

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6148235号
(P6148235)

(45) 発行日 平成29年6月14日(2017.6.14)

(24) 登録日 平成29年5月26日(2017.5.26)

(51) Int.Cl. F I
G O 6 K 9/62 (2006.01) G O 6 K 9/62 6 1 0 Z

請求項の数 26 (全 34 頁)

| | | | |
|---------------|-------------------------------|-----------|---------------------|
| (21) 出願番号 | 特願2014-528410 (P2014-528410) | (73) 特許権者 | 507364838 |
| (86) (22) 出願日 | 平成24年8月6日(2012.8.6) | | クアルコム, インコーポレイテッド |
| (65) 公表番号 | 特表2014-529822 (P2014-529822A) | | アメリカ合衆国 カリフォルニア 921 |
| (43) 公表日 | 平成26年11月13日(2014.11.13) | | 21 サン ディエゴ モアハウス ドラ |
| (86) 国際出願番号 | PCT/US2012/049786 | | イブ 5775 |
| (87) 国際公開番号 | W02013/032639 | (74) 代理人 | 100108453 |
| (87) 国際公開日 | 平成25年3月7日(2013.3.7) | | 弁理士 村山 靖彦 |
| 審査請求日 | 平成26年3月25日(2014.3.25) | (74) 代理人 | 100163522 |
| 審判番号 | 不服2016-3223 (P2016-3223/J1) | | 弁理士 黒田 晋平 |
| 審判請求日 | 平成28年3月2日(2016.3.2) | (72) 発明者 | キューン・ホアン |
| (31) 優先権主張番号 | 61/528,741 | | アメリカ合衆国・カリフォルニア・921 |
| (32) 優先日 | 平成23年8月29日(2011.8.29) | | 21-1714・サン・ディエゴ・モアハ |
| (33) 優先権主張国 | 米国(US) | | ウス・ドライブ・5775 |
| (31) 優先権主張番号 | 13/450,016 | | |
| (32) 優先日 | 平成24年4月18日(2012.4.18) | | |
| (33) 優先権主張国 | 米国(US) | | |

最終頁に続く

(54) 【発明の名称】 コンテキスト情報を用いるカメラOCR

(57) 【特許請求の範囲】

【請求項1】

OCRを実行するための方法であって、
 デバイスに結合されたカメラを使用して画像を取得するステップと、
 少なくとも1つのグラフィカルオブジェクトを含む前記画像の一部分を識別するステッ
 プと、
 前記デバイスに結合されたマイクロフォンを使用して、音響入力を受信するステップと

、
 前記デバイスのロケーションを推定するステップと、
 前記デバイスの前記推定されたロケーションに基づいて、前記画像に関連するコンテキ
 ストを推測するステップと、

前記画像に関連するコンテキストに基づいてグラフィカルオブジェクトの群を選択する
 ステップと、

前記グラフィカルオブジェクトの群を使用して前記少なくとも1つのグラフィカルオブ
 ジェクトのOCR結果を改良するステップと

を含み、

前記デバイスの前記ロケーションが、前記マイクロフォンからの前記音響入力を使用し
 て、前記ロケーションに関連付けられた署名トーンを解析して推定される、方法。

【請求項2】

前記OCR結果を改良するステップが、前記グラフィカルオブジェクトの群を使用してOCR

10

20

を実行するステップを含む、請求項1に記載の方法。

【請求項3】

前記OCR結果を改良するステップが、

前記少なくとも1つのグラフィカルオブジェクトに対してOCRを実行することによって、前記少なくとも1つのグラフィカルオブジェクトに対する複数のOCR候補を生成するステップと、

前記グラフィカルオブジェクトの群を使用して、前記複数のOCR候補から1つのOCR候補を選択するステップと

を含む、請求項1に記載の方法。

【請求項4】

前記OCR結果を使用して、前記推測されたコンテキストを改良するステップをさらに含む、請求項1に記載の方法。

【請求項5】

前記少なくとも1つのグラフィカルオブジェクトが、シンボル、文字、単語、標識、および数のうちの1つまたは複数を含む、請求項1に記載の方法。

【請求項6】

前記グラフィカルオブジェクトの群が、共通の特性を通じて互いに関連する複数のグラフィカルオブジェクトを含む、請求項1に記載の方法。

【請求項7】

前記グラフィカルオブジェクトの群が辞書を表現する、請求項1に記載の方法。

【請求項8】

前記グラフィカルオブジェクトの群が辞書の改良版を表現する、請求項1に記載の方法。

。

【請求項9】

前記デバイスの前記ロケーションが、前記デバイスのGPS要素を使用して推定される、請求項1に記載の方法。

【請求項10】

前記デバイスの前記ロケーションが、セルタワーに関する信号の強度を使用して推定される、請求項1に記載の方法。

【請求項11】

プロセッサと、

画像を取得するためのカメラと、

前記プロセッサに結合された非一時的コンピュータ可読記憶媒体と

を備え、前記非一時的コンピュータ可読記憶媒体が、

デバイスに結合された前記カメラを使用して画像を取得するステップと、

少なくとも1つのグラフィカルオブジェクトを含む前記画像の一部分を識別するステップと、

前記デバイスに結合されたマイクロフォンを使用して、音響入力を受信するステップと

、

前記デバイスのロケーションを推定するステップと、

前記デバイスの前記推定されたロケーションに基づいて、前記画像に関連するコンテキストを推測するステップと、

前記画像に関連するコンテキストに基づいてグラフィカルオブジェクトの群を選択するステップと、

前記グラフィカルオブジェクトの群を使用して前記少なくとも1つのグラフィカルオブジェクトのOCR結果を改良するステップと

を含む方法を実装するために、前記プロセッサによって実行可能なコードを含み、

前記デバイスの前記ロケーションが、前記マイクロフォンからの前記音響入力を使用して、前記ロケーションに関連付けられた署名トーンを解析して推定される、デバイス。

【請求項12】

10

20

30

40

50

前記OCR結果を改良するステップが、前記グラフィカルオブジェクトの群を使用してOCRを実行するステップを含む、請求項 1 1 に記載のデバイス。

【請求項 1 3】

前記OCR結果を改良するステップが、

前記少なくとも1つのグラフィカルオブジェクトに対してOCRを実行することによって、前記少なくとも1つのグラフィカルオブジェクトに対する複数のOCR候補を生成するステップと、

前記グラフィカルオブジェクトの群を使用して、前記複数のOCR候補から1つのOCR候補を選択するステップと

を含む、請求項 1 1 に記載のデバイス。

10

【請求項 1 4】

前記OCR結果を使用して、前記推測されたコンテキストを改良するステップをさらに含む、請求項 1 1 に記載のデバイス。

【請求項 1 5】

前記少なくとも1つのグラフィカルオブジェクトが、シンボル、文字、単語、標識、および数のうちの1つまたは複数を含む、請求項 1 1 に記載のデバイス。

【請求項 1 6】

前記グラフィカルオブジェクトの群が、共通の特性を通じて互いに関連する複数のグラフィカルオブジェクトを含む、請求項 1 1 に記載のデバイス。

【請求項 1 7】

20

前記グラフィカルオブジェクトの群が辞書を表現する、請求項 1 1 に記載のデバイス。

【請求項 1 8】

前記グラフィカルオブジェクトの群が辞書の改良版を表現する、請求項 1 1 に記載のデバイス。

【請求項 1 9】

前記デバイスの前記ロケーションが、前記デバイスのGPS要素を使用して推定される、請求項 1 1 に記載のデバイス。

【請求項 2 0】

前記デバイスの前記ロケーションが、セルタワーに関する信号の強度を使用して推定される、請求項 1 1 に記載のデバイス。

30

【請求項 2 1】

プロセッサに結合された非一時的コンピュータ可読記憶媒体であって、

デバイスに結合されたカメラを使用して画像を取得するステップと、

少なくとも1つのグラフィカルオブジェクトを含む前記画像の一部を識別するステップと、

前記デバイスに結合されたマイクロフォンを使用して、音響入力を受信するステップと

、前記デバイスのロケーションを推定するステップと、

前記デバイスの前記推定されたロケーションに基づいて、前記画像に関連するコンテキストを推測するステップと、

40

前記画像に関連するコンテキストに基づいてグラフィカルオブジェクトの群を選択するステップと、

前記グラフィカルオブジェクトの群を使用して前記少なくとも1つのグラフィカルオブジェクトのOCR結果を改良するステップと

を含む方法を実装するために、前記プロセッサによって実行可能なコードを含み、

前記デバイスの前記ロケーションが、前記マイクロフォンからの前記音響入力を使用して、前記ロケーションに関連付けられた署名トーンを解析して推定される、非一時的コンピュータ可読記憶媒体。

【請求項 2 2】

前記OCR結果を改良するステップが、前記グラフィカルオブジェクトの群を使用してOCR

50

を実行するステップを含む、請求項 2 1 に記載の非一時的コンピュータ可読記憶媒体。

【請求項 2 3】

前記OCR結果を改良するステップが、

前記少なくとも1つのグラフィカルオブジェクトに対してOCRを実行することによって、前記少なくとも1つのグラフィカルオブジェクトに対する複数のOCR候補を生成するステップと、

前記グラフィカルオブジェクトの群を使用して、前記複数のOCR候補から1つのOCR候補を選択するステップと

を含む、請求項 2 1 に記載の非一時的コンピュータ可読記憶媒体。

【請求項 2 4】

OCRを実行するための装置であって、

デバイスに結合されたカメラを使用して画像を取得するための手段と、

少なくとも1つのグラフィカルオブジェクトを含む前記画像の一部を識別するための手段と、

前記デバイスに結合されたマイクロフォンを使用して、音響入力を受信するための手段と、

前記デバイスのロケーションを推定するための手段と、

前記デバイスの前記推定されたロケーションに基づいて、前記画像に関連するコンテキストを推測するための手段と、

前記画像に関連する前記コンテキストに基づいてグラフィカルオブジェクトの群を選択するための手段と、

前記グラフィカルオブジェクトの群を使用して前記少なくとも1つのグラフィカルオブジェクトのOCR結果を改良するための手段と

を含み、

前記デバイスの前記ロケーションが、前記マイクロフォンからの前記音響入力を使用して、前記ロケーションに関連付けられた署名トーンを解析して推定される、装置。

【請求項 2 5】

前記OCR結果を改良するための手段が、前記グラフィカルオブジェクトの群を使用してOCRを実行するための手段を含む、請求項 2 4 に記載の装置。

【請求項 2 6】

前記OCR結果を改良するための手段が、

前記少なくとも1つのグラフィカルオブジェクトに対してOCRを実行することによって、前記少なくとも1つのグラフィカルオブジェクトに対する複数のOCR候補を生成するための手段と、

前記グラフィカルオブジェクトの群を使用して、前記複数のOCR候補から1つのOCR候補を選択するための手段と、

前記OCR結果を改良するのに前記選択されたOCR候補を使用するための手段と

を含む、請求項 2 4 に記載の装置。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

関連出願の相互参照

本出願は、参照により本明細書に組み込まれる、2011年8月29日に出願した「CAMERA OCR WITH CONTEXT INFORMATION」と題する米国仮出願第61/528,741号の優先権を主張する。

【背景技術】

【0 0 0 2】

光学式文字認識(OCR)は、手で書かれた、タイプライタで打たれた、または印刷されたテキスト、グラフィックスまたはシンボルを走査した画像を機械的に符号化されたテキストに機械的または電子的に変換することである。OCRは、コンピュータビジョンおよび拡

10

20

30

40

50

張現実の適用における研究の成長分野である。コンピュータビジョンは、デバイスが、その近傍の環境を知覚することを可能にする。コンピュータビジョンは、エンドユーザの実世界との対話による感覚入力が仮想世界に及ぶことを可能にすることによって、拡張現実における適用を可能にする。実生活の例示的な用途では、コンピュータビジョン対応OCRは、視覚障害者が、スマートフォンなど、個人用モバイルデバイスのカメラの視野の中の書かれたテキストおよびシンボルを知覚するのを助け得る。

【 0 0 0 3 】

アナログデジタル変換の時代はまた、OCRの適用とその技術の改良を活気づけた。大部分の書類棚および大きな書類資料編成システムは、デジタル探索可能媒体(digital searchable media)に変換されている。OCR技術がいくら進歩しても、OCRは、依然として、限定された処理能力のモバイルデバイスには適していない。モバイルデバイスによってキャプチャされたカメラ画像からの文字およびシンボルを無制限に認識することは、異なる言語および異なる文化における単語、文字、およびシンボルの選択が莫大であるので、処理することは困難である。OCRは、コンピュータビジョンおよび拡張現実の適用をサポートするために、モバイルデバイス上で実時間でまたはほぼ実時間で動作する(run)必要があるので、計算の複雑度もまた、問題である。

【先行技術文献】

【特許文献】

【 0 0 0 4 】

【特許文献 1】米国特許出願第12/898,647号

【特許文献 2】米国仮出願第61/449,475号

【発明の概要】

【発明が解決しようとする課題】

【 0 0 0 5 】

本発明の実施形態は、これらおよび他の問題に対処する。

【課題を解決するための手段】

【 0 0 0 6 】

コンテキスト依存OCRを実行するための技法を提供する。本明細書で説明する技法は、限られた処理能力のモバイルデバイスに対して特に有用である。しかしながら、本明細書で説明する技法の用途は、モバイルデバイスに限定されず、すべてのOCR用途に適用可能である。音声、タッチ、およびユーザの使用パターンを含むユーザ入力とともに、マイクロフォン、GPS、およびカメラを含む様々なセンサからの入力が、ユーザコンテキストを推測し、推測されたコンテキストに最も関連する辞書を選択するために使用される。

【 0 0 0 7 】

OCRを実行するための方法の例は、デバイスに結合されたカメラを使用して画像を取得するステップと、少なくとも1つのグラフィカルオブジェクトを含む画像の一部分を識別するステップと、画像に関連するコンテキストを推測するステップと、画像に関連するコンテキストに基づいてグラフィカルオブジェクトの群を選択するステップと、グラフィカルオブジェクトの群を使用して少なくとも1つのグラフィカルオブジェクトのOCR結果を改良するステップとを含む。いくつかの実施形態では、OCR結果を改良するステップは、グラフィカルオブジェクトの群を使用してOCRを実行するステップを含む。別の実施形態では、OCR結果を改良するステップは、少なくとも1つのグラフィカルオブジェクトに対してOCRを実行することによって少なくとも1つのグラフィカルオブジェクトに対する複数のOCR候補を生成するステップと、グラフィカルオブジェクトの群を使用して複数のOCR候補から1つのOCR候補を選択するステップとを含む。方法は、OCR結果を使用して推測されたコンテキストを改良するステップをさらに含み得る。

【 0 0 0 8 】

そのような方法の実装形態は、以下の特徴のうちの1つまたは複数を含み得る。グラフィカルオブジェクトは、シンボル、文字、単語、標識、および数のうちの1つまたは複数を含み得る。グラフィカルオブジェクトの群は、共通の特性を介して互いに関連する複数

10

20

30

40

50

のグラフィカルオブジェクトを含み得、辞書または辞書の改良版を表現し得る。本発明のいくつかの態様では、画像のコンテキストは非テキストである。他の態様では、コンテキストはグラフィカル表示を使用して推測され得、グラフィカル表示は、オブジェクトのジェスチャ表現およびグラフィカル表現を含む群のうちの少なくとも1つからであってよい。画像のコンテキストはまた、センサ入力を使用して推測され得る。センサ入力は、カメラ、マイクロフォン、光センサ、時計、およびGPS要素のうちの1つまたは複数によって与えられ得る。コンテキストはまた、カメラと少なくとも1つのグラフィカルオブジェクトとの間の距離を推定するステップと、カメラと少なくとも1つのグラフィカルオブジェクトとの間の推定された距離に基づいてコンテキストを推測するステップとによって推測され得る。コンテキストを推測するステップはまた、画像が取得される環境内の複数の光線に関連する少なくとも1つの特性を検出するステップと、複数の光線に関連する少なくとも1つの特性に基づいてコンテキストを推測するステップとを含み得る。さらに、複数の光線に関連する少なくとも1つの特性は明度(brightness)であってよく、しきい値より低い明度は、画像を屋内で取得しているとの推測をもたらし、しきい値を超える明度は、画像を屋外で取得しているとの推測をもたらす。

【0009】

他の実装形態では、コンテキストを推測するステップは、マイクロフォンを使用して音響入力を受信するステップと、音響入力に対してパターン認識を実行するステップと、音響入力内で認識されたパターンに基づいてコンテキストを推測するステップとを含み得る。認識されたパターンは、ユーザからの音声コマンド、または音響入力を受信した時点のカメラのロケーションの表示であってよい。コンテキストを推測するステップはまた、ユーザからコンテキスト関連情報を受信するステップと、ユーザから受信したコンテキスト関連情報に基づいてコンテキストを推測するステップとを含み得る。コンテキストはまた、デバイスのロケーションを推定するステップと、推定されたデバイスのロケーションに基づいてコンテキストを推測するステップとを含み得るデバイスのロケーションに基づいて推測され得る。デバイスのロケーションは、デバイスのGPS要素、セルタワーに関する信号の強度、またはマイクロフォンからの音響入力を使用して推定され得る。

【0010】

いくつかの実施形態では、コンテンツ意味解析および構造解析が、同様に、画像に関連するコンテキストを解析するのに使用され得る。構造レイアウト解析(structural layout analysis)では、グラフィカルオブジェクト間の関係が、互いに、またはグラフィカル表示と比較され得る。グラフィカル表示は、オブジェクトのジェスチャ表現またはグラフィカル表現であってよい。コンテンツ意味解析では、グラフィカルオブジェクト相互間、またはグラフィカルオブジェクトとグラフィカル表示との共起が解析され得る。

【0011】

さらに、コンテンツの語義は、画像中に複数の言語で表示されるテキストなど、グラフィカルオブジェクトを使用して解析され得る。一実施形態では、第1の言語における第1のグラフィカルオブジェクトと第2の言語における第2のグラフィカルオブジェクトとを含む画像の1つまたは複数の部分が識別され、第1のグラフィカルオブジェクトおよび第2のグラフィカルオブジェクトは、類似する意味を有する。第1のグラフィカルオブジェクトに対する第1の複数の候補および第2のグラフィカルオブジェクトに対する第2の複数の候補がまた識別され、グラフィカルオブジェクトが、第2の複数の候補のうちの少なくとも1つと類似の意味を有する第1の複数の候補から選択される。さらに、第1の複数の候補からのグラフィカルオブジェクトは、第2のグラフィカルオブジェクトに対する第2の複数の候補を第1の言語に変換するステップと、変換された第2の複数の候補に類似する意味を有する少なくとも1つのグラフィカルオブジェクトを第1の複数の候補から発見するステップとを含み得る。

【0012】

OCRを実行するためのデバイスの一例は、プロセッサと、画像を取得するためのカメラと、プロセッサに結合された非一時的コンピュータ可読記憶媒体とを含み得、非一時的コ

10

20

30

40

50

ンピュータ可読記憶媒体は、デバイスに結合されたカメラを使用して画像を取得するステップと、少なくとも1つのグラフィカルオブジェクトを含む画像の一部分を識別するステップと、画像に関連するコンテキストを推測するステップと、画像に関連するコンテキストに基づいてグラフィカルオブジェクトの群を選択するステップと、グラフィカルオブジェクトの群を使用して少なくとも1つのグラフィカルオブジェクトのOCR結果を改良するステップとを含む方法を実装するための、プロセッサによって実行可能なコードを含む。OCR結果を改良するステップは、グラフィカルオブジェクトの群を使用してOCRを実行するステップを含み得る。OCR結果を改良するステップはまた、少なくとも1つのグラフィカルオブジェクトに対してOCRを実行することによって少なくとも1つのグラフィカルオブジェクトに対する複数のOCR候補を生成するステップと、グラフィカルオブジェクトの群を使用して複数のOCR候補から1つのOCR候補を選択するステップとを含み得る。デバイスは、OCR結果を使用して推測されたコンテキストをさらに改良し得る。そのようなデバイスの実装形態は、他の例に関して上記で説明した特徴のうちの1つまたは複数を含み得る。

【0013】

本発明の実施形態の追加の一例は、プロセッサに結合された非一時的コンピュータ可読記憶媒体を含み、非一時的コンピュータ可読記憶媒体は、デバイスに結合されたカメラを使用して画像を取得するステップと、少なくとも1つのグラフィカルオブジェクトを含む画像の一部分を識別するステップと、画像に関連するコンテキストを推測するステップと、画像に関連するコンテキストに基づいてグラフィカルオブジェクトの群を選択するステップと、グラフィカルオブジェクトの群を使用して少なくとも1つのグラフィカルオブジェクトのOCR結果を改良するステップとを含み得る方法を実装するための、プロセッサによって実行可能なコードを含む。OCR結果を改良するステップは、グラフィカルオブジェクトの群を使用してOCRを実行するステップを含む。OCR結果を改良するステップはまた、少なくとも1つのグラフィカルオブジェクトに対してOCRを実行することによって少なくとも1つのグラフィカルオブジェクトに対する複数のOCR候補を生成するステップと、グラフィカルオブジェクトの群を使用して複数のOCR候補から1つのOCR候補を選択するステップとを含み得る。そのような装置の実装形態は、他の例に関して上記で説明した特徴のうちの1つまたは複数を含み得る。

【0014】

OCRのための方法を実行するための装置のさらに別の例は、デバイスに結合されたカメラを使用して画像を取得するための手段と、少なくとも1つのグラフィカルオブジェクトを含む画像の一部分を識別するための手段と、画像に関連するコンテキストを推測するための手段と、画像に関連するコンテキストに基づいてグラフィカルオブジェクトの群を選択するための手段と、グラフィカルオブジェクトの群を使用して少なくとも1つのグラフィカルオブジェクトのOCR結果を改良するための手段とを含み得る。OCR結果を改良するステップは、グラフィカルオブジェクトの群を使用してOCRを実行するステップを含む。OCR結果を改良するステップはまた、少なくとも1つのグラフィカルオブジェクトに対してOCRを実行することによって少なくとも1つのグラフィカルオブジェクトに対する複数のOCR候補を生成するステップと、グラフィカルオブジェクトの群を使用して複数のOCR候補から1つのOCR候補を選択するステップとを含み得る。そのような装置の実装形態は、他の例に関して上記で説明した特徴のうちの1つまたは複数を含み得る。

【0015】

上記は、以下の詳細な説明をよりよく理解できるように、本開示による実施例の特徴および技術的利点をかなり広く概説したものである。さらなる特徴および利点は、以下で説明される。本開示の同じ目的を実施するために他の構造を修正または設計する際の基礎として、開示される概念および特定の実施例が容易に利用され得る。そのような等価な構造は、添付の特許請求の範囲の趣旨および範囲から逸脱しない。構成と動作方法の両方に関する、本明細書で開示される概念の特徴であると考えられる特徴は、関連する利点とともに、添付の図面とともに検討すると以下の説明からよりよく理解されよう。しかし、図面の各々は例示および説明のみのために与えられ、特許請求の範囲の限定範囲を定めるため

に与えられるものではない。

【0016】

以下の説明は、図面を参照しながら与えられ、全体を通して、同じ参照番号は同じ要素を指すために使用される。本明細書で1つまたは複数の技法の様々な詳細を説明するが、他の技法も同様に可能である。いくつかの場合には、様々な技法の説明を容易にするために、周知の構造およびデバイスがブロック図の形式で示されている。

【0017】

明細書の残りの部分および図面を参照することによって、本開示によって与えられる実施例の性質および利点をさらに理解することができ、同様の参照番号が、同様の構成要素を指すために、いくつかの図面全体で用いられる。いくつかの場合には、複数の同様の構成要素の1つを指すために、サブラベルが参照番号と関連付けられる。存在するサブラベルを特定することなく、参照番号に言及する場合、参照番号はそのようなすべての同様の構成要素を指す。

【図面の簡単な説明】

【0018】

【図1】本発明の実施形態の実行に使用されるデバイスの部分を組み込む例示的なコンピュータシステムを示す図である。

【図2A】コンテキスト依存OCRのためにモバイルデバイスによってキャプチャされた画像の例示的な表現を示す図である。

【図2B】コンテキスト依存OCRのためにモバイルデバイスによってキャプチャされた画像の例示的な表現を示す別の図である。

【図3】OCRのためのグラフィカルオブジェクトの群を選択するための、コンテキスト情報を使用する、非限定的な例示的デシジョンツリーを示すフロー図である。

【図4】コンテキスト依存OCRのための例示的な実施形態を説明するフロー図である。

【図5】コンテキスト依存OCRのための方法を実行するためのさらに別の例示的な実施形態を説明するフロー図である。

【図6】コンテキスト依存OCRのための方法を実行するための例示的な実施形態を説明するフロー図である。

【図7】光の状態を使用するコンテキスト依存OCRのための別の例示的な実施形態を説明するフロー図である。

【図8】OCRのためのグラフィカルオブジェクトの群を選択するためのコンテキストがマイクロフォンからの入力に基づく、例示的な実施形態を例示するフロー図である。

【図9】コンテキスト依存OCRのために手動入力を使用するための実施形態を説明するフロー図である。

【図10】画像がキャプチャされるロケーションを使用してコンテキスト依存OCRを実施するための方法を実行するための別の実施形態を説明するフロー図である。

【図11】コンテキスト依存OCRのための辞書を選択するためにマイクロフォンを使用して環境を認識する、別の例示的な実施形態を説明するフロー図である。

【図12】コンテンツ解析を使用するコンテキスト識別のための一実施形態を説明するフロー図である。

【図13】画像の構造レイアウトを解析することによるコンテキスト識別のための一実施形態を説明するフロー図である。

【図14】コンテンツ解析を使用するコンテキスト識別のための一実施形態を説明するフロー図である。

【図15】挙動要素を使用するコンテキスト依存OCRのための一実施形態を説明するフロー図である。

【図16】複数の言語における例示的な交通標識のブロック図である。

【図17】標識が2つ以上の言語におけるときの、OCR結果の正確さおよび速さを向上するための方法を説明するフロー図である。

【発明を実施するための形態】

【 0 0 1 9 】

図1に示すコンピュータシステムは、前に説明したコンピュータ化デバイスの一部分として組み込まれ得る。たとえば、コンピュータシステム1200は、モバイルデバイスの構成要素のいくつかを表現し得る。モバイルデバイスは、カメラのような入力感覚ユニットとディスプレイユニットとを有する任意のコンピューティングデバイスであってよい。モバイルデバイスの例には、限定はしないが、ビデオゲーム機、タブレット、スマートフォン、および任意の他のハンドヘルドデバイスがある。図1は、本明細書で説明する様々な他の実施形態によって提供される方法を実行し得、かつ/あるいは、ホストコンピュータシステム、リモートキオスク/端末、販売時点管理(point-of-sale)デバイス、モバイルデバイス、セットトップボックスおよび/またはコンピュータシステムとして機能し得る、コンピュータシステム100の一実施形態の概略図を与える。図1は、単に、様々な構成要素の一般化された図を提供することを意図されており、様々な構成要素の一部または全部は、適宜利用され得る。したがって、図1は、個別のシステム要素が、比較的分離方式でまたは比較的より一体方式で、どのように実装され得るかを広範囲に示している。

10

【 0 0 2 0 】

バス105を介して電氣的に結合され得る(または場合によっては適宜通信中であり得る)ハードウェア要素を備えている、コンピュータシステム100が示されている。ハードウェア要素には、1つまたは複数の汎用プロセッサおよび/または1つまたは複数の専用プロセッサ(デジタル信号処理チップ、グラフィックス高速化プロセッサなど)を限定なしに含む1つまたは複数のプロセッサ110と、カメラ、センサ(慣性センサを含む)、マウス、キーボードなどを限定なしに含む1つまたは複数の入力デバイス115と、ディスプレイユニット、プリンタなどを限定なしに含む1つまたは複数の出力デバイス120とがあり得る。

20

【 0 0 2 1 】

コンピュータシステム100は、さらに、ローカル記憶装置および/またはネットワークアクセス可能な記憶装置を限定なしに備え得、かつ/あるいはプログラム可能、フラッシュアップデート可能などであり得るディスクドライブ、ドライブアレイ、光ストレージデバイス、ランダムアクセスメモリ(「RAM」)および/またはリードオンリーメモリ(「ROM」)などのソリッドステートストレージデバイスを限定なしに含む1つまたは複数の非一時的ストレージデバイス125をさらに含む(かつ/または非一時的ストレージデバイス125と通信中であり得る)。そのようなストレージデバイスは、様々なファイルシステム、データベース構造などを限定なしに含む任意の適したデータストレージを実装するように構成され得る。

30

【 0 0 2 2 】

コンピュータシステム100はまた、モデム、ネットワークカード(ワイヤレスまたはワイヤード)、赤外通信デバイス、ワイヤレス通信デバイス、および/またはチップセット(Bluetooth(登録商標)デバイス、802.11デバイス、Wi-Fiデバイス、WiMaxデバイス、セルラー通信機器など)などを限定なしに含む通信サブシステム130を含む。通信サブシステム130は、データが、ネットワーク(一例をあげると、以下で説明するネットワークなど)、他のコンピュータシステム、および/または本明細書で説明する任意の他のデバイスと交換されることを可能にし得る。多くの実施形態では、コンピュータシステム100は、上記で説明したRAMまたはROMを含む非一時的作業メモリ135をさらに備える。

40

【 0 0 2 3 】

コンピュータシステム100はまた、本明細書で説明するように、オペレーティングシステム140、デバイスドライバ、実行可能ライブラリ、および/または1つまたは複数のアプリケーションプログラム145などの他のコードを含む作業メモリ135内に現在配置されているように示されるソフトウェア要素を備え得、アプリケーションプログラム145は、様々な実施形態によって提供されるコンピュータプログラムを含み得、かつ/または方法を実装するように設計され得、かつ/または他の実施形態によって提供されるシステムを構成し得る。単に例として、上述の方法に関して説明した1つまたは複数のプロシージャは、コンピュータ(および/またはコンピュータ内のプロセッサ)によって実行可能なコードお

50

よび/または命令として実装され得、したがって、一態様では、そのようなコードおよび/または命令は、説明した方法による1つまたは複数の動作を実行するために汎用コンピュータ(または他のデバイス)を構成および/または適用するために使用され得る。

【0024】

これらの命令および/またはコードのセットは、上記で説明したストレージデバイス125など、コンピュータ可読記憶媒体上に記憶され得る。場合によっては、ストレージ媒体は、コンピュータシステム100などのコンピュータシステム内に組み込まれ得る。他の実施形態では、ストレージ媒体が、命令/コードを記憶された汎用コンピュータをプログラム、構成、および/または適用するために使用され得るように、ストレージ媒体は、コンピュータシステムから分離され得(たとえば、コンパクトディスクなどの取外し可能媒体)、かつ/またはインストールパッケージ内に備えられ得る。これらの命令は、コンピュータシステム100で実行可能であり、かつ/または、(たとえば、様々な一般的に利用できるコンパイラ、インストールプログラム、圧縮/解凍ユーティリティなどのいずれかを使用して)コンピュータシステム100上にコンパイルおよび/またはインストールすると、実行コードの形態を取るソースコードおよび/またはインストール可能コード(installable code)の形態を取り得る、実行コードの形態を取り得る。

10

【0025】

大幅な変形が、特定の要件に従って作成され得る。たとえば、カスタマイズされたハードウェアもまた、使用されることがあり、および/または特定の要素が、ハードウェア、ソフトウェア(アプレットなどのポータブルソフトウェアを含む)、または両方で実装されることもある。さらに、ネットワーク入力/出力デバイスなどの他のコンピューティングデバイスへの接続が、使用され得る。

20

【0026】

いくつかの実施形態は、本開示による方法を実行するために、コンピュータシステム(コンピュータシステム100など)を使用し得る。たとえば、説明する方法の一部または全部のプロシーダは、作業メモリ135に含まれた(オペレーティングシステム140および/またはアプリケーションプログラム145などの他のコードに組み込まれ得る)1つまたは複数の命令の1つまたは複数のシーケンスを実行するプロセッサ110にตอบสนองして、コンピュータシステム100によって実行され得る。そのような命令は、ストレージデバイス125のうちの1つまたは複数など、別のコンピュータ可読媒体から作業メモリ135内に読み込まれ得る。単に例として、作業メモリ135内に含まれる命令のシーケンスの実行では、プロセッサ110に、本明細書で説明する方法の1つまたは複数のプロシーダを実行させ得る。

30

【0027】

本明細で使用する「機械可読媒体」および「コンピュータ可読媒体」という用語は、機械を特定の方式で動作させるデータを与えることに関与する任意の媒体を指す。コンピュータシステム100を使用して実装される一実施形態では、様々なコンピュータ可読媒体が、実行のためにプロセッサ110に命令/コードを与えることに関与し、かつ/またはそのような命令/コード(たとえば、信号)を記憶および/または搬送するために使用されることがある。多くの実装形態では、コンピュータ可読媒体は、物理的および/または有形のストレージ媒体である。そのような媒体は、限定はしないが、不揮発性媒体、揮発性媒体、および伝送媒体を含む多くの形態を取り得る。不揮発性媒体は、たとえば、ストレージデバイス125などの光ディスクおよび/または磁気ディスクを含む。揮発性媒体は、限定はしないが、作業メモリ135などの動的メモリを含む。伝送媒体は、限定はしないが、バス105、ならびに通信サブシステム130(および/または通信サブシステム130が他のデバイスとの通信を提供する媒体)の様々な構成要素を備える線を含めて、同軸ケーブル、銅線、および光ファイバーを含む。したがって、伝送媒体はまた、(限定はしないが、無線波通信および赤外線データ通信中に生成されるような無線波、音波、および/または光波を含めて)波の形態を取り得る。

40

【0028】

物理的および/または有形のコンピュータ可読媒体の一般的形態は、たとえば、フロッ

50

ピー（登録商標）ディスク、フレキシブルディスク、ハードディスク、磁気テープもしくは任意の他の磁気媒体、CD-ROM、任意の他の光媒体、パンチカード、紙テープ、任意の他の穴のパターンを有する物理的媒体、RAM、PROM、EPROM、FLASH-EPROM、任意の他のメモリチップもしくはカートリッジ、以下で説明する搬送波、またはコンピュータが命令および/またはコードを読み得る任意の他の媒体を含む。

【0029】

様々な形態のコンピュータ可読媒体は、実行のためにプロセッサ110に1つまたは複数の命令の1つまたは複数のシーケンスを搬送することに関与し得る。単に例として、命令は、最初に、リモートコンピュータの磁気ディスクおよび/または光ディスク上で搬送され得る。リモートコンピュータは、命令をその動的メモリ内にロードし、コンピュータシステム100によって受信されかつ/または実行される伝送媒体上の信号として、命令を送信し得る。電磁信号、音響信号、光信号などの形態であり得るこれらの信号は、すべて、本発明の様々な実施形態による、命令が符号化され得る搬送波の例である。

【0030】

通信サブシステム130(および/またはそれらの構成要素)は、一般に、信号を受信し、次いでバス105は作業メモリ135に信号(および/またはその信号によって搬送されるデータ、命令など)を搬送し、作業メモリ135から、プロセッサ110は命令を検索して実行し得る。作業メモリ135によって受信された命令は、場合によっては、プロセッサ110による実行の前または後のいずれかに、非一時的ストレージデバイス125に記憶され得る。

【0031】

上記で論じた方法、システム、およびデバイスは、例である。様々な実施形態において、様々なプロシージャまたは構成要素を、適宜、省略し、置換し、または加えることができる。たとえば、代替構成では、説明する本方法は、説明する順序とは異なる順序で実行されてもよく、ならびに/または、様々なステージが加えられ、省略され、および/もしくは組み合わせられてもよい。また、いくつかの実施形態に関して説明する特徴が、様々な他の実施形態と組み合わせられてもよい。実施形態の様々な態様および要素を同様に組み合わせることができる。また、技術は発展し、したがって、要素の多くは、本開示の範囲をそれらの特定の例に限定しない例である。

【0032】

実施形態の完全な理解を与えるために、説明に具体的な詳細が与えられる。しかしながら、実施形態は、これらの具体的な詳細なしに実践することができる。たとえば、実施形態を不明瞭にすることを避けるために、よく知られている回路、プロセス、アルゴリズム、構造、および技法は、不要な詳細なしに示してきた。この説明は、例示的な実施形態のみを提供し、本発明の範囲、適用可能性、または構成を限定しない。むしろ、これらの実施形態の上述の説明は、本発明の実施形態を実装することを可能にする説明を当業者に提供することになる。本発明の趣旨および範囲から逸脱することなく、要素の機能および構成に様々な変更を行うことができる。

【0033】

また、いくつかの実施形態を、フロー図またはブロック図として示すプロセスとして説明した。各々は動作を逐次プロセスとして説明し得るが、動作の多くは並行してまたは同時に実行され得る。加えて、動作の順序は並び替えられ得る。プロセスは、図に含まれていない追加のステップを有することができる。さらに、本方法の実施形態は、ハードウェア、ソフトウェア、ファームウェア、ミドルウェア、マイクロコード、ハードウェア記述言語、またはそれらの任意の組合せによって実装され得る。ソフトウェア、ファームウェア、ミドルウェア、またはマイクロコードで実装されるとき、関連するタスクを実行するプログラムコードまたはコードセグメントは、記憶媒体などのコンピュータ可読媒体に記憶され得る。プロセッサは関連するタスクを実行し得る。

【0034】

いくつかの実施形態について説明してきたが、様々な変更、代替構成、および均等物は、本開示の趣旨から逸脱することなく使用され得る。たとえば、上記の要素は、単により

10

20

30

40

50

大きいシステムの構成要素であり得、他のルールは、本発明の適用よりも優先するか、そうでなければ本発明の適用を変更することができる。また、上記の要素が考慮される前、間、または後に、いくつかのステップを行うことができる。したがって、上記の説明は、本開示の範囲を制限しない。

【0035】

本明細書で説明するように、画像からのグラフィカルオブジェクトは、OCRに対するターゲットであり、限定はしないが、1つまたは複数のシンボル、英数字、単語、標識、数、またはテキストを含み得る。たとえば、画像中のグラフィカルオブジェクトは、任意の言語における単語または文であってよい。単語は、いくつかの言語におけるシンボルとして表現されることがある。同様に、1つの言語または文化における文は、単に、別の言語または文化におけるシンボルを使用して表現されることがある。別の例では、グラフィカルオブジェクトは、STOP、YIELDなどの道路標識であってよい。異なるグラフィカルオブジェクトは、同じ意味を表現するために異なる地方、言語、または文化において使用されることがある。加えて、グラフィカル表示はまた、画像から識別されることがある。グラフィカル表示には、原文語(text word)、シンボルなど、オブジェクトのジェスチャ表現またはグラフィカル表現が含まれることがある。たとえば、リンゴの画像は、「リンゴ」という単語のグラフィカル表示である。全体的に、グラフィカル表示は、グラフィカルオブジェクトに対するOCR結果を改良するのを助ける。

【0036】

本明細書で説明するように、グラフィカルオブジェクトの群は、共通の特性を通じて互いに関連する複数のグラフィカルオブジェクトを含み得る。一実施形態では、グラフィカルオブジェクトの群は辞書を表現する。別の実施形態では、グラフィカルオブジェクトの群は辞書の改良版を表現する。さらに別の実施形態では、グラフィカルオブジェクトの群は1つまたは複数の特性を共有してともに群化されるオブジェクトのインデックスであってよい。本明細書で説明する方法では、いくつかの実施形態を、例示の目的で辞書を使用して説明するが、これらの実施形態における辞書の使用は、利用され得るグラフィカルオブジェクトのいずれの方法および群にも制約されない。

【0037】

一実装形態では、グラフィカルオブジェクトの群に対する群化は、あらかじめ規定されてよくかつあらかじめ分類されてもよい。たとえば、グラフィカルオブジェクトは、異なる地方および異なる言語にあらかじめ分類されてよい。別の実装形態では、グラフィカルオブジェクトの群に対する群化は、実時間またはほぼ実時間で実行され得る。群化は、連結リスト、配列、データベース、または任意の他の適切な手段を使用して編成および実装され得る。

【0038】

図2Aおよび図2Bは、コンテキスト依存OCRのためにモバイルデバイスを使用して取得される画像の例示的な表現を有する図を示している。コンテキスト依存OCRは、OCRのために選択されたグラフィカルオブジェクトの群を狭めることによって、より正確でより速いOCR結果を可能にする。一態様では、グラフィカルオブジェクトの群は、シンボルまたは単語の辞書であってよい。OCRが実行されると、いくつかの適切な単語が、検出されたコンテキストを使用して生成または選択され得る。OCR結果は、検出されたコンテキストに対する正しい単語を発見するのに利用可能な辞書と照合され得る。コンテキスト依存OCRは、OCRのために最も適切な辞書または辞書の改良版を選択するために、本明細書で説明する本発明の多数の異なる実施形態を、単独でまたは互いに連携して使用し得る。

【0039】

図2Aは、韓国レストランからのメニューである。たとえば、モバイルデバイスのユーザのロケーションが、モバイルデバイスに結合されたGPS要素またはマイクロフォンを介して取得され得、言語およびおそらくは特定のレストランをも識別するのを助け得る。文書の構造、照明、メニューのテキストのカメラレンズからの焦点距離、および時刻(time of the day)が、すべて、ユーザが夕食メニューに焦点を当てているものと判断することを助け

得る。このことが、モバイルデバイスができるだけ狭く辞書を選択すること、またはさらに辞書を改良することを可能にする。コンテキスト情報を照合することで、OCRの正確さと速さが向上し、同時に、OCRに必要な処理能力を潜在的に低減することができる。さらに、図2Aでは、メニュー項目がまた、より小さなフォントの英語で記載されている。本発明の様子はまた、画像の英語の部分に対してOCRを実行し、英語のOCR結果を韓国語に変換し、韓国語のOCRに最も適合するものを選択するために、韓国語OCRおよび英語OCRからの可能性のある候補を比較することができる。

【0040】

同様に、図2Bは、日本の交通システムの画像である。ロケーション、言語、照明、文書の構造、環境、雑音状態、および多くの他の類似の入力が、ユーザのための画像内のテキストおよび標識を認識するために、コンテキストを決定し、OCRプロセスを高速化するのを助け得る。

【0041】

図3は、OCRのためのグラフィカルオブジェクトの群を選択するための、コンテキスト情報を使用する非限定的な例示的デシジョンツリーを示すフロー図である。グラフィカルオブジェクトの群は、辞書または辞書の改良版であってよい。図3に関して説明する実施形態では、辞書は、最も広い基準から最も狭いかまたは最も収束された基準に狭められる。他の実施形態では、唯一のソースまたは感覚入力が、コンテキストおよび関連する辞書の選択に到達するのに使用され得る。それでも、他の実施形態では、辞書の選択を狭めるのに取られる判断のシーケンスは、順序が変わることがある。方法300は、(回路、専用論理回路などの)ハードウェア、(汎用コンピューティングシステムまたは専用機械上で実行されるような)ソフトウェア、(ソフトウェアに埋め込まれる)ファームウェア、またはそれらの任意の組合せを備える処理用論理回路によって実行される。一実施形態では、方法300は、図1のデバイス100によって実行される。

【0042】

図3を参照すると、ブロック302で、モバイルデバイスは、マイクロフォンからの音響入力などの様々なセンサ入力、地理的ロケーション、または画像から数個のグラフィカルオブジェクトを認識することに基づいて言語を決定し得る。言語を決定することで、OCRは、OCR結果を生成するための正しい単語の探索空間を著しく狭めることが可能になる。

【0043】

ブロック304で、モバイルデバイスは、ロケーション識別子からの入力に基づいてロケーションを決定し得る。ロケーションは、モバイルデバイス上のGPS要素からの入力、携帯電話の中継塔に関する信号強度、ユーザによるマイクロフォン入力または手動選択を使用して導出され得る。ロケーション情報は、異なる粒度(granularity)で与えられ、使用され得る。たとえば、ロケーションは、人が都市の中にいるのかまたは外れにいるのかを判断するのを助け得る。ロケーション情報はまた、レストラン名または街路名ほどに具体的であり得る。一例として、ロケーションは街路であるとモバイルデバイスが判断する場合、その特定の都市のすべての街路名を含む辞書または改良された辞書が、より速いOCRのために使用され得る。

【0044】

ブロック306で、グラフィカルオブジェクトの群に対するドメインが選択される。ドメインは、コンテキストの下位範疇化であってよい。たとえば、1つのロケーションが、複数のドメインを持ち得る。ブロック304で、ロケーションが街路である場合、ドメインは道路標識であり得る。同様に、ロケーションがレストランである場合、ドメインは、クレジットカードまたは夕食メニューであり得る。後でより詳細に説明する文書構造解析およびコンテンツ意味解析のようないくつかの技法もまた、ドメインを決定するために使用され得る。

【0045】

ブロック308で、探索は、ドメイン中の興味を中心を発見することによってさらに狭められ得る。たとえば、クレジットカード上の興味を中心は、人名またはクレジットカード

10

20

30

40

50

番号であり得る。ユーザは、興味を中心を選択するために音声コマンドを与えるか、または興味を中心を選択するために画像中のテキストに触ることができる。

【0046】

本発明の様々な実施形態では、コンテキストを決定し、OCRに対して最も適切なグラフィカルオブジェクトの群を選択するために、異なるコンテキストソースが、単独でまたは互いに連携して使用され得る。コンテキストソースのわずかな例は、ロケーション、ユーザ、時計、カメラ、およびユーザパターンを含み得る。各コンテキストソースは、さらに、センサ入力デバイスに関連付けられ得る。たとえば、ロケーションは、モバイルデバイス内にあるGPS要素、携帯電話の中継塔に関する信号強度、または音響入力を介する信号強度を使用して決定され得る。ユーザはまた、モバイルデバイス上に表示されるメニューからコンテキストまたは辞書を手動で選択することによって、またはデバイスへの音声コマンドを介して、コンテキストソースを与え得る。さらに、ユーザの挙動パターンが、ユーザから生起するコンテキストを構成するために使用され得る。同様に、時計および日付(lock and date)が、夜間もしくは日中の環境、または季節に関する重要な情報を与え得る。カメラは、距離、文字サイズ、光の状態などに対するコンテキストを与え得るので、カメラもまたコンテキストの重要なソースである。カメラに加えて、モバイルデバイスもまた、光の状態をよりよく測定するために光センサを有することがある。

10

【0047】

図3に示す具体的なステップは、本発明の一実施形態による、動作のモード間を切り替える特定の方法を提供することを諒解されたい。ステップの他の順序も、代替的な実施形態に従って実行され得る。たとえば、本発明の代替的な実施形態は、異なる順序で、上で概説されたステップを実行することができる。例示のために、ユーザは、第3の動作モードから第1の動作モードに、第4のモードから第2のモードに、またはそれらの間の任意の組合せに変えることを選択し得る。その上、図3に示す個々のステップは、個々のステップに対して様々な順序で適宜実行され得る、複数のサブステップを含み得る。さらに、追加のステップが、具体的な用途に応じて、追加または削除され得る。当業者は、方法300の多くの変形、修正、および代替を認識し、諒解するだろう。

20

【0048】

図4は、コンテキスト依存OCRのための方法を実行するための例示的な実施形態を説明するフロー図である。方法400は、(回路、専用論理回路などの)ハードウェア、(汎用コンピューティングシステムまたは専用機械上で実行されるような)ソフトウェア、(ソフトウェアに埋め込まれる)ファームウェア、またはそれらの任意の組合せを備える処理用論理回路によって実行される。一実施形態では、方法400は、図1のデバイス100によって実行される。

30

【0049】

図4を参照すると、ブロック402で画像が取得される。一実施形態では、画像は、モバイルデバイスに結合されたカメラを使用して取得される。画像が取得されると、ブロック404で、グラフィカルオブジェクトを含む画像の一部分が識別される。画像内のグラフィカルオブジェクトには、限定はしないが、1つまたは複数のシンボル、英数字、単語、標識、または数が含まれ得る。一実施形態では、長方形のテキストボックスが、グラフィカルオブジェクトの周りに置かれる。ブロック406で、グラフィカルオブジェクトは部分的にまたは完全に認識される。一態様では、OCRは、グラフィカルオブジェクトを認識するために利用される。ブロック408で、グラフィカルオブジェクトの群(1つまたは複数の辞書など)が識別される。一態様では、グラフィカルオブジェクトの群は、辞書または辞書の改良版であってよく、画像のコンテキストに基づいて生成される。

40

【0050】

一実施形態では、画像のコンテキストに基づく辞書が、画像の処理と並行して選択され、それにより、OCRからの結果がグラフィカルオブジェクトの群をクエリする(query)準備が整う前に、コンテキストおよびグラフィカルオブジェクトの群は、準備が整うかまたはほぼ整う。ブロック410で、コンテキストセレクトは、異なるセンサからの入力および可

50

能なユーザ入力に基づいてコンテキストを決定し得る。コンテキストは、他のファクタの中でも、ロケーション、カメラ入力、時間入力、および履歴に基づくことができる。ブロック412で、正しいコンテキストが選択されると、適正な辞書または辞書の改良版が、選択されたコンテキストに関連付けられ得、ブロック408に送られる。ブロック408で、画像のコンテキストに基づいて選択された辞書または辞書の改良版を使用して、探索が実行される。ブロック414で、OCR結果が出力される。

【0051】

別の実施形態では、テキストのカメラレンズからの距離および物理的な文字サイズが、テキストのコンテキストを決定するのに使用され得る。標識のモバイルデバイスからの距離を知ること、グラフィカルオブジェクトの群の選択範囲を狭めるために使用され得る視野が可能になる。たとえば、街路の標識は、遠く離れている。レストラン内のメニューは、近距離にある。本のテキストは、おそらく、さらに近い。実施形態のそのような実装形態に対して、テキストのカメラからの正確な距離は必要でないことがある。代わりに、距離は、近い、部屋サイズ、遠いなど、より広いカテゴリーに分割されてよい。別の例示的なカテゴリー化では、距離は、カメラがオートフォーカスされときの焦点距離から、近接、ノーマル、無限大など、別々のカテゴリーに近似されてよい。さらに、いくつかの単語の群化もまた、主題のヒントを生成するのを助け得る。たとえば、道路標識は、約2~4の単語を有することがある。対照的に、テキストブックは、より大きい20~30の単語の群化を有することがある。投影されるテキストのサイズは、OCRが考察するのに必要なテキストのブロックサイズに限定する(constrain)のを助け得る。

【0052】

図4に示す具体的なステップは、本発明の一実施形態による、動作のモードの間を切り替える特定の方法を提供することを諒解されたい。ステップの他の順序も、代替的な実施形態に従って実行され得る。たとえば、本発明の代替的な実施形態は、異なる順序で、上で概説されたステップを実行することができる。例示のために、ユーザは、第3の動作モードから第1の動作モードに、第4のモードから第2のモードに、またはそれらの間の任意の組合せに変えることを選択し得る。その上、図4に示される個々のステップは、個々のステップに対して様々な順序で適宜実行され得る、複数のサブステップを含み得る。さらに、追加のステップが、具体的な用途に応じて、追加または削除され得る。当業者は、方法400の多くの変形、修正、および代替を認識し、諒解するだろう。

【0053】

図5は、コンテキスト依存OCRのための方法を実行するためのさらに別の例示的な実施形態を説明するフロー図である。方法500は、(回路、専用論理回路などの)ハードウェア、(汎用コンピューティングシステムまたは専用機械上で実行されるような)ソフトウェア、(ソフトウェアに埋め込まれる)ファームウェア、またはそれらの任意の組合せを備える処理用論理回路によって実行される。一実施形態では、方法500は、図1のデバイス100によって実行される。

【0054】

図5を参照すると、ブロック502で画像が取得される。一実施形態では、画像は、モバイルデバイスに結合されたカメラを使用して取得される。画像が取得されると、ブロック504で、グラフィカルオブジェクトを含む画像の一部が識別される。画像内のグラフィカルオブジェクトには、限定はしないが、1つまたは複数のシンボル、英数字、単語、標識、または数が含まれ得る。一実装形態では、長方形のテキストボックスが、グラフィカルオブジェクトの周りに置かれる。ブロック506で、グラフィカルオブジェクトは部分的にまたは完全に認識される。一態様では、OCRは、グラフィカルオブジェクトを認識するために利用される。ブロック508で、1つまたは複数のOCR候補が、グラフィカルオブジェクトに対してOCRを実行することによって生成される。ブロック510で、OCR候補が、グラフィカルオブジェクトの群を使用して複数のOCR候補から選択される。たとえば、OCRを実行することで、10個のOCR候補が生成され得る。並行して、デバイスはまた、コンテキストを構築し、コンテキストに基づいてグラフィカルオブジェクトの群(1つまたは複数の辞書

など)を選択し得る。一実施形態では、最良の候補、またはグラフィカルオブジェクトの群からのグラフィカルオブジェクトの1つに適合する候補が、OCR結果と見なされ得る。

【0055】

一実施形態では、画像のコンテキストに基づく辞書が、画像の処理と並行して選択され、それにより、OCRからの結果がグラフィカルオブジェクトの群をクエリする準備が整う前に、コンテキストおよびグラフィカルオブジェクトの群は、準備が整うかまたはほぼ整う。ブロック512で、コンテキストセレクタは、異なるセンサからの入力および可能なユーザ入力に基づいてコンテキストを決定し得る。コンテキストは、他のファクタの中でも、ロケーション、カメラ入力、時間入力、および履歴に基づくことができる。ブロック514で、正しいコンテキストが選択されると、適正な辞書または辞書の改良版が、選択されたコンテキストに関連付けられ得、ブロック510に送られる。ブロック510で、上記で説明したように、最良の候補が、グラフィカルオブジェクトの群を使用して選択される。ブロック516で、OCR結果が出力される。

【0056】

図5に示す具体的なステップは、本発明の一実施形態による、動作のモードの間を切り替える特定のを提供することを諒解されたい。ステップの他の順序も、代替的な実施形態に従って実行され得る。たとえば、本発明の代替的な実施形態は、異なる順序で、上で概説されたステップを実行することができる。例示のために、ユーザは、第3の動作モードから第1の動作モードに、第4のモードから第2のモードに、またはそれらの間の任意の組合せに変えることを選択し得る。その上、図5に示す個々のステップは、個々のステップに対して様々な順序で適宜実行され得る、複数のサブステップを含み得る。さらに、追加のステップが、具体的な用途に応じて、追加または削除され得る。当業者は、方法500の多くの変形、修正、および代替を認識し、諒解するだろう。

【0057】

図6は、コンテキスト依存OCRのための方法を実行するための例示的な実施形態を説明するフロー図である。方法600は、(回路、専用論理回路などの)ハードウェア、(汎用コンピューティングシステムまたは専用機械上で実行されるような)ソフトウェア、(ソフトウェアに埋め込まれる)ファームウェア、またはそれらの任意の組合せを備える処理用論理回路によって実行される。一実施形態では、方法600は、図1のデバイス100によって実行される。

【0058】

図6を参照すると、ブロック602で、グラフィカルオブジェクトに対するコンテキストが、本明細書で説明する機構を使用して推測され得る。OCRコンテキストが推測されると、ブロック604で、OCR結果が、推測されたコンテキストを使用して改良され得る。その上、ブロック606で、推測されたコンテキストが、OCR結果を使用することによってさらに改良され得る。

【0059】

図6に示す具体的なステップは、本発明の一実施形態による、動作のモードの間を切り替える特定のを提供することを諒解されたい。ステップの他の順序も、代替的な実施形態に従って実行され得る。たとえば、本発明の代替的な実施形態は、異なる順序で、上で概説されたステップを実行することができる。例示のために、ユーザは、第3の動作モードから第1の動作モードに、第4のモードから第2のモードに、またはそれらの間の任意の組合せに変えることを選択し得る。その上、図6に示される個々のステップは、個々のステップに対して様々な順序で適宜実行され得る、複数のサブステップを含み得る。さらに、追加のステップが、具体的な用途に応じて、追加または削除され得る。当業者は、方法600の多くの変形、修正、および代替を認識し、諒解するだろう。

【0060】

図7は、画像が取得されるとききの光の状態に基づいてOCRに対するコンテキストを選択するための、本発明の一実施形態を示すフロー図である。方法700は、(回路、専用論理回路などの)ハードウェア、(汎用コンピューティングシステムまたは専用機械上で実行される

ような)ソフトウェア、(ソフトウェアに埋め込まれる)ファームウェア、またはそれらの任意の組合せを備える処理用論理回路によって実行される。一実施形態では、方法700は、図1のデバイス100によって実行される。

【0061】

図7を参照すると、一実施形態では、OCRのための辞書に対するコンテキストを選択することは、画像が取得されたときの光の状態に基づく。辞書は、図7を説明する間に、説明の目的で使用されるグラフィカルオブジェクトの群の一例である。ブロック702で、画像が取得される。一実施形態では、画像は、モバイルデバイスに結合されたカメラを使用して取得される。画像が取得されると、ブロック704で、グラフィカルオブジェクト(単語など)を含む画像の一部が識別される。画像内のグラフィカルオブジェクトには、限定はしないが、1つまたは複数のシンボル、英数字、単語、標識、または数が含まれ得る。一実装形態では、長方形のテキストボックスが、グラフィカルオブジェクトの周りに置かれる。ブロック706で、グラフィカルオブジェクトは部分的にまたは完全に認識される。一態様では、OCRは、グラフィカルオブジェクトを認識するために利用される。ブロック708で、グラフィカルオブジェクトの群が識別される。一態様では、グラフィカルオブジェクトの群は、辞書または辞書の改良版であってよく、画像のコンテキストに基づいて生成される。

10

【0062】

ブロック710で、モバイルデバイスは、光センサから入力を受信する。一実装形態では、光電池が光センサとして使用され得る。光センサからの入力に基づいて、モバイルデバイスは、写真が屋外で撮られたかまたは屋内で撮られたかを判断し得る。たとえば、複数の光線に関連する明度が、光センサの出力を使用して決定され得る。一実施形態では、しきい値より低い明度が、屋内で画像を取得したという推測をもたらす、しきい値を超える明度が、屋外で画像を取得したという推測をもたらす。光線の明度は、輝度(luminance)、色温度、または両方を指すことがある。さらに、いくつかの実施形態では、決定の品質を高めるために、モバイルデバイスはまた、単に非常に明るい屋内環境とは異なる日光の特性を検出し得る。たとえば、部屋は、蛍光灯を使用して明るくなっている可能性がある。ブロック712で、モバイルデバイスは、画像が屋内でキャプチャされているか外でキャプチャされているかを判断する。画像がどこでキャプチャされるかを知ることによって、辞書の選択は著しく狭められる。たとえば、屋内で取得された画像に対するOCRは、概して、レストランメニューまたは本のように、より小さい印刷物を有する辞書または辞書の改良版を目標設定する。一方、屋外での画像のキャプチャに対するOCRは、概して、より大きな印刷物を有し、街路名および店舗標識のように外で発見される辞書を目録設定する。ブロック714で、辞書が選択されて、ブロック708に提供される。ブロック708で、適正な単語が、選択された辞書から選択される。ブロック716で、OCR結果が出力される。

20

30

【0063】

図7に示す具体的なステップは、本発明の一実施形態による、動作のモードの間を切り替える特定の手法を提供することを諒解されたい。ステップの他の順序も、代替的な実施形態に従って実行され得る。たとえば、本発明の代替的な実施形態は、異なる順序で、上で概説されたステップを実行することができる。例示のために、ユーザは、第3の動作モードから第1の動作モードに、第4のモードから第2のモードに、またはそれらの間の任意の組合せに変えることを選択し得る。その上、図7に示される個々のステップは、個々のステップに対して様々な順序で適宜実行され得る、複数のサブステップを含み得る。さらに、追加のステップが、具体的な用途に応じて、追加または削除され得る。当業者は、方法700の多くの変形、修正、および代替を認識し、諒解するだろう。

40

【0064】

図8は、OCRに対する辞書を選択するためのコンテキストがマイクロフォンからの入力に基づく、一実施形態を例示するフロー図である。方法800は、(回路、専用論理回路などの)ハードウェア、(汎用コンピューティングシステムまたは専用機械上で実行されるような)ソフトウェア、(ソフトウェアに埋め込まれる)ファームウェア、またはそれらの任意の

50

組合せを備える処理用論理回路によって実行される。一実施形態では、方法800は、図1のデバイス100によって実行される。

【0065】

図8を参照すると、ブロック802で画像が取得される。一実施形態では、画像は、モバイルデバイスに結合されたカメラを使用して取得される。画像が取得されると、ブロック804で、グラフィカルオブジェクトを含む画像の一部が識別される。画像内のグラフィカルオブジェクトには、限定はしないが、1つまたは複数のシンボル、英数字、単語、標識、または数が含まれ得る。一実施形態では、長方形のテキストボックスが、グラフィカルオブジェクトの周りに置かれる。ブロック806で、グラフィカルオブジェクトは部分的にまたは完全に認識される。一態様では、OCRは、グラフィカルオブジェクトを認識するために利用される。

10

【0066】

ブロック810で、マイクロフォンは、マイクロフォンを介して音響入力を受信して、音響を記録する。ブロック812で、モバイルデバイスは、音響記録について音声認識を実行する。いくつかの実施形態では、ブロック802で、ユーザは、キャプチャされた画像のコンテキストを言葉で選択することができる。たとえば、ユーザは言語を指定し得、その言語から辞書が選択される。他の実施形態では、ユーザはまた、画像から、ユーザが興味を持つ特定の分野を言葉で選択し得る。たとえば、ユーザは、画像中で視認できる正しい名前または数字を選択し得るか、またはユーザは、OCR処理のための項目のクレジットカード番号もしくはコストを選択するために、より具体的なバーバルコマンドを与え得る。ブロック814で、一実施形態では、認識された音響パターンから、キーワードが検出され得る。別の実施形態では、コンテキストが、認識された音響パターンに基づいて推測される。ブロック808で、ブロック814から検出されたキーワードおよび推測されたコンテキストが、OCR結果の改良にさらに使用され得る。一実施形態では、ブロック808で、ディスプレイは、ユーザによって選択された単語をスクリーン上にオーバーレイする。ブロック816で、OCR結果が出力され得る。

20

【0067】

図8に示す具体的なステップは、本発明の一実施形態による、動作のモードの間を切り替える特定の方法を提供することを諒解されたい。ステップの他の順序も、代替的な実施形態に従って実行され得る。たとえば、本発明の代替的な実施形態は、異なる順序で、上で概説されたステップを実行することができる。例示のために、ユーザは、第3の動作モードから第1の動作モードに、第4のモードから第2のモードに、またはそれらの間の任意の組合せに変えることを選択し得る。その上、図8に示される個々のステップは、個々のステップに対して様々な順序で適宜実行され得る、複数のサブステップを含み得る。さらに、追加のステップが、具体的な用途に応じて、追加または削除され得る。当業者は、方法800の多くの変形、修正、および代替を認識し、諒解するだろう。

30

【0068】

図9は、コンテキスト依存OCRのために手動入力を使用するための実施形態を説明するフロー図である。方法900は、(回路、専用論理回路などの)ハードウェア、(汎用コンピューティングシステムまたは専用機械上で実行されるような)ソフトウェア、(ソフトウェアに埋め込まれる)ファームウェア、またはそれらの任意の組合せを備える処理用論理回路によって実行される。一実施形態では、方法900は、図1のデバイス100によって実行される。

40

【0069】

図9を参照すると、ブロック902で画像が取得される。一実施形態では、画像は、モバイルデバイスに結合されたカメラを使用して取得される。画像が取得されると、ブロック904で、グラフィカルオブジェクトを含む画像の一部が識別される。画像内のグラフィカルオブジェクトには、限定はしないが、1つまたは複数のシンボル、英数字、単語、標識、または数が含まれ得る。一実施形態では、長方形のテキストボックスが、グラフィカルオブジェクトの周りに置かれる。ブロック906で、グラフィカルオブジェクトは部分的に

50

または完全に認識される。一態様では、OCRは、グラフィカルオブジェクトを認識するために利用される。

【0070】

図9の実施形態は、本明細書で説明する他の実施形態と併せて使用され得る。一実施形態では、異なる辞書が、異なるコンテキストソースを使用して選択され、最終選択のためにユーザに提示され得る。さらに別の実施形態では、ユーザは、コンテキスト内で使用されるべき辞書を、モバイルデバイスによって提供されるユーザインターフェースを介して、ユーザによって手動で選択し得る。ブロック910で、様々なソースからの候補辞書が選択され、最終選択のためにユーザに提示される。ブロック912で、ユーザは辞書を選択する。ブロック914で、ユーザ入力からのフィードバックが、辞書の選択において評価される。いくつかの実施形態では、ユーザは、辞書そのものではなく、辞書に対するコンテキストを選択し得る。他の実施形態では、ユーザは、辞書の群化またはカテゴリーを選択し得る。ブロック908で、グラフィカルオブジェクトは、ユーザ選択に基づいて認識され、OCRの最終結果が出力される(ブロック916)。

【0071】

図9に示す具体的なステップは、本発明の一実施形態による、動作のモードの間を切り替える特定の方法を提供することを諒解されたい。ステップの他の順序も、代替的な実施形態に従って実行され得る。たとえば、本発明の代替的な実施形態は、異なる順序で、上で概説されたステップを実行することができる。例示のために、ユーザは、第3の動作モードから第1の動作モードに、第4のモードから第2のモードに、またはそれらの間の任意の組合せに変えることを選択し得る。その上、図9に示される個々のステップは、個々のステップに対して様々な順序で適宜実行され得る、複数のサブステップを含み得る。さらに、追加のステップが、具体的な用途に応じて、追加または削除され得る。当業者は、方法900の多くの変形、修正、および代替を認識し、諒解するだろう。

【0072】

図10は、画像が取得されたロケーションを使用してコンテキスト依存OCRを実施するための方法の別の実施形態を説明するフロー図である。たとえば、ユーザが特定のレストラン内にいる場合、選択される辞書は、そのレストランに固有のものであり得る。方法1000は、(回路、専用論理回路などの)ハードウェア、(汎用コンピューティングシステムまたは専用機械上で実行されるような)ソフトウェア、(ソフトウェアに埋め込まれる)ファームウェア、またはそれらの任意の組合せを備える処理用論理回路によって実行される。一実施形態では、方法1000は、図1のデバイス100によって実行される。

【0073】

図10を参照すると、ブロック1002で画像が取得される。一実施形態では、画像は、モバイルデバイスに結合されたカメラを使用して取得される。画像が取得されると、ブロック1004で、グラフィカルオブジェクトを含む画像の一部分が識別される。画像内のグラフィカルオブジェクトには、限定はしないが、1つまたは複数のシンボル、英数字、単語、標識、または数が含まれ得る。一実装形態では、長方形のテキストボックスが、グラフィカルオブジェクトの周りに置かれる。ブロック1006で、グラフィカルオブジェクトは部分的にまたは完全に認識される。一態様では、OCRは、グラフィカルオブジェクトを認識するために利用される。

【0074】

ブロック1010で、データに対するロケーションが取得される。一実施形態では、画像をキャプチャする人のロケーションが、モバイルデバイス内のGPS要素を使用して確認され得る。代替として、ユーザロケーションは、セルタワーに関する信号、またはWiFiのようなワイヤレスアクセスポイント接続もしくはは任意の他の適切な手段を介する信号の強度を使用して決定され得る。ブロック1012で、モバイルデバイスは、ロケーションに対する詳細(specifics)を決定するためにロケーションデータを処理する。ロケーションの粒度もまた、選択される辞書の粒度を決定し得る。ブロック1014で、ロケーションに基づく辞書が選択される。ブロック1012で選択されたロケーションが、国または国の特定の地方であ

る場合、ブロック1014で、その地方に対して適正な言語の辞書が選択され得る。たとえば、韓国内のユーザに対して、韓国語が選択され得る。

【0075】

ブロック1012で確認されたロケーションの粒度が増すにつれて、さらに改良された辞書がブロック1014で選択され得る。別の実施形態では、モバイルデバイスは、特定のロケーションに関してより多くの情報を獲得するために、ネットワークにアクセスし得る。たとえば、一実施形態では、モバイルデバイスのGPS座標が取得されると、モバイルデバイスは、ロケーションに関連する詳細を確認するために、ローカルにまたはリモートに位置するデータベースにクエリし得る。一例では、ブロック1012で、特定のロケーションに対するGPS座標が、ギリシャレストランに関連付けられる。ブロック1014で、OCRに対して選択された辞書は、そのレストランに対して固有であり得る。そのレストランに固有の辞書が利用できない場合、ギリシャ料理に関連するより一般的な辞書が選択され得る。ブロック1008で、ブロック1014からのロケーションコンテキストに基づくグラフィカルオブジェクトの群から、識別されたグラフィカルオブジェクトが選択される。ブロック1018で、選択されたOCRグラフィカルオブジェクトが出力され得る。

10

【0076】

ブロック1012で特定のロケーションを決定するために、マイクロフォンセンサもまた、ブロック1016でロケーション情報を提供し得る。一例では、レストランは、そのロケーションに関連付けられた署名トーン(signature tone)を有し得る。モバイルデバイスはそのトーンを記録して解析し、ローカルにモバイルデバイス上でまたはリモートサーバ上で、トーンに関連付けられたロケーションを決定する。関連特許出願「Mobile device location estimation using environmental information」、米国特許出願第12/898,647号が、レストランを決定するためのこの実施形態を、より詳細に記載している。

20

【0077】

図10に示す具体的なステップは、本発明の一実施形態による、動作のモードの間を切り替える特定の方法を提供することを諒解されたい。ステップの他の順序も、代替的な実施形態に従って実行され得る。たとえば、本発明の代替的な実施形態は、異なる順序で、上で概説されたステップを実行することができる。例示のために、ユーザは、第3の動作モードから第1の動作モードに、第4のモードから第2のモードに、またはそれらの間の任意の組合せに変えることを選択し得る。その上、図10に示される個々のステップは、個々のステップに対して様々な順序で適宜実行され得る、複数のサブステップを含み得る。さらに、追加のステップが、具体的な用途に応じて、追加または削除され得る。当業者は、方法1000の多くの変形、修正、および代替を認識し、諒解するだろう。

30

【0078】

図11は、コンテキスト依存OCRに対するグラフィカルオブジェクトの群を選択するためにマイクロフォンを使用して環境を認識するための別の例示的な実施形態を説明するフロー図である。辞書は、図11を説明する間に、説明の目的で使用されるグラフィカルオブジェクトの群の一例である。方法1100は、(回路、専用論理回路などの)ハードウェア、(汎用コンピューティングシステムまたは専用機械上で実行されるような)ソフトウェア、(ソフトウェアに埋め込まれる)ファームウェア、またはそれらの任意の組合せを備える処理用論理回路によって実行される。一実施形態では、方法1100は、図1のデバイス100によって実行される。

40

【0079】

図11を参照すると、ブロック1102で画像が取得される。一実施形態では、画像は、モバイルデバイスに結合されたカメラを使用して取得される。画像が取得されると、ブロック1104で、グラフィカルオブジェクトを含む画像の一部分が識別される。画像内のグラフィカルオブジェクトには、限定はしないが、1つまたは複数のシンボル、英数字、単語、標識、または数が含まれ得る。一実施形態では、長方形のテキストボックスが、グラフィカルオブジェクトの周りに置かれる。ブロック1106で、グラフィカルオブジェクトは部分的にまたは完全に認識される。一態様では、OCRは、グラフィカルオブジェクトを認識する

50

ために利用される。

【0080】

ブロック1110で、マイクロフォンは、周囲からの入力を受信する。ブロック1112で、ブロック1110から受信された音響入力解析される。ユーザの環境が、音響入力に基づいて認識される。たとえば、ブロック1112で、モバイルデバイスは、環境がレストランであるか、街路であるか、または事務所であるかを判断し得る。ブロック1114で、モバイルデバイスは、マイクロフォンからの音響入力を使用して検出された環境に基づいて適正な辞書を選択し得、その辞書をブロック1108に提供する。ブロック1108で、適正なグラフィカルオブジェクトが、推測されたコンテキストに基づいて辞書から選択され、ブロック1116で出力される。関連仮特許出願「Recognizing environmental sound on local device and server」、米国仮出願第61/449,475号が、環境ファクタを検出するためのこの実施形態を、より詳細に記載している。

10

【0081】

図11に示す具体的なステップは、本発明の一実施形態による、動作のモードの間を切り替える特定の手法を提供することを諒解されたい。ステップの他の順序も、代替的な実施形態に従って実行され得る。たとえば、本発明の代替的な実施形態は、異なる順序で、上で概説されたステップを実行することができる。例示のために、ユーザは、第3の動作モードから第1の動作モードに、第4のモードから第2のモードに、またはそれらの間の任意の組合せに変えることを選択し得る。その上、図11に示される個々のステップは、個々のステップに対して様々な順序で適宜実行され得る、複数のサブステップを含み得る。さらに、追加のステップが、具体的な用途に応じて、追加または削除され得る。当業者は、方法1100の多くの変形、修正、および代替を認識し、諒解するだろう。

20

【0082】

図12は、コンテンツ解析を使用するコンテキスト識別のための一実施形態を説明するフロー図である。この実施形態では、最終のOCR結果は、OCR結果を出力する前に、文書構造解析、コンテンツ意味解析、および従来のOCR結果を考慮する。方法1200は、(回路、専用論理回路などの)ハードウェア、(汎用コンピューティングシステムまたは専用機械上で実行されるような)ソフトウェア、(ソフトウェアに埋め込まれる)ファームウェア、またはそれらの任意の組合せを備える処理用論理回路によって実行される。一実施形態では、方法1200は、図1のデバイス100によって実行される。

30

【0083】

図12を参照すると、ブロック1202で画像が取得される。一実施形態では、画像は、モバイルデバイスに結合されたカメラを使用して取得される。画像が取得されると、ブロック1204で、テキストなどのグラフィカルオブジェクトを含む画像の一部が識別される。画像内のグラフィカルオブジェクトには、限定はしないが、1つまたは複数のシンボル、英数字、単語、テキスト、標識、または数が含まれ得る。テキストは、図12を説明する間に、説明の目的で使用されたグラフィカルオブジェクトの一例であるが、任意のグラフィカルオブジェクトが、類似の方法と併せて使用され得る。加えて、ブロック1204で、グラフィカル表示もまた識別されることがある。グラフィカル表示には、原文語、シンボルなど、オブジェクトのジェスチャ表現またはグラフィカル表現が含まれることがある。たとえば、リンゴの画像は、「リンゴ」という単語のグラフィカル表示である。全体的に、グラフィカル表示は、グラフィカルオブジェクトに対するOCR結果を改良するのを助ける。一実施形態では、長方形のテキストボックスが、テキストの周りに置かれる。ブロック1204で、テキスト領域が検出されると、ブロック1206で、テキスト領域の文書構造解析およびコンテンツ意味解析が並行して開始し得る。

40

【0084】

ブロック1208で、文書構造解析が開始する。文書構造解析は、グラフィカルオブジェクトのコンテキストを決定するために、テキストおよび任意の識別されたグラフィカル表示など、グラフィカルオブジェクトの構造およびレイアウトに焦点を当てる。たとえば、レストランの夕食メニューは、本のカバーとかなり異なるテキストレイアウトを有する。文書

50

構造解析は、フォント高さのような、テキスト領域の単項素性に、また、テキスト領域の2項素性にも焦点を当て得る。2項素性は、共直進性ならびに水平および垂直オーバーラップの次数のような特殊な配列の解析を含み得る。

【0085】

テキストを有する部分が検出されると、文書のコンテンツ意味解析が、同様に開始し得る。ブロック1210で、暫定OCRが実行される。OCRから出力された認識されたテキスト(ブロック1212)が、コンテンツ意味解析のために使用される。たとえば、数個の単語がOCRを使用して部分的に認識されると、それらの単語を使用してコンテキストを推測する様々な方式が実装され得る(ブロック1214)。一実施形態では、最も頻繁に検出されたカテゴリーに属する単語がコンテキストを構築するために使用される、多数決投票方式が使用される。別の実施形態では、ドメイン固有のコンテキストが、コンテキストに基づいて推測され得る。いくつかの実施形態では、コンテンツ意味解析は、コンテキストを推測するために、文書構造解析と併せて動作する。たとえば、レストランメニューでは、食品名と値段とが一緒に掲載される。同様に、本の題名では、人名が、比較的小さいフォントで掲載される。

10

【0086】

ブロック1216で、文書構造解析およびコンテンツ意味解析からのコンテキストが、辞書を改良するために使用される。辞書は、図12において、説明の目的で使用されるグラフィカルオブジェクトの群の一例であるが、グラフィカルオブジェクトの任意の群を使用する他の方法も利用され得る。ブロック1218で、別のOCRが実行される。OCRからの結果は、ブロック1216で生成される、文書構造解析およびコンテンツ意味解析からの改良された辞書を使用する。ブロック1206でテキスト領域が検出されると、ブロック1218で、OCRは、文書構造解析およびコンテンツ意味解析と並行して開始し得る。ブロック1220で、コンテキスト依存OCRに対する出力テキストが生成される。

20

【0087】

図12に示す具体的なステップは、本発明の一実施形態による、動作のモードの間を切り替える特定の方法を提供することを諒解されたい。ステップの他の順序も、代替的な実施形態に従って実行され得る。たとえば、本発明の代替的な実施形態は、異なる順序で、上で概説されたステップを実行することができる。例示のために、ユーザは、第3の動作モードから第1の動作モードに、第4のモードから第2のモードに、またはそれらの間の任意の組合せに変えることを選択し得る。その上、図12に示される個々のステップは、個々のステップに対して様々な順序で適宜実行され得る、複数のサブステップを含み得る。さらに、追加のステップが、具体的な用途に応じて、追加または削除され得る。当業者は、方法1200の多くの変形、修正、および代替を認識し、諒解するだろう。

30

【0088】

図13は、画像の構造レイアウトを解析することによるコンテキスト識別のための一実施形態を説明するフロー図である。一実施形態では、最終のOCR結果は、OCR結果を出力する前に、文書構造解析、および従来のOCR結果を考慮する。方法1300は、(回路、専用論理回路などの)ハードウェア、(汎用コンピューティングシステムまたは専用機械上で実行されるような)ソフトウェア、(ソフトウェアに埋め込まれる)ファームウェア、またはそれらの任意の組合せを備える処理用論理回路によって実行される。一実施形態では、方法1300は、図1のデバイス100によって実行される。

40

【0089】

図13を参照すると、ブロック1302で画像が取得される。一実施形態では、画像は、モバイルデバイスに結合されたカメラを使用して取得される。画像が取得されると、ブロック1304で、テキストなどのグラフィカルオブジェクトを含む画像の一部が識別される。画像内のグラフィカルオブジェクトには、限定はしないが、1つまたは複数のシンボル、英数字、単語、テキスト、標識、または数が含まれ得る。テキストは、図13を説明する間に、説明の目的で使用されたグラフィカルオブジェクトの一例であるが、任意のグラフィカルオブジェクトが、類似の方法と併せて使用され得る。加えて、ブロック1304で、グラフ

50

ィカル表示もまた識別されることがある(図示せず)。グラフィカル表示には、原文語、シンボルなど、オブジェクトのジェスチャ表現またはグラフィカル表現が含まれることがある。たとえば、リンゴの画像は、「リンゴ」という単語のグラフィカル表示である。全体的に、グラフィカル表示は、グラフィカルオブジェクトに対するOCR結果を改良するのを助ける。一実施形態では、長方形のテキストボックスが、グラフィカルオブジェクトの周りに置かれる。グラフィカルオブジェクトおよび/またはグラフィカル表示を有する画像の1つまたは複数の部分が検出されると、ブロック1304で、文書構造解析がブロック1308で並行して開始し得る。

【0090】

ブロック1308で、文書構造解析が開始する。文書構造解析は、グラフィカルオブジェクトのコンテキストを決定するために、テキストおよび任意の識別されたグラフィカル表示など、グラフィカルオブジェクトの構造およびレイアウトに焦点を当てる。たとえば、レストランの夕食メニューは、本のカバーとかなり異なるテキストレイアウトを有する。文書構造解析は、フォント高さのような、テキスト領域の単項素性に、また、グラフィカルオブジェクト間またはグラフィカルオブジェクトとグラフィカル表示との間の2項素性にも焦点を当て得る。余白および表など、文書の付加的な特徴もまた、文書構造解析を生成する間に考察され得る。2項素性は、共直進性ならびに水平および垂直オーバーラップの次数のような特殊な配列の解析を含み得る。

【0091】

ブロック1310で、文書構造解析からのコンテキストが、辞書を改良するために使用される。辞書は、図13において、説明の目的で使用されるグラフィカルオブジェクトの群の一例であるが、グラフィカルオブジェクトの任意の群を使用する他の方法も利用され得る。ブロック1306で、画像からのグラフィカルオブジェクトが識別される。別のOCRが、グラフィカルオブジェクトを識別するために実行され得る。グラフィカルオブジェクトを識別することにおいて、方法は、ブロック1310で生成された文書構造解析からの改良された辞書を使用し得る。ブロック1312で、コンテキスト依存OCRに対する出力テキストが生成される。

【0092】

図13に示す具体的なステップは、本発明の一実施形態による、動作のモードの間を切り替える特定の方法を提供することを諒解されたい。ステップの他の順序も、代替的な実施形態に従って実行され得る。たとえば、本発明の代替的な実施形態は、異なる順序で、上で概説されたステップを実行することができる。例示のために、ユーザは、第3の動作モードから第1の動作モードに、第4のモードから第2のモードに、またはそれらの間の任意の組合せに変えることを選択し得る。その上、図13に示される個々のステップは、個々のステップに対して様々な順序で適宜実行され得る、複数のサブステップを含み得る。さらに、追加のステップが、具体的な用途に応じて、追加または削除され得る。当業者は、方法1300の多くの変形、修正、および代替を認識し、諒解するだろう。

【0093】

図14は、コンテンツ解析を使用するコンテキスト識別のための一実施形態を説明するフロー図である。一実施形態では、最終のOCR結果は、OCR結果を出力する前に、コンテンツ意味解析、および従来のOCR結果を考慮する。方法1400は、(回路、専用論理回路などの)ハードウェア、(汎用コンピューティングシステムまたは専用機械上で実行されるような)ソフトウェア、(ソフトウェアに埋め込まれる)ファームウェア、またはそれらの任意の組合せを備える処理用論理回路によって実行される。一実施形態では、方法1400は、図1のデバイス100によって実行される。

【0094】

図14を参照すると、ブロック1402で画像が取得される。一実施形態では、画像は、モバイルデバイスに結合されたカメラを使用して取得される。画像が取得されると、ブロック1404で、テキストなどのグラフィカルオブジェクトを含む画像の一部が識別される。画像内のグラフィカルオブジェクトには、限定はしないが、1つまたは複数のシンボル、英

10

20

30

40

50

数字、単語、テキスト、標識、または数が含まれ得る。テキストは、図14を説明する間に、説明の目的で使用されたグラフィカルオブジェクトの一例であるが、任意のグラフィカルオブジェクトが、類似の方法と併せて使用され得る。加えて、ブロック1404で、グラフィカル表示もまた識別されることがある。グラフィカル表示には、原文語、シンボルなど、オブジェクトのジェスチャ表現またはグラフィカル表現が含まれることがある。たとえば、リンゴの画像は、「リンゴ」という単語のグラフィカル表示である。全体的に、グラフィカル表示は、グラフィカルオブジェクトに対するOCR結果を改良するのを助ける。一実施形態では、長方形のテキストボックスが、グラフィカルオブジェクトの周りに置かれる。ブロック1404で、グラフィカルオブジェクト領域が検出されると、ブロック1408で、テキスト領域のコンテンツ意味解析が開始し得る。

10

【0095】

グラフィカルオブジェクトを有する部分が検出されると、文書のコンテンツ意味解析がグラフィカルオブジェクトを識別することによって開始し得る(ブロック1408)。ブロック1410で、暫定OCRが実施され得る。OCRから出力された認識されたグラフィカルオブジェクト(ブロック1410)が、ブロック1412で、コンテンツ意味解析のために使用される。たとえば、数個の単語がOCRを使用して部分的に認識されると、それらの単語を使用してコンテキストを推測する様々な方式が実装され得る。一実施形態では、最も頻繁に検出されたカテゴリーに属する単語がコンテキストを構築するために使用される、多数決投票方式が使用される。別の実施形態では、ドメイン固有のコンテキストが、コンテキストに基づいて推測され得る。いくつかの実施形態では、コンテンツ意味解析は、コンテキストを推測するために、文書構造解析と併せて動作する。たとえば、レストランメニューでは、食品名と値段とが一緒に掲載される。

20

【0096】

一実施形態では、ブロック1412で、コンテンツの語義を解析することは、グラフィカルオブジェクト中の共起の確率を解析することに基づく。本発明の一態様では、コンテキストは、画像内で見出される単語などのグラフィカルオブジェクトの関連に基づいて決定され得る。たとえば、1つのイタリアンのメニュー項目は、しばしば、他のイタリアンのメニュー項目とともに見られる。同様に、地下鉄の駅名は一緒に見られる。これらの群内の単語間の相関関係は、比較的非常に高い。高い相関関係の単語およびテキストブロックを有する共起行列は、あらかじめ生成されてよく、またはデバイスの訓練によって生成または更新されてよい。

30

【0097】

一実施形態では、共起行列に対する最大化された全共起率(co-occurrence rate)は、下式で表現され得る。

【0098】

【数1】

$$\max_{(\bar{k}, \bar{m})} \left(\sum_{(TB_i, TB_j) \in TB} C(W(TB_i)_{k_i}, W(TB_j)_{k_j}) \right)$$

40

【0099】

ここで、

Cは共起行列を表現し、Wは単語を表現し、TBはテキストブロックを表現する。通常、OCRは、各テキストブロックに対してN個の最良候補を作成する。OCRがテキストブロックに対してN個の最良候補を選択すると、共起行列は、全共起率を最大化するテキストブロックを選択するために参照され得る。

【0100】

加えて、グラフィカル表示はまた、ブロック1412で、画像内のコンテンツの語義を解析

50

するのに使用され得る。グラフィカル表示には、原文語、シンボルなど、オブジェクトのジェスチャ表現またはグラフィカル表現が含まれることがある。たとえば、リンゴの画像は、「リンゴ」という単語のグラフィカル表示である。たとえば、「リンゴ」という単語は、「オレンジ」よりも「リンゴ」を表現する画像と共起する確率がより高い。

【0101】

さらに、ブロック1412で、コンテンツの語義は、複数の言語に基づいて解析される。複数の言語を論じる本発明の実施形態を、図16および図17でさらに説明する。請求項28に記載の方法では、コンテンツの語義が、複数の言語に基づいて解析される。

【0102】

ブロック1414で、コンテンツ意味解析からのコンテキストが、辞書を改良するために使用される。辞書は、図14において、説明の目的で使用されるグラフィカルオブジェクトの群の一例であるが、グラフィカルオブジェクトの任意の群を使用する他の方法も利用され得る。ブロック1406で、別のOCRが実行され得、グラフィカルオブジェクトが識別される。OCRからの結果は、ブロック1412で生成されるコンテンツ意味解析からの改良された辞書を使用する。ブロック1416で、コンテキスト依存OCRに対する出力テキストが生成される。

【0103】

図14に示す具体的なステップは、本発明の一実施形態による、動作のモードの間を切り替える特定の方法を提供することを諒解されたい。ステップの他の順序も、代替的な実施形態に従って実行され得る。たとえば、本発明の代替的な実施形態は、異なる順序で、上で概説されたステップを実行することができる。例示のために、ユーザは、第3の動作モードから第1の動作モードに、第4のモードから第2のモードに、またはそれらの間の任意の組合せに変えることを選択し得る。その上、図14に示される個々のステップは、個々のステップに対して様々な順序で適宜実行され得る、複数のサブステップを含み得る。さらに、追加のステップが、具体的な用途に応じて、追加または削除され得る。当業者は、方法1400の多くの変形、修正、および代替を認識し、諒解するだろう。

【0104】

図15は、挙動要素を使用するコンテキスト依存OCRのための一実施形態を説明するフロー図である。この実施形態は、辞書を選択するためにユーザのコンテキストを改良するために、ユーザの挙動要素を使用する。辞書は、図15において、説明の目的で使用されるグラフィカルオブジェクトの群の一例であるが、グラフィカルオブジェクトの任意の群を使用する他の方法も利用され得る。方法1500は、(回路、専用論理回路などの)ハードウェア、(汎用コンピューティングシステムまたは専用機械上で実行されるような)ソフトウェア、(ソフトウェアに埋め込まれる)ファームウェア、またはそれらの任意の組合せを備える処理用論理回路によって実行される。一実施形態では、方法1500は、図1のデバイス100によって実行される。

【0105】

図15を参照すると、ブロック1502で画像が取得される。一実施形態では、画像は、モバイルデバイスに結合されたカメラを使用して取得される。画像が取得されると、ブロック1504で、グラフィカルオブジェクトを含む画像の一部分が識別される。画像内のグラフィカルオブジェクトには、限定はしないが、1つまたは複数のシンボル、英数字、単語、標識、または数が含まれ得る。一実装形態では、長方形のテキストボックスが、グラフィカルオブジェクトの周りに置かれる。ブロック1506で、グラフィカルオブジェクトは部分的にまたは完全に認識される。一態様では、OCRは、グラフィカルオブジェクトを認識するために利用される。

【0106】

ブロック1510で、センサ入力、GPS、マイクロフォン、およびカメラなど、様々な入力ソースから受信される。時刻のような他のパラメータもまた、探索のさらなる改良に組み込まれ得る。ブロック1512で、類似のコンテキストを有するユーザの前の活動が、挙動データベースまたは挙動表を使用して探索される(ブロック1518)。挙動データベースは、

ユーザによって最も頻繁に使用されるコンテキストとコンテキストに関してユーザに固有の関連情報とのコレクションである。ブロック1514で、十分な数の活動が生成されるかまたは特定のコンテキストに関連する情報が集められる場合、コンテキストが選択され、対応するコンテキストを表現する1つまたは複数の辞書が選択される。いくつかの実施形態では、OCRが実行されて結果が生成される(ブロック1508)と、ユーザは、OCR結果内の誤りを検証して修正する能力を有し得る(ブロック1516)。フロー図に示すように、ユーザからのこのフィードバックは、ブロック1518で、挙動表を更新するのに利用され得る。

【0107】

下記の表(Table 1(表1))は、日常的に人が遭遇する挙動の活動(behavioral activity)の例示的なリスティングを説明する。Table 1(表1)はまた、例示的な挙動データベースまたは挙動表の部分的な表現を表す(ブロック1518)。

【0108】

【表1】

| 履歴 | ロケーション (GPS) | ロケーション (音響) | 音声コマンド | 時間 | 焦点距離 (テキスト の物理的 サイズ) | 光の状態 | ユーザ活動 |
|----|-----------------|----------------|--------|-------|-------------------------------|------|-------------------|
| 1 | 中心街A | レストラン | 騒がしい | 午後6時 | 30cm | 蛍光灯 | 夕食メニューを選択する |
| 2 | 乗り換え (バス) | 自動車 | 騒がしい | 午前8時 | 30cm (小さい) | 太陽光 | スクリーンを拡大して単語を選択する |
| 3 | 町外れ | 屋外 | 静寂 | 午後10時 | Inf | 暗い | HDR機能を活性化 |

Table I

【0109】

Table 1(表1)はまた、正確さを向上し、選択される辞書をさらに改良するために、本発明の複数の異なる実施形態がどのように一緒に使用され得るかの一例である。ユーザからのフィードバックは、辞書をよりよく選択することにおいてモバイルデバイスを訓練するのを助ける。

【0110】

Table 1(表1)を参照すると、履歴1が、センサによって感知された状態を、特定の時刻における挙動入力とともに表現する。第1の表項目において、GPSがモバイルデバイスのロケーションを決定する。加えて、マイクロフォンの入力は、モバイルデバイスが特定のレストランの内部にあることを決定するのを助ける。時刻は、モバイルデバイスのユーザが夕食のためにレストランにいる可能性が最も高いと判断するのを助ける。モバイルデバイスの構成要素は、環境ファクタをユーザの通常パターンと比較することによってこの選択をさらに改良し得る。たとえば、ユーザが毎夜午後6時に夕食を取る場合、ロケーション、光、騒音レベルのような環境ファクタが履歴1に対する日常的な環境入力に類似している限り、ユーザが今夜も夕食を取っている確率が高い。読んでいる資料のユーザからの距離もまた、テキストをさらに限定するのを助ける。この場合、読んでいる資料は、ユーザから約30cm離れている夕食メニューである。

【0111】

表中の履歴2は、週日の朝ほぼ8時におけるユーザの挙動の活動の履歴である。ユーザロ

ケーションは、GPSおよびマイクロフォンによって決定される。ユーザは、通常、ほぼこの時刻にバスを乗り換えている。センサ入力もまた、明るい太陽光を伴う騒がしい環境を検出する。これらの状況の下で、モバイルデバイスが画像を取得すると、ユーザはバスを乗り換える途上であり、それがコンテキストである。そのコンテキストに従って、ユーザは、スクリーンを拡大して単語を選択することを伴う追加の機能を実行し得る。

【0112】

表中の履歴3は、午後10時のユーザの履歴である。GPSおよび音響入力、ユーザが町外れに居り、非常に静かな環境の戸外に居ることを示している。光の条件は暗く、可視のテキストは視野内にない。これらの状況の下でのユーザの普通の活動は、彼らのモバイルデバイスのHDR(ハイダイナミックレンジ)機能を活性化することである。

10

【0113】

ユーザの活動の履歴は、挙動データベースを訓練するために使用される。ユーザが、特定の環境条件下で特定の活動のセットを繰り返すことを続ける場合、モバイルデバイスは、それらの環境条件をユーザの活動に関係づけることを開始し、獲得された情報が、将来、類似の環境刺激に対して反応するために使用され得る。

【0114】

図15に示す具体的なステップは、本発明の一実施形態による、動作のモードの間を切り替える特定の手法を提供することを諒解されたい。ステップの他の順序も、代替的な実施形態に従って実行され得る。たとえば、本発明の代替的な実施形態は、異なる順序で、上で概説されたステップを実行することができる。例示のために、ユーザは、第3の動作モードから第1の動作モードに、第4のモードから第2のモードに、またはそれらの間の任意の組合せに変えることを選択し得る。その上、図15に示される個々のステップは、個々のステップに対して様々な順序で適宜実行され得る、複数のサブステップを含み得る。さらに、追加のステップが、具体的な用途に応じて、追加または削除され得る。当業者は、方法1500の多くの変形、修正、および代替を認識し、諒解するだろう。

20

【0115】

図16は、例示的な交通標識を示すブロック図である。街路名は、隣り合う2つ以上の言語によるものである。多言語の交通標識、メニュー、および人気のある観光アトラクションは、多言語の国において、および同様に外人観光客の流通が著しい町において一般的である。同じ関心事を記述するために類似の意味を持つ2つの異なる言語における2つの単語の存在が、図17でさらに説明するように、OCRを実行する正確さと速さを向上するために使用され得る。

30

【0116】

図17は、標識が2つ以上の言語におけるときの、OCR結果の正確さおよび速さを向上するための方法を説明するフロー図である。方法1700は、(回路、専用論理回路などの)ハードウェア、(汎用コンピューティングシステムまたは専用機械上で実行されるような)ソフトウェア、(ソフトウェアに埋め込まれる)ファームウェア、またはそれらの任意の組合せを備える処理用論理回路によって実行される。一実施形態では、方法1300は、図1のデバイス100によって実行される。

【0117】

図17を参照すると、ブロック1702で画像が取得される。一実施形態では、画像は、モバイルデバイスに結合されたカメラを使用して取得される。画像が取得されると、ブロック1704で、グラフィカルオブジェクトを含む画像の一部分が識別される。画像内のグラフィカルオブジェクトには、限定はしないが、1つまたは複数のシンボル、英数字、単語、標識、または数が含まれ得る。一実施形態では、長方形のテキストボックスが、グラフィカルオブジェクトの周りに置かれる。ブロック1706で、グラフィカルオブジェクトは部分的にまたは完全に認識される。一態様では、OCRは、グラフィカルオブジェクトを認識するために利用される。モバイルデバイスは、画像から1つまたは複数の異なる言語(図17のAおよびB)を検出する。

40

【0118】

50

フレームが3つ以上の言語を有する場合、各言語に対して利用可能な辞書における言語を検出するために、より多くのプロセスが分岐し得る。ブロック1708で、言語Aにおいて最高の信頼レベルで感知された単語が検出される。同様に、ブロック1710で、言語Bにおいて最高の信頼レベルで感知された単語が検出される。ブロック1708およびブロック1710で、単語の検出が、従来のOCR機構を使用して実行され得る。単語は、図17を説明する間に、説明の目的で使用されたグラフィカルオブジェクトの一例であるが、他のグラフィカルオブジェクトが、限定なしに使用され得る。たとえば、異なる文化におけるシンボルは異なって見えることがある、が同じオブジェクトを指し得る。ブロック1712で、言語Bで検出された単語が、言語Aに変換される。異なる言語を使用する同じ単語の両OCR検出から最良の結果を選択するために、ブロック1714で、1708および1712からの入力組み合わせられる。一実施形態では、検出された共通の単語が直接出力され、両OCR結果間で一致しない単語は、さらに処理される。

10

【0119】

図17に示す具体的なステップは、本発明の一実施形態による、動作のモードの間を切り替える特定の方法を提供することを諒解されたい。ステップの他の順序も、代替的な実施形態に従って実行され得る。たとえば、本発明の代替的な実施形態は、異なる順序で、上で概説されたステップを実行することができる。例示のために、ユーザは、第3の動作モードから第1の動作モードに、第4のモードから第2のモードに、またはそれらの間の任意の組合せに変えることを選択し得る。その上、図17に示される個々のステップは、個々のステップに対して様々な順序で適宜実行され得る、複数のサブステップを含み得る。さらに、追加のステップが、具体的な用途に応じて、追加または削除され得る。当業者は、方法1700の多くの変形、修正、および代替を認識し、諒解するだろう。

20

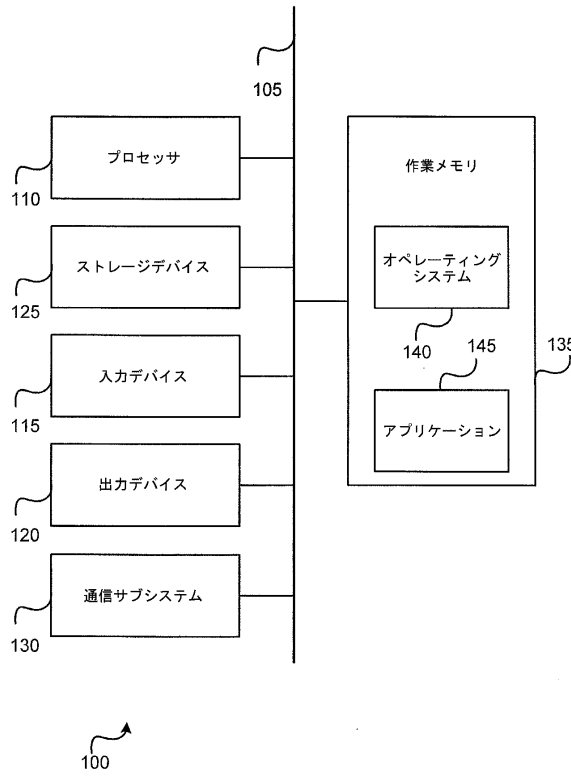
【符号の説明】

【0120】

- 100 コンピュータシステム
- 105 バス
- 110 プロセッサ
- 115 入力デバイス
- 120 出力デバイス
- 125 非一時的ストレージデバイス
- 130 通信サブシステム
- 135 非一時的作業メモリ
- 140 オペレーティングシステム
- 145 アプリケーションプログラム
- 300 方法
- 302 ブロック
- 304 ブロック
- 306 ブロック
- 308 ブロック

30

【図 1】



【図 2 A】

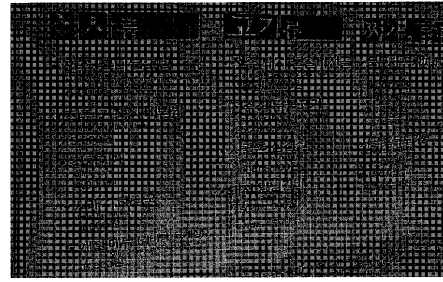


FIG. 2A

【図 2 B】

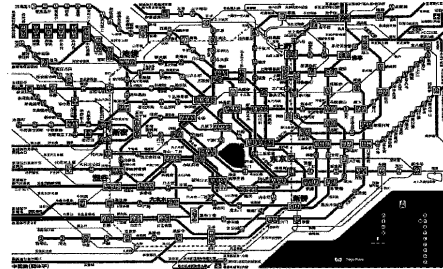
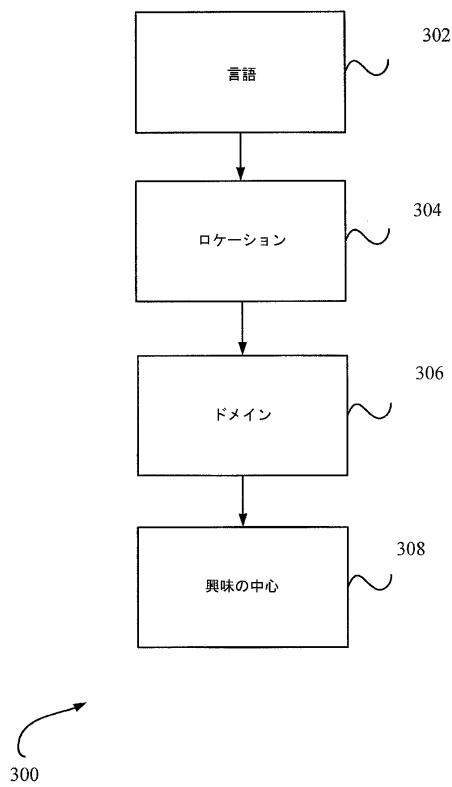
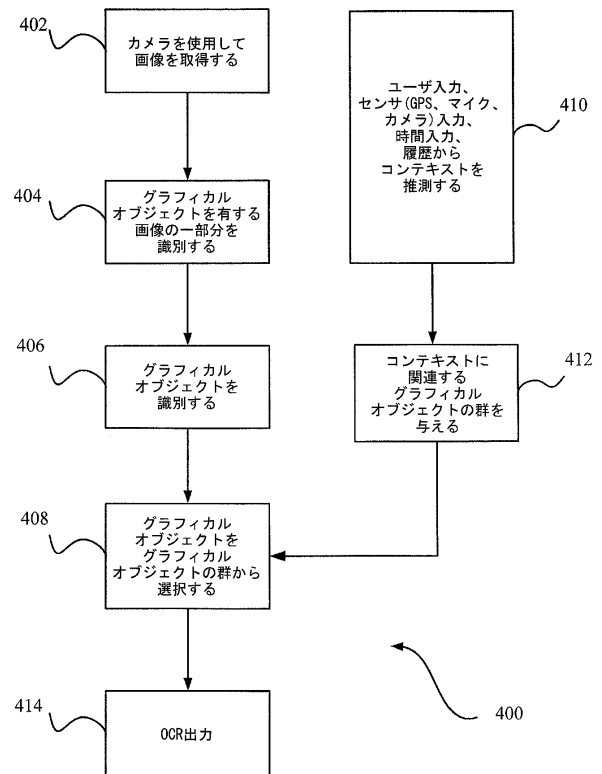


FIG. 2B

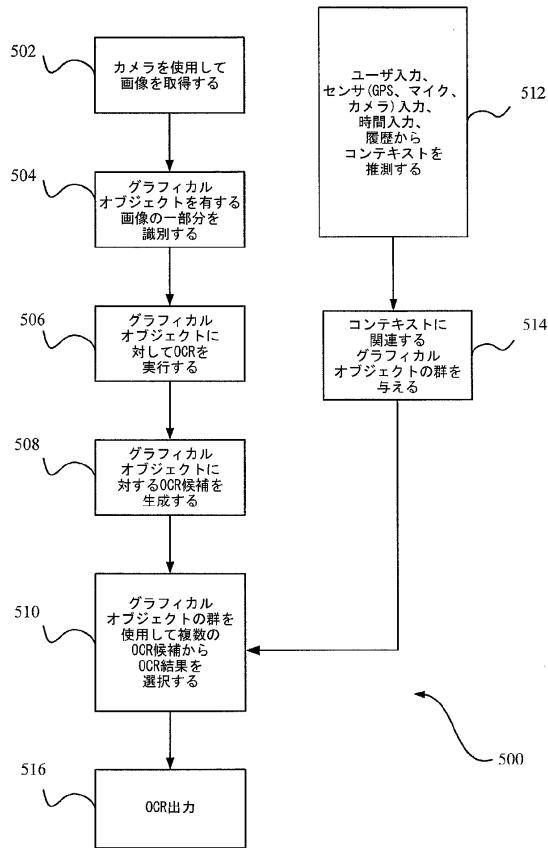
【図 3】



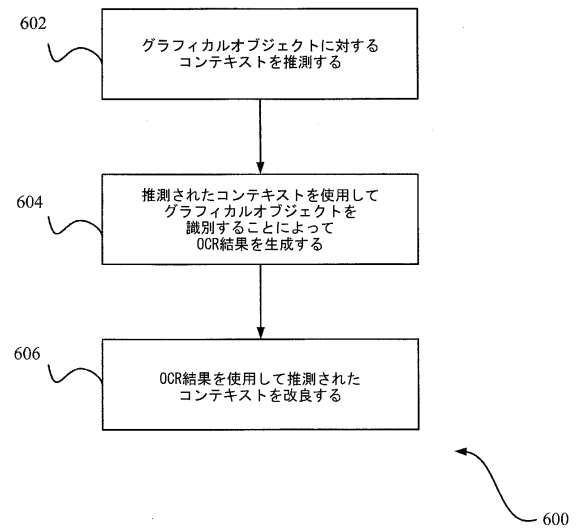
【図 4】



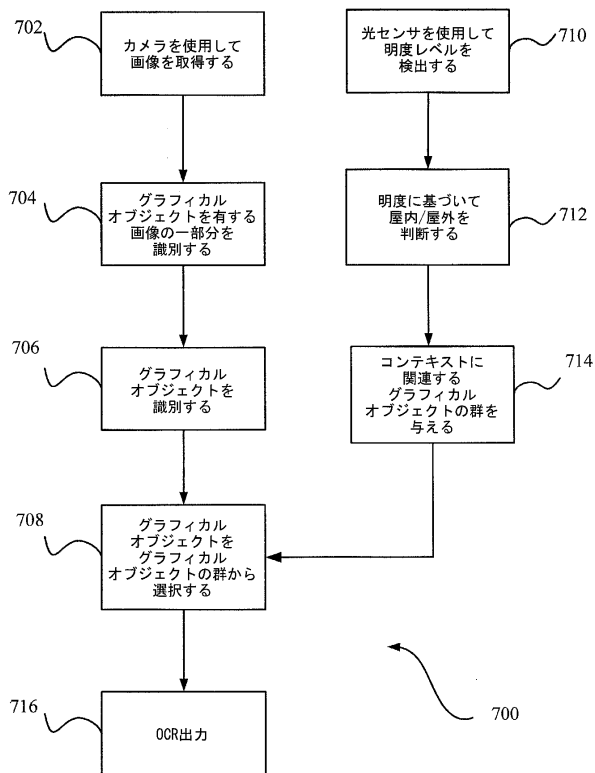
【図 5】



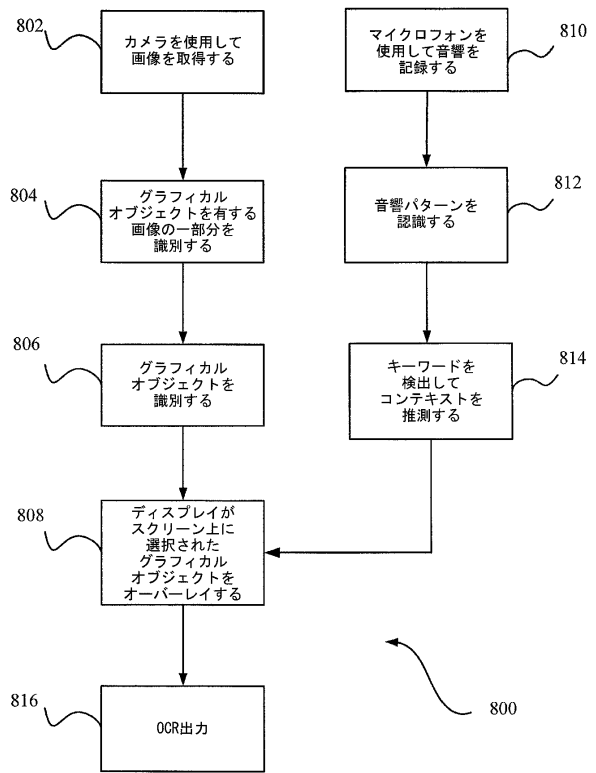
【図 6】



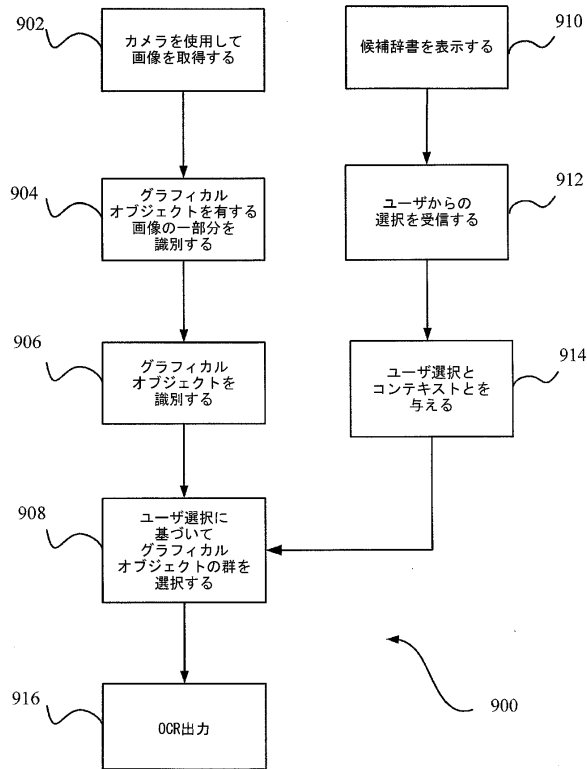
【図 7】



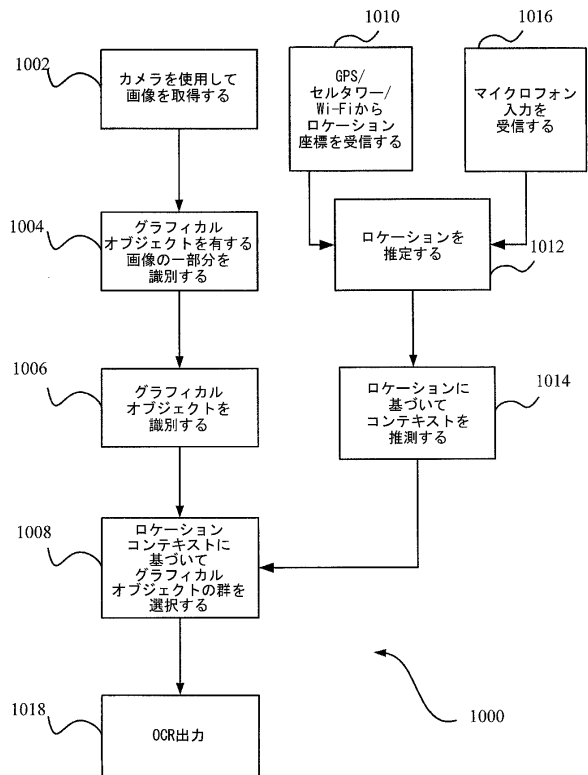
【図 8】



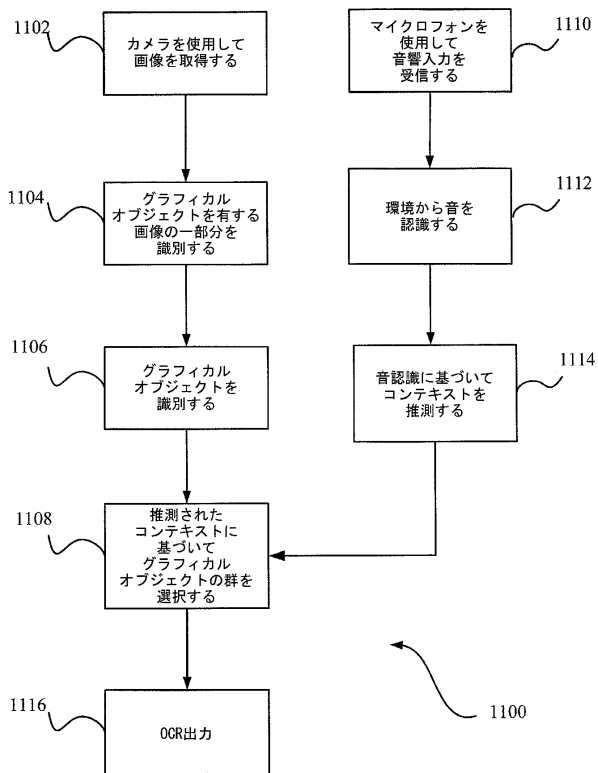
【図 9】



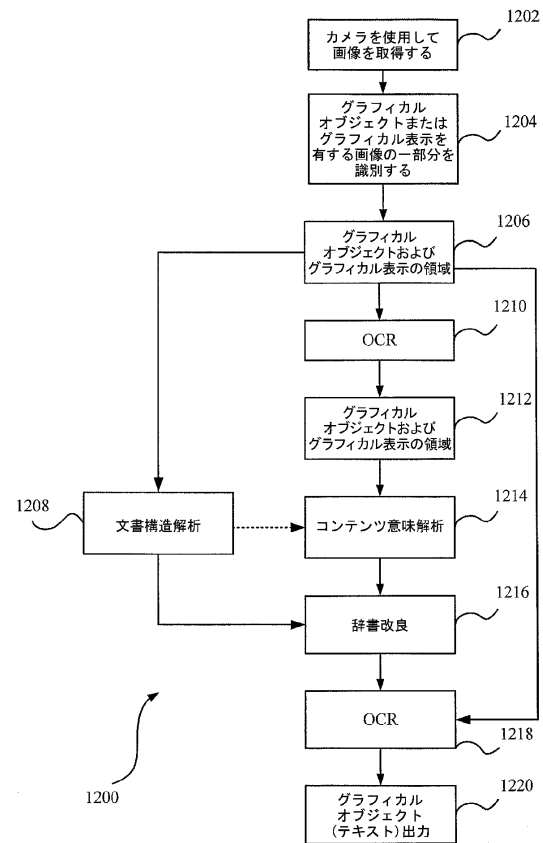
【図 10】



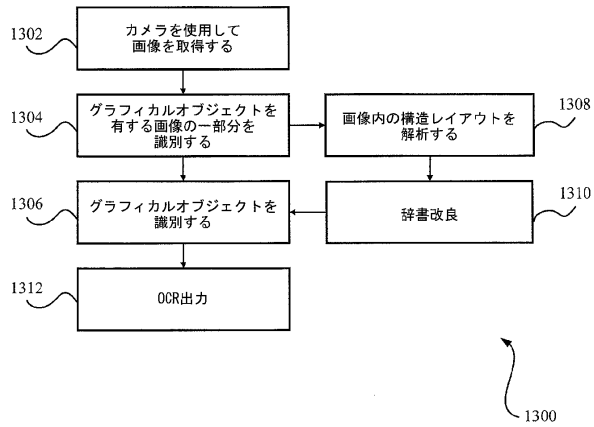
【図 11】



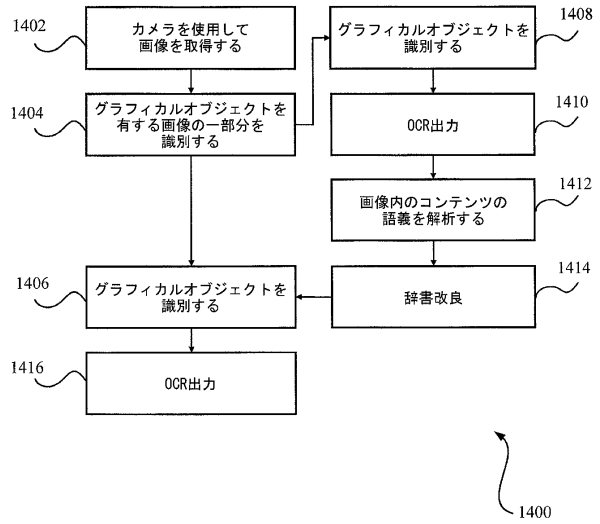
【図 12】



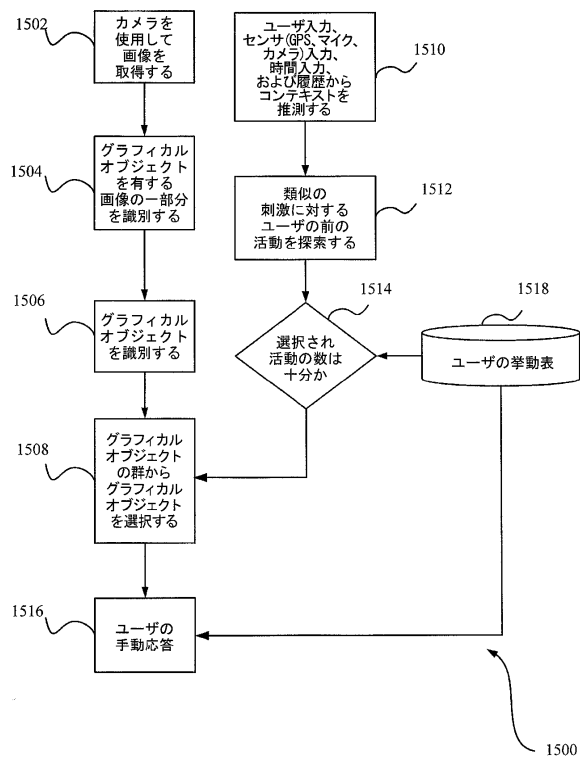
【図 13】



【図 14】



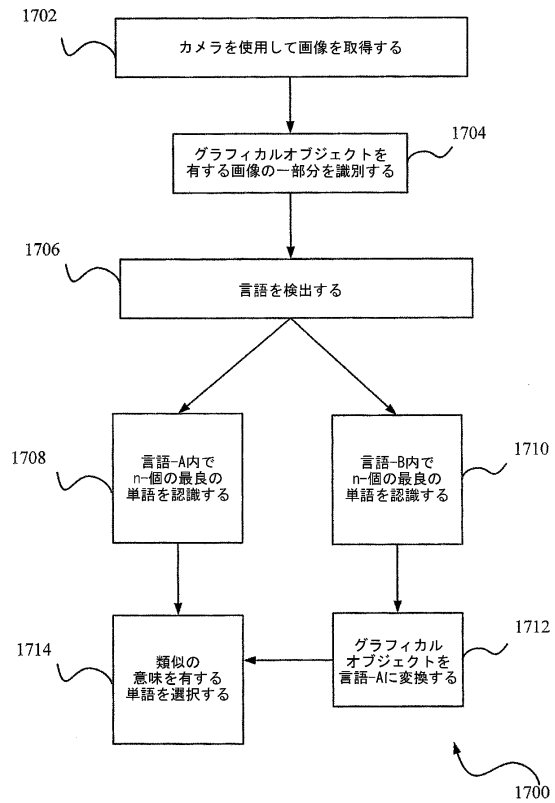
【図 15】



【図 16】



【図 17】



フロントページの続き

- (72)発明者 テ・ウォン・イ
アメリカ合衆国・カリフォルニア・92121-1714・サン・ディエゴ・モアハウス・ドライ
ヴ・5775
- (72)発明者 ダク・フン・キム
アメリカ合衆国・カリフォルニア・92121-1714・サン・ディエゴ・モアハウス・ドライ
ヴ・5775
- (72)発明者 キスン・ユ
アメリカ合衆国・カリフォルニア・92121-1714・サン・ディエゴ・モアハウス・ドライ
ヴ・5775
- (72)発明者 ミンホ・ジン
アメリカ合衆国・カリフォルニア・92121-1714・サン・ディエゴ・モアハウス・ドライ
ヴ・5775
- (72)発明者 テス・キム
アメリカ合衆国・カリフォルニア・92121-1714・サン・ディエゴ・モアハウス・ドライ
ヴ・5775
- (72)発明者 ヒュン・ムク・チョ
アメリカ合衆国・カリフォルニア・92121-1714・サン・ディエゴ・モアハウス・ドライ
ヴ・5775

合議体

審判長 藤井 浩
審判官 小池 正彦
審判官 渡辺 努

- (56)参考文献 特開2003-178067(JP,A)
特開2003-108551(JP,A)
特開2000-348142(JP,A)
特開2006-65477(JP,A)
特開2011-134144(JP,A)
特開2009-86349(JP,A)
特開2009-258871(JP,A)

- (58)調査した分野(Int.Cl., DB名)

G06T 1/00- 7/00