

(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2015-522877
(P2015-522877A)

(43) 公表日 平成27年8月6日(2015.8.6)

(51) Int.Cl.	F I			テーマコード (参考)
G06K 9/20 (2006.01)	G06K	9/20	340L	5B029
G06K 9/38 (2006.01)	G06K	9/38	A	5C077
G06T 7/40 (2006.01)	G06T	7/40	100A	5L096
H04N 1/40 (2006.01)	H04N	1/40	F	

審査請求 有 予備審査請求 未請求 (全 29 頁)

(21) 出願番号 特願2015-516152 (P2015-516152)
 (86) (22) 出願日 平成25年6月4日 (2013.6.4)
 (85) 翻訳文提出日 平成27年1月30日 (2015.1.30)
 (86) 国際出願番号 PCT/US2013/044182
 (87) 国際公開番号 W02013/184726
 (87) 国際公開日 平成25年12月12日 (2013.12.12)
 (31) 優先権主張番号 13/491, 398
 (32) 優先日 平成24年6月7日 (2012.6.7)
 (33) 優先権主張国 米国 (US)

(71) 出願人 506223509
 アマゾン・テクノロジーズ、インコーポレイテッド
 アメリカ合衆国、ネバダ州 89507、
 レノ、ピー、オー、ボックス 8102
 P. O. Box 8102, Reno, Nevada 89507, U. S. A.
 (74) 代理人 100108855
 弁理士 蔵田 昌俊
 (74) 代理人 100103034
 弁理士 野河 信久
 (74) 代理人 100075672
 弁理士 峰 隆司
 (74) 代理人 100153051
 弁理士 河野 直樹

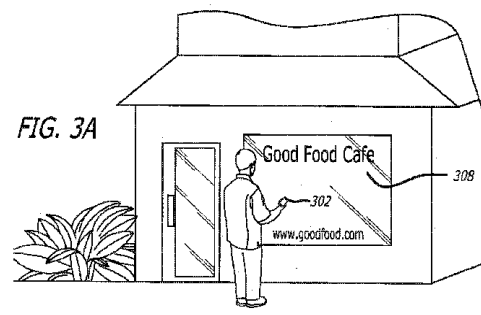
最終頁に続く

(54) 【発明の名称】 画像認識のための適応閾値処理

(57) 【要約】

アプリケーション、システム、またはサービスにテキスト情報を提供するための種々の手法が開示される。具体的には、種々の実施形態は、ユーザが携帯型コンピューティングデバイスのカメラによって画像をキャプチャすることを可能にする。コンピューティングデバイスは、テキストを転送してアクションを実質的にリアルタイムで実施するために、画像を取り込み、これを処理してテキストを認識、特定、および/または分離する能力を持つ。テキストは、Eメール、電話番号、URL、アドレスおよび同等物を含み得、アプリケーションまたは機能は、電話番号をダイヤルすること、URLへ誘導すること、アドレス帳を開いて連絡先情報を保存すること、地図を表示してアドレスを示すことなどであり得る。適応閾値処理は、テキスト認識工程の正確性および効率性を向上させるために、画像全体にわたる変化を説明するために用いられ得る。

【選択図】 図3A



【特許請求の範囲】**【請求項 1】**

コンピュータ実装方法であって、

実行可能な命令で構成される 1 つ以上のコンピュータシステムの制御下において、複数のピクセルについてのピクセル値を含む 1 つの画像中の 1 つの領域の 1 つ以上の輝度勾配を分析して、1 つ以上の可能性のあるエッジ位置を特定することと、

前記 1 つ以上の可能性のあるエッジ位置の所定距離の範囲内にある 1 組のピクセルについてのピクセル値を分析することと、

前記 1 つ以上の可能性のあるエッジ位置のそれぞれの前記所定距離の範囲内にある前記ピクセル値に少なくとも部分的に基づいて、前記領域のための文字ピクセル値の閾を判定することと、

前記文字ピクセル値の閾と少なくとも合致するピクセル値を有する前記 1 組のピクセルの一部分に少なくとも部分的に基づいて、1 つ以上の文字部分を判定することと、を含む、コンピュータ実装方法。

【請求項 2】

文字認識アルゴリズムを用いて前記 1 つ以上の文字部分を分析して、前記 1 つ以上の文字部分に関連付けられる少なくとも一文字を認識することをさらに含む、請求項 1 に記載のコンピュータ実装方法。

【請求項 3】

エッジ位置が、前記所定距離の範囲内の前記文字ピクセル値の閾に少なくとも合致する値を有するピクセル輝度勾配またはピクセル飽和勾配に対応する、請求項 1 に記載のコンピュータ実装方法。

【請求項 4】

前記ピクセル値を分析することは、前記 1 つ以上の可能性のあるエッジ位置の第 1 の側にある第 1 の組のピクセルを分析することと、前記 1 つ以上の可能性のあるエッジ位置の第 2 の側にある第 2 の組のピクセルを分析することとを含む、請求項 1 に記載のコンピュータ実装方法。

【請求項 5】

前記ピクセル値は、グレースケールピクセル輝度値、カラーピクセル輝度値、ピクセル飽和値、またはピクセル飽和レベルのうちの少なくとも 1 つに対応する、請求項 1 に記載のコンピュータ実装方法。

【請求項 6】

テキストライン構成アルゴリズムを用いて前記 1 つ以上の文字部分の関連する個別の文字要素を判定することをさらに含む、請求項 1 に記載のコンピュータ実装方法。

【請求項 7】

関連する個別の文字要素を判定した後に、前記個別の文字要素の支配的方向を判定することをさらに含む、請求項 6 に記載のコンピュータ実装方法。

【請求項 8】

前記個別の文字要素が少なくとも 1 つの既定の文字制約を満たすかどうかを判定することと、

個別の文字要素が既定の文字制約のうちの前記少なくとも 1 つを満たさない場合に、個別の文字要素を検討の対象から外すことと、をさらに含む、請求項 6 に記載のコンピュータ実装方法。

【請求項 9】

前記既定の文字制約は、文字要素間の間隔、隣接する文字要素の高さと幅の比率、または文字要素の垂直配列のうちの 1 つを含む、請求項 8 に記載のコンピュータ実装方法。

【請求項 10】

結合コンポーネントアルゴリズムを用いて、前記文字ピクセル値の閾に少なくとも合致するピクセル値を有する前記複数のピクセルのピクセルを関連付けることと、

ストローク幅変換アルゴリズムを前記関連付けられたピクセルに適用することと、

10

20

30

40

50

前記ストローク幅変換アルゴリズムを適用した結果を文字認識アルゴリズムに提供することと、をさらに含む、請求項5に記載のコンピュータ実装方法。

【請求項11】

プロセッサと、

前記プロセッサによって実行されると、前記プロセッサにさらに、

1つの画像内の1つの領域の1つ以上の輝度勾配を分析して、1つ以上の可能性のある文字エッジ位置を特定させ、

前記1つ以上の可能性のあるエッジ位置の所定距離の範囲内にある1組のピクセルについてのピクセル値を分析させ、

前記1つ以上のエッジ位置のそれぞれの前記所定距離の範囲内にある前記ピクセル値に少なくとも部分的に基づいて、前記領域のための文字ピクセル値の閾を判定させ、

前記文字ピクセル値の閾と少なくとも合致するピクセル値を有する前記1組のピクセルの一部分に少なくとも部分的に基づいて、1つ以上の文字部分を判定させる命令を含む記憶デバイスと、を備える、携帯型コンピューティングデバイス。

10

【請求項12】

前記命令は、前記プロセッサによって実行されると、

前記携帯型コンピューティングデバイスが、文字認識アルゴリズムを用いて前記文字部分を分析して、前記1つ以上の文字部分に関連付けられる少なくとも1つの文字を認識することをさらに可能にする、請求項11に記載の携帯型コンピューティングデバイス。

【請求項13】

20

前記命令は、前記プロセッサによって実行されると、

前記携帯型コンピューティングデバイスが、

前記認識された少なくとも1つの文字に対応する文字パターンを特定し、

前記文字パターンに関連付けられる前記携帯型コンピューティングデバイス上のアプリケーションを判定し、

前記判定されたアプリケーションに前記認識された少なくとも1つの文字を提供することをさらに可能にする、請求項12に記載の携帯型コンピューティングデバイス。

【請求項14】

ピクセル値を分析する前記命令は、

前記プロセッサによって実行されると、前記携帯型コンピューティングデバイスが、前記1つ以上の可能性のあるエッジ位置の第1の側にある第1の組のピクセルを分析することと、前記1つ以上の可能性のあるエッジ位置の第2の側にある第2の組のピクセルを分析することとをさらに可能にする命令をさらに含む、請求項11に記載の携帯型コンピューティングデバイス。

30

【請求項15】

前記命令は、前記プロセッサによって実行されると、

前記携帯型コンピューティングデバイスが、結合コンポーネントアルゴリズムを用いて、前記文字ピクセル値の閾に少なくとも合致するピクセル値を有する前記複数のピクセルの隣接するピクセルを関連付けて、文字の少なくとも一部分を形成することをさらに可能にする、請求項11に記載の携帯型コンピューティングデバイス。

40

【発明の詳細な説明】

【背景技術】

【0001】

人々がタブレットコンピュータおよびスマートフォンなどの携帯型デバイスを含むさまざまなコンピューティングデバイスをますます利用していることに伴い、人々がこれらのデバイスとやりとりする方途に適応することが好都合であり得る。ユーザは、電話番号を手動で入力して電話をかけること、Eメールアドレスを手動でタイプしてEメールを送信すること、ウェブアドレスをウェブブラウザに手動でタイプしてウェブページを閲覧することなどに慣れている。これらのタスクはしばしば単調であり、かつ時間がかかる。種々のデバイスにおいて、電話番号をお気に入りリストに割り当てる、ウェブアドレスをブッ

50

クマークするなど、ユーザの時間を節約するための種々の方法が提供されてきた。技術が進歩するにつれて、および携帯型コンピューティングデバイス上に提供される機能およびサービスが進歩および拡張するにつれて、手早い方法および他の時間短縮方法が提供される方途は変化している。

【図面の簡単な説明】

【0002】

本開示に従う種々の実施形態が図面を参照して記載される。

【図1A】一実施形態に従う、携帯型コンピューティングデバイスを保持するユーザが地図アプリケーションにテキストを提供する実施例を説明したものである。

【図1B】一実施形態に従う、携帯型コンピューティングデバイスを保持するユーザが地図アプリケーションにテキストを提供する実施例を説明したものである。

【図1C】一実施形態に従う、携帯型コンピューティングデバイスを保持するユーザが地図アプリケーションにテキストを提供する実施例を説明したものである。

【図2A】一実施形態に従う、携帯型コンピューティングデバイスを保持するユーザが電話アプリケーションに電話番号を提供する実施例を説明したものである。

【図2B】一実施形態に従う、携帯型コンピューティングデバイスを保持するユーザが電話アプリケーションに電話番号を提供する実施例を説明したものである。

【図2C】一実施形態に従う、携帯型コンピューティングデバイスを保持するユーザが電話アプリケーションに電話番号を提供する実施例を説明したものである。

【図3A】一実施形態に従う、携帯型コンピューティングデバイスを保持するユーザがブラウザにテキストを提供する実施例を説明したものである。

【図3B】一実施形態に従う、携帯型コンピューティングデバイスを保持するユーザがブラウザにテキストを提供する実施例を説明したものである。

【図3C】

【図4】種々の実施形態に従う、携帯型コンピューティングデバイス上のアプリケーションにテキストを提供するための工程の例を説明したものである。

【図5】種々の実施形態に従う、テキスト検出のための画像を前処理する実施例を説明したものである。

【図6】種々の実施形態に従う、ピクセル領域の拡大図を示したものである。

【図7A】種々の実施形態に従う、距離対勾配の輝度グラフの例を説明したものである。

【図7B】種々の実施形態に従う、距離対勾配の輝度グラフの例を説明したものである。

【図7C】種々の実施形態に従う、距離対勾配の輝度グラフの例を説明したものである。

【図8】種々の実施形態に従う、テキスト検出の工程の例を説明したものである。

【図9】種々の実施形態に従う、テキスト検出のための画像を前処理する実施例を説明したものである。

【図10A】種々の実施形態に従って利用され得るコンピューティングデバイスの正面図および後面図の例を示したものである。

【図10B】種々の実施形態に従って利用され得るコンピューティングデバイスの正面図および後面図の例を示したものである。

【図11】種々の実施形態に従って利用され得るコンピューティングデバイスのコンポーネントレベルの例を説明したものである。

【図12】種々の実施形態が実施され得る環境を説明したものである。

【発明を実施するための形態】

【0003】

本開示の種々の実施形態に従うシステムおよび方法は、テキスト情報をアプリケーション、システム、またはサービスに提供するための従来手法において経験される、前述の、および他の欠陥の1つ以上を克服し得る。具体的には、種々の実施形態は、携帯型コンピューティングデバイスが、テキストを含む画像情報を取得すること、画像内のテキストの場所を特定すること、テキストの種類またはパターン（例えば、Eメール、電話番号、URLなど）を特定すること、テキストに関連付けられる機能またはアプリケーション（

10

20

30

40

50

例えば、電話番号に電話する、インターネットブラウザを開くなど)を判定すること、および/またはテキストに関連付けられる機能またはアプリケーションを実施することなどのタスクを実施することを可能にする。少なくともこれらのタスクのうちのいくつかは、ネットワークを通して利用可能な少なくとも1つの資源をも用いて実施され得る。

【0004】

種々の実施形態は、ユーザがテキストを含むオブジェクトに、このオブジェクトの画像をキャプチャするためにカメラを向けることを可能にする。カメラは、テキストを認識、特定、および/または分離してテキストをアプリケーションまたは機能に送信するために、画像を取り込み、かつ画像を処理する(または処理のために画像を提供する)能力をもつ携帯型コンピューティングデバイスと一体化され得る。アプリケーションまたは機能は次に、テキストを利用して実質的にリアルタイムでアクションを実施することができる。例えば、通りを歩いているユーザは、電話番号を含む広告を見る可能性がある。ユーザは、本明細書に記載の教示を用いて、例えば、スマートフォンのカメラを例えば番号に向けて、この番号を手動で入力することなく、この番号を電話アプリケーションに入力することにより、この電話番号に電話することができる。一例において、デバイスは、画像から番号を入力し、通話を開始するというユーザの指示を待ち得、または他の例において、電話はデバイスによって自動的にかけられ得る。本例は、ウェブアドレスまたはURLをウェブブラウザ、新規のメッセージウィンドウを開くためのEメールアドレス、地図アプリケーションを開いて道順を表示するための物理的な住所、およびさまざまな他のアクション可能なデータ種類を入力することに拡張され得る。

【0005】

種々の実施形態は、照明、陰影、コントラスト、または他のかかる側面における変形に対してロバストな手法を用いて1つの画像の中のテキストを検出および認識することができる。例えば、カメラは、部分的に影で覆われた看板の画像をキャプチャし得る。少なくとも部分的にこの影によって、テキスト文字の色値と背景領域との間の分離は、影になっている領域と影になっていない領域との間で異なる。さらに、分離に影響を与え得る陰影における変化が存在し得る。ゆえに、その画像についての単一の閾は、ピクセル色値などの側面に基づいてテキスト領域から背景領域を分離するには不十分であり得る。本明細書に記載の手法は、それぞれの場所のための適切な閾を判定するために、1つの画像にわたってさまざまな場所で分離値を分析し得る。閾は、よって、1つの画像にわたってさまざまに異なり得、および照明条件、コントラストなどにおける変更に適合し得る。かかる手法は、いくつかの状況においてテキスト検出の正確性を向上させ得、また他の状況においてテキスト認識工程の速度および効率性を少なくとも増加させ得る。

【0006】

種々の他のアプリケーション、処理、および使用が種々の実施形態に関して以下に提示される。

【0007】

図1は、イベントのチラシ106の上で携帯型コンピューティングデバイス102を保持しているユーザを示す状況の例100を説明したものである。イベントのチラシ106は、イベントの場所の物理的な住所を含み、かつこの例において、ユーザはその場所への道順を探している。道順を取得するために、ユーザはデバイス102のカメラをアドレスを含むチラシの一部に向けて、図1Aの画面104を通して見るように、そのアドレスの画像をキャプチャまたは取得し得る。一実施形態において、画像を取得し、および/またはテキストの存在を示すプロパティを有する1つ以上の部分を特定すると、デバイス102上のアプリケーションが光学的文字認識(OCR)アルゴリズムを自動的に実行して、チラシの画像化されたテキストを認識する。OCRアルゴリズムは、種々の実施形態において、種々の技術を活用する機械視覚アルゴリズムおよび他の画像前処理アルゴリズムを含み得る。OCRアルゴリズムは、テキストまたは文字の文字列を含み得る1つの画像の1つ以上の領域を特定および分析するために実行する。任意の特定された文字列は、Eメールアドレス、URL/ウェブアドレス、電話番号、および同等物など

10

20

30

40

50

の関与するデータオブジェクトまたは種類の存在を示し得るパターンをさらに特定するために分析される。再び図 1 を参照して、物理的な住所の存在を示す文字列が特定されている。この例において、ユーザは、図 1 B に示されるように、特定のアクションを実施する、または物理的な住所に関連付けられる既定のルーチンまたはワークフローを実行するユーザの意図を確定するように促される。この例において、既定のルーチンは地図アプリケーションを開かせ、住所への道順を表示させる。この例において、ユーザは第 1 に、「はい」ボタンを押すことによって意図を確定することができ、これにより図 1 C に示されるように、イベントへの道順を記した地図が携帯型コンピューティングデバイス 102 の画面 104 上に表示される。

【0008】

図 2 は、図 2 A に示される名刺 206 の上で携帯型コンピューティングデバイスを保持しているユーザ 202 を示す状況の例 200 を説明したものである。上述のとおり、デバイス 202、またはデバイスと通信しているサービスは、キャプチャした名刺の画像の中のテキストの場所を特定し、テキストの種類またはパターン（例えば、Eメールアドレス、電話番号、URL など）を特定し、テキストの種類に関連付けられる機能またはアプリケーション（例えば、電話番号に電話する、インターネットブラウザを開くなど）を判定し、アプリケーションまたは機能に場所を特定されたテキストの少なくとも関連のある一部分を送信してそれとともに操作を実行する。この例において、名刺 206 は、関与する複数のデータオブジェクトを含み、これらは電話番号、Eメール、および物理的な住所である。一例において、ユーザは、複数のデータ種類が特定された場合に、1つのデータ種類を別のものよりも優先させる、種々の機能またはアプリケーションに優先度を割り当てることができる。別の例において、デバイス 202 は、さまざまな利用可能な機能またはアプリケーションの中からの選択をユーザに促し得る。

【0009】

再び図 2 を参照して、2つ以上のデータオブジェクトを特定するデバイスは、図 2 B に示されるように、オブジェクトを連絡先情報として保存するためのアドレス帳アプリケーションを立ち上げるといった選択肢をユーザに促し得る。代替的に、ユーザは、「この番号に電話するには 1 を、道順を探すには 2 を、Eメールを送信するには 3 を、連絡先を保存するには 4 を押してください」などのメッセージを持つ画面 204 を介して促され得る。一例において、データ種類のテキストは、自動的に各々のデータフィールドの種類へと追加され得る。さらに、種々の実施形態において、デバイス 202 は、名刺などのオブジェクトを認識し、これにデフォルトのワークフローに関連付けるためのダウンロード可能なプラグインをプログラムされ、あらかじめ組み込まれ、または含み得、これはいくつかの例において、データを連絡先として保存するためであり得る。図 2 の例において、ユーザは、図 2 B に示されるように、名刺 206 に関連付けられる連絡先を保存するように促される。この場合、しかしながら、ユーザは名刺上の電話番号に電話することを所望し、「いいえ」を選択する。一例において、デバイス 202 は、個々の選択肢をユーザに再び促し得るか、またはユーザは、既定のルーチンを自動的に実行するか、または所定の操作を実行するというアプリケーションの優先度を設定した場合がある。ここで図 2 C を参照して、ユーザは、この電話番号に電話をかけるという選択肢を選択したか、またはユーザは、電話番号をダイヤルすることを優先操作として割り当て、この番号に電話をかけているデバイス 202 が示される。

【0010】

図 3 は、ユーザが携帯型コンピューティングデバイス 302 を窓 308 の方向に向けている状況の例 300 を説明したものである。窓 308 は、この例において、図 3 A に示される、その上に印刷されたウェブアドレスを有する。種々の実施形態に従い、道を歩いているユーザは、アドレスを手動でウェブブラウザの中へとタイピングする代わりに、ウェブアドレスに移動するための近道として、デバイス 302 に関連付けられるカメラをウェブアドレスに向け得る。この例において、ユーザは、ウェブアドレスを含む窓の画像を手動でキャプチャするために、画面 304 上のアイコン 306 を選択するか、または別のかが

10

20

30

40

50

る入力を提供しなければならない。前の例において、カメラは、画像をキャプチャしてルーチンを開始する、または特定のデータオブジェクトに関連付けられる動作を実行するというユーザの意図を予測するために、例えば、ジャイロスコプ、加速度計、光センサ、またはそれらの組み合わせを含んでいた場合がある。この例において、ユーザはデバイス302をかざし、カメラで窓308を狙い、アイコン306を押してウェブアドレスの画像をキャプチャする。次に、上述のとおり、デバイスまたはサービスは、図3Bに示されるように、OCRアルゴリズムを通じて実行し、画像の中のテキストの場所を特定し、テキストの種類またはパターン、この例においてはウェブアドレスを特定し、ブラウザアプリケーションを立ち上げ、ユーザをレストランのウェブアドレスへと誘導する。

【0011】

図4は、種々の実施形態に従うコンピューティングデバイス上のアプリケーションまたは機能に文字情報を提供するための工程の例400を説明したものである。本明細書において検討されるあらゆる工程について、別様に記載のない限り、種々の実施形態の範囲内で、追加的な、より少ない、または代替的なステップが、類似のまたは代替的な順序で、または並行して実施され得るということが理解されるべきである。この例において、画像または画像情報（例えば、ビデオストリーム）が取得される402。種々の実施形態において、画像または画像情報は、携帯型コンピューティングデバイスのカメラアプリケーションから取得される。一例において、カメラに面したオブジェクトの上にデバイスを空中停止すると、カメラに少なくとも1つの画像を自動的にキャプチャさせるか、または一連の画像を記録し得る。別の例において、カメラは、画像をキャプチャするために、例えば、ユーザがタッチ画面上のアイコンを選択するなどによるユーザからの入力を必要とする。取得された画像情報は、処理されてテキストまたは文字の文字列のプロパティを有する少なくとも1つの領域の場所を特定する404。一例において、処理は画像情報を受信すると自動的に開始される。カメラはまた、連続画像キャプチャモードであることもできる。これは、しかしながら、エネルギー集約型であるため、ユーザは、デバイスを連続的に画像キャプチャおよび処理モードで動作させる、またはこれを無効にする、さまざまな選択肢または環境を選択し得る。例えば、設定は、バッテリーが特定のレベルまで消耗したときに、連続的なキャプチャおよび処理モードが自動的にオフにされるように選択され得る。この例において、テキストの文字列は、テキストの文字列406の中のテキストを認識するための光学式文字認識アルゴリズムを用いて分析される。OCRアルゴリズムは、機械視覚アルゴリズムおよび他の画像処理技術またはアルゴリズムを含むことができる。認識されたテキスト408に対応するテキストパターン（例えば、Eメール、電話番号、URLなど）が特定される。この例において、判定されたテキストパターン410に関連付けられるアプリケーションおよび認識されたテキストが自動的にアプリケーションに提供される。

【0012】

一実施形態において、デバイスは、テキストの一部分内の関与するデータオブジェクトを認識しない場合がある。この場合、デバイスは、ユーザにウェブ検索を実施する、プラグインを検索して認識されていないパターンを認識するなどのさまざまな他の選択肢を提供するようにプログラムされ得る。ユーザは、テキストの一部分を選択して、このテキストをEメール、ノートパッド、または種々の他のアプリケーションにコピー/ペーストするように促される場合がある。

【0013】

種々の実施形態は、照明、陰影、コントラスト、または他のかかる側面における変化に対してロバストな画像の中のテキストを検出および認識する手法を提供する。例えば、図5は、コンピューティングデバイスのカメラを用いてキャプチャされる（または別様に取得される）ことができる画像の一部分の例500を説明したものである。画像部分の例500は、図3に関して上述した、「Good Food Cafe」のURLに対応する。図示される画像部分500は、文字「o」502aおよび502bの2つの例を含む。図示されるように、キャプチャされた画像部分は、文字のうち的一方502bの一部分に

10

20

30

40

50

かかる影の領域 504 を含む。この例において、部分 500 は、他のかかる選択肢の中でもとりわけ、画像がカラー画像であるか、またはグレースケール画像に変換されたカラー画像であっても、グレースケール画像であると仮定される。領域が論理的に複数のセグメント 506 に分割されると、影の領域の外にあり、かつテキスト文字の一部を持たないセグメントは、黒のテキスト文字であれば 0 の色値を有する場合、実質的に白（例えば、1 の色値を有する）であると見なされ得る（他のかかる選択肢の中でもとりわけ）。影の領域 504 内のセグメントにおいて、しかしながら、背景色は白ではなくグレーに見える。ゆえに、一例において、黒の値はすべての領域において実質的に同一（例えば、0）のまま留まるが、背景の色値は、影になっていない領域（例えば、0 の色値）と影になっている領域 504（例えば、0.5 の色値）との間で実質的に変化し得る。単一の閾が用いられ得る従来の閾値処理手法を用いる場合、影になっていない領域については 0.4 の閾が十分であり得るが、すべての色値がカットオフ閾を超える場合、すべての影になっている領域の中のすべてのセグメントが背景を含むように判定されることを招き得るので、情報の一部分が失われる。

10

20

30

40

50

【0014】

それにより、種々の実施形態に従う手法は、セグメントまたは領域のために適切な局所の閾を判定することを試みるために、1つの画像にわたって複数のセグメントまたは領域のうちのいくつかまたはすべてを分析することができる。かかる適応できる手法は、照明、彩色、陰影、または他のかかる変化が大きい画像での、改善されたテキスト検出を提供することができる。検討されたように、他の手法においては、1つの画像にわたる位置の複数の領域が代替的に選択されるが、1つのかかる手法は、他のかかる選択肢の中でもとりわけ、1つの画像（または、例えば、テキストを含む可能性が高いと判定された画像部分）を複数のセグメントにセグメント分けすることができる。

【0015】

画像が複数のセグメントにセグメント分けされる一例において、それぞれのセグメントが、そのセグメントが文字またはテキストに対応し得るエッジ領域を含むかどうかを判定することを試みるために分析され得る。例えば、図 6 の拡大図 600 において説明されているセグメント 508 を考察する。セグメントは、背景部分 602 および文字「o」の一部に対応する文字部分 604 の 2つの主要部分を含む。この例では背景部分 602 と文字部分 604 との間に明らかなエッジ境界が存在するが、アンチエイリアス処理、ぼやけ、および他のかかる問題のために、移行箇所がそれほど明確でないことが多い。画像セグメントの 1組の勾配が分析され得、この勾配はセグメントのピクセル 606 の間の色値の量の変化を表す。この例において、領域間の移行箇所の近くに 1組の勾配の大きい値が存在し得る。それにより、このセグメントは、テキスト部分を含むか分析される候補として選択され得る。少なくともいくつかの実施形態において、勾配の閾は、セグメントをテキストの候補とするためにどのくらいの変更が必要であるかを判定するように設定され得る。検討されたように、画像は、この判定を支援するために、分析前にグレースケールまたは単色の画像に変換され得る。

【0016】

エッジ位置を含むと判定されたそれぞれのセグメントまたは領域について、さまざまなピクセルのための色値のサンプリング（または「ピクセル値」）が判定され得、このサンプリングは、エッジ境界のそれぞれの側のピクセルを含む。例えば、エッジの第 1 の側の一連のいくつかのピクセルおよびエッジの他方側の一連のいくつかのピクセル（同一のまたは異なる）が判定され得る。これは、例えば、エッジ領域を中心とした 20×20 のピクセル配列、エッジの所与の距離の範囲内の無作為の選択などを含むことができる。ピクセルの数、エッジ領域からの距離、および他のかかる側面は、実施形態の間でさまざまであることができる。ピクセル値は、セグメントに対する背景ピクセル値およびセグメントに対する文字ピクセル値を判定することを試みるために分析され得る。例えば、理想的な背景が 1 の値を持つ白である場合、エッジの背景側のピクセル値は 0.4 ~ 0.8 の範囲の値を戻し得る。同様に、テキスト側のピクセル値は、0.0 に近い値を戻し得る。それ

により、システムは、背景の値がこれらのピクセル値の最大値（例えば、0.9）（または平均値、加重値など）であること、およびテキストの値が約0.0であることを判定することができる。ゆえに、0.5の閾は、背景領域からテキストを分離するために十分でない場合がある。この例において、アルゴリズムは、値の間の差異の割合に閾を設定することができる。いくつかの実施形態において、アルゴリズムは、最大点の中間点に閾を設定し得、ここで約0.4の色値を設定する。システムが保守的であることを望み、背景としてより少ないデータを排除する実施形態において、閾はより高く、およびその逆に設定され得る。エッジ領域を含むセグメントのそれぞれのピクセルが次に分析され得、対応するピクセル値がそれぞれ、セグメント閾に合致するか、セグメント閾を超過するか、またはセグメント閾未満であるかどうかに応じて（カラースケールなどの要因に応じる）、ピクセルがテキストまたは背景ピクセルとして指定され得る。

10

【0017】

セグメント分けは、1つの画像のより小さい領域上で照明および他のかかる側面がより均一となるという仮定の上で機能する。セグメントの大きさまたはピクセル領域は、しかしながら、十分な範囲の前景および背景ピクセルを覆う程度に大きくなければならず、そうでなければ芳しくない閾値が選ばれ得る。他方、大きすぎる領域を選ぶと、大きい領域がかなりの変化を含み得る場合、変化を的確に説明する処理能力を制限する場合がある。

【0018】

図7A、7B、および7Cは、走査線、ピクセル行、または種々の実施形態に従って生成され得る、異なる画像セグメントを通じた他のかかる横断面について取得され得るピクセル値の例を説明している。種々の実施形態において、ピクセル値は、グレースケールピクセル輝度値、カラーピクセル輝度値、ピクセル飽和値、ピクセル飽和レベル、および同等物を指す。コンピュータ内に保存された画像を表すそれぞれのピクセルが、そのピクセルがどのくらい明るい、および/またはピクセルがどの色であるべきかを記述するピクセル値を有する。グレースケール画像について、ピクセル値は典型的に、ピクセルの明るさを表す単数である。例えば、一般的なピクセル形式はバイト画像であり、ピクセル値の数が0~255の範囲の可能な値を付与する8ビットの整数として保存される。この例において、ゼロは黒と解釈され、255は白と解釈され、その間の値は異なるグレーを構成する。図7Aは、図5の領域508を水平に横切る個々のピクセルによって測定される、輝度の変化を説明するグレースケール画像のための距離対輝度のグラフの例である。左から開始して右へ移動する（図5において線で示される）領域508は、例えば、高い輝度の白いピクセル領域で開始し、文字の一部に対応する低い輝度の暗いまたは黒いピクセル領域で終了する。簡単にするために、以下の例についての可能なピクセル値の範囲は0~1の値で表される。700Aのグラフを参照して、白いピクセル領域の輝度は実質的に完全に明るくしたものに近く、1.0の値で示される。領域508を右へさらに移動すると、最終的には0.0に近いレベルに対応する輝度における急激な低下がある。この例において、0.5の輝度以上のすべてが画像背景の一部と見なされ、かつ0.5以下のすべてが、テキストであり得る関与するオブジェクトの一部と見なされるように、0.5の閾値702Aが選ばれ得る。

20

30

【0019】

ここで、図5の領域510に対応する図7Bを参照して、この例において、領域510の最も左側は影の領域504で開始し、これは最も右の端のテキスト領域よりも相対的に高いが、影の領域のために、図7Aに関して記載される初期の輝度と比較すると輝度が初期にはそれほど高くない。グラフ700Bを参照して、より明るいピクセル領域の輝度は、領域の影の領域504との近さのため、輝度において低く、これは0.5の初期値を付与する。領域510を右へ移動すると、最終的には0.0に近い値に対応する輝度における急激な低下がある。この例において、0.3の輝度より大きいもののすべてが画像背景の一部と見なされ、かつ0.3未満のすべてが、テキストであり得る関与するオブジェクトの一部と見なされるように、0.3の閾値702Bが選ばれ得る。閾値はよって、ピクセル輝度のより低い相対変化の結果として、領域508のものよりも低くなる。

40

50

【 0 0 2 0 】

図 7 C は、図 5 の領域 5 1 2 を水平に横切る個々のピクセルによって測定される輝度の変化を図で説明したものである。この例において、グラフ 7 0 0 C は、領域 5 1 2 からの初期の輝度の値 1 . 0 を示し、領域 5 1 2 を右に移動すると、やはり 0 . 0 に近い輝度の値に対応して輝度の急激な低下がある。継続して右に沿って、文字部位の他方側の輝度に別の大きな上昇があり、結果として再び約 1 . 0 の輝度の値となる。図 7 C において説明されるように、線が影の領域 5 0 4 に到達すると、輝度は再度下落して領域 5 1 2 の右側へと移動する。この例において、照射された部位とより暗い文字部位との間に著しく十分なコントラストがあるため、0 . 4 の輝度の値を有する閾 7 0 2 C が選ばれる。図 5 の例に広範囲な閾値処理が適用され、かつ、例えば、全体の画像部分について 0 . 5 の閾が選

10

【 0 0 2 1 】

図 8 は、種々の実施形態に従って用いられ得るテキスト認識アルゴリズムまたはエンジンにテキストデータを提供するための処理の例 8 0 0 を説明したものである。本明細書において検討されるあらゆる工程について、別様に記載のない限り、種々の実施形態の範囲内で、追加的な、より少ない、または代替的なステップが、類似のまたは代替的な順序で、または並行して実施され得るということが理解されるべきである。この例において、携帯型コンピューティングデバイスのカメラによってキャプチャされた画像が取得される 8 0 2 。少なくともいくつかの実施形態において、画像はグレースケール画像に変換され得、他のかかる前処理は、テキスト認識および他のかかる目的のために当技術分野で既知であるように実施され得る。画像における複数の領域の画像輝度勾配が分析されて 1 組のエッジ位置 8 0 4 を特定する。エッジを輝度における変化がいくつかのピクセルにわたって発生していると解釈すると、例えば、この輝度変化の導関数を計算することおよび計算された値がエッジ選択閾に合致するか超過する領域を選択することにより、エッジ検出アルゴリズムがエッジを判定することができる。1 組のエッジ位置のそれぞれに隣接する、または 1 組のエッジ位置のそれぞれからの所定距離の範囲内のピクセル値が分析される 8 0 6 。複数の領域の少なくとも一部分のための、画像の文字部分および背景部分についての相対ピクセル値が判定される 8 0 8 。この相対ピクセル値に少なくとも部分的に基づいて

20

30

【 0 0 2 2 】

適用可能な局所の閾を用いて可能性のある文字領域が判定された後、結合コンポーネントアルゴリズムを用いて、領域が複数の結合されたコンポーネントに組み立てられ得る。結合されたコンポーネントは、例えば、ピクセルの輝度および関与するピクセルの間の既定の距離を基本として結合されたピクセルであることができる。ゆえに、その領域内に小さい点または染みを含むセグメントは、点が類似の輝度または色値を有する他の領域の中のピクセルと近接していないため、隣接する領域の中のエッジと連結されない可能性が高い。結合コンポーネントアルゴリズム、または他のかかる処理はまた、いったん判定されると、それぞれの結合されたコンポーネントが少なくとも 1 つの文字制約を満たすかどうかを判定するために、結合されたコンポーネントを分析することもできる。この制約は、形状および寸法などの情報を含むことができ、隣接するテキストのプロパティに少なくとも部分的に基づくことができる。隣接するテキスト文字がとりわけ、かかる文字の間の相対的に一定の間隔、類似のアスペクト比、および類似の垂直および水平配列などのある特定のプロパティを共有すると推測され得る。共有のプロパティは、例えば、隣接する結合されたコンポーネントの重心（またはそれぞれの文字の上部または下部）が同一の線にほぼ沿っているかどうか、隣接する文字の高さおよび幅が割合の特定の範囲に収まるかどうか、および文字の間隔が実質的に類似するかどうかを判定することによって確認され

40

50

得る。これらの決定を行うためのパラメータは、経験的であるか、そうであることができる。上記に記載したプロパティおよび技術に基づいて、結合コンポーネントアルゴリズムはまた、形状、結合の配向、および正面画像から取り込まれていない、または画像にわたり対角線である1つの画像の中のテキストの行の方向をほぼ判定することもできる。

【0023】

図9は、画像900にわたり対角線上にキャプチャされたテキストの複数の行の画像部分900を説明したものである。画像部分900はやはり、図3において説明される「Good Food Cafe」のURLに対応する。この例において、テキストの一番上の行904には「Good Food Cafe」と記載され、テキストの2番目の行908には「www.goodfood.com」と記載されている。この例において、アルゴリズムは、それぞれの結合されたコンポーネントの重心の位置を判定することができ、かつ相対位置およびこれらの重心の間隔を分析して、画像部分の中の可能性のあるテキストの配向を判定することができる。例えば、隣接する結合されたコンポーネントの重心（図において点線で表される）は、線が所定量の偏差内でこれらの点を通り抜けることができた場合、同一の線上にあると見なされ得、この適合線に沿った点は許容可能な偏差の量の範囲内と一致する。しかしながら、文字の上部または下部の点は、他の実施形態において、重心の点の代わりに、またはこれを用いて判定され、用いられ得る。ある特定の文字の組み合わせは、異なる点の位置を使用するとよりうまく処理し得るので、十分な結果が別様には得られない場合、異なる点が用いられ判定される。

10

【0024】

この例において、「Good Food Cafe」についてのテキストまたは文字の文字列は、重心の平均線906の周りに大まかに配列し、「www.goodfood.com」についての文字列は、重心の平均線908の周りに大まかに配列する。テキストライン構成アルゴリズムは、テキストの方向およびどのコンポーネントが同一の線に沿うかを判定するために実装され得る。中心の平均線906および908の近くの対角線上にあるより多くの重心点の存在に基づいて、アルゴリズムは、例えば、ほぼ水平である、「Good」の1番目の「o」とURLの2番目の「w」の2つの重心点を通る仮想線を無視することを知り得る。

20

【0025】

加えて、文字が2つ以上の結合されたコンポーネントに分かれることを可能にするために、単一のユニットとして扱われて上記の制約を満たすことができる複数のコンポーネントを判定するための追加的な確認が行われることができる。文字制約のうちの少なくとも1つを満たさないコンポーネントは除去される。文字は近くにあるとも語を形成する他の文字に近く、かつ類似していなければならないと推測されるため、長い直線および長方形ならびに交差線およびボックスなどの制約のうちの少なくとも1つを満たさないコンポーネントは除去される。交差線およびボックスの除去は、長い直線、長方形、および交差線ならびにボックスがテキストに対応しないという考えに基づく。加えて、結合アルゴリズムは、画像中に表れるテキストの形状をほぼ判定するために適用され得る。

30

【0026】

いったん結合されたコンポーネントが分離されると、ストローク幅変換（SWT）アルゴリズムまたは他のかかる処理は、光学式文字リーダの入力を削除するために、1つの画像からのテキストの可能性のある領域のセグメント分けを試みるために提供され得る。アルゴリズムは、ピクセル毎に値を含むものからもっとも可能性の高いストローク幅を含むものへと画像データを転換することができ、その結果は、縮尺、方向、フォント、および言語に関わらずテキストを検出することができる。ある場面の他の要素からテキストを分離する1つの機能は、そのほぼ一定のストローク幅である。これは、テキストを含む可能性のある領域を正確に特定するために利用され得る。続いて、文字領域内のフィルタされかつ処理されたテキストを認識するために、光学式文字認識アルゴリズムを用いて文字の領域が分析される。テキスト位置および配列データをSWT（または他のそのようなもの）アルゴリズムに提供することの利点は、データがこの処理に開始点を提供し、よって処

40

50

理の速度および正確性を向上させることである。

【0027】

種々の実施形態において、オブジェクトの複数の画像が分析されて、ノイズ、ぼやけ、およびかかる他の同種のものなどの画像の不完全さを補正することができる。この技術は、統計的に、ノイズ、ぼやけ、およびさまざまな照明の変化が、複数の画像フレームにわたり一定でなく、かつこれらの不完全さが容易に廃棄またはフィルタで取り除くことができ、よって画像処理の効率性を増大させるという仮定に基づいて用いられる。いくつかの実施形態において、1組の画像が分析され、次にその結果がもっとも可能性のある入力へと進む。他の実施形態において、画像は、許容できる結果が判定されるまで一度に1つ分析され得、その後画像は廃棄される。種々の実施形態の範囲内で種々の他の手法も用いられ得る。

10

【0028】

種々の実施形態において、図2に関して上述した通り、機能またはアプリケーションは、テキストパターンが電話番号を示す場合、番号をダイヤルすることを、またはテキストパターンが電話番号および物理的な住所の存在を示す場合、連絡先情報を保存するためのアドレス帳を開くことを伴い得る。アプリケーションは、図3で説明されるようにテキストパターンがウェブアドレスを示す場合、URLへの誘導であり得る。アプリケーションは、図1で説明されるようにテキストパターンが物理的な住所を示す場合、地図を表示してアドレスを示し得る。機能またはアプリケーションはまた、種々の言語翻訳をも実施し得る。アプリケーションはまた、計算機であり得、数学の方程式を解く能力を有し得、または友人とレストランで食事をする場合に、チップを計算するまたは複数の方途で勘定を折半するための勘定書の特性を特定するためのプラグインに関連付けられ得る。一実施形態において、ユーザは、一次的アクションと、特定のデータオブジェクトに関連付けられる少なくとも二次的アクションから選ぶことを促され得る。名刺の例を再び参照して、名刺の認識に関連付けられる所定の一次的アクションは、名刺上の情報を新規の連絡先としてアドレス帳アプリケーションの中に保存することであり得る。二次的アクションは、カード上の電話番号に電話をかけることであり得、および三次的アクションは、住所への道順を調べることであり得る。

20

【0029】

一実施形態において、多くの実施形態の携帯型コンピューティングデバイスと通信しているシステムまたはサービスは、ユーザの活動を長期間監視して、アプリケーションによって実施されるアクションを調整することができる。例えば、ユーザが常に名刺上の番号に電話することを選び、この連絡先を保存することを決して選ばない場合、デバイスはアルゴリズムを調整して自動的に電話するか、またはユーザにその番号にまず電話することを促す。この例はまた、他のデータの種類の拡張されることもできる。他の実施形態において、デバイスは、さまざまな状況においてユーザの意図を予測するために、さまざまな挙動予想アルゴリズムまたはインテントウェアをリロードされ得る。例えば、デバイスは、ユーザ対面のカメラ、ジャイロスコープ、加速度計、光センサ、またはそれらの組み合わせを含み得る。ユーザ対面のカメラは、顔面認識およびジェスチャ追跡アルゴリズムを用いてユーザが見詰めている方向を判定し得る。光センサは、ユーザのカバンまたはポケットのなかにある可能性を判定し得る。この例において、デバイスは、エネルギー節約モードに入り得、センサによって十分な量の光が検出されるまで画像をキャプチャまたは処理することを控え得る。ジャイロスコープおよび加速度計は、ユーザがデバイスをどのように保持し、動かしているかを検出するために用いられ得る。いくつかの例において、ユーザは、画像の取り込みまたはキャプチャを暗示する特定の様式で予測運動を行うか、またはデバイスを保持する。一例において、動きを記録するジャイロスコープおよび加速度計は、やはりデバイスエネルギー節約モードにし得る。この例において、よい画像をキャプチャすることが難しい場合があり、デバイスはこれらが正常な画像キャプチャ条件でないことを合理的に推察するようにプログラムされ得る。種々の他の方法またはユーザの意図を予測することもまた本開示の範囲内で利用され得る。

30

40

50

【0030】

種々の実施形態は、画像前処理アルゴリズムおよび/または技術を利用して特定および/または認識のためにテキストを標準化する。実世界のテキストは、異なる大きさの、歪曲した、ぼやけた場合のある多くの色、形状、フォント種類および同等物である。機械視覚画像処理技術が用いられることができ、これは、閾値処理（グレースケール画像を白黒に変換する、またはグレースケール値に基づいて分離を用いる）、セグメント化、斑点抽出、パターン認識、バーコードおよびデータマトリクスコード読込、計測（オブジェクトの直径を測定する）、位置付け、エッジ検出、色分析、フィルタリング（例えば、形態的フィルタリング）、およびテンプレートマッチング（特定のパターンを発見、マッチング、および/または勘定する）を含むことができる。OCRによってより容易に認識可能である、テキストを処理するためのさまざまな他の技術もまた本開示の範囲内で利用され得る。

10

【0031】

図10は、種々の実施形態に従って用いられ得る、携帯型コンピューティングデバイス1000の正面図および後面図の例を説明したものである。1つの種類の携帯型コンピューティングデバイス（例えば、スマートフォン、電子書籍リーダー、またはタブレットコンピュータ）が示されるが、判定、処理、および入力を提供する能力があるさまざまな他の種類の電子デバイスが本明細に記載の種々の実施形態に従って用いられ得るということが理解されるべきである。デバイスは、例えば、ノート型コンピュータ、携帯情報端末、携帯電話、ビデオゲーム機またはコントローラ、および携帯型メディアプレーヤをとりわけ含むことができる。

20

【0032】

この例において、携帯型コンピューティングデバイス1000は、画像コンテンツをデバイスの1人以上のユーザまたは閲覧者に表示するように動作可能な表示画面1002（例えば、液晶画面（LCD）要素）を有する。少なくともいくつかの実施形態において、表示画面は、例えば、容量性または抵抗性のタッチ技術を用いてタッチまたはスワイプをベースとした入力を提供する。かかる表示要素は、例えば、ユーザが、右または左のマウスのボタン、タッチポイントなどのボタンの画像に対応する画面のある部位を押すことによって入力を提供することを可能にするために用いられ得る。デバイスはまた、デバイスの側面または背面上などのデバイスの他の部位上に接触感知式および/または感圧式材料1010を有することもできる。少なくともいくつかの実施形態において、かかる材料に接触するまたは握ることによってユーザが入力を適用することができるが、他の実施形態において、材料は、材料に関するパターン模様入りの表面の動きを通じてデバイスの運動を検出するために用いられ得る。

30

【0033】

携帯型コンピューティングデバイスの例は、従来の画像および/またはビデオキャプチャなどの目的のために1つ以上の画像キャプチャ要素を含むことができる。本明細書に記載したように、画像キャプチャ要素はまた、動きを判定し、ジェスチャ入力を受信する目的のために用いられることもできる。この例における携帯型コンピューティングデバイスは、デバイスの「前面」に1つの画像キャプチャ要素1004を含み、デバイスの「背面」に1つの画像キャプチャ要素1012を含むが、画像キャプチャ要素はまた、または代替的に、デバイスの側面または角に置かれ得、および類似のまたは異なる種類の任意の適切な数のキャプチャ要素が存在することができるということが理解されるべきである。それぞれの画像キャプチャ要素が、例えば、カメラ、電荷結合素子（CCD）、動き検出センサ、または赤外線センサであり得、または別の画像キャプチャ技術を利用することができる。

40

【0034】

携帯型コンピューティングデバイスはまた、少なくとも1つのマイク1006またはある特定の実施形態において位置の変化を判定する、またはユーザ入力を受信するために用いられ得るものなどの音声データキャプチャする能力がある他の音声キャプチャ要素を含

50

むこともできる。いくつかのデバイスにおいて、1つのマイクのみが存在し得るが、他のデバイスにおいては、デバイスのそれぞれの側面および/または角に、または他の適切な場所に少なくとも1つのマイクが存在し得る。

【0035】

この例におけるデバイス1000はまた、デバイスの位置、方向、動き、または配向などの情報を提供するように動作可能な、少なくとも1つの動きまたは位置判定要素1008を含む。これらの要素は、例えば、加速度計、内部センサ、電子ジャイロスコープ、電子コンパス、およびGPS要素を含むことができる。別のデバイスのための少なくとも1つの制御信号をトリガすることができるさまざまな種類の動きまたは配向の変更が、デバイスに入力を提供するために用いられ得る。このデバイスの例はまた、1つ以上の携帯型コンピューティングデバイスと通信するように動作可能な、少なくとも1つの有線または無線コンポーネントを含み得るなどの少なくとも1つの通信機構1014を含む。デバイスはまた、従来のプラグイン手法を通じて、または電力マッドまたは他のかかるデバイスとの近接を通じた容量性充電などの他の手法を通じて再充電されるように動作可能なバッテリーを含み得るなどの電力システム1016を含む。種々の実施形態の範囲内でさまざまな他の要素および/または組み合わせもまた可能である。

10

【0036】

図10、図11に関して記載したような機能性を提供するために、図10に関して記載したデバイス1000などの携帯型コンピューティングデバイス1100の1組の基本コンポーネントの例を説明する。この例において、デバイスは、少なくとも1つの記憶デバイスまたは要素1104内に格納され得る命令を実行するための少なくとも1つのプロセッサ1102を含む。当業者には明白であるように、デバイスは、プロセッサ1102によって実行されるためのプログラム命令のための第1のデータ記憶装置などの多くの種類のメモリ、データ記憶装置またはコンピュータ可読の記憶媒体を含むことができ、画像またはデータのために同一または別個の記憶装置が用いられることができ、取り外し可能な記憶メモリは他のデバイスなどと情報を共有するために利用可能である。

20

【0037】

携帯型メディアプレーヤなどのデバイスは音声スピーカなどの他の手段を介して情報を伝達し得るが、デバイスは典型的に、タッチ画面、電子インク(eインク)、有機発光ダイオード(OLED)または液晶画面(LCD)などのいくつかの種類の表示要素1106を含み得る。検討されたように、多くの実施形態において、デバイスは、ユーザ、人々、またはオブジェクトをデバイスの近くに撮像することができる1つ以上のカメラなどの少なくとも1つの画像キャプチャ要素1108を含み得る。少なくともいくつかの実施形態において、デバイスは、画像情報を用いてユーザによるジェスチャまたはモーションを判定することができ、これはユーザが実際に携帯型デバイスに接触する、および/または動かすことなく、携帯型デバイスを通じて入力を提供することを可能にする。画像キャプチャ要素はまた、本明細書に記載されるように、デバイスの周囲環境を判定するために用いられ得る。画像キャプチャ要素は、ユーザがデバイスを操作しているときに、ユーザの画像をキャプチャするために十分な解像度、集束範囲および可視領域を有するCCD画像キャプチャ要素などの任意の適切な技術を含むことができる。

30

40

【0038】

デバイスは、多くの実施形態において、1つ以上の音声スピーカおよび/またはマイクなどの少なくとも1つの音声要素1110を含み得る。マイクは、声認識、デジタル録音など、声で可能となる機能を促進するために用いられ得る。音声スピーカは音声出力を実施し得る。いくつかの実施形態において、音声スピーカ(複数可)は、デバイスから分離して存在し得る。

【0039】

デバイスは、ユーザから従来の入力を受信することができる少なくとも1つの追加的な入力デバイス1112を含むことができる。この従来の入力は、例えば、これらによってユーザがデバイスにコマンドを入力することができる押しボタン、タッチパッド、タッチ

50

画面、ホイール、ジョイスティック、キーボード、マウス、トラックボール、キーパッドまたは任意の他のかかるデバイスまたは要素を含むことができる。これらのI/Oデバイスは、いくつかの実施形態において、無線赤外線またはBluetooth（登録商標）または他のリンクによってさえも接続され得る。いくつかの実施形態において、しかしながら、かかるデバイスはいかなるボタンも含まない場合があり、ユーザがデバイスと接触することなくデバイスを制御することができるように、視覚および音声コマンドの組み合わせによってのみ制御され得る場合がある。

【0040】

デバイスの例はまた、特定の無線チャネルの通信範囲内で1つ以上の携帯型コンピューティングデバイスと通信するように動作可能な1つ以上の無線コンポーネント1114を含む。無線チャネルは、デバイスが無線で通信することを可能にするために用いられる、Bluetooth、セルラー、またはWi-Fiチャネルなどの任意の適切なチャネルであることができる。デバイスが、当技術分野で既知であるような1つ以上の従来の有線通信接続を有することができるということが理解されるべきである。デバイスの例は、携帯型コンピューティングデバイスに電力を提供するための、当技術分野で既知のさまざまな電力コンポーネント1116を含み、これは電力パッドまたは本明細書に記載に記載される類似のデバイスとともに用いられるための容量性充電要素を含むことができる。デバイスの例はまた、デバイスのケーシングの周りの接触感知式材料などの少なくとも1つの接触感知式および/または感圧式要素1118、握ることに基づく入力をデバイスに提供する能力がある少なくとも1つの領域を含むことができる。いくつかの実施形態において、この材料は、例えば、デバイスのまたはユーザの指などの動きを判定するために用いられ得るが、一方他の実施形態において、材料は、特定の入力またはコマンドを提供するために用いられ得る。

【0041】

いくつかの実施形態において、デバイスは、ユーザまたはアプリケーションからのコマンドを受け取ったとき、または音声入力またはビデオ入力を判定しようと再試行しているときなどに、検出および/またはコマンドモードを起動および/または解除する能力を含むことができる。いくつかの実施形態において、デバイスは、赤外線検出器または運動センサを含むことができ、例えば、これらは1つ以上の検出モードを起動するために用いられ得る。例えば、その部屋にユーザが存在しないとき、あるデバイスは、デバイスを検出したり、またはデバイスと通信したりすることを試みない可能性がある。赤外線検出器（すなわち、状態の変化を検出する1ピクセル分解能を持つ検出器）が、例えば、ユーザがその部屋に入ってきたことを検出した場合、デバイスは、ユーザによって必要とされたときにデバイスの準備ができていようにするが、ユーザが近くにいないときは電力および資源を節約することができるように、検出または制御モードを起動することができる。

【0042】

種々の実施形態に従うコンピューティングデバイスは、デバイスが周辺光に曝露しているか、または相対的または完全な暗闇にあるのかを判定することができる光検出要素を含み得る。かかる要素は、いくつかの方途において有益である場合がある。ある特定の従来のデバイスにおいて、光検出要素は、電話の表示要素を一時的に切る（デバイスをユーザの耳に当てている間、ユーザが表示要素を見ることができないため）アクションをトリガする、いつユーザが携帯電話をユーザの顔に向けて保持しているか（光検出要素を実質的に周辺光から遮断させる）を判定するために用いられる。光検出要素は、デバイスの機能を調整するための他の要素からの情報を合わせて用いられ得る。例えば、デバイスがユーザの視覚位置を検出することができず、ユーザがデバイスを保持していないが、デバイスが周辺光に曝露している場合、デバイスは、ユーザによって下に置かれたということとを判定し、表示要素をオフにし、ある特定の機能性を無効にし得る。デバイスがユーザの視覚位置を検出することができず、ユーザがデバイスを保持しておらず、デバイスがさらに周辺光に曝露している場合、デバイスは、デバイスがユーザの手が届かない可能性があるバッグまたは他の仕切りに置かれたということとを判定し、よって別様には利用可能であった

10

20

30

40

50

追加的な機能をオフまたは無効にし得る。いくつかの実施形態において、ユーザはデバイスを見ているか、デバイスを保持しているか、またはデバイスのあるいくつかの機能性を有効にするためにデバイスを光の中へ取り出したかのいずれかであるはずである。他の実施形態において、デバイスは、反射（明るい状況）および放射（暗い状況）などの異なるモードにおいて動作することができる表示要素を含み得る。検出された光に基づいて、デバイスはモードを変更し得る。マイクを用いて、デバイスは、電気の節約に実質的に関係のない理由により、他の機能を無効にすることができる。例えば、デバイスは、声認識を用いて、子供などのデバイスの近くにいる人を判定し、これに基づいてインターネットアクセスまたはペアレンタルコントロールなどの機能を無効または有効にすることができる。さらに、デバイスは、録音されたノイズを分析して、デバイスが車内にあるか飛行機内にあるかなどの環境を判定することを試みることができ、この判定は他の入力に基づいて、どの機能を有効/無効にするか、またはどのアクションが取られるかを決定することを支援することができる。声認識が用いられる場合、デバイスに直接的に話されるか、または会話を通じて間接的に拾った言葉が入力として用いられ得る。例えば、デバイスが、デバイスは車内にあり、ユーザの方を向いていると判定し、「おなかがすいた」または「食べる」などの言葉を検出した場合、次にデバイスは表示要素をオンにして近くのレストランなどの情報を表示し得る。ユーザは、プライバシーおよび他のかかる目的のために、声の記録および会話のモニタリングをオフにする選択肢を有する。

10

20

30

40

50

【0043】

上記の例のうちのいくつかにおいて、デバイスによって取られたアクションは、電力消費を低減する目的のために、ある特定の機能性を解除にすることに関する。しかしながら、デバイスを用いることについての類似の他の可能性のある課題を調整することができるアクションが他の機能に対応することができることが理解されるべきである。例えば、ウェブページのコンテンツを要求する、ハードドライブ上のコンテンツを検索する、およびさまざまなアプリケーションを開くなどのある特定の機能は、完了までにいくらかの時間を要する。資源の限られたデバイス、または多用されるデバイスについては、いくつかのかかる操作が同時に発生することが、デバイスの動作を遅くしたり、または動かなくなったりすることを招く場合があり、これは非効率性に繋がり、ユーザの体験を悪化させ、より多く電力を使用する場合がある。

【0044】

少なくともこれらおよび他のかかる問題のうちのいくつかに対処するため、種々の実施形態に従う手法はまた、ユーザの見つめる方向などの情報を利用して、処理容量、メモリの空きおよび他のかかる資源の必要性を広げるために用いられる可能性がある資源を起動することもできる。

【0045】

いくつかの実施形態において、デバイスは、十分な処理容量を有することができ、撮像要素および関連付けられる解析アルゴリズム（複数可）は、キャプチャされた画像のみに基づいてデバイスの動き、ユーザの頭の動き、ユーザの目の動き、および他のかかる動きを区別するために十分な程に感受性が高い。極めて単純な撮像要素および分析手法を利用することが処理のために望ましい場合がある他の実施形態において、デバイスの現在の配向を判定することができる、少なくとも1つの配向判定要素を含むことが望ましい場合がある。一例において、この少なくとも1つの配向判定要素は、デバイスの3D位置およびデバイスの動きの大きさおよび方向、ならびに振動、衝撃などの因子を検出することができる、少なくとも1つの単軸または多軸加速度計である。加速度計などの要素を用いてデバイスの配向または動きを判定するための方法もまた当技術分野で既知であり、本明細書では詳細に記載しない。配向および/または動きを判定するための他の要素もまた、種々の実施形態の範囲内で配向判定要素として用いるために用いられ得る。加速度計または同様の要素からの入力がかメラからの入力とともに用いられたとき、相対的な動きはより正確に解釈され得、さらに精密な入力および/またはより複雑性の低い画像分析アルゴリズムを可能にする。

【 0 0 4 6 】

コンピューティングデバイスの撮像要素を用いて、例えばデバイスおよび/またはユーザの動きを検出するとき、コンピューティングデバイスは、画像中の背景を用いて動きを判定することができる。例えば、ユーザがデバイスを固定された配向（例えば、距離、角度など）でユーザにかざし、ユーザが配向を周囲の環境に変更した場合、ユーザの画像のみを分析することはデバイスの配向における変更を検出する結果とならない。そうではなく、いくつかの実施形態において、コンピューティングデバイスは、ユーザの後方にある背景画像における変化を認識することによってデバイスの動きをやはり検出することができる。そのため、例えば、オブジェクト（例えば、窓、絵、木、茂み、建物、車など）が画像の左または右へ移動すると、デバイスは、ユーザに関するデバイスの配向が変更されていないにも関わらず、デバイスが配向を変更したと判定することができる。他の実施形態において、デバイスは、デバイスに関して移動したことを判定し、それに応じて調整し得る。例えば、ユーザがデバイスに関して頭を左または右に傾けた場合、表示要素上にレンダリングされるコンテンツは、同様に傾いてコンテンツをユーザの配向に保ち得る。

10

【 0 0 4 7 】

検討されたように、記載の実施形態に従う異なる手法がさまざまな環境において実装され得る。例えば、図 1 2 は、種々の実施形態に従う側面を実装するための環境の例 1 2 0 0 を説明したものである。理解されるように、説明を目的としてウェブベースの環境が用いられるが、種々の実施形態を実装するために異なる環境が適切に用いられ得る。システムは、電子クライアントデバイス 1 2 0 2 を含み、これは、要求、メッセージまたは情報を適切なネットワーク 1 2 0 4 で送受信し、情報をデバイスのユーザへ戻し伝えるように動作可能な任意の適切なデバイスを含むことができる。かかるクライアントデバイスの例は、パソコン、携帯電話、手持ち式のメッセージングデバイス、ラップトップ型パソコン、セットトップボックス、携帯情報端末、電子書籍リーダーおよび同等物を含む。ネットワークは、イントラネット、インターネット、セルラーネットワーク、ローカルエリアネットワーク、または任意の他のかかるネットワーク、またはそれらの組み合わせを含む任意の適切なネットワークを含むことができる。ネットワークは、「プッシュ」型ネットワーク、「プル」型ネットワーク、またはそれらの組み合わせであり得る。「プッシュ」型ネットワークにおいて、サーバのうちの 1 つ以上がクライアントデバイスにデータをプッシュする。「プル」型ネットワークにおいて、サーバのうちの 1 つ以上が、クライアントデバイスからデータについての要求があると、クライアントデバイスにデータを送信する。かかるシステムのために用いられるコンポーネントは、少なくとも部分的に選択されたネットワークおよび/または環境の種類に依存する。かかるネットワークを介して通信するためのプロトコルおよびコンポーネントはよく知られており、本明細書において詳細に記載されない。ネットワークでの通信は、有線または無線接続およびそれらの組み合わせを介して有効とされ得る。この例において、環境が要求を受信し、それに応じてコンテンツを提供するためのウェブサーバ 1 2 0 6 を含むため、ネットワークはインターネットを含むが、他のネットワークについては、当業者に明確となるように、類似の目的を果たす代替的なデバイスが用いられ得る。

20

30

【 0 0 4 8 】

例示的な環境は、少なくとも 1 つのアプリケーションサーバ 1 2 0 8 およびデータストア 1 2 1 0 を含む。相互にやり取りして適切なデータストアからデータを取得するなどのタスクを実施することができる、繋がれるか別様に構成されるいくつかのアプリケーションサーバ、レイヤ、または他の要素、処理またはコンポーネントが存在することができるということが理解されるべきである。本明細書で用いられる場合、「データストア」という用語は、任意のデバイスまたはデータの格納、アクセス、および読み出しの能力を持つデバイスの組み合わせを指し、これらは任意の標準、分散、またはクラスタ環境において、任意の組み合わせおよび数のデータサーバ、データベース、データ記憶デバイスおよびデータ記憶媒体を含み得る。アプリケーションサーバ 1 2 0 8 は、クライアントデバイスのための 1 つ以上のアプリケーションの側面を実行する必要に応じてデータストア 1 2 1

40

50

0と統合するための、およびアプリケーションのためのデータアクセスおよびビジネス論理の大部分を扱うための、任意の適切なハードウェアおよびソフトウェアを含むことができる。アプリケーションサーバは、データストアと協働してアクセス制御サービスを提供し、テキスト、図形、ユーザに転送される音声および/またはビデオなどのコンテンツを生成することができ、これはこの例においてはウェブサーバ1206によってHTML、XML、または他の適切な構造化言語の形式でユーザに提供される。すべての要求および応答の扱い、ならびにクライアントデバイス1202とアプリケーションサーバ1208との間のコンテンツの配信は、ウェブサーバ1206によって扱われることができる。本明細書に記載に記載されるように、本明細書に記載の構造化コードが任意の適切なデバイスまたはホストマシン上で実行され得るため、ウェブおよびアプリケーションサーバは必要とされず、および単にコンポーネントの例であるということが理解されるべきである。

10

【0049】

データストア1210は、特定の側面に関連するデータを格納するためのいくつかの別個のデータテーブル、データベースまたは他のデータ記憶機構および媒体を含むことができる。例えば、説明されるデータストアは、コンテンツ(例えば、生産データ)1212およびユーザ情報1216を格納するための機構を含み、これらは生産側にコンテンツを提供するために用いられ得る。データストアはまた、ログまたはセッションデータ1214を格納するための機構を含むようにも示される。ページ画像情報およびアクセス権情報などの、データストア内に格納される必要があり得る多くの他の側面が存在することができるということが理解されるべきであり、これらは上記に列挙された機構のうちの任意のものに適切に、またはデータストア1210内の追加的な機構に格納され得る。データストア1210は、これに関連付けられる論理を通じて、アプリケーションサーバ1208から命令を受信し、かつそれに応じてデータを取得、更新、または別様に処理するように動作可能である。一例において、ユーザは、ある特定の種類の項目について検索要求を提出し得る。この場合、データストアは、ユーザ情報にアクセスしてユーザの身元を確認し、得、かつカタログ詳細情報にアクセスしてその種類の項目についての情報を取得することができる。情報は次に、ユーザがユーザデバイス1202上のブラウザを介して閲覧することができるウェブページ上に列挙される結果の中でユーザに戻される。関与する特定の項目についての情報は、専用のページまたはブラウザのウィンドウの中で閲覧され得る。

20

30

【0050】

それぞれのサーバが典型的に、そのサーバの一般管理および操作のために実行可能なプログラム命令を提供するオペレーティングシステムを含み、典型的にサーバのプロセッサによって実行されると、サーバがその意図する機能を実施することを可能にする命令を格納するコンピュータ可読の媒体を含む。オペレーティングシステムのための適切な実装およびサーバの一般的な機能性は既知であるか市販されており、特に本開示を考慮して当業者によって容易に実装される。

【0051】

一実施形態における環境は、1つ以上のコンピュータネットワークまたは直接接続を用いた通信リンクを介して相互接続された、いくつかのコンピュータシステムおよびコンポーネントを利用した分散型コンピューティング環境である。しかしながら、図12において説明されたものよりも少ないまたは多数のコンポーネントを有するシステムにおいて、かかるシステムが同等に良好に動作し得るということは当業者によって認識されるであろう。ゆえに、システム1200の図12における描写は、例示的な性質であり、本開示の範囲を制限しないと見なされるべきである。

40

【0052】

種々の実施形態は、多種多様な動作環境においてさらに実装され得、いくつかの場合において、いくつかのアプリケーションのうちの任意のものを操作するために用いられ得る1つ以上のユーザコンピュータまたはコンピューティングデバイスを含むことができる。ユーザまたはクライアントデバイスは、標準オペレーティングシステムを実行するデスク

50

トップ型またはラップトップ型パソコン、ならびにモバイルソフトウェアを実行し、いくつかのネットワーキングおよびメッセージングプロトコルに対応する能力を持つセルラー、無線および手持ち型デバイスなどのいくつかの汎用パソコンを含むことができる。かかるシステムはまた、開発およびデータベース管理などの目的のために、さまざまな市販のオペレーティングシステムおよび他の知られているアプリケーションを実行するいくつかのワークステーションを含むこともできる。これらのデバイスはまた、疑似ターミナル、シンクライアント、ゲームシステム、およびネットワークを介して通信する能力を持つ他のデバイスなどの他の電子デバイスを含むこともできる。

【0053】

ほとんどの実施形態が、TCP/IP、OSI、FTP、UPnP、NFS、CIFS、およびAppleTalkなどのさまざまな市販のプロトコルのうちの任意のものを利用して通信に対応するために、当業者に周知の少なくとも1つのネットワークを利用する。ネットワークは、例えば、ローカルエリアネットワーク、広域ネットワーク、仮想プライベートネットワーク、インターネット、イントラネット、エクストラネット、公衆交換電話ネットワーク、赤外線ネットワーク、無線ネットワーク、およびそれらの組み合わせであることができる。

【0054】

ウェブサーバを利用する実施形態において、ウェブサーバは、HTTPサーバ、FTPサーバ、CGIサーバ、データサーバ、Java（登録商標）サーバ、およびビジネスアプリケーションサーバを含むさまざまなサーバまたはミッドティアアプリケーションの任意のものを実行することができる。サーバ（複数可）はまた、Java（登録商標）、C、C#もしくはC++などの任意のプログラミング言語、またはPerl、Python、もしくはTCLなどの任意のスクリプト言語、またはそれらの組合せで記述された1つ以上のスクリプトまたはプログラムとして実装され得る1つ以上のウェブアプリケーションを実行することによって、ユーザデバイスからの要求に応じてプログラムまたはスクリプトを実行する能力をもち得る。サーバ（複数可）はまた、Oracle（登録商標）、Microsoft（登録商標）、Sybase（登録商標）、およびIBM（登録商標）から市販されているものを含むデータベースサーバを含み得るがこれらに限定されない。

【0055】

環境は、上述のとおり、さまざまなデータストアおよび他のメモリおよび記憶媒体を含むことができる。これらはコンピュータのうちの1つ以上にとってローカルの（および/または常駐の）、またはネットワークにわたる任意のまたはすべてのコンピュータからリモートの記憶媒体上などのさまざまな場所に存在することができる。特定の1組の実施形態において、情報は当業者に周知のストレージエリアネットワーク（SAN）内に存在し得る。同様に、コンピュータ、サーバまたは他のネットワークデバイスに帰属する機能を実施するために必要な任意のファイルは、ローカルまたはリモートに適切に格納され得る。システムがコンピュータ化されたデバイスを含む場合、それぞれのかかるデバイスが、バスを介して電氣的に接続され得る、例えば、少なくとも1つの中央処理装置（CPU）、少なくとも1つの入力デバイス（例えば、マウス、キーボード、コントローラ、タッチ感応式表示要素、またはキーパッド）、および少なくとも1つの出力デバイス（例えば、表示デバイス、プリンタ、またはスピーカ）を含むハードウェア要素を含むことができる。かかるシステムは、ディスクドライブ、光学式記憶デバイス、およびランダムアクセスメモリ（RAM）または読み取り専用メモリ（ROM）などのソリッドステート記憶デバイス、ならびに取り外し可能な媒体デバイス、メモリカード、フラッシュカードなどの1つ以上の記憶デバイスを含み得る。

【0056】

かかるデバイスはまた、上述の通り、コンピュータ可読の記憶媒体リーダ、通信デバイス（例えば、モデム、ネットワークカード（無線または有線）、赤外線通信デバイス）、およびワーキングメモリを含むこともできる。コンピュータ可読の記憶媒体リーダは、リ

10

20

30

40

50

モート、ローカル、固定、および/または取り外し可能な記憶デバイスを表すコンピュータ可読の記憶媒体、ならびに一時的におよび/または永続的にコンピュータ可読の情報を含む、格納する、送信する、および読み出すための記憶媒体と接続されるか、または受信するように構成され得る。システムおよびさまざまなデバイスはまた典型的に、オペレーティングシステムおよびクライアントアプリケーションまたはウェブブラウザなどのアプリケーションプログラムを含む少なくとも1つのワーキングメモリデバイス内に位置するいくつかのソフトウェアアプリケーション、モジュール、サービス、または他の要素を含み得る。代替の実施形態が上述のものからの多くの変化を有することが理解されるべきである。例えば、カスタマイズされたハードウェアもまた用いられ得、および/またはハードウェア、ソフトウェア（アプレットなどの携帯型ソフトウェアを含む）またはその両方に特定の要素が実装され得る。さらに、ネットワーク入力/出力デバイスなどの他のコンピュータデバイスへの接続が採用され得る。

10

【0057】

コードまたはコードの一部を含む記憶媒体およびコンピュータ可読媒体は、RAM、ROM、EEPROM、フラッシュメモリもしくは他のメモリ技術、CD-ROM、デジタル多用途ディスク（DVD）もしくは他の光学式記憶、磁気カセット、磁気テープ、磁気ディスク記憶もしくは他の磁気記憶デバイスまたは任意の他の媒体を含む、コンピュータ可読命令、データ構造、プログラムモジュールもしくは他のデータなどの情報の記憶および/または送信のための、任意の方法または技術で実装された、これらに限定されないが揮発性および不揮発性の、取り外し可能および取り外し可能でない媒体などの記憶媒体および通信媒体を含む、当技術分野で既知または使用されている任意の適切な媒体を含むことができ、これらは所望の情報を格納するために用いられることができ、これらはシステムデバイスによってアクセスされ得る。本開示および本明細書に提供される教示に基づき、当業者は種々の実施形態を実装するための他の方途および/または方法を認識するであろう。

20

【0058】

明細書および図面は、それにより、制限的意味でなく例示的意味と見なされる。しかしながら、特許請求の範囲に記載の本発明の広範な精神および範囲から逸脱することなくこれらに対しさまざまな修正および変更がなされ得ることは明白となる。

30

【0059】

付記

付記1. コンピュータ実装方法であって、

実行可能な命令で構成される1つ以上のコンピュータシステムの制御下において、携帯型コンピュータデバイスのカメラによってキャプチャされた画像であって、前記カメラの複数のピクセルについてのピクセル値を含む画像を取得することと、前記領域内の1つ以上の画像輝度勾配を分析することにより、前記画像の1つの領域内の1つ以上の可能性のあるエッジ位置を特定することと、前記1つ以上の可能性のあるエッジ位置の両側の1組のピクセルのピクセル値を分析して前記領域のための文字ピクセル値および背景ピクセル値を判定することと、前記領域のための前記文字ピクセル値および前記背景ピクセル値に少なくとも部分的に基づいて、前記領域のための文字ピクセル値の閾を判定することと、前記領域の中のピクセルについてのピクセル値が前記文字ピクセル値の閾より高いか低いに少なくとも部分的に基づいて前記領域内の1つ以上の文字部分を判定することと、を含む、方法。

40

【0060】

付記2. 結合コンポーネントアルゴリズムを用いて、前記文字ピクセル値の閾に少なくとも合致するピクセル値を有する前記複数のピクセルの隣接するピクセルを関連付けて、文字部分の少なくとも一部分を形成することを可能にすることと、前記1つ以上の文字部分の少なくとも一部分を文字認識エンジンに提供することと、をさらに含む、付記1に記載のコンピュータ実装方法。

50

【 0 0 6 1 】

付記 3 . 文字認識エンジンを用いて前記 1 つ以上の文字部分を分析して、前記 1 つ以上の文字部分に関連付けられる少なくとも一文字を認識することを分析することをさらに含む、付記 2 に記載のコンピュータ実装方法。

【 0 0 6 2 】

付記 4 . 前記認識された少なくとも 1 つの文字に対応する文字パターンを特定することと、前記文字パターンに関連付けられる前記携帯型コンピューティングデバイス上のアプリケーションを判定することと、前記認識された少なくとも 1 つの文字を前記判定されたアプリケーションに送信することと、をさらに含む、付記 3 に記載のコンピュータ実装方法。

10

【 0 0 6 3 】

付記 5 . コンピュータ実装方法であって、

実行可能な命令で構成される 1 つ以上のコンピュータシステムの制御下において、1 つの画像であって、複数のピクセルについてのピクセル値を含む、画像中の 1 つの領域の 1 つ以上の輝度勾配を分析して 1 つ以上の可能性のあるエッジ位置を特定することと、前記 1 つ以上の可能性のあるエッジ位置の所定距離の範囲内にある 1 組のピクセルについてのピクセル値を分析することと、

前記 1 つ以上の可能性のあるエッジ位置のそれぞれの前記所定距離の範囲内にある前記ピクセル値に少なくとも部分的に基づいて、前記領域のための文字ピクセル値の閾を判定することと、

20

前記文字ピクセル値の閾と少なくとも合致するピクセル値を有する前記 1 組のピクセルの一部分に少なくとも部分的に基づいて、1 つ以上の文字部分を判定することと、を含む、方法。

【 0 0 6 4 】

付記 6 . 文字認識アルゴリズムを用いて前記 1 つ以上の文字部分を分析して、前記 1 つ以上の文字部分に関連付けられる少なくとも一文字を認識することをさらに含む、付記 5 に記載のコンピュータ実装方法。

【 0 0 6 5 】

付記 7 . エッジ位置が、前記所定距離の範囲内の前記文字ピクセル値の閾に少なくとも合致する値を有するピクセル輝度勾配またはピクセル飽和勾配に対応する、付記 5 に記載のコンピュータ実装方法。

30

【 0 0 6 6 】

付記 8 . 前記ピクセル値を分析することは、前記 1 つ以上の可能性のあるエッジ位置の第 1 の側にある第 1 の組のピクセルを分析することと、前記 1 つ以上の可能性のあるエッジ位置の第 2 の側にある第 2 の組のピクセルを分析することとを含む、付記 5 に記載のコンピュータ実装方法。

【 0 0 6 7 】

付記 9 . 前記ピクセル値は、グレースケールピクセル輝度値、カラーピクセル輝度値、ピクセル飽和値、またはピクセル飽和レベルのうちの少なくとも 1 つに対応する、付記 5 に記載のコンピュータ実装方法。

40

【 0 0 6 8 】

付記 10 . テキストライン構成アルゴリズムを用いて、前記 1 つ以上の文字部分の関連する個別の文字要素を判定することをさらに含む、付記 5 に記載のコンピュータ実装方法。

【 0 0 6 9 】

付記 11 . 結合コンポーネントアルゴリズムを用いて、前記文字ピクセル値の閾に少なくとも合致するピクセル値を有する前記複数のピクセルのピクセルを関連付けることと、ストローク幅変換アルゴリズムを前記関連付けられるピクセルに適用することと、前記ストローク幅変換アルゴリズムを適用した結果を文字認識アルゴリズムに提供することと、をさらに含む、付記 5 に記載のコンピュータ実装方法。

50

【 0 0 7 0 】

付記 1 2 . 関連する個別の文字要素を判定した後に、前記個別の文字要素の支配的方向を判定することをさらに含む、付記 1 0 に記載のコンピュータ実装方法。

【 0 0 7 1 】

付記 1 3 . 前記個別の文字要素が少なくとも 1 つの既定の文字制約を満たすかどうかを判定することと、

個別の文字要素が前記既定の文字制約のうちの前記少なくとも 1 つを満たさない場合に、個別の文字要素を検討の対象から外すことと、をさらに含む、付記 1 0 に記載のコンピュータ実装方法。

【 0 0 7 2 】

付記 1 4 . 前記既定の文字制約は、文字要素間の間隔、隣接する文字要素の高さと幅の比率、または文字要素の垂直配列のうちの 1 つを含む、付記 1 3 に記載のコンピュータ実装方法。

【 0 0 7 3 】

付記 1 5 . プロセッサと、

前記プロセッサによって実行されると、前記プロセッサに、

1 つの画像内の 1 つの領域の 1 つ以上の輝度勾配を分析して、1 つ以上の可能性のある文字エッジ位置を特定させ、

前記 1 つ以上の可能性のあるエッジ位置の所定距離の範囲内にある 1 組のピクセルについてのピクセル値を分析させ、

前記 1 つ以上のエッジ位置のそれぞれの前記所定距離の範囲内にある前記ピクセル値に少なくとも部分的に基づいて、前記領域のための文字ピクセル値の閾を判定させ、

前記文字ピクセル値の閾と少なくとも合致するピクセル値を有する前記 1 組のピクセルの一部分に少なくとも部分的に基づいて、1 つ以上の文字部分を判定させる命令を含む記憶デバイスと、を備える、携帯型コンピューティングデバイス。

【 0 0 7 4 】

付記 1 6 . 前記命令は、前記プロセッサによって実行されると、前記携帯型コンピューティングデバイスが、

文字認識アルゴリズムを用いて前記文字部分を分析して、前記 1 つ以上の文字部分に関連付けられる少なくとも 1 つの文字を認識することをさらに可能にする、付記 1 5 に記載の携帯型コンピューティングデバイス。

【 0 0 7 5 】

付記 1 7 . 前記命令は、前記プロセッサによって実行されると、前記携帯型コンピューティングデバイスが、

少なくとも 1 つの追加の画像を取得し、

少なくとも 1 つの文字が認識され得ない場合に、前記少なくとも 1 つの追加の画像を分析することをさらに可能にする、付記 1 6 に記載の携帯型コンピューティングデバイス。

【 0 0 7 6 】

付記 1 8 . 前記命令は、前記プロセッサによって実行されると、前記携帯型コンピューティングデバイスが、

前記認識された少なくとも 1 つの文字に対応する文字パターンを特定し、

前記文字パターンに関連付けられる前記携帯型コンピューティングデバイス上のアプリケーションを判定し、

前記判定されたアプリケーションに前記認識された少なくとも 1 つの文字を提供することをさらに可能にする、付記 1 6 に記載の携帯型コンピューティングデバイス。

【 0 0 7 7 】

付記 1 9 . プロセッサによって実行されると、前記プロセッサに、

1 つの画像であって、複数のピクセルについてのピクセル値を含む、画像中の 1 つの領域の 1 つ以上の輝度勾配を分析させて 1 つ以上の可能性のある 1 つ以上の可能性のあるエッジ位置を特定させ、

10

20

30

40

50

前記 1 つ以上の可能性のあるエッジ位置の所定距離の範囲内にある前記画像の 1 組のピクセルについてのピクセル値を分析させ、

前記 1 つ以上の可能性のあるエッジ位置のそれぞれの前記所定距離の範囲内にある前記ピクセル値に少なくとも部分的に基づいて、前記領域のための文字ピクセル値の閾を判定させ、

前記文字ピクセル値の閾と少なくとも合致するピクセル値を有する前記 1 組のピクセルの一部分に少なくとも部分的に基づいて、1 つ以上の文字部分を判定させる命令を格納する、持続性コンピュータ可読記憶媒体。

【0078】

付記 20 . プロセッサによって実行されると、前記プロセッサにさらに、文字認識アルゴリズムを用いて前記 1 つ以上の文字部分をさらに分析して、前記 1 つ以上の文字部分に関連付けられる少なくとも一文字を認識させる命令を格納する、付記 19 に記載の持続性コンピュータ可読記憶媒体。

10

【0079】

付記 21 . 前記ピクセル値は、グレースケールピクセル輝度値、カラーピクセル輝度値、ピクセル飽和値、またはピクセル飽和レベルのうち少なくとも 1 つに対応する、付記 19 に記載の持続性コンピュータ可読記憶媒体。

【0080】

付記 22 . 前記ピクセル値を分析することは、前記 1 つ以上の可能性のあるエッジ位置の第 1 の側にある第 1 の組のピクセルを分析することと、前記 1 つ以上の可能性のあるエッジ位置の第 2 の側にある第 2 の組のピクセルを分析することを含む、付記 19 に記載の持続性コンピュータ可読記憶媒体。

20

【0081】

付記 23 . プロセッサによって実行されると、前記プロセッサにさらに、前記 1 つ以上の輝度勾配を分析する前に、前記画像をグレースケール画像にさらに変換させる命令を格納する、付記 19 に記載の持続性コンピュータ可読記憶媒体。

【0082】

付記 24 . プロセッサによって実行されると、前記プロセッサにさらに、命令前記文字ピクセル値の閾を下回るピクセル値を持つ前記領域の一部分をさらに削除させる命令を格納する、付記 19 に記載の持続性コンピュータ可読記憶媒体。

30

【0083】

付記 25 . プロセッサによって実行されると、前記プロセッサが、結合コンポーネントアルゴリズムを用いて、前記文字ピクセル値の閾に少なくとも合致するピクセル値を有する前記複数のピクセルの隣接するピクセルを関連付けて、文字の少なくとも一部分を形成することをさらに可能にする命令を格納する、付記 19 に記載の持続性コンピュータ可読記憶媒体。

【0084】

付記 26 . プロセッサによって実行されると、前記プロセッサにさらに、ストローク幅変換アルゴリズムを前記文字の少なくとも一部分に適用させ、前記ストローク幅変換アルゴリズムを適用した結果を文字認識アルゴリズムに提供させる命令を格納する、付記 25 に記載の持続性コンピュータ可読記憶媒体。

40

【 図 1 A 】

図 1A

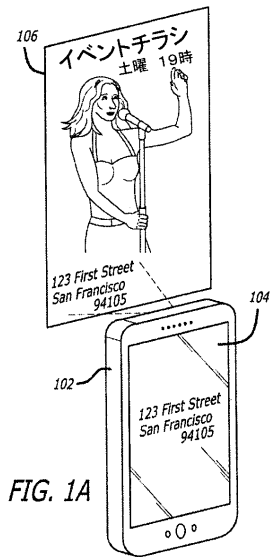


FIG. 1A

【 図 1 B 】

図 1B

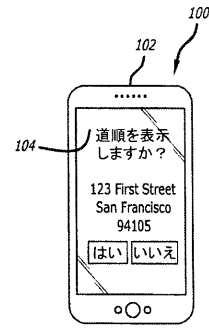


FIG. 1B

【 図 1 C 】

図 1C

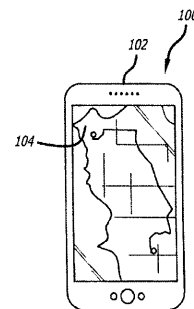


FIG. 1C

【 図 2 A 】

図 2A

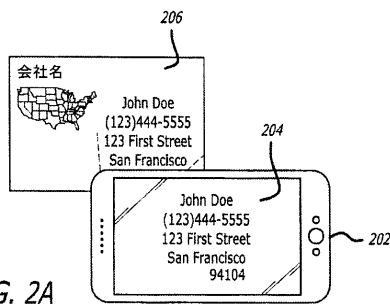


FIG. 2A

【 図 2 C 】

図 2C

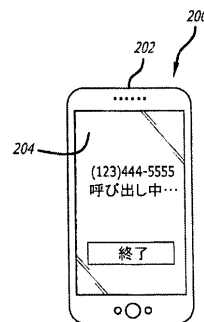


FIG. 2C

【 図 2 B 】

図 2B

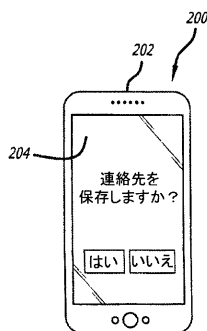


FIG. 2B

【 図 3 A 】

図 3A

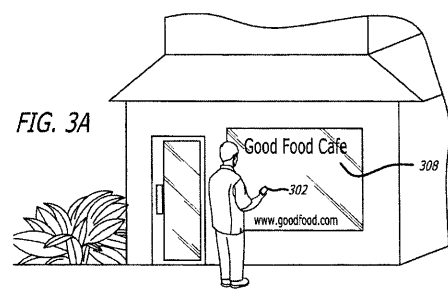


FIG. 3A

【 図 3 B 】

図 3B

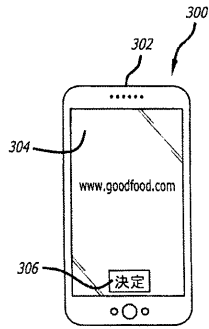


FIG. 3B

【 図 3 C 】

図 3C

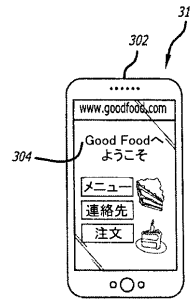
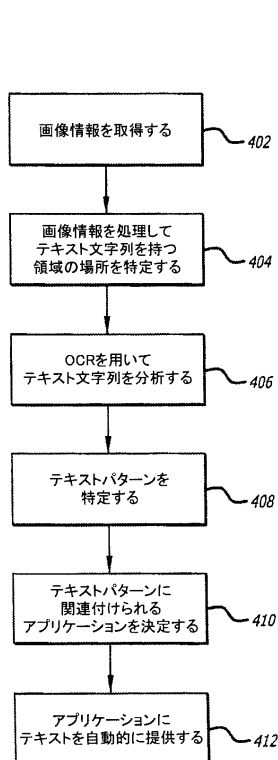


FIG. 3C

【 図 4 】

図 4

FIG. 4



【 図 5 】

図 5

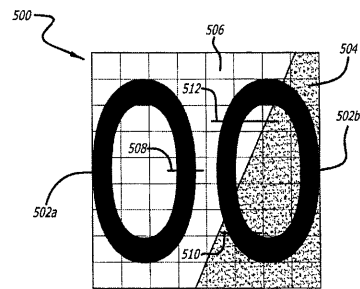


FIG. 5

【 図 6 】

図 6

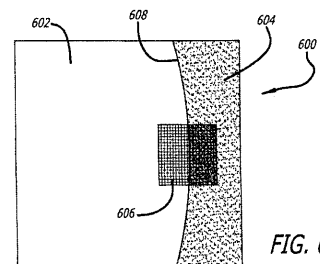
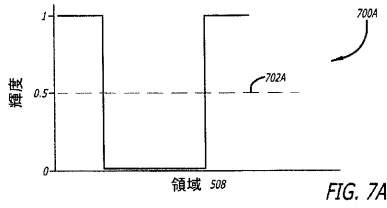


FIG. 6

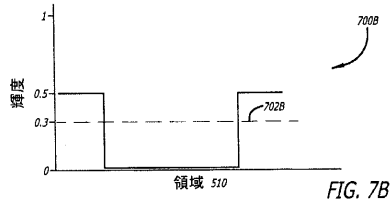
【 図 7 A 】

図 7A



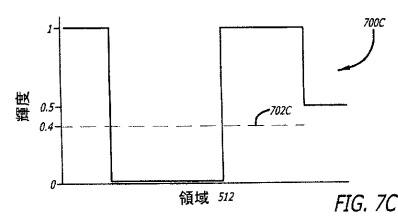
【 図 7 B 】

図 7B



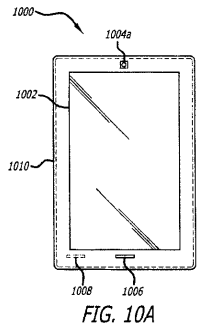
【 図 7 C 】

図 7C



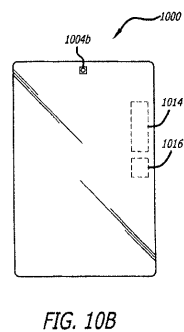
【 図 10 A 】

図 10A



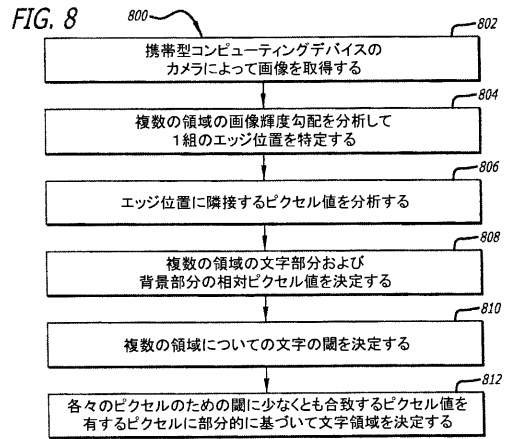
【 図 10 B 】

図 10B



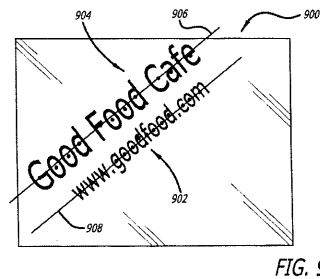
【 図 8 】

図 8



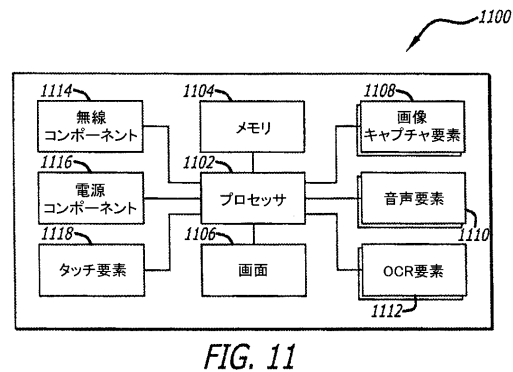
【 図 9 】

図 9



【 図 11 】

図 11



【 図 1 2 】

図 12

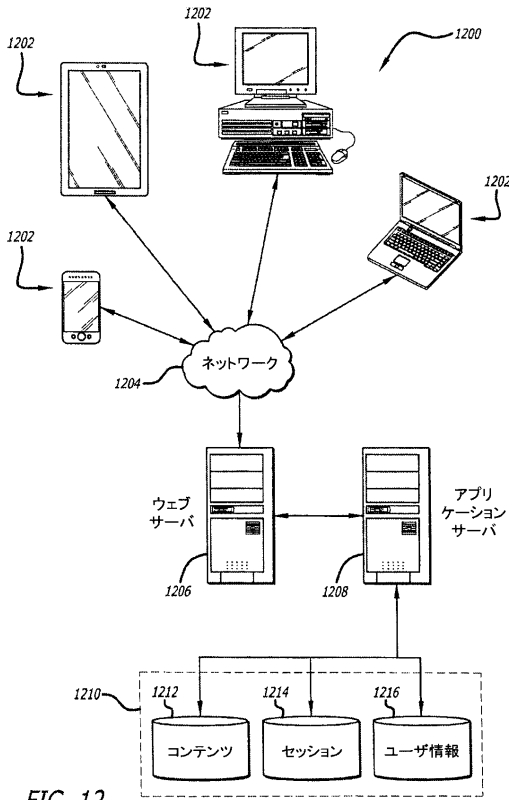


FIG. 12

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT

international application No.
PCT/US2013/044182

A. CLASSIFICATION OF SUBJECT MATTER IPC(8) - G06K 9/34 (2013.01) USPC - 382/176 According to International Patent Classification (IPC) or to both national classification and IPC			
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC(8) - G06K 7/10, 9/00, 9/20, 9/34, 9/46, 9/48, 9/66; G06F 3/00, 13/00, 15/18, 17/00; G06N 5/02; H04N 7/16, 7/173 (2013.01) USPC - 250/559.05-559.08, 559.42; 382/ 140-141, 175-177, 181, 189, 194-195, 199, 312-313, 321; 706/14, 16, 20-24, 47, 48; 715/205-206, 738, 740-744, 868; 725/19, 37, 42-45, 51-54, 59-61, 112 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched CPC - G01N 21/86; G01V 8/00 (2013.01) Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) Orbit, Google Patents, Google Scholar, PatentLens			
C. DOCUMENTS CONSIDERED TO BE RELEVANT			
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	
A	US 7,003,161 B2 (Tessadro) 21 February 2006 (21.02.2006) entire document	1-15	
A	US 5,717,781 A (Ebel et al.) 10 February 1998 (10.02.1998) entire document	1-15	
A	US 4,876,457 A (Bose) 24 October 1998 (24.10.1998) entire document	1-15	
A	US 6,941,016 B1 (Wagman et al.) 06 September 2005 (06.09.2005) entire document	1-15	
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/>			
* Special categories of cited documents:			
"A"	document defining the general state of the art which is not considered to be of particular relevance	"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
"E"	earlier application or patent but published on or after the international filing date	"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
"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)	"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
"O"	document referring to an oral disclosure, use, exhibition or other means	"&"	document member of the same patent family
"P"	document published prior to the international filing date but later than the priority date claimed		
Date of the actual completion of the international search 02 January 2014		Date of mailing of the international search report 31 JAN 2014	
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201		Authorized officer: Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774	

フロントページの続き

(81)指定国 AP(BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, 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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC

(74)代理人 100140176

弁理士 砂川 克

(74)代理人 100124394

弁理士 佐藤 立志

(74)代理人 100112807

弁理士 岡田 貴志

(74)代理人 100111073

弁理士 堀内 美保子

(72)発明者 イバンチェンコ、ボロディミル・ブイ .

アメリカ合衆国、ワシントン州 4 8 1 0 9、シアトル、テリー・アベ・ノース 4 1 0

(72)発明者 ヘラー、ジェフリー・スコット

アメリカ合衆国、ワシントン州 4 8 1 0 9、シアトル、テリー・アベ・ノース 4 1 0

(72)発明者 スプリー・ザ・サード、リチャード・ハワード

アメリカ合衆国、ワシントン州 4 8 1 0 9、シアトル、テリー・アベ・ノース 4 1 0

(72)発明者 ビビレータ、ダニエル

アメリカ合衆国、ワシントン州 4 8 1 0 9、シアトル、テリー・アベ・ノース 4 1 0

Fターム(参考) 5B029 AA01 AA02 BB02 BB17 CC28 CC29 DD05 DD06

5C077 MP01 MP08 PP22 PP27 PP28 PP47 PQ08 PQ20 SS05 SS07

TT09

5L096 AA02 AA06 BA08 BA17 EA35 EA43 FA44 FA60 GA32 GA51