

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6403766号  
(P6403766)

(45) 発行日 平成30年10月10日 (2018. 10. 10)

(24) 登録日 平成30年9月21日 (2018. 9. 21)

(51) Int. Cl. F I  
**G 0 6 F 17/28 (2006. 01)** G O 6 F 17/28 6 9 0  
**G 0 6 F 17/30 (2006. 01)** G O 6 F 17/30 3 1 0 C

請求項の数 19 (全 47 頁)

(21) 出願番号	特願2016-525041 (P2016-525041)	(73) 特許権者	502208397
(86) (22) 出願日	平成26年11月5日 (2014. 11. 5)		グーグル エルエルシー
(65) 公表番号	特表2016-535335 (P2016-535335A)		アメリカ合衆国 カリフォルニア州 94
(43) 公表日	平成28年11月10日 (2016. 11. 10)		043 マウンテン ビュー アンフィシ
(86) 国際出願番号	PCT/US2014/064086		アター パークウェイ 1600
(87) 国際公開番号	W02015/069737	(74) 代理人	100108453
(87) 国際公開日	平成27年5月14日 (2015. 5. 14)		弁理士 村山 靖彦
審査請求日	平成29年6月20日 (2017. 6. 20)	(74) 代理人	100110364
(31) 優先権主張番号	61/901, 996		弁理士 実広 信哉
(32) 優先日	平成25年11月8日 (2013. 11. 8)	(74) 代理人	100133400
(33) 優先権主張国	米国 (US)		弁理士 阿部 達彦
(31) 優先権主張番号	14/075, 956		
(32) 優先日	平成25年11月8日 (2013. 11. 8)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 画像内に描かれたテキストの翻訳の提示

(57) 【特許請求の範囲】

【請求項 1】

データ処理装置によって実行される方法であって、  
 画像を受信するステップと、  
 前記画像内に描かれたテキストをそれぞれが含む複数のはっきりと異なるテキストブロックを特定するステップと、  
 関連するテキストブロック内に描かれたテキストの視覚的特徴と前記関連するテキストブロック内に描かれた前記テキストの配列とに基づいて前記関連するテキストブロックの複数の集合を特定するステップであって、関連するテキストブロックの各集合が、一致している視覚的特徴を持つテキストを含む、ステップと、  
 前記画像に関して、前記特定された複数の集合に基づいて、複数の提示コンテキストから提示コンテキストを選択するステップであって、各提示コンテキストが、前記特定された複数の集合における特定の集合内に含まれた前記テキストの少なくとも一部の翻訳を提示するための少なくとも1つの対応するユーザインターフェースを有し、各提示コンテキストのための前記ユーザインターフェースが、その他の提示コンテキストのための前記ユーザインターフェースとは異なり、前記複数の提示コンテキストにおける第1の提示コンテキストのための前記ユーザインターフェースが、前記複数の提示コンテキストにおける第2の提示コンテキストのための前記ユーザインターフェースとは異なるテキストブロックの集合の翻訳を含む、ステップと、  
 前記選択された提示コンテキストにおける前記特定の集合内に含まれた前記テキストの

10

20

前記少なくとも一部の翻訳に関する読みやすさの尺度に基づいて、前記選択された提示コンテキストに対応する前記ユーザインターフェースから、特定のユーザインターフェースを選択するステップであって、前記読みやすさの尺度は、前記特定のユーザインターフェースを用いて提示され得る前記翻訳の文字数と、前記翻訳の総文字数との間の比である、ステップと、

前記選択されたユーザインターフェースを用いて前記選択された提示コンテキストにおける前記特定の集合内に含まれた前記テキストの前記少なくとも一部の翻訳を提示するステップとを含む、方法。

【請求項2】

前記画像に関する前記提示コンテキストを選択するステップが、  
少なくとも前記画像内に描かれた前記テキストの前記配列に基づいて、前記テキストの第1の部分が前記テキストの少なくとも1つのその他の部分よりも顕著に提示されると決定するステップを含み、

前記特定のユーザインターフェースを選択するステップが、前記テキストにおける前記第1の部分の翻訳を提示するユーザインターフェースを選択するステップを含む請求項1に記載の方法。

【請求項3】

前記選択された提示コンテキストにおける前記特定の集合内に含まれた前記テキストの前記少なくとも一部の翻訳を提示するステップが、前記テキストの前記第1の部分上のオーバーレイ内に前記テキストの前記第1の部分の前記翻訳を提示するステップを含む請求項2に記載の方法。

【請求項4】

前記提示コンテキストを選択するステップが、  
前記画像内に描かれた2つ以上のテキストブロックが互いに関連しないと決定するステップと、

前記関連しないとの前記決定に応じて、集合提示コンテキストとして複数項目提示コンテキストを選択するステップであって、複数項目提示コンテキストが、決定に応じて別個の表示ページに各テキストブロックの言語の翻訳を提示する対応するユーザインターフェースを有する、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5】

2つ以上のテキストブロック内に描かれた前記テキストにおける複数の視覚的特徴間の類似性に基づいて関連付けられる前記2つ以上のテキストブロックを決定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6】

2つ以上のテキストブロックのそれぞれの対の間の距離が閾値の距離よりも短いことに基づいて関連付けられる前記2つ以上のテキストブロックを決定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7】

前記画像に関して同じ向きを持っている2つ以上のテキストブロック内に描かれたテキストに基づいて関連付けられる前記2つ以上のテキストブロックを決定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8】

前記提示コンテキストを選択するステップが、  
前記テキストブロックの配列および提示の評価に基づき、前記選択された提示コンテキストにおける前記特定の集合内に含まれた前記テキストがメニューのテキストを含むと決定するステップと、

メニューのテキストを含む前記選択された提示コンテキストにおける前記特定の集合内に含まれた前記テキストを決定するステップに基づき、各メニュー項目の翻訳を独立に提示する対応するユーザインターフェースを持っている提示コンテキストを選択するステッ

10

20

30

40

50

ブであって、前記対応するユーザインターフェースは、ユーザがその他のメニュー項目の翻訳を要求することを可能にするユーザインターフェースコントロールを含む、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9】

前記提示コンテキストを選択するステップが、

前記テキストブロックの配列および提示の評価に基づき、前記選択された提示コンテキストにおける前記特定の集合内に含まれた前記テキストが複数の異なる道路標識のテキストを含むと決定するステップと、

複数の異なる道路標識のテキストを含む前記選択された提示コンテキストにおける前記特定の集合内に含まれた前記テキストを決定するステップに基づき、前記道路標識上のオーバーレイ内に各道路標識の翻訳を提示する対応するユーザインターフェースを持っている提示コンテキスト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10】

データ処理装置と、

前記データ処理装置とデータ通信するメモリストレージ装置であって、前記データ処理装置によって実行可能であり、そのように実行されると、前記データ処理装置に、

画像を受信する動作、

前記画像内に描かれたテキストをそれぞれが含む複数のはっきりと異なるテキストブロックを特定する動作、

関連するテキストブロック内に描かれたテキストの視覚的特徴と前記関連するテキストブロック内に描かれた前記テキストの配列とに基づいて前記関連するテキストブロックの複数の集合を特定する動作であって、関連するテキストブロックの各集合が、一致している視覚的特徴を持つテキストを含む、特定する動作、

前記画像に関して、前記特定された複数の集合に基づいて、複数の提示コンテキストから提示コンテキストを選択する動作であって、各提示コンテキストが、前記特定された複数の集合における特定の集合内に含まれた前記テキストの少なくとも一部の翻訳を提示するための少なくとも1つの対応するユーザインターフェースを有し、各提示コンテキストのための前記ユーザインターフェースが、その他の提示コンテキストのための前記ユーザインターフェースとは異なり、前記複数の提示コンテキストにおける第1の提示コンテキストのための前記ユーザインターフェースが、前記複数の提示コンテキストにおける第2の提示コンテキストのための前記ユーザインターフェースとは異なるテキストブロックの集合の翻訳を含む、選択する動作、

前記選択された提示コンテキストにおける前記特定の集合内に含まれた前記テキストの前記少なくとも一部の翻訳に関する読みやすさの尺度に基づいて、前記選択された提示コンテキストに対応する前記ユーザインターフェースから、特定のユーザインターフェースを選択する動作であって、前記読みやすさの尺度は、前記特定のユーザインターフェースを用いて提示され得る前記翻訳の文字数と、前記翻訳の総文字数との間の比である、選択する動作、および

前記選択されたユーザインターフェースを用いて前記選択された提示コンテキストにおける前記特定の集合内に含まれた前記テキストの前記少なくとも一部の翻訳を提示する動作を含む動作を実行させる命令を記憶する、メモリストレージ装置とを含むシステム。

【請求項11】

前記画像に関する前記提示コンテキストを選択する動作が、

少なくとも前記画像内に描かれた前記テキストの前記配列に基づいて、前記テキストの第1の部分が前記テキストの少なくとも1つのその他の部分よりも顕著に提示されると決定することを含み、

前記特定のユーザインターフェースを選択する動作が、前記テキストにおける前記第1の部分の翻訳を提示するユーザインターフェースを選択する動作を含む請求項10に記載の

10

20

30

40

50

システム。

【請求項 1 2】

前記選択された提示コンテキストにおける前記特定の集合内に含まれた前記テキストの前記少なくとも一部の翻訳を提示する動作が、前記テキストの前記第1の部分上のオーバーレイ内に前記テキストの前記第1の部分の前記翻訳を提示することを含む請求項11に記載のシステム。

【請求項 1 3】

前記提示コンテキストを選択する動作が、  
前記画像内に描かれた2つ以上のテキストブロックが互いに関連しないと決定する動作と、

前記関連しないとの前記決定に応じて、集合提示コンテキストとして複数項目提示コンテキストを選択する動作であって、複数項目提示コンテキストが、決定に応じて別個の表示ページに各テキストブロックの言語の翻訳を提示する対応するユーザインターフェースを有する、選択する動作と、  
を含む請求項10に記載のシステム。

【請求項 1 4】

前記命令が前記データ処理装置に実行させる前記動作が、さらに、  
2つ以上のテキストブロック内に描かれた前記テキストにおける複数の視覚的特徴間の類似性に基づいて関連付けられる前記2つ以上のテキストブロックを決定する動作を含む請求項10に記載のシステム。

【請求項 1 5】

前記命令が前記データ処理装置に実行させる前記動作が、さらに、  
2つ以上のテキストブロックのそれぞれの対の間の距離が閾値の距離よりも短いことに基づいて関連付けられる前記2つ以上のテキストブロックを決定する動作を含む請求項10に記載のシステム。

【請求項 1 6】

前記命令が前記データ処理装置に実行させる前記動作が、さらに、  
前記画像に関して同じ向きを持っている2つ以上のテキストブロック内に描かれたテキストに基づいて関連付けられる前記2つ以上のテキストブロックを決定する動作を含む請求項10に記載のシステム。

【請求項 1 7】

前記提示コンテキストを選択する動作が、  
前記テキストブロックの配列および提示の評価に基づき、前記選択された提示コンテキストにおける前記特定の集合内に含まれた前記テキストがメニューのテキストを含むと決定する動作と、

メニューのテキストを含む前記選択された提示コンテキストにおける前記特定の集合内に含まれた前記テキストを決定する動作に基づき、各メニュー項目の翻訳を独立に提示する対応するユーザインターフェースを持っている提示コンテキストを選択する動作であって、前記対応するユーザインターフェースは、ユーザがその他のメニュー項目の翻訳を要求することを可能にするユーザインターフェースコントロールを含む、選択する動作と  
を含む請求項10に記載のシステム。

【請求項 1 8】

前記提示コンテキストを選択する動作が、  
前記テキストブロックの配列および提示の評価に基づき、前記選択された提示コンテキストにおける前記特定の集合内に含まれた前記テキストが複数の異なる道路標識のテキストを含むと決定する動作と、

複数の異なる道路標識のテキストを含む前記選択された提示コンテキストにおける前記特定の集合内に含まれた前記テキストを決定する動作に基づき、前記道路標識上のオーバーレイ内に各道路標識の翻訳を提示する対応するユーザインターフェースを持っている提示コンテキストを選択する動作と

10

20

30

40

50

を含む請求項10に記載のシステム。

【請求項19】

コンピュータプログラムを符号化された非一時的コンピュータストレージ媒体であって、前記プログラムが、データ処理装置によって実行されるときに、前記データ処理装置に、

画像を受信する動作と、

前記画像内に描かれたテキストをそれぞれが含む複数のはっきりと異なるテキストブロックを特定する動作と、

関連するテキストブロック内に描かれたテキストの視覚的特徴と前記関連するテキストブロック内に描かれた前記テキストの配列とに基づいて前記関連するテキストブロックの複数の集合を特定する動作であって、関連するテキストブロックの各集合が、一致している視覚的特徴を持つテキストを含む、特定する動作と、

前記画像に関して、前記特定された複数の集合に基づいて、複数の提示コンテキストから提示コンテキストを選択する動作であって、各提示コンテキストが、前記特定された複数の集合における特定の集合内に含まれた前記テキストの少なくとも一部の翻訳を提示するための少なくとも1つの対応するユーザインターフェースを有し、各提示コンテキストのための前記ユーザインターフェースが、その他の提示コンテキストのための前記ユーザインターフェースとは異なり、前記複数の提示コンテキストにおける第1の提示コンテキストのための前記ユーザインターフェースが、前記複数の提示コンテキストにおける第2の提示コンテキストのための前記ユーザインターフェースとは異なるテキストブロックの

集合の翻訳を含む、選択する動作と、  
前記選択された提示コンテキストにおける前記特定の集合内に含まれた前記テキストの前記少なくとも一部の翻訳に関する読みやすさの尺度に基づいて、前記選択された提示コンテキストに対応する前記ユーザインターフェースから、特定のユーザインターフェースを選択する動作であって、前記読みやすさの尺度は、前記特定のユーザインターフェースを用いて提示され得る前記翻訳の文字数と、前記翻訳の総文字数との間の比である、選択する動作と、

前記選択されたユーザインターフェースを用いて前記選択された提示コンテキストにおける前記特定の集合内に含まれた前記テキストの前記少なくとも一部の翻訳を提示する動作とを含む動作を実行させる命令を含む、非一時的コンピュータストレージ媒体。

【発明の詳細な説明】

【技術分野】

【0001】

画像内に描かれたテキストの翻訳の提示。

【背景技術】

【0002】

スマートフォンなどの多くのユーザデバイスは、写真およびデジタル画像を撮影するためのカメラを装備されて提供される。ユーザは、旅行中にカメラを用いて写真を撮影し、マルチメディアメッセージングまたはソーシャルネットワーキングサイトを介して写真を共有することが多い。ユーザは、画像を用いて対象または場所に関連する追加的な情報を取得する可能性もある。たとえば、ユーザは、名所旧跡の画像を撮影し、画像をクエリとして用いてインターネット検索を開始する可能性がある。

【0003】

ユーザと異なる言語を話す別の国を旅行しているとき、ユーザは、ユーザが遭遇するテキストの翻訳を得たい可能性がある。たとえば、ユーザは、別の国を旅行している間に外国語の道路標識に遭遇するか、または外国語の項目のリストに遭遇する可能性がある。

【発明の概要】

【課題を解決するための手段】

【0004】

本明細書は、画像に関連する提示コンテキスト(presentation context)に基づいて画像

10

20

30

40

50

内に描かれたテキストに関連する追加的な情報(たとえば、テキストの言語の翻訳)を提示することに関連するテクノロジーを説明する。

【0005】

概して、本明細書に記載の対象の1つの革新的な態様は、画像を受信する行為と、画像内に描かれたテキストを特定する行為と、画像に関して、画像によって描かれたテキストの配列に基づいて提示コンテキストから提示コンテキストを選択する行為であって、各提示コンテキストが、画像内のテキストの特定の配列に対応し、各提示コンテキストが、画像内に描かれたテキストに関連する追加的な情報を提示するための対応するユーザインターフェースを有し、各提示コンテキストのためのユーザインターフェースが、その他の提示コンテキストのためのユーザインターフェースとは異なる、選択する行為と、選択された提示コンテキストに対応するユーザインターフェースを特定する行為と、特定されたユーザインターフェースを用いて画像内に描かれたテキストの少なくとも一部に関する追加的な情報を提示する行為であって、ユーザインターフェースが、画像上のオーバーレイ内に追加的な情報を提示する、提示する行為とを含む方法において具現化され得る。この態様のその他の実施形態は、コンピュータストレージデバイス上に符号化された方法の行為を実行するように構成された対応するシステム、装置、およびコンピュータプログラムを含む。

10

【0006】

これらのおよびその他の実施形態は、それぞれ、任意で、以下の特徴のうちの1つまたは複数を含む可能性がある。追加的な情報は、特定されたテキストの少なくとも一部の言語の翻訳を含み得る。画像に関する提示コンテキストを選択する行為は、画像内に描かれた個々のテキストブロックの数を特定することと、画像内に描かれた個々のテキストブロックの数に基づいて画像に関する提示コンテキストを選択することとを含み得る。

20

【0007】

画像に関する提示コンテキストを選択する行為は、画像内に描かれたテキストの配列に基づいて、テキストの第1の部分がテキストの少なくとも1つのその他の部分よりも顕著に提示されると決定することと、決定に応じて提示コンテキストから顕著さコンテキスト(prominence context)を選択することとを含み得る。

【0008】

選択された提示コンテキストに対応するユーザインターフェースを特定する行為は、顕著さコンテキストに対応する顕著さユーザインターフェース(prominence user interface)を特定することを含み得る。画像内に描かれたテキストの少なくとも一部に関する追加的な情報を提示する行為は、テキストの第1の部分上のオーバーレイ内にテキストの第1の部分の言語の翻訳を提示することを含み得る。

30

【0009】

画像に関する提示コンテキストを選択する行為は、画像内に描かれたテキストが住所を含むと決定することと、決定に応じて提示コンテキストから地図コンテキスト(map context)を選択することとを含み得る。選択された提示コンテキストに対応するユーザインターフェースを特定する行為は、地図コンテキストに対応する地図ユーザインターフェース(map user interface)を特定することを含み得る。画像内に描かれたテキストの少なくとも一部に関する追加的な情報を提示する行為は、地図の画像内に住所の言語の翻訳を提示することを含み得る。

40

【0010】

選択された提示コンテキストに対応するユーザインターフェースを特定する行為は、地図コンテキストに対応する地図ユーザインターフェースを特定することを含み得る。画像内に描かれたテキストの少なくとも一部に関する追加的な情報を提示する行為は、地図の画像内に住所に対応する見出しを提示することを含み得る。見出しは、地図の画像を提示するデバイスに関連するユーザの言語で提示され得る。

【0011】

画像に関する提示コンテキストを選択する行為は、画像内に描かれた個々のテキストブ

50

ロックを特定することと、個々のテキストブロックの配列および個々のテキストブロックの提示に基づいて、個々のテキストブロックがテキストの集合(collection)に属すると決定することと、決定に応じて提示コンテキストから集合コンテキスト(collection context)を選択することとを含み得る。

【 0 0 1 2 】

選択された提示コンテキストに対応するユーザインターフェースを特定する行為は、集合コンテキストに対応する集合ユーザインターフェース(collection user interface)を特定することを含み得る。画像内に描かれたテキストの少なくとも一部に関する追加的な情報を提示する行為は、画像上のオーバーレイ内にそれぞれの個々のテキストブロック内に描かれたテキストの言語の翻訳を提示することを含み得る。それぞれの言語の翻訳は、言語の翻訳の対応するテキストブロックの上に提示され得る。

10

【 0 0 1 3 】

概して、本明細書に記載の対象の別の態様は、画像を受信する行為と、画像内に描かれたテキストを特定する行為であって、特定されるテキストが、画像内で特定された1つまたは複数のテキストブロック内にあり、各テキストブロックが、画像内のその他のテキストブロックとはっきりと異なる、特定する行為と、画像に関して、顕著さコンテキストから顕著さ提示コンテキスト(prominence presentation context)を選択する行為であって、各顕著さ提示コンテキストが、画像内でテキストが提示される各テキストブロックの相対的な顕著さに対応し、各顕著さ提示コンテキストが、画像内に描かれた特定されたテキストに関連する追加的な情報を提示するための対応するユーザインターフェースを有し、顕著さ提示コンテキストの選択が、特定されたテキストが画像内で描かれる1つまたは複数のテキストブロックの相対的な顕著さに基づく、選択する行為と、選択された顕著さ提示コンテキストに対応するユーザインターフェースを特定する行為と、特定されたユーザインターフェースを用いて画像内に描かれたテキストの少なくとも一部に関する追加的な情報を提示する行為であって、ユーザインターフェースが、画像上のオーバーレイ内に追加的な情報を提示する、提示する行為とを含む方法において具現化され得る。この態様のその他の実施形態は、コンピュータストレージデバイス上に符号化された方法の行為を実行するように構成された対応するシステム、装置、およびコンピュータプログラムを含む。

20

【 0 0 1 4 】

これらのおよびその他の実施形態は、それぞれ、任意で、以下の特徴のうちの1つまたは複数を含む可能性がある。画像に関する顕著さ提示コンテキストを選択する行為は、画像内に描かれたテキストが、画像内の単一のテキストブロックの位置および単一のテキストブロックに含まれるテキストのサイズに基づいて顕著に表示されるテキストの単一のテキストブロックを含むと決定することと、決定に応じて顕著さ提示コンテキストから単一ブロックコンテキスト(single block context)を選択することとを含み得る。単一ブロックコンテキストは、顕著に表示されるテキストの単一のテキストブロック上のオーバーレイ内に顕著に表示されるテキストの単一のテキストブロックに含まれるテキストの言語の翻訳を提示するユーザインターフェースに対応する可能性がある。

30

【 0 0 1 5 】

画像内に描かれたテキストを特定する行為は、第1のテキストブロックおよび第1のテキストブロックとは異なる第2のテキストブロックを特定することであって、第1のテキストブロックおよび第2のテキストブロックが画像内に描かれる、特定することを含み得る。画像に関する顕著さ提示コンテキストを選択する行為は、第1のテキストブロックが第2のテキストブロックよりも画像内で顕著に表示されると決定することと、決定に応じて顕著さ提示コンテキストから支配的-副次的ブロックコンテキスト(dominant-secondary block context)を選択することとを含み得る。支配的-副次的ブロックコンテキストは、テキストの支配的なブロックの言語の翻訳を提示するユーザインターフェースに対応する可能性がある。

40

【 0 0 1 6 】

50

特定されたユーザインターフェースを用いて画像内に描かれたテキストの少なくとも一部に関する追加的な情報を提示する行為は、画像上のオーバーレイ内に第1のテキストブロックに含まれるテキストの言語の翻訳を提示することを含み得る。オーバーレイは、画像内の第1のテキストブロック上に置かれる可能性がある。

【0017】

態様は、画像内の第2のテキストブロックを描く際に支配的-副次的ユーザインターフェース(dominant-secondary user interface)内に選択可能なユーザインターフェース要素を提示することと、選択可能なユーザインターフェース要素の選択を受信することに応じて、第2のテキストブロックに含まれるテキストの言語の翻訳を提示することとを含み得る。

10

【0018】

第1のテキストブロックがテキストの第2のブロックよりも画像内で顕著に表示されると決定することは、第1のテキストブロックに含まれるテキストが第2のテキストブロックに含まれるテキストよりも大きいと決定することと、第1のテキストブロックが第2のテキストブロックよりも画像の中心の近くに置かれると決定することとを含み得る。

【0019】

画像内に描かれたテキストを特定する行為は、第1のテキストブロックおよび第1のテキストブロックとは異なる第2のテキストブロックを特定することを含み得る。第1のテキストブロックおよび第2のテキストブロックは、画像内に描かれる可能性がある。画像に関する顕著さ提示コンテキストを選択する行為は、第1のテキストブロックが第2のテキストブロックと実質的に同等の顕著さで画像内に表示されると決定することと、決定に応じて顕著さ提示コンテキストから同等顕著さ提示コンテキスト(equal prominence presentation context)を選択することとを含み得る。同等顕著さ提示コンテキストは、実質的に同等の顕著さを有するテキストの複数のブロックの言語の翻訳を提示するユーザインターフェースに対応する可能性がある。

20

【0020】

特定されたユーザインターフェースを用いて画像内に描かれたテキストの少なくとも一部に関する追加的な情報を提示する行為は、画像上の第1のオーバーレイ内に第1のテキストブロックに含まれるテキストの言語の翻訳を提示することと、第1のオーバーレイが、画像内の第1のテキストブロック上に置かれる、提示することと、画像上の第2のオーバーレイ内に第2のテキストブロックに含まれるテキストの言語の翻訳を提示することと、第2のオーバーレイが、画像内の第2のテキストブロック上に置かれる、提示することとを含み得る。

30

【0021】

第1のテキストブロックが第2のテキストブロックと実質的に同等の顕著さで画像内に表示されると決定することは、第1のテキストブロックに含まれるテキストのサイズが第2のテキストブロックに含まれるテキストのサイズに実質的に等しいと決定することを含み得る。

【0022】

概して、本明細書に記載の対象の別の態様は、画像を受信する行為と、画像内に描かれたテキストを特定する行為と、特定されるテキストが、画像内に描かれたテキストブロック内にあり、各テキストブロックが、画像内のその他のテキストブロックとはっきりと異なる、特定する行為と、集合提示コンテキスト(collection presentation context)から集合提示コンテキストを選択する行為と、各集合提示コンテキストが、画像内のテキストブロックの配列に対応し、各集合提示コンテキストが、少なくとも1つのテキストブロックに関連する追加的な情報を提示するための対応するユーザインターフェースを有し、集合提示コンテキストの選択が、画像内に描かれたテキストブロックの配列に基づく、選択する行為と、選択された集合提示コンテキストに対応するユーザインターフェースを特定する行為と、特定されたユーザインターフェースを使用してテキストブロックのうちの少なくとも1つに関連する追加的な情報を提示する行為とを含む方法において

40

50

具現化され得る。この態様のその他の実施形態は、コンピュータストレージデバイス上に符号化された方法の行為を実行するように構成された対応するシステム、装置、およびコンピュータプログラムを含む。

【0023】

これらのおよびその他の実施形態は、それぞれ、任意で、以下の特徴のうちの1つまたは複数を含む可能性がある。集合提示コンテキストから集合提示コンテキストを選択する行為は、それぞれのテキストブロックのテキストのそれぞれのその他のテキストブロックのテキストとの比較に基づいて、テキストブロックの各々が互いに関連すると決定することと、決定に応じて単一集合提示コンテキスト(single collection presentation context)を選択することとを含み得る。単一集合提示コンテキストは、画像上の1つまたは複数のオーバーレイ内に各テキストブロックの言語の翻訳を提示する対応するユーザインターフェースを有する可能性がある。

10

【0024】

集合提示コンテキストから集合提示コンテキストを選択する行為は、それぞれのテキストブロックのテキストのそれぞれのその他のテキストブロックのテキストとの比較に基づいて、テキストブロックのうちの2つ以上が互いに関連しないと決定することと、決定に応じて複数項目提示コンテキスト(multiple item presentation context)を選択することとであって、複数項目提示コンテキストが、決定に応じて別個の表示ページに各テキストブロックの言語の翻訳を提示する対応するユーザインターフェースを有する、選択することとを含み得る。

20

【0025】

対応するユーザインターフェースは、各テキストブロックに関するユーザインターフェース要素を含み得る。対応するユーザインターフェースは、特定のテキストブロックのユーザインターフェース要素が選択されるときに特定のテキストブロックに関する言語の翻訳を提示し得る。

【0026】

集合提示コンテキストから集合提示コンテキストを選択する行為は、それぞれのテキストブロックの背景色がそれぞれのその他のテキストブロックの背景色に実質的に一致すると決定することと、背景色が実質的に一致するという決定に基づいてテキストブロックが互いに関連すると決定することと、決定に応じて単一集合提示コンテキストを選択することとを含み得る。単一集合提示コンテキストは、画像上の1つまたは複数のオーバーレイ内に各テキストブロックの言語の翻訳を提示する対応するユーザインターフェースを有する可能性がある。

30

【0027】

集合提示コンテキストから集合提示コンテキストを選択する行為は、それぞれのテキストブロックのサイズおよび位置をそれぞれのその他のテキストブロックのサイズおよび位置と比較することと、比較に基づいてテキストブロックが関連すると決定することと、決定に応じて単一集合提示コンテキストを選択することとを含み得る。単一集合提示コンテキストは、画像上の1つまたは複数のオーバーレイ内に各テキストブロックの言語の翻訳を提示する対応するユーザインターフェースを有する可能性がある。

40

【0028】

テキストブロックが関連すると決定することは、テキストブロックのサイズが実質的に同様であり、隣接するテキストブロックのそれぞれの対が隣接するテキストブロックのそれぞれのその他の対と実質的に同様の距離に画像内で離間されると決定することを含み得る。

【0029】

集合提示コンテキストから集合提示コンテキストを選択する行為は、テキストブロックの第1の組の配列に基づいて同様のテキストブロックの第1の組を特定することと、テキストブロックの第2の組の配列に基づいて同様のテキストブロックの第2の組を特定することとであって、テキストブロックの第2の組が、テキストブロックの第1の組と異なる、特定す

50

ることと、テキストブロックの第1の組および第2の組を特定することに応じて複数集合提示コンテキスト(multiple collections presentation context)を選択することによって、複数集合提示コンテキストが、テキストブロックの第2の組に関する追加的な情報とは別個の同様のテキストブロックの第1の組に関する追加的な情報を提示するユーザインターフェースに対応する、選択することとを含み得る。

【0030】

態様は、テキストブロックのうちの少なくとも1つのテキストを翻訳することと、少なくとも翻訳されたテキストの文字数に基づいて翻訳されたテキストの読みやすさの尺度を決定することとをさらに含み得る。ユーザインターフェースは、読みやすさの尺度に基づいて選択され得る。

10

【0031】

態様は、画像がメニューの画像であると決定することと、各テキストブロックがメニューの特定の項目に関連付けられると決定することとをさらに含み得る。ユーザインターフェースを選択することは、特定のテキストブロック上のオーバーレイ内にそれぞれの特定のテキストブロックの翻訳を提示するユーザインターフェースを選択することを含み得る。

【0032】

態様は、各テキストブロックに関して、テキストブロックに対応する画像内の価格を特定することによって、価格が、特定の通貨で提示される、特定することと、価格を特定の通貨とは異なる通貨に変換することと、特定された価格の上または近くのオーバーレイ内に変換された価格を提示することとをさらに含み得る。

20

【0033】

本明細書に記載の対象の特定の実施形態は、以下の利点のうちの1つまたは複数を実現するように実装され得る。画像内に描かれたテキストの翻訳(または画像内で特定されたテキストに関連するその他の情報)が、画像に関連してより容易に読むことができるようにして提示され得る。テキストの翻訳を提示するためのユーザインターフェースが、翻訳がユーザに役立つようにして提示されるように、画像に関する提示コンテキストに基づいて動的に選択され得る。翻訳の読みやすさおよび明瞭さが、読みやすさの尺度を用いて、翻訳がより読みやすい形態で提示されることを可能にするユーザインターフェースを選択することによって改善され得る。複数のテキストブロックが画像内で特定されるとき、ユーザインターフェースは、ユーザが異なるテキストブロックの翻訳の間を迅速で容易にナビゲートすることを可能にし得る。

30

【0034】

本明細書に記載の対象の1つまたは複数の実施形態の詳細が、添付の図面および以下の説明に記載されている。対象のその他の特徴、態様、および利点は、説明、図面、および請求項から明らかになるであろう。

【図面の簡単な説明】

【0035】

【図1】画像内に描かれたテキストに関連する追加的な情報を提示する例示的なユーザデバイスのブロック図である。

40

【図2】画像を撮影し、画像内に描かれたテキストの言語の翻訳を提示するための例示的なユーザインターフェースのスクリーンショットである。

【図3】画像内に描かれたテキストの言語の翻訳を提示するための例示的なユーザインターフェースのスクリーンショットである。

【図4】画像内に描かれたテキストの言語の翻訳を提示するための例示的なユーザインターフェースのスクリーンショットである。

【図5】画像を撮影し、画像内に描かれたテキストの言語の翻訳を提示するための例示的なユーザインターフェースのスクリーンショットである。

【図6】画像内に描かれたテキストの言語の翻訳を提示するための例示的なユーザインターフェースのスクリーンショットである。

50

【図7】画像内に描かれたテキストの言語の翻訳を提示するための例示的なユーザインターフェースのスクリーンショットである。

【図8】画像内に描かれたテキストの言語の翻訳を提示するための例示的なユーザインターフェースのスクリーンショットである。

【図9】画像のために選択された顕著さ提示コンテキストに基づいて画像内に描かれたテキストに関連する追加的な情報を提示するための例示的なプロセスの流れ図である。

【図10】画像内に描かれたテキストの言語の翻訳を提示するための例示的なユーザインターフェースのスクリーンショットである。

【図11】画像を撮影し、画像内に描かれたテキストの言語の翻訳を提示するための例示的なユーザインターフェースのスクリーンショットである。

【図12】画像内に描かれたテキストの言語の翻訳を提示するための例示的なユーザインターフェースのスクリーンショットである。

【図13】画像のために選択された集合提示コンテキストに基づいて画像内に描かれたテキストに関連する追加的な情報を提示するための例示的なプロセスの流れ図である。

【図14】地図コンテキストを有するものとして特定された画像内に描かれたテキストの言語の翻訳を提示するための例示的なユーザインターフェースのスクリーンショットである。

【図15】画像のために選択された地図提示コンテキスト(map presentation context)に基づいて画像内に描かれたテキストに関連する追加的な情報を提示するための例示的なプロセスの流れ図である。

【図16】画像のために選択された提示コンテキストに基づいて画像内に描かれたテキストに関連する追加的な情報を提示するための例示的なプロセスの流れ図である。

【発明を実施するための形態】

【0036】

様々な図面における同様の参照番号および参照指示は、同様の要素を示す。

【0037】

概要

システムが、画像内に描かれたテキストを特定し、テキストを翻訳し、ユーザに有用である方法でテキストの翻訳を提示することができる。しかし、画像によって描かれたすべてのテキストの翻訳を単に提示することは、画像が複数の異なるテキストブロックを描くときにユーザにとって分かりにくい可能性がある。たとえば、画像は、モール内のいくつかの異なる店に関する標識を描く可能性がある。テキストの単一のブロック内に各標識の翻訳を提示するユーザインターフェースは、ユーザがどこで1つの店の名前が終わり、別の店の名前が始まるのかを特定することができない可能性があるためユーザを混乱させる可能性がある。さらに、すべての翻訳を自動的に提示することは、ユーザエクスペリエンスを損なう雑然とした表示をもたらす可能性がある。

【0038】

システムは、画像内のテキストの配列および/またはその他の視覚的特徴に基づいて翻訳をどのようにして提示すべきかを決定することができる。たとえば、システムは、どのテキストを翻訳すべきかおよび翻訳されたテキストをどのようにして提示すべきかを決定する際に、複数のテキストブロックの互いに対する近さ、テキストブロックの間の位置揃え、テキストの間の類似性(たとえば、フォントサイズ、ファミリー、スタイル、色、背景色など)、1つもしくは複数の集合内の複数のテキストブロックのグループ化、および/またはテキスト認識の信頼のレベルなどの、画像によって描かれたテキストの特徴を考慮し得る。本明細書において説明されるこれらおよびその他の特徴は、画像およびそのテキストの文脈についての洞察を与え得る。たとえば、画像の中心または焦点に対するテキストブロックの相対的な位置は、ユーザが最も関心のあるテキストを示す可能性がある。

【0039】

一部の実装において、システムは、たとえば、画像内のテキストの配列および/またはその他の視覚的特徴に基づいて複数の提示コンテキストから提示コンテキストを選択する

10

20

30

40

50

ことができる。複数の提示コンテキストは、1つもしくは複数の顕著さ提示コンテキスト、1つもしくは複数の集合提示コンテキスト、および/または1つもしくは複数の地図提示コンテキストを含み得る。顕著さ提示コンテキストは、1つまたは複数のテキストブロックが画像内で顕著に描かれるコンテキストである。たとえば、単一のテキストブロックが、画像の中心付近に配置される可能性があり、または2つの同じように顕著なテキストブロックが、画像の中心から似たような距離に位置付けられる可能性がある。集合提示コンテキストは、複数のはっきりと異なるテキストブロックが画像内に描かれるコンテキストである。たとえば、レストランメニューの画像は、それぞれのメニューの項目が画像内に描かれたそれぞれのその他のメニューの項目に関連する個々のテキストブロックであると考えられ得るので、集合提示コンテキストとして分類され得る。地図提示コンテキストは、1つもしくは複数の住所または場所を特定するその他の情報が画像内に描かれるコンテキストである。

10

【0040】

システムは、選択された提示コンテキストに基づいて、画像内に描かれたテキストに関連する追加的な情報を提示するためのユーザインターフェースを選択し得る。たとえば、システムは、画像に関する選択された提示コンテキストに基づいて、顕著さ提示コンテキストのための複数のユーザインターフェース、集合提示コンテキストのための複数のユーザインターフェース、および/または地図提示コンテキストのための複数のユーザインターフェースの間で選択を行い得る。それぞれの種類の提示コンテキストは、その種類の提示コンテキストを有するものとして特定された画像のためのユーザインターフェースを選択するための規則の異なる組または異なるプロセスを有する可能性がある。

20

【0041】

例示的なユーザデバイス

図1は、画像内に描かれたテキストに関連する追加的な情報を提示する例示的なユーザデバイス100のブロック図である。ユーザデバイス100は、デジタル写真および画像を撮影および/または提示することができる電子デバイスである。例示的なユーザデバイス100は、パーソナルコンピュータ、モバイル通信デバイス(たとえば、スマートフォン)、タブレットコンピュータ、ならびに画像を撮影および/または提示することができるその他のデバイスを含む。画像を提示するために、ユーザデバイス100は、液晶ディスプレイ(「LCD」)、タッチスクリーン、またはその他の種類のディスプレイなどのディスプレイ105を含む。この例において、ディスプレイ105は、画像107内で特定されたテキスト109の翻訳111を提示するためのユーザインターフェースを提示している。

30

【0042】

ユーザデバイス100は、ユーザが標識、メニュー、および広告掲示板などの対象の画像を撮影し、対象によって描かれたテキストの翻訳を受け取れることを可能にするトランスレータ115を含む。トランスレータ115は、画像およびその他の種類のドキュメント内のテキストを特定し得るテキスト識別器120を含む。一部の実装において、テキスト識別器120は、画像によって描かれたテキストを特定するために光学式文字認識(「OCR」)を用いて画像を分析する。テキスト識別器120は、複数の異なる言語のテキストを検出することができる。たとえば、テキスト識別器120は、複数の言語のテキストを認識することができるOCRエンジンまたは複数の異なる言語の各々のためのOCRエンジンを含む可能性がある。

40

【0043】

また、テキスト識別器120は、画像内のテキストの色、背景色、テキストサイズ、フォントの種類、および/またはテキストの位置などのテキストの特徴を検出することができる。これらの特徴は、たとえば、はっきりと異なるテキストブロックを特定するために使用され得る。たとえば、テキスト識別器120は、テキストの2つの部分が異なるフォントの色、異なる背景色を有するか、または互いに離間されている(たとえば、少なくとも閾値の距離離れている)ことに基づいて、画像内に描かれたテキストの2つの部分が2つのはっきりと異なるテキストブロックに含まれると決定する可能性がある。テキスト識別器120は、特定されたテキストに関するデータをトランスレータ115の翻訳エンジン125に送信す

50

ることができる。このデータは、テキスト自体およびテキストの特定された特徴を明記し得る。

【 0 0 4 4 】

翻訳エンジン125は、テキスト(または少なくともその一部)を異なる言語に翻訳することができる。たとえば、翻訳エンジン125は、(中国語の)テキスト109をユーザデバイス100に関連するユーザの話し言葉またはユーザによって選択された別の言語に翻訳し得る。ユーザは、以下でより詳細に説明されるように、ユーザインターフェースを用いて言語を指定し得る。この例において、テキスト109は、英語に翻訳されており、テキスト109の英語翻訳111が、ユーザに提示される。一部の実装において、翻訳エンジン125は、ユーザデバイス100によってまたはユーザデバイス100のユーザインターフェースエンジン130によって指定されたテキストの部分のみを翻訳する可能性がある。たとえば、ユーザデバイス100は、翻訳のために1つまたは複数のテキストブロックを選択する可能性がある。加えて、翻訳エンジン125は、一部の実装において、通貨の値を1つの通貨から別の通貨に翻訳する可能性がある。たとえば、翻訳エンジン125は、外国の通貨で示された金額をユーザによって指定された通貨に翻訳する可能性がある。

【 0 0 4 5 】

ユーザインターフェースエンジン130は、画像内で特定されたテキストについての追加的な情報を提示するためのユーザインターフェースを選択することができる。追加的な情報は、画像内で特定されたテキストの言語の翻訳、画像内で特定された金額(たとえば、レストランメニューの画像内で特定された価格)の通貨の翻訳、および/またはテキストに関連するその他の情報を含み得る。ユーザインターフェースは、追加的な情報がどのように提示されるかを定義することができる。たとえば、図1に示されるように、翻訳111は、画像107上のオーバーレイ内に提示される可能性がある。その他のユーザインターフェースにおいて、翻訳は、たとえば、翻訳がオーバーレイ内に存在しきれないほど多くのテキストを含むためにテキストを示すだけの別個の画面内に提示される可能性がある。ユーザインターフェースエンジン130は、画像に関して特定または選択された提示コンテキストに基づいて1組のユーザインターフェースからユーザインターフェースを選択することができる。以下で説明されるように、コンテキストセレクタ135は、画像によって描かれたテキストの配列またはその他の特徴に基づいて画像に関する提示コンテキストを選択し得る。

【 0 0 4 6 】

ユーザインターフェースインデックス140は、ユーザインターフェースとの提示コンテキストのマッピングを記憶することができる。たとえば、ユーザインターフェースインデックス140は、顕著さ提示コンテキストのためのユーザインターフェース、集合提示コンテキストのためのユーザインターフェース、および/または地図提示コンテキストのためのユーザインターフェースを特定する情報を含み得る。ユーザインターフェースエンジン130は、ユーザインターフェースインデックス140にアクセスして画像に関して選択された特定の提示コンテキストのための適切なユーザインターフェースを特定し、ユーザデバイス100上で提示するためのユーザインターフェースを生成し、ユーザインターフェースを用いて画像によって描かれたテキストに関する追加的な情報(たとえば、テキストの翻訳)を提示することができる。様々な提示コンテキストのための例示的なユーザインターフェースが、図2~図8、図10~図12、および図14に示され、以下で詳細に説明される。

【 0 0 4 7 】

ユーザインターフェースジェネレータ120が、画像内に描かれたテキストに関連する追加的な情報を提示するためのユーザインターフェースを選択するために追加的な情報に関連する読みやすさの尺度を用いる可能性もある。読みやすさの尺度は、特定のユーザインターフェースを用いて提示される場合に追加的な情報がどれだけ読みやすいかを示し得る。ユーザインターフェースに関するおよび特定の追加的な情報に関する読みやすさの尺度は、ユーザインターフェースを用いて提示され得る追加的な情報の文字の数(またはパーセント)に基づく可能性がある。たとえば、追加的な情報の10%のみを表示し得るユーザイ

10

20

30

40

50

ンターフェースは、低い読みやすさの尺度を有する可能性があり、したがって、追加的な情報を提示するために使用されない可能性がある。

【0048】

ユーザインターフェースエンジン130は、複数の好適なユーザインターフェースの各々に関する追加的な情報の読みやすさの尺度を比較することによって画像のためのユーザインターフェースを選択することができる。たとえば、画像に関する選択されたユーザインターフェースコンテキストが特定の顕著さ提示コンテキストであると仮定する。ユーザインターフェースエンジン130は、特定の顕著さコンテキストのために利用可能な各ユーザインターフェースに関連して、画像によって描かれたテキストに関連する追加的な情報に関する読みやすさの尺度を決定し得る。ユーザインターフェースエンジン130は、読みやすさの尺度に基づいて利用可能なユーザインターフェースから選択を行い得る。たとえば、ユーザインターフェースエンジン130は、最も高い読みやすさの尺度を有するユーザインターフェースを選択する可能性がある。

10

【0049】

別の例として、ユーザインターフェースエンジン130は、選択された提示コンテキストのための好ましいユーザインターフェースを特定し、好ましいユーザインターフェースに関連する追加的な情報に関する読みやすさの尺度を決定する可能性がある。たとえば、好ましいユーザインターフェースは、画像上のオーバーレイ内にテキストの翻訳を提示する可能性がある。ユーザインターフェースエンジン130は、読みやすさの尺度を好ましいユーザインターフェースに関する読みやすさの閾値と比較する可能性がある。読みやすさの尺度が読みやすさの閾値を(たとえば、読みやすさの閾値を満たすかまたは超えることによって)満足する場合、ユーザインターフェースエンジン130は、好ましいユーザインターフェースを選択し、画像上のオーバーレイ内に翻訳を提示し得る。読みやすさの尺度が読みやすさの閾値を満足しない場合、ユーザインターフェースエンジン130は、ユーザが翻訳を見るために複数の画面またはページを閲覧またはスクロールすることを可能にする別個の画面内に翻訳を提示するユーザインターフェースなどの代替的なユーザインターフェースを選択する可能性がある。

20

【0050】

コンテキストセレクトア135は、画像によって描かれたテキストの配列および/またはその他の特徴に基づいて画像に関する提示コンテキストを選択し得る。画像に関する提示コンテキストは、複数の提示コンテキストから選択される可能性がある。たとえば、提示コンテキストは、1つもしくは複数の顕著さ提示コンテキスト、1つもしくは複数の集合提示コンテキスト、1つもしくは複数の地図提示コンテキスト、および/またはその他の種類の提示コンテキストから選択される可能性がある。顕著さ提示コンテキストは、テキストの単一のブロックが描かれる画像に関する「単一ブロックコンテキスト」、特定のテキストブロックがその他のテキストブロックよりも顕著に描かれる画像に関する「支配的-副次的コンテキスト(dominant-secondary context)」、および/または2つ以上のテキストブロックが実質的に同等に顕著に描かれる(たとえば、テキストブロックが互いの閾値の量以内の顕著さスコア(prominence score)を有する)「同等顕著さコンテキスト(equal prominence context)」を含み得る。集合提示コンテキストは、複数の関連するテキストブロックが描かれる画像に関する「単一集合」コンテキスト、関連するテキストブロックの複数のグループが描かれる画像に関する「複数集合」コンテキスト、および/またはテキストブロックの翻訳が一緒に提示されるべきである画像に関する「統合」コンテキストを含み得る。

30

40

【0051】

提示コンテキストと提示コンテキストが特定の画像に関してどのようにして選択されるのかを決定するための技術との各々が、以下で詳細に説明される。特に、顕著さコンテキストが、図3~図8に関連して説明され、顕著さコンテキストを扱うためのプロセスが、図9を参照して説明され、集合コンテキストが、図10~図12に関連して説明され、集合コンテキストを扱うためのプロセスが、図13を参照して説明され、地図コンテキストが、図14

50

に関連して説明され、地図コンテキストを扱うためのプロセスが、図15を参照して説明される。最後に、画像を処理し、様々なコンテキストから選択を行い、選択されたコンテキストをそれに応じて処理するためのプロセスが、図16を参照して説明される。

#### 【0052】

例示的なユーザデバイス100がテキスト識別器120、翻訳エンジン125、ユーザインターフェイスエンジン130、コンテキストセレクトラ135、およびユーザインターフェイスインデックス140を含むものとして示されるが、これらの要素のうちの1つまたは複数は、ユーザデバイス100と通信するサーバなどの別のデバイス上に記憶されるおよび/またはその別のデバイスによって実行される可能性がある。たとえば、ユーザデバイス100は、通信ネットワーク(たとえば、LAN、WAN、またはインターネット)を介してサーバに画像を送信する可能性がある。サーバは、画像を処理して、画像のための提示コンテキストおよびユーザインターフェイスを選択し、ユーザデバイス100にデータを送信し、それが、ユーザデバイス100に選択されたユーザインターフェイスを用いて画像内に描かれたテキストに関連する追加的な情報を提示させる。

#### 【0053】

例示的な翻訳ユーザインターフェイス

図2は、画像を撮影し、画像によって描かれたテキストの言語の翻訳を提示するための例示的なユーザインターフェイスのスクリーンショット200A~200Cを示す。例示的なスクリーンショット200Aにおいては、ユーザインターフェイス210が、ユーザが翻訳のための言語を選択することを可能にする。特に、ユーザインターフェイス210は、画像内でテキストが発見されるべき元言語212とテキストが翻訳されるべき翻訳言語214とをユーザが選択することを可能にする。この例においては、ユーザは、中国語の元言語212と、英語の翻訳言語214とを選択した。したがって、この例において、ユーザは、画像内で発見される中国語テキストを英語テキストに翻訳したい。ユーザインターフェイスは、ユーザによって選択されるときにユーザデバイスに、ユーザデバイスに取り付けられたまたはユーザデバイスに通信可能なように結合されたカメラを用いて画像を撮影させるカメラボタン206も含む。

#### 【0054】

スクリーンショット200Bにおいては、ユーザインターフェイス230が、たとえば、カメラボタン206の選択に応じてユーザデバイスによって撮影された画像232を描く。例示的な画像232は、ユーザデバイスによって特定され、白い背景によって強調された中国語の文字のテキストのいくつかの部分を含む。特に、画像232は、画像232の左上の角付近にあるテキストの第1の部分234と、画像232の中心付近にあるテキストの第2の部分236と、画像232の右下付近にあるテキストの第3の部分238とを含む。テキストの各部分234~238は、画像232内の別々の標識の上に描かれる。

#### 【0055】

翻訳画面ユーザインターフェイス250が、スクリーンショット200C内に描かれる。翻訳画面ユーザインターフェイス250は、画像232内で特定されたテキスト252と、画像232内で特定されたテキスト252の英語翻訳254とを描く。テキストの第1の部分234は、単語「Free」に翻訳され、テキストの第2の部分236は、「Big River Inn 2nd Floor」に翻訳され、テキストの第3の部分238は、単語「Last」に翻訳された。この例示的なユーザインターフェイス250において、テキスト252は、別個のはっきりと異なるテキストブロック内に描かれていたが、単一のテキストブロックに含まれているかのように翻訳された。翻訳されたテキスト254によって分かるように、複数のはっきりと異なるテキストブロックを1つにまとめて翻訳することは、ユーザを混乱させるかまたはいら立たせる可能性がある無意味な翻訳をもたらす可能性がある。画像によって描かれたテキストの配列および/または視覚的特徴を評価し、画像によって描かれたテキストに関するコンテキストを選択することによって、より意味のある翻訳が提供され得る。たとえば、図3~図5は、テキストの各部分の顕著さの分析に基づいてテキスト252の翻訳をよりうまく提示し得るユーザインターフェイスを示す。

10

20

30

40

50

## 【 0 0 5 6 】

## 顕著さ提示コンテキストおよびユーザインターフェース

画像内のテキストの顕著さは、ユーザが関心のある対象にそれらのユーザがそれらのユーザのカメラの焦点を合わせることが多いので、どのテキストが翻訳(またはその他の情報)を提供すべきかを決定するために使用され得る。テキストの顕著さを評価することによって、トランスレータ115は、どのテキストに関してユーザが翻訳を得たい可能性が高いかを決定し、そのテキストの翻訳を提供することができる。たとえば、ユーザは、標識の翻訳を得るために特定の標識のテキストの写真を撮る可能性がある。写真において、特定の標識は写真の中心付近にあるが、その他の標識が背景内にある可能性がある。テキストの顕著さを評価することによって、トランスレータは、翻訳のための特定の標識のテキストが背景の標識のテキストよりも顕著に描かれるので、このテキストを特定することができる。また、トランスレータ115は、ユーザが背景の標識の翻訳に関心がない可能性があるもので、それらの標識のテキストを翻訳しないことを決定する可能性がある。

10

## 【 0 0 5 7 】

図3は、画像内に描かれたテキストの言語の翻訳を提示するための例示的なユーザインターフェースのスクリーンショット300Aおよび300Bを示す。スクリーンショット300Aは、スクリーンショット200Bと同様であり、ユーザデバイスによって特定され、白い背景によって強調された中国語の文字のテキストのいくつかの部分334~338を含む画像332を提示するユーザインターフェース330を含む。

## 【 0 0 5 8 】

翻訳画面ユーザインターフェース350が、スクリーンショット300B内に描かれる。画像332内に描かれたテキストのすべての翻訳をその翻訳が1つのテキストブロックに含まれるかのように提示するのではなく、例示的な翻訳画面ユーザインターフェース350は、画像332内で特定されたテキストの部分352およびテキストの部分352の英語翻訳354だけを提示する。特に、翻訳画面ユーザインターフェース350は、テキストの第2の部分336およびテキストの第2の部分336の英語翻訳354(「Big River Inn 2nd Floor」)を提示する。

20

## 【 0 0 5 9 】

翻訳画面ユーザインターフェース350は、ユーザが追加的な情報を閲覧することを可能にするコントロール359も含む。コントロール359は、テキストの各部分334~336に関する翻訳を含む画面の間をナビゲートするか、または各翻訳に関するその他の情報を得るように構成され得る。たとえば、コントロール359の選択に応じて、ユーザデバイスは、テキストの第3の部分338の翻訳を提示する別の画面を表示する可能性がある。コントロール359の別の選択は、ユーザデバイスにテキストの第1の部分334の翻訳を提示する画面を表示させる可能性がある。コントロール359のさらに別の選択は、ユーザデバイスを翻訳354を提示する現在の画面に戻す可能性がある。

30

## 【 0 0 6 0 】

この例においては、図1のトランスレータ115などのトランスレータが、テキストの第1の部分334の顕著さおよびテキストの第3の部分338の顕著さと比べたテキストの第2の部分336の顕著さに基づいて翻訳のためにテキストの第2の部分を選択した。たとえば、トランスレータ115は、テキストの部分の相対的な位置およびテキストの部分の向きに基づいてテキストの各部分がはっきりと異なるテキストブロックに属することを特定し得る。この決定に基づいて、トランスレータ115は、提示コンテキストが顕著さ提示コンテキストであると決定し、テキストの各部分の顕著さを評価し得る。テキストの第2の部分は画像の中心付近にテキストのその他の2つの部分よりも大きなフォントで表示されるので、トランスレータ115は、テキストの第2の部分336がテキストのその他の2つの部分よりも顕著に提示されていると決定し、決定に応じてテキストの第2の部分336を自動的に翻訳する可能性がある。

40

## 【 0 0 6 1 】

概して、ユーザは所望の内容が撮影される画像の中心にまたは中心付近にあるようにそれらのユーザのカメラの焦点を合わせるので、最も顕著なテキストのみの翻訳を提示する

50

ことによって、ユーザデバイスは、ユーザによって探し求められる翻訳をユーザに提供する可能性がより高い。また、これは、背景にあり、ユーザによって意図せずに撮影される可能性があるテキストの翻訳によってユーザを混乱させることを防止する。たとえば、テキストの第1の部分および第3の部分334および338は、ユーザが翻訳のためにテキストの第2の部分336の画像を撮影しようとして試みている間にユーザによって意図せずに撮影された可能性がある。

【0062】

より顕著なテキストを自動的に翻訳することに加えて、トランスレータ115は、画像に関する提示コンテキストを選択し、選択された提示コンテキストに基づいて翻訳を提示するためのユーザインターフェースを選択する可能性もある。この例において、トランスレータ115は、テキストの第2の部分336を含むより顕著なテキストブロック(または支配的なテキストブロック)と、テキストの第2の部分336よりも顕著でないように画像332内で描かれるテキストの第1の部分および第2の部分334および338を含む2つの副次的なテキストブロックとを画像332が含むと決定したので、画像332に関する支配的-副次的コンテキストを選択し得る。

10

【0063】

支配的-副次的コンテキストのためのユーザインターフェースは、たとえば、画像内に描かれた最も顕著なテキストの翻訳を提示する翻訳画面ユーザインターフェース(たとえば、ユーザインターフェース350)、副次的なテキストブロックの翻訳を提示することなく最も顕著なテキストの翻訳を含むオーバーレイを画像上に提示するユーザインターフェース、最も顕著なテキストブロックの翻訳を含むオーバーレイと、選択されたときに副次的なテキストブロックの翻訳を提示するユーザインターフェース要素とを画像上に提示するユーザインターフェース、および/またはオーバーレイの対応するテキストブロックの翻訳を提示する各テキストブロックのためのオーバーレイを画像上に提示するユーザインターフェースを含み得る。

20

【0064】

たとえば、図4は、画像上のオーバーレイ内にテキストブロックの言語の翻訳を提示するための例示的なユーザインターフェースのスクリーンショット400Aおよび400Bを示す。スクリーンショット400Aは、スクリーンショット200Bと同様であり、画像432の左上の角付近にあるテキストの第1の部分434と、画像432の中心付近にあるテキストの第2の部分436と、画像432の右下付近にあるテキストの第3の部分438とを含む画像432を提示するユーザインターフェース430を含む。画像432なしに別個の画面にテキストの第2の部分の翻訳を提示するのではなく、スクリーンショット400B内に描かれたユーザインターフェース450は、画像432と、画像432によって描かれたテキストの翻訳をそれぞれが含む画像432上のオーバーレイ454~458とを提示する。特に、オーバーレイ454は、テキストの第1の部分434の翻訳を提示し、オーバーレイ456は、テキストの第2の部分436の翻訳を提示し、オーバーレイ458は、テキストの第3の部分438の翻訳を提示する。この例において、オーバーレイ454~458は、オーバーレイが対応するテキストの上に(またはそのテキストに代わって)それぞれが置かれる。これは、ユーザが翻訳を用いて画像432内の各標識の対象を決定することを容易にする。たとえば、画像432に描かれた通りを歩くユーザは、容易にユーザインターフェース450を見て、Big River Innが通りのどこにあるかを決定し得る。

30

40

【0065】

各オーバーレイ454~458は、ユーザが翻訳されたテキストに関連する追加的な情報を得ることを可能にする選択可能なアイコンを含み得るかまたはそのアイコンに関連付けられ得る。たとえば、スクリーンショット400Cは、ユーザが元の言語かまたは翻訳された言語かのどちらかのテキストの発音を聞くことを可能にするユーザインターフェース470を示す。テキストに含まれる単語の定義またはテキストによって特定された主体(entity)に関連する情報などのテキストに関連するその他の情報も、提示される可能性がある。ユーザインターフェース470は、ユーザがオーバーレイ454に関連するユーザインターフェース要素を選択することに応じて提示される可能性がある。

50

## 【 0 0 6 6 】

一部の実装において、支配的-副次的コンテキストのためのユーザインターフェースは、初め、より顕著でないテキストブロックに関するオーバーレイでなく、最も顕著なテキストに関するオーバーレイのみを含む可能性がある。そのようなユーザインターフェースは、ユーザが画像内に描かれた副次的なテキストブロックの翻訳を要求することができるユーザインターフェース要素を提供する可能性がある。たとえば、図5は、ユーザが翻訳のためにテキストを選択することを可能にする例示的なユーザインターフェースのスクリーンショット500A~500Dを示す。

## 【 0 0 6 7 】

スクリーンショット500Aにおいては、ユーザインターフェース510が、選択されたときにユーザデバイスに画像512を撮影させるカメラボタンを含む。例示的な画像512は、日本語のテキストを含む2つのはっきりと異なるテキストブロック514および516を含む。スクリーンショット500Bにおいては、ユーザインターフェース520が、画像512上の単一のテキストブロック523内に両方のテキストブロック514と516とのテキストを提示する。また、ユーザインターフェース520は、画像512上のテキストブロック525内に両方のテキストブロック514および516の翻訳を提示する。図1のトランスレータ115などのトランスレータは、テキストブロック514および516を特定することに応じてユーザインターフェース520のために自動的に2つのテキストブロックのテキストを翻訳する可能性がある。代替的に、トランスレータ115は、テキストブロック514のテキストがより大きなフォントでより顕著に表示されるのでそのテキストのみを翻訳する可能性がある。あるいは、トランスレータ115は、両方のテキストブロック514と516とが画像の角にあるのでどちらのテキストブロックも顕著に表示されないと決定することに応じてテキストブロック514または516のどちらも翻訳しない可能性がある。画像によって描かれたテキストブロックの顕著さを決定するためのプロセスが、図9を参照して以下で詳細に説明される。

## 【 0 0 6 8 】

スクリーンショット500Cにおいては、ユーザインターフェース530が、ユーザが翻訳のためにテキストを選択するための命令533を提供する。この例において、ユーザは、翻訳のためにテキストブロック516内のテキストを強調した。それに応じて、トランスレータ115は、選択されたテキストに関する翻訳を特定し、オーバーレイ内に翻訳を提供することができる。たとえば、スクリーンショット500Dは、テキストのユーザの選択に応じて画像512上のオーバーレイ549内に翻訳されたテキストを提示するユーザインターフェース540を示す。ユーザは、翻訳されたテキストに関連する追加的な情報を要求するためにオーバーレイ549を選択することもできる。

## 【 0 0 6 9 】

一部の実装においては、ユーザインターフェースが、ユーザが翻訳したいテキストにドラッグされ得るユーザインターフェース要素を提供する可能性がある。たとえば、ユーザインターフェース要素が、タッチスクリーンデバイスに表示される可能性があり、ユーザが、所望のテキストにユーザインターフェース要素をドラッグする可能性がある。それに応じて、ユーザデバイスは、トランスレータ115にテキストの翻訳を要求し、トランスレータ115によって選択されたユーザインターフェースを用いて翻訳を提示する可能性がある。

## 【 0 0 7 0 】

図6は、顕著なテキストに関する翻訳を提供し、ユーザが翻訳のために副次的なテキストを選択することを可能にする例示的なユーザインターフェースのスクリーンショット600Aおよび600Bを示す。スクリーンショット600Aにおいては、ユーザインターフェース610が、2つの道路標識617および619の画像615を示す。各道路標識617および619は、中国語の文字のテキストを描く。画像615内のテキストを中国語の文字から英語に翻訳するために、トランスレータ115は、中国語の文字を含む画像内のテキストブロックを特定し得る。たとえば、トランスレータ115は、道路標識617内の中国語の文字をテキストブロック620として特定し、道路標識619内の中国語の文字をテキストブロック625として特定し得る。

## 【 0 0 7 1 】

どのテキストを翻訳すべきかおよび翻訳されたテキストをどのようにして提示すべきかを決定するために、トランスレータ115は、それぞれの特定されたテキストブロック620および625の顕著さを評価し得る。この例において、トランスレータ115は、テキストブロック620がテキストブロック625よりも画像615の中心に近いのでテキストブロック620がテキストブロック625よりも顕著に描かれると決定する可能性がある。加えて、テキストブロック620に関するテキストサイズは、テキストブロック625に関するテキストサイズよりも大きい。

## 【 0 0 7 2 】

テキストブロック620がテキストブロック625よりも顕著に描かれると決定するのに応じて、トランスレータ115は、テキストブロック625のテキストを自動的に翻訳しない一方、テキストブロック620のテキストを自動的に翻訳することを決定する可能性がある。スクリーンショット600Bに示されるように、トランスレータ115は、画像615内のテキストブロック620の中国語の文字の上の(またはそれらの中国語の文字の代わりに)オーバーレイ643内にテキストブロック620の翻訳を提示するユーザインターフェース640を選択する可能性がある。

10

## 【 0 0 7 3 】

この例においては、完全な翻訳が、オーバーレイ643に収まらない。ユーザインターフェース640は、選択されたときにユーザインターフェース640に完全な翻訳を提示させるユーザインターフェース要素647を提供し得る。一部の実装において、トランスレータ115は、翻訳に関する読みやすさの尺度が閾値を満たさないと決定し、それに応じて、部分的な翻訳を提示するのではなく別の画面に翻訳を提示する可能性がある。

20

## 【 0 0 7 4 】

ユーザインターフェース640は、自動的に翻訳されなかった副次的なテキストに、強調されたブロックなどのユーザインターフェース要素を提供する可能性もある。たとえば、ユーザインターフェース640は、翻訳されないテキストブロック625を強調するためのユーザインターフェース要素645を含む。ユーザインターフェース要素645のユーザの選択に応じて、ユーザインターフェース640は、テキストブロック625に含まれるテキストの翻訳を提示し得る。この翻訳は、道路標識619もしくはテキストブロック625上のオーバーレイ内または別個の画面もしくはユーザインターフェース内に提示される可能性がある。

30

## 【 0 0 7 5 】

いくつかの画像が、テキストの単一のブロックを顕著に描く可能性がある。たとえば、ユーザは、命令または標識の翻訳を得るためにユーザデバイスを用いて命令または単一の標識の画像を撮影し得る。そのような場合、トランスレータ115は、テキストを特定し、特定のユーザインターフェース内にテキストの翻訳を自動的に提示し得る。たとえば、トランスレータ115は、画像が単一の顕著に表示されるテキストブロックを描くと決定することに応じて画像に関する単一ブロックコンテキストを選択する可能性がある。単一ブロックコンテキストのためのユーザインターフェースは、画像上のオーバーレイ内に単一のテキストブロックに含まれるテキストの翻訳を提示するユーザインターフェース、ユーザが翻訳されたテキストを閲覧もしくはスクロールすることを可能にする画面などの別個の画面内に単一のテキストブロックに含まれるテキストの翻訳を画像なしに提示するユーザインターフェース、および/またはその他の種類のユーザインターフェースを含み得る。

40

## 【 0 0 7 6 】

図7は、画像内に描かれたテキストの言語の翻訳を提示するための例示的なユーザインターフェースのスクリーンショット700Aおよび700Bを示す。スクリーンショット700Aは、画像715を提示するユーザインターフェース710を示す。画像715は、中国語の文字を有するテキストブロック720および英語の文字を有するテキストブロック725を描く。この例において、ユーザは、中国語から英語への翻訳を要求した。したがって、トランスレータ115は、画像を分析して任意の中国語の文字を特定し、テキストブロック720に含まれるテキストを特定する一方、テキストブロック725内の英語の文字を無視する可能性がある。さ

50

らに、トランスレータ115は、テキストブロック720に含まれるテキストを分析して、テキストブロック720が互いに関連するテキストを含む単一のテキストブロックであると決定する可能性がある。たとえば、トランスレータ115は、テキストサイズ、テキストの色、テキストに関する背景色、テキストの位置および近さ、ならびに/またはその他の特徴を比較して、テキストブロック720に含まれるテキストが関連すると決定する可能性がある。それに応じて、トランスレータ115は、画像に関する単一ブロックコンテキストを選択し、選択に基づいてユーザインターフェースを特定する可能性がある。

【0077】

スクリーンショット700Bに示されるように、トランスレータ115は、画像715上のオーバーレイ760内にテキストブロック720の翻訳を提示するユーザインターフェース750を選択した。ユーザインターフェース750は、ユーザによって選択されたときに翻訳されたテキストを一度に1行ずつ提示するユーザインターフェース要素755も含む。たとえば、ユーザは、翻訳されたテキストの各行を閲覧できる可能性があり、各行は、それぞれのその他の行とは別個の表示画面内に提示される。

10

【0078】

一部の画像は、等しいまたは同様の顕著さでテキストの2つ以上のブロックを描く可能性がある。たとえば、ユーザは、どの方向またはどのドアを通り抜けるべきかを決定するために2つの標識の画像を撮影する可能性がある。そのような場合、ユーザを案内するのに助けるために両方の標識の翻訳を提示することが有益である可能性がある。

【0079】

20

図8は、同様の顕著さを有するテキストブロックの言語の翻訳を提示するための例示的なユーザインターフェースのスクリーンショット800A~800Cを示す。スクリーンショット800Aにおいては、ユーザインターフェース810が、中国語の文字を含む2つの標識820および830を含む画像815を提示する。この例において、ユーザは、中国語の文字の英語への翻訳を要求している。したがって、トランスレータ115は、標識820に含まれる中国語の文字および標識830に含まれる中国語の文字を特定する可能性がある。また、トランスレータ115は、各標識820および830に含まれる中国語の文字を分析して、各標識の中国語の文字がはっきりと異なるテキストブロックに含まれると決定する可能性がある。たとえば、トランスレータ115は、標識820に含まれる中国語の文字をテキストブロック825として特定し、標識830に含まれる中国語の文字をテキストブロック825とは異なるテキストブロック835として特定する可能性がある。

30

【0080】

画像815内の複数のテキストブロック825および835を特定することに応じて、トランスレータ115は、テキストブロックの顕著さを評価し、各テキストブロック825および835のテキストサイズおよび色が同様であり、テキストブロック825および835が画像815の中心からほぼ同じ距離にあるのでテキストブロック825および835が同様の顕著さを有すると決定し得る。テキストブロック825および835が同様の顕著さを有するので、トランスレータ115は、画像815に関する同等顕著さコンテキストを選択し、同等顕著さコンテキストに対応するユーザインターフェースを選択し得る。

【0081】

40

スクリーンショット800Bに示されるように、ユーザインターフェース840が、テキストブロックの翻訳を提示するように選択された。ユーザインターフェース840は、テキストブロック825の中国語の文字が特定された標識820上に位置付けられるオーバーレイ845内にテキストブロック825の中国語の文字の翻訳を提示する。同様に、ユーザインターフェース840は、テキストブロック835の中国語の文字が特定された標識830上に位置付けられるオーバーレイ847内にテキストブロック835の中国語の文字の翻訳を提示する。各オーバーレイ845および847は、オーバーレイの対応するテキストの完全な翻訳を要求するためにやはり選択可能である。たとえば、ユーザがオーバーレイ845を選択する場合、スクリーンショット800Cに示されるように、「toilet」に当たる中国語の文字の完全な翻訳がユーザインターフェース860内に提示される。

50

## 【 0 0 8 2 】

トランスレータ115などのトランスレータは、画像によって描かれたテキストの配列および視覚的特徴を評価して、画像に関する顕著さ提示コンテキストを選択し得る。トランスレータ115は、特定されたテキストの少なくとも一部に関する追加的な情報を提示するためのユーザインターフェースを選択する可能性もある。ユーザインターフェースは、選択された顕著さ提示コンテキストに基づいて選択され得る。たとえば、単一ブロックコンテキストが選択される場合、単一ブロックコンテキストに対応するユーザインターフェースが追加的な情報を提示するために選択され、使用され得る。

## 【 0 0 8 3 】

図9は、画像のために特定された顕著さ提示コンテキストに基づいて画像内に描かれたテキストに関連する追加的な情報を提示するための例示的なプロセス900の流れ図である。プロセス900の動作は、たとえば、図1のユーザデバイス100および/またはトランスレータ115などのデータ処理装置によって実装される可能性がある。プロセス900は、コンピュータストレージ媒体に記憶された命令によって実装される可能性もあり、データ処理装置による命令の実行が、データ処理装置にこのプロセス900の動作を実行させる。

## 【 0 0 8 4 】

画像が受信される(902)。たとえば、ユーザが、スマートフォン、タブレットコンピュータ、またはその他のユーザデバイスなどのユーザデバイスを用いて対象の写真を撮る可能性がある。標識が、受信された画像内にやはり描かれるユーザにとって異質な言語のテキストを含む可能性がある。

## 【 0 0 8 5 】

受信された画像内のテキストが特定される(904)。テキストは、OCRまたは別のテキスト特定技術を用いて特定され得る。一部の实装においては、特定の言語のテキストのみが特定される。たとえば、ユーザが、特定の言語のテキストがユーザによって理解される言語などの別の言語に翻訳されるべきであることを指定し得る。この例においては、特定の言語のテキストのみが画像内で特定され得る。

## 【 0 0 8 6 】

実際のテキストに加えて、テキストの特徴(たとえば、視覚的特徴)が特定される可能性がある。これらの特徴は、テキストの色、テキストの下の背景の色、テキストのサイズ、テキストに関するフォントの種類、および/または画像内のテキストの位置を含み得る。

## 【 0 0 8 7 】

一部の实装においては、切れているかまたはそれ以外の点で不完全なテキストは無視される可能性がある。たとえば、標識の一部のみが画像によって描かれ、標識のテキストが「stop」ではなく「sto」などの不完全な単語を含む場合、標識のテキストは、追加的な情報を提供するため、または画像に関する顕著さ提示コンテキストを選択するために特定されない可能性がある。不完全なテキストを翻訳することは、ユーザを混乱させるかまたは無意味な翻訳をもたらす可能性がある。

## 【 0 0 8 8 】

1つまたは複数のはっきりと異なるテキストブロックが、特定されたテキストに基づいて特定される(906)。テキストの特徴が、はっきりと異なるテキストブロックを特定するために評価され得る。たとえば、その他のテキストに近く(たとえば、閾値の距離以内にある)、同様の見た目(たとえば、同様のサイズ、テキストサイズ、および/または背景色)を有するテキストが、テキストブロックAの一部であると考えられる可能性がある。反対に、テキストブロックAから離れており(たとえば、閾値の距離を超えている)、テキストブロックAのテキストとは異なるテキストサイズ、異なるテキストの色、および/または異なる背景色を有するテキストの別の部分は、別のはっきりと異なるテキストブロックBとして特定される可能性がある。テキストの2つの部分の視覚的特徴(たとえば、テキストの色、テキストサイズ、フォントの種類、および/または背景色)の類似性ならびにテキストの2つの部分の間の距離が、テキストの2つの部分が共通のテキストブロックに属するかどうかを決定するために使用され得る。

## 【0089】

画像に関する顕著さ提示コンテキストが、特定されたテキストの顕著さに基づいて選択される(908)。顕著さ提示コンテキストは、複数の顕著さ提示コンテキストから選択される可能性がある。たとえば、複数の顕著さ提示コンテキストは、単一ブロックコンテキスト、支配的-副次的コンテキスト、および/または同等顕著さコンテキストを含み得る。

## 【0090】

1つまたは複数のテキストブロックの顕著さを評価するために、各テキストブロックに関して顕著さスコアが計算される可能性がある。テキストブロックに関する顕著さスコアは、画像内のテキストブロックの位置、テキストブロックに含まれるテキストのサイズ、テキストブロックに含まれるテキストの文字数、テキストブロックに含まれるテキストの色、テキストブロックに関する背景の色、および/またはテキストブロックの顕著さを示し得るその他の因子に基づく可能性がある。位置は、(たとえば、座標系を用いる)画像内の絶対的な位置、または画像の中心からテキストブロックの点(たとえば、テキストブロックの中心、もしくは画像の中心に最も近いテキストブロックの点)までの距離である可能性がある。たとえば、画像の中心により近いテキストブロックは、画像の中心からより遠いテキストブロックよりも高い顕著さスコアを有する可能性がある。同様に、より大きなテキストサイズ(または多くの文字)を有するテキストブロックは、より小さなテキストサイズ(またはより少ない文字)を有するテキストブロックよりも高い顕著さスコアを有する可能性がある。加えて、より明るいテキストまたは背景の色は、より暗いまたはよりぼんやりしたテキストまたは背景の色よりも高い顕著さスコアをもたらす可能性がある。

## 【0091】

単一のはっきりと異なるテキストブロックが画像に関して特定される場合、単一ブロックコンテキストが画像に関して選択され得る。一部の実装において、単一のはっきりと異なるテキストブロックの顕著さが、画像に関して単一ブロックコンテキストを選択する前に評価される可能性がある。たとえば、単一のはっきりと異なるテキストブロックは、画像の角のテキストの小さなブロックである場合、低い顕著さスコア(たとえば、顕著さスコアの閾値未満である顕著さスコア)を有する可能性がある。そのような場合、単一ブロックコンテキストは選択されない可能性があり、テキストは自動的に翻訳されない可能性がある。

## 【0092】

複数のはっきりと異なるテキストブロックが画像に関して特定される場合、支配的-副次的コンテキストを選択すべきかまたは同等顕著さコンテキストを選択すべきかを決定するために、テキストブロックの相対的な顕著さが評価される可能性がある。たとえば、それぞれのはっきりと異なるテキストブロックに関して顕著さスコアが計算される可能性があり、顕著さスコアが比較される可能性がある。

## 【0093】

テキストブロックが画像に関して特定されたそれぞれのその他のテキストブロックよりも著しく高い顕著さスコアを有する場合、支配的-副次的コンテキストが画像のために選択され得る。たとえば、テキストブロックAがそれぞれのその他のテキストブロックよりも少なくとも閾値の量大きい顕著さスコアを有する場合、テキストブロックAは支配的なテキストブロックと考えられる可能性があり、支配的-副次的コンテキストが画像のために選択される可能性がある。画像に関して特定された2つ以上のテキストブロックが同様の顕著さを有する場合、同等顕著さコンテキストが画像のために選択され得る。たとえば、テキストブロックBがテキストブロックAに関する顕著さスコアの閾値の量以内の顕著さスコアを有し、テキストブロックAおよびBに関する顕著さスコアが画像に関して特定されたいずれのその他のテキストブロックに関する顕著さスコアよりも大きい場合、同等顕著さコンテキストが画像のために選択され得る。

## 【0094】

追加的な情報が提示されるべきテキストが、選択された顕著さ提示コンテキストに基づいて選択される(910)。たとえば、単一ブロックコンテキストが画像のために選択される

10

20

30

40

50

場合、画像内で特定された単一のはっきりと異なるテキストブロックに含まれるテキストが、追加的な情報を提供するために選択され得る。この追加的な情報は、単一のはっきりと異なるテキストブロックに含まれるテキストの言語の翻訳である可能性がある。支配的-副次的コンテキストが画像のために選択される場合、支配的なテキストブロック(たとえば、最も大きな顕著さを有するテキストブロック)に含まれるテキストが、追加的な情報を提供するために選択され得る。同等顕著さコンテキストが画像のために選択される場合、同様の顕著さを有するテキストブロックに含まれるテキストが、追加的な情報を提供するために選択され得る。

【0095】

追加的な情報が、選択されたテキストに関して特定される(912)。追加的な情報は、選択されたテキストの言語の翻訳を含む可能性がある。たとえば、ユーザが、そのユーザが画像によって描かれたテキストが翻訳されることを望む言語を選択する可能性がある。通貨の金額を含むテキストに関して、追加的な情報は、ユーザによって指定された通貨への金額の変換を含む可能性がある。たとえば、選択されたテキストが日本円の価格を含む場合、追加的な情報は、円の価格に対応するドルの金額を含む可能性がある。

10

【0096】

ユーザインターフェイスが、選択された顕著さ提示コンテキストおよび/または追加的な情報に関する読みやすさの尺度に基づいて選択される(914)。各顕著さ提示コンテキストは、画像内に描かれたテキストに関する追加的な情報を提示するために使用され得る(または使用される資格がある)1つまたは複数の対応するユーザインターフェイスを有する可能性がある。加えて、1つの顕著さ提示コンテキストにふさわしいユーザインターフェイスは、それぞれのその他の顕著さ提示コンテキストのためのユーザインターフェイスとは異なる可能性がある。

20

【0097】

単一ブロックコンテキストは、画像上のオーバーレイ内に追加的な情報を提示するユーザインターフェイス(たとえば、図7のユーザインターフェイス)および/またはユーザが1つもしくは複数の別個の画面内の追加的な情報を閲覧することを可能にするユーザインターフェイス(たとえば、図3のユーザインターフェイス)を有する可能性がある。それぞれの別個の画面は、ユーザがいっそう多くの追加的な情報を見るための追加的な画面を閲覧することを可能にするコントロールを含み得る。

30

【0098】

支配的-副次的コンテキストは、画像内に描かれた最も顕著なテキストブロックに関する追加的な情報を提示するユーザインターフェイス(たとえば、図3のユーザインターフェイス)、副次的なテキストブロックに関する追加的な情報を提示することなく最も顕著なテキストに関連する追加的な情報を含むオーバーレイを画像上に提示するユーザインターフェイス、最も顕著なテキストブロックに関連する追加的な情報を含むオーバーレイ、および選択されたときに副次的なテキストブロックに関連する追加的な情報を提示するユーザインターフェイス要素を画像上に提示するユーザインターフェイス(たとえば、図6のユーザインターフェイス)、ならびに/またはオーバーレイの対応するテキストブロックに関連する追加的な情報を提示する各テキストブロックのためのオーバーレイを画像上に提示するユーザインターフェイス(たとえば、図4のユーザインターフェイス)を含み得る。

40

【0099】

同等顕著さコンテキストは、同等の顕著さを有するものとして特定されたテキストブロックの各々に関するオーバーレイ内に追加的な情報を提示するユーザインターフェイス(たとえば、図7のユーザインターフェイス)ならびに/またはユーザが1つもしくは複数の別個の画面内の同等顕著さテキストブロック(equal prominence text block)(および画像内で特定された任意の他のテキストブロック)に関する追加的な情報を閲覧することを可能にするユーザインターフェイス(たとえば、図3のユーザインターフェイス)を含み得る。

【0100】

各顕著さ提示コンテキストは、そのコンテキストのための好ましいユーザインターフェ

50

ースを有する可能性がある。たとえば、好ましいユーザインターフェースは、画像上のオーバーレイ内に追加的な情報を提示するユーザインターフェースである可能性がある。別の例として、ユーザは、各顕著さ提示コンテキストのための好ましいユーザインターフェースを選択することができる可能性がある。特定の顕著さ提示コンテキストのための好ましいユーザインターフェースは、特定の顕著さ提示コンテキストの選択に応じて選択される可能性がある。

#### 【0101】

一部の実装においては、追加的な情報に関する読みやすさの尺度が、追加的な情報を提示するためのユーザインターフェースを選択するために使用される。上述のように、読みやすさの尺度は、特定のユーザインターフェースを用いてユーザに提示される場合に追加的な情報がどれだけ読みやすいかを示し得る。ユーザインターフェースに関するおよび特定の追加的な情報に関する読みやすさの尺度は、ユーザインターフェースを用いて提示され得る追加的な情報の文字数に基づく可能性がある。読みやすさの尺度は、(a)ユーザインターフェースを用いて提示され得る追加的な情報の文字数と、(b)追加的な情報内の総文字数との間の比に等しいかまたは比例する可能性がある。たとえば、オーバーレイが、追加的な情報を提示するための限られた量の空間を有する可能性がある。空間の量は、オーバーレイが追加的な情報を提示すべきであるテキストブロックのサイズに基づく可能性がある。翻訳の一部のみがオーバーレイ内に提示される場合、それはユーザの混乱につながる可能性がある。そのような場合、翻訳のすべてまたは少なくともより多くが提示され得る別個の画面内に追加的な情報を提示する方がよい可能性がある。

#### 【0102】

読みやすさの尺度が、選択された顕著さ提示コンテキストのための好ましいユーザインターフェース(たとえば、画像上のオーバーレイを有するユーザインターフェース)を使用すべきかまたは代替的なユーザインターフェース(たとえば、ユーザが追加的な情報を閲覧することを可能にするユーザインターフェース)を使用すべきかを決定するために使用され得る。たとえば、好ましい顕著さ提示コンテキストに関連する追加的な情報に関する読みやすさの尺度が読みやすさの閾値を(たとえば、読みやすさの閾値を満たすかまたは超えることによって)満足する場合、好ましいユーザインターフェースが選択される可能性がある。好ましい顕著さ提示コンテキストに関連する追加的な情報に関する読みやすさの尺度が読みやすさの閾値を満足しない(たとえば、読みやすさの尺度が読みやすさの閾値未満である)場合、追加的な情報のより多くの文字の提示を可能にする代替的なユーザインターフェースが選択される可能性がある。

#### 【0103】

読みやすさの尺度は、好ましいユーザインターフェースとは独立した複数のユーザインターフェースの間で選択を行うために使用される可能性がある。たとえば、選択された顕著さ提示コンテキストにふさわしい、追加的な情報に関する最も高い関連する読みやすさの尺度を有するユーザインターフェースが、追加的な情報を提示するために選択され得る。

#### 【0104】

追加的な情報が、選択されたユーザインターフェースを用いて提示される(916)。たとえば、ユーザデバイスは、ユーザインターフェースを生成し、選択されたユーザインターフェースを用いて追加的な情報を提示する可能性がある。

#### 【0105】

集合提示コンテキストおよびユーザインターフェース

一部の画像は、画像内に様々なレベルの顕著さを有するテキストの多くのはっきりと異なる部分を有する可能性がある。たとえば、レストランメニューの画像が、画像内に描かれたそれぞれのメニューの項目に関するテキストのリストを含み得る。この文脈では、関連するテキストブロックの集合が特定され、各集合の翻訳(またはその他の追加的な情報)と一緒に提示される場合、それはユーザにとってより有益である可能性がある。たとえば、ユーザは、画像の中心に近いメニューの項目のみではなく、どの項目を注文すべきかを

10

20

30

40

50

決定するためにそれぞれのメニューの項目の翻訳を見たい可能性がある。同時に、図2のユーザインターフェース250と同様に、メニューの項目をテキストの1つのブロックとして扱うことは、ユーザを混乱させる可能性がある。図1のトランスレータ115などのトランスレータが、図10～図13を参照してより詳細に説明されるように、画像内の複数のテキストブロックの配列および/または視覚的特徴を評価して、テキストブロックの翻訳をユーザにどのようにして提示すべきかを決定し得る。

【0106】

図10は、画像を撮影し、画像内に描かれたテキストの言語の翻訳を提示するための例示的なユーザインターフェースのスクリーンショット1000A～1000Cを示す。スクリーンショット1000Aにおいては、ユーザインターフェース1010が、たとえば、画像撮影ボタン1012のユーザの選択に応じて撮影されたメニューの画像1015を示す。画像1010は、中国語のテキストのレストランの名前1020およびメニューの項目1022のリストを描く。画像1010は、それぞれのメニューの項目1022の価格1024も描く。

10

【0107】

ユーザインターフェース1010の最上部付近に示されるように、ユーザは、中国語のテキストを英語のテキストに翻訳することを選択した。トランスレータ115は、画像1015内の中国語のテキストを特定し、テキストの配列および提示に基づいて(たとえば、テキストのサイズ、テキストの色、背景色、およびテキストの部分の間の距離に基づいて)テキストのどの部分がはっきりと異なるテキストブロックであるかを決定し得る。この例において、トランスレータ115は、レストランの名前1020をはっきりと異なるテキストブロックとして特定し、それぞれのメニューの項目をはっきりと異なるテキストブロックとして特定した。

20

【0108】

また、トランスレータ115は、テキストブロックの配列および提示を評価して、2つ以上のテキストブロックが関連し、テキストブロックの集合に属するべきかどうかを決定し得る。特に、トランスレータ115は、テキストブロックのサイズおよび位置、テキストブロックの背景色、テキストのスタイル、ならびに言語モデルを評価して、テキストブロックを集合へとグループ化するかどうかを決定する可能性がある。どのテキストブロックを集合へとグループ化すべきかを決定するためのプロセスが、図13に示され、以下で説明される。

30

【0109】

この例において、トランスレータ115は、たとえば、テキストと背景との間の類似性によって、さらに隣接するメニューの項目の間の同様の間隔によって、メニューの項目を単一の集合として特定した。レストランの名前1020は、たとえば、(たとえば、左揃えにされた)メニューの項目に対する(たとえば、メニュー上で中心揃えにされた)そのレストランの名前1020の向きの違いならびにレストランの名前1020とメニューの項目1022との間のテキストおよび背景の色の違いによって集合から除外される可能性がある。

【0110】

トランスレータ115は、画像1010内のテキストブロックの単一の集合を特定することに応じて画像1010に関する単一集合コンテキスト(single collection context)を選択し得る。トランスレータ115は、画像1015によって描かれたテキストの翻訳を提示するためのユーザインターフェースも選択し得る。このユーザインターフェースは、単一集合コンテキストにふさわしいまたは対応するユーザインターフェースから選択される可能性がある。たとえば、スクリーンショット1000Bおよび1000Cに示されるように、単一集合コンテキストのためのユーザインターフェースは、ユーザが集合に含まれる各テキストブロックの翻訳の間を閲覧することを可能にし得る。スクリーンショット1000Bに示されるユーザインターフェース1030は、ボックス1032によってレストランの名前を強調し、ボックス1034によってそれぞれのメニューの項目を強調する。レストランの名前またはメニューの項目などの強調されたテキストブロックの選択に応じて、ユーザインターフェース1050が提示され得る。このユーザインターフェース1050は、レストランの名前をその元の言語(中国

40

50

語)で提示し、レストランの名前の翻訳1054をユーザによって指定された言語(英語)で提示する。ユーザインターフェース1050は、ユーザがそれぞれのメニューの項目の翻訳を閲覧することを可能にするナビゲーションコントロール1056も含む。たとえば、矢印1058の選択は、ユーザインターフェース1050に次の項目の翻訳を提示させる可能性があり、矢印1057の選択は、ユーザインターフェース1050に前の項目の翻訳を提示させる可能性がある。別のユーザインターフェース要素1059は、ユーザが表示された翻訳に関連するさらなる情報を要求することを可能にする。

#### 【0111】

この例においては、ユーザインターフェース1050は、ユーザがメニューの項目の翻訳に加えてレストランの名前の翻訳を閲覧することを可能にするが、一部の実装においては、集合に含まれるテキストブロックのみを含む可能性がある。集合に含まれないその他のテキストブロックに関する翻訳は、ユーザインターフェース1030内の強調されたテキストブロックを選択することによって個々に要求され得る。

10

#### 【0112】

単一集合コンテキストのためのユーザインターフェースは、集合の各テキストブロックに関して、テキストブロックに関する翻訳を含む画像上のオーバーレイを提示するユーザインターフェースも含み得る。たとえば、図11は、画像を撮影し、画像内に描かれたテキストの言語の翻訳を提示するための例示的なユーザインターフェースのスクリーンショット1100A~1100Cを示す。スクリーンショット1100Aは、図10のスクリーンショット1000Aと同様であり、レストランの名前1120、メニューの項目1122、およびそれぞれのメニューの項目1122の価格1124を有する画像1115を提示するユーザインターフェース1110を含む。

20

#### 【0113】

スクリーンショット1100Bにおいては、ユーザインターフェース1130が、画像1115と、レストランの名前の翻訳を含むオーバーレイ1132と、それぞれのメニューの項目のためのオーバーレイ1134とを提示する。各オーバーレイ1132は、そのそれぞれのメニューの項目の翻訳を含む。また、各オーバーレイ1132および1134は、ユーザが翻訳を実際のメニューと比較し、適切な選択を行うことを容易にするためにそのそれぞれのテキストブロックの上に位置付けられる。

#### 【0114】

各オーバーレイ1132および1134は、テキストに関連する追加的な情報を要求するために選択可能である可能性もある。たとえば、スクリーンショット1100Cに示されるように、ユーザインターフェース1150が、オーバーレイの選択に応じて提示され得る。この例において、ユーザインターフェース1150は、画像1115の最下部付近に表示された「nut milk」に関するオーバーレイの選択に応じて提示される。ユーザインターフェースは、「nut milk」の中国語の文字の元のテキスト1152と、翻訳1154と、ユーザがそれぞれのメニューの項目およびレストラン名の翻訳を閲覧することを可能にするナビゲーションコントロール1156とを提示する。

30

#### 【0115】

翻訳を図11に示されたようにオーバーレイ内に提示すべきかまたは図10に示されたように閲覧モードで提示すべきかを決定するために、トランスレータ115は、上述のように読みやすさの尺度を評価し得る。オーバーレイ内の翻訳に関する読みやすさの尺度の各々が読みやすさの尺度を(たとえば、読みやすさの尺度を満たすかまたは超えることによって)満足する場合、トランスレータ115は、オーバーレイ内に翻訳を提示する可能性がある。そうでない場合、トランスレータ115は、閲覧モードで翻訳を提示する可能性がある。

40

#### 【0116】

一部の画像は、テキストブロックの複数の集合を含むものとして特定される可能性がある。たとえば、図12は、テキストブロックの複数の集合を有するものとして特定された画像内に描かれたテキストの言語の翻訳を提示するための例示的なユーザインターフェースのスクリーンショット1200Aおよび1200Bを示す。スクリーンショット1200Aにおいては、ユーザインターフェース1210が、単一の支柱1221に付いた複数の道路標識の画像1215を示

50

す。たとえば、ユーザは、交差点に居り、標識に基づいてどの道に行くべきかを決定しようとしている可能性がある。

【0117】

トランスレータ115は、各道路標識上に描かれたテキストを特定し、テキストを評価して、各道路標識がそれぞれのその他のテキストブロックのテキストブロックとははっきりと異なるテキストブロックを含むと決定する可能性がある。たとえば、トランスレータ115は、各道路標識のテキストとそれぞれのその他の道路標識との間の距離および隣接する道路標識の間の背景色の不連続に基づいて各道路標識のテキストがはっきりと異なるテキストブロックであると決定し得る。

【0118】

トランスレータ115は、特定されたテキストブロックの配列および提示を評価して、関連する可能性があるテキストブロックのグループを特定することもできる。それから、各グループが集合として分類される可能性がある。たとえば、トランスレータ115は、テキストブロックが同じテキストの色、同じ背景色、同じ向き、同様のテキストサイズを有し、同様の距離で垂直方向に離間されているので支柱1221の左側に掲示された道路標識のテキストブロックを第1の集合1220として特定することができる。同様に、トランスレータ115は、支柱1221の右側に掲示された道路標識のテキストブロックを含む第2の集合1225を特定し得る。第2の集合1225は、支柱1221によって与えられるテキストブロック間の離間のために第1の集合1220とは別個の集合として特定され得る。

【0119】

画像1210に関する複数の集合を特定することに応じて、トランスレータ115は、画像1210に関する複数集合コンテキスト(multiple collection context)を選択し得る。トランスレータ115は、テキストブロックに関連する翻訳またはその他の追加的な情報を提示するためのユーザインターフェースを選択する可能性もある。複数集合コンテキストのためのユーザインターフェースは、オーバーレイ内に翻訳を提示するユーザインターフェースおよび/またはユーザがナビゲーションコントロールを用いて翻訳を閲覧することを可能にするユーザインターフェースを含み得る。これらの2種類のユーザインターフェースの間での選択は、読みやすさの尺度にやはり基づく可能性がある。

【0120】

スクリーンショット1200Bにおいては、ユーザインターフェース1240が、暗くされたバージョンの画像1215、道路標識のうちの1つの元のテキスト1242、および道路標識の翻訳1244を提示する。ユーザインターフェース1240は、ユーザが道路標識の翻訳を閲覧することを可能にするナビゲーションコントロール1246も含む。この例において、ナビゲーションコントロールは、ユーザが道路標識のうちの8つの翻訳を閲覧することを可能にする。これらの8つの道路標識は、第1の集合1220の道路標識の各々と第2の集合1225の下の4つの道路標識とを含む。第2の集合1225の最も高い道路標識は、その標識のテキストの一部がその他の道路標識によって切られているので除外される可能性がある。したがって、正確な翻訳が利用可能でない可能性があり、部分的なテキストを翻訳することはユーザの役に立たない可能性がある。

【0121】

一部の実装において、ユーザインターフェース1240は、ユーザが一度に2つの集合のうちの一方を閲覧することを可能にし得る。たとえば、ユーザインターフェースは、第1の集合の中のテキストブロックの選択に応じて第1の集合1220の道路標識に関するテキストブロックの翻訳を提示する可能性がある。そして、ナビゲーションコントロール1246が、ユーザが第2の集合1225のテキストブロックを含まない第1の集合のみのテキストブロックの翻訳を閲覧することを可能にし得る。ユーザインターフェース1240は、選択されたときに画像1215を示すユーザインターフェース1210にユーザを戻すユーザインターフェース要素1249も含む可能性がある。それから、ユーザは、第2の集合の道路標識に関するテキストブロックを選択し、ユーザインターフェース1240と同様のユーザインターフェースを用いて第2の集合の道路標識によって描かれたテキストの翻訳を閲覧することができる。

10

20

30

40

50

## 【 0 1 2 2 】

トランスレータ115などのトランスレータは、画像によって描かれたテキストの配列および視覚的特徴を評価して、画像に関する集合提示コンテキストを選択し得る。トランスレータ115は、特定されたテキストの少なくとも一部に関する追加的な情報を提示するためのユーザインターフェースを選択する可能性もある。ユーザインターフェースは、選択された集合提示コンテキストに基づいて選択され得る。たとえば、複数集合コンテキストが選択される場合、図12のユーザインターフェース1240などの複数集合コンテキストに対応するユーザインターフェースが選択され得る。

## 【 0 1 2 3 】

図13は、画像のために選択された集合提示コンテキストに基づいて画像内に描かれたテキストに関連する追加的な情報を提示するための例示的なプロセス1300の流れ図である。プロセス1300の動作は、たとえば、図1のユーザデバイス100などのデータ処理装置によって実装される可能性がある。プロセス1300は、コンピュータストレージ媒体に記憶された命令によって実装される可能性もあり、データ処理装置による命令の実行が、データ処理装置にこのプロセス1300の動作を実行させる。

## 【 0 1 2 4 】

画像が受信される(1302)。たとえば、ユーザが、スマートフォン、タブレットコンピュータ、またはその他のユーザデバイスなどのユーザデバイスを用いて対象の写真を撮る可能性がある。標識が、受信された画像内にやはり描かれるユーザにとって異質な言語のテキストを含む可能性がある。

## 【 0 1 2 5 】

受信された画像内のテキストが特定される(1304)。テキストは、OCRまたは別のテキスト特定技術を用いて特定され得る。一部の实装においては、特定の言語のテキストのみが特定される。たとえば、ユーザが、特定の言語のテキストがユーザによって理解される言語などの別の言語に翻訳されるべきであることを指定し得る。この例においては、特定の言語のテキストのみが画像内で特定され得る。

## 【 0 1 2 6 】

実際のテキストに加えて、テキストの特徴が特定される可能性がある。これらの特徴は、テキストの色、テキストの下の背景の色、テキストのサイズ、テキストに関するフォントの種類、および/または画像内のテキストの位置を含み得る。加えて、上述のように、切れているかまたはそれ以外の点で不完全なテキストは無視される可能性がある。

## 【 0 1 2 7 】

1つまたは複数のはっきりと異なるテキストブロックが、特定されたテキストに基づいて特定される(1306)。テキストの特徴が、はっきりと異なるテキストブロックを特定するために評価され得る。たとえば、その他のテキストに近く(たとえば、閾値の距離以内にある)、同様の見た目(たとえば、同様のサイズ、テキストサイズ、および/または背景色)を有するテキストが、テキストブロックAの一部であると考えられる可能性がある。反対に、テキストブロックAから離れており(たとえば、閾値の距離を超えている)、テキストブロックAのテキストとは異なるテキストサイズ、異なるテキストの色、および/または異なる背景色を有するテキストの別の部分は、別のはっきりと異なるテキストブロックBとして特定される可能性がある。

## 【 0 1 2 8 】

画像に関する集合提示コンテキストが、画像内のテキストブロックの配列および/または視覚的特徴に基づいて選択される(1308)。集合提示コンテキストは、単一集合コンテキスト、複数集合コンテキスト、および/または統合集合コンテキスト(integrated collection context)からなど、多数の集合提示コンテキストから選択される可能性がある。どの集合コンテキストを選択すべきかを決定するために、画像に関して1つまたは複数の集合が特定され得る。

## 【 0 1 2 9 】

画像に関して特定されたテキストブロックの特徴が、集合を特定するために評価され得

10

20

30

40

50

る。これらの特徴は、テキストブロックを集合へとグループ化するために使用され得る各テキストブロックのサイズおよび位置、テキストブロックの背景色、テキストのスタイル(たとえば、フォント、色、すべて大文字、小文字、テキストサイズなど)、ならびにテキストブロックの向きを含み得る。同様の特徴を有するテキストブロックが、集合と一緒にグループ化され得る。たとえば、同じフォント、同じテキストの色、同じ背景色を有し、互いに近く(たとえば、互いの閾値の距離以内)にあるテキストブロックが、集合へとグループ化され得る。反対に、異なるフォントおよび異なる色を有する2つのテキストブロックは、一緒に集合に含まれない可能性がある。

**【0130】**

複数のテキストブロックの間の相対的な距離が、テキストブロックを集合へとグループ化するために使用される可能性もある。たとえば、メニューおよびリストは、テキストの行を等しくまたは少なくとも非常に似たように離間することが多い。複数のテキストブロックの間の同様の距離は、テキストブロックが似ていることを示す可能性があり、したがって、複数のテキストブロックと一緒に集合に属すると決定するために使用され得る。

10

**【0131】**

テキストブロックの向きも、テキストブロックと一緒に集合に含まれるべきであるかどうかを決定するために使用され得る。たとえば、いくつかのテキストブロックが画像に対して同様の角度で配向されている場合、テキストブロックは関連する可能性がある。反対に、1つのテキストブロックのテキストが水平の向きを有し、別のテキストブロックのテキストが水平に対して上向きの角度で傾いている場合、テキストブロックは関連しない可能性があり、したがって、一緒に集合に含まれない可能性がある。

20

**【0132】**

各テキストブロックの背景色が、テキストブロックを集合に含めるかどうかを決定するために比較される可能性もある。同様の背景色を有するテキストブロックは、関連する可能性が高く、一緒に集合に含まれる可能性がある。反対に、実質的に異なる色を有するテキストブロック(たとえば、1つが青であり、1つが赤である)は関連しない可能性があり、一緒に集合に含まれない可能性がある。たとえば、メニューのリストに提示されるメニューの項目は、同じフォントの色および背景色を有することが多い。この例において、メニューの項目は、一緒に集合へとグループ化され得る。

**【0133】**

言語モデルが、テキストブロックを集合へとグループ化するために使用される可能性もある。たとえば、言語モデルが、2つ以上のテキストブロックのテキストが関連するかどうかを決定するために使用される可能性がある。テキストブロックは、言語モデルに基づいて関連すると決定される場合、一緒に集合に含まれ得る。

30

**【0134】**

集合提示コンテキストが、特定された集合に基づいて選択され得る。たとえば、複数の集合が特定される場合、画像に関して複数集合コンテキストが選択され得る。画像に関して単一の集合が特定される場合、単一集合コンテキストと統合集合コンテキストとの間で選択がなされ得る。単一集合コンテキストは、集合に含まれる各テキストブロックに関する追加的な情報がそれぞれのその他のテキストブロックとは独立に提示され得るときに選択される可能性がある。たとえば、単一集合コンテキストは、メニューの項目の集合に関して選択され得る。統合テキストブロック(integrated text block)は、追加的な情報が一緒に提示されるべきであるように追加的な情報が関連するときに選択され得る。たとえば、ユーザが翻訳をその翻訳の対応する道路標識にマッチングすることができるように道路標識の翻訳を一緒に示す方がよい可能性がある。

40

**【0135】**

単一集合コンテキストと統合集合コンテキストとの間で選択を行うために、言語モデルが使用され得る。たとえば、言語モデルは、テキストがメニューの複数の食べ物の項目を指すかまたはテキストが複数の道路標識を指すと決定するために使用され得る。番号または価格などの識別子の存在も、単一集合コンテキストと統合集合コンテキストとの間で選

50

択を行うために使用され得る。集合の各テキストブロックが識別子を含む場合、各テキストブロックは独立して提示される可能性がある。したがって、統合集合コンテキストではなく単一集合コンテキストが選択され得る。

【0136】

追加的な情報が、テキストに関して特定される(1310)。追加的な情報は、選択されたテキストの言語の翻訳を含む可能性がある。たとえば、ユーザが、そのユーザが画像によって描かれたテキストが翻訳されることを望む言語を選択する可能性がある。通貨の金額を含むテキストに関して、追加的な情報は、ユーザによって指定された通貨への金額の変換を含む可能性がある。たとえば、選択されたテキストが日本円の価格を含む場合、追加的な情報は、円の価格に対応するドルの金額を含む可能性がある。

10

【0137】

ユーザインターフェースが、選択された集合提示コンテキストおよび/または追加的な情報に関する読みやすさの尺度に基づいて選択される(1312)。各集合提示コンテキストは、画像内に描かれたテキストに関する追加的な情報を提示するために使用され得る1つまたは複数の対応するユーザインターフェースを有する可能性がある。加えて、1つの集合提示コンテキストにふさわしいユーザインターフェースは、それぞれのその他の集合提示コンテキストのためのユーザインターフェースとは異なる可能性がある。

【0138】

複数集合コンテキストは、(たとえば、図11のユーザインターフェース1130と同様の)画像上のオーバーレイ内に追加的な情報を提示するユーザインターフェースおよび/またはユーザが1つもしくは複数の別個の画面内の集合の各テキストブロックに関する追加的な情報を閲覧することを可能にするユーザインターフェース(たとえば、図12のユーザインターフェース1240)を有する可能性がある。上述のように、各集合は、ユーザインターフェース1240のようなユーザインターフェースにおいてそれぞれのその他の集合とは別々に閲覧され得る。

20

【0139】

単一集合コンテキストも、(たとえば、図11のユーザインターフェース1130と同様の)画像上のオーバーレイ内に追加的な情報を提示するユーザインターフェースおよび/またはユーザが1つもしくは複数の別個の画面内の集合の各テキストブロックに関する追加的な情報を閲覧することを可能にするユーザインターフェース(たとえば、図12のユーザインターフェース1240)を含む可能性がある。一部の実装において、単一集合コンテキストは、デフォルトで自動的に集合の各テキストブロックに関する追加的な情報を別々に提示するように設定されるユーザインターフェースに対応する可能性がある。たとえば、各テキストブロックが識別子を含む場合、デフォルトのまたは好ましいユーザインターフェースは、各テキストブロックに関する追加的な情報を別々に提示し、ユーザが集合のそれぞれのその他のテキストブロックに関する追加的な情報を閲覧することを可能にするユーザインターフェースである可能性がある。

30

【0140】

統合集合コンテキストは、集合に含まれる各テキストブロックに関する追加的な情報を同時にまたは同じ画面上に提示するユーザインターフェースに(たとえば、ユーザが画面に収まらない情報を閲覧することを可能にするユーザインターフェース要素に)関連付けられる可能性がある。統合集合コンテキストに関するユーザインターフェースは、画像上の1つまたは複数のオーバーレイ内に追加的な情報を提示するユーザインターフェースと、別個の画面上に追加的な情報を提示するユーザインターフェースとを含み得る。オーバーレイユーザインターフェースの一例においては、図12の各道路標識の翻訳が、その翻訳の対応する道路標識上のオーバーレイ内に提示される可能性がある。このようにして、ユーザは、どの翻訳がどの道路標識に対応するのか容易に決定することができる。

40

【0141】

各集合提示コンテキストは、そのコンテキストのための好ましいユーザインターフェースを有する可能性がある。たとえば、複数集合コンテキストおよび統合集合コンテキスト

50

のための好ましいユーザインターフェースは、画像上のオーバーレイ内に追加的な情報を提示するユーザインターフェースである可能性がある。単一集合コンテキストのための好ましいユーザインターフェースは、各テキストブロックに関する追加的な情報を別々に提示するユーザインターフェースである可能性がある。

【0142】

一部の実装においては、追加的な情報に関する読みやすさの尺度が、追加的な情報を提示するためのユーザインターフェースを選択するために使用される。読みやすさの尺度が、選択された集合提示コンテキストのための好ましいユーザインターフェース(たとえば、画像上のオーバーレイを有するユーザインターフェース)を使用すべきかまたは代替的なユーザインターフェース(たとえば、ユーザが追加的な情報を閲覧することを可能にするユーザインターフェース)を使用すべきかを決定するために使用され得る。たとえば、好ましい集合提示コンテキストに関連する追加的な情報に関する読みやすさの尺度が読みやすさの閾値を(たとえば、読みやすさの閾値を満たすかまたは超えることによって)満足する場合、好ましいユーザインターフェースが選択される可能性がある。好ましい集合提示コンテキストに関連する追加的な情報に関する読みやすさの尺度が読みやすさの閾値を満足しない(たとえば、読みやすさの尺度が読みやすさの閾値未満である)場合、追加的な情報のより多くの文字の提示を可能にする代替的なユーザインターフェースが選択される可能性がある。

10

【0143】

読みやすさの尺度は、好ましいユーザインターフェースとは独立した複数のユーザインターフェースの間で選択を行うために使用される可能性がある。たとえば、選択された集合提示コンテキストにふさわしく、最も高い読みやすさの尺度をもたらすユーザインターフェースが、選択される可能性がある。

20

【0144】

追加的な情報が、選択されたユーザインターフェースを用いて提示される(1314)。たとえば、ユーザデバイスは、ユーザインターフェースを生成し、選択されたユーザインターフェース内に追加的な情報を提示する可能性がある。

【0145】

地図提示コンテキストおよびユーザインターフェース

一部の文脈においては、画像内に描かれたテキストに基づいて場所への地図または指示を受信することがユーザにとって有益である可能性がある。たとえば、ユーザは、イベントの住所を含むイベントのポスターの写真を撮る可能性がある。場所の翻訳または住所の翻訳をとまなう地図を含むユーザインターフェースを提示することは、ユーザがイベントを特定し、イベントにたどり着く方法を決定するのを助け得る。

30

【0146】

図1のトランスレータ115などのトランスレータは、地図ユーザインターフェースの提示がユーザにとって有益である可能性があるとき、地図提示コンテキストを選択し得る。たとえば、トランスレータ115は、撮影された画像によって描かれた1つまたは複数の住所を検出することに応じて地図提示コンテキストを選択し得る。地図提示コンテキストのためのユーザインターフェースは、検出された住所に対応する地図、および/またはユーザがトランスレータ115に位置情報を提供した場合にはユーザの現在の位置を含み得る。

40

【0147】

地図提示コンテキストは、単一住所地図コンテキスト(single address map context)および複数住所地図コンテキスト(multiple address map context)を含み得る。単一住所地図コンテキストは、単一の住所が画像内で検出されることに応じて選択され得る。単一住所地図コンテキストのためのユーザインターフェースは、単一の住所に対応する場所を示す地図および/または単一の住所への指示を含み得る。ユーザインターフェースは、住所の言語の翻訳および/または住所によって参照される場所もしくは建物の名前の言語の翻訳を提示する可能性もある。複数住所地図コンテキストのためのユーザインターフェースは、ユーザがそれぞれの検出された住所に関する情報を閲覧することを可能にするコント

50

ルールを含み得る。たとえば、ユーザインターフェースは、ユーザが複数の住所のうちの特定の住所に関する地図および/または指示をそれぞれが示すカードまたは画面を閲覧することを可能にする可能性がある。各カードまたは画面は、カードまたは画面が対応する住所の言語の翻訳および/または住所によって参照される場所もしくは建物の名前の言語の翻訳を提示する可能性もある。地図提示コンテキストユーザインターフェースにおいて提供される指示は、ユーザによって指定された翻訳言語による可能性がある。

【0148】

図14は、地図コンテキストを有するものとして特定された画像内に描かれたテキストに関連する追加的な情報を提示するための例示的なユーザインターフェースのスクリーンショット1400A~1400Cを示す。スクリーンショット1400Aにおいては、たとえば、画像撮影ボタン1414のユーザの選択に応じて画像1412が撮影され、ユーザインターフェース1410に提示される。例示的な画像1412は、イベントに関連する情報およびイベントの住所1416を含む。

10

【0149】

トランスレータ115は、画像1412内のテキストを分析して住所を特定することができる。たとえば、トランスレータ115は、画像内のテキストを特定し、特定されたテキストを知られている住所と比較することができる。一致が存在する場合、トランスレータ115は、テキストを住所として特定し得る。住所1414を特定することに応じて、ユーザが住所への地図および/または指示などの住所に関連する追加的な情報を要求し、取得することを可能にするユーザインターフェースが提供され得る。たとえば、ユーザインターフェース1430が、スクリーンショット1400Bに示されるように提示される可能性がある。

20

【0150】

スクリーンショット1400Bにおいて、ユーザインターフェース1430は、住所1414の周りのボックス1432を用いて住所1414を強調する。また、ユーザインターフェース1430は、ユーザが住所1414に対応する地図を見ることを可能にする「地図上で見る」コントロール1434を提供する。コントロール1434の選択に応じて、ユーザインターフェース1450が、スクリーンショット1400Cに示されるように提示される可能性がある。ユーザインターフェース1450は、住所の場所を含む地図1454と、住所の翻訳を含むオーバーレイ1452と、住所1414を特定する地図1454上のマーカ1456と、住所を特定するテキスト1458とを提示する。追加的にまたは代替的に、ユーザインターフェース1450は、イベントが行われるであろう建物または場所の名前の翻訳を提示する可能性がある。

30

【0151】

一部の実装においては、地図または指示を住所または場所の翻訳と一緒に提示するユーザインターフェースが、画像内の住所を特定することに応じて自動的に提示される可能性がある。たとえば、ユーザインターフェース1450が、ユーザがコントロール1434を選択することなく、アドレス1414を検出することに応じて提示される可能性がある。

【0152】

例示的なユーザインターフェース1450は、画像1412内で特定された別の住所があった場合、ユーザが別のカードまたは画面を閲覧することを可能にする閲覧コントロール1458も含む。たとえば、閲覧コントロール1458の選択は、ユーザデバイスに画像1412内で特定された別の住所に関する地図および翻訳を提示させる可能性がある。

40

【0153】

図15は、画像のために選択された地図提示コンテキストに基づいて画像内に描かれたテキストに関連する追加的な情報を提示するための例示的なプロセス1500の流れ図である。プロセス1500の動作は、たとえば、図1のユーザデバイス100および/またはトランスレータ115などのデータ処理装置によって実装される可能性がある。プロセス1500は、コンピュータストレージ媒体に記憶された命令によって実装される可能性もあり、データ処理装置による命令の実行が、データ処理装置にこのプロセス1500の動作を実行させる。

【0154】

画像が受信される(1502)。たとえば、ユーザが、スマートフォン、タブレットコンピュ

50

ータ、またはその他のユーザデバイスなどのユーザデバイスを用いて対象の写真を撮る可能性がある。標識が、受信された画像内にやはり描かれるユーザにとって異質な言語のテキストを含む可能性がある。

【0155】

受信された画像内のテキストが特定される(1504)。テキストは、OCRまたは別のテキスト特定技術を用いて特定され得る。一部の実装においては、特定の言語のテキストのみが特定される。たとえば、ユーザが、特定の言語のテキストがユーザによって理解される言語などの別の言語に翻訳されるべきであることを指定し得る。この例においては、特定の言語のテキストのみが画像内で特定され得る。

【0156】

住所が、特定されたテキスト内で特定される(1506)。たとえば、特定されたテキストが住所を指すかどうかを決定するために、特定されたテキストが知られている住所と比較される可能性がある。画像が複数の住所を描く場合、それぞれの住所が特定され得る。

【0157】

地図提示コンテキストが、特定された住所に基づいて特定される(1508)。一部の実装においては、地図コンテキストが、画像内で特定された住所の数に基づいて選択される。たとえば、単一の住所が画像内で特定されることに応じて単一住所地図コンテキストが選択される可能性がある一方、2つ以上の住所が画像内で特定されることに応じて複数住所地図コンテキストが選択される可能性がある。

【0158】

特定された住所に関連する追加的な情報を提示するためのユーザインターフェースが、選択される。上述のように、地図提示コンテキストのためのユーザインターフェースが、住所に対応する場所を示す地図、および/またはユーザが現在の位置へのアクセスを提供した場合はユーザの現在の位置から住所への指示を示し得る。地図提示コンテキストのためのユーザインターフェースは、地図または指示の上のオーバーレイ内に住所または住所に対応する場所の名前の言語の翻訳を含む可能性もある。単一住所地図コンテキストに関して、ユーザインターフェースは、単一の住所に関する地図および/または指示を含み得る。複数地図コンテキストに関して、ユーザインターフェースは、図14を参照して上で説明されたように、それぞれの特定された場所に関する地図および指示を閲覧するためのコントロールを含み得る。

【0159】

住所に関する追加的な情報が、選択されたユーザインターフェースを用いて提示される(1512)。たとえば、住所への地図または指示と、住所または住所に対応する場所の名前の翻訳を提示するオーバーレイとを提示するユーザインターフェースが、生成される可能性がある。このユーザインターフェースは、画像内の住所を特定することに応じて自動的に提示される可能性がある。一部の実装においては、画像内の住所を特定することに応じて、ユーザが地図ユーザインターフェースを見ることを可能にするユーザインターフェースが提示される可能性がある。たとえば、図14のユーザインターフェース1430と同様のユーザインターフェースが、生成され、提示される可能性がある。「地図上で見る」コントロール1434を選択すると、地図または指示を有するユーザインターフェースが提示される可能性がある。

【0160】

画像に関する提示コンテキストを選択するための例示的なプロセス

上述のように、提示コンテキストの各カテゴリ内の(たとえば、顕著さプレゼンテーションカテゴリ(prominence presentation category)または集合プレゼンテーションカテゴリ(collection presentation category)内の)提示コンテキストを選択することに加えて、トランスレータ115は、画像に関する提示コンテキストのカテゴリを選択する可能性がある。たとえば、画像を受信することに応じて、トランスレータ115が、画像に関して顕著さ提示コンテキストを選択するのか、画像に関して集合提示コンテキストを選択するのか、または画像に関して地図コンテキストを選択するのかを決定し得る。この選択は、画

10

20

30

40

50

像によって描かれたテキストの配列および/または視覚的特徴に基づいて行われ得る。

【0161】

また、トランスレータ115は、選択された提示コンテキストに基づいて、画像によって描かれたテキストに関連する追加的な情報を提示するためのユーザインターフェースを選択し得る。上述のように、各提示コンテキストは、1つまたは複数のユーザインターフェースに関連付けられる可能性があり、1つの提示コンテキストのためのユーザインターフェースは、その他の提示コンテキストのためのユーザインターフェースとは異なる可能性がある。たとえば、オーバーレイ内に支配的なテキストブロックの翻訳を提示するためのユーザインターフェースは、単一集合コンテキストのそれぞれのテキストブロックの翻訳を別に提示するためのユーザインターフェースとは異なる。

10

【0162】

図16は、画像のために選択された提示コンテキストに基づいて画像内に描かれたテキストに関連する追加的な情報を提示するための例示的なプロセス1600の流れ図である。プロセス1600の動作は、たとえば、図1のユーザデバイス100および/またはトランスレータ115などのデータ処理装置によって実装される可能性がある。プロセス1600は、コンピュータストレージ媒体に記憶された命令によって実装される可能性もあり、データ処理装置による命令の実行が、データ処理装置にこのプロセス1600の動作を実行させる。

【0163】

画像が受信される(1602)。たとえば、ユーザが、スマートフォン、タブレットコンピュータ、またはその他のユーザデバイスなどのユーザデバイスを用いて対象の写真を撮る可能性がある。標識が、受信された画像内にやはり描かれるユーザにとって異質な言語のテキストを含む可能性がある。

20

【0164】

受信された画像内のテキストが特定される(1604)。テキストは、OCRまたは別のテキスト特定技術を用いて特定され得る。一部の实装においては、特定の言語のテキストのみが特定される。たとえば、ユーザが、特定の言語のテキストがユーザによって理解される言語などの別の言語に翻訳されるべきであることを指定し得る。この例においては、特定の言語のテキストのみが画像内で特定され得る。

【0165】

実際のテキストに加えて、テキストの特徴が特定される可能性がある。これらの特徴は、テキストの色、テキストの下の背景の色、テキストのサイズ、テキストに関するフォントの種類、および/または画像内のテキストの位置を含み得る。加えて、上述のように、切れているかまたはそれ以外の点で不完全なテキストは無視される可能性がある。

30

【0166】

1つまたは複数のはっきりと異なるテキストブロックが、特定されたテキストに基づいて特定される(1606)。テキストの特徴が、はっきりと異なるテキストブロックを特定するために評価され得る。たとえば、その他のテキストに近く(たとえば、閾値の距離以内にある)、同様の見た目(たとえば、同様のサイズ、テキストサイズ、および/または背景色)を有するテキストが、テキストブロックAの一部であると考えられる可能性がある。反対に、テキストブロックAから離れており(たとえば、閾値の距離を超えている)、テキストブロックAのテキストとは異なるテキストサイズ、異なるテキストの色、および/または異なる背景色を有するテキストの別の部分は、別のはっきりと異なるテキストブロックBとして特定される可能性がある。

40

【0167】

提示コンテキストが、テキストブロックの配列および/またはテキストブロックの特徴に基づいて画像に関して選択される(1608)。テキストブロックに基づいて、画像が提示コンテキストのうちの1つとよく一致するかどうか決定され得る。たとえば、画像が1つまたは複数の集合へとグループ化され得る複数のテキストブロックを含む場合、集合提示コンテキストが選択され得る。この例においては、画像が、画像内の任意の集合と、いずれかの集合が特定される場合には特定された集合の数とを特定するために分析される可能性

50

がある。集合が特定されない場合、それぞれの特定されたテキストブロックの顕著さが、顕著さ提示コンテキストが画像に関して選択されるべきかどうかを決定するために評価され得る。テキストブロックのうちの1つまたは複数が画像内で顕著に描かれる場合、顕著さ提示コンテキストが画像に関して選択され得る。

【0168】

また、特定されたテキストブロックは、テキストブロックに住所が含まれるかどうかを決定するために分析される可能性がある。住所が検出される場合、地図提示コンテキストが、住所を検出することに応じて自動的に特定され得る。一部の実装において、地図提示コンテキストは、テキストブロック内に検出された閾値未満の数の住所が存在するときに選択され得る。たとえば、画像は、多くの住所を含む可能性があり、集合コンテキストまたは顕著さコンテキストが、より適切である可能性がある。住所を有するテキストブロックが集合へとグループ化され得る場合、集合提示コンテキストが選択され得る。住所のうちの1つがその他の住所よりもずっと顕著に描かれる(たとえば、住所がその他の住所の顕著さスコアよりも閾値の量を超えて大きい顕著さスコアを有する)場合、顕著さ提示コンテキストが選択され得る。

10

【0169】

一部の実装においては、上述のプロセス900、1300、および1500のうちの1つまたは複数、画像に関して実行される可能性があり、各提示コンテキストが、提示コンテキストが画像にどれだけよく一致するかどうかに基づいてスコアを付けられる可能性がある。そして、最も高いスコアを有する提示コンテキストが、画像のために選択され得る。

20

【0170】

提示コンテキストは、画像内で特定されたはっきりと異なるテキストブロックの数に基づいて選択される可能性がある。たとえば、画像内で特定されたテキストブロックの数がテキストブロックの閾値の数を超える場合、集合提示コンテキストが選択される可能性があり、テキストブロックが集合に含まれる可能性がある。閾値未満の数のテキストブロックが画像内で特定される場合、顕著さ提示コンテキストが選択され得る。選択された顕著さ提示コンテキストは、上述のように、特定された各テキストブロックの顕著さに基づく可能性がある。

【0171】

追加的な情報が提示されるべきテキストが、選択された顕著さ提示コンテキストに基づいて選択される(1610)。一部の提示コンテキストにおいては、追加的な情報が、特定されたテキストの一部のみに関して提示される。たとえば、支配的-副次的テキストブロック(dominant-secondary text block)が選択される場合、支配的なテキストブロックであるものとして特定されたテキストブロックに関する追加的な情報が特定され得る。したがって、支配的なテキストブロックのテキストが、追加的な情報を提示するために選択され得る。集合提示コンテキストが選択される場合、特定された集合の各テキストブロックが特定され得る。地図提示コンテキストが選択される場合、画像内のそれぞれの特定された住所が、追加的な情報を提示するために選択され得る。

30

【0172】

追加的な情報が、選択されたテキストに関して特定される(1612)。追加的な情報は、選択されたテキストの言語の翻訳を含む可能性がある。たとえば、ユーザが、そのユーザが画像によって描かれたテキストが翻訳されることを望む言語を選択する可能性がある。通貨の金額を含むテキストに関して、追加的な情報は、ユーザによって指定された通貨への金額の変換を含む可能性がある。たとえば、選択されたテキストが日本円の価格を含む場合、追加的な情報は、円の価格に対応するドルの金額を含む可能性がある。

40

【0173】

ユーザインターフェースが、選択された提示コンテキストおよび/または追加的な情報に関する読みやすさの尺度に基づいて選択される(1614)。上述のように、各提示コンテキストは、提示コンテキストが選択される画像内に描かれたテキストに関連する追加的な情報を提示するために使用される資格がある1つまたは複数のユーザインターフェースに関

50

連付けられる可能性がある。選択された提示コンテキストのためのユーザインターフェースのうちの一つが、上述のように、たとえば、選択された提示コンテキストのための好ましいユーザインターフェースならびに/または資格があるユーザインターフェースおよび追加的な情報に関する読みやすさの尺度に基づいて選択され得る。

【0174】

追加的な情報が、選択されたユーザインターフェースを用いて提示される(1616)。たとえば、ユーザデバイスは、ユーザインターフェースを生成し、選択されたユーザインターフェースを用いて追加的な情報を提示する可能性がある。

【0175】

さらなる実装の詳細

本明細書に記載の対象の実施形態および動作は、本明細書において開示された構造およびそれらの構造的均等物を含む、デジタル電子回路、またはコンピュータソフトウェア、ファームウェア、もしくはハードウェア、またはそれらのうちの一つもしくは複数の組合せで実装され得る。本明細書に記載の対象の実施形態は、一つまたは複数のコンピュータプログラム、すなわち、データ処理装置による実行のために、またはデータ処理装置の動作を制御するためにコンピュータストレージ媒体上に符号化されたコンピュータプログラム命令の一つまたは複数のモジュールとして実装され得る。代替的にまたは追加的に、プログラム命令は、データ処理装置による実行のために好適な受信機装置に送信するために情報を符号化するように生成される人為的に生成された伝播信号、たとえば、機械によって生成された電気的信号、光学的信号、または電磁的信号上に符号化され得る。コンピュータストレージ媒体は、コンピュータ可読ストレージデバイス、コンピュータ可読ストレージ基板、ランダムもしくはシリアルアクセスメモリアレイもしくはデバイス、またはそれらの一つもしくは複数の組合せであるか、あるいはそれらに含まれる可能性がある。さらに、コンピュータストレージ媒体は、伝播信号ではないが、人為的に生成された伝播信号に符号化されたコンピュータプログラム命令の送信元または送信先である可能性がある。コンピュータストレージ媒体は、一つまたは複数の別個の物理コンポーネントまたは媒体(たとえば、複数のCD、ディスク、またはその他のストレージデバイス)であるか、またはそれらに含まれる可能性もある。

【0176】

本明細書に記載の動作は、一つもしくは複数のコンピュータ可読ストレージデバイスに記憶された、またはその他のソースから受信されたデータに対してデータ処理装置によって実行される動作として実装され得る。

【0177】

用語「データ処理装置」は、例として、一つのプログラム可能なプロセッサ、一台のコンピュータ、一つのシステムオンチップ、またはそれらの複数もしくは組合せを含む、データを処理するためのすべての種類の装置、デバイス、および機械を包含する。装置は、専用の論理回路、たとえば、FPGA(フィールドプログラマブルゲートアレイ)またはASIC(特定用途向け集積回路)を含み得る。装置は、ハードウェアに加えて、問題にしているコンピュータプログラムのための実行環境を生成するコード、たとえば、プロセッサのファームウェア、プロトコルスタック、データベース管理システム、オペレーティングシステム、クロスプラットフォームランタイム環境、仮想マシン、またはそれらのうちの一つもしくは複数の組合せを構成するコードも含み得る。装置および実行環境は、ウェブサーバインフラストラクチャ、分散コンピューティングインフラストラクチャ、およびグリッドコンピューティングインフラストラクチャなどの、様々な異なるコンピューティングモデルインフラストラクチャを実現することができる。

【0178】

コンピュータプログラム(プログラム、ソフトウェア、ソフトウェアアプリケーション、スクリプト、またはコードとしても知られる)は、コンパイラ型言語もしくはインタープリタ型言語、宣言型言語もしくは手続き型言語を含む任意の形態のプログラミング言語で記述可能であり、独立型プログラムとしての形態、またはモジュール、コンポーネント

10

20

30

40

50

、サブルーチン、オブジェクト、もしくはコンピューティング環境での使用に好適なその他の単位としての形態を含む任意の形態で展開され得る。コンピュータプログラムは、ファイルシステム内のファイルに対応する可能性があるが、ファイルシステム内のファイルに対応することは必須ではない。プログラムは、その他のプログラムもしくはデータを保持するファイルの一部(たとえば、マークアップ言語の文書に記憶された1つもしくは複数のスクリプト)、問題にしているプログラムに専用の単一のファイル、または複数の組織されたファイル(たとえば、1つもしくは複数のモジュール、サブプログラム、もしくはコードの一部を記憶するファイル)に記憶され得る。コンピュータプログラムは、1つのコンピュータ上で、または1つの場所に置かれるか、もしくは複数の場所に分散され、通信ネットワークによって相互に接続されている複数のコンピュータ上で実行されるように展開され得る。

10

## 【0179】

本明細書に記載のプロセスおよび論理フローは、入力データに対して演算を行い、出力を生成することによって行為を行うために1つまたは複数のコンピュータプログラムを1つまたは複数のプログラミング可能なプロセッサが実行することによって実行され得る。また、プロセスおよび論理フローは、専用の論理回路、たとえば、FPGA (フィールドプログラマブルゲートアレイ)またはASIC (特定用途向け集積回路)によって実行される可能性があり、さらに、装置は、それらの専用の論理回路として実装される可能性がある。

## 【0180】

コンピュータプログラムの実行に好適なプロセッサは、例として、汎用マイクロプロセッサと専用マイクロプロセッサとの両方、および任意の種類のデジタルコンピュータの任意の1つまたは複数のプロセッサを含む。概して、プロセッサは、読み出し専用メモリ、またはランダムアクセスメモリ、またはそれらの両方から命令およびデータを受け取る。コンピュータの必須の要素は、命令に従って行為を実行するためのプロセッサ、ならびに命令およびデータを記憶するための1つまたは複数のメモリデバイスである。また、概してコンピュータは、データを記憶するための1つまたは複数の大容量ストレージデバイス、たとえば、磁気ディスク、光磁気ディスク、もしくは光ディスクを含むか、またはそれらの大容量ストレージデバイスからデータを受信するか、もしくはそれらの大容量ストレージデバイスにデータを転送するか、もしくはそれら両方を行うために動作可能なように結合される。しかし、コンピュータは、そのようなデバイスを備えていなくてもよい。さらに、コンピュータは、別のデバイス、たとえば、ほんのいくつか例を挙げるとすれば、モバイル電話、携帯情報端末(PDA)、モバイルオーディオもしくはビデオプレーヤー、ゲームコンソール、全地球測位システム(GPS)受信機、またはポータブルストレージデバイス(たとえば、ユニバーサルシリアルバス(USB)フラッシュドライブ)に組み込まれる可能性がある。コンピュータプログラム命令およびデータを記憶するのに好適なデバイスは、例として、半導体メモリデバイス、たとえば、EPROM、EEPROM、およびフラッシュメモリデバイス、磁気ディスク、たとえば、内蔵ハードディスクまたはリムーバブルディスク、光磁気ディスク、ならびにCD-ROMディスクおよびDVD-ROMディスクを含む、すべての形態の不揮発性メモリ、媒体、およびメモリデバイスを含む。プロセッサおよびメモリは、専用の論理回路によって補完されるか、または専用の論理回路に組み込まれる可能性がある。

20

30

40

## 【0181】

ユーザとのインタラクションを行うために、本明細書に記載の対象の実施形態は、ユーザに対して情報を表示するためのディスプレイデバイス、たとえば、CRT (ブラウン管)またはLCD (液晶ディスプレイ)モニタ、ならびに、ユーザがコンピュータに入力を与えることができるキーボードおよびポインティングデバイス、たとえば、マウスまたはトラックボールを有するコンピュータ上に実装され得る。その他の種類のデバイスが、ユーザとのインタラクションを行うためにやはり使用される可能性があり、たとえば、ユーザに提供されるフィードバックは、任意の形態の感覚フィードバック、たとえば、視覚的フィードバック、聴覚的フィードバック、または触覚的フィードバックである可能性があり、ユー

50

ザからの入力、音響、発話、または触覚による入力を含む任意の形態で受け取られる可能性がある。加えて、コンピュータは、ユーザによって使用されるデバイスに文書を送信し、そのデバイスから文書を受信することによって、たとえば、ウェブブラウザから受信された要求に回答してユーザのクライアントデバイスのウェブブラウザにウェブページを送信することによってユーザとインタラクションすることができる。

【0182】

本明細書に記載の対象の実施形態は、バックエンドコンポーネントを、たとえば、データサーバとして含むか、またはミドルウェアコンポーネント、たとえば、アプリケーションサーバを含むか、またはフロントエンドコンポーネント、たとえば、それによってユーザが本明細書に記載の対象の実装とインタラクションすることができるグラフィカルユーザインターフェースもしくはウェブブラウザを有するクライアントコンピュータを含むか、または1つもしくは複数のそのようなバックエンドコンポーネント、ミドルウェアコンポーネント、もしくはフロントエンドコンポーネントの任意の組合せを含むコンピューティングシステムに実装され得る。システムのコンポーネントは、任意の形態または媒体のデジタルデータ通信、たとえば、通信ネットワークによって相互に接続され得る。通信ネットワークの例は、ローカルエリアネットワーク(「LAN」)および広域ネットワーク(「WAN」)、インターネットワーク(たとえば、インターネット)、ならびにピアツーピアネットワーク(たとえば、アドホックピアツーピアネットワーク)を含む。

10

【0183】

コンピューティングシステムは、クライアントおよびサーバを含み得る。クライアントおよびサーバは、概して互いに離れており、通常は通信ネットワークを介してインタラクションする。クライアントとサーバとの関係は、それぞれのコンピュータ上で実行されており、互いにクライアント-サーバの関係にあるコンピュータプログラムによって生じる。一部の実施形態において、サーバは、(たとえば、クライアントデバイスとインタラクションするユーザに対してデータを表示し、そのようなユーザからユーザ入力を受信する目的で)クライアントデバイスにデータ(たとえば、HTMLページ)を送信する。クライアントデバイスにおいて生成されたデータ(たとえば、ユーザのインタラクションの結果)が、サーバにおいてクライアントデバイスから受信される可能性がある。

20

【0184】

本明細書は多くの特定の实装の詳細を含むが、これらは、いかなる発明の範囲または特許請求され得るものの範囲に対する限定とも見なされるべきではなく、むしろ、特定の発明の特定の实施形態に固有の特徴の説明と見なされるべきである。別個の实施形態に関連して本明細書に記載されている特定の特征が、単一の实施形態で组合せて実装される可能性もある。反対に、単一の实施形態に関連して説明されている様々な特征が、複数の实施形態に別々に、または任意の適切な部分的组合せで実装される可能性もある。さらに、特征は、特定の组合せで動作するものとして上で説明されている可能性があり、最初にそのように主張されてさえいる可能性があるが、主張された组合せの1つまたは複数の特征は、場合によってはその组合せから削除される可能性があり、主張された组合せは、部分的组合せ、または部分的组合せの変形を対象とする可能性がある。

30

【0185】

同様に、動作が図中に特定の順序で示されているが、これは、所望の結果を達成するために、そのような動作が示された特定の順序でもしくは逐次的順序で実行されること、または示されたすべての動作が実行されることを必要とするものと理解されるべきでない。特定の状況においては、マルチタスクおよび並列処理が、有利である可能性がある。さらに、上述の実施形態における様々なシステムコンポーネントの分割は、すべての実施形態においてそのような分割を必要とするものと理解されるべきでなく、説明されたプログラムコンポーネントおよびシステムは、概して、単一のソフトウェア製品と一緒に統合されるか、または複数のソフトウェア製品にパッケージングされる可能性があると理解されるべきである。

40

【0186】

50

このように、対象の特定の実施形態が説明された。その他の実施形態は、添付の特許請求の範囲内にある。場合によっては、特許請求の範囲に挙げられた行為は、異なる順序で実行され、それでも所望の結果を達成し得る。加えて、添付の図面に示されたプロセスは、所望の結果を達成するために、必ずしも、示された特定の順序または逐次的順序である必要はない。特定の実装においては、マルチタスクおよび並列処理が、有利である可能性がある。

【符号の説明】

【 0 1 8 7 】

100	ユーザデバイス	
105	ディスプレイ	10
107	画像	
109	テキスト	
111	翻訳	
115	トランスレータ	
120	テキスト識別器	
125	翻訳エンジン	
130	ユーザインターフェースエンジン	
135	コンテキストセクタ	
140	ユーザインターフェースインデックス	
200A	スクリーンショット	20
200B	スクリーンショット	
200C	スクリーンショット	
206	カメラボタン	
210	ユーザインターフェース	
212	元言語	
214	翻訳言語	
230	ユーザインターフェース	
232	画像	
234	テキストの第1の部分	
236	テキストの第2の部分	30
238	テキストの第3の部分	
250	翻訳画面ユーザインターフェース	
252	テキスト	
254	英語翻訳	
300A	スクリーンショット	
300B	スクリーンショット	
330	ユーザインターフェース	
332	画像	
334	テキストの第1の部分	
336	テキストの第2の部分	40
338	テキストの第3の部分	
350	翻訳画面ユーザインターフェース	
352	テキストの部分	
354	英語翻訳	
359	コントロール	
400A	スクリーンショット	
400B	スクリーンショット	
400C	スクリーンショット	
430	ユーザインターフェース	
432	画像	50

434	テキストの第1の部分	
436	テキストの第2の部分	
438	テキストの第3の部分	
450	ユーザインターフェース	
454	オーバーレイ	
456	オーバーレイ	
458	オーバーレイ	
470	ユーザインターフェース	
500A	スクリーンショット	
500B	スクリーンショット	10
500C	スクリーンショット	
500D	スクリーンショット	
510	ユーザインターフェース	
512	画像	
514	テキストブロック	
516	テキストブロック	
520	ユーザインターフェース	
523	テキストブロック	
525	テキストブロック	
530	ユーザインターフェース	20
533	命令	
540	ユーザインターフェース	
549	オーバーレイ	
600A	スクリーンショット	
600B	スクリーンショット	
610	ユーザインターフェース	
615	画像	
617	道路標識	
619	道路標識	
620	テキストブロック	30
625	テキストブロック	
640	ユーザインターフェース	
643	オーバーレイ	
645	ユーザインターフェース要素	
647	ユーザインターフェース要素	
700A	スクリーンショット	
700B	スクリーンショット	
710	ユーザインターフェース	
715	画像	
720	テキストブロック	40
725	テキストブロック	
750	ユーザインターフェース	
755	ユーザインターフェース要素	
760	オーバーレイ	
800A	スクリーンショット	
800B	スクリーンショット	
800C	スクリーンショット	
810	ユーザインターフェース	
815	画像	
820	標識	50

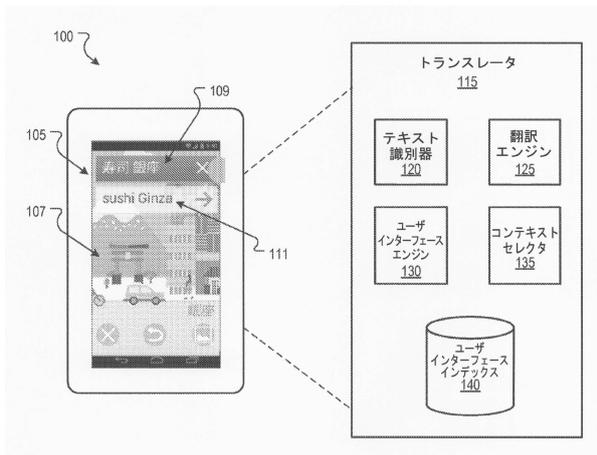
825	テキストブロック	
830	標識	
835	テキストブロック	
840	ユーザインターフェース	
845	オーバーレイ	
847	オーバーレイ	
860	ユーザインターフェース	
900	プロセス	
1000A	スクリーンショット	
1000B	スクリーンショット	10
1000C	スクリーンショット	
1010	ユーザインターフェース	
1012	画像撮影ボタン	
1015	画像	
1020	レストランの名前	
1022	メニューの項目	
1024	価格	
1030	ユーザインターフェース	
1032	ボックス	
1034	ボックス	20
1050	ユーザインターフェース	
1054	翻訳	
1056	ナビゲーションコントロール	
1057	矢印	
1058	矢印	
1059	ユーザインターフェース要素	
1100A	スクリーンショット	
1100B	スクリーンショット	
1100C	スクリーンショット	
1110	ユーザインターフェース	30
1115	画像	
1120	レストランの名前	
1122	メニューの項目	
1124	価格	
1130	ユーザインターフェース	
1132	オーバーレイ	
1134	オーバーレイ	
1150	ユーザインターフェース	
1152	元のテキスト	
1154	翻訳	40
1156	ナビゲーションコントロール	
1200A	スクリーンショット	
1200B	スクリーンショット	
1210	ユーザインターフェース	
1215	画像	
1220	第1の集合	
1221	支柱	
1225	第2の集合	
1240	ユーザインターフェース	
1242	元のテキスト	50

- 1244 翻訳
- 1246 ナビゲーションコントロール
- 1249 ユーザインターフェース要素
- 1300 プロセス
- 1400A スクリーンショット
- 1400B スクリーンショット
- 1400C スクリーンショット
- 1410 ユーザインターフェース
- 1412 画像
- 1414 画像撮影ボタン
- 1414 住所
- 1416 住所
- 1430 ユーザインターフェース
- 1432 ボックス
- 1434 「地図上で見る」コントロール
- 1450 ユーザインターフェース
- 1452 オーバーレイ
- 1454 地図
- 1456 マーカ
- 1458 テキスト、閲覧コントロール
- 1500 プロセス
- 1600 プロセス

10

20

【図1】



【図2】

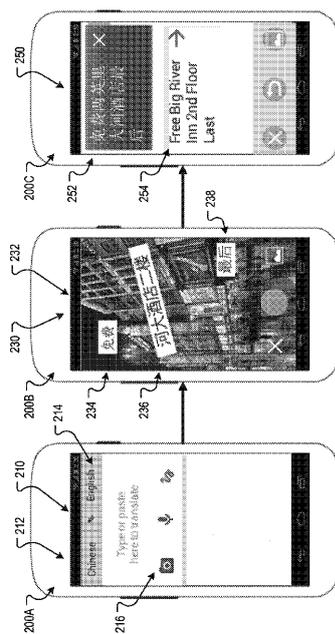


FIG. 2

【 図 3 】

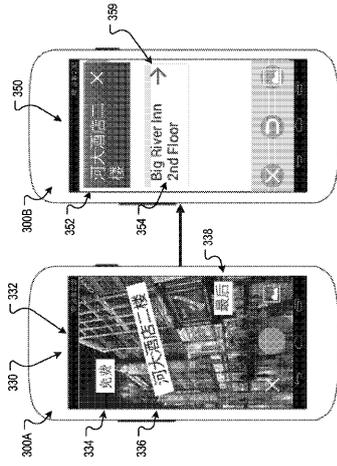


FIG. 3

【 図 4 】

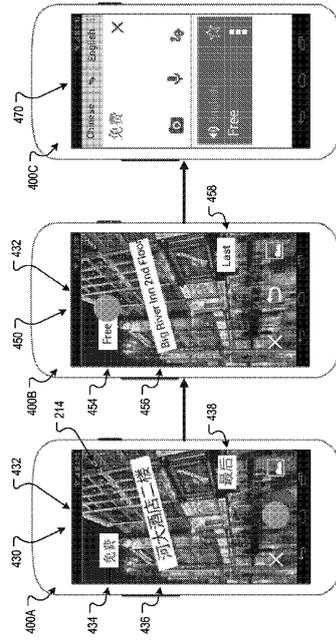


FIG. 4

【 図 5 】

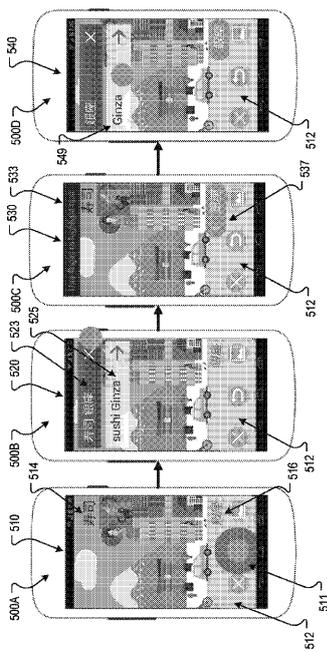


FIG. 5

【 図 6 】

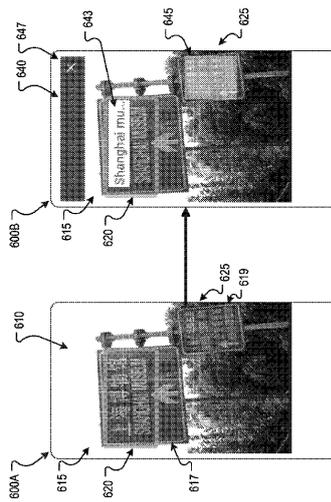


FIG. 6

【図 7】

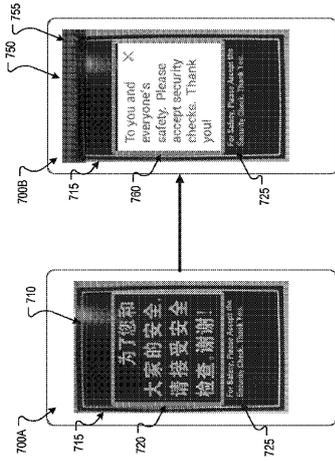


FIG. 7

【図 8】

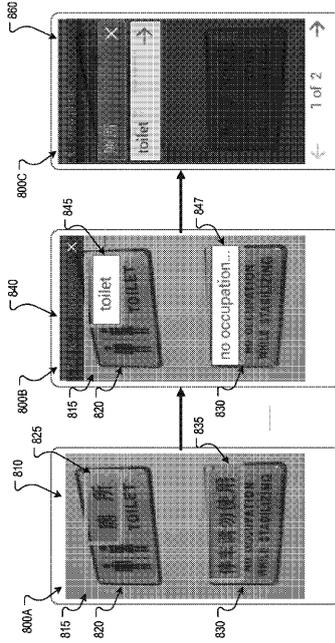
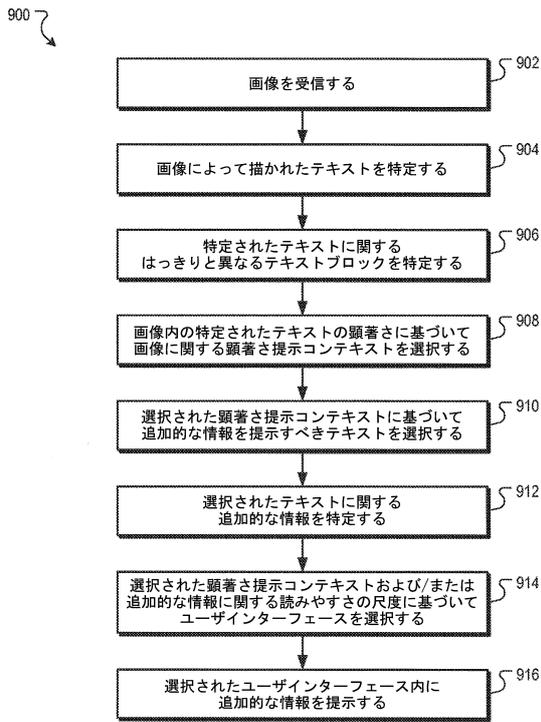


FIG. 8

【図 9】



【図 10】

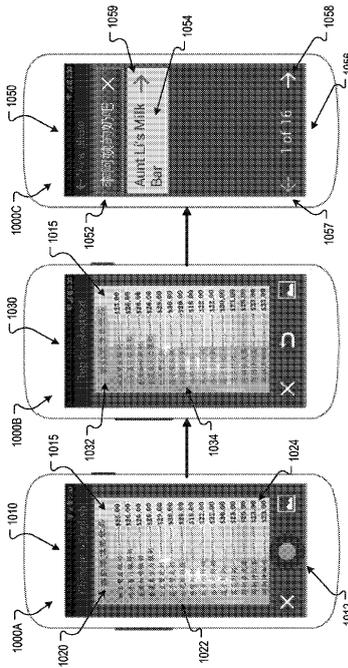


FIG. 10

【図 1 1】

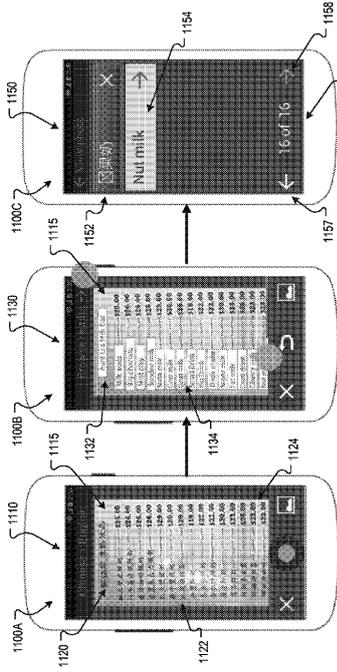


FIG. 11

【図 1 2】

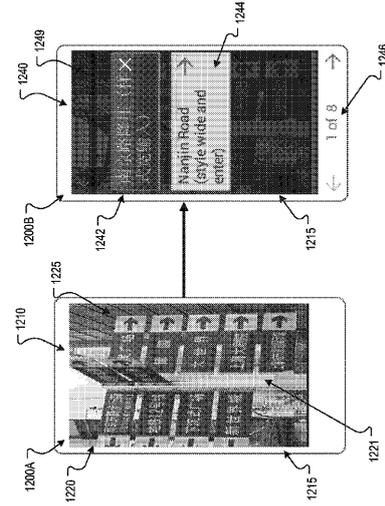
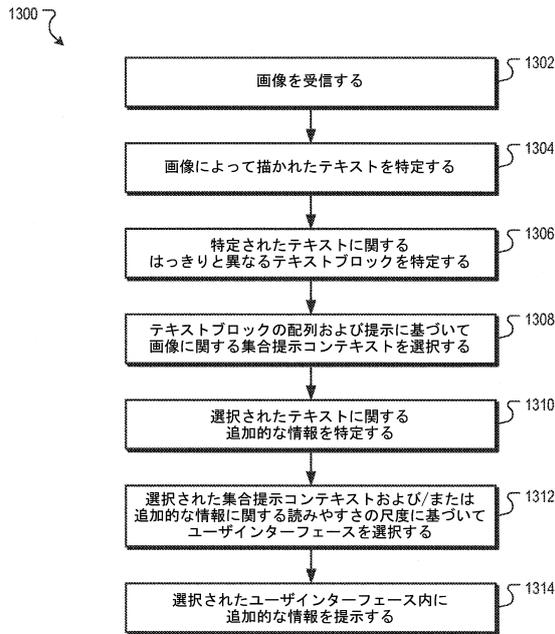


FIG. 12

【図 1 3】



【図 1 4】

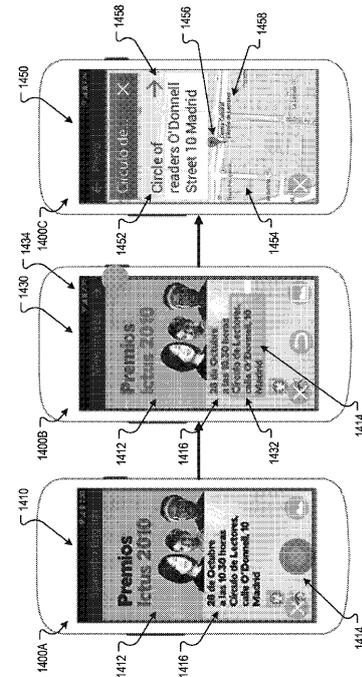
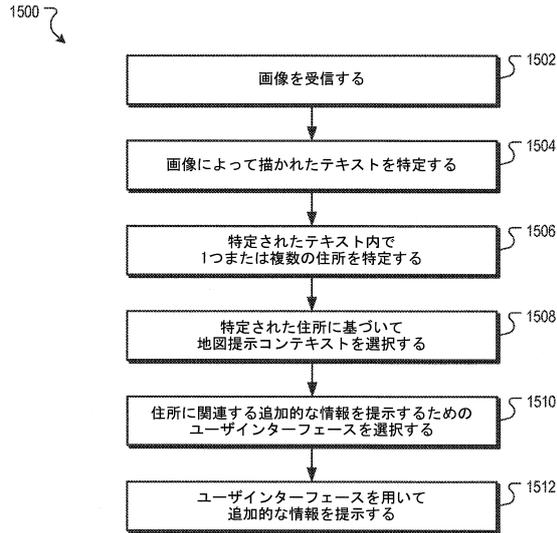
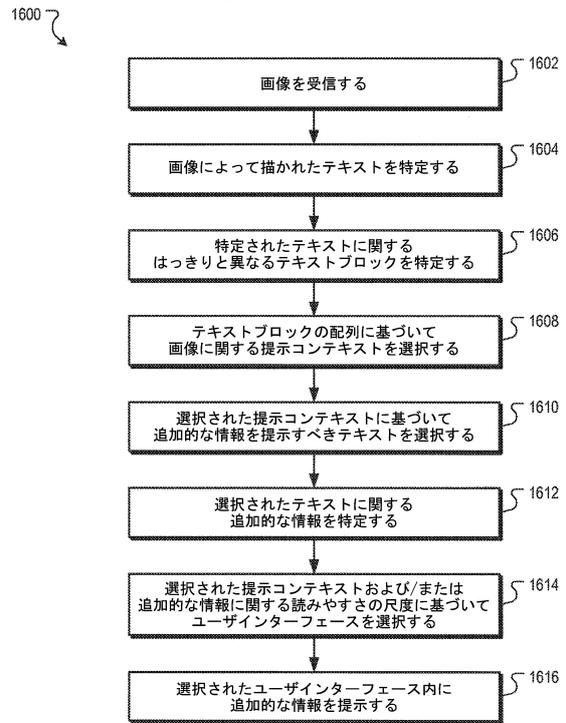


FIG. 14

【図15】



【図16】



## フロントページの続き

- (31)優先権主張番号 14/076,029  
(32)優先日 平成25年11月8日(2013.11.8)  
(33)優先権主張国 米国(US)

## 早期審査対象出願

- (72)発明者 アレクサンダー・ジェイ・カスパート  
アメリカ合衆国・カリフォルニア・94043・マウンテン・ビュー・アンフィシアター・パーク  
ウェイ・1600・グーグル・インコーポレイテッド内  
(72)発明者 ジャシュア・ジェイ・エステル  
アメリカ合衆国・カリフォルニア・94043・マウンテン・ビュー・アンフィシアター・パーク  
ウェイ・1600・グーグル・インコーポレイテッド内

審査官 成瀬 博之

- (56)参考文献 特開平11-120185(JP,A)  
特開2000-132639(JP,A)  
特開2012-008979(JP,A)  
特開2012-079076(JP,A)  
特開平02-249064(JP,A)  
特開2011-065573(JP,A)

- (58)調査した分野(Int.Cl., DB名)  
G06F 17/28 - 17/30  
G06K 9/00 - 9/82