

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2020-521219

(P2020-521219A)

(43) 公表日 令和2年7月16日(2020.7.16)

(51) Int.Cl.  
G06F 16/532 (2019.01)

F I  
G06F 16/532

テーマコード(参考)  
5B175

審査請求 有 予備審査請求 未請求 (全 46 頁)

(21) 出願番号 特願2019-563199 (P2019-563199)  
 (86) (22) 出願日 平成29年9月13日 (2017. 9. 13)  
 (85) 翻訳文提出日 令和1年11月22日 (2019. 11. 22)  
 (86) 国際出願番号 PCT/US2017/051333  
 (87) 国際公開番号 WO2019/054999  
 (87) 国際公開日 平成31年3月21日 (2019. 3. 21)

(71) 出願人 502208397  
 グーグル エルエルシー  
 アメリカ合衆国 カリフォルニア州 94  
 043 マウンテン ビュー アンフィシ  
 アター パークウェイ 1600  
 (74) 代理人 100108453  
 弁理士 村山 靖彦  
 (74) 代理人 100110364  
 弁理士 実広 信哉  
 (74) 代理人 100133400  
 弁理士 阿部 達彦  
 (72) 発明者 チャールズ・ヤン  
 アメリカ合衆国・カリフォルニア・940  
 43・マウンテン・ビュー・アンフィシ  
 アター・パークウェイ・1600

最終頁に続く

(54) 【発明の名称】 関連コンテンツを使用した画像の効率的な増強

(57) 【要約】

本明細書の主題は、一般に、画像に描かれたテキストに関連するコンテンツを提供することに関する。一態様では、システムは、画像からテキストを抽出するように構成されたデータ処理装置を含む。抽出されたテキストは、複数のブロックに分割される。複数のブロックは、第1のズームレベルにおいてユーザインターフェース上にそれぞれの第1のユーザ選択可能なターゲットとして提示される。複数のブロックのうちの第1のブロックのユーザ選択が検出される。第1のブロックのユーザ選択の検出にตอบสนองして、第1のブロック内の抽出されたテキストの部分は、第1のズームレベルより大きい第2のズームレベルにおいて、ユーザインターフェース上にそれぞれの第2のユーザ選択可能なターゲットとして提示される。第1のブロック内の抽出されたテキストの一部のユーザ選択の検出にตอบสนองして、ユーザ選択のテキストのコンテンツに基づいてアクションが開始される。

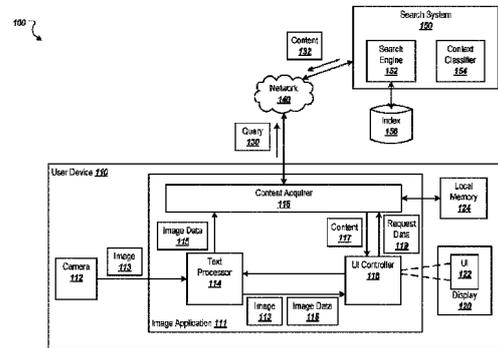


FIG. 1

**【特許請求の範囲】****【請求項1】**

ユーザデバイスによって、キャプチャされた画像からテキストを抽出するステップと、前記ユーザデバイスによって、前記抽出されたテキストを複数のブロックに分割するステップと、

前記ユーザデバイスによって、第1のズームレベルにおいて、前記ユーザデバイスのユーザインターフェース上にそれぞれのユーザ選択可能なターゲットとして前記ブロックを提示するステップと、

前記複数のブロックのうちの第1のブロックのユーザ選択を検出するステップと、

前記第1のブロックの前記ユーザ選択の検出に回答して、前記ユーザデバイスによって、前記第1のズームレベルより大きい第2のズームレベルにおいて、前記第1のブロック内の前記抽出されたテキストの一部を、前記ユーザデバイスの前記ユーザインターフェース上にそれぞれの第2のユーザ選択可能なターゲットとして提示するステップと、

前記第1のブロック内の前記抽出されたテキストの一部のユーザ選択の検出に回答して、前記ユーザデバイスによって、前記ユーザ選択のテキストのコンテンツに基づいてアクションを開始するステップと

を備える、方法。

**【請求項2】**

前記アクションを開始するステップが、前記ユーザ選択のテキストによって識別される宛先アドレスへの通信を開始するステップを備える、請求項1に記載の方法。

**【請求項3】**

前記アクションを開始するステップが、

前記ユーザ選択のテキストの前記コンテンツに基づいて検索クエリを生成するステップと、

前記ユーザデバイスによって、検索エンジンに前記検索クエリを送信するステップとを備える、請求項1に記載の方法。

**【請求項4】**

前記アクションを開始するステップが、

前記ユーザ選択のテキストの前記コンテンツに基づいて、2つ以上の候補検索クエリを生成するステップと、

前記ユーザインターフェースに前記候補検索クエリを表示するステップと、

前記候補クエリのうちの1つのユーザ選択に応じて、前記ユーザデバイスによって、選択された候補クエリを検索エンジンに送信するステップとを備える、請求項1に記載の方法。

**【請求項5】**

前記ユーザデバイスによって、前記画像からコンテキストを抽出するステップをさらに備え、

前記検索クエリまたは前記候補検索クエリを生成するステップが、前記画像の前記抽出されたコンテキストにさらに基づいて、前記検索クエリまたは前記候補検索クエリを生成するステップを備える、請求項3または4に記載の方法。

**【請求項6】**

前記ユーザデバイスによって、前記画像からコンテキストを抽出するステップと、

前記ユーザデバイスによって、前記画像の前記抽出されたコンテキストを使用して、前記画像が複数のあらかじめ定義されたカテゴリの1つに該当するかどうかを決定するステップと

をさらに備え、

前記検索クエリまたは前記候補検索クエリを生成するステップが、前記画像が前記あらかじめ定義されたカテゴリの1つに該当する場合、前記抽出されたコンテキストにさらに基づいて前記検索クエリまたは前記候補検索クエリを生成するステップを備え、該当しない場合、前記検索クエリまたは前記候補検索クエリを生成するステップが、前記ユーザ選

10

20

30

40

50

択のテキストの前記コンテンツのみに基づく、請求項3または4に記載の方法。

【請求項7】

ユーザが前記ブロックのうちのいずれかを選択する前に、  
少なくとも1つのブロックについて、前記ユーザデバイスによって、前記ブロックの前記抽出されたテキストに基づいて1つまたは複数の検索クエリを生成するステップと、  
前記ユーザデバイスから検索エンジンに前記1つまたは複数の検索クエリを送信するステップと、  
をさらに備える、請求項1から6のいずれか一項に記載の方法。

【請求項8】

前記ユーザデバイスが、前記ユーザインターフェース上にユーザ選択可能なそれぞれのターゲットとして前記ブロックを提示する前に、前記1つまたは複数の検索クエリを生成して送信する、請求項7に記載の方法。

10

【請求項9】

前記ユーザデバイスにおいて、前記1つまたは複数の検索クエリに基づいて前記検索エンジンから検索結果コンテンツを受信するステップと、  
前記受信した検索結果コンテンツを前記ユーザデバイスのローカルメモリに記憶するステップと  
をさらに備える、請求項7または8に記載の方法。

【請求項10】

前記ブロックの前記抽出されたテキストに基づいて前記1つまたは複数の検索クエリを生成するステップが、  
前記ブロックの前記抽出されたテキストを、第1のレベルのテキストベースの粒度においてテキストの複数の第1のセットに解析するステップと、  
テキストの第1のセットごとに検索クエリを生成するステップと  
を備える、請求項7、8、または9に記載の方法。

20

【請求項11】

前記ブロックの前記ユーザ選択の検出に応答して、  
前記ブロックの前記抽出されたテキストを、テキストベースの粒度の前記第1のレベルよりも大きい第2のレベルのテキストベースの粒度においてテキストの複数の第2のセットに解析するステップと、  
前記第1のブロック内の前記抽出されたテキストの一部の前記ユーザ選択を受信する前に、  
テキストの第2のセットごとに追加の検索クエリを生成するステップと、  
前記ユーザデバイスから前記検索エンジンに各追加の検索クエリを送信するステップと、  
前記ユーザデバイスによって、および前記検索エンジンから、前記追加の検索クエリに基づく追加の検索結果コンテンツを受信するステップと、  
前記追加の検索結果コンテンツを前記ユーザデバイスのローカルメモリに記憶するステップと  
をさらに備える、請求項10に記載の方法。

30

40

【請求項12】

前記抽出されたテキストを前記ブロックに分割するステップが、前記抽出されたテキストの意味解析に少なくとも部分的に基づき、請求項1から11のいずれか一項に記載の方法。

【請求項13】

前記ユーザインターフェース上に、検索エンジンから前記ユーザデバイスにおいて受信したコンテンツを表示するステップをさらに備える、請求項3から11のいずれか一項、または請求項3から11のいずれか一項に従属する請求項12に記載の方法。

【請求項14】

1つまたは複数のコンピュータによって実行されると、前記1つまたは複数のコンピュー

50

タに、請求項1から13のいずれか一項において定義された方法を実行させる命令でエンコードされたコンピュータストレージ媒体。

【請求項15】

請求項1から13のいずれか一項において定義された方法を実行するように構成された、プロセッサ。

【請求項16】

ユーザデバイスによって、画像からテキストを抽出するステップと、  
前記ユーザデバイスによって、前記抽出されたテキストを複数のブロックに分割するステップと、

前記ユーザデバイスによって、第1のズームレベルにおいて、前記ユーザデバイスのユーザインターフェース上にそれぞれの第1のユーザ選択可能なターゲットとして前記複数のブロックを提示するステップと、

10

前記複数のブロックのうちの第1のブロックのユーザ選択を検出するステップと、

前記第1のブロックの前記ユーザ選択の検出にตอบสนองして、前記ユーザデバイスによって、前記第1のズームレベルより大きい第2のズームレベルにおいて、前記第1のブロック内の前記抽出されたテキストの一部を、前記ユーザデバイスの前記ユーザインターフェース上にそれぞれの第2のユーザ選択可能なターゲットとして提示するステップと、

前記第1のブロック内の前記抽出されたテキストの一部のユーザ選択の検出にตอบสนองして、前記ユーザデバイスによって、前記ユーザ選択のテキストのコンテンツに基づいてアクションを開始するステップと

20

を備える、方法。

【請求項17】

前記アクションを開始するステップが、前記ユーザ選択のテキストによって識別される宛先アドレスへの通信を開始するステップを備える、請求項16に記載の方法。

【請求項18】

前記アクションを開始するステップが、  
前記ユーザ選択のテキストの前記コンテンツに基づいて検索クエリを生成するステップと、

前記ユーザデバイスによって、検索エンジンに前記検索クエリを送信するステップとを備える、請求項16に記載の方法。

30

【請求項19】

前記ユーザデバイスによって、前記画像からコンテキストを抽出するステップと、  
前記画像の前記抽出されたコンテキストを使用して、前記画像が複数のあらかじめ定義されたカテゴリの1つに属すると決定するステップと、

前記画像が前記あらかじめ定義されたカテゴリの1つに属すると決定するステップにตอบสนองして、前記抽出されたコンテキストに基づいて前記検索クエリを調整するステップとをさらに備える、請求項18に記載の方法。

【請求項20】

前記ユーザデバイスによって、および前記検索エンジンから、前記検索クエリに基づいて選択された検索結果コンテンツを受信するステップをさらに備え、前記検索エンジンが、前記画像が属する前記1つのあらかじめ定義されたカテゴリに基づいて検索結果をランク付けする、請求項19に記載の方法。

40

【請求項21】

前記ユーザデバイスによって、前記画像からコンテキストを抽出するステップと、  
前記抽出された前記画像のコンテキストを使用して、前記画像が複数のあらかじめ定義されたカテゴリの1つに属しないと決定するステップと

をさらに備え、

前記検索クエリを生成するステップが、前記ユーザ選択のテキストのみに基づいて一般的な検索クエリを生成するステップを備える、請求項18に記載の方法。

【請求項22】

50

前記アクションを開始するステップが、  
前記ユーザ選択のテキストの前記コンテンツに基づいて、2つ以上の候補検索クエリを生成するステップと、  
前記ユーザインターフェースに前記候補検索クエリを表示するステップと、  
前記候補クエリのうちの1つのユーザ選択に応じて、前記ユーザデバイスによって、前記ユーザ選択の候補クエリを検索エンジンに送信するステップと  
を備える、請求項16に記載の方法。

【請求項23】

ユーザが前記ブロックのうちのいずれかを選択する前に、  
前記第1のブロックについて、前記ユーザデバイスによって、前記第1のブロックの前記抽出されたテキストに基づいて1つまたは複数の検索クエリを生成するステップと、  
前記ユーザデバイスから検索エンジンに前記1つまたは複数の検索クエリを送信するステップと、  
をさらに備える、請求項16に記載の方法。

10

【請求項24】

前記ユーザデバイスが、前記複数のブロックを前記ユーザインターフェース上にそれぞれの第1のユーザ選択可能なターゲットとして提示する前に、前記1つまたは複数の検索クエリを生成して送信する、請求項23に記載の方法。

【請求項25】

前記1つまたは複数の検索クエリに基づいて前記検索エンジンから検索結果コンテンツを受信するステップと、  
前記受信した検索結果コンテンツを前記ユーザデバイスのローカルメモリに記憶するステップと  
をさらに備える、請求項23に記載の方法。

20

【請求項26】

前記第1のブロックの前記抽出されたテキストに基づいて前記1つまたは複数の検索クエリを生成するステップが、  
前記第1のブロックの前記抽出されたテキストを、第1のレベルのテキストベースの粒度においてテキストの複数の第1のセットに解析するステップと、  
テキストの第1のセットごとに検索クエリを生成するステップと  
を備える、請求項23に記載の方法。

30

【請求項27】

前記第1のブロックの前記ユーザ選択の検出にตอบสนองして、  
前記第1のブロックの前記抽出されたテキストを、前記第1のレベルのテキストベースの粒度よりも大きい第2のレベルのテキストベースの粒度においてテキストの複数の第2のセットに解析するステップと、  
前記第1のブロック内の前記抽出されたテキストの一部の前記ユーザ選択を受信する前に、

テキストの第2のセットごとに追加の検索クエリを生成するステップと、  
前記ユーザデバイスから前記検索エンジンに各追加の検索クエリを送信するステップと、

40

前記ユーザデバイスによって、および前記検索エンジンから、前記追加の検索クエリに基づく追加の検索結果コンテンツを受信するステップと、

前記追加の検索結果コンテンツを前記ユーザデバイスのローカルメモリに記憶するステップと  
をさらに備える、請求項26に記載の方法。

【請求項28】

前記抽出されたテキストを前記ブロックに分割するステップが、前記抽出されたテキストの意味解析に少なくとも部分的に基づく、請求項16に記載の方法。

【請求項29】

50

前記ユーザインターフェース上に、検索エンジンから前記ユーザデバイスにおいて受信したコンテンツを表示するステップをさらに備える、請求項16に記載の方法。

【請求項30】

前記ユーザインターフェース上に、前記検索エンジンから前記ユーザデバイスにおいて受信したコンテンツを表示するステップが、前記画像の少なくとも一部とともに前記コンテンツを前記ユーザデバイスのビューポートに表示するステップを備える、請求項29に記載の方法。

【請求項31】

データ処理装置と、  
前記データ処理装置とデータ通信し、前記データ処理装置によって実行可能な命令を記憶するメモリ装置であって、前記命令が、そのような実行時に、前記データ処理装置に、  
画像からテキストを抽出することと、  
前記抽出されたテキストを複数のブロックに分割することと、  
第1のズームレベルにおいて、ユーザインターフェース上にそれぞれの第1のユーザ選択可能なターゲットとして前記複数のブロックを提示することと、  
前記複数のブロックのうちの第1のブロックのユーザ選択を検出することと、  
前記第1のブロックの前記ユーザ選択の検出に応答して、前記第1のズームレベルより大きい第2のズームレベルにおいて、前記第1のブロック内の前記抽出されたテキストの一部を、前記ユーザインターフェース上にそれぞれの第2のユーザ選択可能なターゲットとして提示することと、  
前記第1のブロック内の前記抽出されたテキストの一部のユーザ選択の検出に応答して、前記ユーザ選択のテキストのコンテンツに基づいてアクションを開始することと  
を備える動作を実行させる、メモリ装置と  
を備える、システム。

10

20

【請求項32】

前記アクションを開始することが、前記ユーザ選択のテキストによって識別される宛先アドレスへの通信を開始することを備える、請求項31に記載のシステム。

【請求項33】

前記アクションを開始することが、  
前記ユーザ選択のテキストの前記コンテンツに基づいて検索クエリを生成することと、  
検索エンジンに前記検索クエリを送信することと  
を備える、請求項31に記載のシステム。

30

【請求項34】

前記動作が、  
前記画像からコンテキストを抽出することと、  
前記画像の前記抽出されたコンテキストを使用して、前記画像が複数のあらかじめ定義されたカテゴリの1つに属すると決定することと、  
前記画像が前記あらかじめ定義されたカテゴリの1つに属すると決定することに対応して、前記抽出されたコンテキストに基づいて前記検索クエリを調整することと  
をさらに備える、請求項33に記載のシステム。

40

【請求項35】

前記動作が、前記検索エンジンから、前記検索クエリに基づいて選択された検索結果コンテンツを受信することをさらに備え、前記検索エンジンが、前記画像が属する前記1つのあらかじめ定義されたカテゴリに基づいて検索結果をランク付けする、請求項34に記載のシステム。

【請求項36】

前記動作が、  
前記画像からコンテキストを抽出することと、  
前記抽出された前記画像のコンテキストを使用して、前記画像が複数のあらかじめ定義されたカテゴリの1つに属しないと決定することと

50

をさらに備え、

前記検索クエリを生成することが、前記ユーザ選択のテキストのみに基づいて一般的な検索クエリを生成することを備える、請求項33に記載のシステム。

【請求項37】

前記アクションを開始することが、

前記ユーザ選択のテキストの前記コンテンツに基づいて、2つ以上の候補検索クエリを生成することと、

前記ユーザインターフェースに前記候補検索クエリを表示することと、

前記候補クエリのうちの1つのユーザ選択に応じて、選択された候補クエリを検索エンジンに送信することと

10

を備える、請求項31に記載のシステム。

【請求項38】

前記動作が、ユーザが前記ブロックのうちのいずれかを選択する前に、

前記第1のブロックについて、前記第1のブロックの前記抽出されたテキストに基づいて1つまたは複数の検索クエリを生成することと、

検索エンジンに前記1つまたは複数の検索クエリを送信することと、

をさらに備える、請求項31に記載のシステム。

【請求項39】

前記データ処理装置が、前記複数のブロックを前記ユーザインターフェース上にそれぞれの第1のユーザ選択可能なターゲットとして提示する前に、前記1つまたは複数の検索クエリを生成して送信する、請求項38に記載のシステム。

20

【請求項40】

前記動作が、

前記1つまたは複数の検索クエリに基づいて前記検索エンジンから検索結果コンテンツを受信することと、

前記受信した検索結果コンテンツを前記データ処理装置のローカルメモリに記憶することと

をさらに備える、請求項38に記載のシステム。

【請求項41】

前記第1のブロックの前記抽出されたテキストに基づいて前記1つまたは複数の検索クエリを生成することが、

30

前記第1のブロックの前記抽出されたテキストを、第1のレベルのテキストベースの粒度においてテキストの複数の第1のセットに解析することと、

テキストの第1のセットごとに検索クエリを生成することと

を備える、請求項38に記載のシステム。

【請求項42】

前記動作が、

前記第1のブロックの前記ユーザ選択の検出に回答して、

前記第1のブロックの前記抽出されたテキストを、前記第1のレベルのテキストベースの粒度よりも大きい第2のレベルのテキストベースの粒度においてテキストの複数の第2のセットに解析することと、

40

前記第1のブロック内の前記抽出されたテキストの一部の前記ユーザ選択を受信する前に、

テキストの第2のセットごとに追加の検索クエリを生成することと、

前記検索エンジンに各追加の検索クエリを送信することと、

前記検索エンジンから、前記追加の検索クエリに基づく追加の検索結果コンテンツを受信することと、

前記追加の検索結果コンテンツを前記データ処理装置のローカルメモリに記憶することと

をさらに備える、請求項41に記載のシステム。

50

## 【請求項 4 3】

前記抽出されたテキストを前記ブロックに分割することが、前記抽出されたテキストの意味解析に少なくとも部分的に基づき、請求項31に記載のシステム。

## 【請求項 4 4】

前記動作が、検索エンジンから受信したコンテンツを前記ユーザインターフェース上に表示することをさらに備える、請求項31に記載のシステム。

## 【請求項 4 5】

前記検索エンジンから受信したコンテンツを前記ユーザインターフェース上に表示することが、前記画像の少なくとも一部とともに前記データ処理装置のビューポートに前記コンテンツを表示することを備える、請求項44に記載のシステム。

10

## 【請求項 4 6】

データ処理装置によって実行されると、前記データ処理装置に、  
画像からテキストを抽出することと、  
前記抽出されたテキストを複数のブロックに分割することと、  
第1のズームレベルにおいて、ユーザインターフェース上にそれぞれの第1のユーザ選択可能なターゲットとして前記複数のブロックを提示することと、  
前記複数のブロックのうちの第1のブロックのユーザ選択を検出することと、  
前記第1のブロックの前記ユーザ選択の検出にตอบสนองして、前記第1のズームレベルより大きい第2のズームレベルにおいて、前記第1のブロック内の前記抽出されたテキストの一部を、前記ユーザインターフェース上にそれぞれの第2のユーザ選択可能なターゲットとして提示することと、  
前記第1のブロック内の前記抽出されたテキストの一部のユーザ選択の検出にตอบสนองして、前記ユーザ選択のテキストのコンテンツに基づいてアクションを開始することと  
を備える動作を実行させる命令を記憶した、非一時的コンピュータ可読ストレージ媒体。

20

## 【請求項 4 7】

前記アクションを開始することが、前記ユーザ選択のテキストによって識別される宛先アドレスへの通信を開始することを備える、請求項46に記載の非一時的コンピュータ可読ストレージ媒体。

## 【請求項 4 8】

前記アクションを開始することが、  
前記ユーザ選択のテキストの前記コンテンツに基づいて検索クエリを生成するステップと、  
検索エンジンに前記検索クエリを送信することと  
を備える、請求項46に記載の非一時的コンピュータ可読ストレージ媒体。

30

## 【請求項 4 9】

前記動作が、  
前記画像からコンテキストを抽出することと、  
前記画像の前記抽出されたコンテキストを使用して、前記画像が複数のあらかじめ定義されたカテゴリの1つに属すると決定することと、  
前記画像が前記あらかじめ定義されたカテゴリの1つに属すると決定することにตอบสนองして、前記抽出されたコンテキストに基づいて前記検索クエリを調整することと  
をさらに備える、請求項48に記載の非一時的コンピュータ可読ストレージ媒体。

40

## 【請求項 5 0】

前記動作が、前記検索エンジンから、前記検索クエリに基づいて選択された検索結果コンテンツを受信することをさらに備え、前記検索エンジンが、前記画像が属する前記1つのあらかじめ定義されたカテゴリに基づいて検索結果をランク付けする、請求項49に記載の非一時的コンピュータ可読ストレージ媒体。

## 【請求項 5 1】

前記動作が、  
前記画像からコンテキストを抽出することと、

50

前記抽出された前記画像のコンテキストを使用して、前記画像が複数のあらかじめ定義されたカテゴリの1つに属しないと決定することと

をさらに備え、

前記検索クエリを生成することが、前記ユーザ選択のテキストのみに基づいて一般的な検索クエリを生成することを備える、請求項48に記載の非一時的コンピュータ可読ストレージ媒体。

【請求項52】

前記アクションを開始することが、

前記ユーザ選択のテキストの前記コンテンツに基づいて、2つ以上の候補検索クエリを生成することと、

前記ユーザインターフェースに前記候補検索クエリを表示することと、

前記候補クエリのうちの1つのユーザ選択に応じて、前記ユーザ選択の候補クエリを検索エンジンに送信することと

を備える、請求項46に記載の非一時的コンピュータ可読ストレージ媒体。

【請求項53】

前記動作が、ユーザが前記ブロックのうちのいずれかを選択する前に、

前記第1のブロックについて、前記第1のブロックの前記抽出されたテキストに基づいて1つまたは複数の検索クエリを生成することと、

検索エンジンに前記1つまたは複数の検索クエリを送信することと、

をさらに備える、請求項46に記載の非一時的コンピュータ可読ストレージ媒体。

【請求項54】

前記データ処理装置が、前記複数のブロックを前記ユーザインターフェース上にそれぞれの第1のユーザ選択可能なターゲットとして提示する前に、前記1つまたは複数の検索クエリを生成して送信する、請求項53に記載の非一時的コンピュータ可読ストレージ媒体。

【請求項55】

前記動作が、

前記1つまたは複数の検索クエリに基づいて前記検索エンジンから検索結果コンテンツを受信することと、

前記受信した検索結果コンテンツを前記データ処理装置のローカルメモリに記憶することと

をさらに備える、請求項53に記載の非一時的コンピュータ可読ストレージ媒体。

【請求項56】

前記第1のブロックの前記抽出されたテキストに基づいて前記1つまたは複数の検索クエリを生成することが、

前記第1のブロックの前記抽出されたテキストを、第1のレベルのテキストベースの粒度においてテキストの複数の第1のセットに解析することと、

テキストの第1のセットごとに検索クエリを生成することと

を備える、請求項53に記載の非一時的コンピュータ可読ストレージ媒体。

【請求項57】

前記動作が、

前記第1のブロックの前記ユーザ選択の検出に応答して、

前記第1のブロックの前記抽出されたテキストを、前記第1のレベルのテキストベースの粒度よりも大きい第2のレベルのテキストベースの粒度においてテキストの複数の第2のセットに解析することと、

前記第1のブロック内の前記抽出されたテキストの一部の前記ユーザ選択を受信する前に、

テキストの第2のセットごとに追加の検索クエリを生成することと、

前記検索エンジンに各追加の検索クエリを送信することと、

前記検索エンジンから、前記追加の検索クエリに基づく追加の検索結果コンテンツを受信することと、

10

20

30

40

50

前記追加の検索結果コンテンツを前記データ処理装置のローカルメモリに記憶することと

をさらに備える、請求項56に記載の非一時的コンピュータ可読ストレージ媒体。

【請求項58】

前記抽出されたテキストを前記ブロックに分割することが、前記抽出されたテキストの意味解析に少なくとも部分的に基づき、請求項46に記載の非一時的コンピュータ可読ストレージ媒体。

【請求項59】

前記動作が、検索エンジンから受信したコンテンツを前記ユーザインターフェース上に表示することをさらに備える、請求項46に記載の非一時的コンピュータ可読ストレージ媒体。

10

【請求項60】

前記検索エンジンから受信したコンテンツを前記ユーザインターフェース上に表示することが、前記画像の少なくとも一部とともに前記データ処理装置のビューポートに前記コンテンツを表示することを備える、請求項59に記載の非一時的コンピュータ可読ストレージ媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、関連コンテンツを使用した画像の効率的な増強に関する。

20

【背景技術】

【0002】

インターネットは、画像ファイル、オーディオファイル、ビデオファイル、およびウェブページなどの様々なリソースへのアクセスを提供する。検索システムは、クエリに応じてリソースを識別することができる。クエリは、1つまたは複数の検索用語またはフレーズを含むテキストクエリ、画像を含む画像クエリ、またはテキストクエリと画像クエリとの組合せであり得る。検索システムはリソースをランク付けし、識別されたりソースにリンクしてもよく、クエリに関連するコンテンツを提供してもよい検索結果を提供する。

【0003】

カメラおよびカメラファーストアプリケーションを備えたスマートフォンが広く普及していることからわかるように、画像(たとえば、デジタル写真)は、ユーザが情報を収集し、アクションを起こすための一般的な方法になった。たとえば、自分のスマートフォンまたは他のデバイスで写真を撮影したユーザは、写真のコンテンツに関するさらなる情報を取得したい場合がある。場合によっては、ユーザが写真において識別したコンテンツに基づいてクエリを入力するために異なるアプリケーションを開く必要があり得るが、これは面倒であり、不正確または無効なクエリをもたらす可能性がある。

30

【発明の概要】

【課題を解決するための手段】

【0004】

本明細書は、画像に描かれたテキストに関連するコンテンツを提示するためのシステム、方法、デバイス、および技法について説明する。

40

【0005】

一般に、本明細書に記載される主題の革新的な一態様は、ユーザデバイスによって、キャプチャされた画像からテキストを抽出するステップを含む方法において実装することができる。ユーザデバイスは、抽出されたテキストを複数のブロックに分割することができる。ユーザデバイスは、第1のズームレベルにおいて、ユーザデバイスのユーザインターフェース上にそれぞれのユーザ選択可能なターゲットとしてブロックを提示することができる。複数のブロックのうちの第1のブロックのユーザ選択を検出することができる。第1のブロックのユーザ選択の検出にตอบสนองして、ユーザデバイスは、第1のズームレベルより大きい第2のズームレベルにおいて、第1のブロック内の抽出されたテキストの一部を、ユ

50

ーザデバイスのユーザインターフェース上にそれぞれの第2のユーザ選択可能なターゲットとして提示することができる。第1のブロック内の抽出されたテキストの一部のユーザ選択の検出に応答して、ユーザデバイスは、ユーザ選択のテキストのコンテンツに基づいてアクションを開始することができる。この態様の他の実装形態は、コンピュータストレージデバイス上にエンコードされた方法のアクションを実行するように構成された、対応する装置、方法、およびコンピュータプログラムを含む。

【0006】

これらおよび他の実装形態は、任意で次の機能のうちの1つまたは複数を含むことができる。いくつかの態様では、アクションを開始するステップは、ユーザ選択のテキストによって識別される宛先アドレスへの通信を開始するステップを含むことができる。

10

【0007】

アクションを開始するステップは、ユーザ選択のテキストのコンテンツに基づいて検索クエリを生成するステップと、ユーザデバイスによって検索エンジンに検索クエリを送信するステップとを含む。アクションを開始するステップは、ユーザ選択のテキストのコンテンツに基づいて、2つ以上の候補検索クエリを生成するステップと、ユーザインターフェースに候補検索クエリを表示するステップと、候補クエリのうちの1つのユーザ選択に応じて、ユーザデバイスによって、選択された候補クエリを検索エンジンに送信するステップとを含むことができる。

【0008】

いくつかの態様は、ユーザデバイスによって、画像からコンテキストを抽出するステップを含むことができる。検索クエリまたは候補検索クエリを生成するステップは、画像の抽出されたコンテキストにさらに基づいて、検索クエリまたは候補検索クエリを生成するステップを含むことができる。

20

【0009】

いくつかの態様は、ユーザデバイスによって、画像からコンテキストを抽出するステップと、ユーザデバイスによって、画像の抽出されたコンテキストを使用して、画像が多数のあらかじめ定義されたカテゴリのいずれかに該当するかどうかを決定するステップとを含み得る。検索クエリまたは候補検索クエリを生成するステップは、画像があらかじめ定義されたカテゴリのいずれかに該当する場合、抽出されたコンテキストに基づいて検索クエリまたは候補検索クエリを生成するステップを含むことができ、該当しない場合、検索クエリまたは候補検索クエリを生成するステップは、ユーザ選択のテキストのコンテンツのみに基づく。

30

【0010】

いくつかの態様では、ユーザデバイスは、少なくとも1つのブロックについて、またユーザがブロックのうちのいずれかを選択する前に、ブロックの抽出されたテキストに基づいて1つまたは複数の検索クエリを生成することができる。ユーザデバイスは、ユーザデバイスから検索エンジンに1つまたは複数の検索クエリを送信することができる。ユーザデバイスは、ユーザインターフェース上にユーザ選択可能なそれぞれのターゲットとしてブロックを提示する前に、1つまたは複数の検索クエリを生成して送信することができる。

40

【0011】

いくつかの態様では、ユーザデバイスは、1つまたは複数の検索クエリに基づいて検索エンジンから検索結果コンテンツを受信し、受信した検索結果コンテンツをユーザデバイスのローカルメモリに記憶することができる。ブロックの抽出されたテキストに基づいて1つまたは複数の検索クエリを生成するステップは、ブロックの抽出されたテキストを、第1のレベルのテキストベースの粒度においてテキストの複数の第1のセットに解析するステップと、テキストの第1のセットごとに検索クエリを生成するステップとを含むことができる。

【0012】

いくつかの態様は、ブロックのユーザ選択の検出に応答して、ブロックの抽出されたテ

50

キストを、第1のレベルのテキストベースの粒度よりも大きい第2のレベルのテキストベースの粒度においてテキストの複数の第2のセットに解析するステップと、第1のブロック内の抽出されたテキストの一部のユーザ選択を受信する前に、テキストの第2のセットごとに追加の検索クエリを生成するステップと、ユーザデバイスから検索エンジンに各追加の検索クエリを送信するステップと、ユーザデバイスによって、および検索エンジンから、追加の検索クエリに基づく追加の検索結果コンテンツを受信するステップと、追加の検索結果コンテンツをユーザデバイスのローカルメモリに記憶するステップとを含むことができる。

【0013】

いくつかの態様では、抽出されたテキストをブロックに分割するステップは、抽出されたテキストの意味解析に少なくとも部分的に基づく。いくつかの態様は、ユーザインターフェース上に、検索エンジンからユーザデバイスにおいて受信したコンテンツを表示するステップを含むことができる。

10

【0014】

一般に、本明細書に記載される主題の革新的な別の態様は、ユーザデバイスによって、画像からテキストを抽出するステップを含む方法において実装することができる。ユーザデバイスは、抽出されたテキストを複数のブロックに分割することができる。ユーザデバイスは、第1のズームレベルにおいて、ユーザデバイスのユーザインターフェース上にそれぞれの第1のユーザ選択可能なターゲットとして複数のブロックを提示することができる。複数のブロックのうちの第1のブロックのユーザ選択を検出することができる。第1のブロックのユーザ選択の検出にตอบสนองして、ユーザデバイスは、第1のズームレベルより大きい第2のズームレベルにおいて、第1のブロック内の抽出されたテキストの一部を、ユーザデバイスのユーザインターフェース上にそれぞれの第2のユーザ選択可能なターゲットとして提示することができる。第1のブロック内の抽出されたテキストの一部のユーザ選択の検出にตอบสนองして、ユーザデバイスは、ユーザ選択のテキストのコンテンツに基づいてアクションを開始することができる。

20

【0015】

いくつかの態様では、アクションを開始するステップは、ユーザ選択のテキストによって識別される宛先アドレスへの通信を開始するステップを含むことができる。アクションを開始するステップは、ユーザ選択のテキストのコンテンツに基づいて検索クエリを生成するステップと、検索エンジンに検索クエリを送信するステップとを含む。

30

【0016】

いくつかの態様は、ユーザデバイスによって、画像からコンテキストを抽出するステップと、画像の抽出されたコンテキストを使用して、画像が多数のあらかじめ定義されたカテゴリの1つに属すると決定するステップと、画像があらかじめ定義されたカテゴリの1つに属すると決定するステップにตอบสนองして、抽出されたコンテキストに基づいて検索クエリを調整するステップとを含む。

【0017】

いくつかの態様は、ユーザデバイスによって、および検索エンジンから、検索クエリに基づいて選択された検索結果コンテンツを受信するステップを含むことができる。検索エンジンは、画像が属する1つのあらかじめ定義されたカテゴリに基づいて検索結果をランク付けすることができる。いくつかの態様は、ユーザデバイスによって、画像からコンテキストを抽出するステップと、抽出された画像のコンテキストを使用して、画像が多数のあらかじめ定義されたカテゴリの1つに属しないと決定するステップとを含むことができる。検索クエリを生成するステップは、ユーザ選択のテキストのみに基づいて一般的な検索クエリを生成するステップを含むことができる。

40

【0018】

いくつかの態様では、アクションを開始するステップは、ユーザ選択のテキストのコンテンツに基づいて、2つ以上の候補検索クエリを生成するステップと、ユーザインターフェース上に候補検索クエリを表示するステップと、候補クエリのうちの1つのユーザ選択に

50

応じて、ユーザデバイスによって、ユーザ選択の候補クエリを検索エンジンに送信するステップとを含むことができる。

【0019】

いくつかの態様は、ユーザがブロックのうちのいずれかを選択する前に、また第1のブロックについて、ユーザデバイスによって、第1のブロックの抽出されたテキストに基づいて1つまたは複数の検索クエリを生成するステップと、1つまたは複数の検索クエリをユーザデバイスから検索エンジンに送信するステップとを含むことができる。

【0020】

いくつかの態様では、ユーザデバイスは、複数のブロックをユーザインターフェース上にそれぞれの第1のユーザ選択可能なターゲットとして提示する前に、1つまたは複数の検索クエリを生成して送信する。いくつかの態様は、1つまたは複数の検索クエリに基づいて検索エンジンから検索結果コンテンツを受信するステップと、受信した検索結果コンテンツをユーザデバイスのローカルメモリに記憶するステップとを含むことができる。

10

【0021】

いくつかの態様では、第1のブロックの抽出されたテキストに基づいて1つまたは複数の検索クエリを生成するステップは、第1のブロックの抽出されたテキストを、第1のレベルのテキストベースの粒度においてテキストの複数の第1のセットに解析するステップと、テキストの第1のセットごとに検索クエリを生成するステップとを含むことができる。

【0022】

いくつかの態様は、第1のブロックのユーザ選択の検出に 응답して、第1のブロックの抽出されたテキストを、第1のレベルのテキストベースの粒度よりも大きい第2のレベルのテキストベースの粒度においてテキストの複数の第2のセットに解析するステップと、第1のブロック内の抽出されたテキストの一部のユーザ選択を受信する前に、テキストの第2のセットごとに追加の検索クエリを生成するステップと、ユーザデバイスから検索エンジンに各追加の検索クエリを送信するステップと、ユーザデバイスによって、および検索エンジンから、追加の検索クエリに基づく追加の検索結果コンテンツを受信するステップと、追加の検索結果コンテンツをユーザデバイスのローカルメモリに記憶するステップとを含むことができる。

20

【0023】

いくつかの態様では、抽出されたテキストをブロックに分割するステップは、抽出されたテキストの意味解析に少なくとも部分的に基づく。いくつかの態様は、ユーザインターフェース上に、検索エンジンからユーザデバイスにおいて受信したコンテンツを表示するステップを含むことができる。ユーザインターフェース上に、検索エンジンからユーザデバイスにおいて受信したコンテンツを表示するステップは、画像の少なくとも一部とともにコンテンツをユーザデバイスのビューポートに表示するステップを含むことができる。

30

【0024】

物理的な文書からテキストを抽出する(特に、スキャンされた文書を介して)ために、OCR(光学文字認識)を使用することが提案されている。しかしながら、抽出されたテキストは通常、テキストのブロックとして扱われる。対照的に、本開示の態様は、画像内のすべての(または、いくつかの)単語、フレーズ、および/またはテキストのブロック(たとえば、電話番号などの数字を含む)をそれぞれのユーザ選択可能なターゲットにする。ユーザは、テキスト項目を含む/備えるブロックを選択することによって、画像内のテキスト項目に基づいてクエリを起動することによって、画像のコンテンツに関するより多くの情報を取得することができる。クエリによって返されたコンテンツは、ユーザデバイスのユーザインターフェースに表示される。クエリを開始するためにユーザが画像からテキストを検索エンジンに手動で入力する必要がある従来のプロセスと比較して、本発明の態様は、ユーザが検索語を手動で入力する際にエラーを起こし、検索語を再入力しなければならない可能性を排除する。これにより、誤って入力されたクエリに検索エンジンが応答し、ユーザがエラーを認識して訂正されたクエリを再送信するときにクエリと取得プロセスを繰り返す必要がある可能性を避けるため、ネットワークと検索エンジンのリソースをより効

40

50

率的に使用できるようになる。

【0025】

本明細書に記載される主題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することができる。ユーザが手動でテキストクエリを入力することを必要とするのではなく、画像内に描かれ、ユーザによって選択されたテキストに基づいてコンテンツを識別する(たとえば、クエリを介して)ことによって、ネットワークを通過する提出されたクエリ数を減らすことにより、検索エンジンによって処理される検索クエリ数を減らすことにより、および検索クエリに回答してネットワークを介して提供される検索結果の数を減らすことにより、ネットワークおよび検索エンジンリソースをより効率的に使用できるようになる。たとえば、ユーザが画像内に見えるテキストをクエリとして誤って入力することもあり、正しいクエリがユーザの情報ニーズを満たす検索結果をもたらすまで、1つまたは複数のクエリ補正と後続のネットワーク通信および検索エンジン処理とを必要とする画像のコンテンツを正確に反映しないクエリを提出することもある。これらの追加のクエリはネットワークと検索エンジンのリソースを消費するため、他のネットワークトラフィックおよび他のクエリを処理するために使用できるネットワークと検索エンジンの帯域幅が少なくなり、ならびに/または検索エンジンの多くのユーザにわたってスケールアップした場合は特に検索エンジンの応答時間が長くなる。したがって、ユーザの情報ニーズを正確に反映するテキストのユーザ選択に応じてコンテンツを識別すると、ネットワーク帯域幅の増加、ネットワークおよび検索エンジンの応答時間の短縮、および/または検索エンジンによって処理されるクエリの増加につながる。

10

20

【0026】

ユーザがテキストを再入力するのではなくテキストを選択できるようにすることも、ユーザにとってより効率的であり、ユーザがテキストに関連するコンテンツをより迅速に要求することを可能にする。また、これにより、テキストの異なる部分に対して異なるターゲットを単に選択することによって、ユーザが複数のクエリを提出することがより容易になる。

【0027】

本明細書で説明するズーム技法により、ユーザは、ユーザが追加コンテンツを受信したいテキストを迅速かつ効率的に選択することが可能になる。たとえば、ユーザデバイスが低ズームレベル(たとえば、テキストが最小である最低ズームレベル)にあるときにテキストの部分ごとにユーザ選択可能なターゲットを作成すると、ユーザは、ユーザが選択しようとしたテキストに隣接するテキストなどの間違っただテキストを選択する可能性がある。これにより、意図しないクエリがネットワークを介して検索エンジンに送信される可能性があり、上記の欠点があるが、いずれもユーザのフラストレーションを引き起こす可能性がある。適切なズームレベルにおいてテキストのブロックのユーザ選択可能なターゲットを作成し、次いでより高いズームレベルにおいてテキストのより小さい部分のユーザ選択可能なターゲットを作成することによって、ユーザは正しいテキストをより容易に選択できる。ズーム技法により、ネットワークを介して検索エンジンに送信される意図しないクエリ数が減少するため、ネットワーク帯域幅の増加、ネットワーク応答時間の短縮、検索エンジンによって処理されるクエリの増加、および検索エンジンの応答時間の短縮をもたらすことができる。

30

40

【0028】

画像のコンテンツは、ズームレベル、および/またはそのズームレベルにおいてユーザデバイスのビューポートに提示される画像内のテキストに基づいて、ユーザデバイスにおいてプリキャッシュすることができる。ズームレベルが変更されると、更新されたズームレベルおよび/または更新されたズームレベルにおいてビューポートに提示されるテキストに基づいて、追加の(および/または代替の)コンテンツをプリキャッシュすることができる。このようにして、テキストのユーザ選択に応じて、画像内のテキストに関連するコンテンツをより迅速に提示することができる。

【0029】

50

前述の主題の様々な特徴および利点は、図面に関して以下で説明される。追加の特徴および利点は、本明細書に記載の主題および特許請求の範囲から明らかである。

【0030】

誤解を避けるために、本明細書で使用される「テキスト」は、たとえば、電話番号または電子メールアドレスがテキストのブロックを構成し得るように、文字、数字、および/または記号を包含することが意図される。

【図面の簡単な説明】

【0031】

【図1】ユーザデバイスが画像に描かれたテキストに関連するコンテンツを提示する例示的な環境を示す図である。

10

【図2】ユーザ選択のテキストに関連する追加コンテンツを要求するために、ユーザが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例示的なスクリーンショットを示す図である。

【図3】ユーザがテキストを選択し、選択されたテキストに関連するコンテンツを閲覧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例示的なスクリーンショットを示す図である。

【図4】ユーザがテキストを選択し、選択されたテキストに関連するコンテンツを閲覧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例示的なスクリーンショットを示す図である。

【図5】ユーザ選択のテキストのコンテンツに基づいてアクションを開始するための例示的なプロセスのフローチャートである。

20

【図6】検索クエリに応答してコンテンツを提供するための例示的なプロセスのフローチャートである。

【発明を実施するための形態】

【0032】

一般に、本明細書で説明されるシステムおよび技法は、画像に描かれたテキストおよび/または他のコンテンツに関連する追加コンテンツを識別および提供する。ユーザは、画像をキャプチャするために、スマートフォンまたは他のユーザデバイスのカメラを使用することができる。次いで、追加コンテンツを識別するために、画像内のテキストを使用することができる。たとえば、画像にレストランのメニューが描かれている場合がある。ユーザデバイスは、ユーザがテキストを選択し、選択したテキストに関連する追加コンテンツを閲覧することを可能にするために、テキストまたはその周辺にユーザ選択可能なターゲットを提示することができる。たとえば、ユーザは、メニュー項目を選択し、メニュー項目に関連するコンテンツ(たとえば、栄養情報)、メニュー項目の画像、またはメニュー項目に関連する他の種類のコンテンツを閲覧し得る。追加コンテンツは、以下で説明するプリキャッシュ技法を使用して、選択に応じて瞬時に(または、ほぼ瞬時に)提示することができる。

30

【0033】

ユーザデバイスは、たとえば、テキストの意味解析、および/または画像内のテキストの配置を使用して、画像に描かれたテキストをブロックに分割し、各ブロックをユーザ選択可能なターゲットとして提示することができる。たとえば、ユーザデバイスは、画像内のテキストのレイアウトおよび意味解析に基づいて、メニューの画像内のテキストを、メニューのセクションごとのブロック(たとえば、前菜のブロック、主菜のブロックなど)を含む複数のブロックに分割することができる。ユーザがブロックのうちの1つを選択すると、ユーザデバイスは選択されたブロックをより高いズームレベルにおいて提示し、ブロック内のテキストの一部をユーザ選択可能なターゲットとして提示することができる。たとえば、ユーザデバイスは、フレーズ、メニュー項目、またはテキストの他の関連部分を識別し、テキストの関連部分のうち少なくともいくつかを、より高いズームレベルにおいてユーザ選択可能なターゲットとして提示することができる。ズームレベルと、様々なズームレベルに基づいてユーザ選択可能なターゲットを変更することによって、ユーザは

40

50

、追加コンテンツについて閲覧したいテキストをより容易に選択することができる。

【0034】

各ズームレベルにおいて、ズームレベルにおいてユーザデバイスによって提示されるテキストに基づいて、ユーザデバイスにおいてコンテンツをプリキャッシュすることができる。ユーザがズームレベルを調整すると、更新されたズームレベルにおいてユーザデバイスによって提示されるテキストに基づいて、追加および/または異なる(代替)コンテンツを、(たとえば、ユーザがテキストを選択する前に)ユーザデバイスにおいてプリキャッシュすることができる。このようにして、ユーザが選択したテキストに関連するコンテンツを、ユーザが選択を行った後に瞬時に、または、少なくともネットワークを介して要求を送信するよりもより迅速に提供することができる。

10

【0035】

図1は、ユーザデバイス110が画像に描かれたテキストに関連するコンテンツを提示する例示的な環境100を示す図である。ユーザデバイス110は、スマートフォン、タブレットコンピュータ、ウェアラブルデバイス、あるいは他のタイプのモバイルまたは非モバイル電子デバイスであり得る。ユーザデバイス110は、画像113および/またはビデオをキャプチャするカメラ112を含む。ユーザデバイス110はまた、画像113および/またはビデオ、ならびに画像113に関連するコンテンツを提示するディスプレイ120を含む。以下により詳細に説明するように、ディスプレイ120は、ユーザがディスプレイ120に提示される画像113に描かれたテキストと対話することを可能にするタッチスクリーンなどの入力デバイスを含むことができる。

20

【0036】

ユーザデバイス110は、ユーザがカメラ112を使用して画像をキャプチャすること、またはビューファインダを使用してカメラのレンズの視野内のオブジェクトを閲覧することを可能にする画像アプリケーション111を含む。画像アプリケーション111はまた、ユーザが、画像に描かれたテキストに関連するコンテンツを閲覧および/または聞くことも可能にする。本明細書の目的上、画像は、カメラ112によってキャプチャされた画像、およびカメラのレンズの視野を示すビューファインダ内の可視コンテンツを指す。また、画像アプリケーション111により、ユーザが、たとえばユーザデバイス110のローカルメモリ124などのユーザデバイスに記憶された画像、または異なる場所に記憶された画像に描かれたテキストに関する追加コンテンツを閲覧することが可能になる。たとえば、ユーザは、ユーザデバイス110(または、ネットワーク化された場所)に記憶された画像を選択し、画像および/または画像に描かれたテキストに関するコンテンツを閲覧するために、画像アプリケーション111を使用することができる。ローカルメモリ124は、ユーザデバイス110の内部メモリ(たとえば、ユーザデバイス110のフラッシュメモリ)またはユーザデバイス110に接続された外部メモリ(たとえば、ユーザデバイス110のユニバーサルシリアルバス(USB)ポートに接続されたメモリ)を含むことができる。

30

【0037】

画像アプリケーション111は、画像に描かれたテキストを処理するテキストプロセッサ114を含む。テキストプロセッサ114は、たとえば、光学文字認識(OCR)を使用して、画像に描かれたテキストを認識することができる。いくつかの実装形態では、テキストプロセッサ114は、たとえば、ネットワーク化された場所にあるフルサービスOCRアプリケーションと比較して、画像においてテキストが識別される速度を上げるために、ユーザデバイス110上で実行される軽量OCRエンジンを含む。軽量OCRエンジンは少量のメモリ(たとえば、メモリのしきい値量よりも少ない)を使用するか、および/またはユーザデバイス110の最小の処理能力を使用して画像内のテキストを識別するOCRエンジンである。軽量OCRエンジンは、ネットワーク化された場所のOCRエンジンに画像を送信する場合と比較して、ユーザデバイス110の追加の処理能力を使用し得るが、軽量OCRエンジンにより、画像がネットワークを介して送信され、OCRエンジンによって処理され、テキストがユーザデバイス110に返送される場合よりも迅速にテキスト認識を実行することが可能になる。これにより、テキストを認識し、テキストに関連するコンテンツを提示する際の待ち時間が短縮される。

40

50

軽量OCRエンジンは、画像内のテキストを識別するための機能を含むソフトウェア開発キット(SDK)として実装することができる。

【0038】

テキストプロセッサ114はまた、関連するテキストのブロック、フレーズ、または他の部分を識別するために、画像に描かれたテキストを処理することができる。いくつかの実装形態では、テキストプロセッサ114は、画像内のテキストの配置に基づいて関連するテキストを識別することができる。たとえば、テキストプロセッサ114は、関連するテキストとして、一緒にグループ化され、ならびに/あるいは同じフォントタイプ、サイズ、および/または色などの同じ(または、類似の)視覚特性を有するテキストを識別することができる。テキストのグループは、グループ内の他のテキストのしきい値距離内にあるテキスト、および/または画像内の可視エッジまたは境界内にあるテキストを含むことができる。たとえば、テキストのグループは、画像に描かれた可視ブロック(たとえば、正方形、長方形、または可視境界を有する他の適切な形状)内の、または画像に描かれた分割線の片側にあるテキストを含み得る。

10

【0039】

いくつかの実装形態では、テキストプロセッサ114は、関連するテキストを識別するために、画像に描かれたテキストの意味解析を実行する。たとえば、テキストプロセッサ114は、関連するテキストとして、画像に描かれたテキスト内のフレーズ、句、文、または段落を識別するために、意味解析を使用することができる。テキストプロセッサ114はまた、関連するフレーズ、句、文、段落、または個々の単語のブロックを識別するために、意味解析および/または画像内のテキストの配置を使用することができる。たとえば、画像内の認識されたすべての単語(または、その大部分)を分析することによって、テキストプロセッサ114は、以下で説明するように画像のカテゴリを決定し、複数の意味を持つ場合がある単語の意味を明確にするために画像のカテゴリを使用することができる。別の例では、テキストプロセッサ114は、特定の見出しを含む画像の特定の部分に単語が含まれることを決定し、単語の意味を決定するために見出しを使用するために、テキストの配置を使用することができる。たとえば、「頬」という単語がメニューの画像にあり、見出し「BBQ主菜」の下にある場合、テキストプロセッサ114は、「頬」という単語が身体部分ではなく食品を指していると決定することができる。

20

【0040】

場合によっては、テキストのブロックは複数の行のテキストを含む。たとえば、メニューの前菜のテキストのブロックは、前菜ごとにテキストの1つまたは複数の行を含むことができる。他の場合、テキストのブロックは、単一のフレーズまたは単一の単語を含み得る。たとえば、テキストのブロックは、ボートの横に描かれたテキスト、または公園において看板に刻まれたテキストを含み得る。

30

【0041】

テキストプロセッサ114は、処理された画像において識別されたテキストと、画像において識別されたテキストの各ブロック(または、他の部分)を指定する画像データ115を出力することができる。テキストのブロックごとに、画像データ115は、ブロックに含まれる実際のテキスト(たとえば、実際の単語およびフレーズ)を指定することができる。画像データ115はまた、画像内のテキストの各部分の位置を指定することができる。たとえば、画像データ115は、テキストのブロックについて、テキストのブロックを囲む画像内の座標、またはテキストの長方形ブロックの各コーナーの座標を含むことができる。以下に説明するように、画像アプリケーション111のコンテンツ取得部116は、画像において識別されたテキストに関連するコンテンツを取得するために、たとえば、ローカルメモリ124におけるコンテンツをプリキャッシュするために、画像データ115を使用することができる。

40

【0042】

画像アプリケーション111のユーザインターフェースコントローラ118は、画像113を受信し、ディスプレイ120上に画像113を提示するユーザインターフェース122を生成するこ

50

とができる。いくつかの実装形態では、画像アプリケーション111は、カメラ112によってキャプチャされた画像113を、たとえば画像がテキストプロセッサ114によって処理される前に、ユーザデバイス110のディスプレイ120によって表示するためにユーザインターフェースコントローラ118に直接提供する。このようにして、画像113は、画像がキャプチャされた直後に表示される。

#### 【0043】

ユーザインターフェースコントローラ118はまた、テキストプロセッサ114によって出力された画像113の画像データ115を受信することができる。ユーザインターフェースコントローラ118は、テキストプロセッサ118によって識別されたテキストのブロックのユーザ選択可能なターゲットを生成するために、画像データ115(たとえば、テキストのブロックと画像113内のそれらの位置を示すデータ)を使用することができる。ユーザ選択可能なターゲットにより、ユーザは、ブロック内のテキストに関連するコンテンツを取得するためにテキストのブロックを選択することと、ブロック内のテキストを選択するためにブロックにズームインすることと、および/または別のアクションを開始することとが可能になる。ユーザ選択可能な各ターゲットは、ディスプレイの選択可能な部分(たとえば、選択可能なタッチスクリーン領域)と、ブロックが選択可能であることを示す視覚インジケータを含むことができる。たとえば、テキストのブロックのユーザ選択可能なターゲットは、視覚インジケータとして、選択可能なアイコン、ブロックの周囲の強調表示されたボックス(または、他の形状)、ブロック内のテキストの異なる背景色、ブロック内のテキストの異なるフォントの色、強調表示または太字のテキスト、あるいはブロックが選択可能であることをユーザに示す他の適切な視覚インジケータを含むことができる。ユーザは、ターゲットと対話することによって、たとえば、ターゲットをクリックすること、ターゲット上にポインタを置くこと、ターゲットの場所でタッチスクリーンにタッチすること、または他の適切なユーザ対話を使用することによって、ユーザ選択可能なターゲットを選択することができる。

10

20

#### 【0044】

ユーザインターフェースコントローラ118は、ユーザインターフェース122においてユーザ選択可能なターゲットを画像113に提示することができる。ユーザインターフェースコントローラ118はまた、ブロックに含まれるテキストの部分に対してユーザ選択可能なターゲットを生成することができる。たとえば、ユーザインターフェース122は、テキストのブロックについて、ユーザが、ブロック内のテキストをより大きくするためにブロックにズームインすることを可能にするユーザ選択可能なターゲットを含むことができる。ユーザインターフェース122はまた、ブロック内のテキストの1つまたは複数の部分ごとに、ユーザがテキストの部分に関連するコンテンツを要求するためにテキストの部分を選択することを可能にするユーザ選択可能なターゲットを含むことができる。上述のように、画像がキャプチャされた直後に画像113が表示されるいくつかの実装形態では、画像がテキストプロセッサによって処理されると、ユーザインターフェースコントローラ118は、最初に表示された画像をユーザ選択可能なターゲットで補完し得る。

30

#### 【0045】

いくつかの実装形態では、ユーザインターフェースコントローラ118は、テキストブロックに2つのユーザ選択可能なターゲットを含むことができる。たとえば、第1のユーザ選択可能なターゲットは、ユーザがテキストのブロックにズームインすることを可能にし、第2のユーザ選択可能なターゲットは、ユーザがブロック内のテキストに関連するコンテンツを要求することを可能にする。別の例では、ユーザがテキストを含まないブロック内の場所、たとえば、テキスト間のスペースを選択することに応答して、ユーザインターフェースコントローラ118はテキストのブロックにズームインすることができる。ユーザがブロック内のテキストを選択すると、ユーザインターフェースコントローラ118は、ユーザ選択のテキスト、またはブロック内のテキストに関連するコンテンツの要求を開始することができる。

40

#### 【0046】

50

ユーザインターフェースコントローラ118は、テキストブロックのユーザ選択に応じて、テキストのブロックにズームインするか、コンテンツを提示するかを決定することができる。いくつかの実装形態では、ユーザインターフェースコントローラ118は、選択されたテキストブロック内のテキストに基づいて、たとえば、選択されたテキストブロック内のテキストの意味コンテンツに基づいて、コンテンツにズームインするか、または提示するかを決定することができる。たとえば、テキストブロックが1つの特定のエンティティまたは特定のタイプのコンテンツに関連するテキストを含む場合、ユーザインターフェースコントローラ118は、テキストブロックにズームインするのではなく、エンティティに関連するコンテンツを提示することを決定することができる。別の例では、テキストブロックが特定の料理に関連するテキストを含む場合、ユーザインターフェースコントローラ118は、テキストブロックのユーザ選択に応じて特定の料理に関連するコンテンツを提示することを決定することができる。コンテンツは、たとえば、アニメーション化された吹出しで料理のレビューを含むテキストブロックの横にあるキャラクターアバターであってよい。テキストブロックが複数のエンティティまたは複数の種類のコンテンツに関連するテキストを含む場合、ユーザインターフェースコントローラ118は、テキストブロックのユーザ選択に応じてテキストブロックにズームインすることを決定することができる。

10

## 【0047】

ユーザインターフェースコントローラ118は、たとえば、特定のズームレベルに達するまで、ユーザがテキストの一部に対してユーザ選択可能なターゲットを選択するたびにズームレベルを調整することができる。ユーザがテキストのブロックを選択した場合、ユーザインターフェースコントローラ118は、ブロック内のテキストがディスプレイ120上でより大きくなるようにテキストのブロックにズームインすることができる。たとえば、ユーザインターフェースコントローラ118は、ユーザ選択のブロックが少なくとも1つの次元(たとえば、水平および/または垂直)においてディスプレイ120のビューポートを満たすように、ディスプレイ120上の画像にズームインすることができる。

20

## 【0048】

ユーザインターフェースコントローラ118はまた、ユーザ選択のブロック内のテキストの少なくとも一部をユーザ選択可能なターゲットとして提示することができる。このようにして、ユーザは、単語またはフレーズに関連するコンテンツを要求するために、テキストの一部、たとえば単語またはフレーズを選択することができる。いくつかの実装形態では、ユーザが複数の単語を含むフレーズまたはテキストの他の部分を選択した場合、ユーザインターフェースコントローラ118は、テキストの部分にさらにズームインし、選択された部分のサブ部分のユーザ選択可能なターゲットを提示し得る。たとえば、ユーザがフレーズを選択した場合、ユーザインターフェースコントローラ118は、フレーズが少なくとも1つの次元(たとえば、水平および/または垂直)でディスプレイのビューポートを満たすようにズームインすることができる。ユーザインターフェースコントローラ118はまた、テキストのサブ部分をユーザ選択可能なターゲットとして提示することができる。たとえば、ユーザがメニューの画像において「赤ワイン」というフレーズを選択した場合、ユーザインターフェースコントローラ118はそのフレーズにズームインし、「赤」という単語を1つのユーザ選択可能なターゲットとして提示し、「ワイン」という単語を第2のユーザ選択可能なターゲットとして提示することができる。

30

40

## 【0049】

ユーザインターフェースコントローラ118は、テキストの部分のコンテンツに基づいて、ユーザ選択可能なターゲットとして提示するために、ユーザ選択のブロック内のテキストの部分を選択することができる。いくつかの実装形態では、ユーザインターフェースコントローラ118は、テキスト部分に利用可能なコンテンツに基づいてテキストを選択することができる。たとえば、ユーザインターフェースコントローラ118は、特定のタイプのコンテンツが利用可能な、知られているエンティティ(たとえば、エンティティに関する情報の要約を含む知識パネルが利用可能なエンティティ)のリストにアクセスすることができる。テキストの一部が、名前、または知られているエンティティのリスト内のエンテ

50

ィティへの他の参照を含む場合、ユーザインターフェースコントローラ118は、テキストの一部をユーザ選択可能なターゲットとして提示できるため、ユーザは、エンティティに関連する追加コンテンツを要求するために、テキストの一部を選択することができる。

【0050】

いくつかの実装形態では、ユーザインターフェースコントローラ118は、ユーザ選択のブロック内のテキストの部分をランク付けし、テキストの上位ランク部分をユーザ選択可能なターゲットとして提示し、テキストの下位ランク部分をユーザ選択可能なターゲットとして提示しない。たとえば、ユーザインターフェースコントローラ118は、最高ランクを有するテキストの部分の特定の数または割合、またはしきい値スコアを満たす(たとえば、超える)ランクスコアを有する部分を選択することができる。ユーザインターフェースコントローラ118は、テキストの部分ごとに利用可能なコンテンツの量、テキストの各部分と画像のカテゴリとの間の関連性、テキストの部分によって参照されるエンティティの人気度(たとえば、エンティティごとに受信したクエリの数に基づいて)、および/または他の適切な基準に基づいて、テキストの部分をランク付け(および/またはランクスコアを決定)することができる。

10

【0051】

ユーザインターフェースコントローラ118はまた、ユーザ対話に応じてテキストからズームアウトすることができる。たとえば、ユーザインターフェースコントローラ118は、タッチスクリーンにおけるユーザピンチまたはズームアウトのためのアイコンの選択に応じて、より多くの画像を提示するためにズームアウトすることができる。

20

【0052】

たとえば、テキストの一部に対してユーザ選択可能なターゲットを選択することによってユーザがコンテンツの要求を開始すると、UIコントローラ118は要求データ123をコンテンツ取得部116に送信することができる。要求データ119は、選択されたテキストを指定することができる。たとえば、ユーザがユーザインターフェース122において提示される画像に描かれたユーザ選択可能なフレーズを選択した場合、ユーザインターフェースコントローラ118は、フレーズ内の各単語を指定する要求データ119を生成することができる。別の例では、ユーザがテキストのブロックを選択した場合、ユーザインターフェースコントローラ118は、テキストのブロック内の各単語を指定する要求データ119を生成することができる。

30

【0053】

いくつかの実装形態では、要求データ119はコンテキストデータを含む。たとえば、要求データ119は、選択されていない他のテキストを含むことができる。特定の例では、要求データ119は、画像において識別される各単語(または、単語のうちのいくつか)、または現在のズームレベルに基づいてユーザインターフェース122において可視である各単語(または、単語のうちのいくつか)を含むことができる。以下で説明するように、画像のカテゴリを識別するためにコンテキストデータを使用することができ、次いで、コンテキストデータは、選択した単語を明確にするために、またユーザ選択に応じて提示するコンテンツを選択するために使用することができる。たとえば、「肩」などの選択された単語が食品に対応することを決定し、適切なコンテンツを選択するために、メニューの画像内の追加の単語を使用することができる。

40

【0054】

コンテンツ取得部116は、ユーザインターフェースコントローラ118から要求データ119を受信し、要求データ119に基づいて、検索システム150(または、ローカルメモリ124)からコンテンツを取得することができる。いくつかの実装形態では、コンテンツ取得部118は、要求データ119に基づいて1つまたは複数の検索クエリを指定するクエリデータを生成し、たとえば、ローカルエリアネットワーク(LAN)、ワイドエリアネットワーク(WAN)、モバイルネットワーク、インターネット、またはそれらの組合せなどのデータ通信ネットワーク140を介して、クエリデータ130を検索システム150に送信する。1つまたは複数のクエリは、ユーザによって選択されたテキストを含むことができる。クエリデータ130はまた

50

、コンテキストデータを含むことができる。

【0055】

検索システム150は、検索エンジン152およびコンテキスト分類器154を含む。検索エンジン150は、クエリデータ130に含まれる1つまたは複数の検索クエリに含まれる単語に基づいてコンテンツを識別するために、インデックス156を使用することができる。インデックス156は、クエリに回答して提供することができる、またはクエリに回答して提供された検索結果において参照することができる、電子リソース(たとえば、ウェブページ、画像、ビデオ、および/または他のコンテンツ)を指定するデータを含むことができる。

【0056】

コンテキスト分類器154は、画像または画像の一部に描かれたテキストに基づいて画像(または、ユーザインターフェース122において可視である画像の一部)を分類するために、1つまたは複数の機械学習モデルを使用することができる。機械学習モデルは、ラベル付きトレーニングデータを使用してトレーニングすることができる。ラベル付きトレーニングデータは、トレーニング画像のセットと、画像ごとに、画像に描かれている単語を指定するデータと、画像のカテゴリを指定するラベルとを含むことができる。例示的なカテゴリは、レストランのメニュー、標識(たとえば、道路標識または店舗の看板)、雑誌、映画のポスター、音楽のポスターなどを含むことができる。図1は、コンテキスト分類器154を検索システム150の一部として示しているが、いくつかの実装形態では、コンテキスト分類器154をユーザデバイス110上に実装することができる。

【0057】

クエリデータが受信されると、コンテキスト分類器154は、画像を分類するために機械学習モデルがトレーニングされたあらかじめ定義されたカテゴリのセットの1つに画像を分類しようと試みるために、クエリデータ130のコンテキストデータと機械学習モデルを使用することができる。いくつかの実装形態では、検索エンジン152は、分類、および/または画像が正常に分類されたかどうかに基づいて、検索クエリを調整し、かつ/またはリソースのランキングを調整する。

【0058】

コンテキスト分類器154が画像をカテゴリのうちの1つに分類する場合、検索エンジン152(または、コンテキスト分類器154がユーザデバイス110に実装されている場合はコンテンツ取得部116)は、画像が分類されたカテゴリに基づいて1つまたは複数の用語を含めるために、検索クエリを調整することができる。たとえば、画像が映画のポスターとして分類される場合、検索エンジン152は、「映画」という単語を検索クエリに追加することができる。特定の例では、ユーザ選択のテキストは映画のタイトル「エイリアン」であり得る。コンテキスト分類器154は、画像を映画のポスターとして分類するために、映画のポスターに描かれた追加の単語、たとえば、俳優の名前または映画の説明を使用することができる。次いで、検索エンジン152は、「エイリアン映画」というクエリを形成するために、「映画」という単語を「エイリアン」というクエリに追加することができる。このようにして、検索結果は一般的にエイリアンに関連する他のコンテンツよりも、「エイリアン」という映画に関連する可能性が高くなる。

【0059】

いくつかの実装形態では、検索エンジン152は、画像の分類に基づいて検索結果のランキングを調整する。たとえば、検索エンジン152は、検索クエリを修正するのではなく、画像のカテゴリに関連するリソースの検索結果を高める(たとえば、検索結果のランクを上げる)、および/または画像のカテゴリに関連しない検索結果のランキングを下げるることができる。いくつかの実装形態では、検索エンジン152は、画像のカテゴリに基づいてリソースのタイプのランキングを調整することができる。たとえば、画像のカテゴリが音楽のポスターである場合、検索エンジン152はオーディオおよびビデオリソースを高め、ウェブページリソースのランキングを下げるることができる。

【0060】

コンテキスト分類器154が画像をカテゴリのうちの1つに分類しない場合、たとえば、コ

10

20

30

40

50

ンテキストデータがカテゴリのうちの1つに関連していないため、検索エンジン152は、ユーザ選択のテキストを使用して一般的な検索を実行することができる。これは、コンテキスト分類器154によって受信されたコンテキストデータに基づいてカテゴリを決定できない画像の代替オプションであり得る。たとえば、店舗の看板の画像は、画像を分類するために使用することができる追加のコンテキストテキストがない単一の単語のみを含む場合がある。

#### 【0061】

検索エンジン152は、検索クエリと、検索クエリに応答して識別されたリソースのランキングに基づいて、コンテンツ132をコンテキスト取得部116に提供することができる。たとえば、検索エンジン152は、コンテンツ132に、特定の数の最高ランクのリソースを提供することができる。別の例では、検索エンジン152は、コンテンツ132に、最高ランクのリソースを参照する、たとえば、それにリンクする、検索結果を提供することができる。たとえば、以下で説明するように、コンテンツを画像で提示することができ、または、ユーザ選択のテキストに基づいて選択した検索結果を含む検索結果ページが提示され得る。コンテンツ取得部116は、コンテンツ132を受信し、コンテンツ132をユーザインターフェースコントローラ118に提供することができる。次に、ユーザインターフェースコントローラ118は、ユーザインターフェース122におけるコンテンツ132に含まれるリソースの少なくとも一部を、検索結果ページ上の画像または検索結果とともに提示することができる。たとえば、ユーザインターフェースコントローラ118は、コンテンツとして、ウェブページリソースから抽出されたテキストおよび/または画像、ウェブサイトによってホストされるビデオへのリンク、ならびに/あるいは、ユーザによって選択された場合、ユーザインターフェースにおいてビデオの再生を開始するビデオのサムネイル画像を提示することができる。

#### 【0062】

コンテンツ取得部116はまた、たとえば、画像に描かれたテキストの選択を受信する前に、コンテンツをローカルメモリにプリキャッシュすることもできる。いくつかの実装形態では、コンテンツ取得部116は、テキストプロセッサ114によって出力された画像データ115に基づいてクエリデータ130を検索システム150に送信する。クエリデータ130は、テキストプロセッサ114によって画像内で識別されたすべてのテキストを含むことができる。検索エンジン152は、コンテキスト分類器154によって決定された画像のテキストおよび/またはカテゴリに基づいてコンテンツを識別することができる。コンテンツ取得部116は、検索エンジン152から識別されたコンテンツを受信し、ローカルメモリ124にコンテンツを記憶することができる。このようにして、ユーザが画像全体に関連するコンテンツを要求した場合、コンテンツ取得部116は、ユーザ要求後にネットワーク140を介して追加のクエリデータを検索システム150に送信するのではなく、ローカルメモリからコンテンツを取得することができる。これにより、ユーザにコンテンツを提示する際の待ち時間を短縮し、選択したテキストに関連する追加コンテンツをユーザが瞬時に、またはほぼ瞬時に(たとえば、1秒などのしきい値未満の時間で)閲覧することを可能にする。

#### 【0063】

いくつかの実装形態では、コンテンツ取得部116は、プリキャッシュのために、ユーザインターフェース122のズームレベルに基づいて異なるテキストベースの粒度のクエリを生成する。テキストベースの粒度は、各クエリを生成するために使用されるテキストの量を定義する。たとえば、あるテキストベースの粒度では、クエリを生成するために個々の単語が使用され得る。より高いテキストベースの粒度では、クエリを生成するために、個々のフレーズまたは他の複数の単語のセットをしきい値の単語数まで使用することができる。

#### 【0064】

ズームレベルにより、ユーザインターフェースにおいて画像全体が閲覧可能である場合、コンテンツ取得部は、ユーザ選択可能なターゲットが提示されるテキストのブロックごとに1つまたは複数のクエリを生成することができる。たとえば、テキストプロセッサ114

10

20

30

40

50

は、たとえば、ブロックのうちの1つのユーザ選択を受信する前に、画像内のテキストの多数の別個のブロックを識別し、ブロックごとに別個のコンテンツを要求することができる。特定の例では、メニューの画像は、前菜用のテキストの第1のブロック、主菜用のテキストの第2のブロック、および飲み物用のテキストの第3のブロックを含み得る。この例では、クエリデータ130は、第1のブロックにおける前菜のリスト、第2のブロックにおける主菜のリスト、および第3のブロックにおける飲み物のリストを含むことができる。検索エンジン152は、ブロックごとにコンテンツ(たとえば、電子リソース)を個別に識別し、コンテンツをコンテンツ取得部116に提供することができる。コンテンツ取得部116は、コンテンツをローカルメモリ124に記憶することができる。このようにして、ユーザがブロックのうちの1つを選択すると、コンテンツ取得部116は、ユーザ選択後にネットワーク140を介して追加のクエリデータを検索システム150に送信することなしに、ローカルメモリ124からユーザ選択のブロックのコンテンツを取得することができる。

10

#### 【0065】

コンテンツ取得部116は、ユーザインターフェース122のズームレベルを変更するユーザインターフェース122とのユーザ対話に応答して、ローカルメモリ124に記憶されたコンテンツを更新することができる。たとえば、ユーザがブロックのうちの1つを選択すると、ユーザインターフェースコントローラ118は選択されたブロックにズームインし、ユーザ選択のブロック内のテキストの少なくとも一部をユーザ選択可能なターゲットとして提示することができる。コンテンツ取得部116はまた、更新されたズームレベルに基づいて、異なるテキストベースの粒度でクエリを生成することができる。たとえば、コンテンツ取得部は、選択されたブロックにおいて識別されたフレーズごとに、または、ユーザインターフェースコントローラ118がユーザ選択可能なターゲットとして提示する選択されたブロック内のフレーズごとにクエリを生成することができる。

20

#### 【0066】

前のメニュー例を続けると、ユーザが前菜のブロックを選択した場合、コンテンツ取得部116は、前菜を表すフレーズごとに1つまたは複数のクエリを生成することができる。コンテンツ取得部116は、たとえば、ユーザが前菜のうちの1つを選択する前、または前菜がユーザ選択可能なターゲットとしてユーザインターフェース122において提示される前に、クエリを指定するクエリデータ130を検索エンジン152に送信することができる。検索エンジン152は、前菜ごとのコンテンツを識別し、識別したコンテンツをコンテンツ取得部116に提供することができる。コンテンツ取得部116は、コンテンツをローカルメモリ124に記憶することができる。このようにして、ユーザが前菜のうちの1つを選択した場合、コンテンツ取得部116は、ユーザ選択後にネットワーク140を介して追加のクエリデータを検索システム150に送信することなしに、ローカルメモリ124からユーザ選択の前菜のコンテンツを取得することができる。

30

#### 【0067】

いくつかの実装形態では、ユーザインターフェースコントローラ118は、テキストのコンテンツに基づいて、ユーザ選択のテキストに関連するコンテンツを要求する以外のアクションを実行することができる。たとえば、画像が電話番号を示している場合、ユーザインターフェースコントローラ118は、スマートフォンによる電話番号表示のユーザ選択に応じて、スマートフォン上で通話を開始することができる。別の例では、画像が住所を示す場合、ユーザインターフェースコントローラ118は、地図アプリケーションを開き、住所のユーザ選択に応じて、地図アプリケーションが住所によって参照される場所の地図を表示することを要求することができる。別の例では、画像がウェブリソースの宛先アドレス、たとえばユニバーサルリソースロケータ(URL)を示す場合、ユーザインターフェースコントローラ118は、アプリケーション(たとえば、ウェブブラウザ)を開き、アプリケーションに、ネットワークアドレスによって参照される電子リソースにナビゲートするよう要求することができる。別の例では、画像がウェブリソースの電子メールアドレスを示す場合、ユーザインターフェースコントローラ118は、電子メールアドレスのユーザ選択に応じて、たとえば、電子メールアプリケーションを開き、その電子メールアドレス宛ての

40

50

空の電子メールを表示することによって、その電子メールアドレスへの電子メールを開始することができる。

【0068】

いくつかの実装形態では、ユーザインターフェースコントローラ118は、ビューファインダ内に表示されるコンテンツを増強するユーザインターフェースを生成する。たとえば、ユーザは、カメラ112をメニューまたは他のオブジェクトに向けることができる。ユーザインターフェースコントローラ118は、ビューファインダ内に現在表示されているテキストブロックおよび/またはテキストの他の部分のユーザ選択可能なターゲットを生成し、ユーザ選択可能なターゲットにビューファインダ内のテキストを提示することができる。ユーザがカメラ112を移動させる場合、ユーザインターフェースコントローラ118は、ユーザ選択可能なターゲットが生成されたテキストとともに、ユーザ選択可能なターゲットをそれらの適切な位置に移動させることができる。ビューファインダ内において新しいテキストが可視になると、ユーザインターフェースコントローラ118は、ユーザ選択可能なターゲットを生成し、ユーザ選択可能なターゲットのユーザ選択に応じてコンテンツを提示するために、上述のようにテキストを処理することができる。

10

【0069】

図2は、ユーザ選択のテキストに関連する追加コンテンツを要求するために、ユーザが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205の例示的なスクリーンショットを示す図である。ユーザインターフェース205の第1のスクリーンショット210は、例示的なレストランからのメニューの画像211を提示するユーザインターフェース205を示している。たとえば、メニューの画像211をキャプチャするために、ユーザはスマートフォンのカメラを使用した可能性がある。例示的なメニューは、前菜のセクション、サンドイッチのセクション、主菜のセクション、およびデザートセクションを含む。

20

【0070】

ユーザインターフェース205の第2のスクリーンショット220は、メニューの異なるセクション用に作成されたユーザ選択可能なターゲット221~224を示している。たとえば、図1のユーザインターフェースコントローラ118などのユーザインターフェースコントローラは、メニューの各セクションが関連テキストのブロックであることを識別し、各セクションをユーザ選択可能なターゲットとして提示することができる。具体的には、ユーザ選択可能なターゲット221により、ユーザはメニューの前菜セクションを選択することができ、ユーザ選択可能なターゲット222により、ユーザはメニューの主菜セクションを選択することができ、ユーザ選択可能なターゲット223により、ユーザはメニューのサンドイッチセクションを選択することができ、ユーザ選択可能なターゲット224により、ユーザはメニューのデザートセクションを選択することができる。

30

【0071】

この例では、ユーザ選択可能な各ターゲット221~224は、そのセクションの周囲にメニューのセクションが選択可能であることをユーザに示す太い線を含む。いくつかの実装形態では、セクションを選択することができるという事実にユーザの注意を引くために、各セクションの周囲の線が特定の色であってもよく、特定の明るさを有していてもよい。他の例では、セクションが選択可能であることを示すために、各セクションのテキストの背景色が強調表示されても変更されてもよく、または、セクションが選択可能であることを示すためにテキストの色が変更されてもよい。

40

【0072】

例示的なユーザインターフェース205はまた、スクリーンショット220にアレルギーインジケータ227を含むように更新されている。いくつかの実装形態では、画像アプリケーション、たとえば、画像アプリケーション111は、ユーザに関連し、画像に描かれたテキストに関連する追加コンテンツを提示するためにユーザによって提供されたユーザデータにアクセスすることができる。ユーザデータの例は、ユーザがアレルギーを持っている食品のリストである。画像アプリケーションは、メニューの食品をアレルギーのリストと比較し、一致する場合は、メニュー内の項目の隣にアレルギーインジケータを提示することが

50

できる。この例では、ユーザはエビにアレルギーがあり、アレルギーインジケータ227はメニュー項目をユーザに示す。

【0073】

ユーザは、ユーザ選択可能なターゲット内でタッチすることによって(ユーザデバイスにタッチスクリーンがある場合)、ポインタおよびマウスを使用してユーザ選択可能なターゲットをクリックすることによって、または別の適切な方法で、ユーザ選択可能なターゲットを選択することができる。この例では、ユーザ選択可能なターゲット221を選択し、メニューの前菜セクションにズームインするために、ユーザはポインタ226を使用している。

【0074】

ユーザ選択可能なターゲット221の選択にตอบสนองして、例示的なスクリーンショット230に示されるように、前菜セクションのズームインされた表示を提示するためにユーザインターフェイス205が更新される。たとえば、メニューの画像211は、第1のズームレベル、たとえば、デフォルトズームレベルにおいて表示することができる。ユーザ選択可能なターゲットが選択されると、選択したセクション内のテキストをより高いズームレベルにおいて提示するために、ユーザインターフェイスのズームレベルを第1のズームレベルよりも大きい第2のズームレベルに更新することができる。これにより、選択したセクション内のテキストをユーザがより適切に閲覧し、ユーザがさらにズームインしたり、コンテンツを要求したりしたいテキストをより適切に選択することが可能になる。

【0075】

例示的なスクリーンショット230では、ユーザインターフェイス205はまた、各前菜をユーザ選択可能なターゲットとして示している。たとえば、前菜のチップスアンドサルサは、ユーザ選択可能なターゲット231として提示され、前菜のスタフドマッシュルームは、ユーザ選択可能なターゲット232として提示され、前菜のバッファローウィングは、ユーザ選択可能なターゲット233として提示され、前菜のモツァレラスティックは、ユーザ選択可能なターゲット234として提示される。上述のように、ユーザがテキストのブロックを選択すると、選択されたブロック内のテキストの少なくとも一部をユーザ選択可能なターゲットとして提示することができる。

【0076】

ユーザがユーザ選択可能なターゲットを選択すると、さらにズームインするか、ユーザ選択のターゲットのコンテンツに関連するコンテンツを提示するために、ユーザインターフェイス205を更新することができる。この例では、ユーザは、ポインタ236を使用してチップスアンドサルサのユーザ選択可能なターゲット231を選択し、それに応じて、例示的なスクリーンショット240に示されるように、チップスアンドサルサというフレーズにさらにズームインするために、ユーザインターフェイス205が更新される。

【0077】

また、ユーザインターフェイス205はまた、「チップスアンドサルサ」というフレーズの単語に対してユーザ選択可能なターゲットを提示するように更新されている。具体的には、ユーザインターフェイス205は、「チップス」という単語に対するユーザ選択可能なターゲット241と、「サルサ」という単語に対するユーザ選択可能なターゲット243とを含む。ユーザインターフェイス205は、「アンド」という単語、またはユーザによって選択される可能性が低い他の単語に対するユーザ選択可能なターゲットを含まなくてもよい。たとえば、上述のように、ユーザインターフェイスコントローラは、テキストのコンテンツに利用可能なコンテンツ(たとえば、テキストによって参照されるエンティティ)、または他の基準に基づいて、選択したブロックまたはテキストの他の部分のどの単語またはフレーズをユーザ選択可能なターゲットとして提示されるべきかを選択することができる。

【0078】

ユーザがユーザ選択可能なターゲット241を選択した場合、チップスの栄養情報、チップスの画像、チップスのレシピなどの、チップスに関連するコンテンツを提示するために、ユーザインターフェイス205を更新することができる。同様に、ユーザがユーザ選択可

10

20

30

40

50

能なターゲット243を選択した場合、サルサの栄養情報、サルサの画像、サルサのレシピなどの、サルサに関連するコンテンツを提示するために、ユーザインターフェース205を更新することができる。

【0079】

図3は、ユーザがテキストを選択し、選択されたテキストに関連するコンテンツを閲覧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305の例示的なスクリーンショット310および320を示す図である。スクリーンショット310は、図2のスクリーンショット230に類似している。スクリーンショット310において、ユーザインターフェース305は、メニューの画像の部分的なビュー(たとえば、ズームインしたビュー)を提示する。具体的には、ユーザインターフェース305は、メニューの前菜セクションを含む画像の部分を提示する。ユーザインターフェース305はまた、メニュー内の前菜項目のユーザ選択可能なターゲット311~314を提示する。この例では、ユーザは、ポインタ316を使用して、前菜「チップスアンドサルサ」のユーザ選択可能なターゲット311を選択している。

10

【0080】

ユーザ選択に応答して、スクリーンショット320に示されるように、メニューの画像上にチップスアンドサルサに関連するコンテンツを提示するように、ユーザインターフェース305が更新される。この例では、コンテンツは、チップスアンドサルサに関する情報と、チップスアンドサルサの画像323および324とを含むテキストブロック322を含む。テキストブロック322と画像323および324は、選択されたユーザ選択可能なターゲット311が可視のままであるように、スクリーンショット310において可視であったメニューの画像の部分の上に提示することができる。たとえば、テキストブロック322と画像323および324は、チップスアンドサルサのテキストが可視のままである間、他の前菜を描写する画像の部分の上に提示される。

20

【0081】

いくつかの実装形態では、選択されたテキストが可視のままであり、選択されたテキストに関連するコンテンツが選択されたテキストの下に提示されるように、画像の異なる部分が可視であるように、ユーザインターフェース305を更新することができる。いくつかの実装形態では、画像の可視部分は変更されず、コンテンツは、選択されたテキストを含まない画像の部分、たとえば選択されたテキストの上、または選択されたテキストの側面に提示される。

30

【0082】

図4は、ユーザがテキストを選択し、選択されたテキストに関連するコンテンツを閲覧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405の例示的なスクリーンショット410および420を示す図である。スクリーンショット410は、図2のスクリーンショット230に類似している。スクリーンショット410において、ユーザインターフェース405は、メニューの画像の部分的なビュー(たとえば、ズームインしたビュー)を提示する。具体的には、ユーザインターフェース405は、メニューの前菜セクションを含む画像の部分を提示する。ユーザインターフェース405はまた、メニュー内の前菜項目のユーザ選択可能なターゲット411~414を提示する。ユーザは、ポインタ416を使用して、前菜「チップスアンドサルサ」のユーザ選択可能なターゲット411を選択している。

40

【0083】

この例では、図3の例のように画像上にコンテンツを提示するのではなく、スクリーンショット420に示されるように検索結果ページ421を提示するためにユーザインターフェース405が更新されている。検索結果ページ421は、ユーザによって選択されたテキストに基づいて検索を開始するために使用することができる。たとえば、検索結果ページ421は、ユーザが「チップスアンドサルサ」というテキストに対してユーザ選択可能なターゲット411を選択したことに応答して、たとえば画像アプリケーションによって「チップスアンドサルサ」というテキストを自動的に入力された検索ボックス422を含む。次いで、ユーザは、「チップスアンドサルサの栄養情報」などの修正されたクエリを提出し、修正されたクエリの検索結果を検索結果ページ421上で閲覧するために、検索ボックス422内のクエ

50

りを修正することができる。

【0084】

検索結果ページ421はまた、「チップスアンドサルサ」というクエリに基づいて識別された検索結果423~425を含む。たとえば、画像アプリケーションはクエリを検索エンジンに提出し、ユーザ選択可能なターゲット411の選択の検出にตอบสนองして、検索エンジンから受信した検索結果を検索結果ページに提示することができる。ユーザがクエリを修正し、修正されたクエリを提出すると、修正されたクエリの新しい検索結果を、たとえば、検索結果423~425の代わりに検索結果ページに提示することができる。

【0085】

図5は、ユーザ選択のテキストのコンテンツに基づいてアクションを開始するための例示的なプロセス500のフローチャートである。プロセス500の動作は、たとえば、図1のユーザデバイス110などの1つまたは複数のデータ処理装置を含むユーザデバイスによって実装することができる。プロセス500はまた、データ処理装置を含むシステムによる命令の実行によりデータ処理装置にプロセス500の動作を実行させる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することができる。

10

【0086】

ユーザデバイスは、キャプチャされた画像からテキストを抽出する(502)。画像は、ユーザデバイスのカメラを使用してキャプチャされてもよく、ユーザデバイスに送信される別のデバイスのカメラによってキャプチャされてもよく、ストレージから取得されてもよい。ユーザデバイスは、OCR技法を使用して、キャプチャした画像からテキストを抽出

20

【0087】

ユーザデバイスは、抽出されたテキストを複数のブロックに分割する(504)。ユーザデバイスは、キャプチャされた画像内の抽出されたテキストの配置に基づいて、抽出されたテキストの1つまたは複数の意味解析に基づいて、および/あるいは関連テキストのグループを識別するための他の適切な技法を使用して、抽出されたテキストを関連テキストの複数のブロックに分割することができる。

【0088】

ユーザデバイスは、ユーザデバイスのユーザインターフェース上にそれぞれの第1のユーザ選択可能なターゲットとしてテキストの複数のブロックを提示する(506)。ユーザインターフェースは、複数のブロックをユーザ選択可能なターゲットとして、第1のズームレベルにおいて、たとえば、キャプチャされた画像全体を表示するデフォルトズームレベルにおいて提示することができる。

30

【0089】

ユーザデバイスは、複数のブロックのうちの第1のブロックのユーザ選択を検出する(508)。ユーザは、テキストを選択するために、ユーザ選択可能なターゲットと対話することができる。たとえば、ユーザはテキストの場所でユーザデバイスのディスプレイのタッチスクリーンにタッチするか、マウスを使用してテキストをクリックすることができる。

【0090】

テキストのユーザ選択にตอบสนองして、ユーザデバイスは、第1のブロック内の抽出されたテキストを、ユーザデバイスのユーザインターフェース上にそれぞれの第2のユーザ選択可能なターゲットとして提示する(510)。たとえば、ユーザデバイスは、テキストの第1のブロックにズームインするために、ユーザインターフェースを更新することができる。すなわち、ユーザデバイスは、第1のズームレベルより大きい第2のズームレベルにおいてテキストの第1のブロックを表示することができる。このようにして、第1のブロックのテキストはより大きくなり、ユーザが、そのテキストに関連するコンテンツを要求するために適切なテキストを選択することがより容易になる。

40

【0091】

ユーザデバイスは、テキストの第1のブロック内の抽出されたテキストの一部のユーザ選択を検出する(512)。それに応じて、ユーザデバイスは、選択したテキストのコンテン

50

ツに基づいてアクションを開始する(514)。アクションは、選択したテキストのコンテンツに応じて異なる場合がある。たとえば、選択したテキストのコンテンツが電話番号である場合、アクションは電話番号を使用して電話をかけることであってよい。選択したテキストのコンテンツが住所の場合、アクションはその住所への地図アプリケーションを開くことであってよい。

【0092】

選択したテキストのコンテンツがエンティティであるか、対応するアクションを有しない場合、アクションは選択したテキストのコンテンツに関連するコンテンツを要求することであってよい。たとえば、ユーザデバイスは、選択したテキストのコンテンツに基づいて1つまたは複数の検索クエリを生成し、選択したテキストのコンテンツを使用して検索を開始することができる。

10

【0093】

いくつかの実装形態では、アクションは、ユーザが好みの検索クエリを選択することを可能にするために、ユーザインターフェース上に複数の候補の検索クエリを表示することであってよい。候補の検索クエリは、選択したテキストのコンテンツに基づくことができる。たとえば、候補の検索クエリは、選択したテキストのコンテンツと、選択したテキストのコンテンツに関連する1つまたは複数の推奨される用語を含むことができる。ユーザが候補クエリのうちの1つを選択すると、ユーザデバイスはクエリを検索エンジンに送信し、選択したクエリに関連する検索結果またはコンテンツをユーザインターフェースに提示することができる。

20

【0094】

図6は、検索クエリに応答してコンテンツを提供するための例示的なプロセス600のフローチャートである。プロセス600の動作は、たとえば、図1の検索システム150などの1つまたは複数のデータ処理装置を含むシステムによって実装することができる。プロセス600はまた、データ処理装置を含むシステムによる命令の実行によりデータ処理装置にプロセス600の動作を実行させる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することができる。

【0095】

本システムはクエリデータを受信する(602)。クエリデータは、キャプチャされた画像のクエリとコンテキストデータを含むことができる。クエリは、画像に描かれ、ユーザによって選択されたテキストのコンテンツを含むことができる。コンテキストデータは、画像に描かれているがユーザによって選択されていない他のテキストを含むことができる。

30

【0096】

本システムは、コンテキストデータに基づいて画像の分類を試みる(604)。たとえば、本システムは、コンテキストデータに含まれる画像のテキストに基づいて画像の分類を試みるために、1つまたは複数の機械学習モデルを使用することができる。機械学習モデルの出力は、あらかじめ定義されたカテゴリのセット(たとえば、メニュー、映画、音楽、看板など)の1つであってよく、分類なしであってもよい。

【0097】

画像がカテゴリのうちの1つに属すると本システムが決定した場合、本システムは、ユーザ選択のテキストのコンテンツとカテゴリとの組合せに基づいて、電子リソース(たとえば、ウェブページ、画像、ビデオなど)を識別してランク付けする(608)。たとえば、本システムは、識別されたカテゴリに基づく1つまたは複数の用語を含めるために、選択されたテキストのコンテンツに基づくクエリを修正することができる。別の例では、本システムは、カテゴリに関連するリソースのランクを高める、および/またはカテゴリに関連しないリソースのランクを下げるることができる。

40

【0098】

画像がカテゴリのうちのいずれにも属しないとシステムが決定した場合、本システムは一般的な検索を使用してリソースを識別してランク付けする(610)。たとえば、本システムはクエリを修正したり、リソースのランキングを調整したりすることはできない。代わ

50

りに、本システムは選択されたテキストのコンテンツのみに基づいてリソースを識別することができる。

【0099】

本システムは、クエリデータが受信されたユーザデバイスでの提示のための検索結果コンテンツを提供する(612)。いくつかの実装形態では、検索結果コンテンツは、クエリに基づいて識別されたリソースと、任意で画像がカテゴリのうちの1つに属すると分類された場合の画像のカテゴリを含むことができる。たとえば、ユーザデバイスは、テキストが選択された画像を含むコンテンツ(たとえば、ウェブページのテキスト、画像、埋込みビデオ、またはビデオへのリンクなど)を表示することができる。

【0100】

いくつかの実装形態では、検索結果コンテンツは、クエリに基づいて識別されたリソースの検索結果と、任意で画像がカテゴリのうちの1つに属すると分類された場合の画像のカテゴリを含む。たとえば、ユーザデバイスは、検索結果を検索結果ページに、たとえば画像の代わりに、または画像の上に提示することができる。

【0101】

本明細書に記載される主題および動作の実施形態は、本明細書で開示される構造およびそれらの構造的同等物を含む、デジタル電子回路、あるいはコンピュータソフトウェア、ファームウェア、またはハードウェア、あるいはそれらの1つまたは複数の組合せにおいて実装することができる。本明細書に記載される主題の実施形態は、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にエンコードされた、1つまたは複数のコンピュータプログラム、すなわち、コンピュータプログラム命令の1つまたは複数のモジュールとして実装することができる。代替的または追加的に、プログラム命令は、データ処理装置による実行のために適切な受信機装置に送信するための情報をエンコードするために生成される、人工的に生成された伝播信号、たとえば、機械生成電気、光学、または電磁信号でエンコードすることができる。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、ランダムまたはシリアルアクセスメモリアレイまたはデバイス、あるいはそれらの1つまたは複数の組合せであってもよく、またはそれらに含まれてもよい。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、コンピュータストレージ媒体は、人工的に生成された伝播信号にエンコードされたコンピュータプログラム命令のソースまたは宛先であり得る。コンピュータストレージ媒体はまた、1つまたは複数の別個の物理的コンポーネントあるいは媒体(たとえば、複数のCD、ディスク、または他のストレージデバイス)であってもよく、それらに含まれてもよい。

【0102】

本明細書に記載される動作は、1つまた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データ、または他のソースから受信し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することができる。

【0103】

「データ処理装置」という用語は、例として、プログラマブルプロセッサ、コンピュータ、チップ上のシステム、あるいは前述のものの複数のもの、または組合せを含む、データを処理するためのあらゆる種類の装置、デバイス、および機械を包含する。本装置は、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)などの専用論理回路を含むことができる。本装置はまた、ハードウェアに加えて、問題のコンピュータプログラムの実行環境を作成するコード、たとえば、プロセッサファームウェア、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、あるいはそれらの1つまたは複数の組合せを構成するコードも含むことができる。本装置と実行環境は、ウェブサービス、分散コンピューティング、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。

10

20

30

40

50

## 【0104】

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、コンパイラ型言語またはインタープリタ型言語、宣言型言語または手続き型言語を含む、任意の形式のプログラミング言語で記述することができ、スタンドアロンプログラムとして、あるいはモジュール、コンポーネント、サブルーチン、オブジェクト、またはコンピューティング環境での使用に適した他のユニットとして、を含む任意の形式で展開することができる。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場合があるが、そうである必要はない。本プログラムは、他のプログラムまたはデータを保持するファイルの一部(たとえば、マークアップ言語ドキュメントに記憶された1つまたは複数のスクリプト)、問題のプログラム専用の単一のファイル、または複数の調整ファイル(たとえば、1つまたは複数のモジュール、サブプログラム、あるいはコードの一部を記憶するファイル)に記憶することができる。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つのサイトに配置されているか、複数のサイトにわたって分散され、通信ネットワークで相互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することができる。

10

## 【0105】

本明細書に記載されるプロセスと論理フローは、入力データを操作して出力を生成することによってアクションを実行するために、1つまたは複数のコンピュータプログラムを実行する1つまたは複数のプログラム可能なプロセッサによって実行することができる。プロセスおよび論理フローはまた、特殊用途の論理回路、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)によって実行することもでき、装置はそれらとして実装することもできる。

20

## 【0106】

コンピュータプログラムの実行に適したプロセッサは、例として、汎用および専用の両方のマイクロプロセッサ、およびあらゆる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。一般的に、プロセッサは、読取り専用メモリまたはランダムアクセスメモリ、あるいはその両方から命令およびデータを受信する。コンピュータの重要な要素は、命令に従ってアクションを実行するためのプロセッサと、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。一般に、コンピュータはまた、データを格納するための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、または光ディスクを含むか、それらからデータを受信する、またはそれらにデータを転送するように動作可能に結合される。しかしながら、コンピュータにそのようなデバイスは必要ない。さらに、コンピュータは別のデバイス、たとえば、ほんの数例を挙げると、モバイル電話、携帯情報端末(PDA)、モバイルオーディオまたはビデオプレーヤ、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)に組み込むことができる。コンピュータプログラムの命令およびデータを記憶するために適したデバイスは、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、磁気ディスク、ならびにCD-ROM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、メディア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、専用論理回路によって補完されてもよく、専用論理回路に組み込まれてもよい。

30

40

## 【0107】

ユーザとの対話を提供するために、本明細書に記載される主題の実施形態は、ユーザに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(陰極線管)またはLCD(液晶ディスプレイ)モニタ、ユーザがコンピュータに入力を提供す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で実装することができる。ユーザとの対話を提供するために他の種類のデバイスを使用することもでき、たとえば、ユーザに提供されるフィードバックは、たとえば、視覚フィードバック、聴覚フィードバック、または触覚フィードバックなどの任意の形式の感

50

覚フィードバックであってよく、ユーザからの入力、音響、音声、または触覚の入力を含む任意の形式で受信することができる。さらに、コンピュータは、ユーザによって使用されるデバイスとの間で文書を送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した要求に応じて、ユーザのクライアントデバイス上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、ユーザと対話することができる。

#### 【0108】

本明細書に記載される主題の実施形態は、たとえばデータサーバとしてのバックエンドコンポーネントを含む、またはアプリケーションサーバなどのミドルウェアコンポーネントを含む、またはユーザが本明細書に記載される主題の実装形態と対話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またはウェブブラウザを有するクライアントコンピュータなどのフロントエンドコンポーネントを含むコンピューティングシステムにおいて、あるいは1つまた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、またはフロントエンドコンポーネントの任意の組合せにおいて実装することができる。本システムのコンポーネントは、任意の形式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接続することができる。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」)およびワイドエリアネットワーク(「WAN」)、インターネットワーク(たとえば、インターネット)、および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

#### 【0109】

コンピューティングシステムは、クライアントとサーバを含むことができる。通常、クライアントとサーバは互いにリモートであり、通常は通信ネットワークを通じて対話する。クライアントとサーバの関係は、それぞれのコンピュータで実行され、相互にクライアントとサーバの関係を有するコンピュータプログラムによって発生する。いくつかの実施形態では、サーバは、データ(たとえば、HTMLページ)をクライアントデバイスに送信する(たとえば、クライアントデバイスと対話するユーザにデータを表示し、クライアントデバイスと対話する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おいて生成されたデータ(たとえば、ユーザ対話の結果)は、サーバにおいてクライアントデバイスから受信することができる。

20

#### 【0110】

本明細書は多くの特定の实装形態の詳細を含むが、これらは、発明の範囲または請求されるものの範囲を制限するものではなく、特定の発明の特定の实装形態に特有の特徴の説明として解釈されるべきである。別個の実施形態の文脈において本明細書に記載されている特定の特征是、単一の実施形態において組み合わせて実装することもできる。逆に、単一の実施形態の文脈において説明される様々な特征是、複数の実施形態において別々に、または任意の適切なサブコンビネーションにおいて実装することもできる。さらに、特征是特定の組合せにおいて作用するものとして上記で説明され、最初にそのように主張され得るが、主張された組合せからの1つまたは複数の特征是、場合によっては組合せから削除することができ、主張された組合せはサブコンビネーションまたはサブコンビネーションのバリエーションに向けられ得る。

30

#### 【0111】

同様に、動作は図面において特定の順序で描かれているが、これは、望ましい結果を達成するために、そのような動作が示されている特定の順序または連続した順序で実行されること、あるいはすべての説明された動作が実行されることを必要とするものとして理解されるべきではない。特定の環境では、マルチタスキングおよび並列処理が有利な場合がある。さらに、上記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離は、すべての実施形態においてそのような分離を必要とするものとして理解されるべきではなく、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一般に単一のソフトウェア製品に一緒に統合することもでき、複数のソフトウェア製品にパッケージ化されることが理解されるべきである。したがって、主題の特定の实装形態が説明された。他の实装形態は、以下の特許請求の範囲内にある。場合によっては、特許請求の範囲に記載されているアクショ

40

50

ンを異なる順序で実行しても、依然として望ましい結果を達成することができる。さらに、添付の図面に描かれているプロセスは、望ましい結果を達成するために、示されている特定の順序、または連続した順序を必ずしも必要としない。特定の実装形態では、マルチタスキングおよび並列処理が有利な場合がある。

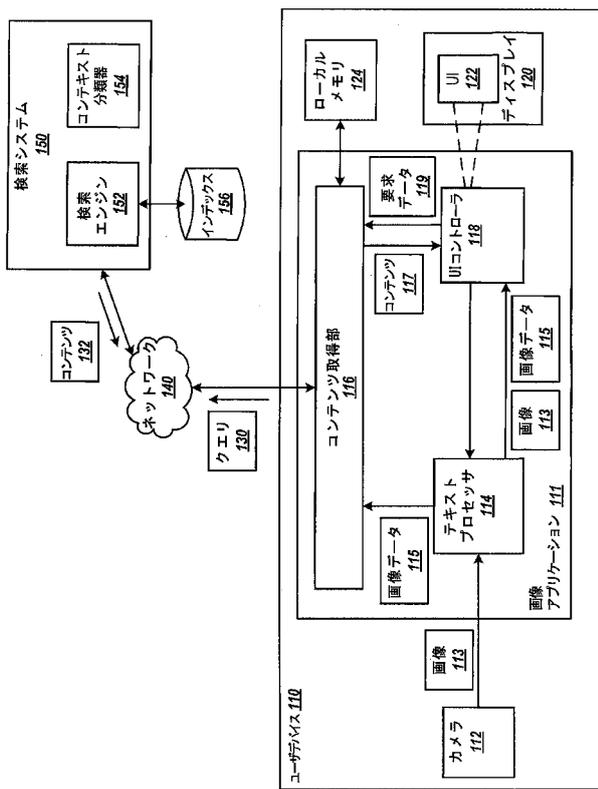
【符号の説明】

【0112】

100	例示的な環境	
110	ユーザデバイス	
111	画像アプリケーション	
112	カメラ	10
113	画像	
114	テキストプロセッサ	
115	画像データ	
116	コンテンツ取得部	
118	ユーザインターフェースコントローラ、コンテンツ取得部	
118	UIコントローラ	
119	要求データ	
120	ディスプレイ	
122	ユーザインターフェース	
123	要求データ	20
124	ローカルメモリ	
130	クエリデータ	
132	コンテンツ	
140	データ通信ネットワーク	
150	検索システム	
152	検索エンジン	
154	コンテキスト分類器	
156	インデックス	
205	ユーザインターフェース	
210	第1のスクリーンショット	30
211	画像	
220	第2のスクリーンショット	
221	ユーザ選択可能なターゲット	
222	ユーザ選択可能なターゲット	
223	ユーザ選択可能なターゲット	
224	ユーザ選択可能なターゲット	
226	ポインタ	
227	アレルギーインジケータ	
230	スクリーンショット	
231	ユーザ選択可能なターゲット	40
232	ユーザ選択可能なターゲット	
233	ユーザ選択可能なターゲット	
234	ユーザ選択可能なターゲット	
236	ポインタ	
240	スクリーンショット	
241	ユーザ選択可能なターゲット	
243	ユーザ選択可能なターゲット	
305	ユーザインターフェース	
310	スクリーンショット	
311 ~ 314	ユーザ選択可能なターゲット	50

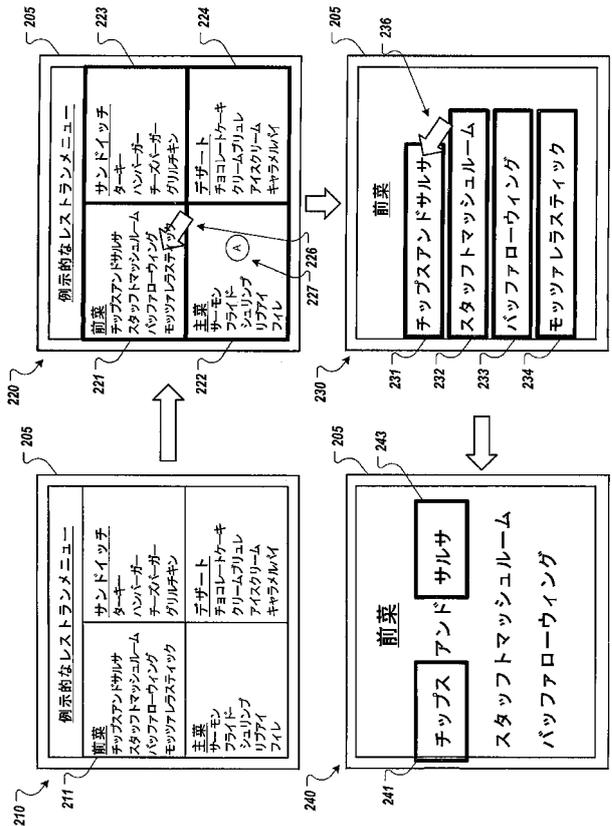
- 311 ユーザ選択可能なターゲット
- 316 ポインタ
- 320 スクリーンショット
- 322 テキストブロック
- 323 画像
- 324 画像
- 405 ユーザインターフェース
- 410 スクリーンショット
- 411 ~ 414 ユーザ選択可能なターゲット
- 411 ユーザ選択可能なターゲット
- 416 ポインタ
- 420 スクリーンショット
- 421 検索結果ページ
- 422 検索ボックス
- 423 ~ 425 検索結果
- 500 プロセス
- 600 プロセス

【 図 1 】



100

【 図 2 】



210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

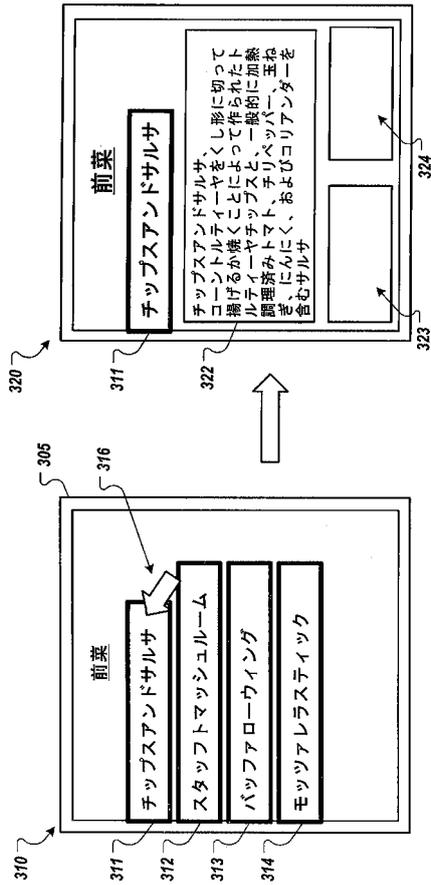
504

505

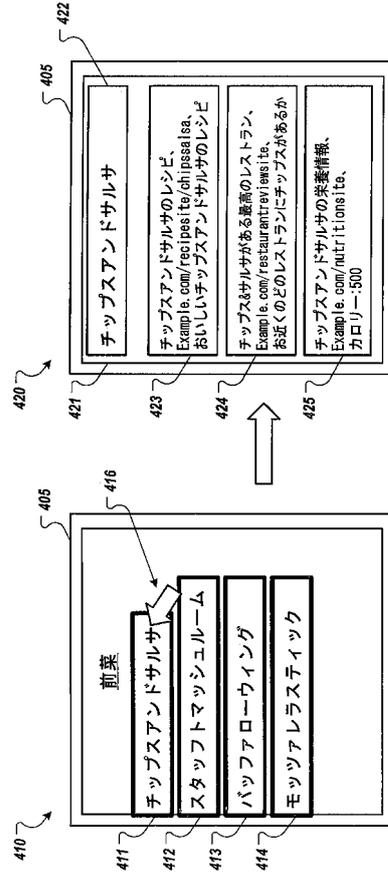
506

507</

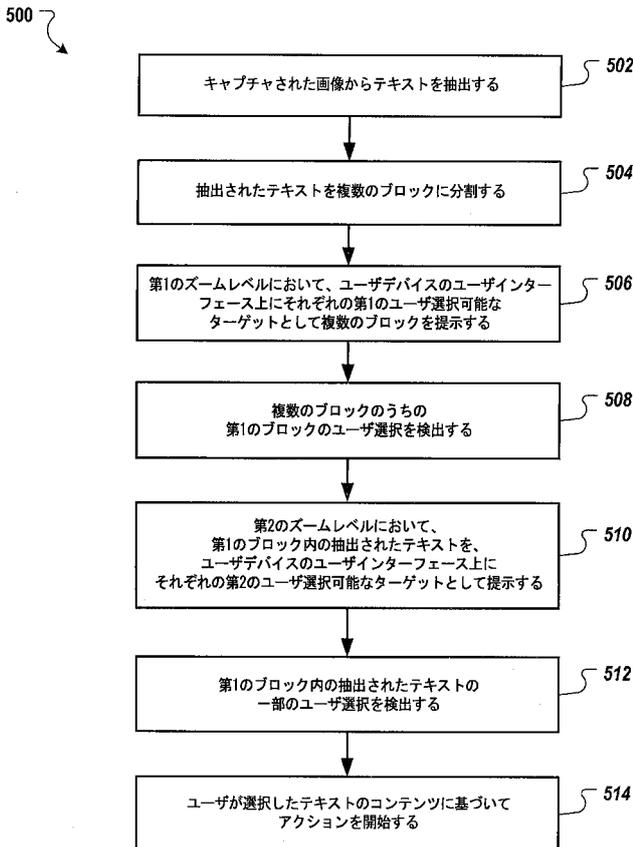
【 図 3 】



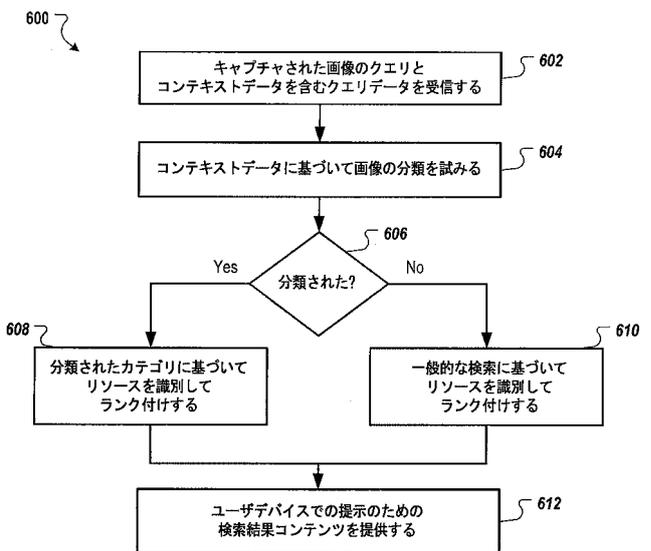
【 図 4 】



【 図 5 】



【 図 6 】



## 【手続補正書】

【提出日】令和1年11月22日(2019.11.22)

## 【手続補正1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】

ユーザデバイスによって、キャプチャされた画像からテキストを抽出するステップと、  
前記ユーザデバイスによって、前記抽出されたテキストを複数のブロックに分割するステップと、

前記ユーザデバイスによって、第1のズームレベルにおいて、前記ユーザデバイスのユーザインターフェース上にそれぞれのユーザ選択可能なターゲットとして前記ブロックを提示するステップと、

前記複数のブロックのうちの第1のブロックのユーザ選択を検出するステップと、

前記第1のブロックの前記ユーザ選択の検出に応答して、前記ユーザデバイスによって、前記第1のズームレベルより大きい第2のズームレベルにおいて、前記第1のブロック内の前記抽出されたテキストの一部を、前記ユーザデバイスの前記ユーザインターフェース上にそれぞれの第2のユーザ選択可能なターゲットとして提示するステップと、

前記第1のブロック内の前記抽出されたテキストの一部のユーザ選択の検出に応答して、前記ユーザデバイスによって、前記ユーザ選択のテキストのコンテンツに基づいてアクションを開始するステップと

を備える、方法。

【請求項2】

前記アクションを開始するステップが、前記ユーザ選択のテキストによって識別される宛先アドレスへの通信を開始するステップを備える、請求項1に記載の方法。

【請求項3】

前記アクションを開始するステップが、

前記ユーザ選択のテキストの前記コンテンツに基づいて検索クエリを生成するステップと、

前記ユーザデバイスによって、検索エンジンに前記検索クエリを送信するステップとを備える、請求項1に記載の方法。

【請求項4】

前記アクションを開始するステップが、

前記ユーザ選択のテキストの前記コンテンツに基づいて、2つ以上の候補検索クエリを生成するステップと、

前記ユーザインターフェースに前記候補検索クエリを表示するステップと、

前記候補検索クエリのうちの1つのユーザ選択に応じて、前記ユーザデバイスによって、選択された候補検索クエリを検索エンジンに送信するステップと

を備える、請求項1に記載の方法。

【請求項5】

前記ユーザデバイスによって、前記画像からコンテキストを抽出するステップをさらに備え、

前記検索クエリまたは前記候補検索クエリを生成するステップが、前記画像の前記抽出されたコンテキストにさらに基づいて、前記検索クエリまたは前記候補検索クエリを生成するステップを備える、請求項3または4に記載の方法。

【請求項6】

前記ユーザデバイスによって、前記画像からコンテキストを抽出するステップと、

前記ユーザデバイスによって、前記画像の前記抽出されたコンテキストを使用して、前

記画像が複数のあらかじめ定義されたカテゴリの1つに該当するかどうかを決定するステップと

をさらに備え、

前記検索クエリまたは前記候補検索クエリを生成するステップが、前記画像が前記あらかじめ定義されたカテゴリの1つに該当する場合、前記抽出されたテキストにさらに基づいて前記検索クエリまたは前記候補検索クエリを生成するステップを備え、該当しない場合、前記検索クエリまたは前記候補検索クエリを生成するステップが、前記ユーザ選択のテキストの前記コンテンツのみに基づく、請求項3または4に記載の方法。

【請求項7】

ユーザが前記ブロックのうちのいずれかを選択する前に、

少なくとも1つのブロックについて、前記ユーザデバイスによって、前記ブロックの前記抽出されたテキストに基づいて1つまたは複数の検索クエリを生成するステップと、

前記ユーザデバイスから検索エンジンに前記1つまたは複数の検索クエリを送信するステップと、

をさらに備える、請求項1から6のいずれか一項に記載の方法。

【請求項8】

前記ユーザデバイスが、前記ユーザインターフェース上にユーザ選択可能なそれぞれのターゲットとして前記ブロックを提示する前に、前記1つまたは複数の検索クエリを生成して送信する、請求項7に記載の方法。

【請求項9】

前記ユーザデバイスにおいて、前記1つまたは複数の検索クエリに基づいて前記検索エンジンから検索結果コンテンツを受信するステップと、

前記受信した検索結果コンテンツを前記ユーザデバイスのローカルメモリに記憶するステップと

をさらに備える、請求項7または8に記載の方法。

【請求項10】

前記ブロックの前記抽出されたテキストに基づいて前記1つまたは複数の検索クエリを生成するステップが、

前記ブロックの前記抽出されたテキストを、第1のレベルのテキストベースの粒度においてテキストの複数の第1のセットに解析するステップと、

テキストの第1のセットごとに検索クエリを生成するステップと

を備える、請求項7、8、または9に記載の方法。

【請求項11】

前記ブロックの前記ユーザ選択の検出に応答して、

前記ユーザデバイスによって、前記ブロックの前記抽出されたテキストを、テキストベースの粒度の前記第1のレベルよりも大きい第2のレベルのテキストベースの粒度においてテキストの複数の第2のセットに解析するステップと、

前記第1のブロック内の前記抽出されたテキストの一部の前記ユーザ選択を受信する前に、

前記ユーザデバイスによって、テキストの第2のセットごとに追加の検索クエリを生成するステップと、

前記ユーザデバイスから前記検索エンジンに各追加の検索クエリを送信するステップと、

前記ユーザデバイスによって、および前記検索エンジンから、前記追加の検索クエリに基づく追加の検索結果コンテンツを受信するステップと、

前記追加の検索結果コンテンツを前記ユーザデバイスのローカルメモリに記憶するステップと

をさらに備える、請求項10に記載の方法。

【請求項12】

前記抽出されたテキストを前記ブロックに分割するステップが、前記抽出されたテキス

トの意味解析に少なくとも部分的に基づく、請求項1から11のいずれか一項に記載の方法。

【請求項13】

前記ユーザインターフェース上に、検索エンジンから前記ユーザデバイスにおいて受信したコンテンツを表示するステップをさらに備える、請求項3から11のいずれか一項、または請求項3から11のいずれか一項に従属する請求項12に記載の方法。

【請求項14】

1つまたは複数のコンピュータによって実行されると、前記1つまたは複数のコンピュータに、請求項1から13のいずれか一項において定義された方法を実行させる命令でエンコードされたコンピュータストレージ媒体。

【請求項15】

請求項1から13のいずれか一項において定義された方法を実行するように構成された、プロセッサ。

【請求項16】

ユーザデバイスによって、画像からテキストを抽出するステップと、  
前記ユーザデバイスによって、前記抽出されたテキストを複数のブロックに分割するステップと、  
前記ユーザデバイスによって、第1のズームレベルにおいて、前記ユーザデバイスのユーザインターフェース上にそれぞれの第1のユーザ選択可能なターゲットとして前記複数のブロックを提示するステップと、  
前記複数のブロックのうちの第1のブロックのユーザ選択を検出するステップと、  
前記第1のブロックの前記ユーザ選択の検出にตอบสนองして、前記ユーザデバイスによって、前記第1のズームレベルより大きい第2のズームレベルにおいて、前記第1のブロック内の前記抽出されたテキストの一部を、前記ユーザデバイスの前記ユーザインターフェース上にそれぞれの第2のユーザ選択可能なターゲットとして提示するステップと、  
前記第1のブロック内の前記抽出されたテキストの一部のユーザ選択の検出にตอบสนองして、前記ユーザデバイスによって、前記ユーザ選択のテキストのコンテンツに基づいてアクションを開始するステップと  
を備える、方法。

【請求項17】

前記アクションを開始するステップが、前記ユーザ選択のテキストによって識別される宛先アドレスへの通信を開始するステップを備える、請求項16に記載の方法。

【請求項18】

前記アクションを開始するステップが、  
前記ユーザ選択のテキストの前記コンテンツに基づいて検索クエリを生成するステップと、  
前記ユーザデバイスによって、検索エンジンに前記検索クエリを送信するステップと  
を備える、請求項16に記載の方法。

【請求項19】

前記ユーザデバイスによって、前記画像からコンテキストを抽出するステップと、  
前記画像の前記抽出されたコンテキストを使用して、前記画像が複数のあらかじめ定義されたカテゴリの1つに属すると決定するステップと、  
前記画像が前記あらかじめ定義されたカテゴリの1つに属すると決定するステップにตอบสนองして、前記抽出されたコンテキストに基づいて前記検索クエリを調整するステップと  
をさらに備える、請求項18に記載の方法。

【請求項20】

前記ユーザデバイスによって、および前記検索エンジンから、前記検索クエリに基づいて選択された検索結果コンテンツを受信するステップをさらに備え、前記検索エンジンが、前記画像が属する前記1つのあらかじめ定義されたカテゴリに基づいて検索結果をランク付けする、請求項19に記載の方法。

## 【請求項 2 1】

前記ユーザデバイスによって、前記画像からコンテキストを抽出するステップと、  
前記抽出された前記画像のコンテキストを使用して、前記画像が複数のあらかじめ定義されたカテゴリの1つに属しないと決定するステップと  
をさらに備え、

前記検索クエリを生成するステップが、前記ユーザ選択のテキストのみに基づいて一般的な検索クエリを生成するステップを備える、請求項18に記載の方法。

## 【請求項 2 2】

前記アクションを開始するステップが、  
前記ユーザ選択のテキストの前記コンテンツに基づいて、2つ以上の候補検索クエリを生成するステップと、  
前記ユーザインターフェースに前記候補検索クエリを表示するステップと、  
前記候補検索クエリのうちの1つのユーザ選択に応じて、前記ユーザデバイスによって、前記ユーザ選択の候補検索クエリを検索エンジンに送信するステップと  
を備える、請求項16に記載の方法。

## 【請求項 2 3】

ユーザが前記ブロックのうちのいずれかを選択する前に、  
前記第1のブロックについて、前記ユーザデバイスによって、前記第1のブロックの前記抽出されたテキストに基づいて1つまたは複数の検索クエリを生成するステップと、  
前記ユーザデバイスから検索エンジンに前記1つまたは複数の検索クエリを送信するステップと、  
をさらに備える、請求項16に記載の方法。

## 【請求項 2 4】

前記ユーザデバイスが、前記複数のブロックを前記ユーザインターフェース上にそれぞれの第1のユーザ選択可能なターゲットとして提示する前に、前記1つまたは複数の検索クエリを生成して送信する、請求項23に記載の方法。

## 【請求項 2 5】

前記ユーザデバイスによって、前記1つまたは複数の検索クエリに基づいて前記検索エンジンから検索結果コンテンツを受信するステップと、  
前記受信した検索結果コンテンツを前記ユーザデバイスのローカルメモリに記憶するステップと  
をさらに備える、請求項23に記載の方法。

## 【請求項 2 6】

前記第1のブロックの前記抽出されたテキストに基づいて前記1つまたは複数の検索クエリを生成するステップが、  
前記第1のブロックの前記抽出されたテキストを、第1のレベルのテキストベースの粒度においてテキストの複数の第1のセットに解析するステップと、  
テキストの第1のセットごとに検索クエリを生成するステップと  
を備える、請求項23に記載の方法。

## 【請求項 2 7】

前記第1のブロックの前記ユーザ選択の検出にตอบสนองして、  
前記ユーザデバイスによって、前記第1のブロックの前記抽出されたテキストを、前記第1のレベルのテキストベースの粒度よりも大きい第2のレベルのテキストベースの粒度においてテキストの複数の第2のセットに解析するステップと、  
前記第1のブロック内の前記抽出されたテキストの一部の前記ユーザ選択を受信する前に、  
前記ユーザデバイスによって、テキストの第2のセットごとに追加の検索クエリを生成するステップと、  
前記ユーザデバイスから前記検索エンジンに各追加の検索クエリを送信するステップと、

前記ユーザデバイスによって、および前記検索エンジンから、前記追加の検索クエリに基づく追加の検索結果コンテンツを受信するステップと、

前記追加の検索結果コンテンツを前記ユーザデバイスのローカルメモリに記憶するステップと

をさらに備える、請求項26に記載の方法。

【請求項28】

前記抽出されたテキストを前記ブロックに分割するステップが、前記抽出されたテキストの意味解析に少なくとも部分的に基づく、請求項16に記載の方法。

【請求項29】

前記ユーザインターフェース上に、検索エンジンから前記ユーザデバイスにおいて受信したコンテンツを表示するステップをさらに備える、請求項16に記載の方法。

【請求項30】

前記ユーザインターフェース上に、前記検索エンジンから前記ユーザデバイスにおいて受信したコンテンツを表示するステップが、前記画像の少なくとも一部とともに前記コンテンツを前記ユーザデバイスのビューポートに表示するステップを備える、請求項29に記載の方法。

【請求項31】

データ処理装置と、

前記データ処理装置とデータ通信し、前記データ処理装置によって実行可能な命令を記憶するメモリ装置であって、前記命令が、そのような実行時に、前記データ処理装置に、

画像からテキストを抽出することと、

前記抽出されたテキストを複数のブロックに分割することと、

第1のズームレベルにおいて、ユーザインターフェース上にそれぞれの第1のユーザ選択可能なターゲットとして前記複数のブロックを提示することと、

前記複数のブロックのうちの第1のブロックのユーザ選択を検出することと、

前記第1のブロックの前記ユーザ選択の検出にตอบสนองして、前記第1のズームレベルより大きい第2のズームレベルにおいて、前記第1のブロック内の前記抽出されたテキストの一部を、前記ユーザインターフェース上にそれぞれの第2のユーザ選択可能なターゲットとして提示することと、

前記第1のブロック内の前記抽出されたテキストの一部のユーザ選択の検出にตอบสนองして、前記ユーザ選択のテキストのコンテンツに基づいてアクションを開始することと

を備える動作を実行させる、メモリ装置と

を備える、システム。

【請求項32】

前記アクションを開始することが、前記ユーザ選択のテキストによって識別される宛先アドレスへの通信を開始することを備える、請求項31に記載のシステム。

【請求項33】

前記アクションを開始することが、

前記ユーザ選択のテキストの前記コンテンツに基づいて検索クエリを生成することと、

検索エンジンに前記検索クエリを送信することと

を備える、請求項31に記載のシステム。

【請求項34】

前記動作が、

前記画像からコンテキストを抽出することと、

前記画像の前記抽出されたコンテキストを使用して、前記画像が複数のあらかじめ定義されたカテゴリの1つに属すると決定することと、

前記画像が前記あらかじめ定義されたカテゴリの1つに属すると決定することに対応して、前記抽出されたコンテキストに基づいて前記検索クエリを調整することと

をさらに備える、請求項33に記載のシステム。

【請求項35】

前記動作が、前記検索エンジンから、前記検索クエリに基づいて選択された検索結果コンテンツを受信することをさらに備え、前記検索エンジンが、前記画像が属する前記1つのあらかじめ定義されたカテゴリに基づいて検索結果をランク付けする、請求項34に記載のシステム。

【請求項36】

前記動作が、  
前記画像からコンテキストを抽出することと、  
前記抽出された前記画像のコンテキストを使用して、前記画像が複数のあらかじめ定義されたカテゴリの1つに属しないと決定することと  
をさらに備え、  
前記検索クエリを生成することが、前記ユーザ選択のテキストのみに基づいて一般的な検索クエリを生成することを備える、請求項33に記載のシステム。

【請求項37】

前記アクションを開始することが、  
前記ユーザ選択のテキストの前記コンテンツに基づいて、2つ以上の候補検索クエリを生成することと、  
前記ユーザインターフェースに前記候補検索クエリを表示することと、  
前記候補検索クエリのうちの1つのユーザ選択に応じて、選択された候補検索クエリを検索エンジンに送信することと  
を備える、請求項31に記載のシステム。

【請求項38】

前記動作が、ユーザが前記ブロックのうちのいずれかを選択する前に、  
前記第1のブロックについて、前記第1のブロックの前記抽出されたテキストに基づいて1つまたは複数の検索クエリを生成することと、  
検索エンジンに前記1つまたは複数の検索クエリを送信することと、  
をさらに備える、請求項31に記載のシステム。

【請求項39】

前記データ処理装置が、前記複数のブロックを前記ユーザインターフェース上にそれぞれの第1のユーザ選択可能なターゲットとして提示する前に、前記1つまたは複数の検索クエリを生成して送信する、請求項38に記載のシステム。

【請求項40】

前記動作が、  
前記1つまたは複数の検索クエリに基づいて前記検索エンジンから検索結果コンテンツを受信することと、  
前記受信した検索結果コンテンツを前記データ処理装置のローカルメモリに記憶することと  
をさらに備える、請求項38に記載のシステム。

【請求項41】

前記第1のブロックの前記抽出されたテキストに基づいて前記1つまたは複数の検索クエリを生成することが、  
前記第1のブロックの前記抽出されたテキストを、第1のレベルのテキストベースの粒度においてテキストの複数の第1のセットに解析することと、  
テキストの第1のセットごとに検索クエリを生成することと  
を備える、請求項38に記載のシステム。

【請求項42】

前記動作が、  
前記第1のブロックの前記ユーザ選択の検出にตอบสนองして、  
前記第1のブロックの前記抽出されたテキストを、前記第1のレベルのテキストベースの粒度よりも大きい第2のレベルのテキストベースの粒度においてテキストの複数の第2のセットに解析することと、

前記第1のブロック内の前記抽出されたテキストの一部の前記ユーザ選択を受信する前に、

テキストの第2のセットごとに追加の検索クエリを生成することと、

前記検索エンジンに各追加の検索クエリを送信することと、

前記検索エンジンから、前記追加の検索クエリに基づく追加の検索結果コンテンツを受信することと、

前記追加の検索結果コンテンツを前記データ処理装置のローカルメモリに記憶することと

をさらに備える、請求項41に記載のシステム。

【請求項43】

前記抽出されたテキストを前記ブロックに分割することが、前記抽出されたテキストの意味解析に少なくとも部分的に基づき、請求項31に記載のシステム。

【請求項44】

前記動作が、検索エンジンから受信したコンテンツを前記ユーザインターフェース上に表示することをさらに備える、請求項31に記載のシステム。

【請求項45】

前記検索エンジンから受信したコンテンツを前記ユーザインターフェース上に表示することが、前記画像の少なくとも一部とともに前記データ処理装置のビューポートに前記コンテンツを表示することを備える、請求項44に記載のシステム。

【請求項46】

データ処理装置によって実行されると、前記データ処理装置に、

画像からテキストを抽出することと、

前記抽出されたテキストを複数のブロックに分割することと、

第1のズームレベルにおいて、ユーザインターフェース上にそれぞれの第1のユーザ選択可能なターゲットとして前記複数のブロックを提示することと、

前記複数のブロックのうちの第1のブロックのユーザ選択を検出することと、

前記第1のブロックの前記ユーザ選択の検出にตอบสนองして、前記第1のズームレベルより大きい第2のズームレベルにおいて、前記第1のブロック内の前記抽出されたテキストの一部を、前記ユーザインターフェース上にそれぞれの第2のユーザ選択可能なターゲットとして提示することと、

前記第1のブロック内の前記抽出されたテキストの一部のユーザ選択の検出にตอบสนองして、前記ユーザ選択のテキストのコンテンツに基づいてアクションを開始することと

を備える動作を実行させる命令を記憶した、非一時的コンピュータ可読ストレージ媒体。

【請求項47】

前記アクションを開始することが、前記ユーザ選択のテキストによって識別される宛先アドレスへの通信を開始することを備える、請求項46に記載の非一時的コンピュータ可読ストレージ媒体。

【請求項48】

前記アクションを開始することが、

前記ユーザ選択のテキストの前記コンテンツに基づいて検索クエリを生成するステップと、

検索エンジンに前記検索クエリを送信することと

を備える、請求項46に記載の非一時的コンピュータ可読ストレージ媒体。

【請求項49】

前記動作が、

前記画像からコンテキストを抽出することと、

前記画像の前記抽出されたコンテキストを使用して、前記画像が複数のあらかじめ定義されたカテゴリの1つに属すると決定することと、

前記画像が前記あらかじめ定義されたカテゴリの1つに属すると決定することにตอบสนองして、前記抽出されたコンテキストに基づいて前記検索クエリを調整することと

をさらに備える、請求項48に記載の非一時的コンピュータ可読ストレージ媒体。

【請求項50】

前記動作が、前記検索エンジンから、前記検索クエリに基づいて選択された検索結果コンテンツを受信することをさらに備え、前記検索エンジンが、前記画像が属する前記1つのあらかじめ定義されたカテゴリに基づいて検索結果をランク付けする、請求項49に記載の非一時的コンピュータ可読ストレージ媒体。

【請求項51】

前記動作が、  
前記画像からコンテキストを抽出することと、  
前記抽出された前記画像のコンテキストを使用して、前記画像が複数のあらかじめ定義されたカテゴリの1つに属しないと決定することと  
をさらに備え、

前記検索クエリを生成することが、前記ユーザ選択のテキストのみに基づいて一般的な検索クエリを生成することを備える、請求項48に記載の非一時的コンピュータ可読ストレージ媒体。

【請求項52】

前記アクションを開始することが、  
前記ユーザ選択のテキストの前記コンテンツに基づいて、2つ以上の候補検索クエリを生成することと、  
前記ユーザインターフェースに前記候補検索クエリを表示することと、  
前記候補検索クエリのうちの1つのユーザ選択に応じて、前記ユーザ選択の候補検索クエリを検索エンジンに送信することと  
を備える、請求項46に記載の非一時的コンピュータ可読ストレージ媒体。

【請求項53】

前記動作が、ユーザが前記ブロックのうちのいずれかを選択する前に、  
前記第1のブロックについて、前記第1のブロックの前記抽出されたテキストに基づいて1つまたは複数の検索クエリを生成することと、  
検索エンジンに前記1つまたは複数の検索クエリを送信することと、  
をさらに備える、請求項46に記載の非一時的コンピュータ可読ストレージ媒体。

【請求項54】

前記データ処理装置が、前記複数のブロックを前記ユーザインターフェース上にそれぞれの第1のユーザ選択可能なターゲットとして提示する前に、前記1つまたは複数の検索クエリを生成して送信する、請求項53に記載の非一時的コンピュータ可読ストレージ媒体。

【請求項55】

前記動作が、  
前記1つまたは複数の検索クエリに基づいて前記検索エンジンから検索結果コンテンツを受信することと、  
前記受信した検索結果コンテンツを前記データ処理装置のローカルメモリに記憶することと  
をさらに備える、請求項53に記載の非一時的コンピュータ可読ストレージ媒体。

【請求項56】

前記第1のブロックの前記抽出されたテキストに基づいて前記1つまたは複数の検索クエリを生成することが、  
前記第1のブロックの前記抽出されたテキストを、第1のレベルのテキストベースの粒度においてテキストの複数の第1のセットに解析することと、  
テキストの第1のセットごとに検索クエリを生成することと  
を備える、請求項53に記載の非一時的コンピュータ可読ストレージ媒体。

【請求項57】

前記動作が、  
前記第1のブロックの前記ユーザ選択の検出に応答して、

前記第1のブロックの前記抽出されたテキストを、前記第1のレベルのテキストベースの粒度よりも大きい第2のレベルのテキストベースの粒度においてテキストの複数の第2のセットに解析することと、

前記第1のブロック内の前記抽出されたテキストの一部の前記ユーザ選択を受信する前に、

テキストの第2のセットごとに追加の検索クエリを生成することと、

前記検索エンジンに各追加の検索クエリを送信することと、

前記検索エンジンから、前記追加の検索クエリに基づく追加の検索結果コンテンツを受信することと、

前記追加の検索結果コンテンツを前記データ処理装置のローカルメモリに記憶することと

をさらに備える、請求項56に記載の非一時的コンピュータ可読ストレージ媒体。

【請求項58】

前記抽出されたテキストを前記ブロックに分割することが、前記抽出されたテキストの意味解析に少なくとも部分的に基づく、請求項46に記載の非一時的コンピュータ可読ストレージ媒体。

【請求項59】

前記動作が、検索エンジンから受信したコンテンツを前記ユーザインターフェース上に表示することをさらに備える、請求項46に記載の非一時的コンピュータ可読ストレージ媒体。

【請求項60】

前記検索エンジンから受信したコンテンツを前記ユーザインターフェース上に表示することが、前記画像の少なくとも一部とともに前記データ処理装置のビューポートに前記コンテンツを表示することを備える、請求項59に記載の非一時的コンピュータ可読ストレージ媒体。

## 【 国際調査報告 】

## INTERNATIONAL SEARCH REPORT

International application No PCT/US2017/051333
---

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> INV. G06F17/30 ADD.		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) G06F		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal, WPI Data		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2011/017557 A1 (GOOGLE INC [US]; PETROU DAVID [US]) 10 February 2011 (2011-02-10) page 1 - page 37 -----	1-60
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 20 October 2017		Date of mailing of the international search report 30/10/2017
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Authorized officer Korkuzas, Valdas

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International application No

PCT/US2017/051333

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2011017557	A1	10-02-2011	
		AU 2010279333	A1 15-03-2012
		AU 2013205924	A1 06-06-2013
		CA 2771094	A1 10-02-2011
		CN 102625937	A 01-08-2012
		EP 2462520	A1 13-06-2012
		JP 5933677	B2 15-06-2016
		JP 6148367	B2 14-06-2017
		JP 2013501975	A 17-01-2013
		JP 2015064901	A 09-04-2015
		JP 2016139424	A 04-08-2016
		KR 20120058538	A 07-06-2012
		KR 20160092045	A 03-08-2016
		US 2011125735	A1 26-05-2011
		US 2014164406	A1 12-06-2014
		WO 2011017557	A1 10-02-2011

## フロントページの続き

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

(72)発明者 ルイス・ワン

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

(72)発明者 チャールズ・ジェイ・ローゼンバーグ

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

Fターム(参考) 5B175 DA02 FA01 GA03