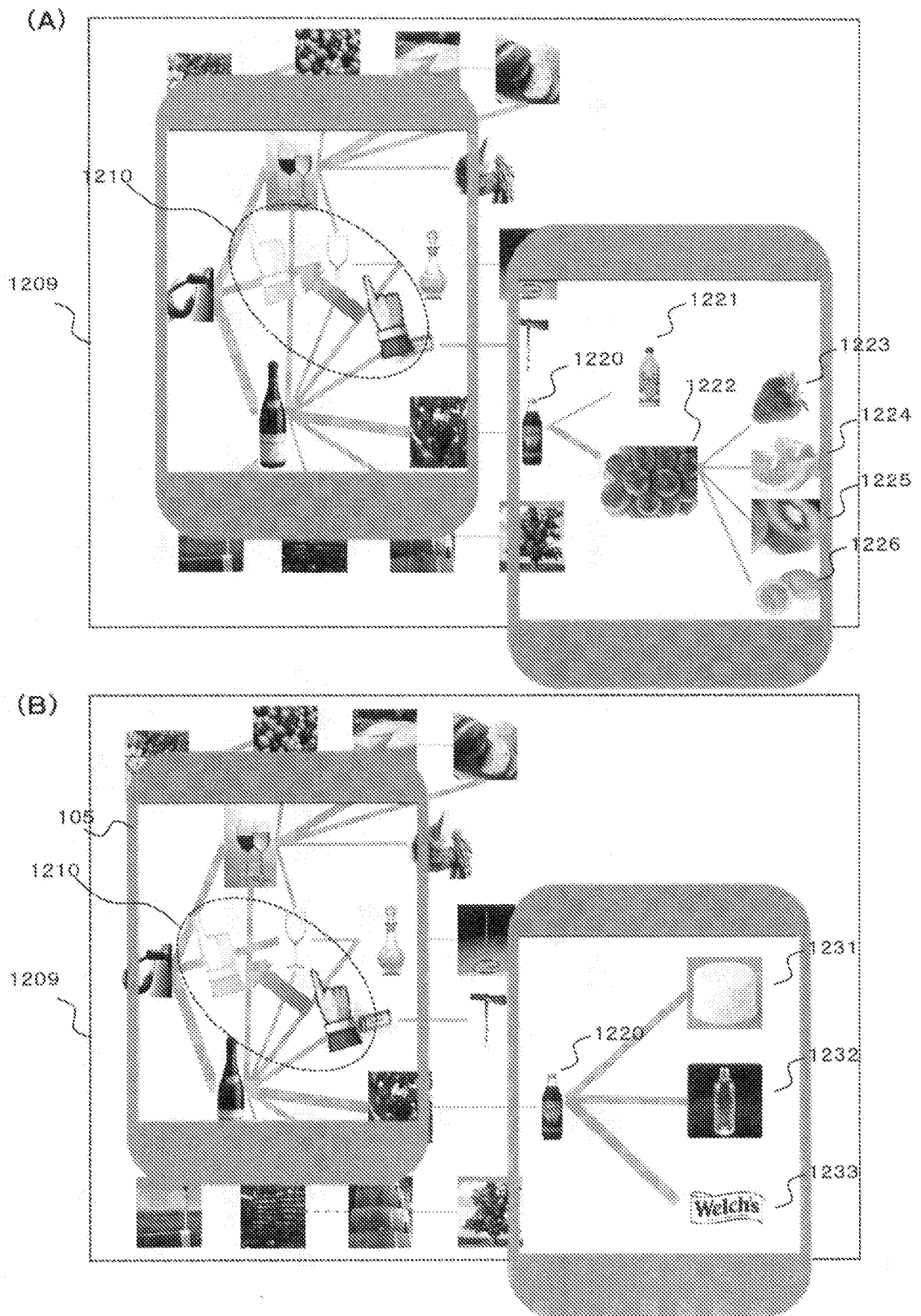
(A)



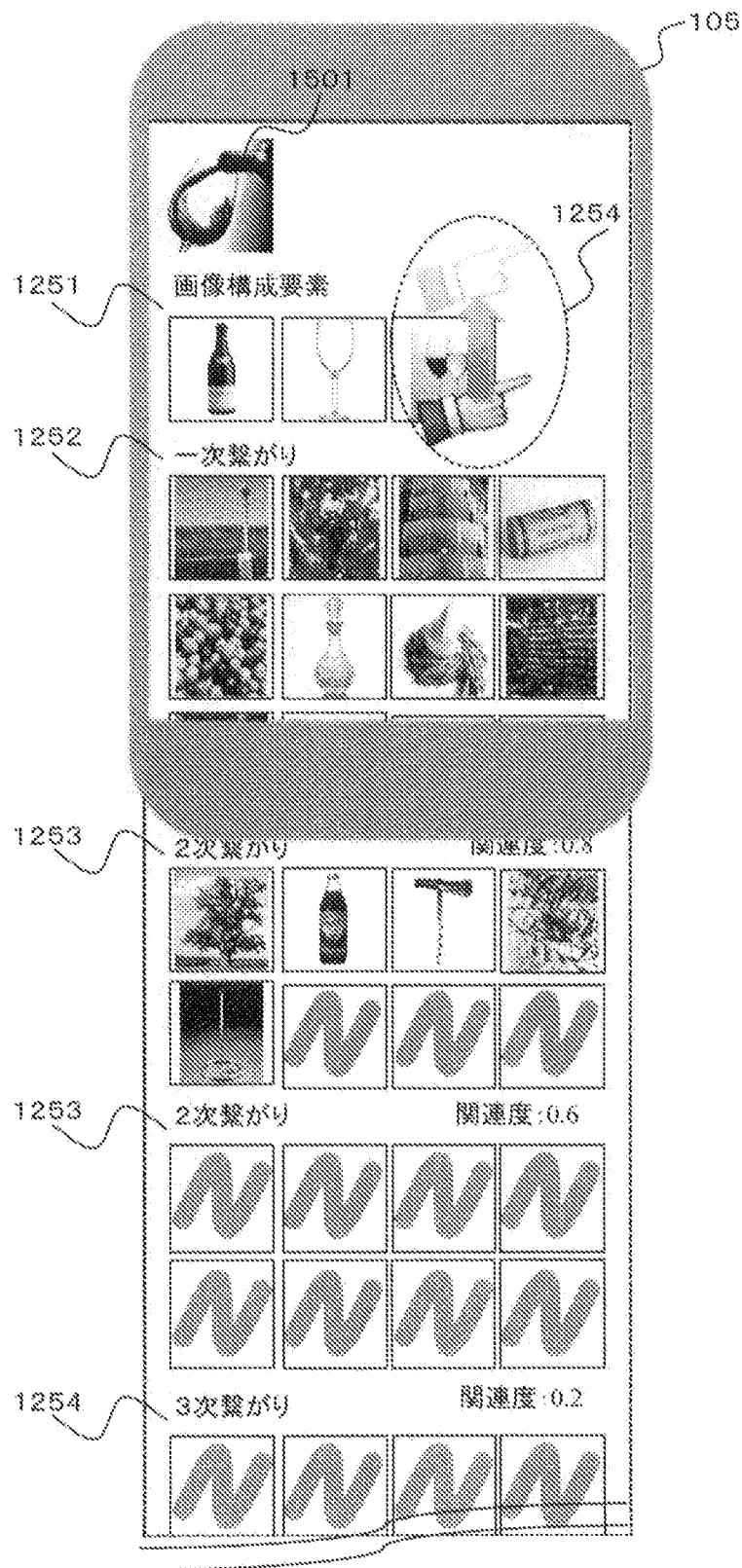
(B)



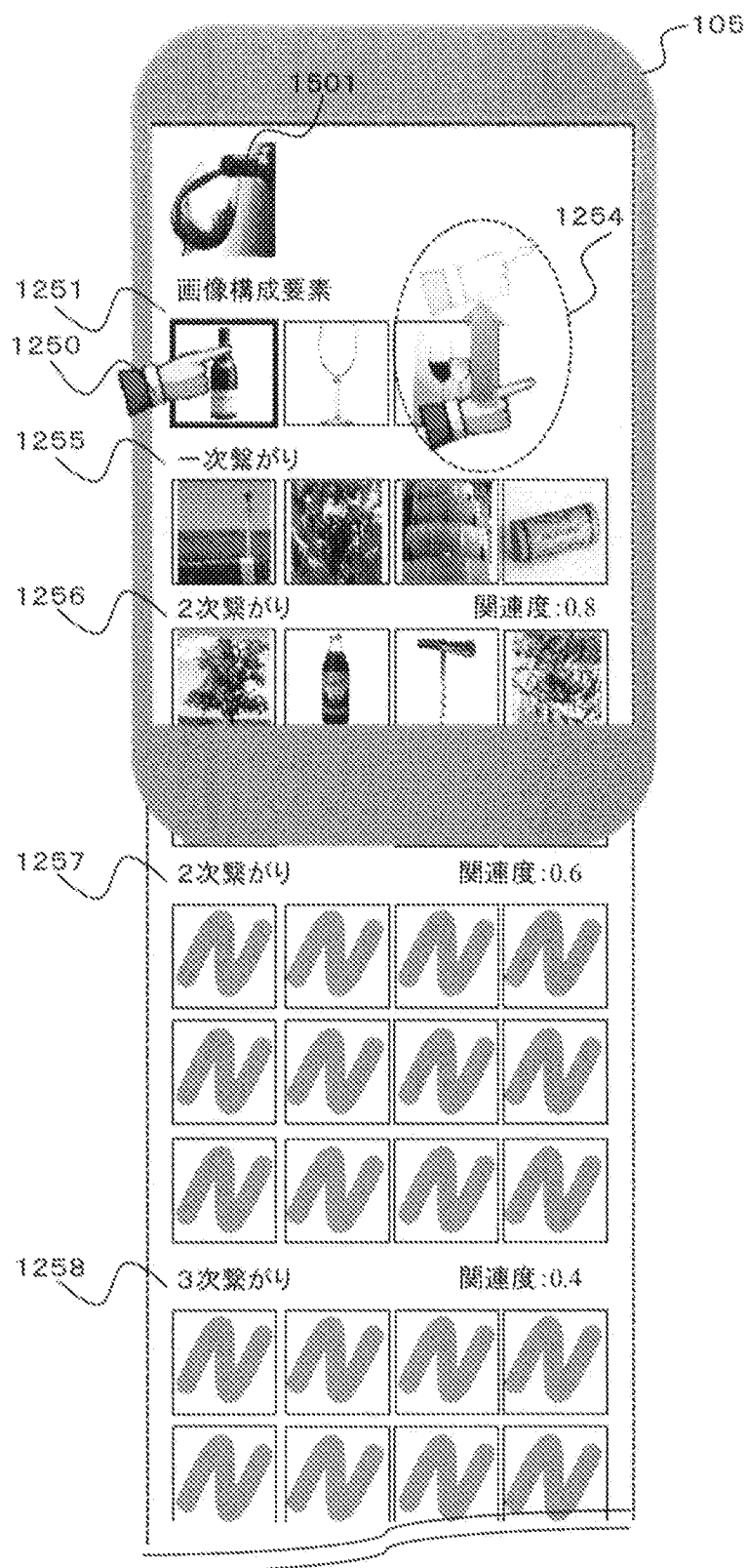
[図12B]



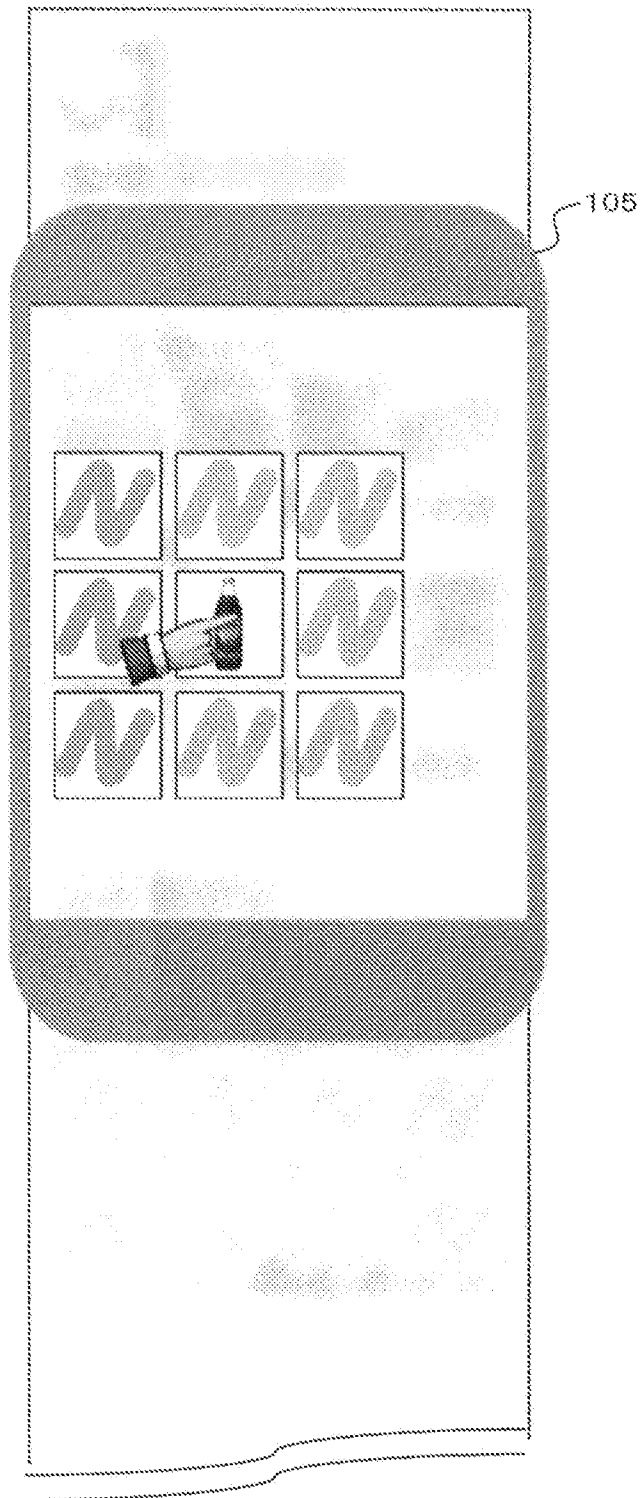
[図12C]



[図12D]

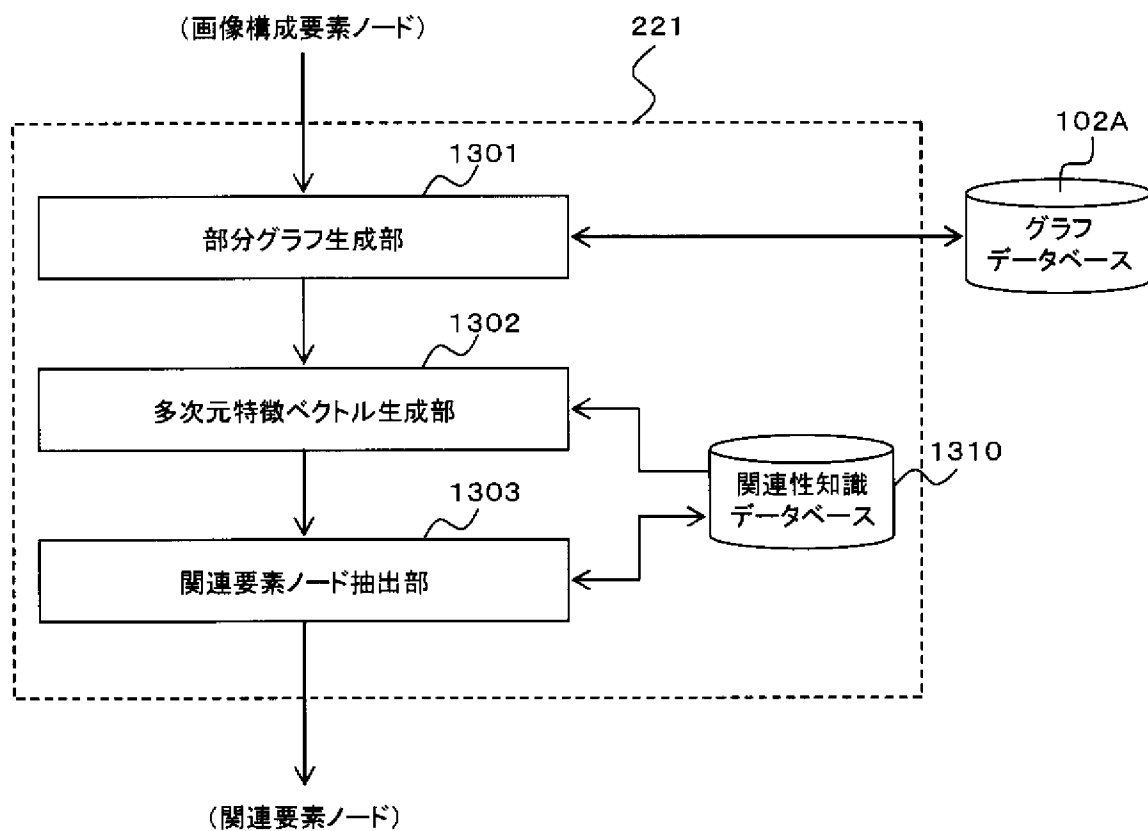


[図12E]



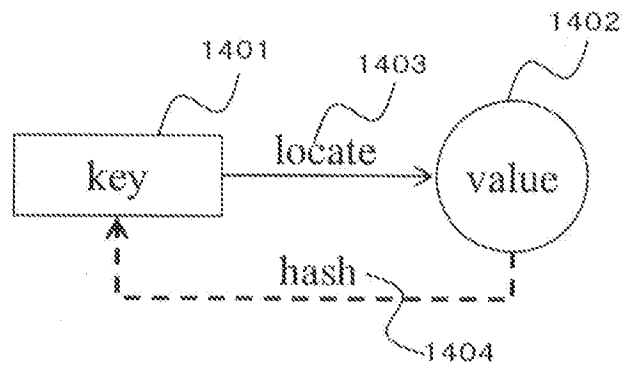
[図13]

## グラフ演算部221



[図14A]

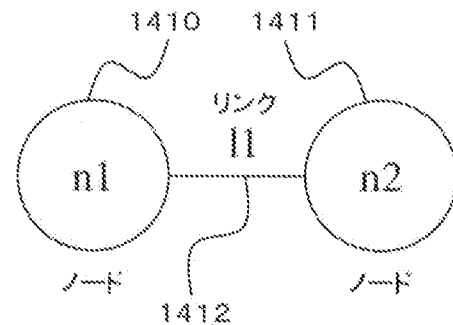
(A)



(B)

(key, {value})

(C)



(D)

| 項目          | 内容                                             |
|-------------|------------------------------------------------|
| タイプ         | ノードのタイプ<br>(USER OBJECT)(META SUBJECT URI EXT) |
| 生成時刻        | ノードが生成された時刻                                    |
| アクセス時刻      | ノードがアクセスされた時刻                                  |
| データ (value) | ノードの保持するデータ                                    |

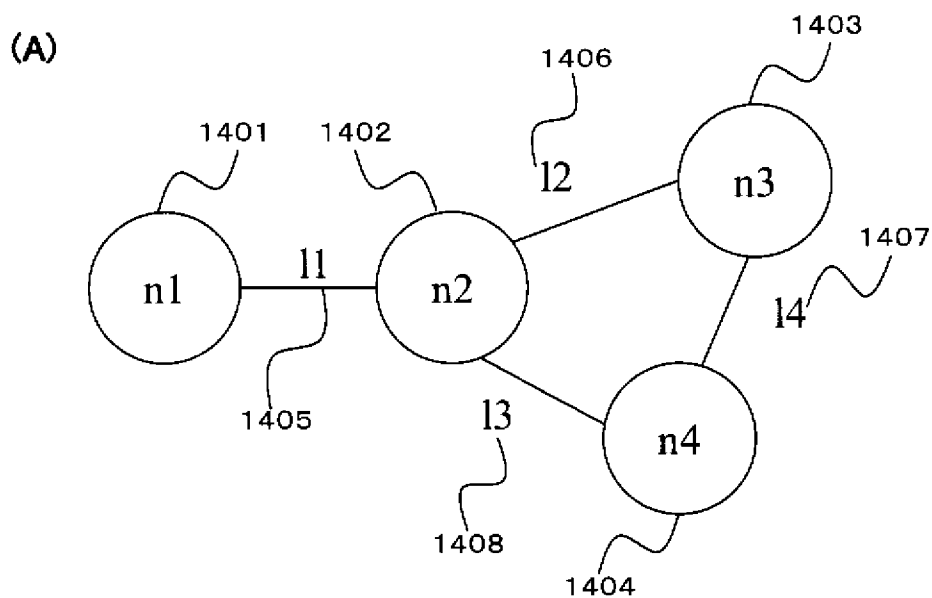
(E)

| 項目          | 内容                                    |
|-------------|---------------------------------------|
| タイプ         | リンクのタイプ<br>(UNDIRECTED DIRECTED NONE) |
| 生成時刻        | リンクが生成された時刻                           |
| アクセス時刻      | リンクがアクセスされた時刻                         |
| データ (value) | 左ノード (key)                            |
|             | 右ノード (key)                            |
|             | 重み (w)                                |
|             | 関数 (f)                                |

[図14B]

| GDB OP.名 | オペレーション                        |
|----------|--------------------------------|
| CREATE   | 指定したタイプのノードを生成する。              |
| CONNECT  | 指定したノードとノード間に指定したタイプのリンクを生成する。 |
| NODE     | keyに対応するノードデータを取得する。           |
| LINK     | keyに対応するリンクデータを取得する。           |
| SUBGRAPH | 指定したノードの部分グラフを取得する。            |

[図14C]



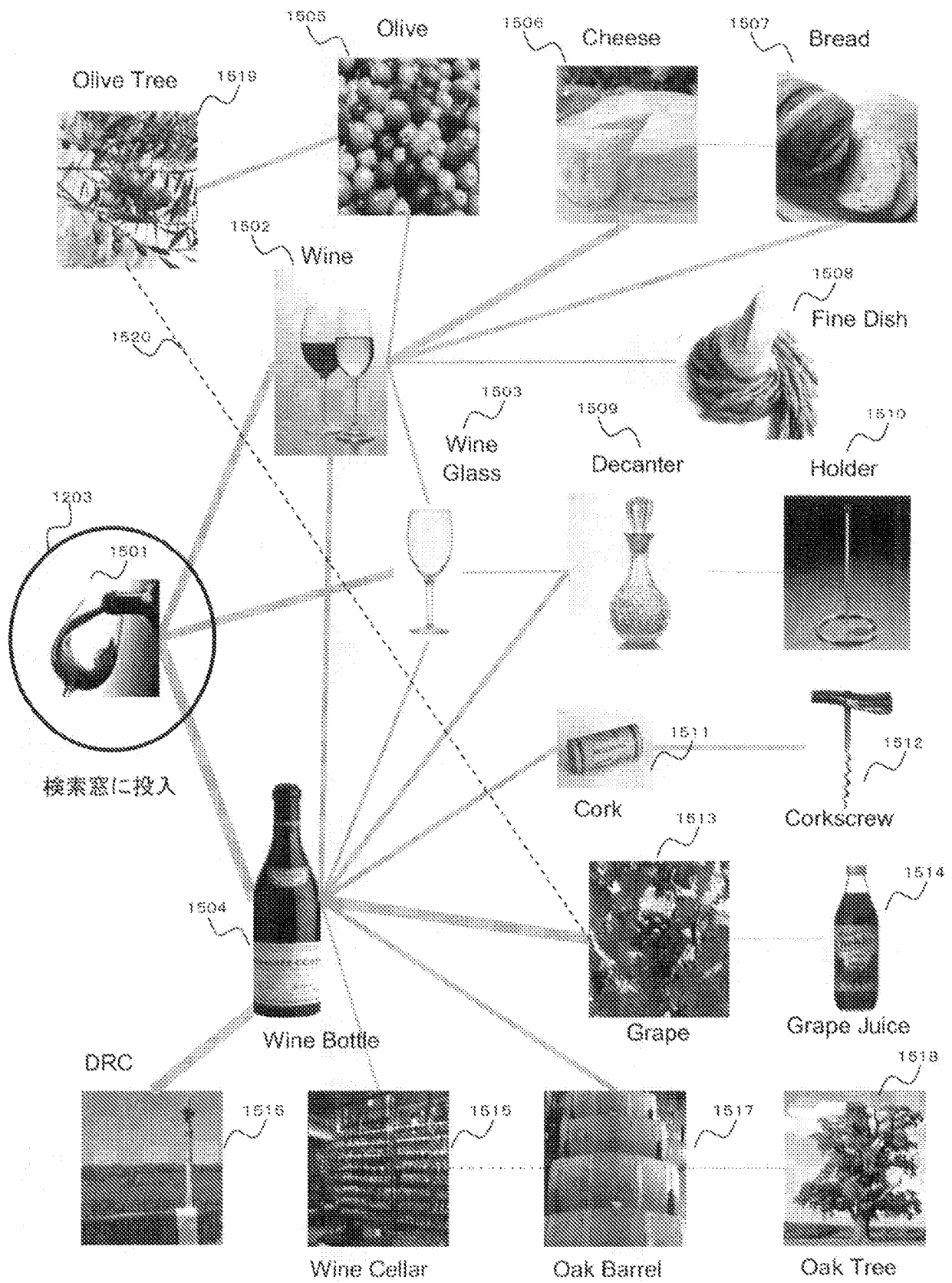
(B)

$(n1, \{\text{node } n1\})$      $(l1, \{n1, n2\})$   
 $(n2, \{\text{node } n2\})$      $(l2, \{n2, n3\})$   
 $(n3, \{\text{node } n3\})$      $(l3, \{n3, n4\})$   
 $(n4, \{\text{node } n4\})$      $(l4, \{n2, n4\})$

(C)

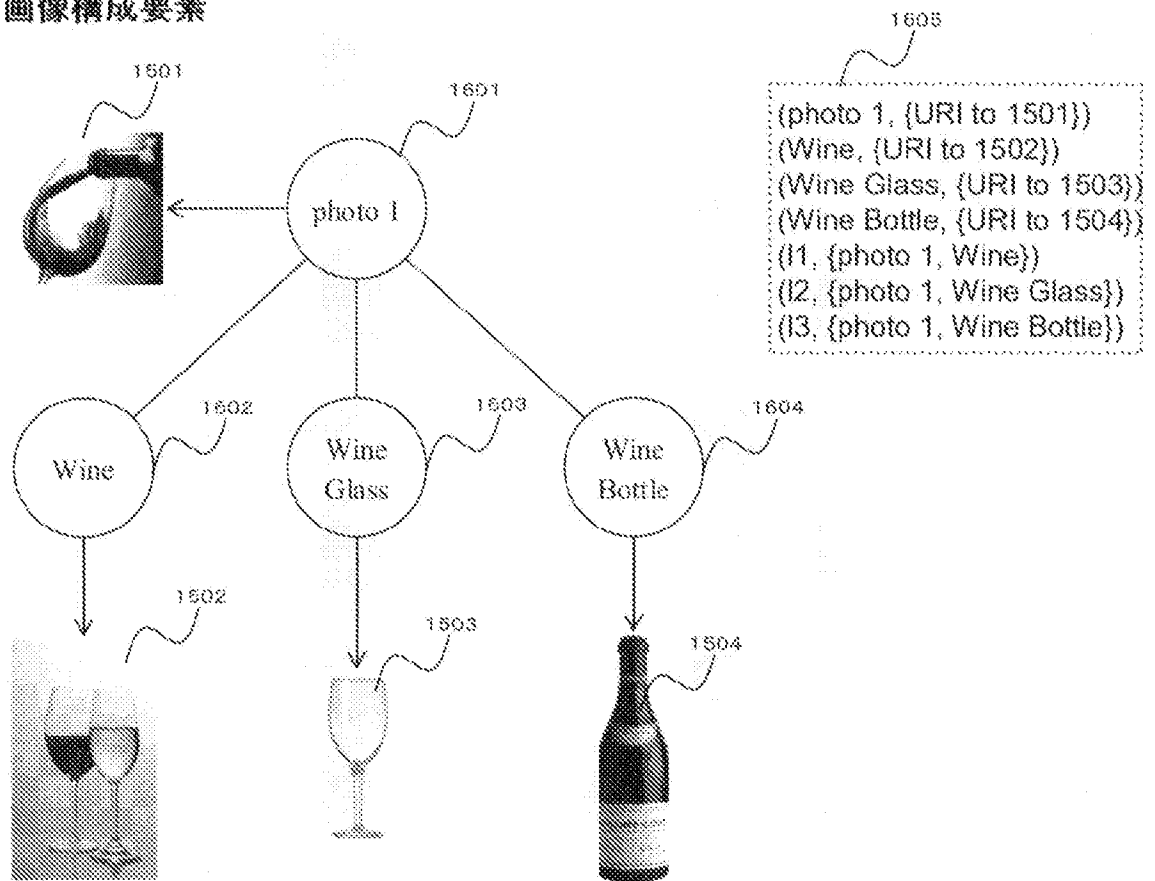
$(n1, \{f(n1, n2)\})$   
 $(n2, \{f(n2, n3)\})$   
 $(n3, \{f(n3, n4)\})$   
 $(n4, \{f(n2, n4)\})$

[図15]

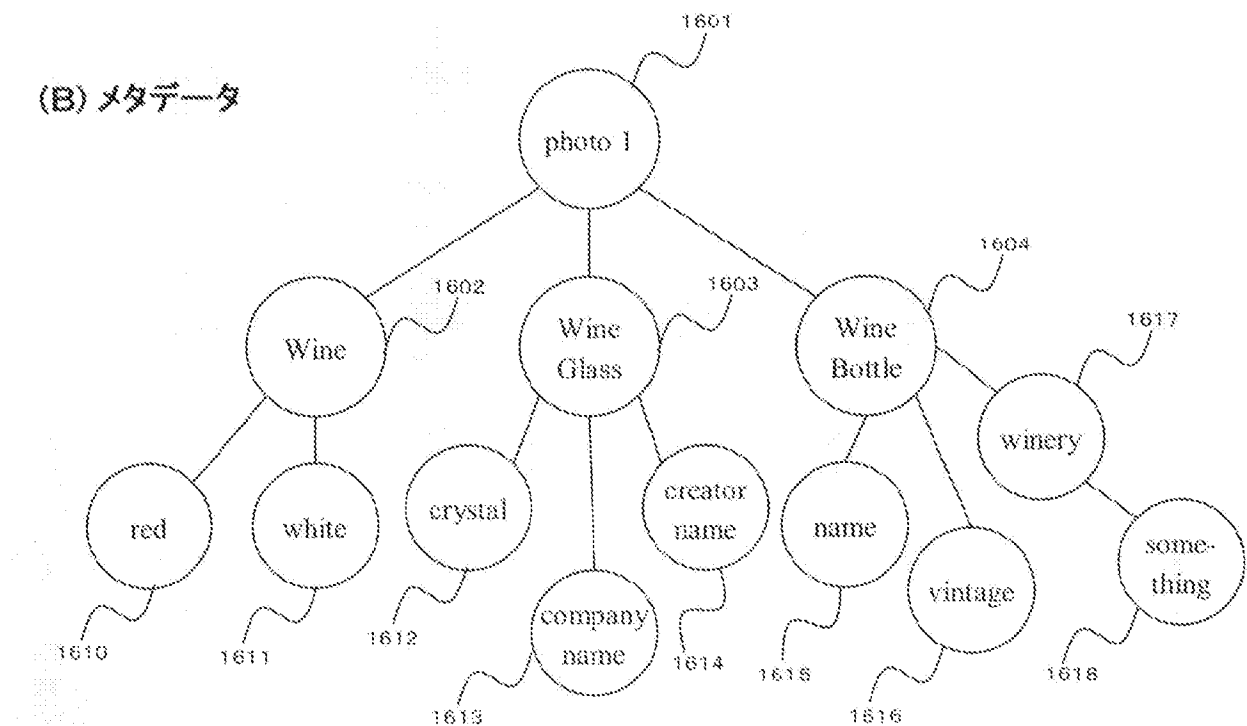


[図16]

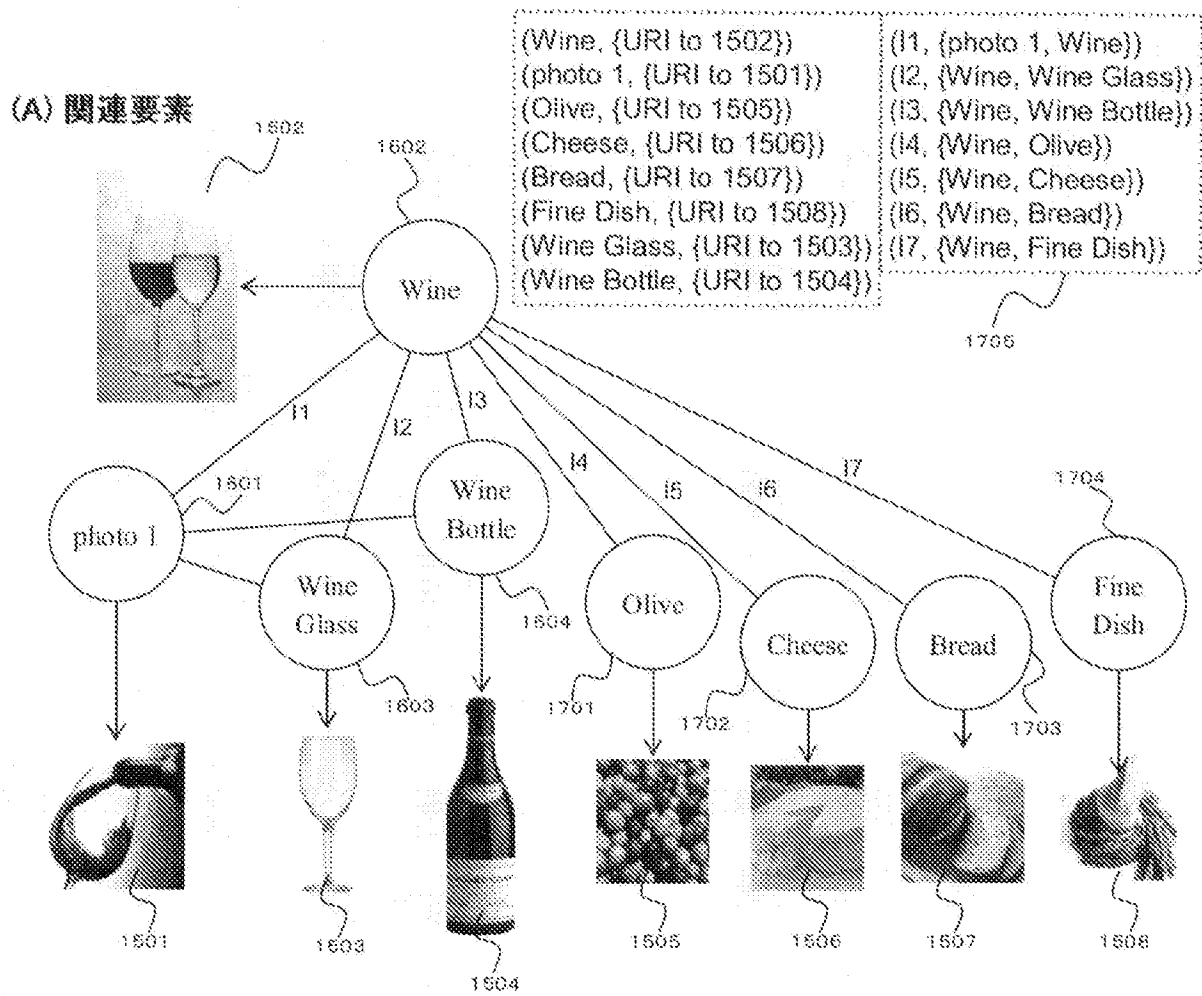
## (A) 画像構成要素



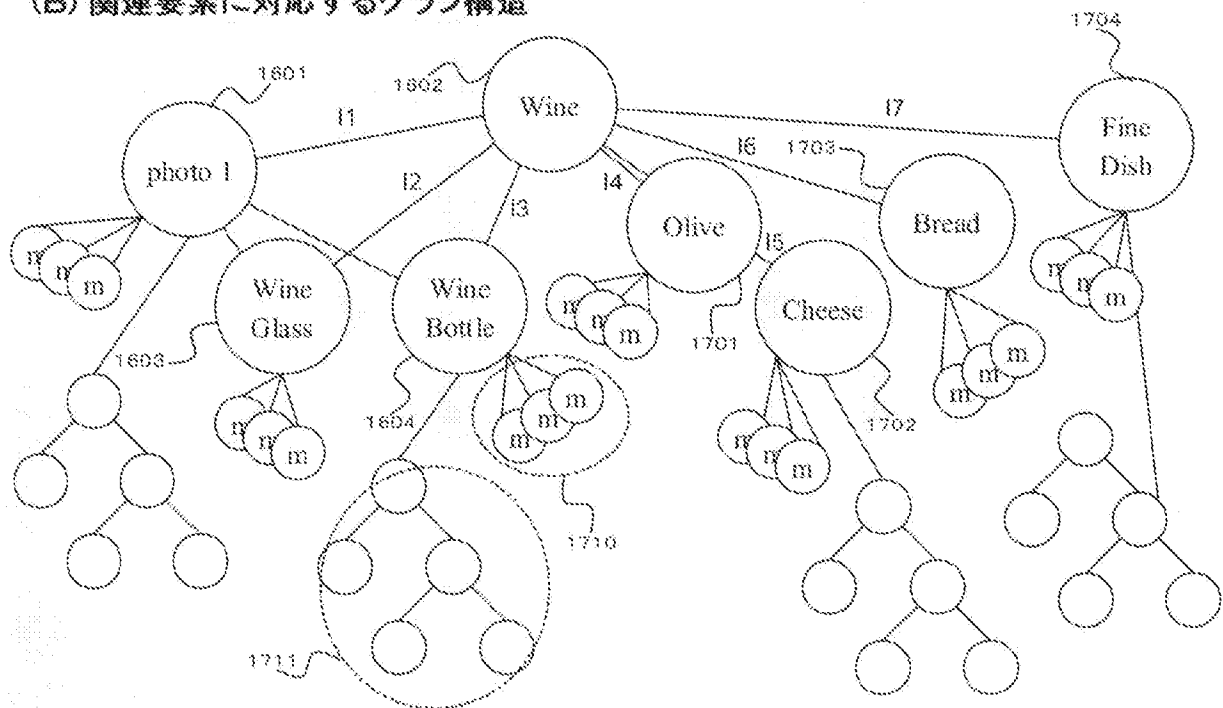
## (B) メタデータ



[図17]

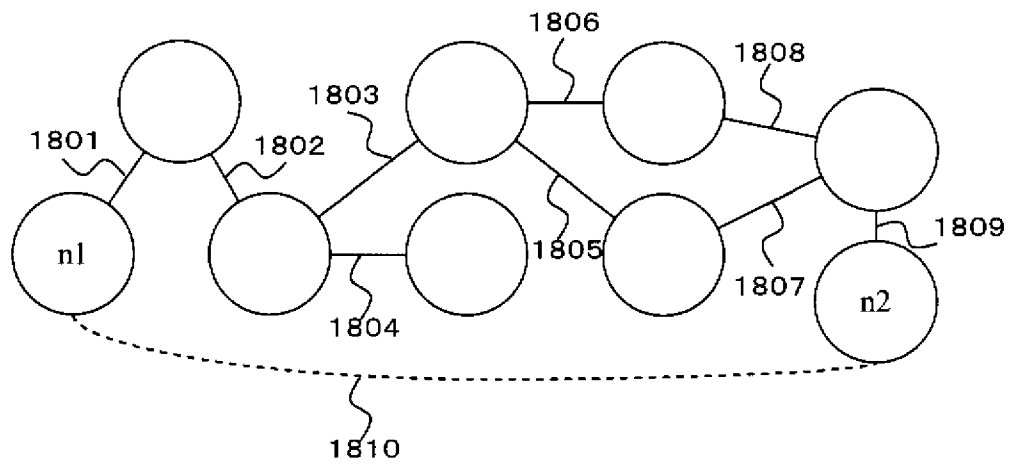


(B) 関連要素に対応するグラフ構造



[図18]

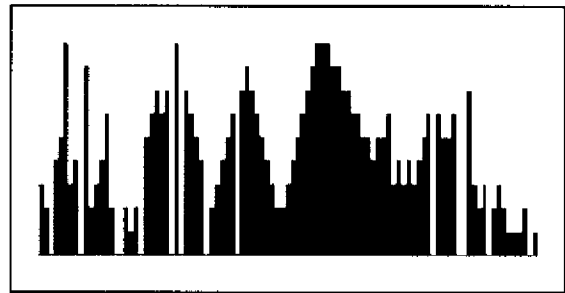
(A)



(B)

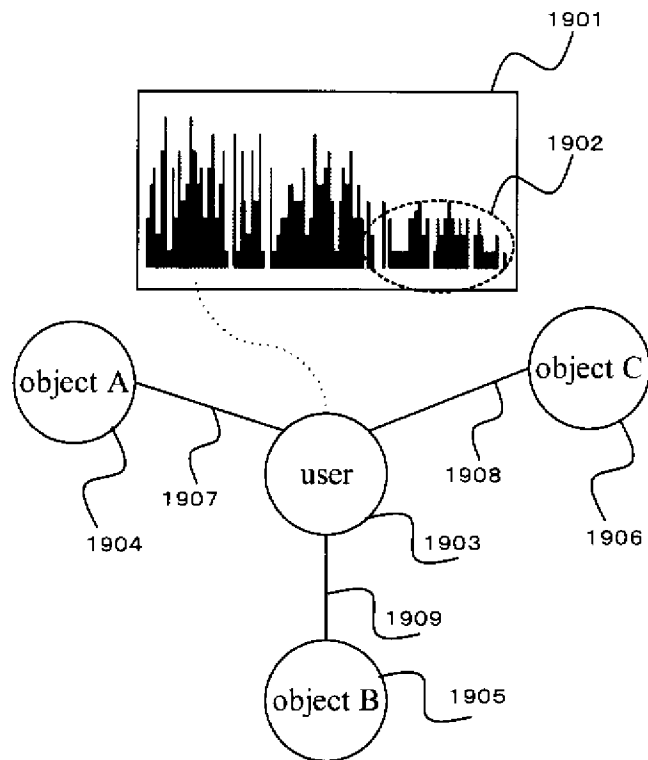
$$v1 = \left\{ \begin{array}{l} f_{1801} \\ f_{1802} \\ f_{1803} \\ f_{1804} \\ f_{1805} \\ f_{1806} \\ f_{1807} \\ f_{1808} \\ f_{1809} \end{array} \right\}$$

(C)

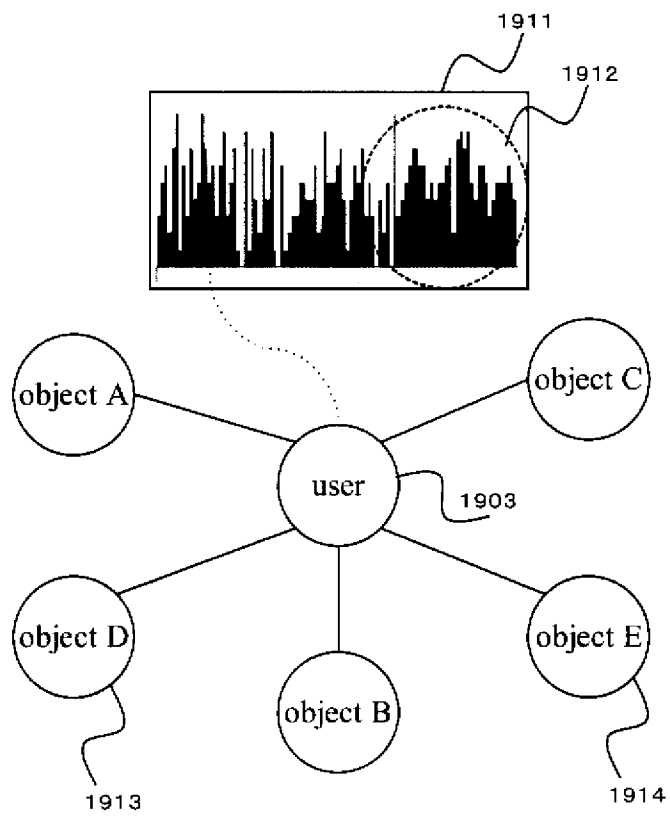


[図19]

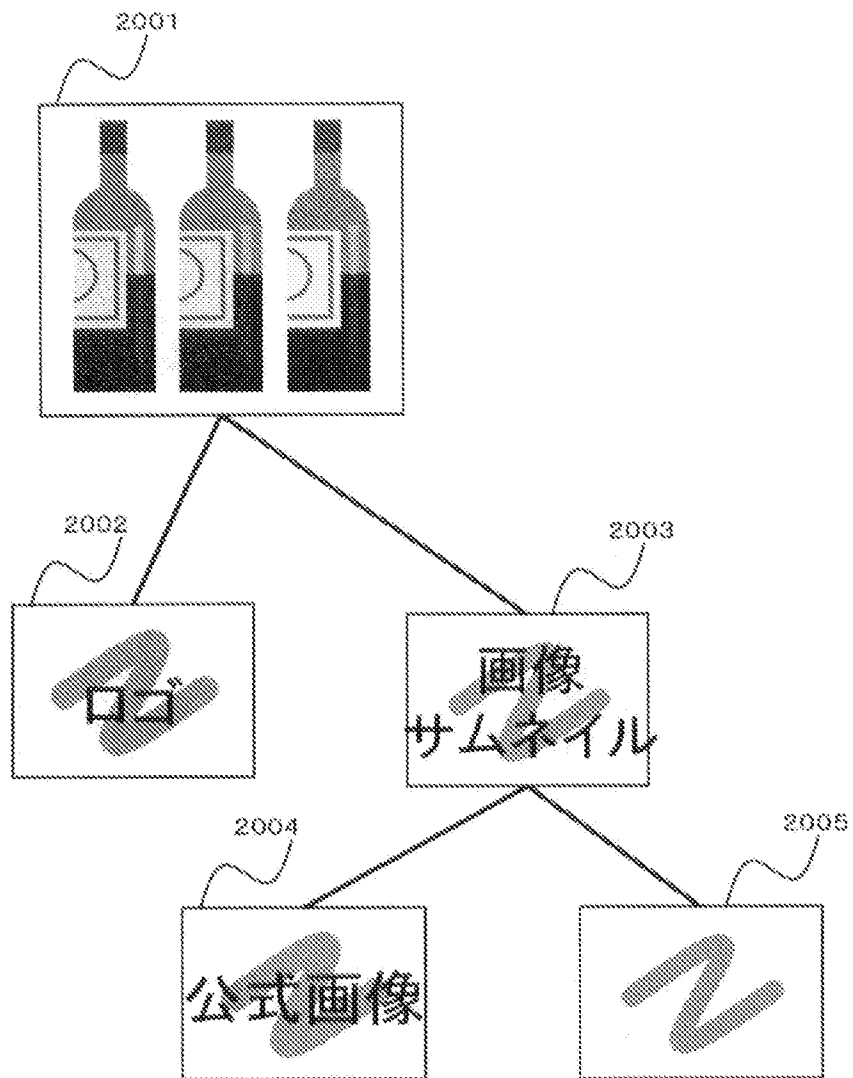
(A)



(B)

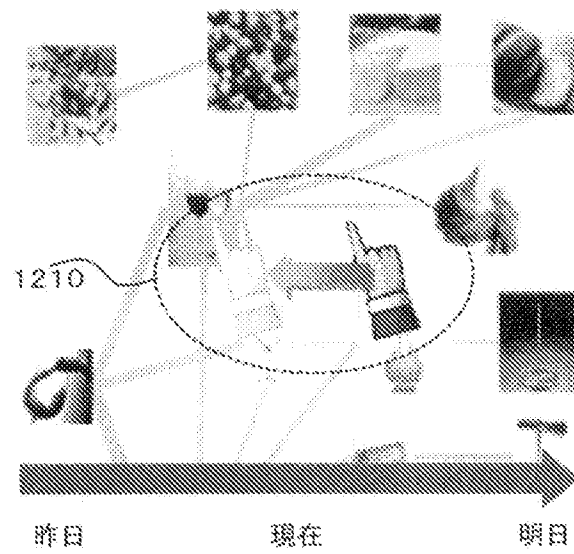


[図20A]

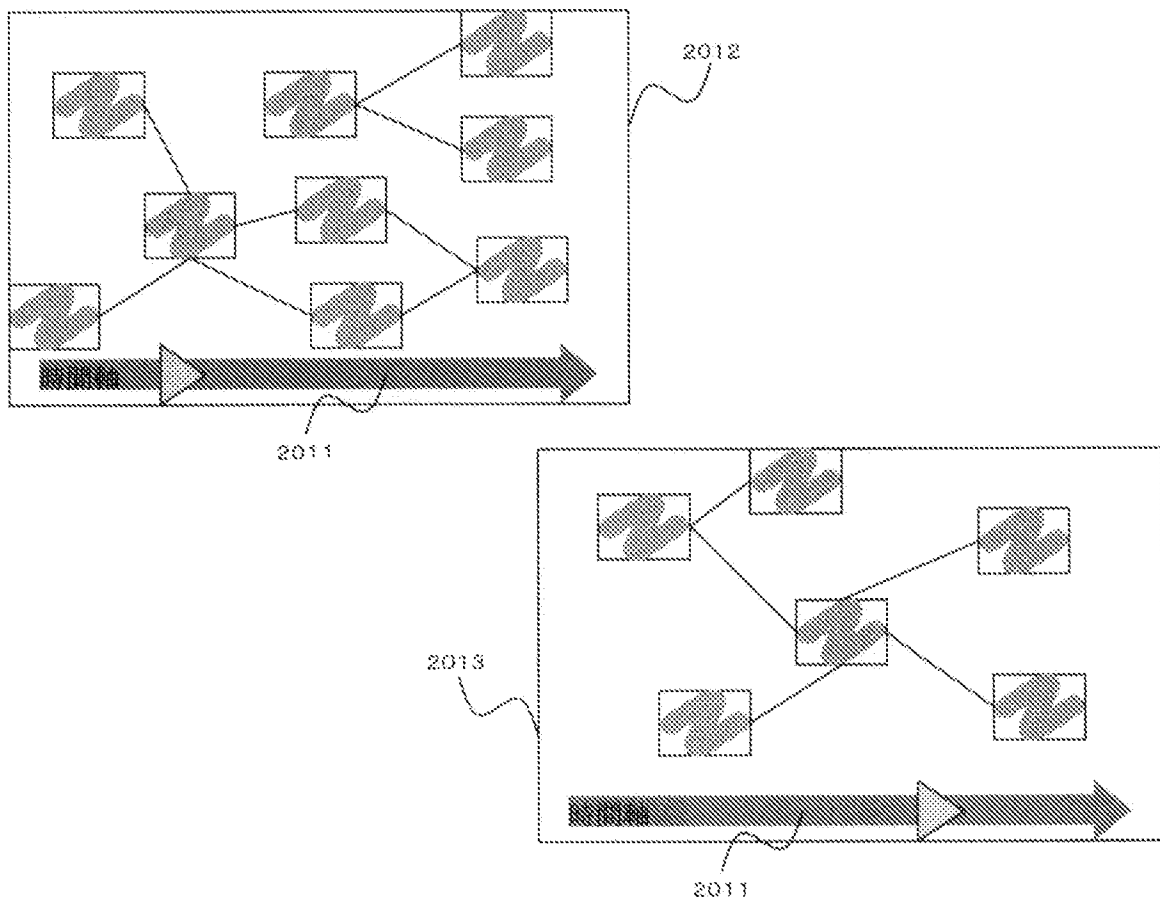


[図20B]

## (A) 時間軸

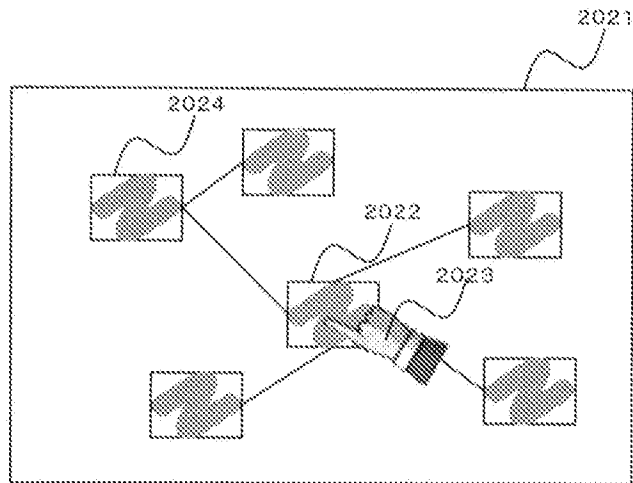


## (B) 当該時刻表示

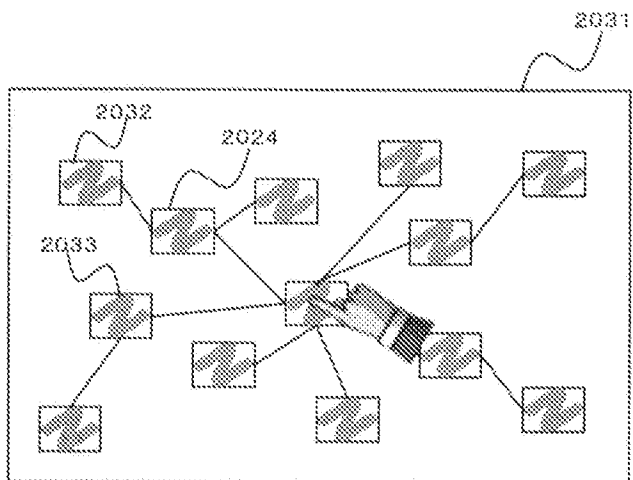


[図20C]

(A)

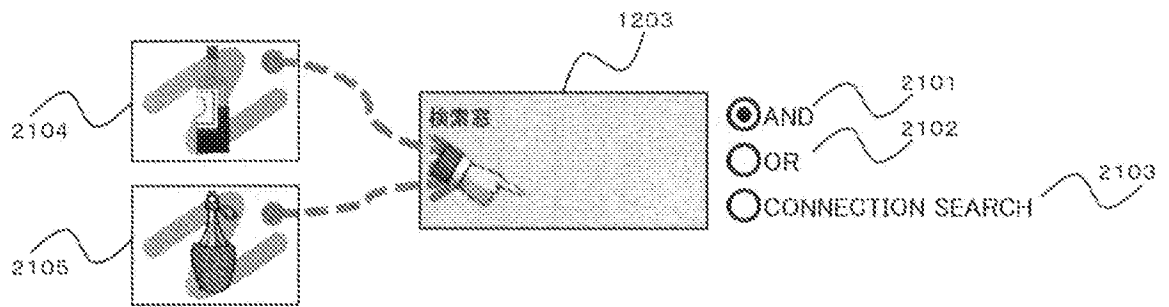


(B)

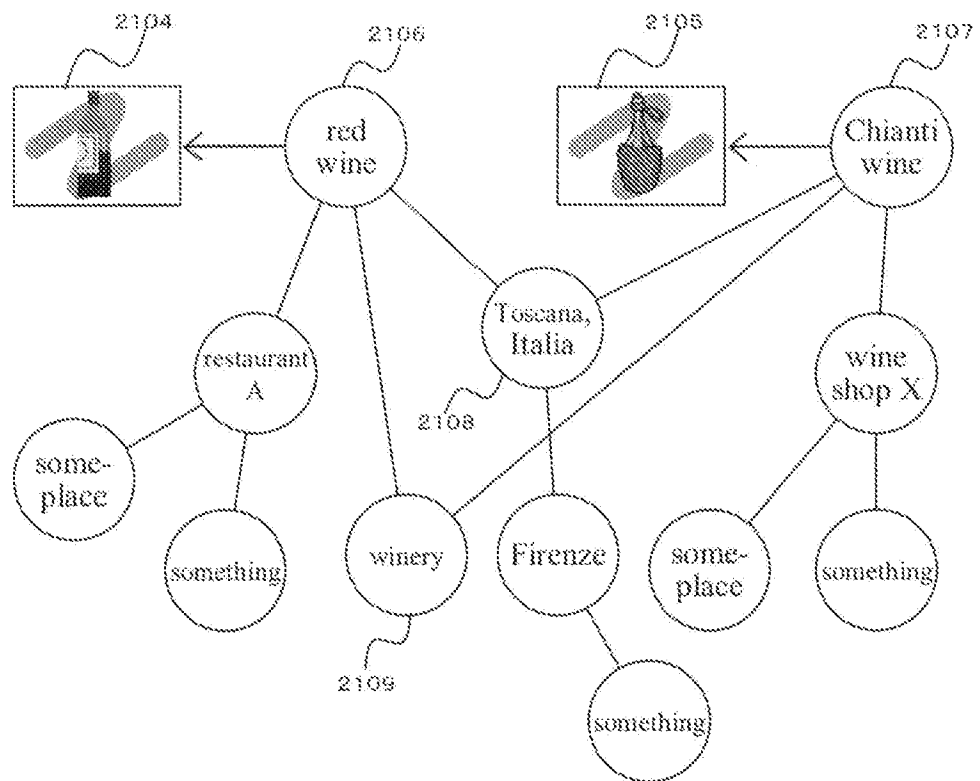


[図21]

(A)

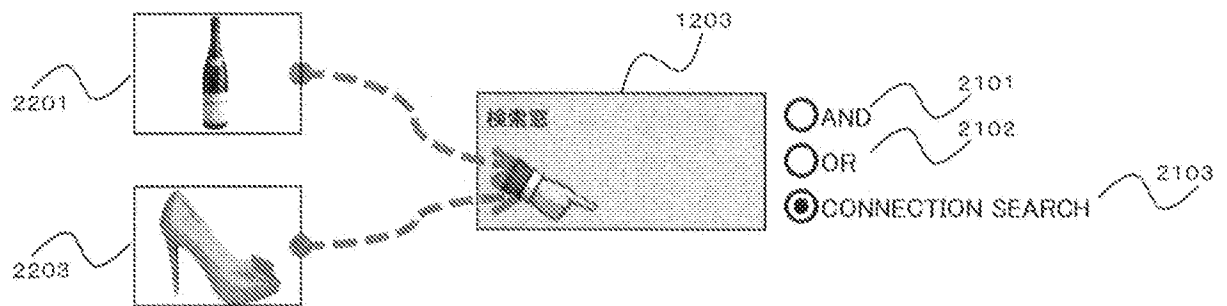


(B)

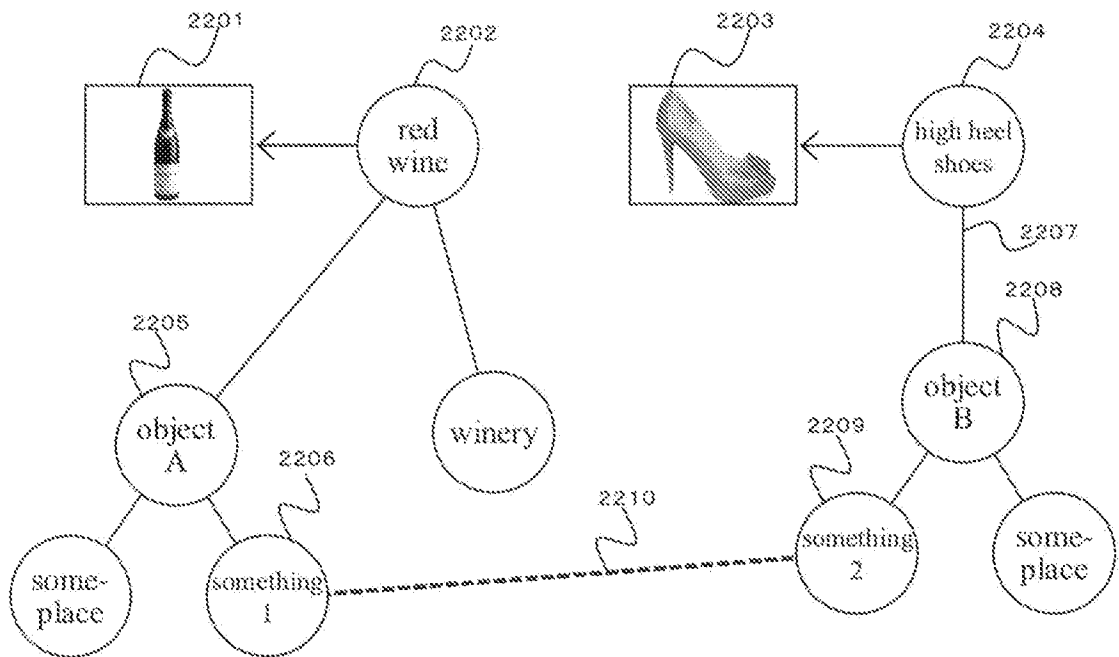


[図22]

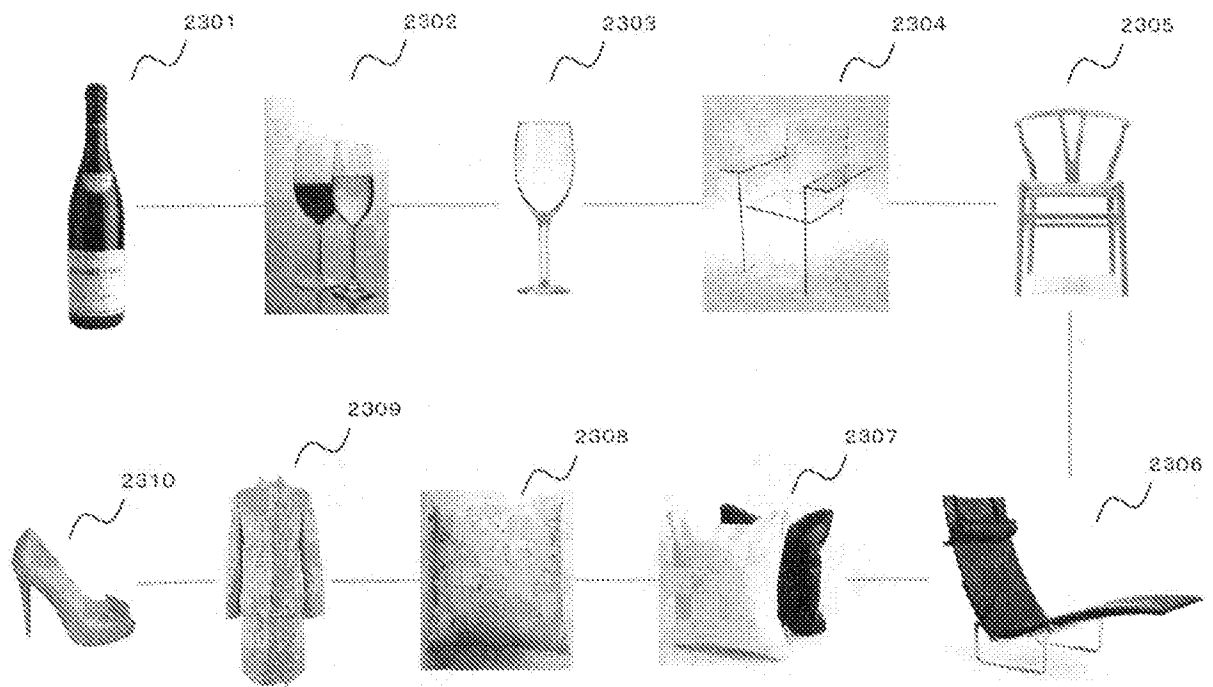
(A)

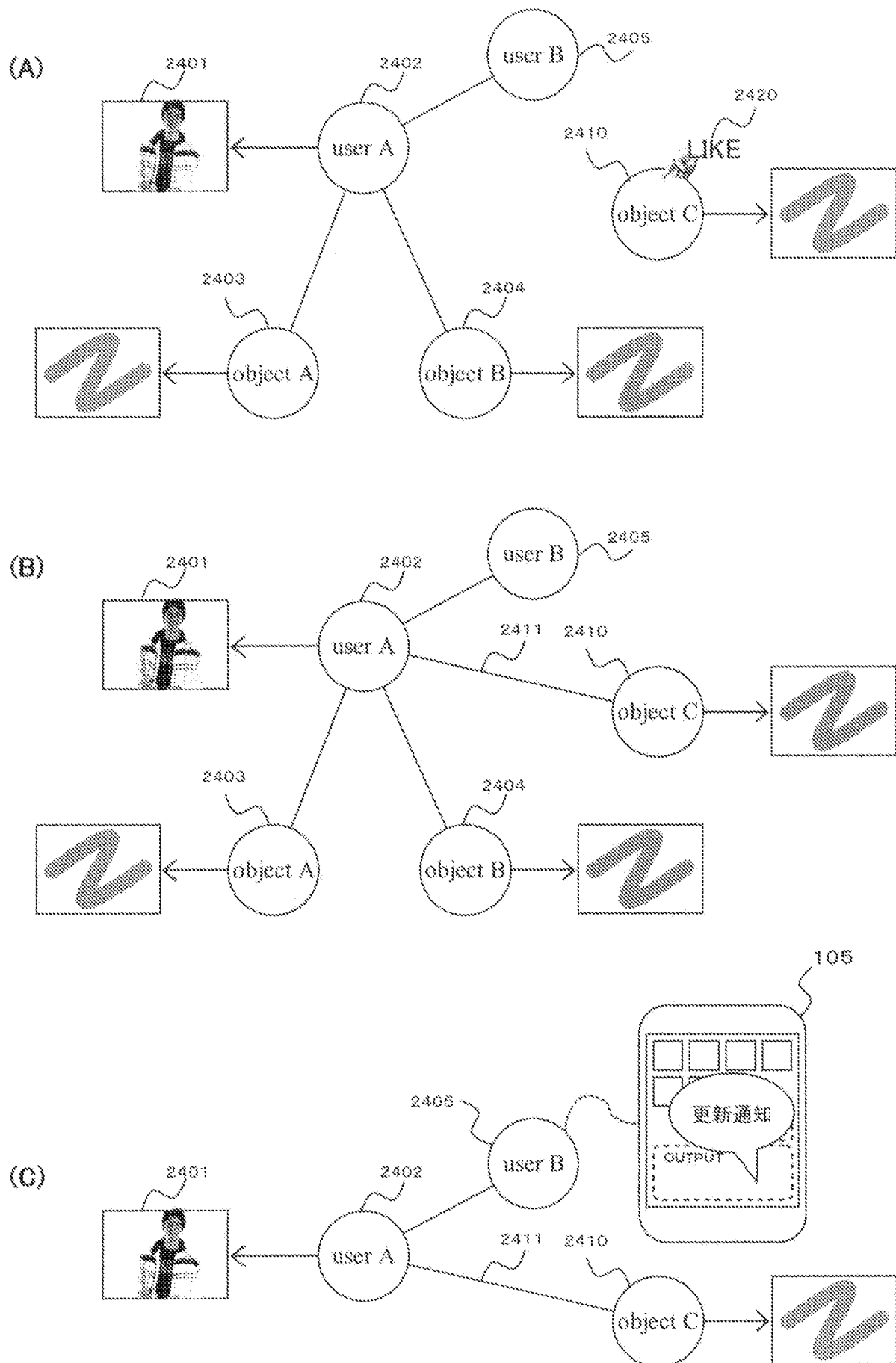


(B)

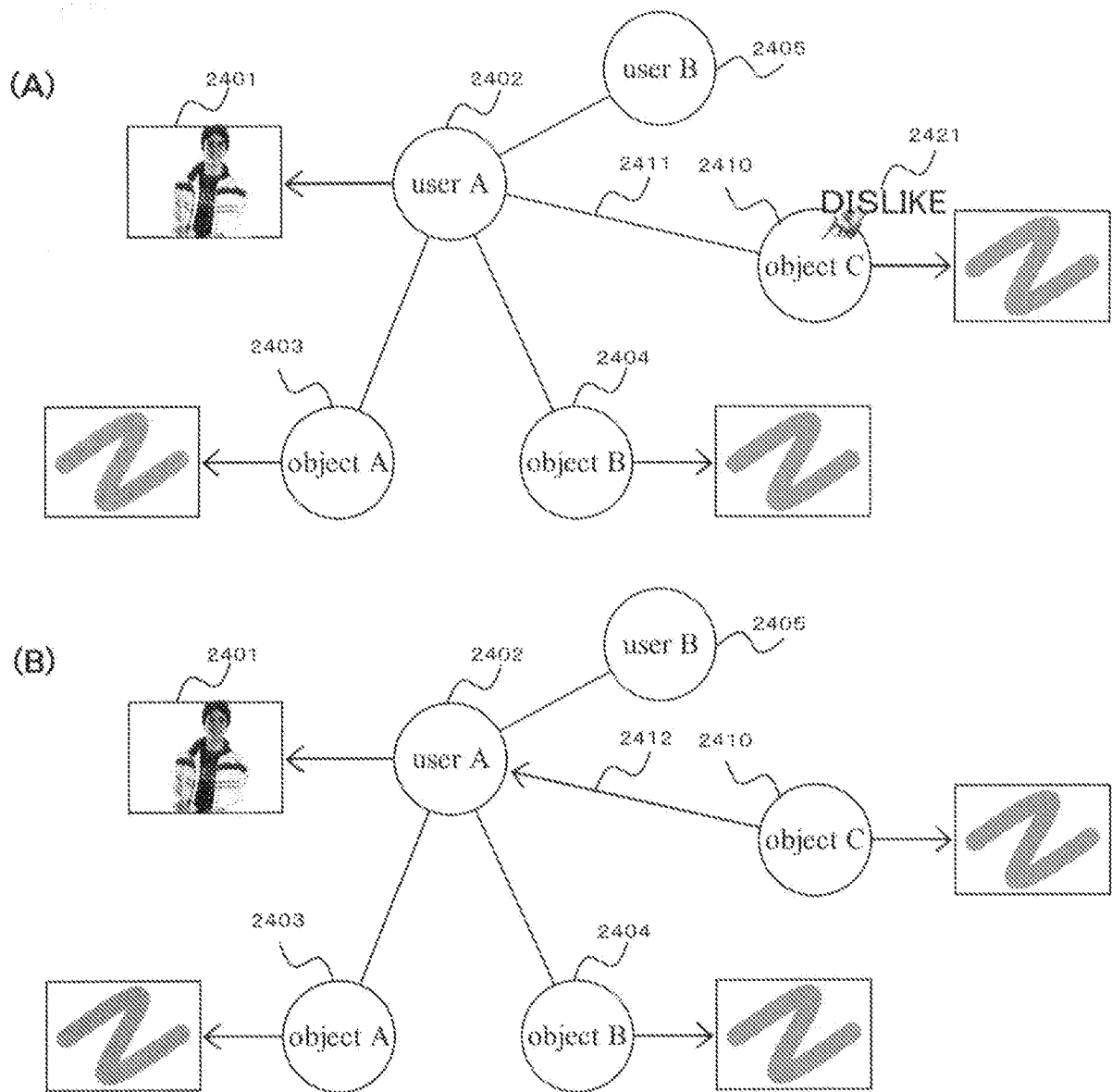


[図23]

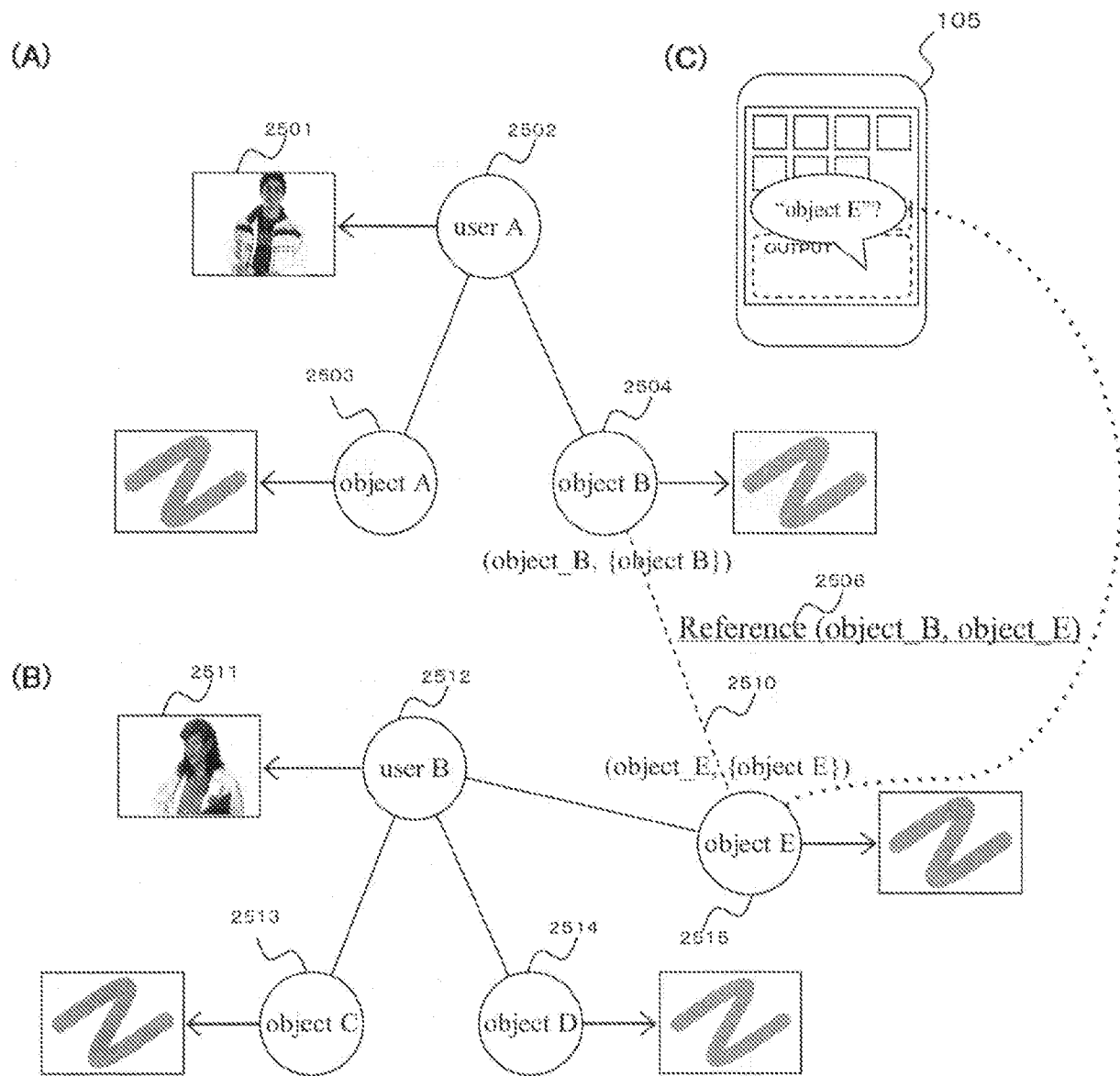




[図24B]

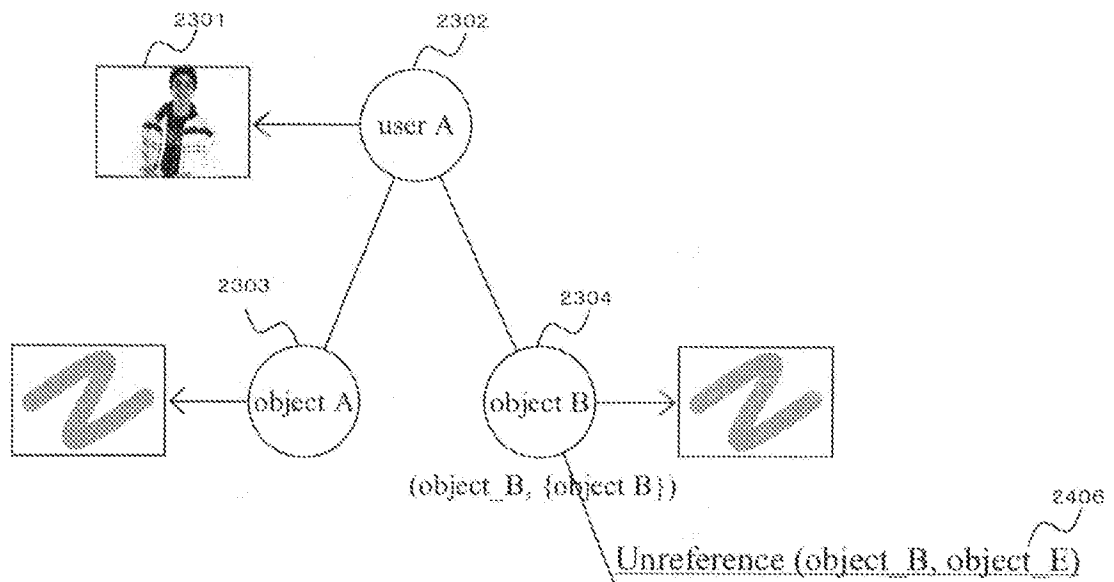


[図25]

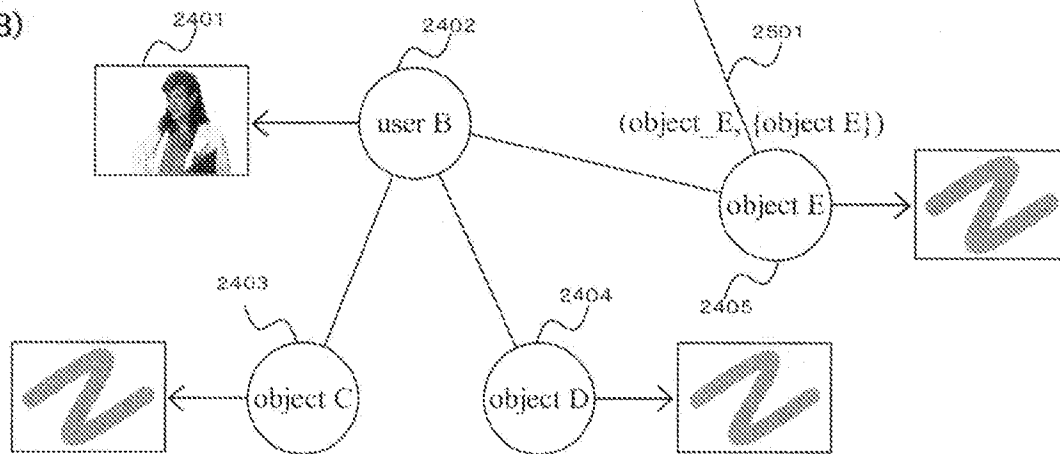


[図26]

(A)

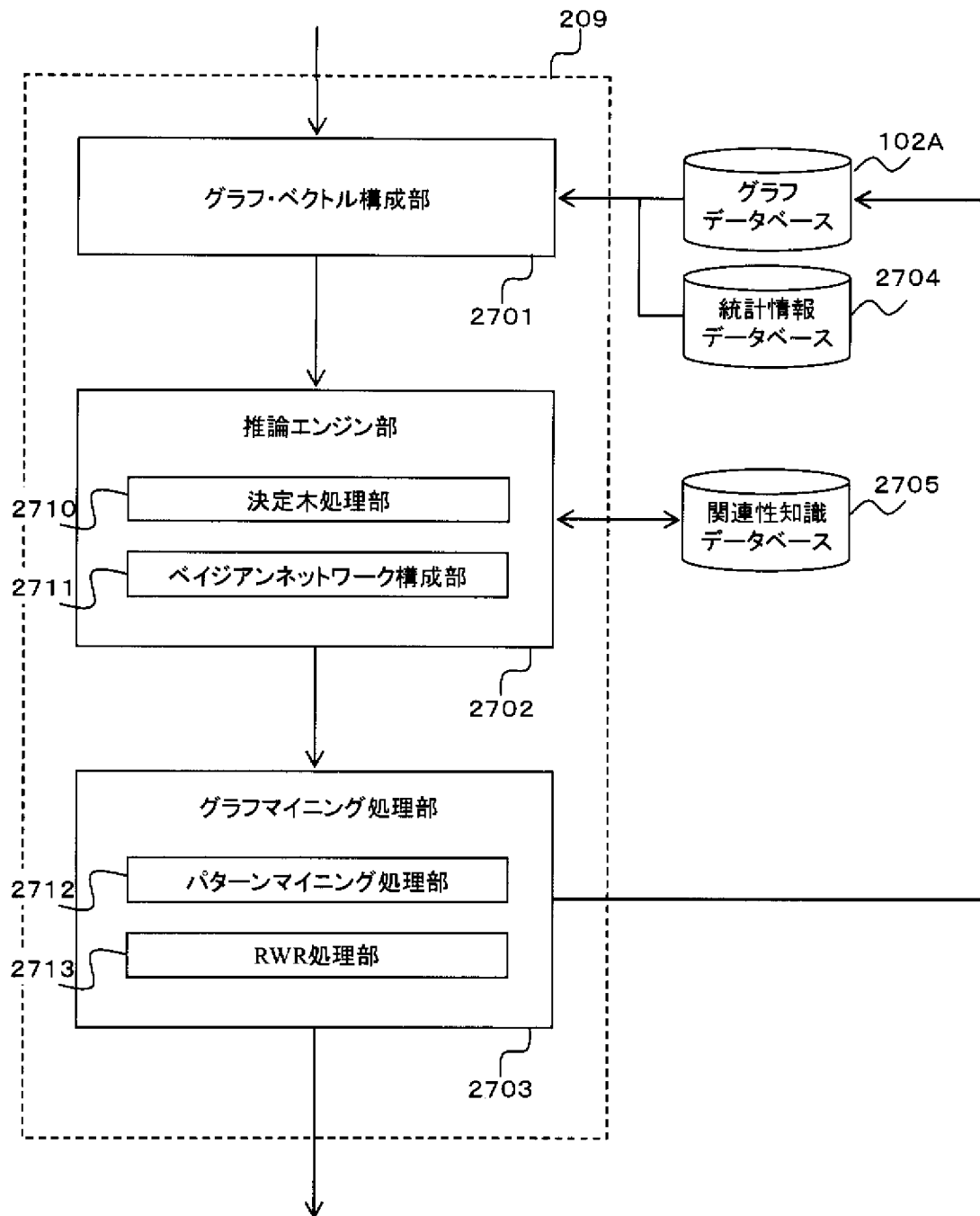


(B)

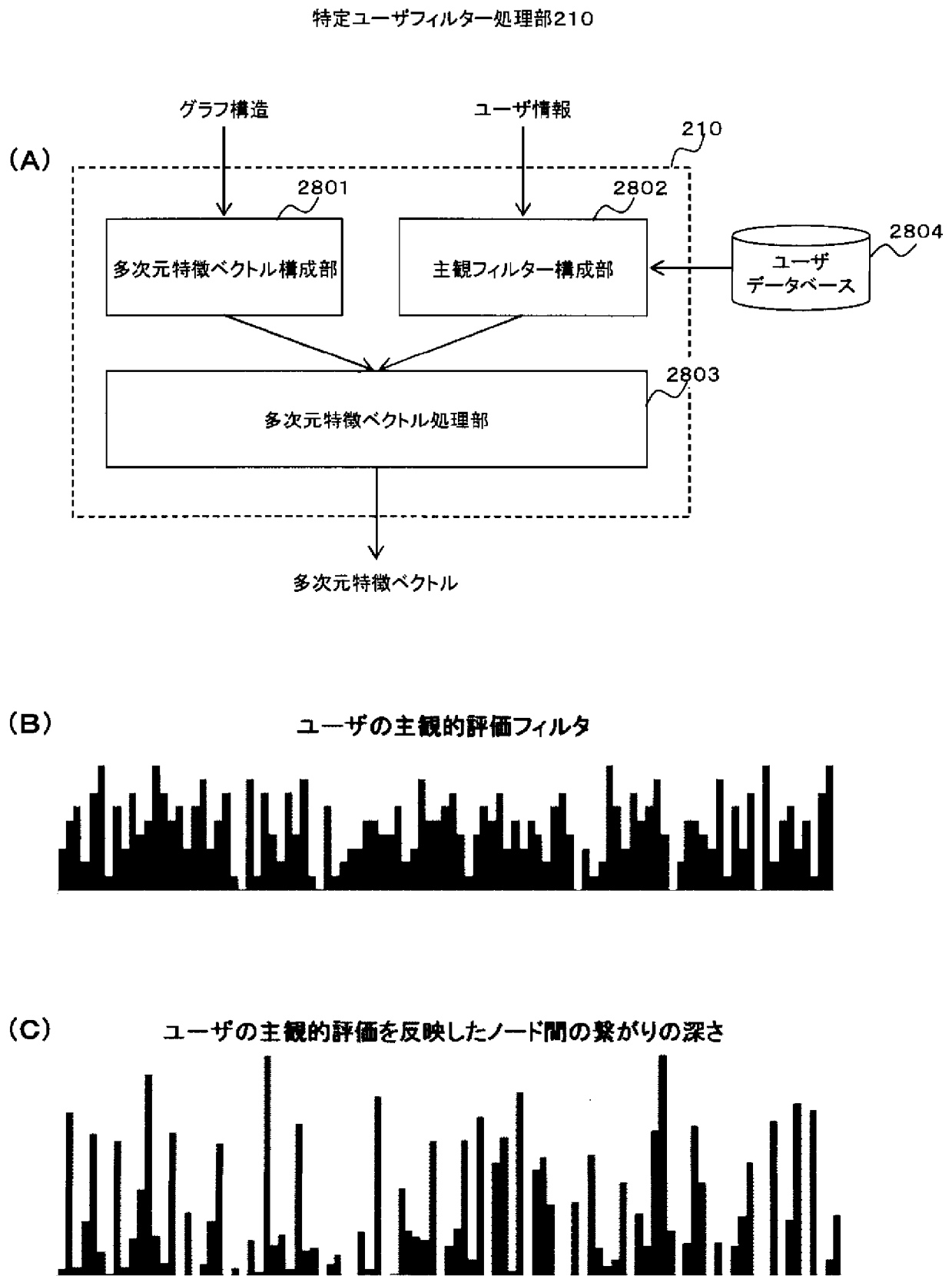


[図27]

## 統計情報処理部209(グラフ・データマイニング)

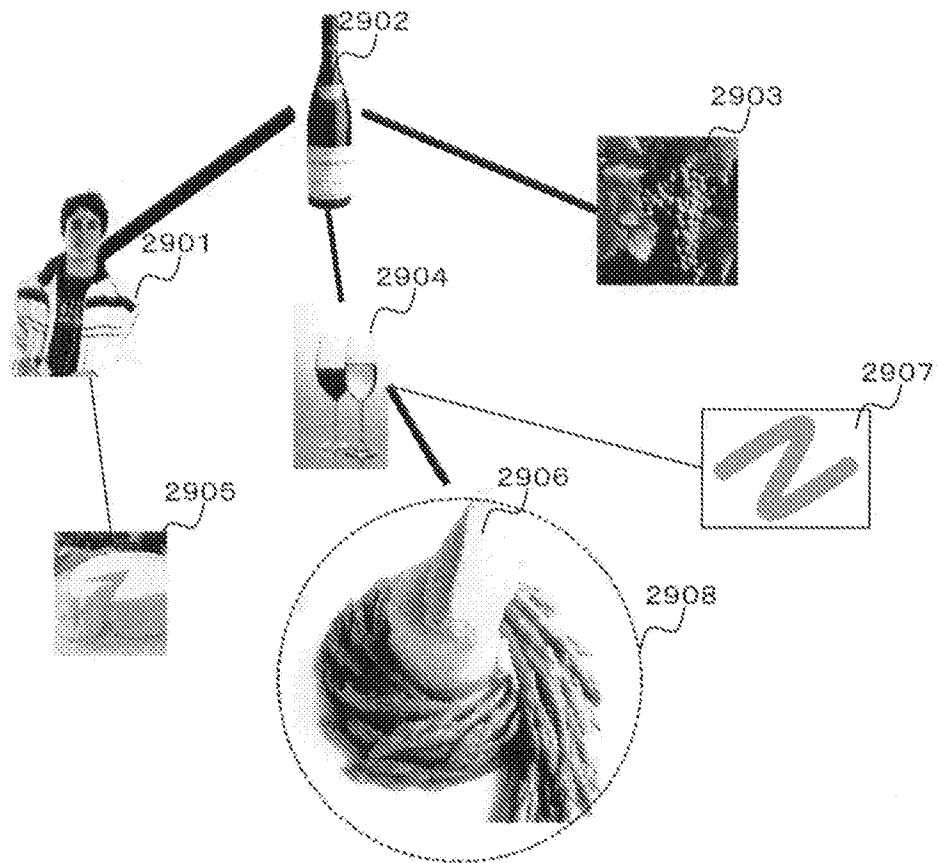


[図28]

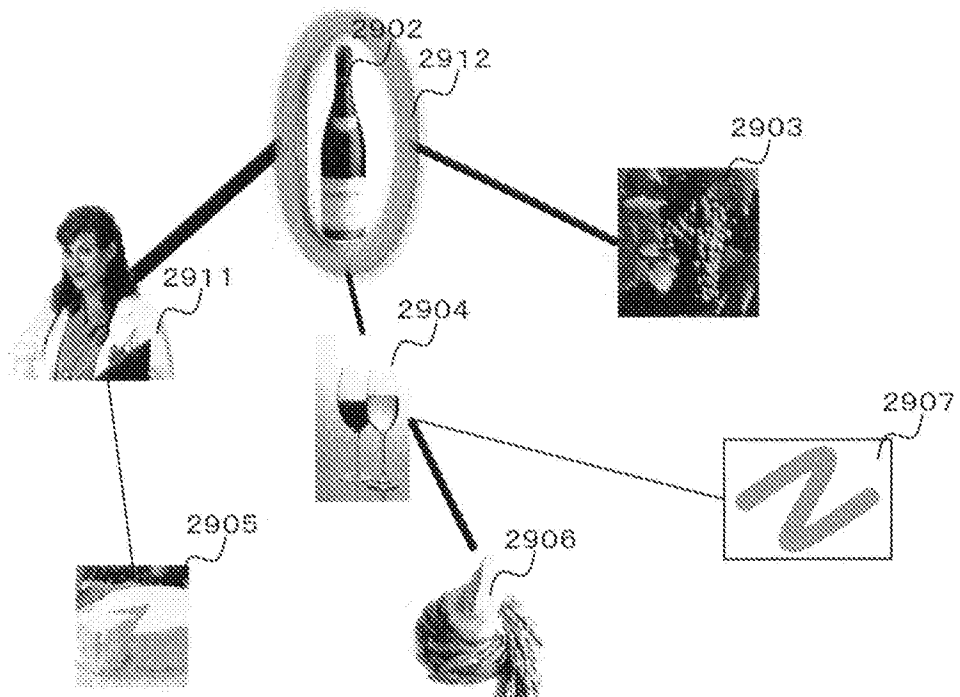


[図29]

(A)

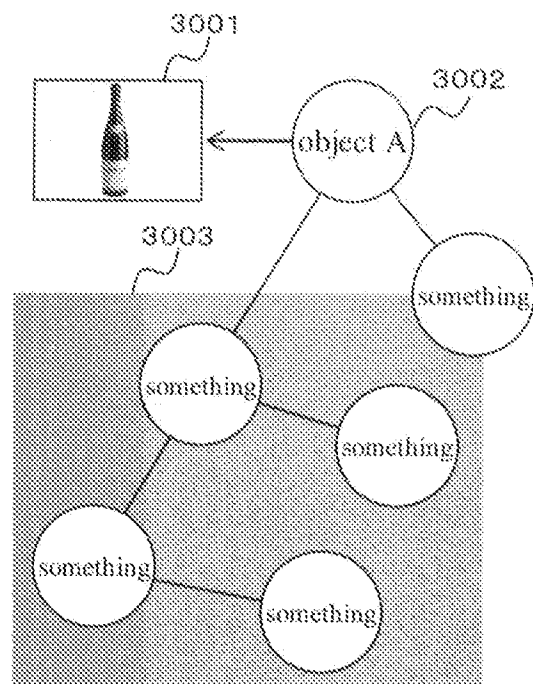


(B)

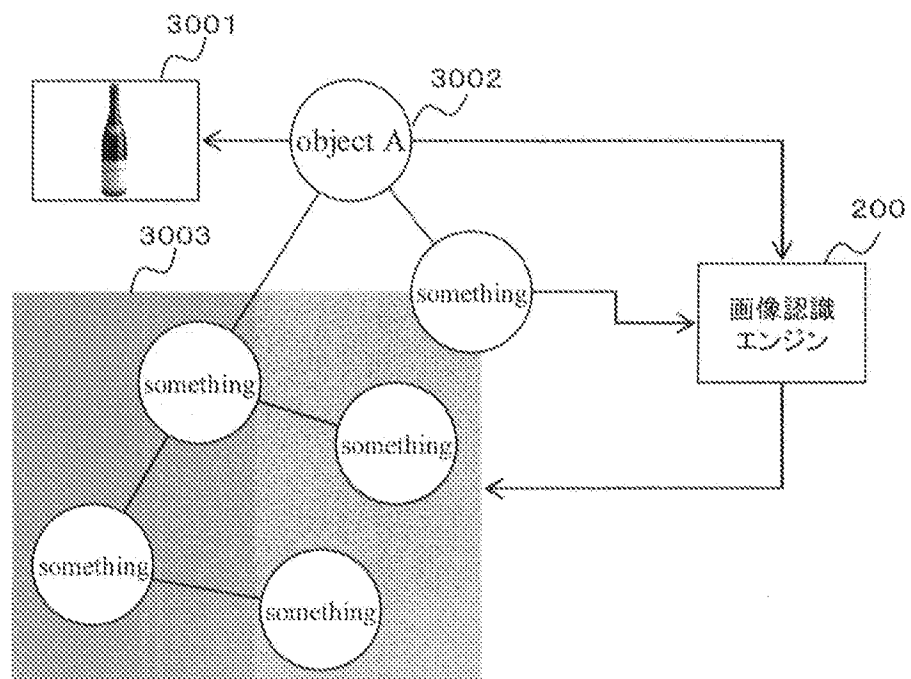


[図30]

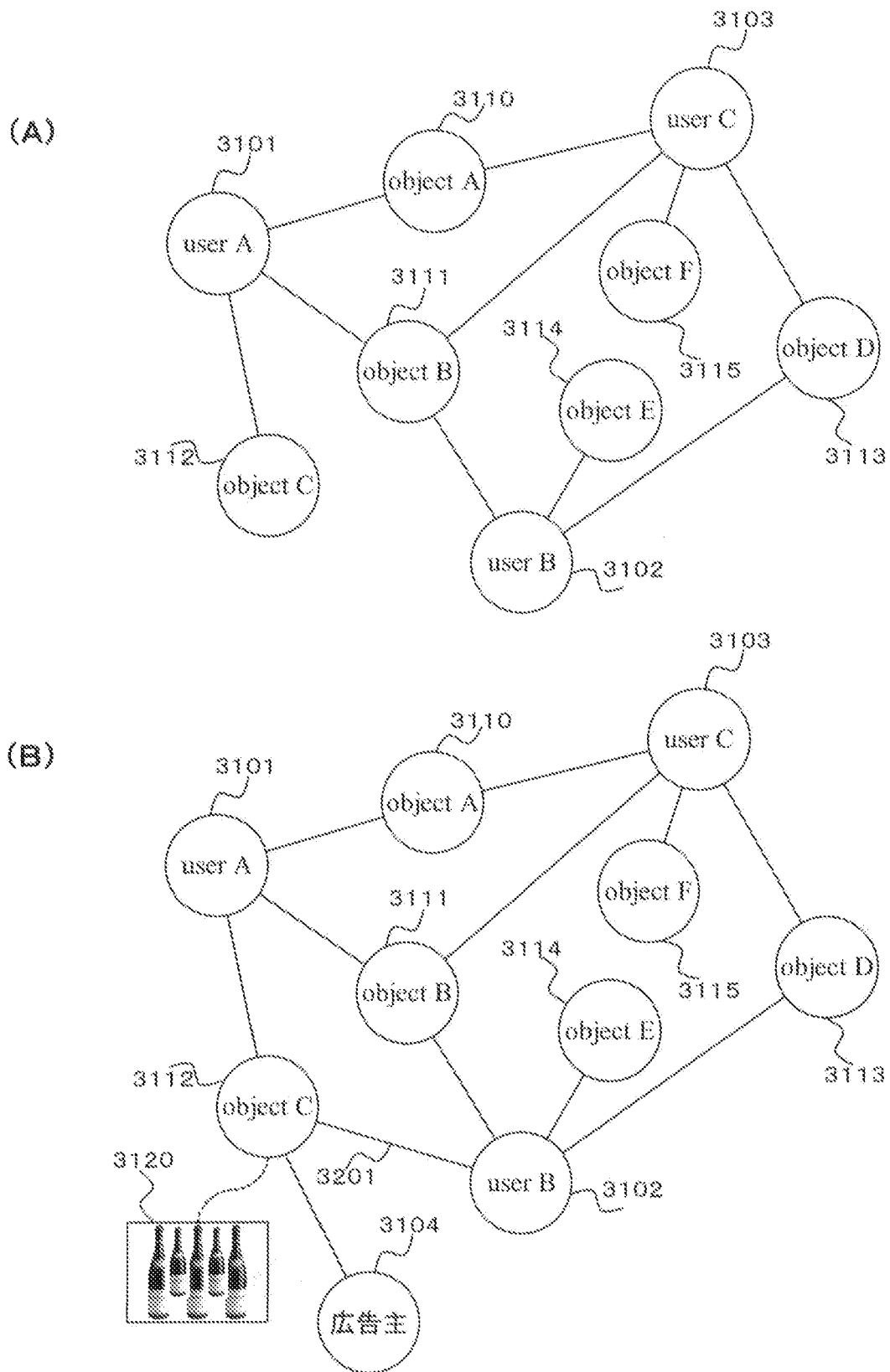
(A)



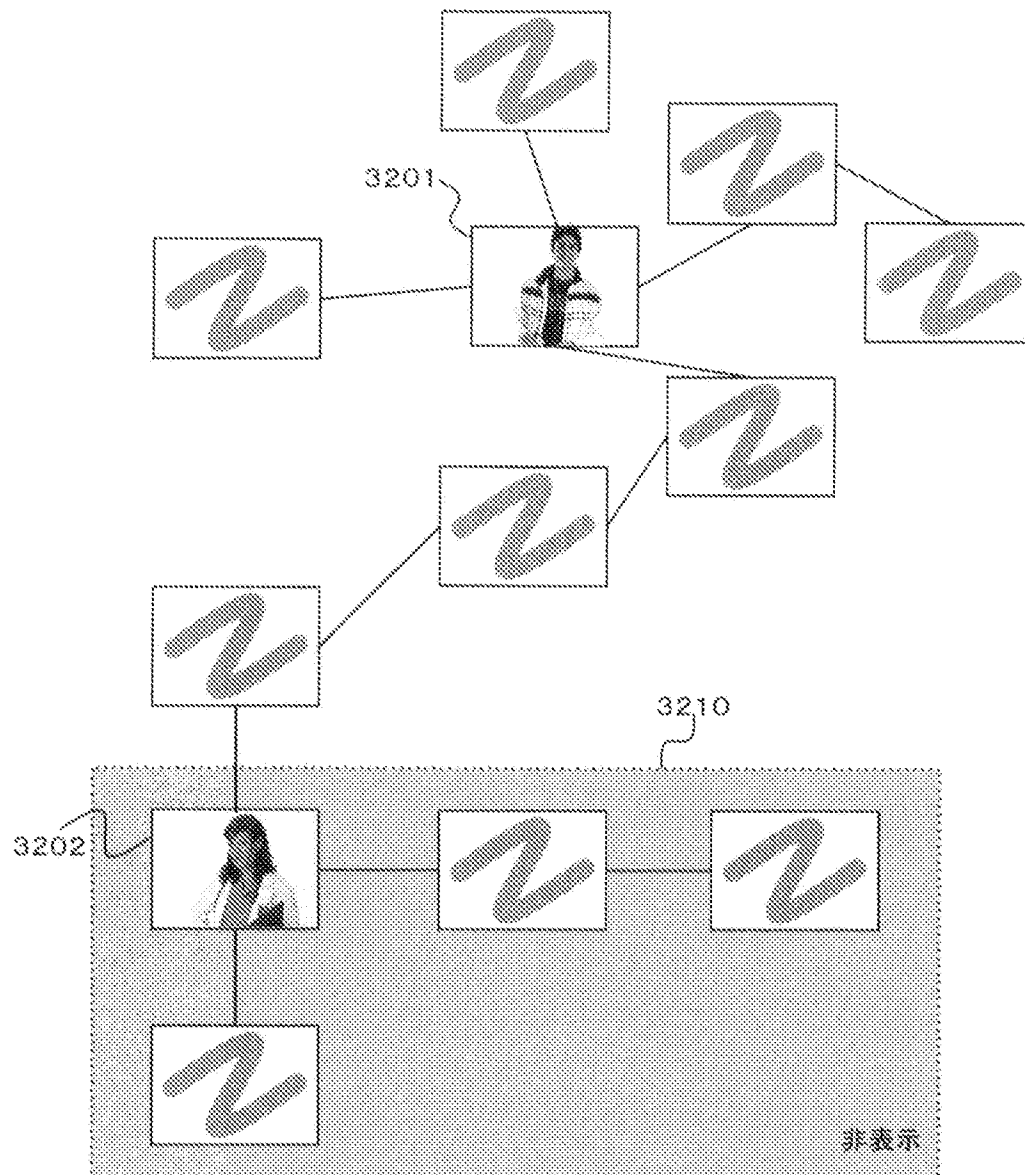
(B)



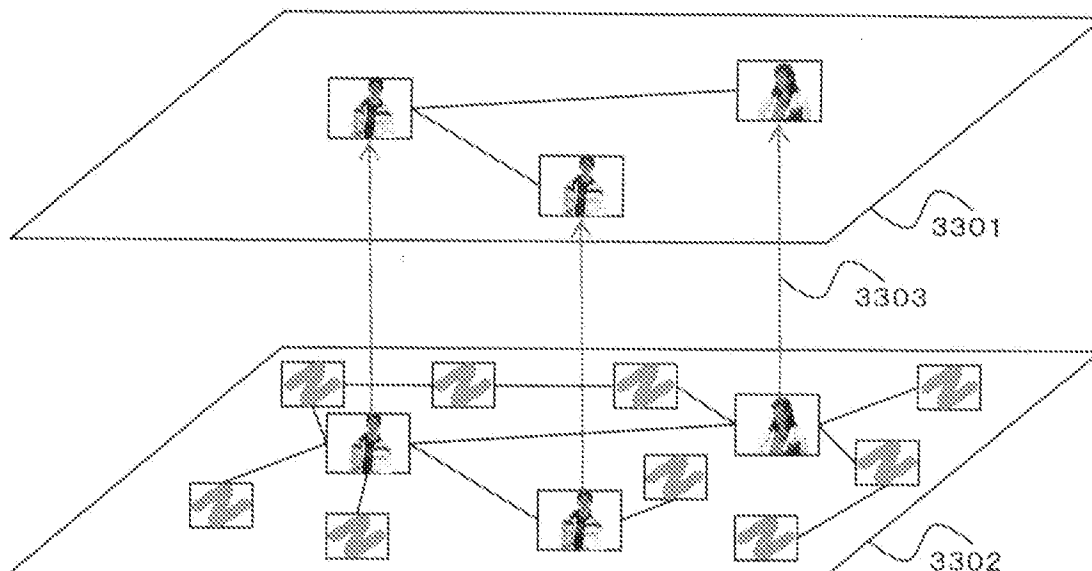
[図31]



[図32]

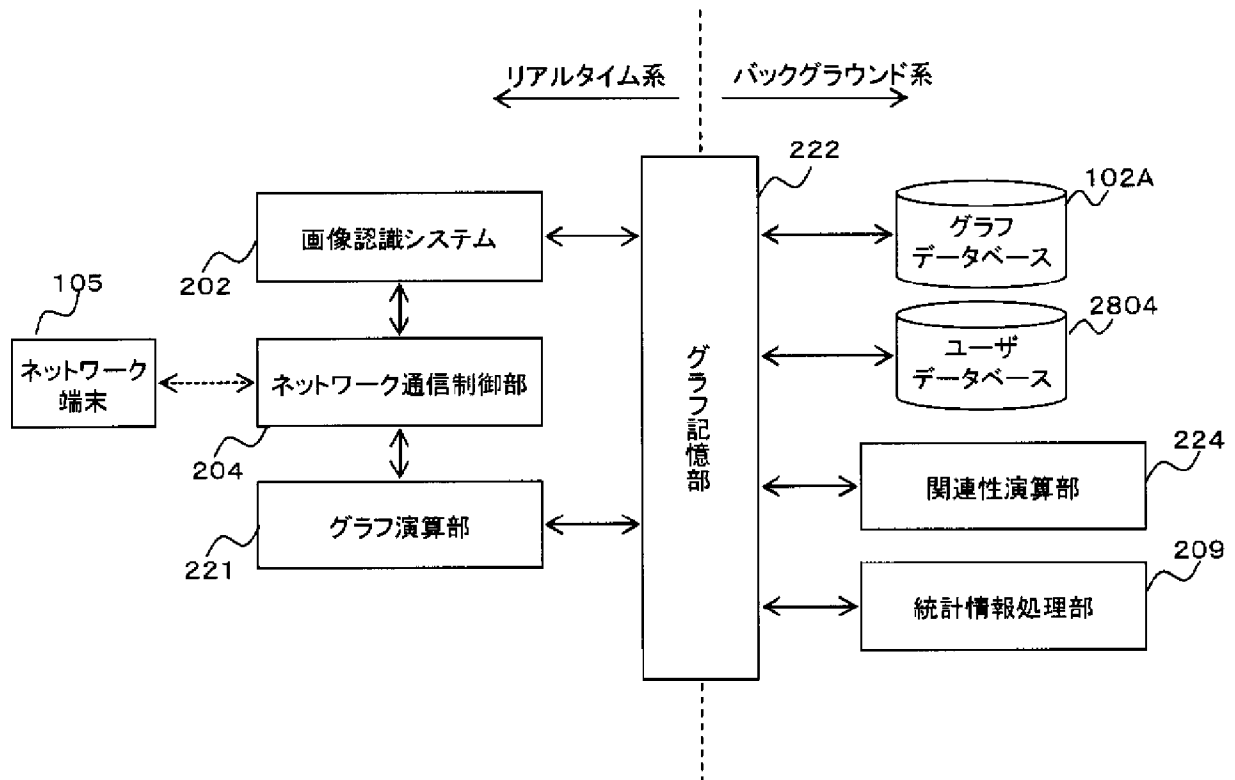


[図33]



[図34]

インタレスト・グラフ収集プロセス図

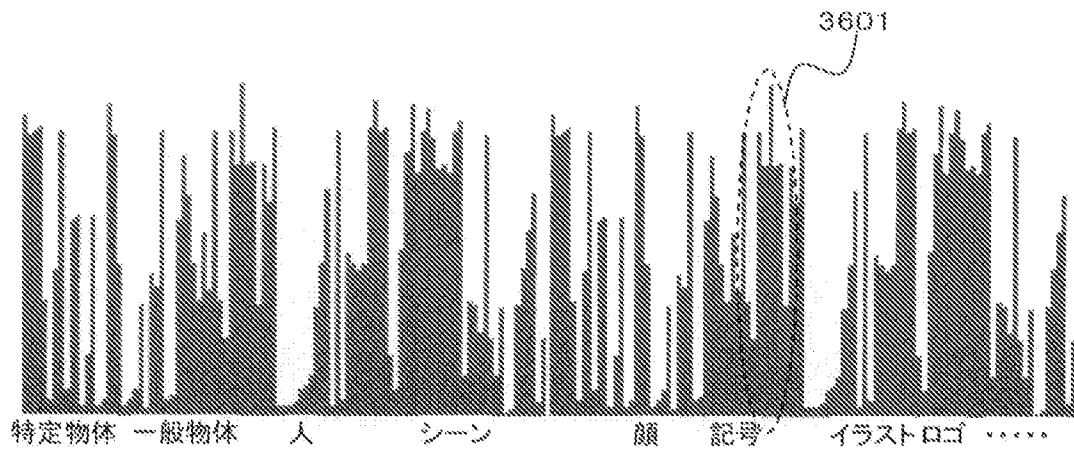


[図35]

ユーザ毎のインタレスト・グラフに対応した  
多次元特徴ベクトルのイメージ

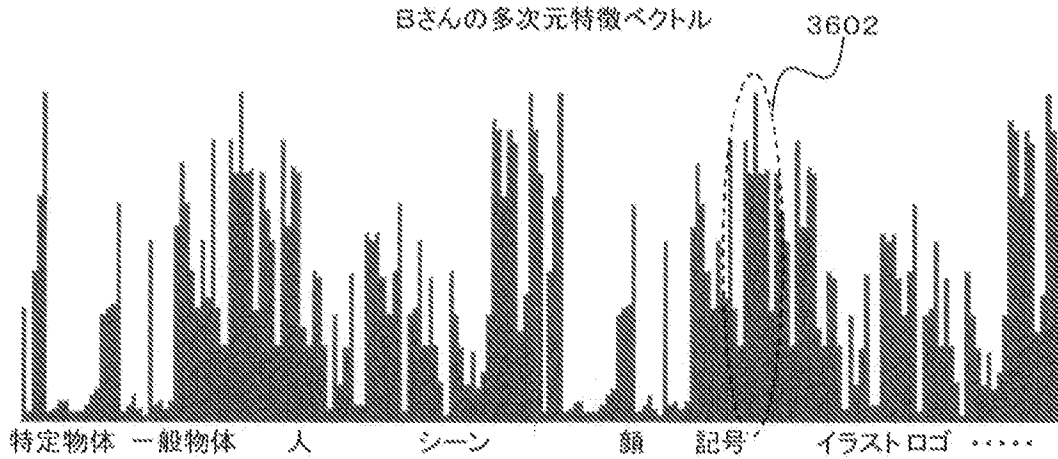
(A)

Aさんの多次元特徴ベクトル



(B)

Bさんの多次元特徴ベクトル



## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2011/064463

## A. CLASSIFICATION OF SUBJECT MATTER

G06T1/00(2006.01)i, G06F17/30(2006.01)i, G06T7/00(2006.01)i

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)

G06T1/00, G06F17/30, G06T7/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2011

Kokai Jitsuyo Shinan Koho 1971-2011 Toroku Jitsuyo Shinan Koho 1994-2011

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages       | Relevant to claim No.    |
|-----------|------------------------------------------------------------------------------------------|--------------------------|
| Y         | Yoshisuke YAMAKAWA et al., "Ruiji Gazo Kensaku Gijutsu o Mochiita Shohin Suisen System", | 1, 2, 5-8, 10, 11, 13-17 |
| A         | Image Lab, 10 August 2008 (10.08.2008), vol.19, no.8, pages 11 to 15                     | 3, 4, 9, 12              |
| Y         | JP 2010-250529 A (Yahoo Japan Corp.), 04 November 2010 (04.11.2010),                     | 1, 2, 5-8, 10, 11, 13-17 |
| A         | paragraphs [0046] to [0053]; fig. 6 (Family: none)                                       | 3, 4, 9, 12              |
| Y         | JP 2007-264718 A (Yahoo Japan Corp.), 11 October 2007 (11.10.2007),                      | 13                       |
|           | paragraphs [0013] to [0015]; fig. 1 (Family: none)                                       |                          |

☒ 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

"&amp;" document member of the same patent family

Date of the actual completion of the international search  
14 September, 2011 (14.09.11)Date of mailing of the international search report  
27 September, 2011 (27.09.11)Name and mailing address of the ISA/  
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2011/064463

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages                                                                                                                      | Relevant to claim No. |
|-----------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| Y         | Noriyuki YAMAZAKI, "03 Life Log & Recommendation ni yoru Jisedai Kensaku no Kanosei (1) Social Graph ga Tsumugu Atarashii Kankeisei", Web Site expert, 25 June 2008 (25.06.2008), no.18, pages 52 to 55 | 17                    |
| A         | JP 2009-205289 A (Nippon Telegraph and Telephone Corp.),<br>10 September 2009 (10.09.2009),<br>paragraph [0050]; fig. 5<br>(Family: none)                                                               | 1-17                  |
| A         | JP 2001-222586 A (Sony Corp.),<br>17 August 2001 (17.08.2001),<br>paragraph [0050]; fig. 4<br>(Family: none)                                                                                            | 1-17                  |
| A         | US 2006/0004892 A1 (Lunt et al.),<br>05 January 2006 (05.01.2006),<br>paragraphs [0057] to [0064]; fig. 7<br>(Family: none)                                                                             | 1-17                  |

## A. 発明の属する分野の分類（国際特許分類（I P C））

Int.Cl. G06T1/00(2006.01)i, G06F17/30(2006.01)i, G06T7/00(2006.01)i

## B. 調査を行った分野

調査を行った最小限資料（国際特許分類（I P C））

Int.Cl. G06T1/00, G06F17/30, G06T7/00

最小限資料以外の資料で調査を行った分野に含まれるもの

|             |                     |
|-------------|---------------------|
| 日本国実用新案公報   | 1 9 2 2 - 1 9 9 6 年 |
| 日本国公開実用新案公報 | 1 9 7 1 - 2 0 1 1 年 |
| 日本国実用新案登録公報 | 1 9 9 6 - 2 0 1 1 年 |
| 日本国登録実用新案公報 | 1 9 9 4 - 2 0 1 1 年 |

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

## C. 関連すると認められる文献

| 引用文献の<br>カテゴリー* | 引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示                                               | 関連する<br>請求項の番号                             |
|-----------------|---------------------------------------------------------------------------------|--------------------------------------------|
| Y<br><br>A      | 山川義介, 外 1 名, " 類似画像検索技術を用いた商品推薦システム", 画像ラボ, 2008.08.10, 第 1 9 巻, 第 8 号, p.11-15 | 1, 2, 5-8, 10,<br>11, 13-17<br>3, 4, 9, 12 |
| Y<br><br>A      | JP 2010-250529 A (ヤフー株式会社)<br>2010.11.04, 段落【0046-0053】, 第 6 図 (ファミリーなし)        | 1, 2, 5-8, 10,<br>11, 13-17<br>3, 4, 9, 12 |

☒ C 欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

## \* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの  
「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの  
「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）  
「O」口頭による開示、使用、展示等に言及する文献  
「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの  
「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの  
「Y」特に関連のある文献であって、当該文献と他の 1 以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの  
「&」同一パテントファミリー文献

国際調査を完了した日

1 4 . 0 9 . 2 0 1 1

国際調査報告の発送日

2 7 . 0 9 . 2 0 1 1

国際調査機関の名称及びあて先

日本国特許庁（I S A / J P）

郵便番号 1 0 0 - 8 9 1 5

東京都千代田区霞が関三丁目 4 番 3 号

特許庁審査官（権限のある職員）

千葉 久博

5 H

3 9 9 1

電話番号 0 3 - 3 5 8 1 - 1 1 0 1 内線 3 5 3 1

| C (続き) . 関連すると認められる文献 |                                                                                                                             |                |
|-----------------------|-----------------------------------------------------------------------------------------------------------------------------|----------------|
| 引用文献の<br>カテゴリー*       | 引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示                                                                                           | 関連する<br>請求項の番号 |
| Y                     | JP 2007-264718 A (ヤフー株式会社)<br>2007.10.11, 段落【0013-0015】, 第1図 (ファミリーなし)                                                      | 13             |
| Y                     | 山崎徳之, ” 0 3 ライフログ&レコメンデーションによる次世代<br>検索の可能性 (1) ソーシャルグラフが紡ぐ新しい関係性”, W<br>e b S i t e e x p e r t, 2008.06.25, 第18号, p.52-55 | 17             |
| A                     | JP 2009-205289 A (日本電信電話株式会社)<br>2009.09.10, 段落【0050】, 第5図 (ファミリーなし)                                                        | 1-17           |
| A                     | JP 2001-222586 A (ソニー株式会社)<br>2001.08.17, 段落【0050】, 第4図 (ファミリーなし)                                                           | 1-17           |
| A                     | US 2006/0004892 A1 (Lunt et al.)<br>2006.01.05, 段落【0057-0064】, 第7図 (ファミリーなし)                                                | 1-17           |