

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5509753号
(P5509753)

(45) 発行日 平成26年6月4日(2014.6.4)

(24) 登録日 平成26年4月4日(2014.4.4)

(51) Int.Cl. F 1
G 0 6 F 17/30 (2006.01)
 G 0 6 F 17/30 3 2 0 Z
 G 0 6 F 17/30 3 8 0 D
 G 0 6 F 17/30 1 1 0 C

請求項の数 22 (全 57 頁)

(21) 出願番号	特願2009-212242 (P2009-212242)	(73) 特許権者	000006747 株式会社リコー 東京都大田区中馬込1丁目3番6号
(22) 出願日	平成21年9月14日(2009.9.14)	(74) 代理人	100070150 弁理士 伊東 忠彦
(65) 公開番号	特開2010-67274 (P2010-67274A)	(72) 発明者	岸 秀信 東京都町田市鶴間1517-1 サンガー デン南町田C203号
(43) 公開日	平成22年3月25日(2010.3.25)	(72) 発明者	高橋 禎郎 神奈川県横浜市港北区篠崎町72-2-3 -205号
審査請求日	平成24年7月5日(2012.7.5)	(72) 発明者	古川 隆弘 神奈川県川崎市宮前区馬絹1371-3- 306号
(31) 優先権主張番号	12/210,511		
(32) 優先日	平成20年9月15日(2008.9.15)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 認識結果を生成するためのシステム及び方法

(57) 【特許請求の範囲】

【請求項1】

検索リクエストを受信して、認識結果を生成するためのシステムであって、
プロセッサと、

少なくとも一部の画像とコンテキスト情報とを含む前記検索リクエストを受信し、前記
 検索リクエストを処理して問い合わせ画像および認識パラメータを生成し、前記認識結果
 を送信するゲートウェイと、

コンピュータが読み取ることが可能な記録媒体に格納されかつ前記プロセッサにより実
 行可能な照合部であって、前記問い合わせ画像および前記認識パラメータを前記ゲートウ
 ェイから受信し、前記問い合わせ画像および前記認識パラメータを解析してインデックス
 テーブルと対比し、前記認識結果に関するページ識別番号及びロケーションを含む前記認
 識結果を生成し、かつ前記認識結果を前記ゲートウェイに送信する照合部と、

前記照合部に結合され、前記認識結果で指定されている前記ページ及び前記ロケーショ
 ンに対応するホットスポット情報を抽出するホットスポットデータベースと
 を備えるシステム。

【請求項2】

少なくとも1つの画像を生成し、前記照合部に結合され、前記画像を前記照合部に与え
 る発行部をさらに有する請求項1に記載のシステムであって、

前記照合部は画像登録部を有し、該画像登録部は、前記画像から画像特徴を抽出するこ
 と及びインデックステーブルを更新することを自動的に行い、該画像登録部は、画像を受

信するように前記発行部に結合されている、システム。

【請求項 3】

前記検索リクエストが、問い合わせ画像、他のコンテキスト情報、及びメタデータのうちの1つ以上を含み、

前記問い合わせ画像が、何らかのフォーマットにおける画像又は該画像内の1つ以上の特徴で表現され、

前記認識結果は、ブール値(真/偽)、携帯可能な文書ファイルで定義されているアクション、アドレス、移動装置上で変換可能な情報、二値データ、追加的なアクションを含むメニュー、ラスト画像及び画像特徴を含む群中の1つ以上を含む、請求項1に記載のシステム。

10

【請求項 4】

プラグインモジュールを有する移動装置をさらに有する請求項1に記載のシステムであって、前記移動装置は、画像を捕捉しかつ前記ゲートウェイと通信することが可能であり、前記プラグインモジュールは、画像を処理し、捕捉した画像を該移動装置上で提示し、該捕捉した画像のユーザによる選択を受け付け、該捕捉した画像を前記移動装置から送信する、システム。

【請求項 5】

前記認識結果に関する前記ホットスポット情報が、前記認識結果に関するリンクを含む、請求項1に記載のシステム。

【請求項 6】

20

前記照合部が、

複数の認識ユニットを有する取得部であって、前記認識ユニット各々は対応するインデックステーブルに結合され、前記認識ユニットは、前記問い合わせ画像の認識を実行し、前記インデックステーブルから前記認識結果を検索する、取得部と、

問い合わせ画像及び認識パラメータを受信し、それらを前記取得部の特定の認識ユニットに送信するディスパッチ部と

をさらに有する、請求項1に記載のシステム。

【請求項 7】

複数の認識ユニット及び対応するインデックステーブルは、インビジブルジャンクショナルアルゴリズムを実行する第1認識ユニット、ブリックウォールコーディングアルゴリズムを実行する第2認識ユニット、及びパスコーディングアルゴリズムを実行する第3認識ユニットを含む、請求項6に記載のシステム。

30

【請求項 8】

複数の認識ユニット及び対応するインデックステーブルは、第1の画像群に関する第1認識ユニット及びインデックステーブルと、第2の画像群に関する第2認識ユニット及びインデックステーブルとを有し、前記第1の画像群は前記第2の画像群と異なる、請求項6に記載のシステム。

【請求項 9】

前記照合部が、前記取得部の前記インデックステーブルを更新する画像登録モジュールを有し、該画像登録モジュールは、画像及びページ番号を受信し、特徴を抽出し、該特徴及びページ番号を前記取得部の前記インデックステーブルに加える、請求項6に記載のシステム。

40

【請求項 10】

前記照合部が、前記インデックステーブルを調整するダイナミック負荷バランスを含み、該ダイナミック負荷バランスは、当該ディスパッチ部から送信された前記問い合わせ画像と、利用度を測定する前記取得部が出力した認識結果とを監視及び評価し、かつ前記画像登録モジュールへの制御信号を生成し、利用度に合うように前記インデックステーブルのコンフィギュレーションを修正する、請求項9に記載のシステム。

【請求項 11】

前記ディスパッチ部が、前記問い合わせ画像を受信する入力と、認識時間を示すオーダ

50

リング信号を出力する出力とを有する画像特徴オーダ部を含み、前記画像特徴オーダ部は、入力された問い合わせ画像を分析し、画像特徴を分析することで該入力された問い合わせ画像を認識するのに要する時間を予測する、請求項 6 に記載のシステム。

【請求項 1 2】

少なくとも 1 つの認識ユニットが、検索部に結合された特徴抽出部を含み、該特徴抽出部は、問い合わせ画像を受信する入力を有し、前記問い合わせ画像から画像特徴を生成し、それらを前記検索部に送信し、前記検索部は、前記画像特徴を前記インデックスと対比し、前記認識結果を生成する、請求項 6 に記載のシステム。

【請求項 1 3】

検索リクエストを受信し、認識結果を生成するためにシステムが実行する方法であって

10

、ゲートウェイが、少なくとも一部の画像とコンテキスト情報とを含む前記検索リクエストを受信し、前記検索リクエストを処理して問い合わせ画像および認識パラメータを生成するステップと、

前記ゲートウェイに結合された照合部が、前記問い合わせ画像および前記認識パラメータを前記ゲートウェイから受信し、前記問い合わせ画像および前記認識パラメータを解析してインデックステーブルと対比し、前記認識結果に関するページ識別番号及びロケーションを含む認識結果を生成し、かつ前記認識結果を前記ゲートウェイに送信するステップと

を有し、前記照合部に結合されたホットスポットデータベースが、前記認識結果で指定されている前記ページ及び前記ロケーションに対応するホットスポット情報を抽出する、方法。

20

【請求項 1 4】

前記照合部のディスパッチ部が、問い合わせ画像及び認識パラメータを受信し、それらを複数の認識ユニットを有する取得部の特定の認識ユニットに送信するステップ

をさらに有し、前記認識ユニット各々は対応するインデックステーブルに結合され、前記認識ユニットは、前記問い合わせ画像の認識を実行し、前記インデックステーブルから前記認識結果を検索する、請求項 1 3 に方法。

【請求項 1 5】

複数の認識ユニット及び対応するインデックステーブルは、インビジブルジャンクションアルゴリズムを実行する第 1 認識ユニット、ブリックウォールコーディングアルゴリズムを実行する第 2 認識ユニット、及びパスコーディングアルゴリズムを実行する第 3 認識ユニットを含む、請求項 1 4 に記載の方法。

30

【請求項 1 6】

複数の認識ユニット及び対応するインデックステーブルは、第 1 の画像群に関する第 1 認識ユニット及びインデックステーブルと、第 2 の画像群に関する第 2 認識ユニット及びインデックステーブルとを有し、前記第 1 の画像群は前記第 2 の画像群と異なる、請求項 1 4 に記載の方法。

【請求項 1 7】

前記照合部の画像登録モジュールにより、前記取得部の前記インデックステーブルを更新するステップを有し、該画像登録モジュールは、画像及びページ番号を受信し、特徴を抽出し、該特徴及びページ番号を前記取得部の前記インデックステーブルに加える、請求項 1 4 に記載の方法。

40

【請求項 1 8】

前記照合部のダイナミック負荷バランスが、前記インデックステーブルを動的に調整するステップを含み、該ダイナミック負荷バランスは、前記ディスパッチ部から送信された前記問い合わせ画像と、利用度を測定する前記取得部が出力した認識結果とを監視及び評価し、かつ前記画像登録モジュールへの制御信号を生成し、利用度に合うように前記インデックステーブルのコンフィギュレーションを修正する、請求項 1 7 に記載の方法。

【請求項 1 9】

50

前記照合部が、前記問い合わせ画像を受信する入力と、認識時間を示すオーダリング信号を出力する出力とを有する画像特徴オーダ部を含み、前記画像特徴オーダ部は、入力された問い合わせ画像を分析し、画像特徴を分析することで該入力された問い合わせ画像を認識するのに要する時間を予測する、請求項 14 に記載の方法。

【請求項 20】

少なくとも 1 つの認識ユニットが、検索部に結合された特徴抽出部を含み、該特徴抽出部は、問い合わせ画像を受信する入力を有し、前記問い合わせ画像から画像特徴を生成し、それらを前記検索部に送信し、前記検索部は、前記画像特徴を前記インデックスと対比し、前記認識結果を生成する、請求項 14 に記載の方法。

【請求項 21】

請求項 13 ないし 20 の何れか 1 項に記載の方法をコンピュータに実行させるコンピュータプログラムを記録した記録媒体。

【請求項 22】

請求項 13 ないし 20 の何れか 1 項に記載の方法をコンピュータに実行させる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、少なくとも 2 種類のメディアから形成された混合メディア文書を作成する技法に関連し、特にミックスドメディアリアリティ (MMR: Mixed Media Reality) システムに画像及び他のデータを登録することに関連し、MMR システムは印刷媒体及び電子媒体を併用して混合メディア文書を検索する。

【背景技術】

【0002】

文書の印刷及び複写の技術は長年にわたって様々な状況で使用されてきた。例えば、プリンタやコピー機は、業務用事務環境において、パーソナルコンピュータを備えた家庭環境において、文書の印刷及び発行サービス環境等において使用されている。しかしながら、印刷及び複写の技術は、静的な印刷媒体 (すなわち、紙の文書) と双方向性のある「仮想世界 (virtual world)」との間のギャップを埋める手段として、かつては考えられておらず、その手段は、デジタル通信、ネットワーク、情報提供、宣伝広告、娯楽及び電子商取引へのリンクを含む。

【0003】

何世紀のもの間、印刷媒体は、新聞や広告情報等のように情報を伝える一次資料であった。パーソナルデジタルアシスタント (PDA) 装置及びセルラ電話機 (例えば、カメラ付き携帯電話) のようなパーソナル電子機器及びパーソナルコンピュータの出現及び普及は、ここ数年の間に印刷媒体の概念を拡張し、その拡張は、印刷媒体を電子的に読取可能にして検索可能な形式にすること、及び従来の紙媒体では類のない双方向のマルチメディア能力を導入することによってなされている。

【0004】

不都合なことに、電子的にアクセス可能な電子マルチメディアの世界と印刷媒体の物理的な世界との間には、ギャップ又は隔たりがある。例えば、先進国のほとんどの人は印刷媒体や電子情報に毎日のようにアクセスしているが、印刷媒体のユーザや電子装置のユーザは、両者の間にリンクを形成する (すなわち、混合メディア文書を促す) のに必要なツールや技術を持っていない。

【0005】

通常の印刷媒体には、手触りの感じがあること、電力を要しないこと、構成及び保存の永続性等のような特に優れた性質があり、これらの性質は仮想的な又はデジタルメディアにはない。逆に、通常のデジタルメディアには、携帯の容易性 (例えば、セルラ電話やラップトップに保存して持ち運べること) 及び送付の容易性 (例えば、電子メール) 等のような優れた性質がある。

10

20

30

40

50

【発明の概要】

【発明が解決しようとする課題】

【0006】

特に問題なことの1つは、コンテンツの印刷バージョンを使って発行者がコンテンツの電子バージョンにアクセスできないことである。例えば、新聞の出版の場合、印刷された新聞を毎日受け取るユーザが、その新聞の画像を利用して、そのオンライン電子コンテンツだけでなく補足的なコンテンツにもアクセスできるようにする手段は、一切無い。さらに、発行者は一般に印刷の前に日々の新聞内容を電子形式で所有しているが、その内容を補足的な内容を伴う電子形式に移す簡易な手段は今のところ無い。

【0007】

従来技術の第2の問題は、移动通信装置（例えば、セルラ電話）の一部として広く普及している一般的な画像捕捉装置は、低品質の画像を生成することである。この低品質画像を元の印刷文書バージョンと比較しようとする、（不可能ではないにしても）認識は非常に困難になる。したがって、印刷された新聞の低品質画像を認識する方法が望まれている。

【0008】

従来技術の第3の問題は、入力された問い合わせ画像から元の文書のページ及び場所を正確に認識する際、画像認識プロセスは演算負担の観点から非常に高価であり、何分もかかるわけではないが何秒間も必要とするおそれがあることである。これは、大きなデータ群、例えば多数のページの文書の場合、特に問題になる。したがって、認識を実行するスピードを改善する手段が望まれている。

【0009】

従来技術の第4の問題は、低品質画像をデータベースの元の画像と比較すると、一致する候補を多数生成してしまうことである。さらに、低品質画像を問い合わせ画像として使用する場合、何らかの一致を発見するために複数の異なる認識アルゴリズムが必要とされるかもしれない。目下の従来技術には、それら複数の認識結果を、ユーザに提示可能な1つの結果に統合する手段が無い。

【0010】

したがって、マスメディア出版の分野においてミックスドメディアリアリティ（MMR）システムの利用を促す技術、方法及びシステムが望まれている。

【0011】

本発明の課題は、そのような技術、方法及びシステムを提供することである。

【課題を解決するための手段】

【0012】

本発明は、新聞出版に使用されるMMRシステムに伴う従来の欠点を克服する。新聞発行者用のMMRシステムは、画像及びコンテンツをMMRシステムに登録する自動手段を新聞発行者に提供できる点で、新聞発行用のMMRシステムは特に有利である。新聞発行用のMMRは、印刷された新聞の画像部分又はページで形成された問い合わせ画像に応答するよう構築された特有のアーキテクチャを有する点でも、新聞発行用のMMRシステムは特に有利である。

【図面の簡単な説明】

【0013】

【図1】ロケーション検索及び画像登録を行う本発明の一実施例による第1形態のミックスドメディアリアリティ（MMR）システムのブロック図。

【図2A】本発明に係る、移動装置、ネットワーク及びMMRゲートウェイの第1実施形態を示すブロック図。

【図2B】本発明に係る、移動装置、ネットワーク及びMMRゲートウェイの第2実施形態を示すブロック図。

【図2C】本発明に係る、移動装置、プラグイン、MMRゲートウェイ及びMMR照合部に関する他の実施形態を示すブロック図。

10

20

30

40

50

【図 2 D】本発明に係る、移動装置、プラグイン、MMR ゲートウェイ及び MMR 照合部に関する他の実施形態を示すブロック図。

【図 2 E】本発明に係る、移動装置、プラグイン、MMR ゲートウェイ及び MMR 照合部に関する他の実施形態を示すブロック図。

【図 2 F】本発明に係る、移動装置、プラグイン、MMR ゲートウェイ及び MMR 照合部に関する他の実施形態を示すブロック図。

【図 2 G】本発明に係る、移動装置、プラグイン、MMR ゲートウェイ及び MMR 照合部に関する他の実施形態を示すブロック図。

【図 2 H】本発明に係る、移動装置、プラグイン、MMR ゲートウェイ及び MMR 照合部に関する他の実施形態を示すブロック図。

10

【図 3】本発明による MMR ゲートウェイの一形態を示すブロック図。

【図 4 A】本発明の第 1 形態による MMR 照合部のブロック図。

【図 4 B】本発明の第 2 形態による MMR 照合部のブロック図。

【図 5】本発明の一形態によるディスパッチ部のブロック図。

【図 6 A】本発明の第 1 形態による画像検索部のブロック図。

【図 6 B】本発明の第 2 形態による画像検索部のブロック図。

【図 7】本発明の一形態による登録部のブロック図。

【図 8】本発明の一形態による MMR 発行部のブロック図。

【図 9】入力画像から文書及びロケーションを検索する本発明の一形態による方法を示すフローチャート。

20

【図 10】画像を MMR 照合部に登録する方法の一実施形態を示すフローチャート。

【図 11】負荷バランスをとる方法の一実施形態を示すフローチャート。

【図 12】画像特徴に基づく順序づけを行う方法の一実施形態を示すフローチャート。

【図 13】不鮮明な画像及び対応するインデックスを生成する方法の一実施形態を示すフローチャート。

【図 14】トレーニング用の不鮮明な画像を生成する方法の一実施形態を示すフローチャート。

【図 15 A】不鮮明画像に関する分類子を初期化及び生成する方法の第一実施形態を示すフローチャート。

【図 15 B】不鮮明画像に関する分類子を初期化及び生成する本発明の第 2 形態による方法の第二実施形態を示すフローチャート。

30

【図 16】最良の画像を認識する分類子を発見する本発明の一形態による方法のフローチャート。

【図 17】認識されていない画像に対して分類子を修正及び検査する本発明の一形態による方法のフローチャート。

【図 18】本発明の一形態による品質予測器のブロック図。

【図 19】本発明の一形態によるベクトル計算器のブロック図。

【図 20】ロバストネス特性を出力する本発明の一形態による方法のプロセス図。

【図 21】予測スコア（品質予測スコア）を生成する本発明の一形態による方法のプロセス図。

40

【図 22】予測器を用いて複数の結果を結合する本発明の第 1 形態によるフローチャート。

【図 23】予測器を用いて複数の結果を結合する本発明の第 2 形態によるフローチャート。

【図 24】複数の結果を結合する本発明の一形態によるフローチャート。

【発明を実施するための形態】

【0014】

一実施例では、新聞発行用の MMR システムは、複数の移動装置と、コンピュータと、MMR ゲートウェイと、MMR 照合部と、MMR 発行部とを有する。移動装置は MMR ゲートウェイと通信可能に結合され、問い合わせ画像（image query）及び他の状況情報（

50

コンテキスト情報)を含む検索リクエストを送信する。MMRゲートウェイは、何千台ではないにしても何百台もの移動コンピュータ装置に結合され、それらの検索リクエストに対してサービスを提供する。MMRゲートウェイはコンピュータにも通信可能に結合され、そのコンピュータは、MMRゲートウェイの管理及び保守を行い且つ業務アプリケーションを動作させる。一実施例では、MMRゲートウェイは、移動装置からの検索リクエストを処理し、ユーザ認証、財務処理、分析その他の通信を行い、MMR照合部に伝送する問い合わせ画像を生成する。MMR照合部はディスパッチ部、複数の認識部及びインデックステーブルを含むことに加えて、画像登録部も含む。MMR照合部は、MMRゲートウェイから問い合わせ画像を受信し、それを1つ以上の認識部に送り、問い合わせ画像に対応する文書、ページ及びページ中の場所を示す認識結果を特定する。その結果はMMRゲートウェイを介して移動装置に返される。MMR照合部の画像認識部は、MMR発行部にも結合され、新たなコンテンツを受信し、それに応じてMMR照合部のインデックステーブルを更新する。

10

【0015】

本発明は多数の新規な方法も含み、例えば、画像及び他のデータをMMR照合部に自動的に登録する方法、動的に負荷を分散してバランスをとる方法、及び画像特徴に基づく順位づけ方法等を含む。

【0016】

本願で説明される特徴及び利点は網羅的なものではなく、多くの追加的な特徴及び利点が、本明細書及び図面により当業者に明らかになるであろう。さらに、本明細書で使用されている言葉は原則として読みやすさや説明を意図して選択されており、本発明の対象範囲を限定するものではない。

20

【0017】

以下、添付図面と共に被限定的な実施例により本発明が説明される。図中、同様な番号は同様な要素を指す。

【実施例1】

【0018】

問い合わせ画像を受信し、文書のページ及び場所に加えて受信画像、ホットスポット及び他のデータを返し、そのような情報をMMRシステムに追加できるミックストメディアリアリティ(MMR)システム100のアーキテクチャが開示される。説明を目的とする以下の詳細な説明では、本発明の実施例の十分な理解を図るため、様々な具体的な詳細が述べられる。しかしながら、本発明はそのような具体的な詳細によらず実施されてもよいことは明白であろう。それ以外には、本発明を不必要に曖昧にするかもしれないことを避けるため、構造や装置はブロック図形式で示される。例えば、通常のマスメディア発行者(特に、新聞発行者)による利用例の観点から、本発明が説明される。しかしながら、本発明は、電子媒体及び印刷媒体を含む複数のメディア形式が使用されている如何なるタイプのコンピュータシステムやデータ処理に適用されてもよい。

30

【0019】

本明細書における「一実施例」又は「形態」は、その実施例に関して説明される特定の特質、構造又は特徴が、本発明の少なくとも1つの実施例に含まれることを意味する。本明細書の様々な場所で「一実施例」なる用語が使用されているが、それらが全て同じ実施例を必ずしも指しているわけではない。特に、本発明は2つの異なるアーキテクチャに関して説明され、ある要素は双方のアーキテクチャで使用可能であるが、他の要素はそうではない。

40

【0020】

以下の詳細な説明の一部は、コンピュータメモリ内のデータビットを処理するアルゴリズム及び記号表現で与えられる。これらのアルゴリズムの説明及び表現は、データ処理技術分野の当業者が他の当業者に彼らの仕事内容を最も効率的に伝えるのに使用される手段である。ここで、アルゴリズムは一般に所望の結果に導く首尾一貫した一連のステップと考えられる。そのステップは物理量の物理的処理を必要とするものである。必須ではない

50

が、通常それらの物理量は、格納、転送、結合、比較その他の処理を施すことの可能な電氣的な又は磁氣的な信号の形態をとる。原則的な一般的な用法の観点から、ビット、値、エレメント、シンボル、キャラクタ、期間、数等としてそれらの信号に言及することが折に触れて便利なが分かる。

【 0 0 2 1 】

しかしながら、これらの及び類似の用語の全ては、適切な物理量に関連しており且つそれらの量に付された便宜的なラベルにすぎないことに留意を要する。特に断りのない限り、以下の説明から明らかなように、本説明を通じて、「処理」、「演算」、「計算」、「決定」又は「表示」等のような用語を用いる説明は、コンピュータシステム又は同様な電子コンピュータ装置の動作や処理に関連し、その動作や処理は、コンピュータシステムのレジスタ及びメモリの中で物理的な（電子的な）量として表現されるデータを、コンピュータシステムメモリやレジスタその他の情報ストレージ、伝送又は表示装置の中で物理量として同様に表現される他のデータに変換及び処理することが、理解されるであろう。

【 0 0 2 2 】

本発明はここで説明される処理を実行する装置にも関連している。その装置は、必要な目的に応じて特別に構築されてもよいし、コンピュータに格納されているコンピュータプログラムによって選択的にアクティブにされる又は再構成される汎用コンピュータで構築されてもよい。そのようなコンピュータプログラムはコンピュータ読取可能な記憶媒体に格納されてもよく、その記憶媒体は、限定ではないが、フロッピディスク、光ディスク、CD-ROM、磁気光ディスク、リードオンリメモリ（ROM）、ランダムアクセスメモリ（RAM）、EPROM、EEPROM、磁気又は光カード等の如何なるタイプのディスクを含んでもよいし、或いは電子的な命令を格納するのに指摘した如何なるタイプの媒体を含んでもよいし、ディスクも媒体もそれぞれコンピュータシステムバスに結合される。

【 0 0 2 3 】

そして、ここで説明されるアルゴリズム及び表示は、特定のコンピュータその他の装置のどれにも固有に関連するものではない。様々な汎用システムがここで教示するものによるプログラムと共に使用されてよく、或いは、必要な方法ステップを実行するように、よりいっそう特化した装置を構築することが便利なが分かるかもしれない。これら様々なシステムに必要な構造は以下の説明から明らかになるであろう。さらに、本発明は特定のプログラミング言語のどれにも依存しない。ここで説明される本発明の教示内容を実現するのに様々なプログラミング言語が使用されてよいことが分かるであろう。

システム概要

図1は、本発明の一実施例によるMMRシステム例100を示す。MMRシステム100は、複数の移動装置102a - 102n、MMRゲートウェイ104、MMR照合部106、MMR発行部108及びコンピュータ110を有する。本発明は、新聞発行に使用するMMRシステム100を提供する。新聞発行用のMMRシステムは、画像及びコンテンツをMMRシステム100に登録する自動手段を新聞発行者に提供できる点で、新聞発行用のMMRシステム100は特に有利である。新聞発行用のMMRは、印刷された新聞の画像部分又はページを含む問い合わせ画像に応答するよう構築された特有のアーキテクチャを有する点でも、新聞発行用のMMRシステム100は特に有利である。

【 0 0 2 4 】

移動装置102a - 102nは、信号線132a - 132nによりそれぞれMMRゲートウェイ104に通信可能に結合され、「検索リクエスト」を送信する。検索リクエストは、1つ以上の「問い合わせ画像（image query）」、他のコンテキスト情報（状況を示す情報）及びメタデータを含む。一実施例では、問い合わせ画像は如何なるフォーマットでよいし、或いは画像の1つ以上の特徴でもよい。問い合わせ画像の具体例は、静止画、ビデオフレーム及び一連のビデオフレーム等を含む。移動装置102a - 102nは、画像を捕捉するカメラを有する携帯電話である。MMRシステム100は、毎日の新聞のよ

うな従来の刊行物を受信する何百人或いは何千人ものユーザにより使用されることが、理解されるべきである。したがって、2つの移動装置102a、102nしか示されていないが、MMRゲートウェイ104は、多数の移動装置102a - 102nからの検索リクエストを受信及び応答するように同時に結合されてよいことを、当業者は認識するであろう。移動装置102a - 102nの代替例は、図2A及び2Bを参照しながら詳細に説明される。

【0025】

上述したように、MMRゲートウェイ104は、何百万台ではないにしても何百台もの移動通信装置102a - 102nと接続し、それらの検索リクエストに応じるサービスを提供できる。MMRゲートウェイ104は、信号線130を介してコンピュータ110に通信可能に結合し、コンピュータ110は、MMRゲートウェイ104の保守及び管理を行い、業務用アプリケーションを実行する。一実施例では、MMRゲートウェイ104はウェブのポータルを作成及び提示し、業務用アプリケーションを走らせることに加えてMMRシステム100のユーザの履歴にアクセスするコンピュータ110によるアクセスに備える。コンピュータ110は、パーソナルコンピュータのような通常の任意のコンピュータ装置とすることができる。MMRゲートウェイ104の主な機能は、移動装置102a - 102nからの検索リクエストを処理すること、及び認識結果を移動装置102a - 102nに返すことである。一実施例では、その認識結果は1つ以上のブール値（真/偽）を含み、真の場合、ページID及びページ上の位置の情報を含む。他の実施例の場合、認識結果は、アクション（処理）、認識が成功した（又は失敗した）ことの応答メッセージ、及び判定に起因する処理等を含む群の中の1つ以上を含み、その処理は、例えば、電子メールメッセージ、文書、ポータル文書ファイルで規定されたアクション（処理）、URLのようなアドレス、ビデオのような二値データ、移動装置102において提示可能な情報、追加的なアクションのメニュー、ラスト画像、画像特徴等を送ることである。MMRゲートウェイ104は、ユーザの認証、財務処理、分析及び他の通信を行うことで、受信した検索リクエストを処理する。MMRゲートウェイ104は、検索リクエストから問い合わせ画像及び認識パラメータを生成し、それらを信号線134を介してMMR照合部に渡す。MMRゲートウェイ104の例及び動作は、図3を参照しながら詳細に説明される。

【0026】

MMR照合部106は、MMRゲートウェイ104から信号線134を介して問い合わせ画像を受信し、（認識）結果を得るためにそれを1つ以上の認識部に送信し、その結果は、問い合わせ画像に対応する文書、ページ及びページ中の場所を含み、このようなプロセスは本願において一般的に「検索プロセス（retrieval process）」と言及される。結果は、信号線134を介して、MMR照合部106からMMRゲートウェイ104へ返される。その結果に加えて、MMR照合部106は、ホットスポットデータのような関連する他の情報を返してもよい。MMR照合部106は、新たなコンテンツを受信し、検索プロセスで使用されるインデックステーブルを更新及び再編成する要素を含む。新たなコンテンツをMMR照合部106に加えるプロセスは、本願では一般的に「登録プロセス（registration process）」と言及される。一実施例では、MMR照合部106は信号線138、140を介してMMR発行部108の出力に結合され、MMR照合部106のインデックステーブルを更新するのに使用される新たなコンテンツを得る。代替実施例では、MMR発行部108は信号線138を介してMMRゲートウェイ104に結合され、MMRゲートウェイ104は信号線136によりMMR照合部106に結合される。この代替実施例の場合、MMRゲートウェイ104は、ホットスポット情報のような補足的データを抽出し、それを保存し、画像、参照ページ及び他の情報をMMR照合部106に渡し、インデックステーブルの更新に備える。MMR照合部106及びその構成要素に関する様々な形態は、図4Aないし7を参照しながら以下でさらに説明される。

【0027】

MMR発行部108は、新聞その他の形式の定期刊行物を作成するのに使用される従来

10

20

30

40

50

の出版用システム（パブリッシャーシステム）を含む。一実施例では、MMR発行部108は、印刷文書の画像をMMRシステム100に登録するのに必要な追加的な情報を生成する要素を含む。MMR発行部108がMMR照合部106に与える情報は、画像ファイル、境界枠（バウンディングボックス）データ、ホットスポットデータ、及びページ固有の識別番号等を含む。或るシステム例の場合、これは、カリフォルニア州サンノゼのアドビコーポレーションによるポータブルドキュメントフォーマットの文書及び境界枠情報である。MMR発行部108の例は、図8を参照しながら以下でさらに説明される。

移動装置102

図2A及び2Bを参照しながら、移動装置102の第1及び第2形態を説明する。

10

【0028】

図2Aは、移動装置102及びMMRゲートウェイ104間の第1の接続形態を示す。この例では、移動装置102は、カメラを含む任意の移動電話である（又は、通信能力を有する他の携帯用コンピュータ装置である。）。例えば、移動装置102は、リサーチインモーション社がブラックベリー（Blackberry）（登録商標）として製造販売するようなスマートフォンでもよい。移動装置102は、通信チャンネル230によりネットワーク202と無線通信するのに適している。ネットワーク202は、無線キャリアにより維持されるセルラネットワークのような通常のタイプのものであり、サーバを含む。この例の場合、移動装置102は、画像を取得し、その画像を通信チャンネル230を介して（例えば、マルチメディアメッセージングサービス（MMS）を使って）ネットワークに送る。ネットワーク202も通信チャンネル230を使って例えばMMSにより又はショートメッセージサービス（SMS）により結果を返す。図示されているように、ネットワーク202は信号線232を介してMMRゲートウェイ104に結合されている。信号線232は、MMSやSMSメッセージを送信するためのチャンネルを表すことに加えて、ハイパーテキストトランスファプロトコル（HTTP）リクエストを受信してHTTPレスポンスを送信するチャンネルを表わす。当業者は、これが、移動装置102及びMMRゲートウェイ104間を結合する一例に過ぎないことを認識するであろう。代替実施例では、移動装置102及びMMRゲートウェイ104間を結合する通信の一部分として、ブルートゥース（登録商標）、WiFi又は適切な他の任意の無線通信プロトコルが使用されてもよい。移動装置102がMMRゲートウェイ104に画像を送信することができ、かつMMRゲートウェイ104が文書識別子、ページ数及びロケーション情報を送信することで応答できる限り、移動装置102及びMMRゲートウェイ104は、当業者に既知の他の如何なる方法で結合されてもよい（例えば、直接的なデータ接続、SMS、WAP、電子メール等が使用されてもよい。）。

20

30

【0029】

図2Bを参照するに、第2形態の移動装置102が示されている。第2形態では、移動装置102は、カリフォルニア州クパチーノのアップルコンピュータ社がiPhone（登録商標）として製造販売するようなスマートフォンである。第2形態は第1形態と同様な多くの要素を有するので、同一の又は類似する機能を備えた同様な要素を参照する際、同様な参照番号が使用される。第1形態及び第2形態間の主な相違は、移動装置102にインストールされている品質予測プラグイン204が含まれていること、及びウェブサーバ206が信号線234によりネットワーク202に結合されていること等である。品質予測プラグイン204は、移動装置102が取得した画像を分析する。品質予測プラグイン204は、その分析に基づいて生成される追加的な情報を用意し、MMRゲートウェイ104に送信される検索要求の一部としてその情報を含め、検索の精度を高める。代替実施例では、品質予測プラグイン204の出力は、移動装置102からMMRゲートウェイ104へ送信される画像を選択するのに使用される。例えば、所定の閾値を超える予測品質を有する画像だけが（例えば、認識可能な画像だけが）、移動装置102からMMRゲートウェイ104に送信される。画像の伝送はかなりの帯域幅を要し、移動装置102及びネットワーク202間の通信チャンネル230は限られた帯域幅しか

40

50

いので、品質予測プラグイン 204 を使って送信画像を選択することは、特に有利である。図 2 B に示される第 2 形態は、MMR ゲートウェイ 104 から返された結果又は品質予測プラグイン 204 により用意された他の情報を、移動装置 102 のユーザが、ウェブサーバ 206 で利用可能な捕捉情報又はホットスポットにアクセスする際、どのように使用できるかをも示している。そのような場合、MMR ゲートウェイ 104 からの結果又は品質予測プラグイン 204 による出力は、ウェブサーバ 206 にアクセスするのに使用可能な情報を含んでおり、そのアクセスは、例えば、通常の HTTP リクエストと共に移動装置 102 のアクセス機能を用いてなされる。

【0030】

移動装置 102 について第 1 形態又は第 2 形態の何れが使用されるかによらず、移動装置 102 は、問い合わせ画像、ユーザ又は装置の ID、コマンド及び他のコンタクト情報を含む検索リクエストを生成し、他のコンタクト情報は、例えば、装置タイプ、ソフトウェア、プラグイン、居場所（例えば、移動装置が GPS 機能を有する場合）、装置及びステータスの情報（例えば、装置モデル、マクロレンズのオン/オフ状態、自動焦点のオン/オフ、バイブレーションのオン/オフ、チルト角等の情報）、コンテキスト関連情報（電話機が位置する場所の天候、時間、日付、電話機で現在動作しているアプリケーション等の情報）、ユーザ関連情報（例えば、id 番号、好み、ユーザの予約 [user subscriptions]、ユーザグループ及び社会的組織、アクション及びアクションに関するメタデータ [例えば、電子メールを送信するアクション、送信する電子メールを待機するアクション等]）等である。

【0031】

図 2 C ないし 2 H を参照するに、移動装置 102 のプラグイン（クライアント 250）、MMR ゲートウェイ 104 及び MMR 照合部 106 に関する様々な形態が示されており、MMR ゲートウェイ 104 及び MMR 照合部 106 は、本発明において可能性のある様々な構成を有するサーバ 252 を含むものとして概略的に表現されている。特に、図 2 C ないし 2 H は、プラグイン又はクライアント 250 の構成要素が様々な機能レベルをどのように備えることができるか、及びサーバ 252 がクライアントの機能に合わせて様々な機能レベルをどのように備えることができるかを示す。図 2 C ないし 2 H の様々な例において、クライアント 250 又はサーバ 252 は以下の要素を含む：MMR データベース 254、画像又はビデオを捕捉する捕捉モジュール 260、より良い認識を行うため特徴抽出の前に品質予測のような画像処理を行う予備処理モジュール 262、画像特徴を抽出する特徴抽出モジュール 264、その特徴を利用して、MMR データベース 254 中の情報を検索する検索モジュール 266、サーバ 252 からクライアント 250 へメッセージを送信するメッセージ送信モジュール 268、処理を実行するアクションモジュール 270、特徴抽出に先立って画像を処理する予備処理及び予測モジュール 272、ユーザに情報を提示し、入力を受信するフィードバックモジュール 274、クライアント 250 からサーバ 252 へ情報を送る送信モジュール 276、及びクライアント 250 からサーバ 252 へビデオをストリーミング送信するストリーミングモジュール 278 が含まれる。

【0032】

図 2 C は、クライアント 250 及びサーバ 252 の一例を示し、クライアント 250 は、画像又はビデオ及び/又はメタデータを、処理を行うサーバ 252 に送信する。この例の場合、クライアント 250 は捕捉モジュール 260 を含む。サーバ 252 は、MMR データベース 254、予備処理モジュール 262、特徴抽出モジュール 264、検索モジュール 266、メッセージ送信モジュール 268 及びアクションモジュール 270 を含む。

【0033】

図 2 D は、クライアント 250 及びサーバ 252 の別の例を示し、クライアント 250 は、画像又はビデオを取得し、品質予測を行い、その画像又はビデオ及び/又はメタデータを、処理を行うサーバ 252 に送信する。この例の場合、クライアント 250 は、捕捉モジュール 260、予備処理及び予測モジュール 272、フィードバックモジュール 274 及び送信モジュール 276 を含む。サーバ 252 は、MMR データベース 254、予備

10

20

30

40

50

処理モジュール262、特徴抽出モジュール264、検索モジュール266、メッセージ送信モジュール268及びアクションモジュール270を含む。この例の場合、サーバ252に送信される画像は、捕捉されたままの画像と異なってもよいことに留意すべきである。例えば、送信される画像は、デジタル的に強調されたり、整形されてもよいし、或いは単なるバイナリデータでもよい。

【0034】

図2Eは、クライアント250及びサーバ252の別の例を示し、クライアント250は、画像又はビデオを取得し、特徴抽出を行い、その画像特徴を、処理を行うサーバ252に送信する。この例の場合、クライアント250は、捕捉モジュール260、特徴抽出モジュール264、予備処理及び予測モジュール272、フィードバックモジュール274及び送信モジュール276を含む。サーバ252は、MMRデータベース254、検索モジュール266、メッセージ送信モジュール268及びアクションモジュール270を含む。この例の場合、特徴抽出は予備処理を含んでいることに留意を要する。特徴が抽出された後、それらの特徴について予備処理及び予測モジュール272が処理を行い、特徴の品質が満足できないものであった場合、別の画像を取得するようにユーザは求められる。

10

【0035】

図2Fは、クライアント250及びサーバ252の別の例を示し、検索プロセス全体がクライアント250で行われる。この例の場合、クライアント250は、捕捉モジュール260、特徴抽出モジュール264、予備処理及び予測モジュール272、フィードバックモジュール274、送信モジュール276、MMRデータベース254及び検索モジュール266を含む。サーバ252は、アクションモジュール270を備えることだけを要する。

20

【0036】

図2Gは、クライアント250及びサーバ252の別の例を示し、クライアント250は、ビデオをサーバ252に流す。この例の場合、クライアント250は、捕捉モジュール260、ストリーミングモジュール278を含む。サーバ252は、MMRデータベース254、予備処理モジュール262、特徴抽出モジュール264、検索モジュール266、メッセージ送信モジュール268及びアクションモジュール270を含む。図示されていないが、クライアント250は、捕捉したビデオストリームについて予測器を走らせ、カメラが指す場所をユーザにフィードバックしてもよいし、或いは検索により良いビデオを如何にして取得するかをユーザにフィードバックしてもよい。この例の変形例では、サーバ252は、捕捉されたビデオに関する情報をストリームで戻し、クライアント250はその情報をビデオのプレビュー画面に重ねて表示する。

30

【0037】

図2Hは、クライアント250及びサーバ252の別の例を示し、クライアント250は認識部を動作させ、サーバ252は、第1の認識結果に基づいて、MMRデータベース情報をクライアント250のローカルデータベースに流す。この例は、図2Fを参照しながら説明したものに似ている。例えば、ある認識アルゴリズムについて、全体的な検索プロセスがクライアント250で行われる。認識アルゴリズムが良い結果をもたらさなかった場合、問い合わせ画像は、さらに複雑な検索プロセスを行うサーバ252の処理に委ねられる。この例の場合、クライアント250は、捕捉モジュール260、特徴抽出モジュール264、予備処理及び予測モジュール272、フィードバックモジュール274、送信モジュール276、MMRデータベース254（ローカルバージョン）及び検索モジュール266を含む。サーバ252は、別の検索モジュール266、アクションモジュール270及びMMRデータベース254（完全な及びより複雑なバージョン）を含む。一実施例では、問い合わせ画像がローカルなMMRデータベース254では認識できなかった場合、クライアント250は、検索用の画像をサーバ252に送り、ローカルなMMRデータベース254の更新を開始する。或いは、クライアント250は或る認識部について更新されたバージョンのデータベースを有するが、問い合わせ画像がそのローカルMMR

40

50

データベース254の中で発見できなかった場合、別の検索アルゴリズム用のデータベースが、ローカルMMRデータベース254にストリーミングされてもよい。

MMRゲートウェイ104

図3を参照するに、MMRゲートウェイ104の一形態が示されている。この例のMMRゲートウェイ104は、サーバ302、ウェブサーバ304、データ格納部306、ポータルモジュール308、履歴(ログ)310、1つ以上のアプリケーション312、認証モジュール314、経理モジュール316、メールモジュール318及び分析モジュール320を有する。

【0038】

上述したように、MMRゲートウェイ104の主な機能の1つは、多数の移動装置102と通信し、検索要求を受信し、応答を送信することであり、その応答は、ステータスインジケータ(真=認識された場合/偽=認識されなかった場合)、ページ識別番号、ページ中の場所、及びホットスポットデータのような他の情報を含む。1つのMMRゲートウェイ104が、何百或いは何百万もの検索リクエストに回答できる。図示及び説明の便宜上、図1及び図3には1つのMMRゲートウェイ104しか示されていないが、当業者は、別の実施例の場合に任意の数のMMRゲートウェイ104が、多数の移動装置102のニーズに応じてサービスするよう使用されてよいことを認識するであろう。特に、MMRゲートウェイ104のサーバ302は、様々な移動装置102と通信するために信号線132a-132nに結合されている。サーバ302は、信号線132a-132nを介して移動装置102からの検索リクエストを受信し、同じ信号線132a-132nを用いて移動装置102に回答を送信する。一実施例では、検索リクエストは、コマンド、ユーザの識別番号、画像及び他のコンテキスト情報を含む。例えば、他のコンテキスト情報は、移動装置102の型、モデル又は製造者のような装置情報、移動装置の一部であるGPSシステムにより又は三角測量により得られた位置情報、日時、温度、天候状態、光、影等のような環境情報、オブジェクト情報、距離、場所、チルト及びジッタのような配置情報等である。

【0039】

サーバ302は、コンピュータ110と通信するために信号線130にも結合されている。再び、図示及び説明の便宜上、図1及び図3には1つのコンピュータ110及び1つの信号線130しか示されていないが、当業者は、任意の数のコンピュータ装置がサーバ302との通信に使用されてよい。サーバ302は、ポータルモジュール308、ログモジュール310及びアプリケーション312と、コンピュータ110との間の通信を促す。サーバ302は、ポータルモジュール308、ログモジュール310及びアプリケーション312と信号線330により結合されている。以下において詳細に説明するように、これらのモジュールはサーバ302と協同してウェブポータルを提供し、ユーザが情報をやり取りできるようにする。ウェブポータル308は、システムの監視、保守及び管理に使用可能である。

【0040】

サーバ302は、検索リクエストを処理し、問い合わせ画像及び認識パラメータを生成し、問い合わせ画像及び認識パラメータは、認識を行うMMR照合部106へ信号線134を介して伝送される。また、サーバ302は、信号線134を介してMMR照合部106からの認識結果(応答)を受信する。サーバ302は、検索リクエストを処理し、その情報を、MMRゲートウェイ104の他の要素へ信号線330を介して送信し、これについては後述される。サーバ302は、信号線138によりMMR発行部108と、及び信号線136によりMMR照合部106と通信することに適合している。信号線138は、ホットスポットのウェブコンテンツをウェブサーバ304に送信する経路及び他の情報をサーバ302に提供する経路をMMR発行部108に提供する。一実施例では、サーバ302は、MMR発行部108から情報を受信し、その情報を信号線136を介してMMR照合部106に登録するために送信する。

10

20

30

40

50

【 0 0 4 1 】

ウェブサーバ304は、通常のタイプのものであり、クライアントからの要求（リクエスト）を受け付けること及びデータコンテンツと共に応答（レスポンス）を送信することの責務を有し、データコンテンツは、例えば、ウェブページ、文書及びリンクされたオブジェクト（画像等）である。ウェブサーバ304は、通常のデータベースのようなデータ格納部306に結合される。ウェブサーバ304は、信号線234を介して通信を行い、インターネットのようなネットワークを介して何らかの通信装置からのHTTPリクエストを受信する。ウェブサーバ304は信号線138にも結合され、上述したように、ホットスポットに関するウェブページを受信し、データ格納部306に格納し、HTTPリクエストに応じて検索及び送信を後に行うことに備える。ウェブサーバ304及びデータ格納部306をMMRゲートウェイ104の一部に含めることは、単なる一例に過ぎず、ウェブサーバ304がインターネットを介して移動装置102及びコンピュータ110にとってアクセス可能である限り、ウェブサーバ304及びデータ格納部306は如何なる代替的な場所又は形態で使用されてもよいことを、当業者は理解するであろう。

10

【 0 0 4 2 】

一実施例では、ポータルモジュール308は、サーバ302において動作するソフトウェア又はルーチンであり、ウェブポータルの作成及び提示を行うためのものである。ポータルモジュール308は、サーバ302と通信するために信号線330と結合される。一実施例では、ウェブポータルは、MMRゲートウェイ104の他の要素の管理及び保守を含む機能に関するアクセスポイントを提供する。他の実施例では、ウェブポータルは、MMR文書に関する経験による情報をユーザ達が共有できるエリア（場所）を提供する。さらに別の実施例では、ウェブポータルは、業務用アプリケーションや利用履歴310をユーザ達が共有できるエリアを提供する。

20

【 0 0 4 3 】

履歴（ログ）310はメモリ又はストレージ領域であり、そのメモリ又はストレージ領域は、サーバ302が移動装置302から受信した検索リクエスト、及びサーバ302が移動装置に送った対応するレスポンス総てのリストを保存する。別の実施例では、履歴310は、生成後にMMR照合部106に送信された問い合わせ画像、及びMMR照合部106から受信した認識応答結果のリストも保存する。履歴310は、サーバ302によるアクセスに備えて信号線330に結合される。

30

【 0 0 4 4 】

1つ以上の業務アプリケーション312は、MMR文書进行处理することに関する機能を提供するソフトウェア及びルーチンである。一実施例では、1つ以上の業務アプリケーション312は、サーバ302上で実行可能である。業務アプリケーション312は、検索要求の処理及び認識応答の配布に関する情報を使用することに適した様々なタイプの業務アプリケーションの内の何れでも良く、例えば、経理、グループウェア（groupware）、顧客関係の管理、人材、アウトソーシング、融資取組（loan origination）、顧客対応、サービス関係等に関するアプリケーションであるがこれらに限定されない。

【 0 0 4 5 】

認証モジュール314は、権限のあるユーザのリストを管理し及びMMRシステム100に対するアクセスを許可するためのソフトウェア及びルーチンである。一実施例では、認証モジュール314は、ある人に対応するユーザID及びパスワードのリストを維持し、その人は、システム100にアカウントを作成しており、それ故に、MMRゲートウェイ104及びMMR照合部106を使用して検索リクエストを処理することが、その人に認められている。認証モジュール314は、信号線330によりサーバ302に通信可能に結合されている。しかしながら、サーバ302が検索リクエストを受信した際、関連する問い合わせ画像を生成して信号線134により送信する前に、検索リクエストを処理して認証モジュール314中の情報と比較することができる。一実施例では、移動装置102が認証されなかった場合、移動装置がアカウントを設定していなかった場合、移動装置102のアカウントが不正使用又は未払い等の理由に起因して凍結されていた等の場合に

40

50

、移動装置 102 に返すサーバ 302 のメッセージを、認証モジュール 314 が生成する。

【0046】

経理モジュール 316 は、ユーザのアカウント及び MMR システム 100 の利用に関する財務処理を行うソフトウェア及びルーチンである。一実施例では、検索サービスが多種多様なビジネスモデル（経済モデル）の下で提供され、例えば、予約のモデル、検索リクエスト毎に課金するモデル又は他の様々な課金モデルの下で MMR システム 100 を利用することを含むが、これらに限定されない。一実施例では、MMR システム 100 は、多種多様な課金モデルを提供し、セルラ電話やデータネットワークについて現在提供されているものと同様である。経理モジュール 316 は、信号線 330 によりサーバ 302 に結合され、サーバ 302 が受信した何らかの検索リクエストの指示を受ける。一実施例では、経理モジュール 316 は、サーバ 302 により処理された取引記録（検索リクエスト/認識レスポンスの記録）を移動装置 102 各々について維持する。図示されていないが、経理モジュール 316 は、電子的な又は紙の請求書を発行する通常の課金システムに結合されていてもよい。

10

【0047】

メールモジュール 318 は、電子メール及び他の種類の通信を行うためのソフトウェア及びルーチンである。メールモジュール 318 は信号線 330 によりサーバ 302 に結合される。一実施例では、移動装置 102 は、あるコマンドを含む検索リクエストを発することができ、そのコマンドは、電子メール、ファクシミリ又は他の通常の電子通信手段を介して、文書、文書の一部又は他の情報を配ることを指示する。メールモジュール 318 は、そのような情報を生成し、MMR ゲートウェイ 104 から、ユーザが指定するアドレスへ送信する。一実施例では、各ユーザプロファイルは関連するアドレスを有し、それらのアドレスは、検索された情報を受け取る可能性のある受け手である。

20

【0048】

分析モジュール 320 は、MMR システム 100 のユーザの動作を分析するソフトウェア及びルーチンである。分析モジュール 320 は、MMR 照合部 106 によって行われた特徴抽出及び認識の有効性や精度を測るソフトウェア及びルーチンでもある。分析モジュール 320 は、以下のような MMR システム 100 の利用性を評価し、例えば、どの画像が検索リクエストの一部として最も頻繁に含まれているか、どのホットスポットデータが最も頻繁にアクセスされたか、画像が検索された順序、検索プロセスにおける最初の画像、MMR 動作を改善するのに使用される他の重要な実効性指標（パフォーマンスインジケータ）、及び/又はマーケティングキャンペーンの視聴者の反応等を評価する。一実施例では、分析モジュール 320 は、MMR システム 100 のメトリックを測定し、ホットスポット及び/又はホットスポットデータの有効性を測るのに使用されるメトリックを分析する。分析モジュール 320 は、サーバ 302、認証モジュール 314 及び経理モジュール 316 に信号線 330 を介して結合される。分析モジュール 320 は、サーバ 302 を介して信号線 134 にもつながっており、MMR 照合部 106 の要素にアクセスし、認識パラメータ、画像特徴、品質認識スコア、及び MMR 照合部 106 により生成又は使用される任意の他の情報を検索することができる。分析モジュール 320 は、ユーザのパラメータや好み、移動装置 102、ページ ID、ロケーション等に基づいて、様々なデータの検索や区分け（セグメンテーション）を実行することもできる。

30

40

【0049】

一実施例では、MMR ゲートウェイ 104 はホットスポットデータベース 404 も含む。図 3 ではホットスポットデータベース 404 が破線で示されているが、これは、ホットスポットデータベースを MMR ゲートウェイ 104 に含めることが代替例であることを反映している。ホットスポットデータベース 404 は、信号線 134 を介して認識応答を受信するように、信号線 436 により結合される。ホットスポットデータベース 404 は、これらの認識応答を使ってデータベースに問い合わせを行い、その認識応答に対応するホットスポットコンテンツを信号線 432 を介して出力する。ホットスポットコンテンツは

50

サーバ 302 に送られ、ホットスポットコンテンツが、認識応答に包含されるようにし、かつ要求を行っている移動装置 102 に送信されるようにする。

MMR 照合部 106

図 4 A 及び 4 B を参照しながら、MMR 照合部 106 に関する 2 つの例を説明する。MMR 照合部 106 の基本的な機能は、問い合わせ画像を受信すること、認識（検索）用の問い合わせ画像を送信すること、問い合わせ画像中の画像を認識すること、ホットスポット情報を検索すること、認識結果をホットスポット情報に結び付けること、及びそれを MMR ゲートウェイ 104 に返すことである。

【0050】

図 4 A は第 1 形態の MMR 照合部 106 を示す。第 1 形態の MMR 照合部 106 は、ディスパッチ部 402、ホットスポットデータベース 404、取得部 406、画像登録部 408 及び負荷バランサ 418 を有する。取得部 406 は、複数の認識部 410a - 410n 及び複数のインデックステーブル 412a - 412n をさらに有する。画像登録部 408 は、インデックス部 414 及びマスターインデックステーブル 416 をさらに有する。

【0051】

ディスパッチ部 402 は信号線 134 に結合され、MMR ゲートウェイ 104 から問い合わせ画像を受信し、MMR ゲートウェイ 104 に認識結果を送信する。ディスパッチ部 402 は、認識部 410a - 410n 各々に対する問い合わせ画像の割当及び送付を行う責務を有する。一実施例では、ディスパッチ部 402 は、問い合わせ画像を受信し、認識部の識別番号を生成し、その認識部の識別番号及び問い合わせ画像を取得部 406 に送信し、さらなる処理に備える。ディスパッチ部 402 は信号線 430 に結合され、認識部の識別番号及び問い合わせ画像を認識部 410a - 410n に送る。ディスパッチ部 402 は、信号線 430 を介して認識部 406 から認識結果も受信する。ディスパッチ部 402 の一例は、図 5 を参照しながらさらに説明される。

【0052】

図 3 を参照しながらホットスポットデータベース 404 の代替例が説明され、そこではホットスポットデータベースが MMR ゲートウェイ 104 の一部分であった。しかしながら、好適実施例のホットスポットデータベース 404 は、図 4 A に示されるように MMR 照合部 106 の一部分である。実施形態によらず、ホットスポットデータベース 404 は同様な機能を有する。ホットスポットデータベース 404 はホットスポット情報を格納するのに使用される。問い合わせ画像が認識され、認識結果が生成されると、これらの認識結果は、ホットスポットデータベース 404 の問い合わせ（クエリ）の一部として使用され、その問い合わせは、認識結果に関連するホットスポット情報を検索するためのものである。検索されたホットスポット情報は信号線 134 により MMR ゲートウェイ 104 に出力され、パッケージ化及び移動装置 102 への配布に備える。図 4 A に示されるように、ホットスポットデータベース 404 は、信号線 436 によりディスパッチ部 402 に結合され、認識結果を含む問い合わせを受信する。ホットスポットデータベース 404 は、問い合わせ結果を配布するために、信号線 432 及び信号線 134 を介して MMR ゲートウェイ 104 にも結合される。ホットスポットデータベース 404 は、信号線 136 にも結合され、保存する新しいホットスポット情報を MMR 発行部 108 から受信する。

【0053】

取得部 406 は、複数の認識部 410a - 410n 及び複数のインデックステーブル 412a - 412n を有する。認識部 410a - 410n の各々は、対応するインデックステーブル 412a - 412n を有し、それに結合されている。一実施例では、認識部 410 / インデックステーブル 412 のペア（対又は組）が、同じサーバ上にある。ディスパッチ部 402 は、問い合わせ画像を 1 つ以上の認識部 410a - 410n に送信する。冗長性を有する実施例の場合、問い合わせ画像は、認識及び検索用にディスパッチ部 402 から複数の認識部 410 へ送信され、インデックステーブル 412a - 412n は同じデータを索引付ける。順次的な実施例の場合、問い合わせ画像はディスパッチ部 402 から

10

20

30

40

50

第1の認識部410aに送信される。第1の認識部410aにおいて認識が成功しなかった場合、その問い合わせ画像は第2の認識部410bに送信され、以下同様に処理される。さらに別の実施例の場合、ディスパッチ部402は、問い合わせ画像について何らかの予備的な分析を行い、その問い合わせ画像を認識するのに最適な最も成功しやすい認識部410a - 410nを選択する。複数の認識部410a - 410n及び複数のインデックステーブル412a - 412nについて様々な構成例が存在することを、当業者は理解するであろう。取得部406の具体例は、図6A及び6Bを参照しながらさらに後述される。マスターインデックステーブル416からの破線によって示されているように、インデックステーブル412a - 412nは、様々な時点で更新されてもよい。

【0054】

画像登録部408は、インデックス部414及びマスターインデックステーブル416を有する。画像登録部408は、MMR発行部108からの更新情報を受信するために信号線136に結合された入力と、負荷バランサ418からの更新情報を受信するために信号線438に結合された入力とを有する。画像登録部408は、マスターインデックステーブル416を維持すること、及びマスターインデックステーブル416の全部又は一部を取得部406のインデックステーブル412a - 412nに移すことを行う責務を有する。一実施例では、インデックス部414は、画像、固有のページID及び他の情報を受信し、それらを、マスターインデックステーブルで保存するインデックステーブル情報に変換する。インデックス部414は、MMR発行部108と協同し、MMR発行部108により生成された画像ページにわたって一貫した固有のページ識別番号体系、マスターインデックステーブル416に保存されている画像ページ、及びホットスポットデータベース404の参照データに使用されるページ番号を管理する。

【0055】

画像登録部408の一例は、図7を参照しながらさらに詳細に図示及び説明される。

【0056】

負荷バランサ418は信号線430に結合され、ディスパッチ部402からの問い合わせ画像と、認識部406からの対応する認識結果とを受信する。負荷バランサ418の出力は、信号線438により、画像登録部408の入力に結合される。負荷バランサ418は、画像登録部408に対する入力を提供し、その入力は、取得部406のインデックステーブル412a - 412nを動的に調整するために使用される。特に、負荷バランサ418は、問い合わせ画像を監視及び評価し、その問い合わせ画像は、ディスパッチ部402から取得部406へ所与の期間の間に送信されたものである。利用度に基づいて、負荷バランサ418は、インデックステーブル412a - 412nを調整するための入力を提供する。例えば、負荷バランサ418は一日の問い合わせ画像を測定する。その日の測定された利用度に基づいて、インデックステーブルは、負荷バランサ418により測定された利用度に合うように、取得部406内で修正及び構築される。負荷バランサ418は、図11を参照しながら後述される方法を使用する。

【0057】

図4Bは第2形態のMMR照合部106を示す。第2形態では、MMR照合部106の多くの要素が、第1形態の対応する要素と同一又は類似する機能を有する。したがって、同一の又は同様の機能を備えた要素を指すために、同様な参照番号が使用されている。第2形態のMMR照合部106は、ディスパッチ部402、ホットスポットデータベース404及び負荷バランサ418を、第1形態のMMR照合部106と同様に含む。しかしながら、取得部406及び画像登録部408は、図4Aを参照しながら説明したものと異なる。特に、取得部406及び画像登録部408は、インデックステーブル及びマスターインデックステーブルについて共有SQLデータベースを利用する。より具体的には、マスターインデックステーブル416と、ローカルインデックステーブル412a - nを含むミラーデータベース(mirrored database)とがある。さらに、SQLデータベース複製の通常の機能を使用して、認識に使用するインデックステーブル412a - nに保存されているマスターインデックステーブル416のミラー画像を生成する。新たな画像がマス

10

20

30

40

50

ターインデックステーブル 4 1 6 に追加された場合、それらが総ての認識部 4 1 0 で速やかに利用可能になるように、画像認識部 4 0 8 は構築される。これは、大規模な R A M (図示せず) 及びデータベースのミラーリング技術を利用して、総てのローカルインデックステーブル 4 1 2 a - n にわたってマスターインデックステーブル 4 1 6 をミラーリングすることによってなされる。

ディスパッチ部 4 0 2

図 5 を参照するに一形態によるディスパッチ部 4 0 2 が示されている。ディスパッチ部 4 0 2 は、品質予測器 5 0 2、画像特徴オーダ部 5 0 4 及び分配部 5 0 6 を有する。品質予測器 5 0 2、画像特徴オーダ部 5 0 4 及び分配部 5 0 6 は、信号線 5 3 2 に結合され、MMR ゲートウェイ 1 0 4 から問い合わせ画像を受信する。分配部 5 0 6 は、品質予測器 5 0 2 及び画像特徴オーダ部 5 0 4 の出力を受けるようにも結合されている。分配部 5 0 6 は F I F O キュー 5 0 8 及びコントローラ 5 1 0 を含む。分配部 5 0 6 は信号線 5 3 4 上で出力を生成し、その出力は、問い合わせ画像及び認識部の識別番号 (R U I D) を含む。R U I D 以外の様々な手段を用いて、問い合わせ画像がどの特定の認識部に仕向けられてもよいことを、当業者は理解するであろう。問い合わせ画像が信号線 5 3 2 により受信されると、分配部 5 0 6 はその問い合わせ画像を受信し、それらが受信された順序でそれらを F I F O キュー 5 0 8 に入れる。コントローラ 5 1 0 は、各問い合わせ画像についての認識スコア (recognizability score) を品質予測器 5 0 2 から受信し、さらにオーダ信号を画像特徴オーダ部 5 0 4 から受信する。品質予測器 5 0 2 及び画像特徴オーダ部 5 0 4 からのこれらの情報を使用することで、コントローラ 5 1 0 は、F I F O キュー 5 0 8 から問い合わせ画像を選択し、それらを特定の認識部 4 1 0 に割り当て、その問い合わせ画像を割り当てられた認識部 4 1 0 に送信して処理に備える。コントローラ 5 1 0 は、確認部 4 1 0 に割り当てられた問い合わせ画像と、各画像について完了に予測される時間 (画像特徴オーダ部 5 0 4 により予測されるような時間) とのリストを保持する。各認識部 4 1 0 についてキューを空にするのに予測される総時間は、認識部に割り当てられた画像について予測された時間の総和である。コントローラ 5 1 0 は、いくつものキュー管理法を実行できる。最も簡易な割当法の場合、問い合わせ画像は、それらが到来した順序で F I F O キュー 5 0 8 から削除され、最初の利用可能な認識部 4 1 0 に割り当てられる。バランス応答法 (balanced response strategy) の場合、各問い合わせについて予測される総応答時間が均一な水準に維持され、問い合わせ画像をそれらが到来した順序で F I F O キュー 5 0 8 から削除し且つ認識部の F I F O キュー 5 0 8 に割り当ての際、予測される総応答時間が他の認識部とできるだけ近くなるようにする。イージーファースト法 (easy-fast strategy) の場合、画像は、それらの予測される完了時間によって決まる順序で F I F O キュー 5 0 8 から削除され、予測される最短の完了時間しか要しない画像が、最初の利用可能な認識部に割り当てられる。このように、ユーザが認識に容易な画像を提示した場合、ユーザは速やかな応答時間の恩恵を受ける。このことは、ユーザが提示する画像をユーザが注意深く選択することを促す。他のキュー管理法も可能である。ディスパッチ部 4 0 2 は、信号線 5 3 0 を介して認識部 4 1 0 から認識結果を受信する。認識結果は、ブル値 (真 / 偽)、真の場合はページ I D 及びページ中の場所の情報を含む。一実施例の場合、ディスパッチ部 4 0 2 はデータを単に受信し、データを MMR ゲートウェイ 1 0 4 に再送信する。

【 0 0 5 8 】

品質予測器 5 0 2 は、問い合わせ画像を受信し、ディスパッチ部 4 0 2 により使用される認識スコアを生成し、その問い合わせ画像を複数の認識部 4 1 0 の内の何れかにルーティングする。一実施例では、品質予測器 5 0 2 は、コンテキスト情報及び装置パラメータを入力として受信する。品質予測器 5 0 2 の動作は、図 1 8 ないし 2 1 を参照しながら後にさらに説明される。一実施例の場合、認識スコアは、有効な認識結果をもたらす最も確からしい認識アルゴリズムのタイプを指定する情報を含む。

【 0 0 5 9 】

10

20

30

40

50

画像特徴オーダ部504は、問い合わせ画像を受信し、オーダ信号を出力する。画像特徴オーダ部504は、入力された問い合わせ画像を分析し、そして、それが含む画像特徴を分析することで画像を認識するのに要する時間を予測する。実際の認識時間と予測された時間との間の相違は、将来的な予測の仕方を調整するのに使用され、それにより精度を改善する。最も簡易な例の場合、僅かな特徴しか伴っていない簡易な画像は、軽い負荷の認識部410に割り当てられ、それらの画像が速やかに認識されかつユーザが回答を速やかに得られるようにする。一実施例では、画像特徴オーダ部504が時間を予測するのに使用する特徴は、実際に認識する認識部410により使用される特徴と異なる。例えば、画像の中で検出されるコーナー（corner）の数が、画像分析に要する時間を予測するのに使用される。予測に使用される特徴群は、実際の認識時間と関連付けられる必要があるだけである。一実施例では、いくつかの異なる特徴群が使用され、いくつかの期間に関して測定された認識時間に関連付けられる。そして、最良の予測及び最低のコスト（最高の効率）の特徴群が判別され、他の特徴群は破棄される。画像特徴オーダ部504の動作は、図12を参照しながら後にさらに説明され、さらに理解できるであろう。

10

取得部406

図6A及び6Bを参照しながら、取得部406の一例を説明する。

【0060】

図6Aは取得部406の一例を示し、認識部410及びインデックステーブル412のペアは、それらが索引付けするコンテンツ又は画像に基づいて区分けされている。この構成は、定期的にコンテンツを提供するマスメディア出版に特に有利である。インデックステーブル412内のコンテンツの組織は或る方法で区分けすることができ、その方法は、最もアクセスされやすいコンテンツが、認識部410及びインデックステーブル412のペアの多数で利用可能であるようにする。以下で説明される区分けは単なる一例に過ぎないこと、及び時間とともに測定された実際の利用統計に基づく他の様々な区分けが使用可能であることを、当業者は認識するであろう。図6Aに示されるように、取得部406は、複数の認識部410a-h及び複数のインデックステーブル412a-hを有する。複数の認識部410a-hは信号線430に結合され、ディスパッチ部402からの問い合わせ画像を受信する。複数の認識部410a-hの各々は、対応するインデックステーブル412a-hに結合される。認識部410は、問い合わせ画像から特徴を抽出し、その画像特徴をインデックステーブルに格納済みの特徴と対比し、一致するページ及びそのページ中の位置を特定する。認識及び検索システム並びに方法の具体例については、例えば以下の出願で説明されており、これらの出願の内容全体が本願のリファレンスに組み入れられる。

20

30

【0061】

西暦2006年7月31日付け出願の“System And Methods For Creation And Use Of A Mixed Media Environment”と題する米国特許出願第11/461,017号；

西暦2006年7月31日付け出願の“Method And System For Image Matching In A Mixed Media Environment”と題する米国特許出願第11/461,279号；

西暦2006年7月31日付け出願の“Method And System For Document Fingerprinting Matching In A Mixed Media Environment”と題する米国特許出願第11/461,286号；

40

西暦2006年7月31日付け出願の“Method And System For Position-Based Image Matching In A Mixed Media Environment”と題する米国特許出願第11/461,294号；

西暦2006年7月31日付け出願の“Method And System For Multi-Tier Image Matching In A Mixed Media Environment”と題する米国特許出願第11/461,300号；

西暦2006年7月31日付け出願の“Data Organization and Access for Mixed Media Document System”と題する米国特許出願第11/461,147号；

西暦2006年7月31日付け出願の“Database for Mixed Media System”と題する米国特許出願第11/461,164号；

50

西暦2006年7月31日付け出願の“Searching Media Content For Objects Specified Using Identifiers”と題する米国特許出願第11/461,109号;

西暦2008年3月31日付け出願の“ Invisible Junction Feature Recognition For Document Security Or Annotation ”と題する米国特許出願第12/059,583号;

西暦2008年5月15日付け出願の“ Web-Based Content Detection In Images, Extraction And Recognition ”と題する米国特許出願第12/121,275号;

西暦2007年7月11日付け出願の“ Invisible Junction Features For Patch Recognition ”と題する米国特許出願第11/776,510号;

西暦2007年7月11日付け出願の“ Information Retrieval Using Invisible Junctions and Geometric Constraints ”と題する米国特許出願第11/776,520号;

10

西暦2007年7月11日付け出願の“ Recognition And Tracking Using Invisible Junctions ”と題する米国特許出願第11/776,530号;

西暦2007年7月12日付け出願の“ Retrieving Documents By Converting Them to Synthetic Text ”と題する米国特許出願第11/777,142号;及び

西暦2007年1月18日付け出願の“ Synthetic Image and Video Generation From Ground Truth Data ”と題する米国特許出願第11/624,466号。

【 0 0 6 2 】

図6Aに示されているように、認識部410/インデックステーブル412のペアは、インデックステーブル412中の内容に応じてグループ化されている。特に、認識部410a-d及びインデックステーブル412a-dの第1グループ612は、今日の新聞のような刊行物のページを索引付けるのに使用される。例えば、8つの認識部410の内の4つが、今日の新聞の内容(コンテンツ)を索引付けるのに使用される。なぜなら、検索要求のほとんどは、直近の24時間以内に発行された新聞に関連する傾向があるからである。認識部410e-g及びインデックステーブル412e-gの第2グループ614は、例えばここ1週間のような最近の過去の日付の新聞ページを格納するのに使用される。認識部410h及びインデックステーブル412hの第3グループ616は、例えば昨年のようなさらに古い日付の新聞ページを格納するのに使用される。これは、取得部406の組織構造が、受信した検索リクエストの性質に合うように最適化されることを可能にする。さらに、所与の問い合わせ画像が認識用の第1グループ612に先ず送信され、第1グループ612がその問い合わせ画像を認識できなかった場合、それが認識用の第2グループ614に送信され、以下同様に続くように、取得部406の動作が修正されてもよい。

20

30

【 0 0 6 3 】

認識部410及びインデックステーブル412の4つを第1グループ612として使用することは、単なる一例に過ぎず、第2グループ614及び第3グループ616の認識部410及びインデックステーブル412の数に対する相対的な例を示すに過ぎないことに、留意すべきである。特定のどのグループ612、614及び616における認識部410及びインデックステーブルの数も、認識部410及びインデックステーブル412の総数に基づいて比率を変えてよい。さらに、所与の刊行物に関して取得部406に検索リクエストを送っている全ユーザーの特徴に合うように、特定のどのグループ612、614及び616における認識部410及びインデックステーブルの数も適合させてよい。

40

【 0 0 6 4 】

図6Bは、第2形態による取得部406を示し、認識部410及びインデックステーブル412は、それらが使用する認識アルゴリズムのタイプに応じて区分けされている。第2形態の場合、特定の認識部が認識リクエストを生成することに失敗すると、入力された問い合わせ画像が、処理用の別の認識部に送付されることを引き起こすように、認識部410は結合されている。さらに第2実施例では、インデックステーブル412は、様々な装置及び画像捕捉装置の環境要因(例えば、ぼやけ、不鮮明さ)に応じて異なる特徴群を含む。

【 0 0 6 5 】

50

第2形態の取得部406は、複数の認識部410a - 410e、複数のインデックステーブル412a - 412e、及び結果合成部610を含む。この例の場合、認識部410a - 410eの各々は異なるタイプの認識アルゴリズムを使用する。例えば、認識部410a、410b及び410cはインビジブルジャンクション(Invisible junction)アルゴリズムを、認識部410dはブリックウォールコーディング(Brick Wall Coding)アルゴリズムを、認識部410eはパスコーディング(Pass Coding)アルゴリズムを使用して、ページ番号及び場所の認識及び検索を行う。認識部410a、410b及び410cの各々は、信号線630を介して問い合わせ画像を受信するように結合された入力

を有する。複数の認識部410a - 410e各々からの認識結果は、信号線636、638、640、642及び644を介して結果合成部610に送られる。合成部610の出力結果は信号線430に与えられる。

10

【0066】

一実施例では、認識部410a、410b及び410cはインデックステーブル1、2及び3(412a - 412c)と協働し、そのインデックステーブル各々は、同じページの画像特徴であるが、装置及び環境要因に起因する不鮮明さの程度が異なる画像特徴を保存している。例えば、インデックステーブル1(412a)は、PDF文書から得られるような元の画像ページの画像特徴を保存する一方、インデックステーブル2(412b)は、第1水準の不鮮明さを伴う同じページの画像を保存し、インデックステーブル3(412c)は、第2水準の不鮮明さを伴う同じページの画像を保存する。一実施例の場合、インデックステーブル1、2及び3(412a - 412c)は量子化ツリーになっている。

第1認識部410aは信号線630を介して問い合わせ画像を受信する。第1認識部410aは、インビジブルジャンクション特徴抽出部602及び検索部604aを有する。インビジブルジャンクション特徴抽出部602は、問い合わせ画像を受信し、インビジブルジャンクション特徴(不可視のつながりに関する特徴)を取り出し、それらを検索部604aに与える。検索部604aは、取り出されたインビジブルジャンクション特徴を使用して、それらをインデックステーブル1(412a)と対比する。検索部604aが一致を確認すると、検索部604aは、その認識結果を信号線636を介して結果合成部610に送る。しかしながら、検索部604aが一致を確認できなかった場合、或いは確実性に乏しい一致しか確認できなかった場合、検索部604aは、取り出されたインビジブルジャンクション特徴を、信号線632を介して第2認識部410bの検索部604bに送信する。インビジブルジャンクション特徴は既に抽出されているので、第2認識部410bはインビジブルジャンクション特徴抽出部602を要しないことに留意を要する。第2認識部410bは第2認識部410aと同様な検索機能を実行するが、その際、僅かにぼけた画像に対するインビジブルジャンクション特徴を有するインデックステーブル2(412b)を使用する。検索部604bが一致を確認すると、検索部604bは、その認識結果を信号線636を介して結果合成部610に送る。第2認識部410bの検索部604bが一致を確認できなかった場合、或いは確実性に乏しい一致しか確認できなかった場合、検索部604bは、取り出されたインビジブルジャンクション特徴を、信号線634を介して第3認識部410cの検索部604cに送信する。そして、検索部604cは同様な検索機能を実行するが、その際インデックステーブル3(412c)を使用する。

1つの元の画像特徴及び2段階の不鮮明さの組み合わせが使用されているが、これは単なる一例に過ぎず、0からnに及ぶ追加的なレベル(段階)が幾つでも使用されてよいことを、当業者は理解するであろう。

20

30

40

【0067】

認識部410d及び410eは、他の認識部410a - cと並列的に動作する。認識部410dは、ブリックウォールコーディング特徴抽出部606及び検索部604dを有する。ブリックウォールコーディング特徴抽出部606は、問い合わせ画像及び境界枠(bounding box)を受信し、その境界枠を分析し、ブリックウォールコーディング特徴を生成する。これらのブリックウォールコーディング特徴は検索部604dに与えられ、検索部604dはそれらをインデックステーブル4(412d)に保存されている

50

特徴と比較する。一実施例では、インデックステーブル4(412d)はハッシュテーブルである。検索部604dは、何らかの一致するページを特定し、その認識結果を信号線642を介して結果合成部610に返す。第5認識部410eは同様に動作するが、その際パスコーディングを使用する。第5認識部410eは、パスコーディング特徴抽出部608及び検索部604eを有する。パスコーディング特徴抽出部608は、問い合わせ画像及び境界枠を受信し、その画像を分析し、検索部604eに与えるパスコーディング特徴を生成し、検索部604eはそれらをインデックステーブル5(412e)に保存されている特徴と比較する。一実施例では、インデックステーブル5(412e)は、キャラクタ文字列のSQLデータベースである。検索部604eは、何らかの一致する文字列を特定し、その認識結果を信号線644を介して結果合成部610に返す。

10

【0068】

結果合成部610は、複数の認識部410a-410eからの認識結果を受信し、合致する結果に関する1つ又は少数のリストを生成する。一実施例の場合、認識結果の各々は、関連する信頼性因子(confidence factor)を含む。別の実施例では、日付、時間、場所、個人的プロフィール又は検索履歴のようなコンテキスト情報が、結果合成部610に与えられる。これらの信頼性因子は他の情報と共に結果合成部610により使用され、入力された問い合わせ画像に最も確からしく一致する認識結果を選択する。結果合成部610は、後述の図22ないし24の方法による機能を実行する。

画像登録部408

20

図7は、一形態による画像登録部408を示す。画像登録部408は、不鮮明度生成部702、複数のインビジブルジャンクション特徴抽出部704a-c、複数のインビジブルジャンクションインデックステーブル更新部706a-c、ブリックウォールコーディング特徴抽出部708、ブリックウォールコーディングインデックステーブル更新部710、パスコーディング特徴抽出部712、パスコーディングインデックステーブル更新部714及び複数のマスターインデックステーブル416a-eを有する。画像登録部408は他の制御論理装置(図示せず)も含み、その制御論理装置(制御ロジック)は、マスターインデックステーブル416により、ワーキングインデックステーブル412の更新を制御する。画像登録部408は、様々な基準に基づいて様々な方法で取得部406のインデックステーブル412を更新することができ、例えば、定期的に更新を実行してもよいし、新たな内容が追加された場合に更新を実行してもよいし、利用度に応じて更新を実行してもよいし、記憶効率の観点から更新を行ってもよい。

30

【0069】

不鮮明度生成部702は、信号線730を介して画像及びページ識別番号を受信するように結合された入力を有する。不鮮明度生成部702は複数の出力を有し、各出力は、信号線732、734及び736によりインビジブルジャンクション特徴抽出部704a-cにそれぞれ結合される。不鮮明度生成部702は、元の画像及びページ識別番号を出力の信号線732に通す。不鮮明度生成部702は、第1水準の不鮮明さを伴う画像を生成し、それとページ識別番号とをインビジブルジャンクション特徴抽出部704bに至る信号線734に出力し、不鮮明度生成部702は、第2水準の不鮮明さを伴う別の画像を生成し、それとページ識別番号とをインビジブルジャンクション特徴抽出部704cに至る信号線736に出力する。不鮮明度生成部702は、図13及び14に関連して後述される機能を実行する。

40

【0070】

インビジブルジャンクション特徴抽出部704は、画像及びページIDを受信し、その画像からインビジブルジャンクション特徴を抽出し、ページIDと共にそれらを、各自のインビジブルジャンクションインデックステーブル更新部706に送る。複数のインビジブルジャンクション特徴抽出部704a-cの出力は、複数のインビジブルジャンクションインデックステーブル更新部706a-cの入力に結合される。例えば、インビジブルジャンクション特徴抽出部704aの出力は、インビジブルジャンクションインデックス

50

テーブル更新部 706 a の入力に結合される。残りのインビジブルジャンクション特徴抽出部 704 b - c も同様に各自のインビジブルジャンクションインデックステーブル更新部 706 b - c に結合される。インビジブルジャンクションインデックステーブル更新部 706 は、抽出された特徴のフォーマットを整えること、及びそれらに対応するマスターインデックステーブル 416 に保存することの責務を有する。マスターインデックステーブル 416 は 5 つの別々のマスターインデックステーブル 416 a - e として図示されているが、総てのマスターインデックステーブルが 1 つのマスターインデックステーブルに統合されてもよいこと、或いは少数のマスターインデックステーブルに統合されてもよいことを、当業者は認めるであろう。インビジブルジャンクションインデックステーブル更新部 706 が、抽出された特徴をインデックステーブル 416 に保存すると、それらの更新部は確認信号を発行し、確認信号は信号線 740 及び 136 を介して MMR 発行部 108 に送り返される。

10

【0071】

ブリックウォールコーディング特徴抽出部 708 及びパスコーディング特徴抽出部 712 は、同様に機能し、画像、ページ識別番号及び境界枠情報を受信するように信号線 738 に結合される。ブリックウォールコーディング特徴抽出部 708 は、関連するインデックステーブル 416 d を更新するのに必要な情報を入力から抽出する。ブリックウォールコーディングインデックステーブル更新部 710 は、ブリックウォールコーディング特徴抽出部 708 から抽出された情報を受信し、それをインデックステーブル 416 d に保存する。パスコーディング特徴抽出部 712 及びパスコーディングインデックステーブル更新部 714 は、同様に動作するが、パスコーディングを使用する。パスコーディング特徴抽出部 712 も、画像、ページ番号及び境界枠情報を信号線 738 を介して受信する。パスコーディング特徴抽出部 712 は、パスコーディング情報を抽出し、それをパスコーディングインデックステーブル更新部 714 にわたす。パスコーディングインデックステーブル更新部 714 は、その情報をインデックステーブル 5 (416 e) に保存する。登録部 408 のこのアーキテクチャは、ある環境を提供する点で特に有利であり、その環境では、MMR 発行部 108 は、画像及びページ番号を画像登録部 408 に単に与えることでインデックステーブルを自動的に更新でき、インデックステーブルは自動的に更新される。

20

30

MMR 発行部 108

図 8 には、一形態による MMR 発行部 108 が示されている。MMR 発行部 108 は、従来の出版用ソフトウェア 802、登録及び注釈用のプリプレスソフトウェアプラグイン 804、及びホットスポット生成部 806 を有する。MMR 発行部 108 は、画像登録部 408 及びホットスポットデータベース 404 に（例えば、サーバ 302 を介して）つながるように結合されている。

【0072】

従来の出版用ソフトウェア 802 は、文字（テキスト）、図形及び画像を含むデジタルコンテンツを生成するソフトウェアである。新聞の生成及び拡布に使用される従来の出版用ソフトウェアの具体例は、カリフォルニア州サンノゼのアドビインコーポレーションにより製造販売されているアドビアクロバット（Adobe Acrobat）や、コロラド州デンバーのクオークインコーポレーションにより製造販売されているクオークイクスプレス（QuarkXPress）等がある。従来の出版用ソフトウェア 802 は、大手出版社が所有する他のタイプの専用ソフトウェアでもよい。

40

【0073】

登録及び注釈用のプリプレスソフトウェアプラグイン 804 は、従来の出版用ソフトウェア 802 との間でインターフェース機能を発揮し、かつ画像登録部 408 で使用される追加的な情報を生成するソフトウェア及びルーチンであり、その追加的な情報によりコンテンツは取得部 406 に自動的にアップロードできるようになる。例えば、登録及び注釈用のプリプレスソフトウェアプラグインは、画像登録部 408 と通信するインターフェー

50

ス機能を発揮し、特定のページ識別番号体系を維持し、そのページ識別番号体系は、従来の出版用ソフトウェア 802 により生成された画像ページ、マスターインデックステーブル 416 に保存されている画像ページ、及びホットスポットデータベース 404 の参照データで使用されているページ番号の間で矛盾無く整合している。登録用のプリプレスプラグイン 804 は、従来の出版用ソフトウェア 802 によって生成された画像ページの境界枠も生成し、その境界枠は、ある種のエンコーディング用に画像登録部 408 により以後使用される。登録及び注釈用のプリプレスソフトウェアプラグイン 804 は、従来の出版用ソフトウェア 802 から与えられた画像を分析し、ファイルのタイプが特徴抽出部 704、708 及び 712 で利用可能であることを保証する。

【0074】

ホットスポット生成部 806 は、ホットスポットを生成するオーサリングツール (authoring tool) である。ホットスポットは、文書画像中のポリゴンの (x, y) 座標を特定し、各ポリゴンに関連する URL のような電子データをリンクする。一実施例の場合、ホットスポットはアクトバット PDF ファイルの注釈であり、文書画像中のポリゴンの (x, y) 座標を含み、URL のような電子データをリンクする。ホットスポットを生成すること (「ホットスポット (hotspot)」は画像における多角形である)、ホットスポットを文書中の特定のページやページ中の特定の場所に関連付けること、ゼロ個以上のリンクを各ホットスポットに関連付けること、そしてホットスポットファイルを生成してサーバ 302 に送信し、ホットスポットデータベース 404 に保存すること等のために、ホットスポット生成部 806 は使用可能である。ホットスポットファイルは、各々のホットスポット、ポリゴンの座標、及びそれに関連する何らかのリンクをリスト化する。ホットスポットファイルは、別個のファイルとして保存されてもよいし、或いは等価な同じデータが文書ソースファイルに組み込まれてもよい。ホットスポット生成部 806 は、ワード (Word)、PDF 又はティフ (tiff) 等の画像フォーマットのようなソース文書を読み取り、ユーザインターフェース機能を発揮するスタンドアローンアプリケーションであり、ユーザが、文書上で多角形 (ポリゴン) を描くこと、ホットスポット各々をリンクさせること及びホットスポットファイルを保存することを可能にする。ホットスポット生成部 806 は、ワード (Word)、アクトバット (Acrobat)、クオークイクスプレス (Quark Express) 等のような既存の文書オーサリングアプリケーション用のプラグインとしてパッケージ化することが可能であり、かつユーザが文書上でポリゴンを規定し、リンクを加え、そしてホットスポットファイル (或いは、組み込まれたホットスポットデータを伴う等価的なソースファイル) を保存できるように設計可能である。ホットスポット生成部 806 は、あるソフトウェアアプリケーションとしてパッケージ化することができ、そのアプリケーションは既存の文書からリンクを抽出してそれらをホットスポットファイルに保存する。例えば、ポリゴン座標を含み且つ電子データにリンクする PDF ファイルの注釈が、PDF ファイルから除去され、ホットスポットファイルに別途保存することができる。

方法

図 9 は、検索リクエストを生成及び送信し、MMR システム 100 と共に検索リクエストを処理する一般的な方法のフローチャートを示す。本方法は、移動装置 102 が画像を捕捉することから始まる (902)。画像、ユーザ識別子及び他のコンテキスト情報を含む検索リクエストが移動装置 102 により生成され、MMR ゲートウェイ 104 に送信される (904)。MMR ゲートウェイ 104 は、検索リクエストからユーザ識別子を抽出し、それが正当なユーザに関連することを確認することで、その検索リクエストを処理する (906)。MMR ゲートウェイ 104 は他の処理も実行し、その処理は、履歴 (ログ) 310 に検索リクエストを記録すること、検索リクエストに関して必要な精算処理を実行すること、及び MMR 分析メトリックを分析すること等である。次に、MMR ゲートウェイ 104 は、問い合わせ画像を生成し、それをディスパッチ部 402 に送信する (904)。ディスパッチ部 402 は、負荷バランスをとる処理を実行し、問い合わせ画像を取得部 406 に送信する。一実施例では、ディスパッチ部 402 は、問い合わせ画像を処理

10

20

30

40

50

すべき取得部 406 の特定の認識部 410 を指定する。取得部 406 は画像認識処理を実行し、認識結果を生成する (912)。認識結果は、ディスパッチ部 402 へ返され、ひいては MMR ゲートウェイ 104 に返される。認識結果は、その認識家かで特定されているページ及び場所に対応するホットスポットを検索するのに使用される (916)。そして、ホットスポットデータ及び認識結果は、MMR ゲートウェイ 104 から移動装置 102 へ送信される (918)。

【0075】

図 10 を参照しながら、画像を登録する方法を説明する。本方法は、従来の出版用ソフトウェア 802 を用いて画像を生成することで開始される (1002)。登録及び注釈用のプリプレスソフトウェアプラグイン 804 を用いて、画像は、境界枠、ホットスポットデータ及びページ識別番号と共に補足 (補強) される (1004)。補足された画像は、MMR ゲートウェイ 104 に送信される (1006)。MMR ゲートウェイ 104 は、ホットスポットデータを抽出し、それをホットスポットデータベースに加える。MMR ゲートウェイ 104 は、画像、ページ番号及び境界枠を画像認識部 408 に送信する (1008)。図 10 で破線によって示されている代替例では、補足された画像は、MMR ゲートウェイ 104 を通らずに画像登録部 408 に送られてもよい。そして、インデックス部 414 は、その画像について特徴抽出処理を実行する。実施形態に依存して、特徴抽出処理は、1つの認識アルゴリズムや複数の異なる認識アルゴリズムを使用してもよいし、異なるデータセットに同じ認識アルゴリズムを使用してもよい。ステップ 1010 により抽出された特徴は、マスターインデックステーブル 416 a - e を更新するのに使用される。そして、マスターインデックステーブル 416 a - e に適用された変更は、取得部 406 のワーキングインデックステーブル 412 a - n に反映される。画像登録部 408 に画像が与えられることだけを必要とし、画像登録部 408 は、マスターインデックステーブル 416 を自動的に更新し、その変更内容を、取得部 406 のワーキングインデックステーブル 412 に自動的に移せる点で、本方法は特に有利である。

【0076】

図 11 を参照しながら、動的に負荷バランスをとる方法を説明する。本方法は、概して、図 4A の負荷バランサ 418 により実行される処理に関連する。ディスパッチ部 402 及び取得部 406 間の信号線 430 上で伝送される問い合わせ画像及び認識結果を監視することから本方法は始まる (1102)。負荷バランサ 418 は、ある所定の期間 (例えば、1日) の間、問い合わせ画像及び対応する認識結果を監視する。良好に発見された画像の実際の流通及びそれらが発見されたインデックステーブルが監視される。負荷バランサ 418 は負荷バランス制御信号を生成し (1104)、その信号は、取得部 406 のワーキングインデックステーブル 412 a - n のコンフィギュレーションを指定する。例えば、前の日について負荷バランサ 418 により測定された利用度に基づいて、期待される利用度が予測され、取得部 406 のワーキングインデックステーブル 412 a - n に適切なコンテンツの負荷がかかるようにする。例えば、n 個のインデックステーブルに本日の新聞のコンテンツが積み込まれ、昨日の新聞のコンテンツが m 個のインデックステーブルに積み込まれていたとする。この場合、合計 n + m 個のサーバがある。過去の履歴に基づいて、本発明は、本日の新聞について $n / (n + m)$ 回の問い合わせを予測する。画像登録部 408 のインデックス部 414 は、本日の新聞について新たな画像を受信する。ステップ 1104 で生成された負荷バランス信号に基づいて、ワーキングインデックステーブル 412 a - n の内の幾つか (例えば、n 個) に新たな画像が積み込まれ、ワーキングインデックステーブル 412 a - n 中の既存の情報は、新たな画像を保存するのに使用されていない残りのワーキングインデックステーブル 412 a - n (例えば、 $m = (\text{全ワーキングインデックステーブル}) - n$ 個) に保存できるように再構築される。

【0077】

一実施例では、取得部 406 の処理は、ワーキングインデックステーブル 412 a - n のコンフィギュレーションに合うように修正される又は最適化される。特に、取得部 406 は、問い合わせ画像を受信すると、まず、本日の新聞の画像と共に構築されたインデッ

10

20

30

40

50

クステーブル412及び認識部410を認識しようとする(1108)。次に、本方法は認識が成功したか否かを判定する(1110)。成功した場合、本方法は結果を返し(1118)、本方法は終了する。成功しなかった場合、本方法は、過去の新聞の画像と共に構築されたインデックステーブル412及び認識部410と共に問い合わせ画像を認識しようとする(1112)。再び、本方法は認識の成否を判定する(1114)。成功した場合、本方法は結果を返し(1118)、本方法は終了する。成功しなかった場合、本方法は、認識が成功しなかったことを示す信号を返し(1116)、本方法は終了する。新聞のデータを周期的に分割する観点から、負荷バランス法が説明されたが、任意の他の利用度の基準に基づいてデータを分割することに負荷バランス法が適用されてもよいこと、及び画像データは新聞以外の任意のタイプの画像でもよいことを、当業者は理解するであ

10

【0078】

図12を参照しながら、画像特徴に基づく順序づけを行う方法の一形態を説明する。本方法の機能は、概して、ディスパッチ部402の画像特徴オーダ部504により発揮される。本方法は、問い合わせ画像を受信することから始まる(1202)。次に、ディスパッチ部402の画像特徴オーダ部504は、問い合わせ画像中の画像特徴を分析する(1204)。認識ステップ(1204)で使用される画像特徴は、認識部410で使用される画像特徴と必ずしも同じでなくてよいことに留意を要する。認識する画像特徴と関連していさえすればよい。さらに別の例の場合、いくつもの異なる特徴群が使用され、時間をかけて相関性が測定される。最終的には、最良の予測をもたらし、かつ最低の演算コストしかかからない特徴群が決定され、他の特徴群は破棄される。画像特徴オーダ部504は、画像特徴を認識するのに要する時間を測定し(1206)、それにより予測時間を算出する。次に、本方法は、特徴と予測時間との間に対応関係をつける(1208)。次に、本方法は、問い合わせ画像を認識する認識部406が実際にかけた時間を測定する(1210)。認識部406がかけたその時間は、「実時間」として言及される。そして、画像特徴オーダ部504は、ステップ1208で生成された対応関係を、その実時間で調整する(1212)。調整された対応関係は、問い合わせ画像を並べ替えたり、認識部に割り当てたりするのに使用される(1214)。例えば、僅かな特徴しか備えていない簡易な画像は、負荷の軽いサーバ(認識部410及びインデックステーブル412のペア)に割り当てられ、それらが速やかに認識され、ユーザが回答を速やかに受け取れるようにする。図12に示される方法は、1つの画像又は小さな画像群に関するプロセスを説明している。しかしながら、多くの画像が上記の方法で処理された場合、多数の対応関係が生成され、画像特徴オーダ部504は、処理時間の観点から画像特徴の分配法を先ず求め、分配部506のコントローラ501は、その分配法を使って負荷のバランスをとり、それに応じて、特定の画像特徴を伴う問い合わせ画像を仕向け直すことができることを、当業者は認識するであろう。対応関係が調整された後(1212)、新たな画像の特徴オーダ指標が、説明済みのキュー管理法で速やかに使用可能になる。最も簡易な割当法は、単に新たな予測を使用するものであろう。バランス応答法は、各認識部のリクエストのキューのバランスを取り直し、キューの間でリクエストを選択的に動かし、認識部410各々について一様な予測応答を維持する。イージーファースト法は、到来するFIFOキュー508

20

30

40

不鮮明画像

図13ないし17を参照しながら、画像捕捉装置の特性にMMRシステム100を自動的に適合させる方法を説明する。認識システムは、認識すべきオブジェクトの元画像になじむようにしばしばトレーニングされる。しかしながら、何百もの捕捉装置がある場合や、システムが使用される環境が多くの異なる画像を生成する場合がある。図13ないし17は、当初は元のオブジェクト画像用に設計されたMMRシステム100を、一般的に使用される様々な捕捉装置で生成された画像に自動的に適合させる方法を示す。基本的には、本方法は、認識アルゴリズムのN個のインスタンスをN個の異なるデータセットに合わ

50

せる。装置トレーニング (DT: Device Training) 画像群を良好に認識するように認識部が生成した N 個のデータセットが認識部の能力に基づいて選択され、その DT 画像群は、実際に使用された場合に装置が生成する画像を表現する。そして、それらの 1 つが結果を生成するまで、問い合わせ画像は各インスタンスと共に認識される。これらの調整は、MMR システム 100 の認識精度を大幅に改善する。本発明は、MMR システムをばやけに対して合わせる状況の観点から説明されるが、ばやけ (不鮮明さ) は、本発明原理を説明する具体例として使用された 1 つの特性に過ぎない。画像捕捉装置の如何なる特性も、その特性に関する実際の変形例、生成されたデータセット、及びそのような特性の変形例を伴う画像を認識するように適合させられたシステム 100 をシミュレートするように代替的にモデル化可能である。

10

【0079】

図 13 は、移動装置 102 により導入された不鮮明さに対して MMR システム 100 を合わせる一般的な方法例を示す。本プロセスは、装置のトレーニング及びデータ捕捉から始まる (1302)。装置のトレーニングの一例は、図 14 を参照しながら以下でさらに説明される。トレーニングデータ群が生成されると、本方法は初期化を行い、分類子を生成する (1304)。初期化及び分類子生成 (1304) の具体例は、図 15 A 及び 15 B を参照しながら以下で説明される。一群の分類子が生成されると、本方法は、最多数の画像を適切に認識する分類子を特定する (1306)。一実施例では、これは、図 16 を参照しながら後述されるグリーディ探索法又は貪欲法 (Greedy search) を実行することで達成可能である。発見された分類子が仮定されて検査される (1308)。言い換えれば、分類子は、認識可能な画像数を増やすように修正され、それらの修正が有効化される。この検査法の一例は、図 17 を参照しながら後述される。そして、本方法は、インデックステーブル 412 及び認識部 410 を、仮定され検査されたその分類子に基づいて更新する (1310)。別の実施例の場合、図 13 の方法は認識アルゴリズムの実効性を監視する方法と組み合わせられ、認識アルゴリズムの新たなインスタンスが追加されるべきか否か、或いは現在のインスタンスが図 13 の方法を実行することで置換されるべきか否かを判定する。

20

【0080】

図 14 を参照するに、トレーニングデータ群を生成する方法の一例を説明する。一般に、トレーニングデータ群を生成することは、配置属性の範囲にわたるデータ群と、オブジェクトに関する環境因子とを生成することを要する。本方法は、配置属性の範囲にわたる装置トレーニング (DT) データ群と、オブジェクトに関する環境因子とを収集する。配置属性及び環境因子各々についての指標を伴う画像各々が、そのデータ群に加えらる。

30

【0081】

本方法は、画像捕捉装置、オブジェクト群 (O)、環境属性群 (E) 及び装置配置属性群 (P) を決定することから始まる (1402)。本発明は以下の事項が特定されていることを仮定している：所与の捕捉装置、代表的な一群のオブジェクト (例えば、インビジブルジャンクションインの場合における紙文書)、一群の環境因子 (例えば、照明、影、オブジェクトの変形、温度等)、及び一群の配置属性 (例えば、距離、回転、傾斜 (チルト)、ジッタ (jitter) 等)。例えば、所与の捕捉装置は、カメラを備えた商業的に入手可能なセルラ電話でもよく、例えば、アップルコンピュータ社の iPhone (iPhone) や、リサーチインモーション社のブラックベリー (Blackberry) 等でもよい。オブジェクト群は、各ページに異なる X - Y 座標を有する新聞紙のページでもよい。環境属性及び関連する値は、例えば、光 = 蛍光、日光等；影 = 薄い、鮮明等である。配置属性は、距離 = 1"、2"、. . . ; チルト = 0 度、5 度、7 度、. . . 等である。

40

【0082】

次に、本方法は、オブジェクト群 (O) から或るオブジェクト (o) を取り出し (1404)、環境属性群 (E) から或る環境属性 (e) を取り出し (1406)、装置配置属性群 (P) から或る装置配置属性 (p) を取り出す (1408)。そして、その装置配置属性 (p) の各値について、その環境属性 (e) におけるオブジェクト (o) の画像を、

50

本方法は捕捉する(1410)。画像は、 o 、 e 及び p の値と共にタグ付けされる(1412)。そして、タグ付けされた画像は、データ群(DT)に保存される(1414)。本方法はステップ1416に進み、環境属性群(E)の中に何らかの別の環境属性が有るか否かを判定する。もしも有れば、本方法はステップ1406に戻り、別の環境属性を取り出し、その環境属性の画像を、捕捉し、タグ付けし、データ群に保存することが、データ群の総ての環境属性が処理されるまで行われる。一方、ステップ1416において、さらなる環境属性が無かった場合、本方法はステップ1418に進み、オブジェクト群(O)の中に何らかの別のオブジェクトが有るか否かを判定する。もしも有れば、本方法はステップ1404に戻り、別のオブジェクトを取り出し、取り出したオブジェクトについて、ステップ1406ないし1416を反復するように処理が進む。無ければ本方法は終了し、オブジェクト、環境属性及び装置配置属性の総ての可能な組み合わせと共に、データ群の生成及び保存の処理を終了する。本方法は、新たな装置が追加又は検出された時はいつでも図14のステップを自動的に実行するように自動化できることを、当業者は認識するであろう。

10

【0083】

図15Aを参照するに、初期化及び分類子の生成に関する一般的な方法が示されている。本方法は、ノイズ因子に関する範囲及びサンプリング間隔(例えば、ピンぼけ、動きによるぼやけ、動的な範囲等)を決定することで開始される(1502)。本方法は、多次元空間中の各点で分類子を生成する(1504)。そして、各分類子はデータ群(DT)に適用される(1506)。最終的には、分類子のパフォーマンスはデータ群に記録される(1508)。

20

【0084】

図15Bを参照するに、分類子の生成及び初期化を行う別の方法例(1302)が示されている。図15Bの方法は分類子生成の一例を与え、ノイズは、ガウシアンぼかし、 x モーションブラー(motion blur)、及び y モーションブラーである。本方法は、ガウシアンぼかし、 x モーションブラー及び y モーションブラーについて、範囲及びサンプリング間隔を決定することから始まる(1520)。例えば、ガウシアンぼかしは、 s_i 間隔で s_0 ないし s_g に等しく、 x 方向のモーションブラーは d_{x_i} 間隔で d_{x_0} ないし d_{x_x} に等しく、及び y 方向のモーションブラーは d_{y_i} 間隔で d_{y_0} ないし d_{y_y} に等しい。次に、本方法は、ガウシアンぼかし、 x モーションブラー及び y モーションブラーを元の画像群(P)に適用し、画像群P'を得る。例えば、元の画像群P内の元のオブジェクト画像 p の各々について、本ステップは、ガウシアンぼかしを s_0 ないし s_g に間隔 s_i 各々について適用し、 x モーションブラーを d_{x_0} ないし d_{x_x} に間隔 d_{x_i} 各々について適用し、 y モーションブラーを d_{y_0} ないし d_{y_y} に間隔 d_{y_i} 各々について適用し、結果の画像を画像群P'に加える。次に、本方法はP'から分類子を生成する(1524)。例えば、本方法は、多次元空間中のある点について、 d_x 、 d_y に関する分類子Cを生成する。そして、本方法は、 d_x 、 d_y に関する分類子Cを適用し、トレーニング群DT中の画像を認識する(1526)。認識された画像数、画像の身元(識別子)及び平均的な信頼性スコア(例えば、インビジブルジャンクションに対する平均的な正対応(inlier)の数)を含む分類子パフォーマンスを本方法は記録する(1528)。そして、分類子Cは分類子結果群(CR)に加えられる。次に、本方法は、その分類子が最良のパフォーマンスをもたらすか否かを判定し、もしそうであれば、その分類子をそのように特定する[最良分類子=(d_x , d_y)](1530)。例えば、ある点に至るまでに検査された総ての分類子の中で、最大数の認識画像数及び最高の平均信頼性スコアを有する場合、その分類子はベストパフォーマンスである。本方法は、次に、何らかのさらなる分類子が有るか否かを判定する(1532)。言い換えれば、本方法は、ガウシアンぼかし、 x モーションブラー及び y モーションブラーの可能な値に関する多次元空間中に、分類子が生成されていない何らかの点があるか否かを確認する。存在する場合、本方法はステップ1522に戻り、それらの点の内の1つについて分類子を生成し、ステップ1524ないし1532を繰り返す。一方、生成されるべき如

30

40

50

何なる追加的な分類子も無かった場合、本方法は完了し、終了する。

【 0 0 8 5 】

図 1 6 を参照するに、最多数の画像を認識する分類子を発見する方法が説明される。基本的に本方法は、ほとんど「認識されていない (unrecognized)」画像を認識する識別子を発見し、それをソリューション群に加え、それらの画像を「認識された (recognized)」ものとする。N 個の分類子が決定されるまで、或いは新たに「認識された」画像数がしきい値未満になるまで、本プロセスは反復される。

【 0 0 8 6 】

必要な分類子の最小数 N を設定し、装置トレーニング画像を「未認識 (unrecognized)」として設定することで、本方法は始まる (1 6 2 0)。次に、本方法は、分類子群 C R の中から或る分類子を検索する (1 6 0 4)。そして、分類子により認識される装置トレーニング群中の画像及び認識された画像に対する平均信頼性スコアを決める (1 6 0 6)。新たな画像に対する平均信頼性スコア及び新たに認識される画像数を用いて、分類子認識スコアが決定される。その分類子認識スコアが最高であった場合、その分類子のスコアに最大スコアが設定され (1 6 0 8)、その分類子の d_x 及び d_y の値を最良の分類子が有するように設定され、その分類子で認識される画像を含むようにソリューション群が決められる。次に、本方法は、分類子群 C R の中にさらに分類子があるか否かを確認する。有れば、本方法はステップ 1 6 0 4 に戻り、ステップ 1 6 0 6 及び 1 6 0 8 を繰り返すように処理を進める。無かった場合、本方法は、その新たな「最良の」分類子が、現在の解 (ソリューション) について予想される特性を大幅に改善するか否かを確認する。これは、新たに認識された画像数がしきい値未満であるか否か、或いは認識された画像総数が装置トレーニング群に等しいか否か等を確認することでなされる。条件が真ならば本方法は終了する。そうでなければ本方法はステップ 1 6 1 4 に進み、発見された分類子数を表す変数 C F を 1 つだけ増やす。本方法は、新たな「最良の」分類子を分類子群 C S に加え、その分類子により認識された新たな画像をソリューション群に加え、最大スコアをゼロにリセットする。そして、本方法は「最良の」分類子を分類子群 C R から除去する。最終的に、本方法は、発見された分類子数が、ワーキング分類子群について望まれる分類子数 N 以上であるか否かを判定する。そうでなければ、本方法はステップ 1 6 0 4 に戻り、分類子群 C R に関するプロセスを反復し、最良の分類子とその分類子群から除かれる。望まれる分類子数 N 以上であった場合、本方法は完了し、終了する。

【 0 0 8 7 】

図 1 7 を参照しながら、仮定及び検査を行う方法を説明する。最良の分類子が判定された場合、本発明は、分類子を修正することで認識を改善しようとする。本発明は、デバイストレーニングセット内の未認識の画像を決定することで (1 7 0 2)、本方法は始まる。これは例えば、デバイストレーニングセットの中から、図 1 6 の方法で認識された画像総てを除くことで実行可能であり (未認識の画像 = $D T - G S$)、 $G S$ = 図 1 6 で発見されたソリューションセットである。本方法は、選択された未認識の或る画像を抽出する (1 7 0 4)。次に、様々な分類子に合致する認識済み画像のうち最も近いペアを見出し、その場合における距離 (相違) は、配置属性及び環境要因を比較することで決定される。一実施例の場合、このステップは $G S$ 内で画像のペアを見出すことで実行され、次のように行われる: a) 異なる識別子により i_1 及び i_2 が認識され、b) $G S$ の画像全体の中で距離 (u, i_1) は最も小さく、c) 条件 1 の下で $G S$ の画像全体の中で距離 (u, i_1) は次に小さく、及び d) 2 つの画像の距離が、オブジェクト、環境条件及び各々の配置属性を比較することで判定される。最小の場合、総てのオブジェクト及び環境条件は同じであると考えられ、距離は、配置属性に基づく L_2 ノルム ($L_2 \text{ norm}$) により決定され、各次元に関して規格化される。そして、本方法は、ペアのメンバ各々の画像間の中間 (half way) であるノイズ属性から新しい分類子を生成し、新たな分類子を格付けする (1 7 0 8)。一実施例の場合、本方法は画像 i_1 及び i_2 のペアに基づいて新たな分類子を生成する。例えば、 $\sigma = (i_1 \text{ の } \sigma + i_2 \text{ の } \sigma) / 2$ 、 $d_x = (i_1 \text{ の } d_x \text{ 及び } i_2 \text{ の } d_x) / 2$ 、及び $d_y = (i_1 \text{ の } d_y \text{ 及び } i_2 \text{ の } d_y) / 2$ と

共に本方法は始まる。次に、トレーニングセット P 内の下のオブジェクト画像 p 各々について、ガウシアンブラシグマ (Gaussian blur sigma)、モーションブラー dx 及び dy が p に適用され、その結果 p' に加えられる。そして、修正された分類子 C シグマ、dx、dy (P') が生成される。次に、修正された分類子 C シグマ、dx、dy (P') を使用して未認識の画像 (DT-GS) を認識し、本方法は、認識された画像数、画像の識別子及び平均的な確からしさ (IJ に関する平均的なインライア (inliers) の数) を記録する。新たな分類子は新たな分類子群に加えられる (1710)。次に、本方法は、未認識の画像の存否を判定する (1712)。存在していた場合、本方法はステップ 1704 に戻り、別の未認識画像を抽出し、その未認識画像についてステップ 1706、1708、1710 及び 1712 を反復する。次に、本方法は、分類子認識スコアが閾値を

10

【0088】

一実施例の場合、本方法は、(オブジェクト中の 3"、チルト角 = 5 度、回転角 = 0 度の) 画像 u と共に始まり、C (シグマ = 20、dx = 0、dy = 10) により認識される画像 i1 (オブジェクト中の 3"、チルト角 = 10 度、回転角 = 0 度) と、C (シグマ = 10、dx = 10、dy = 0) により認識される画像 i2 (オブジェクト中の 4"、チルト角 = 0 度、回転角 = 10 度) とを見出す。本方法は、新たな分類子 nc シグマ =

20

品質予測器 502

図 18 ないし 21 を参照しながら、一形態による品質予測器 502 及びその動作を詳細に説明する。品質予測器 502 は認識スコア (aka Quality Predictor) を生成し、その認識スコアは、特定の利用可能な画像 / 認識アルゴリズムにとって或る画像が良い候補であるか否かを予測するのに使用可能である。動きによる不鮮明さ、焦点の不鮮明さ、充分でない明るさ、及び情報不足等のような多数の理由に起因して、画像が認識されないかもしれない。認識スコアを計算する目的は、認識可能な画像を「低品質」と分類すること

30

【0089】

品質予測器 502 は、ある形態に関して説明され、その形態における品質予測器は、図 5 に示され説明されたようなディスパッチャ 402 の一部分である。本実施例の場合、品質予測器 502 は認識スコアを分配部 506 への入力として出力し、分配部はどの認識部 410 を使用するか (及びどの認識アルゴリズムを実行するか) を決定する。しかしながら、品質予測器 502 及び認識スコアが有効かつ有利に活用される多くのシステム構成が存在することを当業者は認めるであろう。別の形態 (第 2 形態) の場合、品質予測器 502 は捕捉装置 (例えば、移動電話装置 102、デジタルカメラ、コンピュータ 110 等) 上で動作し、捕捉装置は、捕捉した画像品質が、MMR 照合部 106 の何れかの照合ユニット 410 で認識される程度に充分か否かを判定する。捕捉された画像品質が充分であった場合、それは MMR 照合部 106 に送られ、そうでなかった場合、別の画像を捕捉するようにユーザは単に要求される。或いは、捕捉された画像及び品質予測スコアがユーザに提示され、彼 / 彼女は、その画像が MMR 照合部 106 に送られるべきか否かを決定してもよい。第 3 形態の場合、品質予測器 502 は結果合成部 610 の一部分であり、複数の認識部 410 が存在し、認識スコアは、認識結果がどのように評価されるかを定める。第 4 形態の場合、品質予測器 502 はインデックス部 414 の一部分であり、認識スコアの計算が索引付けプロセスに先行し、そのスコアを使って、入力文書ページを索引付けす

40

50

る際、どのインデクサ（索引作成プログラム）を使用する必要があるかを決定する。例えば、BWCアルゴリズムを使って索引付けされる画像としては認識スコアが低かった場合、その画像はIJアルゴリズムだけを使って索引付けされるかもしれない。さらに、索引付け及び認識の双方について同じ品質予測器が使用されてもよい。第5形態の場合、品質予測器502は移動装置102における「画像捕捉」処理の前に使用される。画像を取得する前に認識スコアが計算され、認識スコアが閾値より高かった場合にのみ装置は画像を捕捉する。品質予測器502は、カメラのチップ内に組み込むことが可能であり、カメラのハードウェア又はソフトウェアを制御するのに使用可能である。例えば、カメラの絞り（開口）、露出時間、フラッシュ、マクロモード、安定化（手ぶれ防止機能）等が、認識部410の条件及び捕捉画像に基づいて調整されてもよい。例えば、BWCは、ぼけたテキスト画像を認識することができ、カメラ付き電話機を振動させることで、ぼかした画像の取得を行うことができる。

【0090】

図18に示されるように、一形態の品質予測器502は、認識アルゴリズムパラメータ1802、ベクトル計算器1804、スコア生成器1806及びスコアモジュール1808を有する。品質予測器502は、信号線532に結合され、問い合わせ画像、コンテキスト及びメタデータ、デバイスパラメータを受信する。問い合わせ画像は、ビデオフレーム、1つのフレーム、又は複数の画像フレームでもよい。コンテキスト及びメタデータは、時間、日付、場所、環境状態等を含む。デバイスパラメータは、ブランド、型、マクロブロックオン/オフ、ジャイロ又は加速度計の測定値、絞り、時間、露出、フラッシュ等の情報を含む。さらに、品質予測器502は、認識アルゴリズムパラメータ1802の或るパラメータを使用する。認識アルゴリズムパラメータ1802のこれらのパラメータは、取得部406又は画像登録部408から品質予測器502に提供されてもよい。ベクトル計算器1804は、画像から品質特徴ベクトルを計算し、コンテンツや歪を計測し、例えば、ぼけ、認識可能な特徴の存在、認識可能な特徴の量、明度等を測る。ベクトル計算器1804は、1つないしn個のいくつかの品質特徴ベクトルを計算してもよい。ベクトル計算器1804の一例は、図19を参照しながら以下で説明される。場合によっては、ベクトル計算器1804は、使用される認識アルゴリズムの情報を必要とし、ベクトル計算器1804は、信号線1820を介して認識アルゴリズムパラメータ1802に結合される。例えば、インビジブルジャンクションアルゴリズムが使用される場合、ベクトル計算器1804は、認識の尺度として、どの程度多くの接合点が画像中に存在するかを算出する。これら算出された特徴の全部又は一部は、信号線1824を介してスコア生成部1806に入力される。スコア生成部1806は信号線1822にも結合され、認識アルゴリズムパラメータ1802に関する認識パラメータを受信する。スコア生成部1806の出力は、スコアモジュール1808に与えられる。スコアモジュール1808は、スコア生成部1806から与えられたスコアを使用して認識スコアを生成し、それらのスコアにウェイトを適用する。一形態では、その結果は一つの認識スコアになる。別の形態の場合、結果は、最高から最低に至る格付けされた複数の認識スコアになる。

【0091】

図19を参照するに、本発明の一形態によるベクトル計算器が示されている。ベクトル計算器1804は、ぼけ属性モジュール1902、コンテンツ量属性モジュール1904、明度属性モジュール1906、ブリードスルー（bleed through）属性モジュール1908、符号化アーチファクト属性モジュール1910、視点歪属性モジュール1912、カメラノイズ属性モジュール1914、テキスト品質属性モジュール1916、オブジェクト検出属性モジュール1918、認識アルゴリズム属性モジュール1920及びロバストネス属性モジュール1922を有する。ベクトル計算器1804は他の様々な属性モジュールを含んでもよく、図19の例は単なる一例に過ぎないことを当業者は理解するであろう。他の形態は、図19に示されているものよりも多数の又は少数のモジュールを含んでもよい。ベクトル計算器1804は、問い合わせ画像、コンテキスト及びメタデータ、デバイスパラメータ、並びに認識アルゴリズムのパラメータを受信する入力

10

20

30

40

50

を有する。一形態では、これらの入力は、ベクトル計算器 1804 のモジュール 1902、1904、1906、1908、1910、1912、1914、1916、1918、1920 及び 1922 に結合される。

【0092】

ぼけ属性モジュール 1902 は、捕捉された画像中の不鮮明さの量に基づいて、品質属性を決定するソフトウェア及びルーチンである。一実施例では、以下の方法の 1 つ以上によりモーションブラーが計算される。

【0093】

グラディエント属性 (Gradient Based Feature)。入力画像の勾配 (gradient) が、(例えば、 $[-1 \ 0 \ 1]$ カーネルと共に) 縦及び横の双方向で算出される。勾配値の大きさを使用して、 n ビンヒストグラム (n -bin histogram) を 2 つ形成し (例えば、 $n = 3$)、1 つは縦方向、1 つは横方向であり、 $2n$ 次元の特徴ベクトルを生成する。入力画像の勾配は、 $2D$ カーネルと共に計算され、 n 個の勾配方向についてヒストグラムが計算される (例えば、 $n = 8$ 方向)。画像全体についてヒストグラムが計算されてもよいし、或いは重複する又は重複しない領域に画像が分割され、各領域について n ビンヒストグラムが計算されてもよい。例えば、 $K \times L$ 個の重複しない四角形領域に画像が分割された場合、結果の特徴ベクトルは $K \times L \times n$ のサイズになる。

【0094】

エッジ検出。エッジ検出については、例えば、以下の非特許文献 1 に記載されている：“A Computational Approach To Edge Detection,” Canny, J., IEEE Trans. Pattern Analysis and Machine Intelligence, 8:679-714, 1986。不鮮明属性を計算すると同様に、エッジ方向及びエッジ数のヒストグラムが使用可能である。

【0095】

動き推定属性 (Motion Estimation Based Features)。入力がビデオフレームの場合、2 つのフレーム間の動きは、例えば “A Novel Four-Step Search Algorithm For Fast Block Motion Estimation” として当該技術分野で知られている動き推定アルゴリズムの 1 つにより推定可能であり、水平方向の変位 (dx)、垂直方向の変位 (dy) 及び時間的なずれ (dt) は、特徴ベクトル $[dx, dy, dt]$ を生成する。

【0096】

周波数領域分析による属性 (Frequency Domains Analysis Based Features)。画像の焦点ぼけは、空間的には不変のローパス空間フィルタとして取り扱うことができ、このフィルタのカットオフ周波数は、ぼけの大きさに反比例する。画像が鮮明でコントラストの高い輪郭 (エッジ) を含んでいた場合、その画像はかなり高い周波数成分を含んでいる。ぼけ属性モジュール 1902 は、画像に関する $2D$ 離散フーリエ変換 (DFT) を行う。ほとんど総ての自然な画像の場合、周波数が増えるにつれてスペクトルの大きさは急速に減少する。カットオフを算出するため、各周波数における (総ての方向における) 平均的なスペクトルの大きさが算出される。すなわち、DC 項 (0 周波数) を中心とし、同心状の長円に沿って平均値が算出される。この平均的な半径の大きさのグラフが得られた後で、カットオフを抽出したい。具体的には、これは、スペクトルがノイズフロアの平坦なレベルに落ち込む最初のポイントである。このカットオフを選出する方法の 1 つは、(スパイクを除去するために) グラフをローパスフィルタに通し、グラフを一定の高さに尺度変換し、所定のグラフ値未満の第 1 ポイント (グラフが (一定の) 非常に平坦な傾斜に達するポイント) を探すことである。別の方法は、一定の距離を超えて最も大きな曲率 (最も大きな角度変化) を伴うグラフ上のポイントを探すことである。これらのカットオフが判定されると、グラフは実際のスケールに変換される。これは、次のようにすることで最も簡易に実行され、テキストブロックの校正ショット (calibration shot) を使用し、カットオフを算出し、校正されたカットオフと新たにサンプリングされたカットオフ各々との比率を関数が返すようにすることである。不鮮明さの特徴と同様に、カットオフ周波数 f 及び何らかの他の属性を使用し、固有ベクトル E_i やスペクトル拡散率 S のような周波数領域分析から演算を行うことができる。

10

20

30

40

50

【 0 0 9 7 】

コンテンツ量属性モジュール 1 9 0 4 は、画像中のコンテンツの量に基づいて品質属性を決定するソフトウェア及びルーチンである。一実施例では、次のような方法の何れかによりコンテンツ特徴量が算出される。

【 0 0 9 8 】

接続コンポーネント属性 (Connected component features)。“Survey Over Image Thresholding Techniques And Quantitative Performance Evaluation”に記載されているような技術を使用して画像は二値化される。“A Simple And Efficient Connected Components Labeling Algorithm”に記載されているような技術を使用して、接続されたコンポーネントが算出される。そして、サイズ、幅、高さ、画素数、アスペクト比、拡散率、固有ベクトル、ホール数 (number of holes)、コンパクト性等のような接続コンポーネントに関し、ヒストグラム ($H = [h_1 \dots h_n]$)、平均 (A) 及び分散 () のような多数の特徴を、コンテンツ量属性モジュール 1 9 0 4 は接続コンポーネントから算出する。

10

【 0 0 9 9 】

エッジ属性。“Optimal Corner Detector”を使用するようなコーナー、画像における算出されたエッジ及びそれらの検出を使用して、コーナー及びエッジのヒストグラムを形成する。

【 0 1 0 0 】

エントロピー。画像中のグレーレベル / カラー画素 (E_g / E_c)、二値化された画素 (E_b)、接続されたコンポーネント (E_{cc}) のエントロピーは、どの程度多くの情報が画像の中にあるかを指す指標になる。画像は、重複する又は重複しないセグメント (部分) に分割可能であり、これらのセグメント各々についてのエントロピーが算出され、エントロピーは、エントロピー特徴ベクトルを計算するのに使用される。エントロピーを計算する前に、カメラのノイズを除去するために、ローパスフィルタリング又はメジアンフィルタリングのステップが使用されてもよい。

20

【 0 1 0 1 】

周波数領域属性。周波数属性分析、例えば 2 D 高速フーリエ変換 (FFT) 又は離散コサイン変換 (DCT) のような変換関数を、画像に適用することができる。周波数領域表現が得られると、それはヒストグラムと共に表現可能である。

30

【 0 1 0 2 】

圧縮画像サイズ。JPEG のような同じ圧縮アルゴリズムを使用し、同じ量子化因子のような同じパラメータと共に、2 つの画像が圧縮されたならば、大抵の場合、高い周波数成分を伴う画像は、大きなファイルサイズの圧縮画像になる。ファイルサイズ S が過剰に大きかった場合、それは、画像の中にスペckルノイズ (speckle noise) が有ることの指標になり、そのノイズは画像の認識を困難にする。ファイルサイズが過剰に小さかった場合、それは、画像の中で重要な情報が不足していることの指標になる。

【 0 1 0 3 】

明度属性モジュール 1 9 0 6 は、明度属性に基づいて品質属性を決定するソフトウェア及びルーチンである。一実施例の場合、明度属性は以下の方法の何れかにより算出される：

40

ヒストグラム属性。グレースケールのヒストグラムやカラーヒストグラムのパターンは、画像がどの程度明るいかを指す良い指標になる。例えば、画像がほとんど暗い画素を有しかつ 5 0 より大きな値を有する輝かしい画素を 1 つも含まない場合 (画素値は 0 ないし 2 5 5 の値をとるものとしている)、その画像は、低い認識スコアを示すであろう。ヒストグラム又はヒストグラムから抽出された特徴 (例えば、ピーク数) に基づいて特徴ベクトルを形成することに加えて、フィッシャー判別分析 (Fisher discriminant analysis) が、画素の明度及び飽和値双方について実行されてもよい。明度サンプルに関する判別分析は、平均値 m_{l_1} 及び m_{l_2} の 2 クラスと、クラス間分散 (interclass variance) v_{l_1} 及び v_{l_2} をもたらす。分離度合い $m_{l_1} - m_{l_2}$ 、クラス間分散の合計

50

$v l_1 + v l_2$ 、それらの比率 $m l_1 - m l_2 / v l_1 + v l_2$ は、特徴ベクトルを算出するのに使用可能である。

【0104】

コントラスト。画像のコントラストは、多数の効率的な方法により算出可能であり、例えば、 $\text{コントラスト} = (\text{明るい画素上位 } X \% \text{ の平均値}) / (\text{暗い画素上位 } Y \% \text{ の平均値})$ である。

【0105】

シャドー（影）。一形態では、“Image Difference Threshold Strategies And Shadow Detection” に記載されているようにして、画像からシャドーマスクを算出することができ、明度属性モジュール1906はシャドーヒストグラムから抽出された特徴を使用する。他の形態では、明度属性モジュール1906は、重複する又は重複しないセグメントに画像を分割し、それらのセグメントの領域のヒストグラムを算出し、ヒストグラム平坦化 (histogram equalization) を実行し、ヒストグラムが類似しているか否かを検査する。類似スコアは1つの属性として使用可能である。

10

【0106】

ブリードスルー属性モジュール1908は、ブリードスルー属性に基づいて品質属性を決めるソフトウェア及びルーチンである。捕捉された画像が紙書類の画像であり、その画像が、視界の中になくページからの何らかの内容（すなわち、そのページの裏側に印刷されている内容）を含んでいた場合、認識の実効性（パフォーマンス）は劣化する。他のページからの内容は、本願では「ブリードスルー (bleed through) 属性」と言及する。ブリードスルー属性は、次のような非特許文献2に記載されている技術を利用して検出できる。“Restoring Ink Bleed-Through Degraded Document Images Using a Recursive Unsupervised Classification Technique,” P.Rosen, et al., Proceedings of the 6th British Machine Vision Conference, 1995.

20

【0107】

符号化アーチファクト属性モジュール1910は、符号化アーチファクトに基づいて品質属性を決めるソフトウェア及びルーチンである。符号化及び圧縮のアーチファクトは、画像の認識を劣化させる。“Detection Of Blocking Artifacts Of Compressed Still Images,” Triantafyllidis et al. に記載されているような方法は、画像中に存在する符号化アーチファクトの量及び確率に基づいて、特徴ベクトルを算出するのに使用可能である。

30

【0108】

視点歪属性モジュール1912は、視点歪 (perspective distortion) に基づいて品質属性を決めるソフトウェア及びルーチンである。重大な視点歪は、画像の認識に影響を及ぼす。視点歪を検出及び修正する多くの方法が存在する。修正は、認識法の一部であってもよいが、そのような歪の検出は、画像が認識できるか否かを確認するのに有用である。視点歪は、画像中の線及び縁（輪郭）の相対的な位置関係を用いる方法や、次の非特許文献3に記載されている方法等を用いることで算出可能である。“Perspective Correction Methods For Camera Based Document Analysis,” by Jagannathan et al., Proc. First Int. Workshop on Camera-based Document Analysis and Recognition, 2005。視点歪のヒストグラム量を表す属性は、認識スコアを算出するのに使用可能である。

40

【0109】

カメラノイズ属性モジュール1914は、カメラノイズに基づいて品質属性を決めるソフトウェア及びルーチンである。カメラの電荷結合素子 (CCD) のサイズが小さすぎた場合、或いは環境の中に十分な光がなかった場合、捕捉された画像又はビデオフレームは、スペckル雑音のようなノイズを含むおそれがある。そのようなノイズを測定するいくつかの方法がある。カメラノイズ属性モジュール1914で使用される方法の1つは、メジアンフィルタを使ってスペckル雑音の低減を図り、縮小画像ノイズ及び元の画像間の差分を求め、その差分画像に基づいて何らかの特徴を算出することである。例えば、閾値

50

を上回る画素の割合が算出される。

【0110】

テキスト品質属性モジュール1916は、画像中のテキストに基づいて、品質属性を決めるソフトウェア及びルーチンである。テキストは、テキストボックス属性、光文字認識(OCR)属性、又はテキストライン属性として測定可能である。

【0111】

テキストボックス属性。テキストボックスは、例えば、次の非特許文献4に記載されている方法を使用して画像の中から検出可能である。“Text Detection from Natural Scene Images: Towards A System For Visually Impaired Persons,” N.Ezaki, et al., Proc. of 17th Int. Conf. on Pattern Recognition(ICPR 2004), IEEE Computer Society, 2004, pp.683-686, vol.II, 23-26 August, Cambridge, UK. テキスト品質属性モジュール1916は、テキストボックス高さのヒストグラム、テキストボックス幅/高さの比率のヒストグラム(nubs)、及びテキストボックス密度のヒストグラム等のような様々な特徴を、テキストボックス検出により算出する。認識に好ましい(例えば、適切な高さ)と考えられるテキストボックスのヒストグラムが算出される。例えば、テキストボックスは、画像高さの1/10未満であって、画像高さの1/50より大きい。

10

【0112】

光文字認識(OCR)属性。画像は、OCR(光学文字認識)により処理可能であり、OCRの信頼性スコアを使用して特徴ベクトルを算出することができる。

20

【0113】

テキストライン属性。一形態では、テキスト品質属性モジュール1916は、行間ヒストグラムを計算することで、テキストライン属性を算出する。これらの属性は、テキストが不鮮明であって、勾配法(gradient based method)が十分な情報をもたらさない場合に特に有用である。(1画素の幅で)縦方向の強度値の第1グラフが算出され、各グラフの局所的な最小値が見出される。次に、局所的な最小値同士の間差分が算出される。これらの差分は、概して、行間の距離に対応する。そして、これらの距離に関するヒストグラムが算出される。非文書画像及びほとんど文字を含んでいない文書画像のヒストグラムは容易に区別可能である。なぜなら、それらは多くのピークを有し、ヒストグラムは非常に一様に分布しているからである。それらは低いピーク値も有する。ピーク数N、ピーク高さ P_i のようなヒストグラムから抽出された属性は、テキストライン属性ベクトルを算出するのに使用される。

30

【0114】

オブジェクト検出属性モジュール1918は、画像オブジェクトに基づいて品質属性を決めるソフトウェア及びルーチンである。線、表、箇条書き項目、リスト、顔及び建物等のような何らかの対象や形状が存在することは、画像認識アルゴリズムに依存する認識性に影響する。例えば、画像の大部分が表を含んでいた場合、IJアルゴリズムは良好には機能しないかもしれない。画像が多くの線を含んでいた場合、BWCアルゴリズムは良好には機能しないかもしれない。一実施例では、オブジェクト検出属性モジュール1918は、画像中のオブジェクトを検出する際、以下の何れかの非特許文献5ないし7の技術を使用してよい: “Use Of The Hough Transformation To Detect Lines And Curves In Pictures,” R.Duda and P.Hart, Communications of the ACM, Vol.15, pp11-15, 1972; “A Survey Of Table Recognition,” by R Zanibbi, D Blostein, JR Cordy, International Journal on Document Analysis and Recognition, 2004;又は“Face Detection: A Survey,” by E Hjelm, BK Low, Computer Vision and Image Understanding, 2001(これらは、画像中のオブジェクト認識法として使用可能である。)。或るオブジェクトを含む画像の確率 $[P(O_1), \dots, P(O_i)]$ は、オブジェクト検出属性モジュール1918により算出される特徴ベクトルである。

40

【0115】

50

認識アルゴリズム属性モジュール1920は、認識アルゴリズムにおける相違に基づいて品質属性を決めるソフトウェア及びルーチンである。(画像認識アルゴリズムの一部である)特徴抽出プロセスの全部又は一部は、画像の認識性を測るために画像について実行される。例えば、認識アルゴリズムにより抽出する固有の特徴が多くなかった場合、その画像は低い認識スコアに割り当てられる。例えば、特徴抽出プロセスは、(IJ、BWC又はPCのような)認識アルゴリズム各々について異なる可能性があり、或いはいくつかの一般的なコンポーネント(例えば、BWC及びSyntextコーディングにおける境界枠抽出コンポーネント)を共有するかもしれない。例えばIJの場合、IJスケルトンの特性から次のような特徴が抽出される:接合点数、領域(ホール(holes))の数、領域サイズのヒストグラム、及び画像内での領域の分布等である。さらに、各接合点における勾配方向の特徴ベクトルが算出された後、これらの特徴ベクトルの相違は、コンテンツの認識性の指標として使用可能である。SIFTの場合、SIFT属性自体だけでなくSIFT属性の検出位置から(例えば、コーナー検出器により)抽出された特徴が、認識性の指標として使用可能である。BWC及びSyntextの場合、(「テキスト品質属性」の説明部分で言及された)境界枠属性に加えて、完全な特徴ベクトルが算出されてもよい。固有認識特徴ベクトル(recognition-specific feature vector)の数及びその空間分布のような特徴は、認識性の指標として使用可能である。また、各認識アルゴリズムにおける特徴ベクトル各々の一意性(uniqueness)が算出され、指標として使用されてもよい。例えば、BWCの場合、[3 22 22]のような特徴ベクトルは非常に一般的であり、[10 67 897]のような特徴ベクトルは非常に特殊である。特徴ベクトルの一意性を測るため、ターム頻度-インバース文書頻度(TF-IDF:term frequency-inverse document frequency)法が適用されてもよい。認識アルゴリズム属性モジュール1920は、これらの何れか又は組み合わせを使用して、特徴ベクトルを生成してもよい。

【0116】

ロバストネス属性モジュール1922は、ロバストネス(robustness)に基づいて品質属性を決めるソフトウェア及びルーチンである。画像から抽出された属性が、配置、環境及び装置変更等によるカメラ歪の影響を受けにくく安定的であった場合、画像から抽出された属性は堅牢である(ロバスト的である)。入力画像はさらに処理され、様々なカメラ歪に対して抽出画像が安定的でロバスト的であるかどうかを判定する。例えば、IJの場合、画像は、ぼけていたり、傾いていたり、歪んでいたりすることがあり(あるいは、何らかの他のノイズシミュレーションが考慮されてもよい)、スケルトン(骨組み)、場所属性及び特徴ベクトルが再び算出される。入力画像の場所属性及び歪んだ画像の場所属性間の距離が、ハウスドルフ距離法(Hausdorff distance)のような技法により測定され、特徴ベクトル間の距離が、ユークリッド距離法のような技法により測定される。これらの距離の結合(例えば、重み付け加算値又は乗算値)が小さかった場合、画像は様々なカメラの影響に対してロバスト的であり、それ故に認識スコアに対して正(positive)の寄与をもたらす。距離の指標及び結合方法は、認識アルゴリズムに依存する。

【0117】

図20を参照するに、ロバストネス特性を出力する方法のプロセス図が示されており、本方法は例えばロバストネス属性モジュール1922により実行される。本方法は入力画像を受信することから始まる(2002)。入力画像は、認識属性を抽出するように処理され(2008)、様々なレベルの歪を適用するように処理される(2004、2006)。認識属性を抽出する際(2008)、画像登録部408や認識ユニット406から抽出されるような一群の認識パラメータ2020が使用される。認識パラメータ2020は、他の属性の抽出(2010、2012)、距離の算出(2014、2016)及びロバストネスベクトルの算出(2018)にも使用される。抽出された画像属性(2008)は、場所属性間の距離計算値をモデル化するように与えられる。同様に、受信した入力画像2002に第1歪が適用され(2004)、歪んだ画像から認識属性が抽出され(2010)、その認識属性は、場所属性間の距離を算出するのに使用される(2014)。画像を受信し(2002)、歪を適用し(2006)、属性を抽出して属性を提供するこれ

10

20

30

40

50

らのプロセスは、好ましくは、多数の様々なタイプの歪について実行される。抽出された総ての属性は、場所属性間の距離を算出し、特徴ベクトルを算出するように与えられる(2014)。次に本方法は特徴ベクトル間の距離を算出する(2016)。そして、特徴ベクトル間の距離に基づいて特徴ロバストネスベクトルが算出される(2018)。

【0118】

図21は、認識スコアを生成する本発明による方法のプロセス図を示す。上述したように、本プロセスは画像又はビデオフレームを捕捉することから始まる(2102)。捕捉された画像は、画像バッファに一時的に格納されてもよいし(2104)、或いは品質ベクトルを計算するために直接的且つ速やかに与えられてもよい(2106)。いくつかの画像/ビデオフレームが、品質特徴ベクトル計算で使用するためにバッファリングされてもよい。ベクトル計算器1804に関して上述したように、画像の認識性を計算するのに使用可能な多種多様な属性が存在する。これらの属性又は指標の何らかの組み合わせが、特定の認識アルゴリズム群を使用して画像の認識性を予測するのに使用可能である(例えば、認識スコアを生成するのに使用可能である)。ベクトル計算器1804の各モジュールは、それらの指標を表す特徴ベクトルを生成する。次に、本方法は、ベクトル計算器1804から受信した特徴ベクトルを規格化(正規化)する(2108)。規格化されたベクトルは、認識スコアを分類及び計算するモジュールに与えられる。ベクトルは、分類及びスコアリングモジュールにより処理される(2110)。一形態では、分類及びスコアリングモジュール2110は、正及び負の画像サンプルと共に予め訓練されており(すなわち、正負の画像サンプルに合わせて調整されており)、分類パラメータを調整するために、認識部410からのフィードバックループ2116を含んでいる。一形態の場合、分類及びスコアリングモジュール2110は、ニューラルネットワークのようなアルゴリズム、ナイーブベースの分類器(naive based classifiers)、ベイジアンベースの分類器(Bayesian based classifiers)、サポートベクトルマシン(SVM: support vector machine)ベースの分類器等に基づいている。これらの分類器は、スコアと、スコアの確からしさ又は誤差のマージンとを出力する。一形態では、スコアが出力され、入力画像の認識を行う際(2114)、どの認識部410が選択されるべきかを決定するために、スコアは、ディスプレイ部402の分配部506のような他のコンポーネントにより使用される。第2形態では、分類及びスコアリングモジュール2110から出力されるスコアと閾値が比較され(2118)、スコアが閾値より上である又は下であるということがブール値に変換され、そのブール値も、どの認識部410が処理に使用されるべきかを決定するために、ディスプレイ部402の分配部506のような他のコンポーネントにより使用可能である。

【0119】

別のアプリケーションの場合、例えば、品質予測器502が移動装置102のクライアントの一部として動作する場合、スコア又は閾値によりブール値に変換されたスコアは、さらなる処理に備えて捕捉画像を移動装置102から送信できる程度に十分な品質を、受信画像が有するか否かを判定するのに使用可能である。一形態では、認識スコアが閾値より低かった場合、ある確認要求と共に認識スコアがユーザに表示され(2112)、その確認要求は、表示された認識スコアを有する捕捉画像について抽出が実行されることの確認である。別の形態では、認識スコアを示すメッセージがユーザに表示され、そのメッセージは、画像を認識できるレベルよりもスコアが低いのでその画像はMMRゲートウェイ104に送信されないことを示す。さらに、表示されるメッセージは、ユーザが別の画像を取得することを要求してもよい。

【0120】

さらに別の例の場合、アルゴリズム各々について別々に分類(及び分類パラメータの調整)が行われる。品質予測器502がディスプレイ部402又は索引付けの前に使用され、最高の認識スコアをもたらす認識アルゴリズムが、ディスプレイ部402において又は索引付けアルゴリズムとして最初に選択される場合、次に高いものが選択され、以下同様に続く。品質予測器502が、検索の実行前に使用される場合、各アルゴリズム(R_n)

10

20

30

40

50

についての認識スコアは、(ウェイト(w_n)と共に)重み付け加算されてもよい:スコア = w₁R₁ + . . . + w_NR_N。ウェイトは、検索時間及び検索精度に基づいて計算されてもよい。ウェイトは、検索結果を使用しながら途中で計算し直してもよい。

結果合成部 610

図 22 ないし 24 を参照しながら、異なる画像検索アルゴリズムからの画像テキスト検索結果を合成するいくつかの方法を説明する。図 22 は、予測器又は認識スコアを用いて認識結果を合成する第 1 形態の方法を示す。図 23 は、予測器を用いて認識結果を合成する第 2 形態の方法を示す。図 24 は、複数の認識結果(予測なし)を合成する本発明の一形態による方法を示す。

10

【0121】

図 22 を参照するに、品質予測器を用いて認識結果を合成する第 1 形態の方法が示されている。より具体的には、図 22 は、結果合成部 610 及び認識部 410 により実行されるプロセスを示す。破線より上側のプロセスは認識部 410 により実行され、破線より下側のプロセスは結果合成部 610 により実行される。所与の入力画像又は画像(I)及びいくつかの画像認識アルゴリズム A₁ . . . A_N の下で、各認識アルゴリズムの出力は、認識結果のリスト = { R_{1, n}, . . . , R_{k, n} } (例えば、上位 K 個の合致する文書、ページ及びページ中の場所)と、それらの結果各々に関する信頼度 = { C_{1, n}, . . . , C_{k, n} } であり、したがって、

【数 1】

20

$$[\vec{C}_n, \vec{R}_n] = A_n(I)$$

である。さらに、予測アルゴリズム P₁ . . . P_N (これらは、所与のアルゴリズム A_n の下で認識可能性を計算する)は、入力画像又は画像(I)及び認識アルゴリズムに基づいて認識スコアを計算する: S_n = P_n(I, A_n)。第 1 形態の方法は、以下の方法を使用して、結果 R (合致する文書に関する情報を含み、例えば、文書 id、ページ、ページ中の場所の情報を含む)又はオーダ結果リスト(ordered result list)を決定する。

30

【0122】

本プロセスは画像を検索することから始まる(2202)。検索された画像は、複数のインデックス 1 ないし n と共に、認識/検索を実行するのに使用される(2204、2206)。一形態では、各々の認識アルゴリズム A_n は、入力画像について結果及び信頼度ベクトルを得るように動作する:

【数 2】

$$[\vec{C}_n, \vec{R}_n] = A_n(I)$$

次に、結果合成部 610 は、複数のインデックス 1 ないし n を用いて認識スコアを計算する(2208、2210)。予測アルゴリズムは、入力画像各々及びアルゴリズム各々について認識スコアを求めるよう動作する: S_n = P_n(I, A_n)。これら 2 つの信頼度スコアは、新たな信頼度スコア

40

【数 3】

$$\vec{C}'_n$$

を計算するために使用される(2212、2214)。例えば、新たな信頼度スコア

50

【数 4】

$$\vec{C}'_n$$

は、

【数 5】

$$\vec{C}_n \text{ and } S_n$$

10

に基づいて計算され、すなわち、各々のアルゴリズム A_n について、

【数 6】

$$\vec{C}'_n = f(\vec{C}_n, S_n)$$

である。そして、新たな信頼度スコアに基づく最良の結果が選択され (2 2 1 6)、出力される。一形態では、これは、各自の関連する信頼度

【数 7】

20

$$\vec{C}'_1, \dots, \vec{C}'_N$$

に基づいて、

【数 8】

$$\vec{R}_1, \dots, \vec{R}_N$$

を併合して並べ替え、

30

【数 9】

$$\vec{R}_S$$

を得ることで実行される。そして、本方法は、

【数 10】

$$\vec{R}_S$$

40

のうち上位 L 個の結果を出力する。 L は 1 より大きい任意の整数とすることができる。

【0 1 2 3】

図 2 3 を参照するに、品質予測器を用いて認識結果を合成する第 2 形態の方法が示されている。第 1 形態とは異なり、図 2 3 の様々な方法ステップは、MMR 照合部 1 0 6 のコンポーネントにより実行される。本方法は画像を受信することから始まる (2 3 0 2)。次に、これらの画像は、複数のインデックス 1 ないし n の認識スコアを計算するのに使用される (2 3 0 4、2 3 0 6)。言い換えれば、予測アルゴリズムは、入力画像各々及びアルゴリズム各々について認識スコアを求めるように動作する： $S_n = P_n(I, A_n)$ 。次に、本方法は、関連する認識スコアが所定の閾値より大きいか否かを判定する (2 3 0 8、2 3 1 0)。大きくなかった場合、受信画像は破棄され、後続の計算は行われぬ

50

。一方、認識スコアが所定の閾値より大きかった場合、本方法は、インデックス 1 ないし n に関する認識 / 検索を実行するステップに続く (2 3 1 2、2 3 1 4)。例えば、入力画像について結果及び信頼度ベクトルを得るように、各認識アルゴリズム A_n が動作する :

【数 1 1】

$$[\vec{C}_n, \vec{R}_n] = A_n(I),$$

そして、インデックス 1 ないし n 各々について、新たな信頼度スコアが生成される (2 3 1 6、2 3 1 8)。新たな信頼度スコア

【数 1 2】

$$\vec{C}'_n$$

は、

【数 1 3】

$$\vec{C}_n \text{ and } S_n$$

に基づいて計算される :

【数 1 4】

$$\vec{C}'_n = f(\vec{C}_n, S_n).$$

そして、最良の結果が選択され (2 3 1 0)、出力される。一形態では、

【数 1 5】

$$\vec{R}_1, \dots, \vec{R}_n$$

は、各自対応する信頼度

【数 1 6】

$$\vec{C}'_1, \dots, \vec{C}'_n,$$

に基づいて併合及び並べ替えられ、

【数 1 7】

$$\vec{R}_s$$

をもたらし、その結果の内上位 L 個が出力される。

【0 1 2 4】

代替例では、図 2 2 及び 2 3 に関して説明したプロセスが以下のように若干修正される。認識結果、信頼度スコア、修正された又は新たな信頼度スコアが、同様な方法で生成さ

10

20

30

40

50

れる。特に、入力画像及び各アルゴリズムについて、認識スコアを取得するように予測アルゴリズムが動作する： $S_n = P_n(I, A_n)$ 。対応する予測スコア S_n が閾値より大きかった場合にのみ、認識アルゴリズム A_n は、入力画像について結果及び信頼度ベクトルを求めるように動作する：

【数 1 8】

$$[\vec{C}_n, \vec{R}_n] = A_n(I)$$

(信頼度スコアに基づいて、

【数 1 9】

$$\vec{R}_n$$

が格納されており、従って $R_{1, n}$ は照合結果について最良の推定であることを仮定している。いくつかのアルゴリズムは、それらの内的な閾値に起因して、如何なる結果も出力しないかもしれないことに留意を要する。)。上記の方法のように、上位の認識結果を単に併合及び並べ替えるのではなく、代替例は、各アルゴリズム($R_{1, 1}, \dots, R_{1, N}$)からの上位の結果を使用して、最良の結果を得るように投票する。投票は、文書id、(docid+pageid)又は(docid+pageid+patch location)に基づいて勘定することができる。パッチロケーション(patch location)(x, y)は、厳密に同じである必要はなく、各文書パッチロケーションに対する投票を与えるように、近接性指標(proximity measure)が使用されてもよい。そして、当該方法は最も票を多く獲得した結果を出力する。

【0 1 2 5】

さらに、結果同士の間に関連性(tie)があった場合、以下の何れかの方法を使って結果を出力することができる：a) 関連する結果のグループの信頼度 $C_{1, n}$ を合計し、最高の信頼度合計に対応する結果を出力する；b) 修正された信頼度スコア $C'_{1, n} = f(C_{1, n}, S_n)$ を計算し、関連する結果のグループについて修正後の信頼度 $C'_{1, n}$ の合計を求め、最高の信頼度合計に対応する結果を出力する；c) 関連する結果のグループについて予測スコア S_n を合計し、最高の予測スコア合計に対応する結果を出力する；d) 最高の信頼度スコア $C_{1, n}$ 又は最高の修正信頼度スコア $C'_{1, n}$ をもたらしたアルゴリズムの結果 $R_{1, n}$ を出力する；e) 最高の予測スコア S_n をもたらしたアルゴリズムの結果 $R_{1, n}$ を出力する；又はf) 出力するアルゴリズムの順序を(事前に測定した精度に基づいて)予め決定しておく。例えば、最初がI Jであり、次がスネークコーディング(snake coding)であり、等々。

【0 1 2 6】

図 2 4 は、本発明による複数の認識結果(予測無し)を合成する方法を示す。プロセスは画像を検索することから始まる(2 4 0 2)。検索された画像は、複数のインデックス1ないしnと共に認識/検索を実行するのに使用される(2 4 0 6、2 4 0 8)。一形態では、入力画像について結果(もしあれば)及び信頼度ベクトルを求めるように、各認識アルゴリズムが動作する：

【数 2 0】

$$[\vec{C}_n, \vec{R}_n] = A_n(I)$$

(信頼度スコアに基づいて、

10

20

30

40

【数 2 1】

$$\vec{R}_n$$

が格納され、したがって、 $R_{1, n}$ は照合する結果について最良の推定になることを仮定している。いくつかのアルゴリズムは、それらの内的な閾値に起因して、如何なる結果も出力しないかもしれないことに留意を要する。次に、入力画像及びアルゴリズム各々について尤度スコアを求めるように、予測アルゴリズムが動作する(2404、2408)： $S_n = P_n(I, A_n)$ 。尤度スコアは受信されてもよいし、日付、時間、場所、個人プロフィール、検索履歴及び他のコンテキスト・メタデータを使用して、尤度スコア O_n を生成されてもよい。信頼度スコア及び尤度スコアを用いて、修正された信頼度スコアを計算する(2412、2414)。例えば、各アルゴリズム A_n について、修正された信頼度スコア

10

【数 2 2】

$$\vec{C}_n'$$

は、

【数 2 3】

$$\vec{C}_n' = f(O_n, \vec{C}_n)$$

20

に基づいて算出される。そして、新たな信頼度スコアに基づく最良の結果が選択され(2216)、出力される。一形態では、これは、各自対応する信頼度

【数 2 4】

$$\vec{C}_1', \dots, \vec{C}_n'$$

30

に基づいて

【数 2 5】

$$\vec{R}_1, \dots, \vec{R}_n$$

を併合及び並べ替え、

【数 2 6】

$$\vec{R}_s$$

40

の結果を得ることで実行される。そして、本方法は上位 L 個を出力する。 L は 1 より大きい任意の整数である。第 2 形態の場合、最良の結果は、各アルゴリズムから投票によって決定される。最良の結果を投票するため、各アルゴリズムから上位 L_n 個の結果 ($\{R_{1, 1}, \dots, R_{L_1, 1}\}, \dots, \{R_{1, N}, \dots, R_{L_N, N}\}$) が使用される。 L_n はアルゴリズムに依存する。例えば、 IJ は 1 つの結果しか出力しないかもしれないし(すなわち、 $L_n = 1$)、 BMC は上位 10 個の結果を出力するかもしれない($L_n = 10$)。総てのアルゴリズムについて L_n が 1 に設定されていた場合、各アルゴリズムから最上位のみが、結果について票を得る。一形態では、票のウェイトは、結果の各々、 $1 / (\text{検索におけるランク})$ 又は他の指標に対応する信頼度スコアに基づく。投票法及

50

び重み付けは、例えば次のような文献に記載されているブースティング (boosting) アルゴリズムを使って調整可能である: "A Short Introduction To Boosting," by Freund, Y. and Schapire, R., J. Japanese Society of Artificial Intelligence, v.14, no.5, pp.771-780, 1999。さらに、関連性が存在する場合、以下の何れかの方法を使って結果を出力することができる: a) 関連する結果のグループの信頼度を合計し、最高の信頼度スコアをもたらしたアルゴリズムによる結果 $R_{1, n}$ を出力する; 又は c) 出力するアルゴリズムの順序を (事前に測定した精度に基づいて) 予め決定しておく。例えば、最初が I J であり、次がスネークコーディングであり、等々。

【 0 1 2 7 】

他の形態では、検索アルゴリズムの信頼度スコア、検索結果、認識スコア、尤度スコア S_n 等の何らかの組み合わせ、及び / 又は認識スコアを計算するのに使用される属性の何らかの組み合わせにより、分類器 (例えば、ニューラルネットワーク、ナイーブベース分類器、ベイジアンベースの分類器、SVM ベースの分類器) が調整されてもよい。検索時間、算出されたスコア及び入力画像の属性は、認識結果及び信頼度スコアを算出する調整対象の分類器に入力される (認識結果等は、誤差の余裕、分類の確からしさ等に基づいてもよい。)

【 0 1 2 8 】

上記の実施形態に加えて、検索されたページ内の場所及び領域が、複数の認識結果の間で調整されてもよい。文書、ページ及び近似的なパッチロケーションが特定されると、本方法は、そのパッチの正確な場所を算出する。異なるアルゴリズムは異なるロケーション結果をもたらす可能性がある。アルゴリズムは、「ページ中の場所」の情報を生成しないかもしれないし、或いは、 x, y 座標のみを生成する、 x, y 座標と幅及び高さを生成する、四角形ではない (x_1, y_1, x_2, y_2) 座標を生成する、或いは多角形や閉曲線等で規定された任意の形状を指定してもよい。最終的な検索領域、すなわち x, y 中心は、アルゴリズムの精度に基づいてアルゴリズムの優先度を指定し、その優先度に基づいて検索領域を出力することで見出すことができる。代替的に、一形態における方法は、検索領域の質量中心を平均化することで、ページ内の検索位置を見出す、或いは、閉じた領域の縁総てに至るハウスドルフ距離を最小化する x, y 座標を見出す。ページ内の検索領域は、四角形や多角形 (等) で規定される領域を見出すことで規定され、その領域は、検索領域内の全体的なオーバーラップエラーを最小化する。

【 0 1 2 9 】

以上、本発明に関する実施例が、例示及び説明を意図してなされてきた。これらが網羅的であることや、本発明を開示された厳密な形態に限定することは、意図されていない。上記の教示内容により、多くの修正例や変形例が可能である。本発明の範囲は、この詳細な説明によって限定されるのではなく、特許請求の範囲によって規定されるように意図されている。当業者に理解されるように、本発明は、その精神又は本質的特徴から逸脱せず、他の具体的な形式で実現されてもよい。同様に、モジュール、ルーチン、特徴機能、属性、方法及び他の態様に関する具体的な名称や分け方は、必須でもなく重要でもない。本発明又は本発明の特徴機能を実現する手段 (仕組み) は、異なる名称や、異なる分け方及び / 又は形式でもよい。さらに、当業者には明らかであるように、モジュール、ルーチン、特徴、属性、方法及び他の本発明の態様は、ソフトウェア、ハードウェア又はファームウェアにより、又はそれら三者の如何なる組み合わせにより実現されてもよい。本発明のモジュールの一例のコンポーネントがソフトウェアで実現される場合、コンポーネントは、スタンドアロンプログラムとして、大規模なプログラム的一部分として、複数個の個々のプログラムとして、静的に又は動的にリンクしたライブラリとして、カーネルロードダブルモジュールとして、デバイスドライバとして実現されてもよく、及び / 又はコンピュータプログラミングの技術分野における通常の知識を有する者にとって現在又は将来知られる他の如何なる方法で実現されてもよい。さらに、本発明は、特定のどのプログラミング言語で実現されることにも限定されず、また、特手のどのオペレーティングシステム / 環境で実現することにも限定されない。したがって明細書の開示内容は例示的なもので

10

20

30

40

50

あるに過ぎず、特許請求の範囲で規定されている本発明の範囲を限定しようとするものではない。

【0130】

以下、本発明の実施例により教示される手段を例示的に列挙する。

【0131】

(1)

検索リクエストを受信し、認識結果を生成するシステムであって、
前記検索リクエストを受信する複数の入力を有するゲートウェイと、
前記ゲートウェイから問い合わせ画像及び認識パラメータを受信する照合部と
を有し、前記ゲートウェイは、前記問い合わせ画像及び前記認識パラメータを生成する
ように前記検索リクエストを処理し、前記ゲートウェイは、認識結果を送信する複数の出力を有し、

10

前記照合部は、前記問い合わせ画像を分析し、前記認識パラメータをインデックステーブルと対比し、前記認識結果を生成し、前記照合部は、前記認識結果を送信するように前記ゲートウェイに結合されている、システム。

【0132】

(2)

少なくとも1つの画像を生成し、該画像を前記照合部に与えるように結合されている発行部

をさらに有する第1項に記載のシステムであって、

20

前記照合部は画像登録部を有し、該画像登録部は、前記画像から画像特徴を抽出すること及びインデックステーブルを更新することを自動的に実行し、該画像登録部は、画像を受信するように前記発行部に結合されている、システム。

【0133】

(3)

前記検索リクエストが、問い合わせ画像、他のコンテキスト情報、及びメタデータのうちの1つ以上を含み、

前記問い合わせ画像が、何らかのフォーマットにおける画像又は画像内の1つ以上の特徴で表現され、

前記認識結果は、ブール値(真/偽)、ページID、ページ中の場所、アクション、文書、携帯可能な文書ファイルで定義されているアクション、アドレス、移動装置上で変換可能な情報、二値データ、追加的なアクションを含むメニュー、ラスト画像及び画像特徴を含む群中の1つ以上を含む、第1項に記載のシステム。

30

【0134】

(4)

プラグインモジュールを有する移動装置をさらに有する第1項に記載のシステムであって、前記移動装置は、画像を捕捉することが可能であり、前記移動装置は、前記ゲートウェイと通信することが可能であり、前記プラグインモジュールは、画像を処理し、捕捉した画像を該移動装置上で提示し、該捕捉した画像のユーザによる選択を受け付け、該捕捉した画像を前記移動装置から送信する、システム。

40

【0135】

(5)

前記認識結果に関する追加的な情報を有するホットスポットデータベースをさらに有する第1項に記載のシステムであって、前記ホットスポットデータベースは、前記認識結果を使って追加的な情報を検索するアクセスを受け、前記追加的な情報は、前記認識結果とともに前記ゲートウェイにより出力される、システム。

【0136】

(6)

複数の認識ユニットを有する取得部と、

問い合わせ画像及び認識パラメータを受信し、それらを前記取得部の特定の認識ユニッ

50

トに送信するディスパッチ部と

を前記照合部がさらに有し、前記認識ユニット各々は対応するインデックステーブルに結合され、前記認識ユニットは、前記問い合わせ画像の認識を実行し、前記インデックステーブルから前記認識結果を検索する、第1項に記載のシステム。

(7)

複数の認識ユニット及び対応するインデックステーブルは、インビジブルジャンクションアルゴリズムを実行する第1認識ユニット、ブリックウォールコーディングアルゴリズムを実行する第2認識ユニット、及びパスコーディングアルゴリズムを実行する第3認識ユニットを含む、第6項に記載のシステム。

【0137】

10

(8)

複数の認識ユニット及び対応するインデックステーブルは、第1の画像群に関する第1認識ユニット及びインデックステーブルと、第2の画像群に関する第2認識ユニット及びインデックステーブルとを有し、前記第1の画像群は前記第2の画像群と異なる、第6項に記載のシステム。

【0138】

(9)

前記照合部が、前記取得部の前記インデックステーブルを更新する画像登録モジュールを有し、該画像登録モジュールは、画像及びページ番号を受信し、特徴を抽出し、該特徴及びページ番号を前記取得部の前記インデックステーブルに加える、第6項に記載のシステム。

20

【0139】

(10)

前記照合部が、前記インデックステーブルを調整する負荷バランサを含み、該負荷バランサは、前記ディスパッチ部から送信された前記問い合わせ画像と、利用度を測定する前記取得部が出力した認識結果とを監視及び評価するように結合され、かつ前記画像登録モジュールへの制御信号を生成し、利用度に合うように前記インデックステーブルのコンフィギュレーションを修正する、第9項に記載のシステム。

【0140】

(11)

30

前記ディスパッチ部が、前記問い合わせ画像を受信する入力と、認識時間を示すオーダリング信号を出力する出力とを有する画像特徴オーダ部を含み、前記画像特徴オーダ部は、入力された問い合わせ画像を分析し、画像特徴を分析することで該入力された問い合わせ画像を認識するのに要する時間を予測する、第6項に記載のシステム。

【0141】

(12)

少なくとも1つの認識ユニットが、検索部に結合された特徴抽出部を含み、該特徴抽出部は、問い合わせ画像を受信する入力を有し、前記問い合わせ画像から画像特徴を生成し、それらを前記検索部に送信し、前記検索部は、前記画像特徴を前記インデックスと対比し、前記認識結果を生成する、第6項に記載のシステム。

40

【0142】

(13)

ミックスドメディアリアリティシステムにおいて動的に負荷バランスをとる方法であって、

問い合わせ画像及び認識結果を所定の期間監視するステップと、

取得部のインデックステーブルのコンフィギュレーションを指定する負荷バランス制御信号を生成するステップと、

前記負荷バランス制御信号に応じて、第1群のインデックステーブル及び第2群のインデックステーブルにコンテンツをロードするステップと

を有する方法。

50

【 0 1 4 3 】

(1 4)

前記問い合わせ画像及び前記認識結果が、ディスパッチ部及び取得部の間で伝送される、第 1 3 項に記載の方法。

【 0 1 4 4 】

(1 5)

負荷バランス制御信号を生成する前記ステップが、インデックステーブルで良好に発見された画像の分布を判定する、第 1 3 項に記載の方法

(1 6)

前記第 1 群のインデックステーブルを使用して、入力された問い合わせ画像を認識するステップと、

認識が成功したか否かを確認するステップと、

認識が成功しなかった場合、第 2 群のインデックステーブルを使用して前記問い合わせ画像を認識するステップと

をさらに有する第 1 3 項に記載の方法

(1 7)

前記第 1 群のインデックステーブルが現在の日付のコンテンツに対応し、前記第 2 群のインデックステーブルが過去の日付のコンテンツに対応する、第 1 6 項に記載の方法

(1 8)

複数の認識ユニットを有する取得部ヘディスパッチ部が問い合わせ画像を分配する際、画像特徴のオーダリングを行う方法であって、

問い合わせ画像を受信するステップと、

前記問い合わせ画像中の画像特徴を分析するステップと、

前記画像特徴を認識するのに要する時間を測定することで、予想時間を求めるステップと、

画像特徴及び予測時間の間に対応関係を付けるステップと、

前記問い合わせ画像を認識する前記取得部が要した実際の時間を測定するステップと、

前記実際の時間により前記対応関係を調整するステップと、

調整された対応関係を使用して問い合わせ画像の並べ替えを行い、問い合わせ画像を前記取得部の認識ユニットに割り当てるステップと

を有する方法。

【 0 1 4 5 】

(1 9)

前記問い合わせ画像中の画像特徴を分析する際に使用する画像特徴は、認識を実行する際に使用する画像特徴と異なる、第 1 8 項に記載の方法。

【 0 1 4 6 】

(2 0)

前記問い合わせ画像中の画像特徴を分析するステップが、複数の異なる特徴群を分析し、時間経過と共に関連性を測定する、第 1 8 項に記載の方法。

【 0 1 4 7 】

(2 1)

当該方法が、或る特徴群を複数の異なる特徴群から選択及び使用するステップをさらに含み、該或る特徴群は、最良の予測器を提供し、最低の演算コストを有する、第 2 0 項に記載の方法。

【 先行技術文献 】

【 非特許文献 】

【 0 1 4 8 】

【 非特許文献 1 】 “ A Computational Approach To Edge Detection, ” Canny, J., IEEE Trans. Pattern Analysis and Machine Intelligence, 8:679-714, 1986

【非特許文献2】“Restoring Ink Bleed-Through Degraded Document Images Using a Recursive Unsupervised Classification Technique,” P.Rosen, et al., Proceedings of the 6th British Machine Vision Conference, 1995。

【非特許文献3】“Perspective Correction Methods For Camera Based Document Analysis,” by Jagannathan et al., Proc. First Int. Workshop on Camera-based Document Analysis and Recognition, 2005。

【非特許文献4】“Text Detection from Natural Scene Images: Towards A System For Visually Impaired Persons,” N.Ezaki, et al., Proc. of 17th Int. Conf. on Pattern Recognition(ICPR 2004), IEEE Computer Society, 2004, pp.683-686, vol.II, 23-26 August, Cambridge, UK。 10

【非特許文献5】“Use Of The Hough Transformation To Detect Lines And Curves In Pictures,” R.Duda and P.Hart, Communications of the ACM, Vol.15, pp11-15, 1972

【非特許文献6】“A Survey Of Table Recognition,” by R Zanibbi, D Blostein, JR Cordy, International Journal on Document Analysis and Recognition, 2004;

【非特許文献7】“Face Detection: A Survey,” by E Hjelm, BK Low, Computer Vision and Image Understanding, 2001

【符号の説明】

【0149】 20

100 MMRシステム

102 a - 102 n 移動装置

104 MMRゲートウェイ

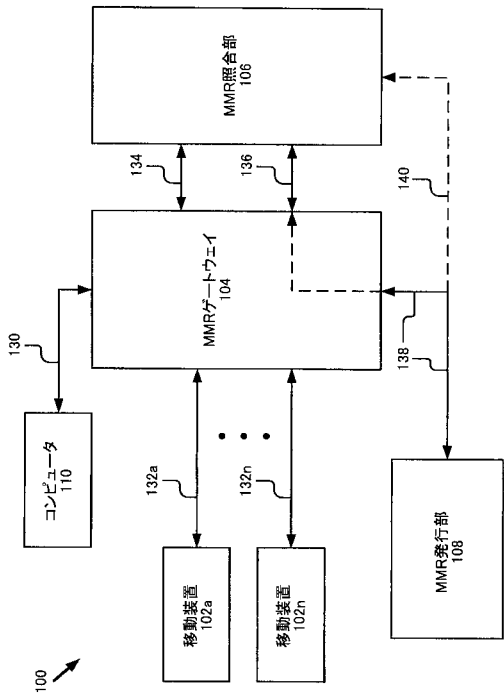
106 MMR照合部

108 MMR発行部

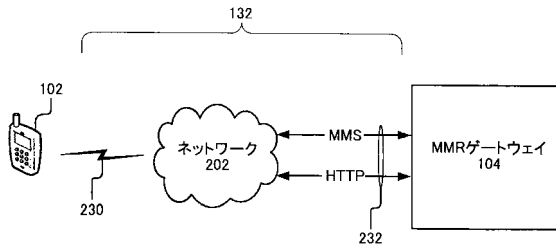
110 コンピュータ

132 a - 132 n 信号線

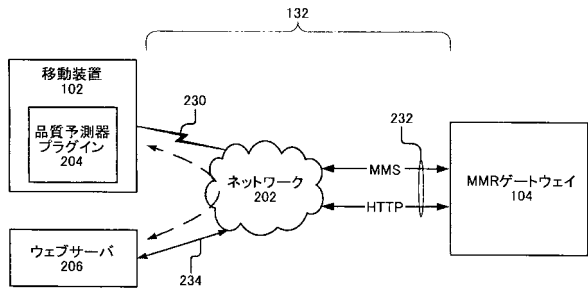
【図1】



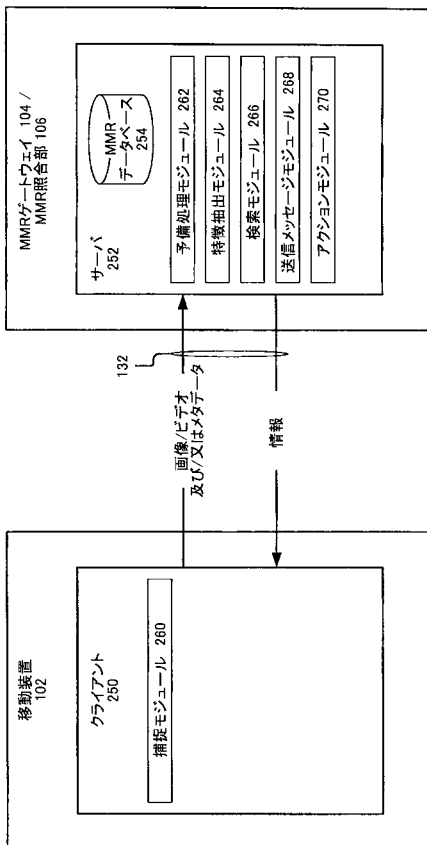
【図2A】



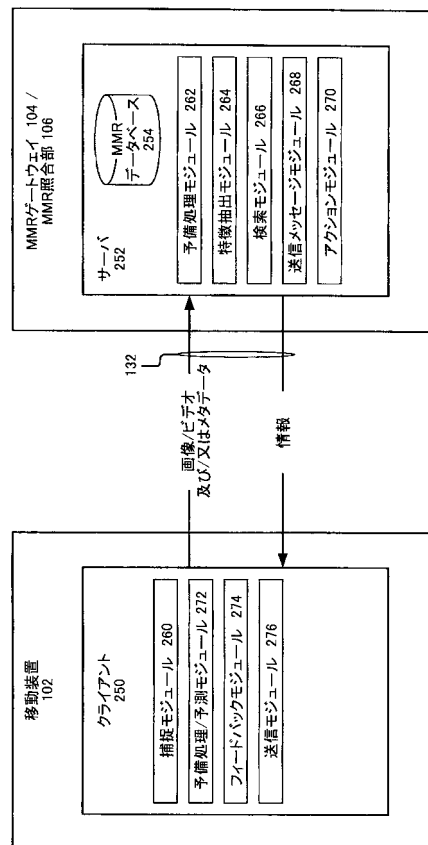
【図2B】



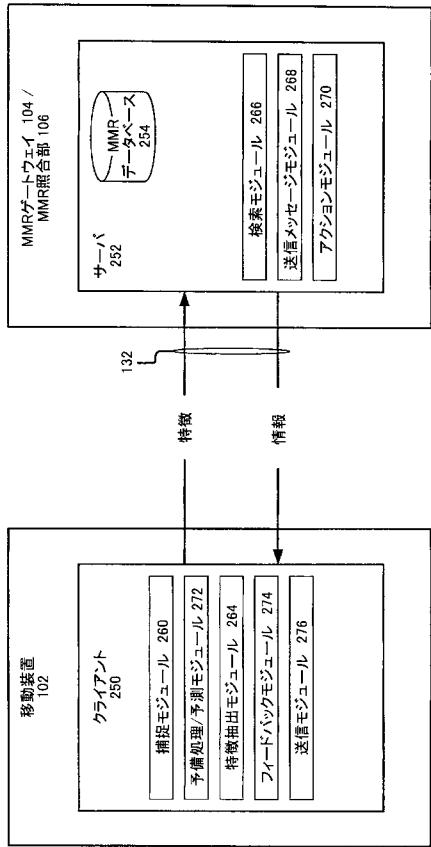
【図2C】



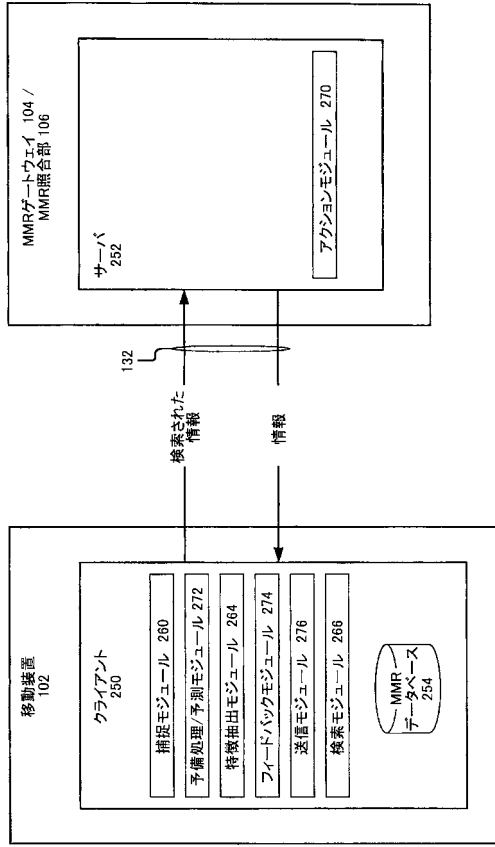
【図2D】



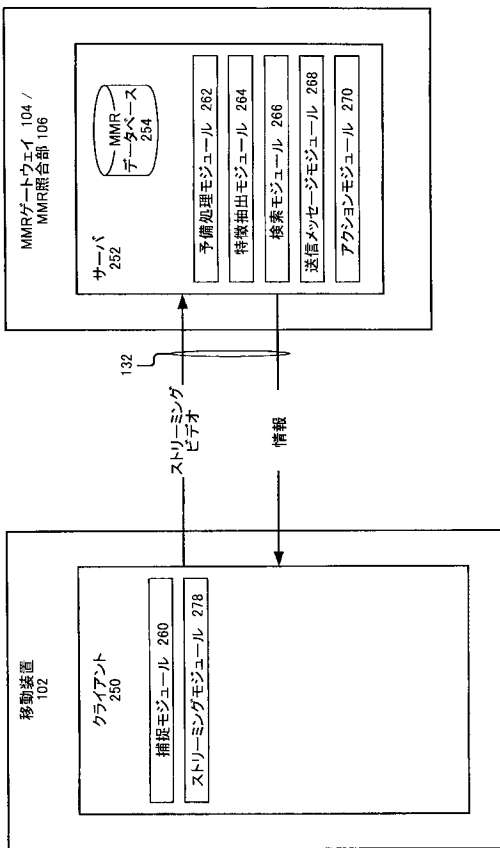
【図 2 E】



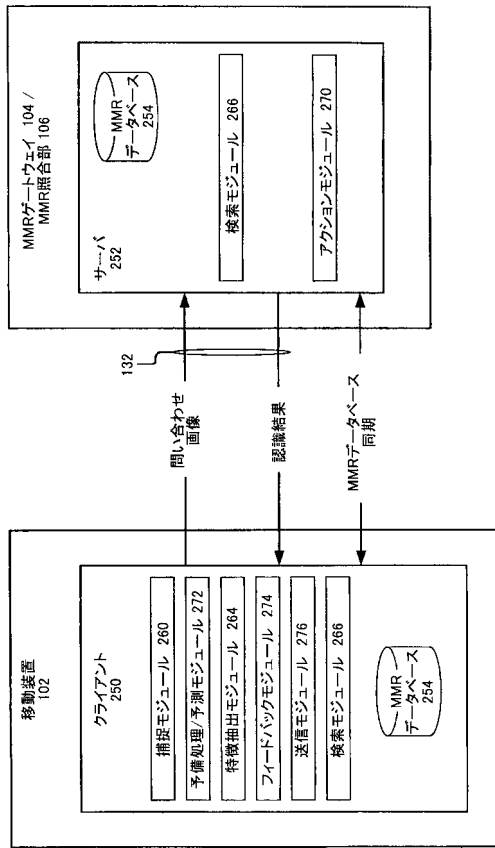
【図 2 F】



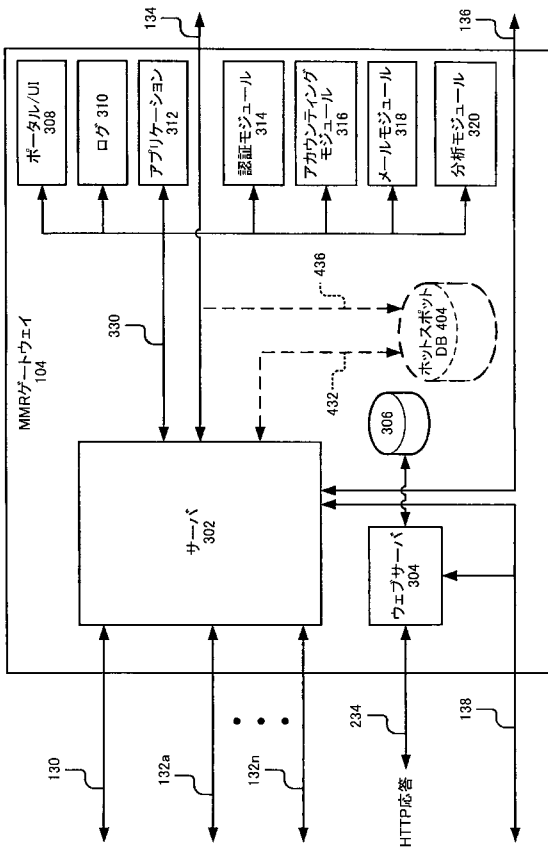
【図 2 G】



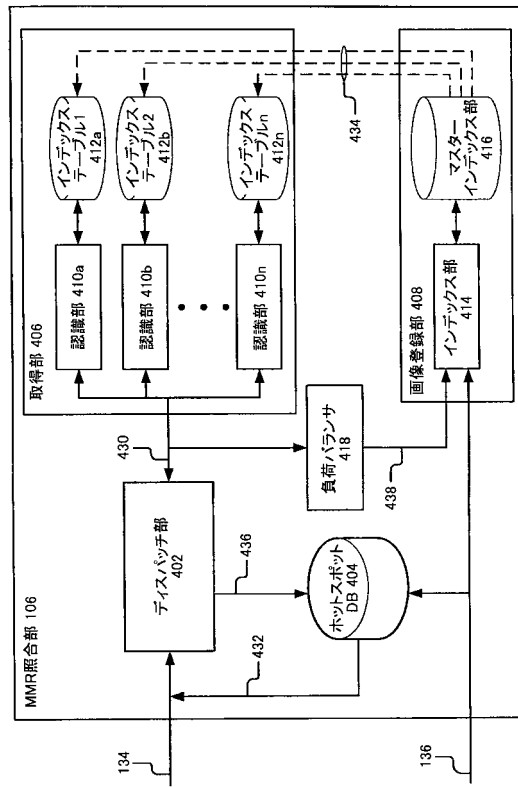
【図 2 H】



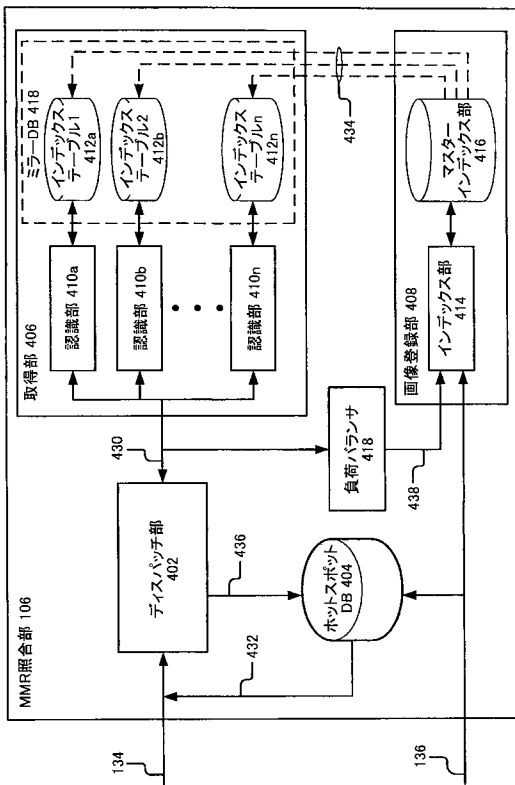
【図3】



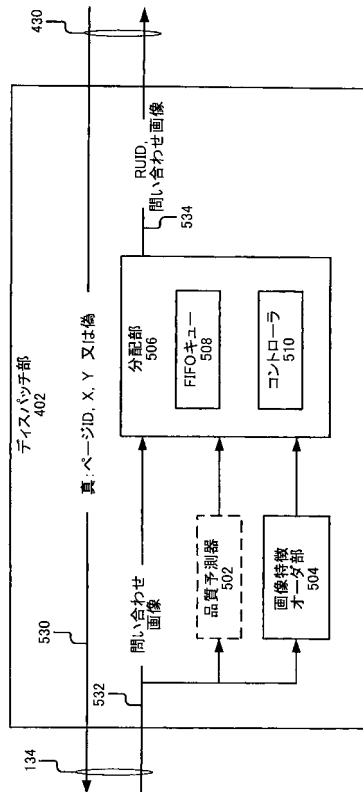
【図4A】



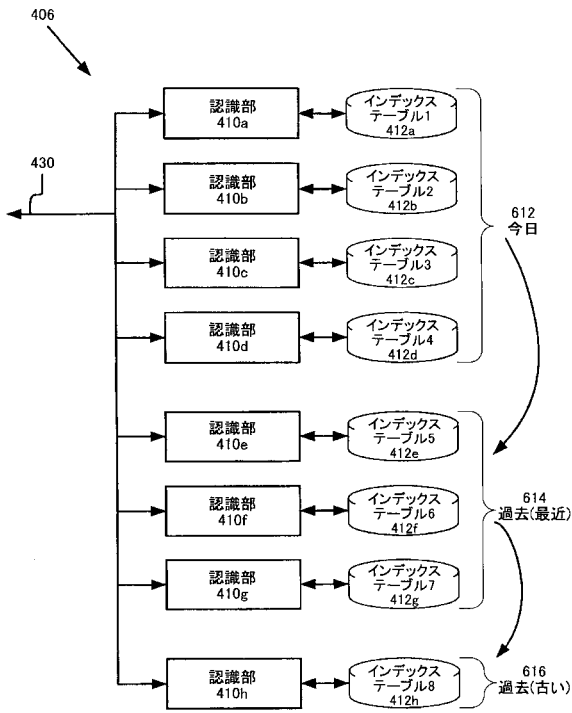
【図4B】



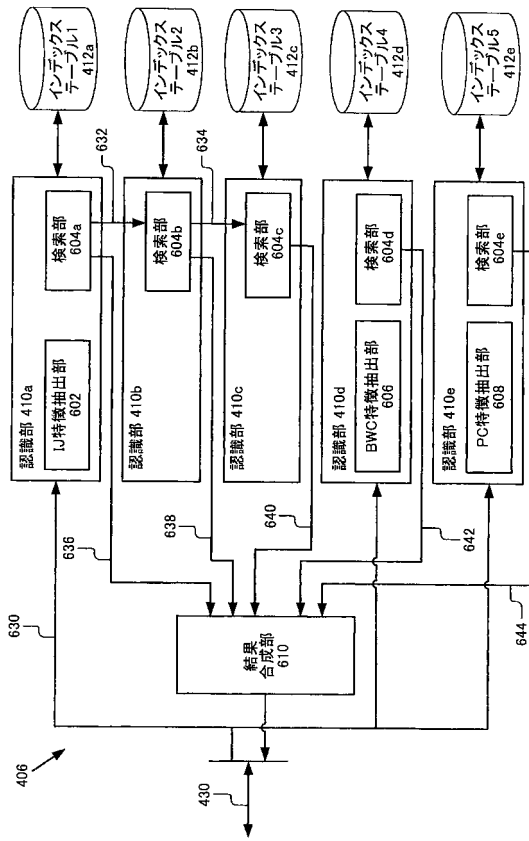
【図5】



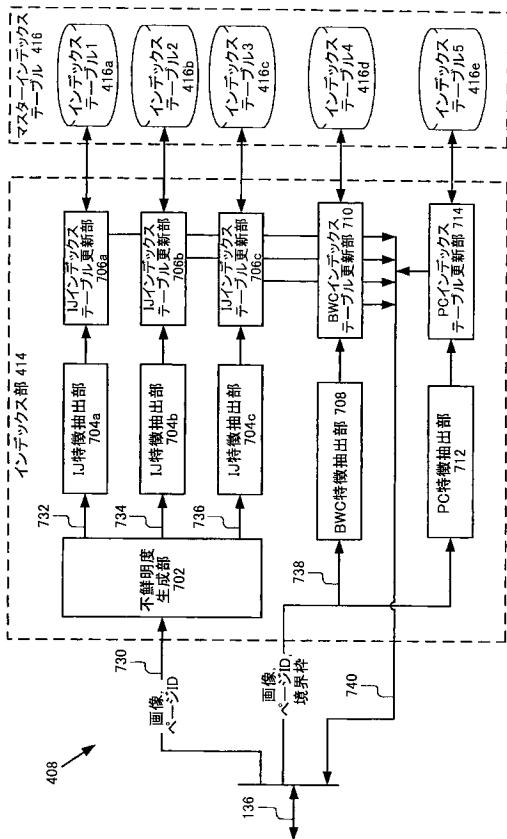
【図6A】



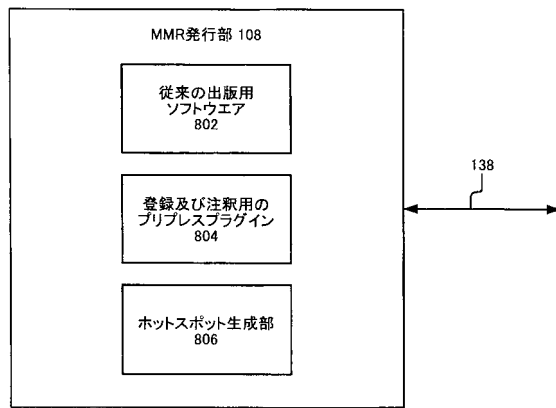
【図6B】



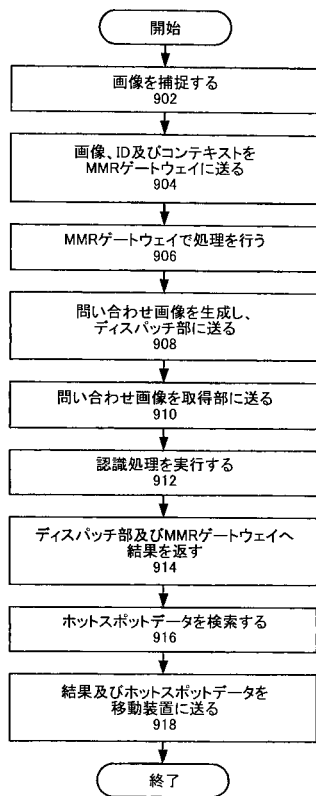
【図7】



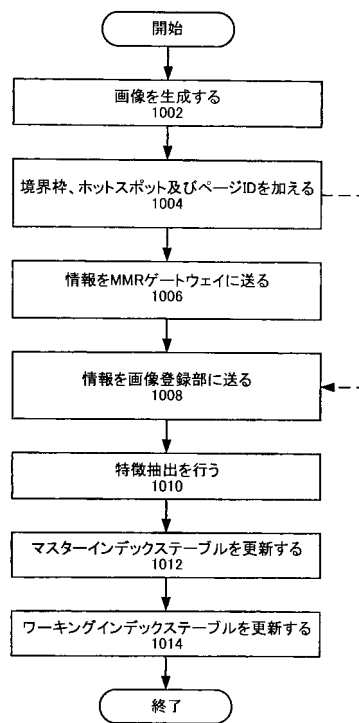
【図8】



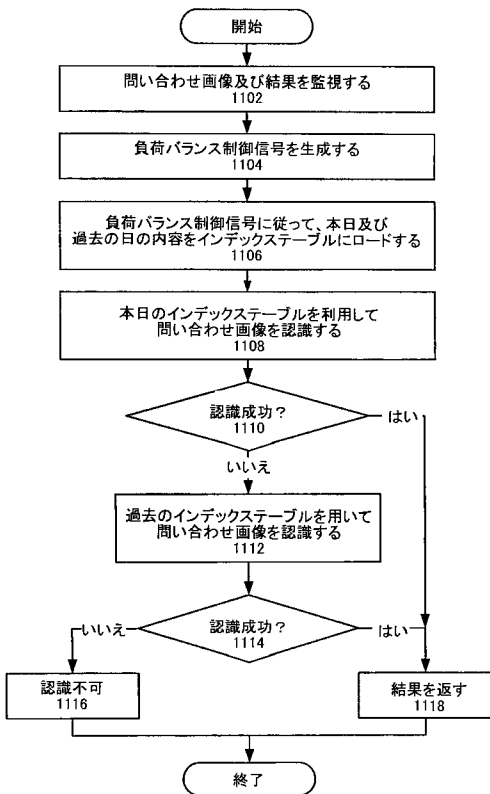
【図9】



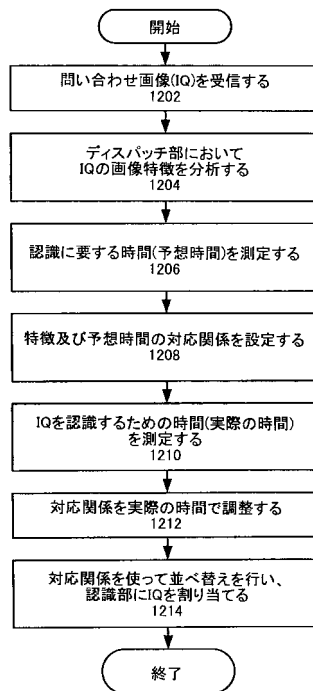
【図10】



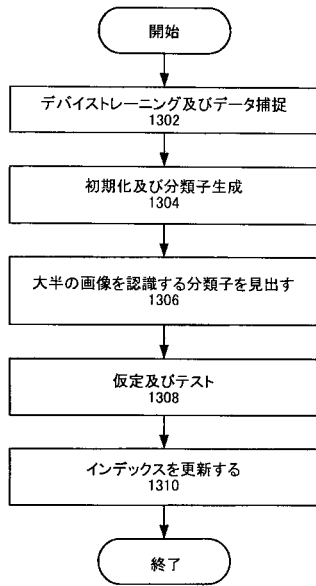
【図11】



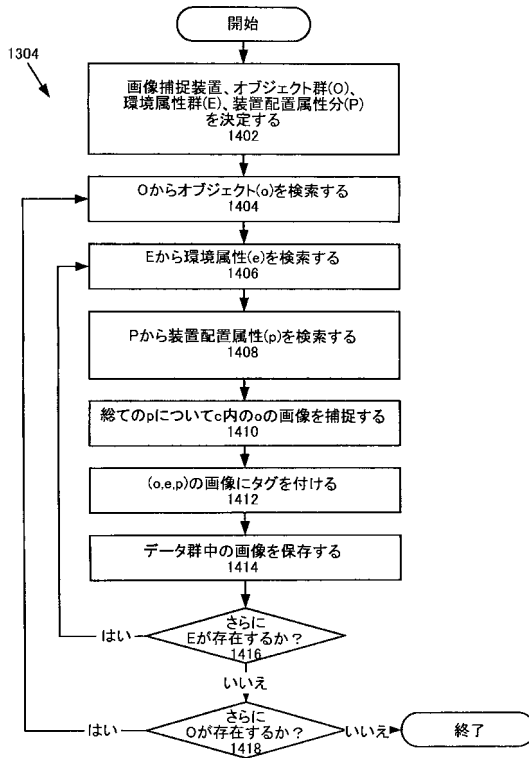
【図12】



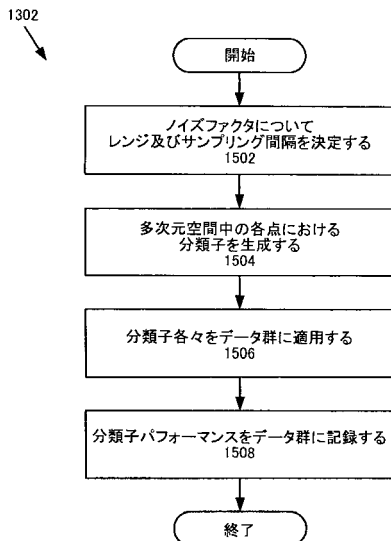
【図13】



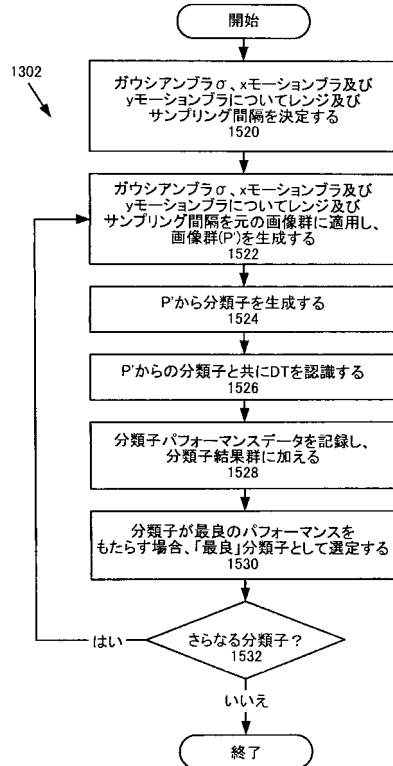
【図14】



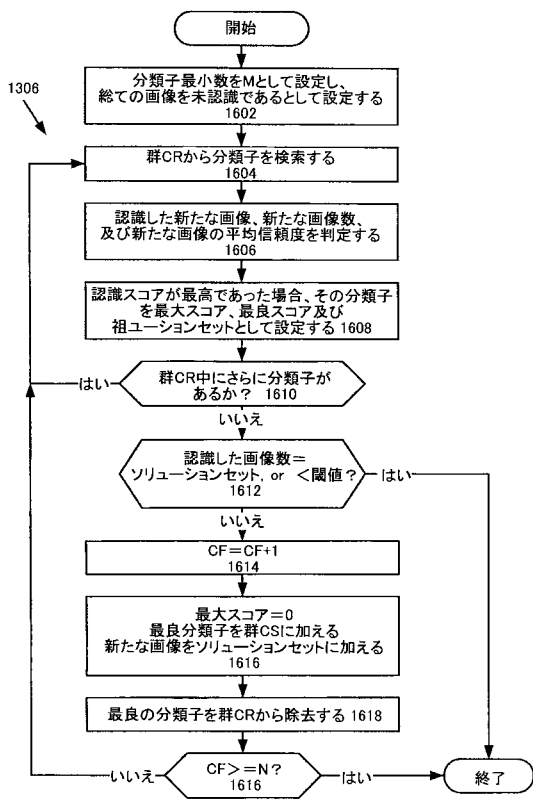
【図15A】



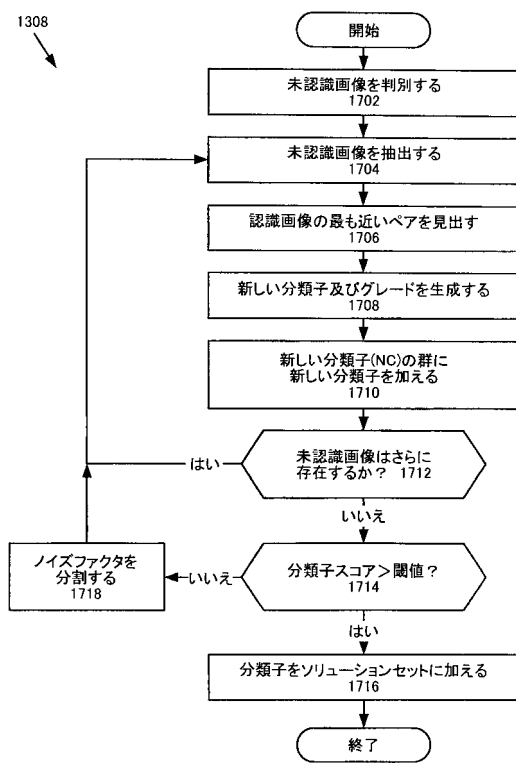
【図15B】



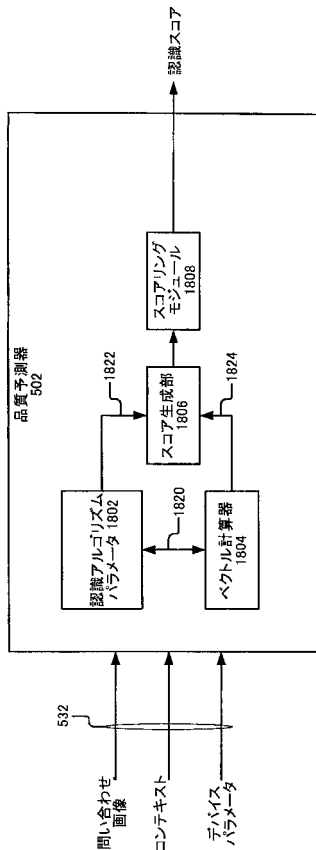
【図16】



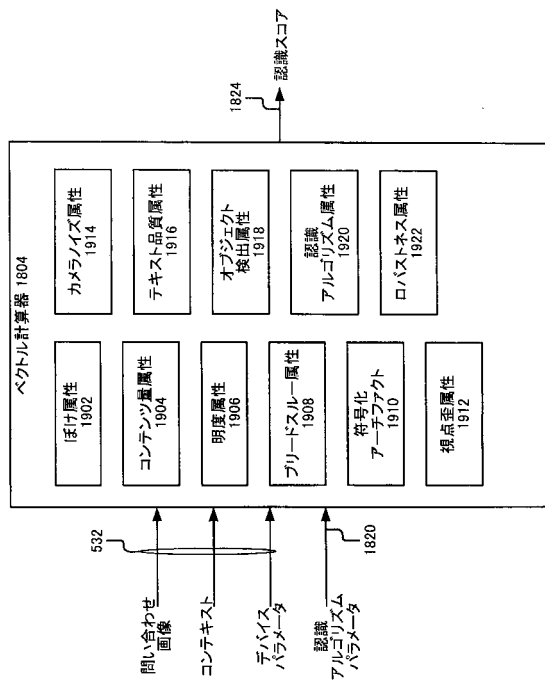
【図17】



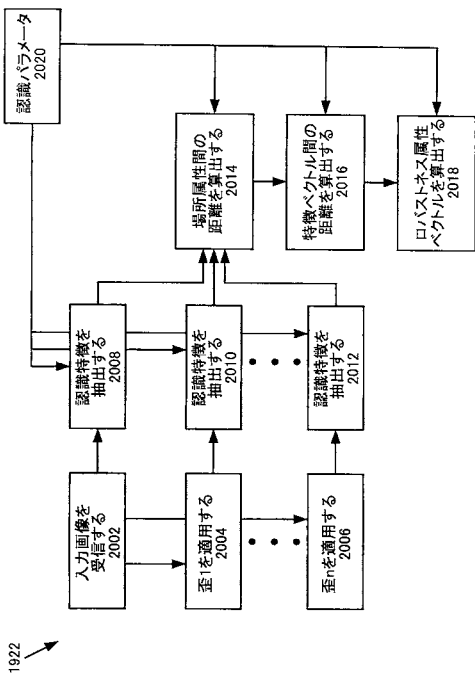
【図18】



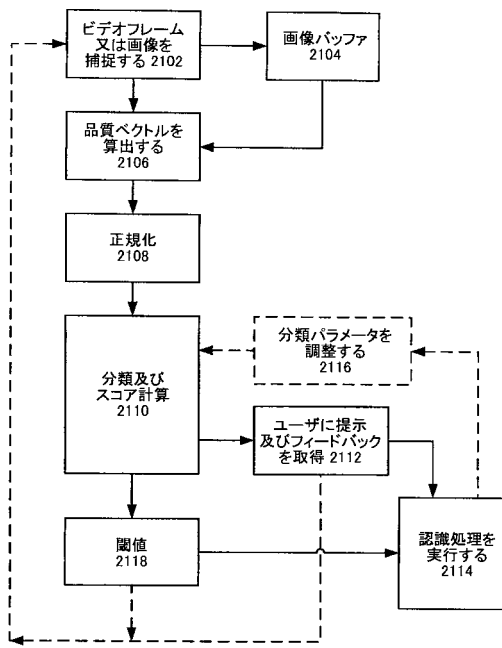
【図19】



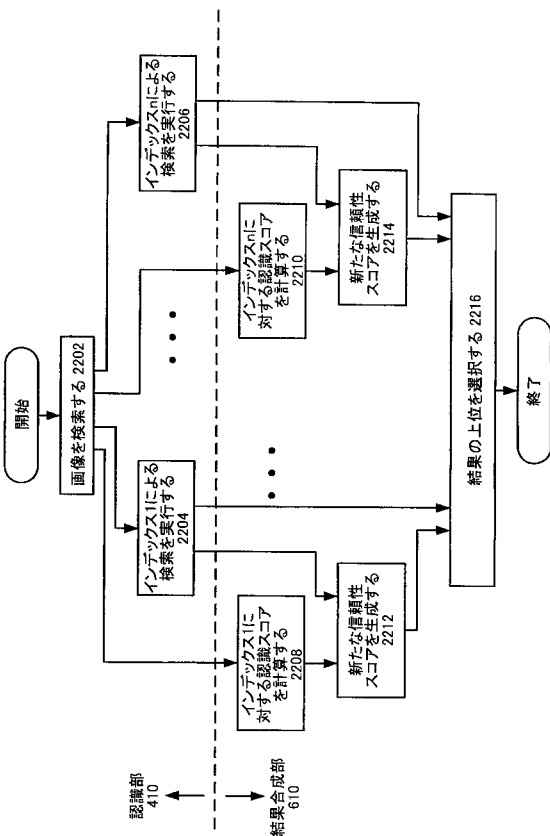
【図20】



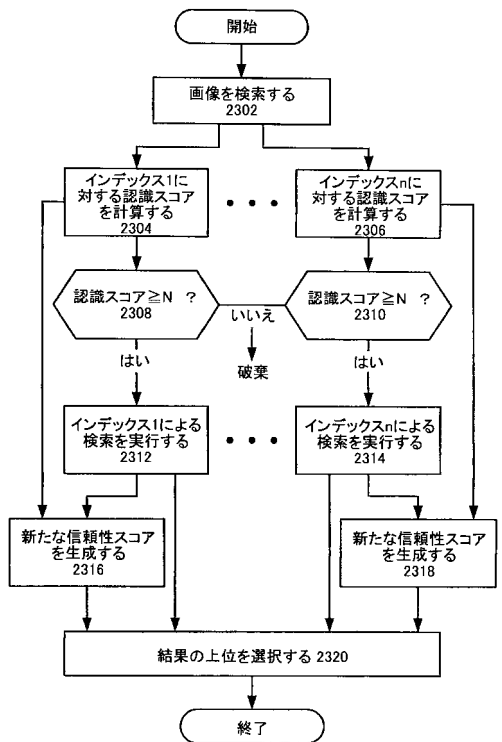
【図21】



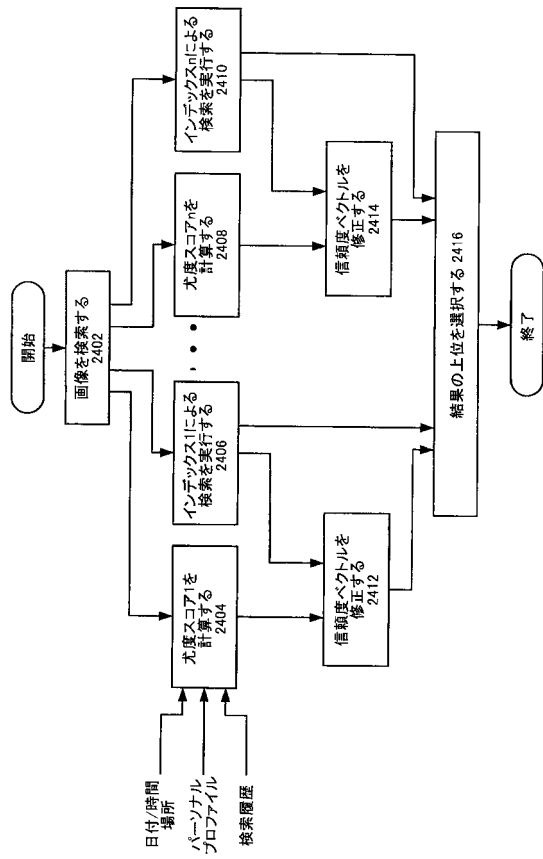
【図22】



【図23】



【 図 2 4 】



フロントページの続き

- (72)発明者 ベルナ エロール
アメリカ合衆国, カリフォルニア 94025 - 7054, メンロ・パーク, サンド・ヒル・ロード 2882番, スイート 115 リコー イノベーション インク内
- (72)発明者 ジョナサン ジェー ハル
アメリカ合衆国, カリフォルニア 94025 - 7054, メンロ・パーク, サンド・ヒル・ロード 2882番, スイート 115 リコー イノベーション インク内
- (72)発明者 ジョージ モラレダ
アメリカ合衆国, カリフォルニア 94025 - 7054, メンロ・パーク, サンド・ヒル・ロード 2882番, スイート 115 リコー イノベーション インク内
- (72)発明者 ジェイミー グラハム
アメリカ合衆国, カリフォルニア 94025 - 7054, メンロ・パーク, サンド・ヒル・ロード 2882番, スイート 115 リコー イノベーション インク内

審査官 久々宇 篤志

- (56)参考文献 特表2002 - 521752 (JP, A)
特表2005 - 157931 (JP, A)
国際公開第2007 / 023992 (WO, A1)
特開2006 - 53622 (JP, A)
国際公開第2006 / 092957 (WO, A1)
特開2001 - 92852 (JP, A)
特開2004 - 220424 (JP, A)

- (58)調査した分野(Int.Cl., DB名)
G06F 17 / 30