

【特許請求の範囲】**【請求項 1】**

セルフチェックアウト車両を操作する買い物客によって選択された商品アイテムの 1 つまたは複数の画像をキャプチャするように構成された 1 つまたは複数のカメラと、

前記セルフチェックアウト車両に取り付けられ、前記セルフチェックアウト車両に置かれたときに前記選択された商品アイテムの重量を測定するように構成された少なくとも 1 つの重量センサと、

を有するセルフチェックアウト車両と、

少なくとも 1 つのプロセッサによって実行されると、少なくとも 1 つのプロセッサに、コンピュータビジョン技術を使用して前記画像から、前記選択された商品アイテムに付加された識別子に基づいて前記選択された商品アイテムを識別させる、非一時的媒体に格納されたコンピュータ実行可能命令を取得するように構成された少なくとも 1 つのプロセッサと、

を備えたセルフチェックアウト車両システム。

【請求項 2】

前記識別子は、グローバルトレードアイテム番号 (GTIN)、価格検索 (PLU) コード、及びバーコードのうち少なくとも 1 つを含む、請求項 1 に記載のセルフチェックアウト車両システム。

【請求項 3】

前記プロセッサは、少なくとも 1 つのプロセッサによって実行されると、前記プロセッサに、前記選択された商品アイテムの前記識別子及び前記測定された重量に基づいて、前記商品アイテムの価格を算出させる、非一時的媒体に格納されたコンピュータ実行可能命令を取得するようにさらに構成される、請求項 1 に記載のセルフチェックアウト車両システム。

【請求項 4】

前記プロセッサは、少なくとも 1 つのプロセッサによって実行されると、前記プロセッサに、前記選択された商品アイテム及びその価格を前記買い物客によって購入される商品アイテムの実行リストに追加させる、非一時的媒体に格納されたコンピュータ実行可能命令を取得するようにさらに構成される、請求項 3 に記載のセルフチェックアウト車両システム。

【請求項 5】

前記カメラの少なくとも 1 つは、その視野が、前記選択された商品が置かれる前記セルフチェックアウト車両の領域を含むように配置され、配向される、請求項 1 に記載のセルフチェックアウト車両システム。

【請求項 6】

前記視野は、前記重量センサの少なくとも 1 つをさらに含む、請求項 5 に記載のセルフチェックアウト車両システム。

【請求項 7】

前記プロセッサは、少なくとも 1 つのプロセッサによって実行されると、前記プロセッサに、前記重量センサによって提供される重量測定値を監視させ、商品アイテムが前記セルフチェックアウト車両から取り出されたことを示す測定重量の減少を検出させる、非一時的媒体に格納されたコンピュータ実行可能命令を取得するようにさらに構成される、請求項 1 に記載のセルフチェックアウト車両システム。

【請求項 8】

前記プロセッサは、少なくとも 1 つのプロセッサによって実行されると、前記プロセッサに、商品アイテムが前記セルフチェックアウト車両から取り出されたことを示す商品アイテムの移動を検出させる、非一時的媒体に格納されたコンピュータ実行可能命令を取得するようにさらに構成される、請求項 1 に記載のセルフチェックアウト車両システム。

【請求項 9】

前記商品アイテムが前記セルフチェックアウト車両から取り出されたことを示す商品ア

10

20

30

40

50

アイテムの移動を検出することは、

予め定められたしきい値距離を超えた前記商品アイテムの垂直方向の動きを検出することと、

前記1つまたは複数のカメラによってキャプチャされた画像に前記商品アイテムが存在しないことを検出することと、前記商品アイテムは、前記1つまたは複数のカメラによって以前にキャプチャされた画像に存在していた、ことと、

の少なくとも1つを含む、請求項8に記載のセルフチェックアウト車両システム。

【請求項10】

前記プロセッサは、少なくとも1つのプロセッサによって実行されると、前記プロセッサに、前記商品アイテムの不在を検出する前にキャプチャされた第1の重量測定値と、前記商品アイテムの不在を検出する際に、または検出後にキャプチャされた第2の重量測定値とを比較させ、前記第1の重量測定値と前記第2の重量測定値の差が、最初にセルフチェックアウト車両に置かれたときの前記選択された商品アイテムの測定重量に一致する場合、前記商品アイテムが取り出されたと判定させる、非一時的媒体に格納されたコンピュータ実行可能命令を取得するようにさらに構成される、請求項9に記載のセルフチェックアウト車両システム。

10

【請求項11】

前記プロセッサは、少なくとも1つのプロセッサによって実行されると、前記プロセッサに、コンピュータビジョンを使用して前記カメラによってキャプチャされた1つまたは複数の画像から、前記商品アイテムが前記セルフチェックアウト車両から取り出されることを識別させる、非一時的媒体に格納されたコンピュータ実行可能命令を取得するようにさらに構成される、請求項7または請求項8に記載のセルフチェックアウト車両システム。

20

【請求項12】

前記プロセッサは、少なくとも1つのプロセッサによって実行されると、前記プロセッサに、前記買い物客が購入する商品アイテムの実行リストから、前記セルフチェックアウト車両から取り出される前記商品アイテムを削除させる、非一時的媒体に格納されたコンピュータ実行可能命令を取得するようにさらに構成される、請求項11に記載のセルフチェックアウト車両システム。

【請求項13】

前記プロセッサは、少なくとも1つのプロセッサによって実行されると、前記プロセッサに、

前記識別子の確率的方向を識別することと、

前記確率的方向に基づいて、前記識別子を所定の方角に向けるために前記識別子を回転させる程度を決定することと、

前記識別子を前記決定された程度まで回転することと、

を実行させる、非一時的媒体に格納されたコンピュータ実行可能命令を取得するようにさらに構成される、請求項1に記載のセルフチェックアウト車両システム。

【請求項14】

前記決定された程度まで前記識別子を回転させることにより、前記識別子を任意の向きにして前記選択された商品を前記セルフチェックアウト車両に置くことを可能にする、請求項13に記載のセルフチェックアウト車両システム。

40

【請求項15】

商品を識別するコンピュータビジョンシステムであって、

少なくとも1つのプロセッサによって実行されると、少なくとも1つのプロセッサに、

買い物客によって選択された商品の1つまたは複数の画像内で、前記選択された商品に付加された識別子を識別することと、

前記画像のうちの少なくとも1つに基づいて、前記選択された商品のアイテムカテゴリを識別することと、

それぞれの画像内で識別された識別子とアイテムカテゴリを比較して、前記商品の最

50

も可能性の高い識別情報を決定することと、

を実行させる、非一時的媒体に格納されたコンピュータ実行可能命令を取得するようにさらに構成される、コンピュータビジョンシステム。

【請求項 16】

前記識別子は、グローバルトレードアイテム番号 (GTIN)、価格検索 (PLU) コード、及びバーコードのうち少なくとも 1 つを含む、請求項 15 に記載のコンピュータビジョンシステム。

【請求項 17】

それぞれの画像内で前記識別子を識別することは、

前記識別子を含む前記画像の部分をローカライズすることと、

前記識別子におけるテキストを識別するために、前記識別子を含む前記画像のローカライズされた部分に光学式文字認識 (OCR) 技術またはバーコード認識技術のいずれかを適用することと、

を含む、請求項 15 に記載のコンピュータビジョンシステム。

10

【請求項 18】

前記識別子を含む前記それぞれの画像の部分をローカライズすることは、

前記画像における前記識別子の位置を検出することと、

前記検出された位置にて、前記識別子を含む画像のピクセルを、前記選択された商品を含む画像のピクセルから区別することと、

前記識別子を含む画像のピクセルを所定の向きに回転することと、

を含む、請求項 17 に記載のコンピュータビジョンシステム。

20

【請求項 19】

前記画像における前記識別子の位置を検出することは、

前記それぞれの画像において、前記識別子が含まれる確率的領域を識別することと、

前記確率的領域を取り囲むバウンディングボックスを生成することと、

を含む、請求項 18 に記載のコンピュータビジョンシステム。

【請求項 20】

前記識別子を含む画像のピクセルを、前記選択された商品を含む画像のピクセルから区別することは、

前記選択された商品のナイーブアイテムカテゴリを生成することと、

前記位置にある画像のピクセルを前記ナイーブアイテムカテゴリに関連付けられたピクセルと比較することと、

前記ピクセルの比較から、前記ナイーブアイテムカテゴリに関連付けられたピクセルとは異なるピクセルを識別することと、

前記異なるピクセルを前記識別子を含むピクセルとして識別することと、

を含む、請求項 18 に記載のコンピュータビジョンシステム。

30

【請求項 21】

前記識別子を含む画像のピクセルを所定の向きに回転することは、

前記画像における前記識別子の確率的方向を識別することと、

前記確率的方向に基づいて、前記識別子を所定の方向に向けるために前記識別子を回転させる度合いを決定することと、

前記識別子を決定された度合いで回転させることと、

を含む、請求項 18 に記載のコンピュータビジョンシステム。

40

【請求項 22】

前記識別子におけるテキストを識別するために、前記識別子を含む前記画像のローカライズされた部分に光学式文字認識 (OCR) 技術を適用することは、

テキスト検出器を使用して、前記識別子を含む前記画像のローカライズされた部分のテキストをローカライズすることと、

前記ローカライズされたテキストを所定の向きに回転させることと、

畳み込みニューラルネットワーク (CNN) を使用して前記テキストの 1 つまたは複数

50

の特徴を抽出することと、

コネクショニスト時系列分類（CTC）を使用して、全ての可能なテキスト出力にわたる出力分布を生成することと、

を含む、請求項 17 に記載のコンピュータビジョンシステム。

【請求項 23】

前記出力分布から、可能性の高い出力を推測することと、

可能性の高い出力では任意の繰り返しを折りたたみ、可能性の高い出力では任意のブランクシンボルを除去することによって、前記識別子におけるテキストを識別することと、

をさらに含む、請求項 22 に記載のコンピュータビジョンシステム。

【請求項 24】

前記出力分布から、所与の出力の確率を評価することと、

最も高い確率を有する出力から前記識別子におけるテキストを識別することと、

をさらに含む、請求項 22 に記載のコンピュータビジョンシステム。

【請求項 25】

前記識別子におけるテキストを識別するために、前記識別子を含む前記画像のローカライズされた部分に光学式文字認識（OCR）技術を適用することは、

テキスト検出器を使用して、前記識別子を含む前記画像のローカライズされた部分のテキストをローカライズすることと、

前記ローカライズされたテキストを所定の向きに回転させることと、

画像二値化及び輪郭検出技術を使用してテキスト内の文字を分割することと、

各文字を認識するためにディープラーニング分類器または機械学習分類器のいずれかを使用して文字のバッチを評価することと、

認識された文字をシーケンスすることと、

を含む、請求項 17 に記載のコンピュータビジョンシステム。

【請求項 26】

前記選択された商品のアイテムカテゴリを識別することは、

前記選択された商品を含む前記画像の部分のローカライズすることと、

前記画像のローカライズされた部分に示された前記選択された商品の特徴表現を生成することと、

前記選択された商品の特徴表現を、データセット内の複数の利用可能な商品の特徴表現と比較して、選択された商品の特徴表現に最も類似した特徴表現を有する利用可能な商品を識別することと、

を含む、請求項 15 に記載のコンピュータビジョンシステム。

【請求項 27】

前記画像のローカライズされた部分に示された前記選択された商品の特徴表現を生成することは、前記選択された商品のローカライズされた部分から識別された前記選択された商品の特徴の多次元ベクトルマップを生成することを含み、

前記選択された商品の特徴表現を、データセット内の複数の利用可能な商品の特徴表現と比較して、選択された商品の特徴表現に最も類似した特徴表現を有する利用可能な商品を識別することは、

前記選択された商品の特徴ベクトルと複数の利用可能な商品の特徴ベクトルとの間の距離を計算することと、

その特徴ベクトルと前記選択された商品の特徴ベクトルとの間の最小距離を有する利用可能な商品を識別することと、

を含む、

請求項 26 に記載のコンピュータビジョンシステム。

【請求項 28】

それぞれの画像内で識別された識別子とアイテムカテゴリを比較して、前記商品の最も可能性の高い識別情報を決定することは、

前記選択された商品アイテムの 2 つ以上の画像において識別された各識別子に関連付け

10

20

30

40

50

られるアイテムカテゴリを識別することと、

関連するアイテムカテゴリ及びそれぞれの画像で直接識別されたアイテムカテゴリの中から、最も頻繁に識別されたアイテムカテゴリを選択することと、

他のものよりも頻繁に現れるアイテムカテゴリに基づいて前記商品を識別することと、を含む、請求項 15 に記載のコンピュータビジョンシステム。

【請求項 29】

それぞれの画像内で識別された識別子とアイテムカテゴリを比較して、前記商品の最も可能性の高い識別情報を決定することは、

前記選択された商品アイテムの 2 つ以上の画像において識別された各識別子に関連付けられるアイテムカテゴリを識別することと、

それぞれの関連する商品カテゴリ及び各画像内で直接識別される各商品カテゴリに重み係数を適用することと、

前記適用された重み係数を考慮して、他のものよりも頻繁に現れるアイテムカテゴリに基づいて前記商品を識別することと、

を含む、請求項 15 に記載のコンピュータビジョンシステム。

【請求項 30】

セルフチェックアウト車両に取り付けられ、選択された商品アイテムが前記セルフチェックアウト車両に置かれたときに前記セルフチェックアウト車両を操作する買い物客によって選択された商品アイテムの重量を測定するように構成された少なくとも 1 つの重量センサと、

その視野が前記重量センサの少なくとも 1 つを含むように前記セルフチェックアウト車両に配置され、配向される 1 つまたは複数のカメラであって、前記 1 つまたは複数のカメラは、少なくとも 1 つの重量センサ上の前記選択された商品アイテムの 1 つまたは複数の画像をキャプチャするように構成されている、1 つまたは複数のカメラと、

を有するセルフチェックアウト車両と、

少なくとも 1 つのプロセッサによって実行されると、少なくとも 1 つのプロセッサに、コンピュータビジョン技術を使用して前記画像から、前記選択された商品アイテムに付加された識別子に基づいて前記選択された商品アイテムを識別させる、非一時的媒体に格納されたコンピュータ実行可能命令を取得するように構成された少なくとも 1 つのプロセッサと、

を備えたセルフチェックアウト車両システム。

【発明の詳細な説明】

【技術分野】

【0001】

本出願は、2021年9月23日に出願された中国特許出願第2021111110492.6号に対する米国特許法第119条に基づく優先権を主張し、その開示は、その全体が参照により本明細書に組み込まれる。

【背景技術】

【0002】

ショッピングは、面倒で、ストレスの多い体験になることもあり得る。現在、商品を購入するには、顧客またはレジ係がショッピングカートから各商品を取り出し、商品のバーコードを見つけて商品をスキャンし、カートに戻す必要がある。レジ係や顧客がバーコードを検索してスキャンする必要があるため、チェックアウトカウンターで長い行列ができる。買い物客は、農産物の価格検索(PLU)コードステッカーを見つけるだけでなく、番号を読み取れるように商品を回転させ、販売時点管理(POS)端末にその番号を手動で入力できるようにする必要があるため、このプロセスは、農産物の場合によくあるように、コードを手動で入力する必要がある場合にさらに時間がかかる。また、正しい価格を決定するために農産物の重量を量る必要があり、プロセスがさらに複雑になる。

【0003】

ショッピング体験を向上させるために、いくつかの企業が「スマート」ショッピングカ

10

20

30

40

50

ートの導入を開始している。例えば、いくつかの企業は、カートに取り付けられたバーコードスキャナ及び販売時点管理（POS）端末を有し、買い物客がレジ待ち行列で待つことなく自分の商品をスキャンして支払うことを可能にする。さらに、農産物を計量するための内蔵スケールを備えているものもある。それでも、多くのそのようなシステムは、買い物客に、任意の所与の買い物中に上記の幾分厄介で時間のかかるプロセスを何度も通過することを要求する。

【0004】

従って、買い物客によって選択された商品アイテムの価格を自動的に識別、計量、及び計算することができるセルフチェックアウト車両システムが必要である。商品アイテムに付加された識別子の種類（例えば、グローバル取引識別番号（GTIN）、PLUコード、バーコードなど）に関係なく商品の識別を可能にするセルフチェックアウト車両と組み合わせ使用することができるコンピュータビジョンシステムのさらなる必要性がある。

10

【発明の概要】

【0005】

一態様では、本開示は、商品アイテムの価格を自動的に識別し、計量し、計算するためのセルフチェックアウト車両システムを対象とする。セルフチェックアウト車両システムは、セルフチェックアウト車両を操作する買い物客によって選択された商品アイテムの1つ以上の画像をキャプチャするように構成された1つ以上のカメラを備え得る、ショッピングカート及び/または搬送ステーションなどのセルフチェックアウト車両を含み得る。様々な実施形態によれば、セルフチェックアウト車両システムは、セルフチェックアウト車両に取り付けられ、セルフチェックアウト車両に置かれたときに選択された商品アイテムの重量を測定するように構成された少なくとも1つの重量センサを含み得る。セルフチェックアウト車両システムは、追加的または代替的に、非一時的媒体に格納されたコンピュータ実行可能命令を取得するように構成された少なくとも1つのプロセッサを含んでもよく、この命令は、少なくとも1つのプロセッサによって実行されると、少なくとも1つのプロセッサに、選択された商品の識別、計量、及び価格の計算のための1つ以上のコンピュータ実行ステップを行わせる。様々な実施形態では、1つ以上のコンピュータ実行ステップ（本明細書では単に「ステップ」とも言う）は、コンピュータビジョン技術を使用して画像から、選択された商品アイテムに付加された識別子に基づいて選択された商品アイテムを識別するステップだけでなく、選択された商品アイテムの識別子及び測定された重量に基づいて、商品アイテムの価格を計算するステップを含む。

20

30

【0006】

様々な実施形態によれば、セルフチェックアウト車両システムの少なくとも1つのカメラは、その視野が、選択された商品が置かれるセルフチェックアウト車両の領域を含むように配置され、配向され得る。いくつかの実施形態では、アイテムは、識別され、計量されるために2つの別々の位置に置かれる必要がないように、視野は、重量センサの少なくとも1つを含む。そのステップは、プロセッサに、識別子の確率的方向を識別することと、確率的方向に基づいて、識別子を所定の方向（例えば、表を上にする）に向けるために識別子の画像を回転させる程度を決定することと、識別子の画像を決定された程度まで回転することをさらに実行させる命令を含む。決定された程度まで識別子を回転させることにより、識別子を任意の向きにして選択された商品をセルフチェックアウト車両に置くことを可能にする。様々な実施形態によれば、識別子は、グローバルトレードアイテム番号（GTIN）、価格検索（PLU）コード、及びバーコードの少なくとも1つを含む。

40

【0007】

様々な実施形態では、セルフチェックアウト車両システムは、重量センサによって提供される重量測定値を監視し、商品アイテムがセルフチェックアウト車両から取り出されたことを示す測定重量の減少を検出するようにさらに構成され得る。追加的または代替的に、システムは、商品アイテムがセルフチェックアウト車両から取り出されていることを示す商品アイテムの動きを検出するように構成され得る。例えば、商品アイテムがセルフチェックアウト車両から取り出されたことを示す商品アイテムの動きを検出することは、予

50

め定められたしきい値距離を超えた商品アイテムの垂直方向の動きを検出することと、1つ以上のカメラによってキャプチャされた画像に商品アイテムが存在しないことを検出することと、当該商品アイテムは、1つ以上のカメラによって以前にキャプチャされた画像に存在していた、ことのうちの少なくとも1つを含み得る。一実施形態では、システムは、商品アイテムの不在を検出する前にキャプチャされた第1の重量測定値と、商品アイテムの不在を検出する際に、または検出後にキャプチャされた第2の重量測定値とを比較し、第1の重量測定値と第2の重量測定値の差が、最初にセルフチェックアウト車両に置かれたときの選択された商品アイテムの測定重量に一致する場合、商品アイテムが取り出されたと判定するように構成され得る。様々な実施形態では、システムは、コンピュータビジョンを使用してカメラによってキャプチャされた1つ以上の画像から、商品アイテムがセルフチェックアウト車両から取り出されたことを識別するように構成され得る。

10

【0008】

別の態様では、本開示は、商品を識別するコンピュータビジョンシステムを対象とする。コンピュータビジョンシステムは、セルフチェックアウト車両と、非一時的記憶媒体に格納されたコンピュータ実行可能命令を取得するように構成された少なくとも1つのプロセッサとを含み、及び/または共有することができ、命令は、少なくとも1つのプロセッサによって実行されると、少なくとも1つのプロセッサに、(i) 買い物客によって選択された商品の1つ以上の画像に、選択された商品に付加された識別子を識別することと、(ii) 少なくとも1つの画像に基づいて、選択された商品のアイテムカテゴリを識別することに向けさせることと、(iii) それぞれの画像において識別された識別子及びアイテムカテゴリを比較して商品の最も可能性の高い識別情報を決定することと、を実行させる。

20

【0009】

様々な実施形態では、コンピュータビジョンシステムは、識別子を含む画像の部分をローカライズし、及び識別子を定義するテキストを識別するために識別子を含む画像のローカライズされた部分に光学式文字認識(OCR)技術を適用することを含み得る。様々な実施形態によれば、命令は、プロセッサに、検出された位置にて識別子を含む画像のピクセルを選択された商品を含む画像のピクセルとは区別すること、及び識別子を含む画像のピクセルを所定の方向に回転させること、のうちの1つ以上を実行させることができる。いくつかの実施形態では、画像における識別子の位置を検出することは、プロセッサが、それぞれの画像において、識別子が含まれる確率的領域を識別し、及び/または確率的領域を取り囲むバウンディングボックスを生成する、ための命令を実行することを備える。

30

【0010】

識別子を含む画像のピクセルを、選択された商品を含む画像のピクセルから区別することは、様々な実施形態において、プロセッサが、選択された商品のナイーブアイテムカテゴリを生成することと、その位置にある画像のピクセルをナイーブアイテムカテゴリに関連付けられたピクセルと比較することと、ピクセルの比較から、ナイーブアイテムカテゴリに関連付けられたピクセルとは異なるピクセルを識別することと、異なるピクセルを識別子を含むピクセルとして識別することと、のうちの1つまたは複数のステップを実行することを備え得る。いくつかの実施形態では、識別子を含む画像のピクセルを所定の向きに回転することは、プロセッサが、画像における識別子の確率的方向を識別することと、確率的方向に基づいて、識別子を所定の方向に向けるために識別子を回転させる度合いを決定することと、識別子を決定された度合いで回転させることと、のうちの1つ以上のステップを実行することを含む。

40

【0011】

いくつかの実施形態によれば、識別子を定義するテキストを識別するために、識別子を含む画像のローカライズされた部分に光学式文字認識(OCR)技術を適用することは、プロセッサが、テキスト検出器を使用して、識別子を含む画像のローカライズされた部分のテキストをローカライズすることと、ローカライズされたテキストを所定の向きに回転させることと、畳み込みニューラルネットワーク(CNN)を使用してテキストの1つま

50

たは複数の特徴を抽出することと、コネクショニスト時系列分類（CTC）を使用して、全ての可能なテキスト出力にわたる出力分布を生成することと、のうちの1つまたは複数のステップを実行することを含む。いくつかの実施形態では、コンピュータビジョンシステムは、出力分布から、可能性の高い出力を推測し、及び可能性の高い出力では任意の繰り返しを折りたたみ、可能性の高い出力では任意のブランクシンボルを除去することのうちの1つまたは複数によって、識別子を定義するテキストを識別するコンピュータ実行可能命令を含むことができる。様々な実施形態では、命令は、出力分布から、所与の出力の確率を評価することと、最も高い確率を有する出力から識別子を定義するテキストを識別することと、のうちの1つまたは複数を含むことができる。

【0012】

いくつかの実施形態では、識別子を定義するテキストを識別するために、識別子を含む画像のローカライズされた部分に光学式文字認識（OCR）技術を適用することは、プロセッサが、テキスト検出器を使用して、識別子を含む画像のローカライズされた部分のテキストをローカライズすることと、ローカライズされたテキストを所定の向きに回転させることと、画像二値化及び輪郭検出技術を使用してテキスト内の文字を分割することと、各文字を認識するためにディープラーニング分類器または機械学習分類器のいずれかを使用して文字のバッチを評価することと、認識された文字をシーケンスすることと、のうちの1つまたは複数のステップを実行することを含むことができる。選択された商品のアイテムカテゴリを識別することは、プロセッサが、選択された商品を含む画像の部分をローカライズすることと、画像のローカライズされた部分に示された選択された商品の特徴表現を生成することと、選択された商品の特徴表現を、データセット内の複数の利用可能な商品の特徴表現と比較して、選択された商品の特徴表現に最も類似した特徴表現を有する利用可能な商品を識別することと、を含む1つまたは複数のステップを実行することを含むことができる。

【0013】

様々な実施形態において、画像のローカライズされた部分に示された選択された商品の特徴表現を生成することは、プロセッサが、選択された商品のローカライズされた部分から識別された選択された商品の特徴の多次元ベクトルマップを生成することを含む。いくつかの実施形態では、選択された商品の特徴表現を、データセット内の複数の利用可能な商品の特徴表現と比較して、選択された商品の特徴表現に最も類似した特徴表現を有する利用可能な商品を識別することは、プロセッサが、選択された商品の特徴ベクトルと複数の利用可能な商品の特徴ベクトルとの間の距離を計算し、及びその特徴ベクトルと選択された商品の特徴ベクトルとの間の最小距離を有する利用可能な商品を識別するための1つまたは複数の命令を実行することを含む。

【図面の簡単な説明】

【0014】

実例となる非限定的な例示的な実施形態は、添付図面と併せて行われる以下の詳細な説明からより明確に理解されるであろう。

【0015】

【図1】図1は、本開示の様々な実施形態によるセルフチェックアウト車両システムを示す。

【図2】図2は、本開示の様々な実施形態によるセルフチェックアウト車両を示す。

【図3】図3は、本開示の様々な実施形態による、セルフチェックアウト車両システムの様々なソフトウェアモジュール及び関連するアーキテクチャを示す。

【図4】図4は、本開示の様々な実施形態による、コンピュータビジョンシステムの様々なソフトウェアモジュール及び関連するアーキテクチャを示す。

【図5】図5は、本開示の一実施形態による、コンピュータビジョンモジュールにおける主な推論データフローを示す。

【図6】図6は、本開示のいくつかの実施形態による識別子ローカリゼーションモジュールの構成要素を示す。

10

20

30

40

50

【図 7】図 7 は、いくつかの実施形態による、識別子ローカリゼーションモジュールの推論パイプラインを示す。

【図 8】図 8 は、本開示のいくつかの実施形態による、識別子ローカリゼーションモジュール内のクラス、`box head`を示す。

【図 9】図 9 は、本開示のいくつかの実施形態による、識別子ローカリゼーションモジュール内のセグメンテーションヘッドを示す。

【図 10】図 10 は、本開示のいくつかの実施形態による識別子ローカリゼーションモジュール内の回転ヘッドを示す。

【図 11】図 11 は、本開示の一実施形態による中心周りの回転角の予測を示す。

【図 12】図 12 は、いくつかの実施形態による、識別子 OCR モジュールの概略を示す 10

【図 13】図 13 は、いくつかの実施形態による、識別子 OCR モジュール内のエンドツーエンド OCR データフローを示す。

【図 14】図 14 は、いくつかの実施形態による、テキスト画像特徴を抽出するためにシステムによって使用される CNN バックボーンを示す。

【図 15】図 15 は、いくつかの実施形態によるシステムによる CTC の実装を示す。

【図 16】図 16 は、いくつかの実施形態による推論における CTC を示す。

【図 17】図 17 は、いくつかの実施形態による CTC 復号の例を示す。

【図 18】図 18 は、本開示の様々な実施形態によるトレーニング中の CTC を示す。

【図 19】図 19 は、いくつかの実施形態による、識別子 OCR モジュールにおける従来の CV + ML の OCR データフローを示す。 20

【図 20】図 20 は、本開示の様々な実施形態によるアイテム埋め込みモジュールのトレーニングフローを示す。

【図 21】図 21 は、いくつかの実施形態によるアイテム埋め込みモジュールの推論フローを示す。

【図 22】図 22 は、いくつかの実施形態による結果融合モジュールを示す。

【図 23】図 23 は、いくつかの実施形態による投票プロセスを示す。

【図 24】図 24 は、いくつかの実施形態による重み付け投票を示す。

【図 25】図 25 は、いくつかの実施形態による、投票のための学習重みを示す。

【図 26】図 26 は、システムのいくつかの実施形態による、システム及び方法を可能にする、またはそれを含むコンピュータシステムを示す。 30

【発明を実施するための形態】

【0016】

本開示は、買い物客によって選択された商品アイテムの価格を自動的に識別、計量、及び計算するように構成されたセルフチェックアウト車両システムを対象とする。構成されるように、セルフチェックアウト車両システムの実施形態は、多くの方法でショッピング体験を改善することができる。例えば、商品アイテムがショッピングカートに置かれたときに商品アイテムの価格を自動的に識別、計量、及び計算することによって、セルフチェックアウトシステムの実施形態は、買い物客が商品アイテムを単に棚から取り出し、ショッピングカートに置くことを可能にし、商品アイテムをチェックアウトするために 1 つまたは複数の固定端末（例えば、店舗出口の近くに位置する販売時点管理（POS）端末、及び/または商品を識別及び計量し、POS 端末でスキャンされるステッカーを印刷するために使用される端末）に別々に持ち込む必要はない。例えば、様々な実施形態では、セルフチェックアウト車両システムは、買い物客が（a）商品アイテムを識別するためのコードを入力する必要がなく、（b）ショッピングカートに入れる前にスケール上で商品アイテムを計量する必要がなく、（c）任意選択で、固定 POS 端末でスキャンされる商品アイテムの識別子及び重量を記載するラベルを印刷する必要がなく、（d）行列に並んで固定 POS 端末で支払う必要がないように構成され得る。後でより詳細に説明されるように、これは、より速く、よりスムーズで、より便利なショッピング体験を提供し得る。

【0017】

本開示は、商品アイテムに付加された識別子に基づいて買い物客によって選択された商品アイテムを識別するコンピュータビジョンシステムをさらに対象とする。コンピュータビジョンシステムのいくつかの実施形態は、しばしば、グローバルトレードアイテム番号（GTIN）、製品検索（PLU）コード、バーコード、または購入手続きのために農産物アイテムを識別する買い物客またはチェックアウトクラークによって通常スキャンまたは手動で固定端末に入力される商品アイテムを識別する他の手段のうちの1つまたは組み合わせを特徴とするステッカーまたは他のラベルを有する農産物アイテムとの使用に特に適している。本開示のコンピュータビジョンシステムは、革新的なアーキテクチャ及び技術を利用して、ショッピングカートに置かれている間に商品アイテムに貼付されたそのような識別子を見つけて読み取り、それによって、買い物客が買い物を継続したり店を出たりする前に、誰か（例えば、買い物客またはチェックアウトクラーク）がチェックアウトプロセスの一部として識別子を別々にスキャンまたは入力する必要性を回避する。さらに、様々な実施形態では、本コンピュータビジョンシステムは、コンピュータビジョンシステムによってキャプチャされた、またはそうでなければコンピュータビジョンシステムに提供される画像内のその向きにかかわらず識別子を検索して読み取るように構成されることができ、それによって、買い物客が、識別子が読み取られるために特定の向きで提示されるように、商品アイテムを操作して保持することなく、商品アイテムをよりカジュアルにショッピングカートに入れることを可能にする。別の言い方をすれば、様々な実施形態では、コンピュータビジョンシステムは、商品アイテムがカメラの前に保持されていなくても、後でより詳細に説明されるように、識別子を右側上向きに提示する方法で識別子を検索して読み取るように構成され得る。

10

20

【0018】

説明を容易にするために、本開示のシステム及び方法は、ショッピングカートを使用してPLUコードでラベル付けされた農産物アイテムの買い物の文脈で説明され得るが、本開示は、そのように限定されることを意図するものではない。代わりに、本システム及び方法は、識別子（すなわち、単なるPLUコードではない）を特徴とする任意の商品アイテム（すなわち、単なる農産物ではない）、並びに選択された商品（すなわち、単なるショッピングカートではない）を運ぶために買い物客によって使用される任意の輸送手段と共に使用され得ることを認識されたい。さらに、本明細書に開示されるシステム及び方法は、価格が重量の関数である商品アイテム（例えば、ポンド当たりの価格を有する商品アイテム）と共に使用するために説明され得るが、いくつかの実施形態では、セルフチェックアウト車両システム及びコンピュータビジョンシステムは、重量を量る必要のない商品アイテム（例えば、アイテム当たりの固定価格を有する商品アイテム）と共に使用され得、そのような場合には、ロードレーバ/スケール/重量センサ及び関連する計量機能は、任意選択で含まれないことを認識されたい。

30

【0019】

セルフチェックアウト車両システム100

図1は、本開示の様々な実施形態によるセルフチェックアウト車両システム100を示す。一般的に言えば、セルフチェックアウト車両システム100は、買い物客によって選択された商品アイテムを収集する際に使用するセルフチェックアウト車両110（例えば、ショッピングカート）と、セルフチェックアウト車両110に取り付けられたカメラ114（後で図2に示される）によってキャプチャされた画像から選択された商品アイテム10を識別するように構成された1つまたは複数のプロセッサ120とを含み得る。プロセッサ120は、一実施形態では、オンボードセルフチェックアウト車両110（図示せず）に配置されてもよく、一方、他の実施形態では、プロセッサ120（または複数のプロセッサ120のうちの1つまたは複数）は、セルフチェックアウト車両110から離れて配置されてもよく、ネットワーク130（図示）を介してそれに接続されてもよい。先に図3を参照すると、代表的な実施形態では、特定の商品アイテム認識機能は、リモートプロセッサ120a上でリモートで実行されてもよく、他の機能は、後でより詳細に説明されるように、オンボードプロセッサ120b上でローカルに実行されてもよい。

40

50

【 0 0 2 0 】

図 2 は、セルフチェックアウト車両 1 1 0 の代表的な実施形態を示す。一般的に言えば、セルフチェックアウト車両 1 1 0 は、ショッピングカートまたは他の搬送手段 1 1 2 と、買い物客によってカート 1 1 2 に配置された商品アイテム 1 0 の画像をキャプチャする 1 つまたは複数のカメラ 1 1 4 と、カート 1 1 2 内の商品アイテム 1 0 を計量するための重量センサ 1 1 6 (例えば、ロードレーバ) と、セルフチェックアウト車両 1 1 0 を使用して買い物客にショッピング関連情報を表示するユーザ端末 1 1 8 とを含み得る。様々な実施形態では、カメラ 1 1 4 は、カート 1 1 2 の貨物エリアに向かって内側に向けられてもよく、カメラの少なくとも 1 つが、商品アイテム 1 0 をカート 1 1 2 に入れるときに商品アイテム 1 0 の識別子 1 2 の画像をキャプチャするように配置されてもよい。構成されているように、買い物客は、カメラの前で商品アイテム 1 0 をぎこちなく保持する必要がなく、またはそうでなければ、例えば、唯一のカメラが配置されているカート 1 1 2 の他の端部 / 側面まで歩き回る必要がなく、商品アイテム 1 0 をカート 1 1 2 にカジュアルに置くことができる。いくつかの実施形態では、カメラ 1 1 4 は、重量センサ 1 1 6 が、後述するように、それぞれの機能を別々に収容するためにアイテム 1 2 を移動させることなく、両方の機能を実行することができるため、アイテム 1 2 の計量及び識別のプロセスをさらに合理化することができるカメラ 1 1 4 の少なくともオンの視野内にあるように配置及び配向され得る。様々な実施形態では、セルフチェックアウト車両 1 1 0 は、商品アイテム 1 0 がカート 1 1 2 に入れられているか、またはカート 1 1 2 から取り出されているかを検出するために、1 つまたは複数のモーションセンサ 1 1 7 をさらに含み得る。セルフチェックアウト車両 1 1 0 及びその特定のコンポーネントのさらなる代表的な実施形態は、2018年4月18日に出版された「セルフチェックアウト盗難防止車両システム及び方法」と題された米国特許出願公開第 2019/0034897号、2020年8月18日に発行された「移動可能なデバイスで操作可能な正確な重量測定システム及び方法」と題された米国特許第 10,745,039号、及び2020年1月10日に出版された「オブジェクト識別及びセルフチェックアウト盗難防止のためのデータ生成を訓練するためのシステム及び方法」と題された米国特許出願公開第 2020/0151692号に示され、説明されており、これらのそれぞれは、全ての目的のためにその全体が参照により本明細書に組み込まれる。

10

20

【 0 0 2 1 】

様々な実施形態では、システム 1 0 0 は、セルフチェックアウト車両に取り付けられ、選択された商品アイテムがセルフチェックアウト車両に置かれたときにセルフチェックアウト車両を操作する買い物客によって選択された商品アイテムの重量を測定するように構成された少なくとも 1 つのロードレーバを有するセルフチェックアウト車両と、その視野がロードレーバのうちの少なくとも 1 つを含むようにセルフチェックアウト車両に配置及び配向された 1 つ以上のカメラとを含むことができ、1 つ以上のカメラは、少なくとも 1 つのロードレーバ上の選択された商品アイテムの 1 つ以上の画像をキャプチャするように構成される。構成されているように、プロセッサは、少なくとも 1 つのプロセッサによって実行されると、コンピュータビジョン技術を使用して画像から、カート 1 1 0 内の別の場所または他の場所でアイテム 1 0 を別個に計量する必要なく、選択された商品アイテムに付加された識別子に基づいて選択された商品アイテムをプロセッサに識別させる、非一時的媒体に格納されたコンピュータ実行可能命令を取得するように構成され得る。

30

40

【 0 0 2 2 】

図 3 は、本開示の様々な実施形態による、セルフチェックアウト車両システム 1 0 0 の様々なソフトウェアモジュール及び関連するアーキテクチャを示す。一般的に言えば、セルフチェックアウト車両システム 1 0 0 は、商品アイテム 1 0 を識別するためのコンピュータビジョンモジュール 1 3 0 と、商品アイテム 1 0 の重量を量り、コストを決定することに関連付けられた様々な機能を実行するためのイベントモジュール 1 4 0 と、ディスプレイ端末 1 1 8 上で買い物客にショッピング関連情報を提示するためのユーザエクスペリエンスモジュール 1 5 0 とを備えて構成されたプロセッサ 1 2 0 を含み得る。

50

【 0 0 2 3 】

コンピュータビジョンモジュール 1 3 0

様々な実施形態では、コンピュータビジョンモジュール 1 3 0 は、カメラ 1 1 4 によってキャプチャされた画像内の商品アイテム 1 0 を識別するのに適した任意のコンピュータモジュールを含み得る。様々な実施形態では、コンピュータビジョンモジュール 1 3 0 は、選択された商品アイテム 1 0 に付加された商品アイテム識別子 1 2 を少なくとも部分的に検索して読み取ることによって商品アイテム 1 0 を識別するように構成され得る。例えば、一実施形態では、コンピュータビジョンモジュール 1 3 0 は、識別子 1 2 を含む画像の一部をローカライズし、識別子 1 2 を含む画像のローカライズされた部分に光学式文字認識 (OCR) 技術を適用して、識別子 1 2 の全部または一部に含まれる、またはそうでなければ定義するテキストを識別するように構成され得る。識別子 1 2 が右側に向いていない場合、または OCR 技術を実行するのに適した別の方向に向いていない場合、様々な実施形態では、コンピュータビジョンモジュール 1 3 0 は、(a) 識別子 1 2 の確率的方向を識別し、(b) 確率的方向に基づいて、識別子 1 2 を所定の方向に向けるために識別子 1 2 を回転させる程度を決定し、(c) 識別子 1 2 の画像を決定された程度まで回転させるように構成され得る。決定された程度で識別子 1 2 を回転させることで、一実施形態では、選択された商品アイテム 1 0 を識別子 1 2 を任意の方向にしてセルフチェックアウト車両 1 1 0 に置くことができるようになる。追加的または代替的に、コンピュータビジョンモジュール 1 3 0 は、選択された商品アイテム 1 0 の商品カテゴリ (例えば、生産物、またはより詳細に言えば、トマトなどの一般的な種類の生産物) を決定し、いくつかの実施形態では、識別子 1 2 と各それぞれの画像で識別された商品カテゴリとを比較して、商品アイテム 1 0 の最も可能性の高い識別子を決定することができる。

10

20

【 0 0 2 4 】

追加的または代替的に、いくつかの実施形態では、コンピュータビジョンモジュール 1 3 0 は、後でより詳細に説明されるように、本開示のコンピュータビジョンシステム 2 0 0 の 1 つまたは複数の態様を含み得る。

【 0 0 2 5 】

イベントモジュール 1 4 0

イベントモジュール 1 4 0 は、様々な実施形態では、関心のあるイベントがいつ発生しているかを判定し、イベントに関連付けられた様々な機能を実行するのに適した任意のコンピュータモジュールを含み得る。代表的なイベントは、限定されないが、買い物客が商品アイテム 1 0 をセルフチェックアウト車両 1 1 0 に入れているか、またはセルフチェックアウト車両 1 1 0 から取り出していることを検出することを含み得る。イベントモジュール 1 4 0 は、様々な実施形態では、セルフチェックアウト車両が建物内のどこに配置されているかを検出するように構成された位置モジュールを含み得る。セルフチェックアウト車両の位置 (例えば、現在のショッピングセクション及び/または通路) を知ることによって、システムは、認識された商品の位置を建物内のその割り当てられた位置と比較することができる。本明細書でより詳細に説明されるように、マッチ (match) は、認識された商品の信頼レベルをさらに改善する。

30

【 0 0 2 6 】

いくつかの実施形態では、イベントモジュール 1 4 0 は、商品アイテム 1 0 が様々な方法でいつセルフチェックアウト車両 1 1 0 に置かれるか、またはセルフチェックアウト車両 1 1 0 から取り出されるかを検出するように構成され得る。一例では、モーションセンサ 1 1 7 は、商品アイテムが置かれるか、または取り出されることを示す、セルフチェックアウト車両 1 1 0 内の動きを検出し得る。一実施形態では、垂直面に配置された複数のセンサ 1 1 7 は、様々なモーションセンサ 1 1 7 が動きを検出する順序を監視することによって、アイテムがセルフチェックアウト車両 1 1 0 内に置かれる状況とセルフチェックアウト車両 1 1 0 から取り出される状況とを区別するように構成され得る。例えば、動きが最初にセルフチェックアウト車両 1 1 0 の上部の近くに位置するモーションセンサ 1 1 7 によって検出され、続いて、セルフチェックアウト車両 1 1 0 の中央または下部の近く

40

50

に位置するモーションセンサ 117 によって検出される場合、イベントモジュール 140 は、商品アイテム 10 がセルフチェックアウト車両 110 に置かれていることを推測し得る、逆の場合もまた同様に推測し得る。イベントモジュールは、様々な実施形態では、セルフチェックアウト車両 110 の位置を位置モジュールから受信し、検出された商品アイテム 10 の予想される位置をスキャンされた位置と比較し得る。位置が一致する場合、イベントモジュールは、以前に検出され、セルフチェックアウト車両 110 に存在する他の商品アイテムのリストに商品アイテム 10 を追加することを含む追加イベントを実行し得る。位置が一致しない場合、いくつかの実施形態によれば、システムは、ユーザに確認を促し、確認が受信されると商品アイテム 10 をリストに追加し、またはシステムが商品アイテムを誤って識別した場合、アイテムをメモリから削除することができる。同様のプロセスがカメラ 114 を使用して実行され得ることを理解されたい。すなわち、イベントモジュール 140 は、アイテムが連続して高くなっている、カメラ 114 によってキャプチャされた一連の画像（またはビデオ）を処理することによって、セルフチェックアウト車両 110 内の商品アイテム 10 の垂直移動を検出するように構成され得る。一実施形態では、イベントモジュール 140 は、アイテム 10 が所定の位置閾値（例えば、カート 110 の底部の上の特定の距離）に達したときにそのような決定を行うように構成され得る。さらに別の実施形態では、イベントモジュール 140 は、カメラ 114 によってキャプチャされた一連の画像（またはビデオ）を処理して、商品アイテム 10 が 1 つまたは複数のカメラ 114 によって以前にキャプチャされた画像に以前に存在していた場合、1 つまたは複数のカメラ 114 によってキャプチャされた画像内のその商品アイテム 10 の不在を検出することによって、セルフチェックアウト車両 110 内の商品アイテム 10 の垂直移動を検出するように構成されてもよい。そのような実施形態では、イベントモジュール 140 は、商品アイテム 10 の不在を検出する前にキャプチャされた第 1 の重量測定値を、商品アイテム 10 の不在を検出したときまたは検出した後にキャプチャされた第 2 の重量測定値とさらに比較し、第 1 の重量測定値と第 2 の重量測定値の差が、セルフチェックアウト車両 110 に最初に置かれたときの選択された商品アイテム 10 の測定された重量に対応する場合、商品アイテム 10 が取り出されたと決定することができる。言い換えれば、イベントモジュール 140 は、後述するように、カート 110 内の重量を使用して、どのカメラ 114 が商品アイテム 10 の取り出しの可能性を考慮したかを検証することができる。前述の実施形態のそれぞれにおいて、アイテム認識モジュールは、取り出されているときにアイテム 10 自体を識別することによって、または特定のアイテム 10 の不在を検出する前に、アイテム 10 がカート 110 内に置かれているときにアイテム 10 の以前にキャプチャされた画像に基づいてアイテム 10 を識別することによって、その取り出しが検出された商品アイテム 10 を識別するために利用され得る。

10

20

30

【0027】

追加的または代替的に、イベントモジュール 140 は、商品アイテムがセルフチェックアウト車両 110 に置かれるか、またはセルフチェックアウト車両 110 から取り出されることを示し得る重量の変化のために、重量センサ 116 によって提供される重量測定値を監視し得る。例えば、重量の増加が検出された場合、イベントモジュール 140 は、商品アイテム 10 がセルフチェックアウト車両 110 に置かれたことを推測することができる、逆の場合もまた同様に推測することができる。一実施形態では、イベントモジュール 140 は、商品アイテムがセルフチェックアウト車両 110 に追加されたか、またはセルフチェックアウト車両 110 から取り出されたかを決定する際に、重量センサ 116 及びモーションセンサ 117 の両方からのそのような入力を考慮し得る。例えば、イベントモジュール 140 は、モーションセンサ 116 を介して動きを検出すると重量測定値を監視し、重量に対応する変化がある場合、商品アイテムが追加または削除されたことを決定するように構成され得る。最初に動きを監視することによって、これは、セルフチェックアウト車両 110 内で回転するアイテムに起因して重量測定値が変化したものなどの誤ったイベント検出を回避するのに役立つ。追加的または代替的に、イベントモジュール 140 は、商品アイテム 10 がセルフチェックアウト車両 110 に追加されたか、またはセル

40

50

フチェックアウト車両 110 から削除されたかを決定する際に、コンピュータビジョンモジュール 130 からの入力を考慮するように構成され得る。例えば、一実施形態では、イベントモジュール 140 は、商品アイテム 10 がコンピュータビジョンモジュール 130 によって検出されたかどうか、そのような商品アイテム 10 が完全に識別されたかどうか、または商品カテゴリのみが決定されたかどうかを単に考慮し得る。商品アイテムが実際にセルフチェックアウト車両 110 の下部にて他のアイテムの上に存在したことを視覚的に確認することによって、イベントモジュール 140 は、商品アイテム 10 が追加または削除されたことを推測し得る。この考慮事項は、商品アイテム 10 がセルフチェックアウト車両 110 に追加または取り出されたかどうかを識別する際に前述の考慮事項（例えば、動き検出、重量変化）の 1 つまたは複数と同様の方法で組み合わせられ得る。

10

【0028】

位置モジュールは、いくつかの実施形態では、モーションセンサ 117 を含む 1 つまたは複数のセンサを使用して、セルフチェックアウト車両 110 が店舗内のどこにあるかを判定するように構成される。位置モジュールは、セルフチェックアウト車両 110 の位置及び/または座標をイベント決定モジュールに提供する。イベント決定モジュールが検出された商品アイテム 10 の位置を受信すると、イベント決定モジュールは、(1) 認識結果を検証するために位置に従って全ての可能な商品を把握するステップと、(2) (a) 認識された結果がその場所であると仮定される場合、成功した追加イベントを、または (b) ユーザに結果の正確性を確認することを要求する保留中の追加イベントをユーザインターフェース上に表示するステップのうちの 1 つまたは複数を実行する。

20

【0029】

様々な実施形態では、イベントモジュール 140 は、さらに、コンピュータビジョンモジュール 130 から商品アイテム 10 のアイデンティティに関する情報を受信し、受信したアイデンティティ及び(任意選択で)重量センサ 116 によって測定される商品アイテム 10 の重量に基づいて商品アイテム 10 のコストを決定するように構成され得る。例えば、イベントモジュール 140 は、コンピュータビジョンモジュール 130 によって提供される識別子に基づいて、商品アイテム 10 の価格(そのような価格がオンボードメモリにローカルに格納されるか、またはリモートメモリにリモートに格納されるか)を調べ、価格が重量の関数として提供される場合、コストに測定された重量を掛けることによって価格を計算することができる。価格が固定価格(すなわち、重量とは無関係)として提供される場合、イベントモジュール 140 は、商品アイテム 10 の価格を検索価格として単純に識別し得る。追加的または代替的に、価格がアイテムごとに固定価格である実施形態では、コンピュータビジョンモジュール 130 は、複数の商品アイテム 10 が同時に追加されるとき(例えば、一度に 1 つではなく 3 つの桃が追加される)を識別するように構成されることができ、その場合、イベント決定モジュール 140 は、一単位当たりのコストに、追加されたコンピュータビジョンモジュール 140 によって識別された商品アイテム 10 の数を掛けることによって、追加された商品アイテムのコストを計算することができる。もちろん、同様のアプローチは、商品アイテム 10 が代わりにセルフチェックアウト車両 110 から取り出されるときのコスト減少を決定するために使用され得る。

30

【0030】**ユーザエクスペリエンスモジュール 150**

様々な実施形態では、ユーザエクスペリエンスモジュール 150 は、ショッピング関連情報をディスプレイ端末 118 上で買い物客に提示するように構成され得る。例えば、一実施形態では、ユーザエクスペリエンスモジュール 150 は、セルフチェックアウト車両 110 内にあると判定された商品アイテム 10 の実行タブを処理することができる。商品アイテム 10 がセルフチェックアウト車両 110 に追加または削除されたかどうかに関する情報をイベント判定モジュール 140 から受信すると、ユーザエクスペリエンスモジュール 150 は、ディスプレイ端末 118 に、検出されたセルフチェックアウト車両 110 からの商品アイテム 10 の追加または削除を買い物客に通知し、一実施形態では、実行タブに追加または実行タブから削除された対応する料金を表示し得る。追加的または代替的

40

50

に、ユーザエクスペリエンスモジュール150は、セルフチェックアウト車両110に追加された商品アイテム10に基づいて、買い物客にレコメンデーション、クーポン、または他の提案を提供するように構成され得る。例えば、ユーザエクスペリエンスモジュール150は、商品アイテムのデータベースから、他の買い物客が選択された商品アイテム10と共に過去に購入したものに基づいて、買い物客が関心を持ち得る1つまたは複数の追加の商品アイテムを識別するように、つまり、例えば、買い物客がセルフチェックアウト車両110にパンケーキミックスを追加するとき、買い物客がシロップを購入することを推奨するように構成され得る。一実施形態では、ユーザエクスペリエンスモジュール150は、推奨されるアイテムが店舗内のどこに配置されているかをさらに識別し、そのような場所にナビゲートするための命令を買い物客に提供し得る。いくつかの実施形態では、ディスプレイ118は、買い物客がユーザエクスペリエンスモジュール150によって提供される特徴と対話することができるように、ユーザインターフェース（例えば、タッチスクリーン）を含み得る。さらに別の実施形態では、ディスプレイ端末118は、代わりに、買い物客によって提供される電子デバイス（例えば、タブレットまたはスマートフォン）であってよく、イベントモジュール150は、買い物客が衛生的な理由で以前の買い物客によって使用されたディスプレイに触れる必要がないように（例えば、Bluetooth（登録商標）を介して）ペアリングする。

【0031】

コンピュータビジョンシステム200

図4は、本開示の様々な実施形態による、コンピュータビジョンシステム200の様々なソフトウェアモジュール及び関連するアーキテクチャを示す。一般的に言えば、コンピュータビジョンシステム200は、商品アイテム10を識別するためのコンピュータビジョンモジュール230と、商品アイテム10の重量を量り、コストを決定することに関連付けられた様々な機能を実行するためのイベントモジュール240と、ディスプレイ端末118上で買い物客にショッピング関連情報を提示するためのユーザエクスペリエンスモジュール250とを備えて構成されたプロセッサ220を含み得る。様々な実施形態では、イベントモジュール240及びユーザエクスペリエンスモジュール250は、イベントモジュール140及びユーザエクスペリエンスモジュール150と実質的に同様であり得、いくつかの実施形態では、コンピュータビジョンモジュール230は、本開示のコンピュータビジョンモジュール130の1つまたは複数の態様を含み得る。それにもかかわらず、本開示のコンピュータビジョンシステム200は、(a) 選択された商品アイテム10に貼付された識別子を見つけて読み取り、(b) 選択された商品アイテム10の商品カテゴリ（例えば、生産物）を決定し、正確でフラストレーションのないショッピング体験を提供するために本明細書にさらに記載される様々な方法に従って(a)及び(b)を融合することに基づいて、選択された商品アイテム10をより正確に識別するためのワークフローで構成され得る。

【0032】

非限定的な例示的なワークフローは、コンピュータビジョンシステム200が、コンピュータビジョンモジュール230、イベントモジュール240、及びユーザエクスペリエンスモジュール250を含むプロセッサ220を起動すること、並びにカメラ114、重量センサ116、モーションセンサ117、及びイメージ署名プロセッサ（ISP）などの1つまたは複数の電子デバイスを起動することから始まる。様々な実施形態では、ISPは、ノイズ除去、オートホワイトバランス、並びに署名イメージ処理に関連付けられた任意の従来の動作を含む単純な画像処理を実行するように構成される。カメラ114は、データ転送コストを低減するために、ノイズ除去、サイズ変更、及びトリミングによってISPによって処理される商品アイテム10の画像ストリームを生成する。コンピュータビジョンモジュールは、ISPによって処理された商品データをフェッチし、商品データを2つのアルゴリズムパイプライン、すなわち、識別子認識モジュールと商品認識モジュールにフィードする。アイテム認識モジュールは、商品アイテム10の1つまたは複数の画像を使用することによってアイテムカテゴリを識別するように構成される。商品アイ

ム10が検出された後、特徴抽出ネットワークモジュールは、色、形状、テクスチャ、及び/または画像によって決定され得る任意の他の物理的属性のうちの一つ以上などの情報を含む特徴表現を抽出するように実装される。次に、特徴マッチング機構モジュールを実装して、抽出された特徴表現間の距離を計算し、それらをギャラリーデータベース内の特徴と比較する。アイテムカテゴリは、最小距離で定義される。選択された商品アイテム10のアイテムカテゴリを識別するシステムによって実装される一つまたは複数のステップは、プロセッサが、選択された商品アイテム10を含む画像の一部をローカライズすることと、画像のローカライズされた部分に示される選択された商品アイテム10の特徴表現を生成することと、選択された商品アイテム10の特徴表現をデータセット内の複数の利用可能な商品の特徴表現と比較して、選択された商品アイテム10の特徴表現に最も類似した特徴表現を有する利用可能な商品を識別することと、を実行することを含む。

10

【0033】

識別子認識モジュールは、商品アイテム10上の一つ以上の識別子12を認識することによって、商品アイテム10を識別する。識別子認識モジュールは、最初に入力フレーム内の一つ以上の識別子12をローカライズし、OCR技術を検出された領域に適用する。いくつかの実施形態では、従来のコンピュータビジョン(CV)及びディープラーニング(DL)のうちの一つ以上が、高速推論及び正確なテキスト認識結果を提供するために用いられる。テキストが認識された後、関連するアイテムカテゴリを取得するために識別子12データベースにクエリされる。識別子結果融合モジュールは、識別子OCRモジュール及びアイテム認識モジュールの両方からの一つ以上の履歴フレーム結果を入力として受け取り、ルールを使用して最も可能性の高い結果を取得する。一つまたは複数の重量センサ116は、セルフチェックアウト車両110内の全ての商品アイテム10のリアルタイムの重量デルタを与える。商品アイテム10の価格は、重量デルタによって計算され得る。イベント決定モジュールは、コンピュータビジョン結果、重量変化、及びモーションデータのうちの一つまたは複数を組み合わせて、顧客が商品アイテム10を入れることを意図するか、または商品アイテム10を取り除くことを意図するかを決定する。一つ以上のイベントが決定されると、システムは、イベント通知をユーザインターフェースに表示するように構成される。

20

【0034】

様々な非限定的な実施形態では、システムは、買い物客によって選択された商品アイテム10の一つまたは複数の画像において、選択された商品アイテム10に付加された識別子12を識別することと、画像のうち少なくとも一つに基づいて、選択された商品アイテム10のアイテムカテゴリを識別することと、識別子12及びそれぞれの画像で識別されたアイテムカテゴリを比較して、商品アイテム10の最も可能性の高い識別を決定することと、のうちの一つまたは複数を実装し得る。識別子12は、いくつかの非限定的な実施形態による、価格検索(PLU)コード、数値コードのうち少なくとも一つを含み得る。

30

【0035】

推論データフロー(Inference Dataflow)

図5は、本開示の一実施形態による、コンピュータビジョンモジュールにおける主な推論データフローを示す。推論データフローには2つのストリームがあり、それぞれ識別子12認識のためのストリーム及びアイテム認識のためのストリームである。識別子認識ストリームは、識別子12のラベルをローカライズし、OCRを用いて識別子12のラベル内の任意のテキスト(数字を含む)を認識する。識別子12がいくつかの実施形態に従って認識されると、商品アイテムカテゴリ(例えば、農産物)は、キー値データベースを使用して識別される。商品認識ストリームは、商品アイテム10がカメラ114から遠すぎる場合、商品アイテム10を検出する。アイテム画像は、キャプチャされた商品アイテム10の特徴表現を生成するために埋め込みネットワークに供給される。いくつかの実施形態では、特徴表現は、フロートまたは整数の多次元ベクトルであり、特徴データセット及び/またはギャラリーと呼ばれる各特徴ベクトルへの最小距離を選択することによってカ

40

50

テゴリを識別するために使用することができる。両方のストリームの出力カテゴリは、結果を融合することを担当する識別子結果融合アルゴリズムに流れ込み、経験に基づいてより信頼性が高く正確な予測を提供する。様々な非限定的な実施形態では、システムプロセッサは、識別子12を含む画像の一部をローカライズすることと、識別子12を含む画像のローカライズされた部分に光学式文字認識技術を適用して識別子12を定義するテキストを識別することと、の1つまたは複数のステップを実装し得る。システムはまた、画像内の識別子12の位置を検出することと、検出された位置で、識別子12を含む画像のそれらのピクセルを選択された商品を含む画像のそれらのピクセルから区別することと、識別子12を含む画像のそれらのピクセルを所定の向きに回転させることと、を含む1つまたは複数のステップを実装するように構成され得る。いくつかの実施形態では、システムは、識別子12が含まれる確率的領域を識別し、確率的領域を取り囲む境界ボックスを生成する。

10

【0036】

非バーコード識別子ローカリゼーションモジュール (Non Barcode Identifier Localization Module)

図6は、本開示のいくつかの実施形態による識別子ローカリゼーションモジュールの構成要素を示す。識別子ローカリゼーションモジュールは、検出、セグメンテーション、回転の3つの主な目標を持つ強力な検出器である。検出は、商品アイテム10及びナイーブな商品アイテムカテゴリの周りの境界ボックスを出力する。セグメンテーションは、0と1からなるマスクを出力することによって、フレーム内の正確な商品アイテムピクセルを予測する。値1は商品アイテムのピクセルを表し、値0は背景を表す。回転出力は、商品アイテム10の角度を-180度から180度の間で回転させる。商品アイテム10をセルフチェックアウト車両110に入れる挙動は予測不可能であり、識別子12がフレーム内で回転される可能性が最も高いため、回転較正はローカリゼーションモジュールで実行される。

20

【0037】

システムは、回転を決定及び/または実装するときに、識別子12の確率的方向を識別することと、確率的方向に基づいて、識別子12を所定の方向に向けるために識別子12を回転させる度合いを決定することと、識別子12を決定された度合いで回転させることと、のうちの1つまたは複数を実装し得る。決定された度合いだけ識別子12を自動的に回転させるシステムは、選択された商品が識別子12を任意の向きでセルフチェックアウト車両に置かれることを可能にする。

30

【0038】

識別子12を含む画像のそれらのピクセルを選択された商品を含む画像のそれらのピクセルから区別することは、1つ以上のプロセッサによって実行されるステップ、すなわち、選択された商品のナイーブアイテムカテゴリを生成するステップと、位置にある画像のピクセルをナイーブアイテムカテゴリに関連付けられたピクセルと比較するステップと、ピクセルの比較から、ナイーブアイテムカテゴリに関連付けられたピクセルとは異なるそれらのピクセルを識別するステップと、識別子12を含むそれらのピクセルとして異なるピクセルを識別するステップとを含む。

40

【0039】

図7は、いくつかの実施形態による、識別子ローカリゼーションモジュールの推論パイプラインを示す。識別子ローカリゼーションモジュールでの推論の間、各フレームは、ニューラルネットワークのバックボーンに供給される。いくつかの実施形態では、領域提案ネットワーク (RPN) 中間層は、バックボーン出力を入力として受け取り、確率的領域提案を生成する。いくつかの実施形態では、RPNは、非限定的な例として、CenterNet2に提示される確率的2段階検出器を含む1つ以上の修正を含む。追加の畳み込みは、提案された領域に接続され、より高いセマンティック特徴を抽出し、これらは、3つの予測ヘッド、すなわち、検出、セグメンテーション及び/または回転に供給される。予測された境界ボックス、マスク、及び回転角度を元のフレームに適用して、識別子12

50

のトリミングされ、セグメント化され、較正された画像を出力する。

【 0 0 4 0 】

図 8 は、本開示のいくつかの実施形態による、識別子ローカリゼーションモジュール内のクラス、`b b o x h e a d`を示す。クラス、`b b o x h e a d`の目的は、識別子ローカリゼーションモジュールで検出を実行することである。いくつかの実施形態では、クラス、`b b o x h e a d`は、予測されたクラス及びバウンディングボックスの両方を含むベクトルを出力するように構成される。ニューラルネットワーク（`N N`）層のスタック（例えば、畳み込み層及び/または完全に接続された層）が実行された後、カテゴリ予測及び`B`ボックス予測である2つの予測が生成される。いくつかの実施形態では、ナイーブなカテゴリ予測は、所与のアイテムが商品アイテム10（例えば、農産物）のタイプまたは背景である確率を提供する。いくつかの実施形態では、フロート閾値はまた、フロート閾値よりも小さい値を有する商品確率を排除することによって商品カテゴリを決定するために使用される。いくつかの実施形態では、システムは、`[x 1 , y 1 , x 2 , y 2]`などのその周囲の長方形座標によって商品アイテム10の予測された位置を与える回帰タスクである境界ボックス予測を実装する。

10

【 0 0 4 1 】

トレーニングの間、クラス、`b b o x h e a d`自体がマルチタスク学習アルゴリズムである。いくつかの実施形態では、ヘッドは、カテゴリ予測からのクロスエントロピー損失及び`b`ボックス予測からのユニオン上の交差（`I o U`）損失の変動から同時に学習するように構成され、定数はそれぞれ損失に異なる重みを与える。

20

【 0 0 4 2 】

【 数 1 】

$$\mathcal{L}_{class, bbox} = \mathcal{L}_{category} + \mathcal{L}_{bbox}$$

【 0 0 4 3 】

いくつかの実施形態では、推論の間、クラス、`b b o x h e a d`は、ベクトルを予測、`N x (4 + 1) x C`として出力する。ここで、`N`は提案の数を表し、4は説明された長方形座標であり、1はカテゴリ確率であり、`C`は背景を含むカテゴリの数である。

【 0 0 4 4 】

図 9 は、本開示のいくつかの実施形態による、識別子ローカリゼーションモジュール内のセグメンテーションヘッドを示す。セグメンテーションヘッドは、`[N , h , w , C]`の形状を持つマスクであるマスク予測のみを提供する。ここで、`N`は提案の数を表し、`h`及び`w`はいくつかの畳み込み層の後の提案の高さ及び幅であり、`C`は背景を含むカテゴリの数である。マスク予測の各ピクセルは0または1で表され、0が背景として、1がカテゴリとして表される。いくつかの実施形態では、トレーニングの間、マスク予測の各ピクセルは、グラウンドトゥルースピクセルと比較され、ネットワーク全体にバックプロパゲートする（`back propagate`）マスク損失を与える。推論の間、いくつかの実施形態に従って、クラス、`b b o x h e a d`内の予測されたカテゴリのみが選択される。

30

【 0 0 4 5 】

図 10 は、本開示のいくつかの実施形態による識別子ローカリゼーションモジュール内の回転ヘッドを示す。回転ヘッドは、中心の周りの度数での回転角度である角度予測を提供する。図 11 に示されるように、識別子12ラベルの中心点は、`[x _ c , y _ c]`によって表され、オレンジ線と黒線との間の角度は、いくつかの実施形態による角度予測である。トレーニングの間、いくつかの実施形態では、角度予測は、`M S E`、`S m o o t h L 1`などの回帰損失を使用して調整されることができる。推論の間、いくつかの実施形態に従って、全てのフォアグラウンドカテゴリは、この角度予測を使用して回転されなければならない。

40

【 0 0 4 6 】

識別子`O C R`モジュール（`Identifier OCR Module`）

50

識別子OCRモジュールは、トリミングされた識別子画像のバッチである識別子ローカリゼーションモジュールから出力を取り出す。有能な (competent) OCRアルゴリズムパイプラインは、入力画像を出力カテゴリに直接変換することができる。図12は、いくつかの実施形態による、識別子OCRモジュールの概略を示す。基本的に、エンドツーエンドのOCRアプローチと従来のコンピュータビジョン (CV) + 機械学習 (ML) アプローチという、同様のOCR結果を生成することができる2つのアプローチがある。エンドツーエンドのOCRアプローチは、パイプライン全体でディープラーニングを採用している。エンドツーエンドのOCRデータフローは、従来のCV + MLアプローチと比較して拡張性が高いという利点がある。

【0047】

エンドツーエンドOCRデータフロー (End to End OCR Dataflow)

エンドツーエンドOCRデータフローでは、識別子画像は、識別子画像内のテキストをローカライズするためにテキスト検出器に供給される。いくつかの実施形態では、テキスト検出器は、テキスト境界ボックスを検出するだけでなく、検出された境界ボックスの方向を合わせる。検出されたテキスト境界ボックスは、いくつかの実施形態に従って、畳み込みニューラルネットワーク (CNN) に供給され、続いて、接続リスト時系列分類 (CTC) デコーダに供給される。テキストが復号された後、それらは識別子12データベースにクエリされ、商品アイテム10の正しいカテゴリを決定する。図13は、いくつかの実施形態による、識別子OCRモジュール内のエンドツーエンドOCRデータフローを示す。いくつかの実施形態では、システムは、テキスト検出器を使用して、識別子12を含む画像のローカライズされた部分内のテキストをローカライズすることと、ローカライズされたテキストを所定の向きに回転させることと、畳み込みニューラルネットワーク (CNN) を使用してテキストの1つ以上の特徴を抽出することと、接続リスト時系列分類 (CTC) を使用して、全ての可能なテキスト出力にわたる出力分布を生成することと、を含む1つ以上のステップを実行するように構成される。これらのステップはまた、出力分布から、可能性の高い出力を推測し、可能性の高い出力では任意の繰り返しを折りたたみ、可能性の高い出力では任意のブランクシンボルを除去することによって、識別子12を定義するテキストを識別することを含み得る。

【0048】

テキスト検出 (Text Detection)

検出されたテキスト境界ボックスから、CNN + CTCデコーダは、正確なテキストコンテンツ (例えば、数字) を予測するように設計される。ビジュアル認識タスクのために、CNNは、しばしば識別子画像内の場合である歪み及びノイズに対するその良好な耐性のために使用される。図14は、いくつかの実施形態による、テキスト画像特徴を抽出するためにシステムによって使用されるCNNバックボーンを示す。CNNはまた、高レベルの特徴を抽出する能力を提供する：それらの境界ボックス化されたテキスト画像は、プーリングパディング及びいくつかのアクティベーションを伴ういくつかの畳み込み層を通過して、形状が $[b, h', w', c]$ の最終的に計算された特徴マップを取得する。ここで、 b は入力のバッチサイズであり、 h' 、 w' は、以前の畳み込みニューラルネットワークにおけるいくつかのストライディング及びプーリング操作の後の高さ及び幅を示し、 c は出力チャンネルの数である。いくつかの実施形態では、CNNは、画像のバッチを高レベル特徴マップのセットに変える通常のバックボーン (例えば、resnet 50、shufflenet v2、mobilenet v3) である。いくつかの実施形態では、抽出された特徴マップは、さらなる予測のために高レベルの意味的特徴としてCTCデコーダに送信される。

【0049】

画像内のテキスト領域サイズが変化し、識別子テキスト長が変化するため、システムは、いくつかの実施形態では、画像及びテキストコンテンツの正確な整列を欠くという問題を回避するために、接続リスト時系列分類 (CTC) デコーダを実装するように構成される。図15は、いくつかの実施形態によるシステムによるCTCの実装を示す。所

10

20

30

40

50

与の境界ボックス化されたテキスト画像について、従来のデコーダは固定長の出力しか予測できないため、CTCアルゴリズムは、数値を直接予測する代わりに、全ての可能なテキスト出力に対する出力分布を与える。いくつかの実施形態では、システムは、この分布を使用して、可能性の高い出力を推測するか、または所与の出力の確率を評価するように構成される。図16は、いくつかの実施形態による推論におけるCTCを示す。前の畳み込みニューラルネットワークによって生成された特徴マップは、再形成され、高密度層に供給される。高密度層の出力は、[b, t, c]に再形成される。ここで、bは入力のバッチサイズであり、tは事前定義されたCTCタイムスタンプ（例えば、図16の8）であり、cはカテゴリの数（可能な桁及び空白記号、例えば、図16の11）である。各CTCタイムスタンプにおける各文字の確率的分布を与える高密度層出力は、CTCデコーダによって復号される。

10

【0050】

推論の間、CTCデコーダは、折り畳みリピートを伴うCTC予測分布から最も可能性の高い出力を取り、空白記号を除去することによってテキスト予測を与え、テキスト予測は、最終的に識別子12データベースに送信され、出力カテゴリに変換される。

【0051】

CTCデコーダは、様々な実施形態に従って、2つの機能、すなわち、(1)各タイムスタンプで全てまたはいくつかの可能な形態にわたって分布（スコア）を計算すること、(2)異なる形態（例えば、466 8 88及び44688 88の両方の出力4688）で同じシーケンスをマージすること、を実行する。様々な実施形態では、システムは、出力分布から所与の出力の確率を評価し、最も高い確率を有する出力から識別子12を定義するテキストを識別するように構成される。

20

【0052】

様々な実施形態によれば、CTCデコーダの異なる実装がある：(1)CTCビームデコーダは、各タイムスタンプでのブルーニングサーチを行い、ビーム候補としてBの可能な形状を最上位にする、(2)CTCグリーディデコーダは、各タイムスタンプで最大確率分布をとり、繰り返しを折りたたみ、ブランクを除去する。CTC復号後、最高スコアを有するシーケンスが選択される（例えば、図16の4688）。図17は、いくつかの実施形態によるCTC復号の例を示す。3つのカテゴリ（0、1及び空白記号）及び2つのタイムスタンプを有することによって、(1)可能なテキスト出力の全てまたはいくつかはスコアとともに導出され、(2)同じシーケンスがマージされ、テキスト出力の繰り返しを折りたたみ、空白を除去した後にスコアが合計され（0、0及び00は全て0に変換され得る）、並びに(3)最高スコアを有するシーケンスが選択される。図18は、トレーニング中のCTCを示す。トレーニングの間、システムは、正式には $p(Y|X)$ を示す、正しい回答に inputs が割り当てられる可能性を最大化する必要がある。ここで、Xは入力画像を表し、Yはグラウンドトゥルースを表す。トレーニングセットDの場合、モデルのパラメータは、負の対数尤度を最小限に抑えるように調整される。

30

【0053】

【数2】

$$\sum_{(x,Y) \in D} -\log p(Y|X)$$

40

【0054】

従来のCV+MLのOCRデータフロー（Traditional CV+ML OCR Dataflow）

従来のCV+MLのOCRアプローチは、従来のCV技術を使用してテキストを文字に分割し、機械学習分類器を使用して各文字を認識する。図19は、いくつかの実施形態による、識別子OCRモジュール内の従来のCV+MLのOCRデータフローを示す。従来のCV+MLのOCRデータフローでは、識別子画像は、最初にテキスト検出器に供給され、識別子画像内のテキストをローカライズする。テキスト検出器は、テキスト境界ボッ

50

クスを検出するだけでなく、検出された境界ボックスの方向を合わせる。従来のコンピュータビジョンテキストスプリッタは、画像の二値化と輪郭検出を使用して、テキスト内の文字を分割する。全ての単一の文字がバッチとして形成され、文字(char)認識器に入力される。文字認識器は、ディープラーニング(DL)分類器またはサポートベクトルマシン(SVM)、ランダムフォレスト、KNN(K近傍)などのML分類器のいずれかであり得る。認識された文字は、識別子テキストを形成するための入力として配列される。テキストが復号された後、それらは識別子データベースにクエリされ、商品アイテム10の正しいカテゴリを決定する。従来のCV+MLのOCRデータフローの利点は、エンドツーエンドのディープラーニングアプローチよりも高速なことである。様々な実施形態によれば、システムは、テキスト検出器を使用して、識別子12を含む画像のローカライズされた部分内のテキストをローカライズすることと、ローカライズされたテキストを所定の向きに回転させることと、画像二値化及び輪郭検出技術を使用してテキスト内の文字を分割することと、各文字を認識するためにディープラーニング分類器または機械学習分類器のいずれかを使用して文字のバッチを評価することと、認識された文字をシーケンスすることと、を含む1つまたは複数のステップを実装するように構成される。

10

【0055】

アイテム埋め込みモジュール(Item Embedding Module)

アイテム埋め込みモジュールは、各商品アイテム10に従って識別的な特徴表現を提供することを目的とする。図20は、本開示の様々な実施形態によるアイテム埋め込みモジュールのトレーニングフローを示す。特徴表現をより識別的にするために、トレーニングフローではシャムネットワーク(siamese network)が使用される。シャムネットワークは、一对の画像を入力として受け取り、それぞれの特徴表現を抽出する。システムによって使用される各特徴表現は、多次元特徴ベクトルである。両方の特徴ベクトルは、対照的損失を計算するために使用される。損失は、共有された重みを持つ両方のネットワークで逆伝播される。

20

【0056】

アイテム埋め込みモジュールの推論はより簡単である。ネットワークは、アイテム画像を入力として受け取り、それに応じて特徴表現を出力する。予測された特徴ベクトルは、カテゴリ特徴ベクトルからなる特徴ギャラリーにクエリされ、距離最小で正しいカテゴリを取得する。図21は、いくつかの実施形態によるアイテム埋め込みモジュールの推論フローを示す。

30

【0057】

非バーコード識別子結果融合モジュール(Non Barcode Identifier Result Fusion Module)

結果融合モジュールは、識別子認識モジュール及びアイテム認識モジュールの両方からだけでなく、履歴結果からもカテゴリを受け入れる。結果を融合することによって、識別子結果融合モジュールは、はるかに信頼性が高く、堅牢な予測を提供することができる。図22は、いくつかの実施形態による結果融合モジュールを示す。結果融合モジュールのコア機能は、異なるモデルの予測を比較して精度を高めることである。いくつかの実施形態では、非限定的な例として、袋詰め(bagging)、ブースト、スタッキングなどのいくつかの関連する研究がある。いくつかの実施形態では、結果融合モジュールは、識別子認識モジュール及びアイテム認識モジュールからの結果を組み立てるように構成される。一度に1つの商品アイテムタイプ(例えば、1つのタイプの農産物)のみがシステム200に提示されるという事実などの履歴情報が考慮される。いくつかの実施形態では、履歴期間は、商品アイテム10をカメラ114から配置する時間よりも短い。システムは、モジュールの精度とタイムスタンプに応じて各予測に異なる重みを与えるように構成される。重み付けは、機械学習の方法で学習することができる。

40

【0058】

重み付け投票(Weighted Voting)

投票は、全ての分類器の大多数の予測によって融合予測を決定するアンサンブルメカニ

50

ズムである。図 2 3 は、いくつかの実施形態による投票プロセスを示す。非限定的な例で投票プロセスを説明するために、4つの予測があるとする。それらは、Red Sweet Bell Pepperに2つ、Orange Sweet Bell Pepperに1つ、Yellow Sweet Bell Pepperに1つある。投票結果は多数決なので、Red Sweet Bell Pepperである。重み付けされた投票は、各投票の重み付けにより、より良いモデルで結果をもたらす。図 2 4 は、いくつかの実施形態による重み付け投票を示す。

【0059】

いくつかの実施形態では、システムは、重み付け投票を実装し、現在のタイムスタンプまたは履歴タイムスタンプのいずれかからの全ての予測を投票として取得するが、それらは異なる重みを有するように構成される。

10

【0060】

投票するカテゴリ j 及び c の基礎分類器の数を考慮すると、各サンプルの重み付け投票の予測カテゴリ F は、以下の通りとなる。

【0061】

【数 3】

$$F = \operatorname{argmax}_j \sum_i^c p_{ji} \times w_i$$

【0062】

ここで、予測はバイナリである。 i 番目の基底分類器が p サンプルを j 番目のカテゴリに分類する場合、 $p_{ji} = 1$ であり、それ以外の場合、 $p_{ji} = 0$ である。ここで、 w_i は i 番目の基底分類器の重みである。重みは、現在のタイムスタンプに近いほど重みが大きくなり、モデルが優れているほど重みが大きくなるなど、何らかの戦略に従って事前に定義された定数にすることができる。しかしながら、重み付けは、機械学習によっても学習することができる。図 2 5 は、いくつかの実施形態による投票のための学習重みを示す。

20

【0063】

図 2 6 は、システムのいくつかの実施形態による、システム及び方法を可能にする、またはそれを含むコンピュータシステム 4 1 0 を示す。いくつかの実施形態では、コンピュータシステム 4 1 0 は、前述のシステム及び方法の 1 つ以上のソフトウェアモジュールのコンピュータ実行可能コードを操作及び / または処理することができる。さらに、いくつかの実施形態では、コンピュータシステム 4 1 0 は、システムと統合された、またはシステムに結合された 1 つまたは複数のグラフィカルユーザインタフェース (例えば、HMI) 内で情報を操作及び / または表示することができる。

30

【0064】

いくつかの実施形態では、コンピュータシステム 4 1 0 は、少なくとも 1 つのプロセッサ 4 3 2 を含むことができる。いくつかの実施形態では、少なくとも 1 つのプロセッサ 4 3 2 は、1 つ以上の従来のサーバプラットフォーム (図示せず) に常駐し、または結合することができる。いくつかの実施形態では、コンピュータシステム 4 1 0 は、ネットワークインターフェース 4 3 5 a 及び少なくとも 1 つのオペレーティングシステム 4 3 4 を処理することができる少なくとも 1 つのプロセッサ 4 3 2 に結合されたアプリケーションインターフェース 4 3 5 b を含むことができる。さらに、いくつかの実施形態では、少なくとも 1 つのプロセッサ 4 3 2 に結合されたインターフェース 4 3 5 a、4 3 5 b は、ソフトウェアモジュール (例えば、エンタープライズアプリケーション 4 3 8 など) のうちの 1 つ以上を処理するように構成され得る。いくつかの実施形態では、ソフトウェアアプリケーションモジュール 4 3 8 は、サーバベースのソフトウェアを含むことができ、少なくとも 1 つのユーザアカウント及び / または少なくとも 1 つのクライアントアカウントをホストするように動作することができ、少なくとも 1 つのプロセッサ 4 3 2 を使用してこれらのアカウントのうちの 1 つ以上の間でデータを転送するように動作することができる。

40

50

【 0 0 6 5 】

上記の実施形態を念頭に置くと、システムは、コンピュータシステムに格納されたデータを含む様々なコンピュータ実装動作を用いることができることが理解される。さらに、本開示を通して説明される上述のデータベース及びモデルは、様々な実施形態による、コンピュータシステム 4 1 0 内のコンピュータ可読記憶媒体及びコンピュータシステム 4 1 0 に結合されたコンピュータ可読記憶媒体に、分析モデル及び他のデータを記憶することができる。加えて、いくつかの実施形態では、システムの上述のアプリケーションは、コンピュータシステム 4 1 0 内のコンピュータ可読記憶媒体及びコンピュータシステム 4 1 0 に結合されたコンピュータ可読記憶媒体に記憶され得る。いくつかの実施形態では、これらの動作は、物理量の物理的操作を必要とする。通常、必ずしもそうではないが、いくつかの実施形態では、これらの量は、記憶され、転送され、結合され、比較され、さもなければ操作されることができる電氣的、電磁的、磁氣的、光学的、または磁気光学的信号のうちの一つ以上の形態をとる。いくつかの実施形態では、コンピュータシステム 4 1 0 は、少なくとも一つのデータソース 4 3 7 a、少なくとも一つのデータストレージ 4 3 7 b、及び/または少なくとも一つの入力/出力 4 3 7 c のうちの一つに結合された少なくとも一つのコンピュータ可読媒体 4 3 6 を含むことができる。いくつかの実施形態では、コンピュータシステム 4 1 0 は、コンピュータ可読媒体 4 3 6 上のコンピュータ可読コードとして具現化され得る。いくつかの実施形態では、コンピュータ可読媒体 4 3 6 は、データを格納することができ、その後、コンピュータ(コンピュータ 4 4 0 など)によって読み取ることができる任意のデータストレージであり得る。いくつかの実施形態では、コンピュータ可読媒体 4 3 6 は、所望の情報またはデータまたは命令を有形的に記憶するために使用することができ、コンピュータ 4 4 0 またはプロセッサ 4 3 2 によってアクセスすることができる任意の物理媒体または材料媒体であり得る。いくつかの実施形態では、コンピュータ可読媒体 4 3 6 は、ハードドライブ、ネットワークアタッチドストレージ(NAS)、読み取り専用メモリ、ランダムアクセスメモリ、フラッシュベースメモリ、CD-ROM、CD-R、CD-RW、DVD、磁気テープ、他の光学及び非光学データストレージを含むことができる。いくつかの実施形態では、様々な他の形態のコンピュータ可読媒体 4 3 6 は、ルータ、プライベートネットワークもしくはパブリックネットワーク、または有線及び無線の両方の他の伝送もしくはチャンネルを含む、リモートコンピュータ 4 4 0 及び/または少なくとも一つのユーザ 4 3 1 に命令を伝送または搬送することができる。いくつかの実施形態では、ソフトウェアアプリケーションモジュール 4 3 8 は、データベース(例えば、データベースを含むことができるデータソース 4 3 7 a 及びデータストレージ 4 3 7 b を含むコンピュータ可読媒体 4 3 6 から)からデータを送受信するように構成され得、データは、ソフトウェアアプリケーションモジュール 4 3 8 によって少なくとも一つの他のソースから受信され得る。いくつかの実施形態では、ソフトウェアアプリケーションモジュール 4 3 8 のうちの一つは、少なくとも一つのデジタルディスプレイ上にレンダリングされた少なくとも一つのグラフィカルユーザインタフェースを介して少なくとも一つのユーザ 4 3 1 にデータを出力するようにコンピュータシステム 4 1 0 内に構成され得る。

10

20

30

40

【 0 0 6 6 】

いくつかの実施形態では、コンピュータ可読媒体 4 3 6 は、ネットワークインターフェース 4 3 5 a を介して従来のコンピュータネットワークを通して分散されることができ、コンピュータ可読コードによって具現化されたシステムは、分散された方法で格納され、実行され得る。例えば、いくつかの実施形態では、コンピュータシステム 4 1 0 の一つまたは複数の構成要素は、ローカルエリアネットワーク(LAN) 4 3 9 a 及び/またはインターネット結合ネットワーク 4 3 9 b (例えば、無線インターネットなど)を介してデータを送信及び/または受信するように結合され得る。いくつかの実施形態では、ネットワーク 4 3 9 a、4 3 9 b は、ワイドエリアネットワーク(WAN)、直接接続(例えば、ユニバーサルシリアルバスポートを介して)、または他の形態のコンピュータ可読媒体 4 3 6、またはそれらの任意の組み合わせを含むことができる。

50

【 0 0 6 7 】

いくつかの実施形態では、ネットワーク 4 3 9 a、4 3 9 b の構成要素は、例えば、デスクトップコンピュータ、及び/またはラップトップコンピュータ、または LAN 4 3 9 a を介して結合された任意の固定された、一般的に非モバイルインターネット機器を含む任意の数のパーソナルコンピュータ 4 4 0 を含むことができる。例えば、いくつかの実施形態は、管理者を含む任意のタイプのユーザのために構成することができる LAN 4 3 9 a を介して結合されたパーソナルコンピュータ 4 4 0、データベース 4 4 1、及び/またはサーバ 4 4 2 のうちの 1 つ以上を含む。いくつかの実施形態は、ネットワーク 4 3 9 b を介して結合された 1 つ以上のパーソナルコンピュータ 4 4 0 を含むことができる。いくつかの実施形態では、コンピュータシステム 4 1 0 の 1 つまたは複数の構成要素は、インターネットネットワーク（例えば、ネットワーク 4 3 9 b など）を介してデータを送信または受信するように結合され得る。例えば、いくつかの実施形態は、少なくとも 1 つのユーザ 4 3 1 a、4 3 1 b を含み、無線で結合され、入力及び出力（「I/O」）4 3 7 c を介して少なくとも 1 つのエンタープライズアプリケーション 4 3 8 を含むシステムの 1 つ以上のソフトウェアモジュールにアクセスする。いくつかの実施形態では、コンピュータシステム 4 1 0 は、少なくとも 1 つのユーザ 4 3 1 a、4 3 1 b が、LAN 4 3 9 a を介して I/O 4 3 7 c を介してエンタープライズアプリケーション 4 3 8 にアクセスするように結合されることを可能にする。いくつかの実施形態では、ユーザ 4 3 1 は、デスクトップコンピュータ、及び/またはラップトップコンピュータ、またはインターネット 4 3 9 b を介して結合された任意の固定された、一般的に非モバイルインターネット機器を使用してコンピュータシステム 4 1 0 に結合されたユーザ 4 3 1 a を含むことができる。いくつかの実施形態では、ユーザは、コンピュータシステム 4 1 0 に結合されたモバイルユーザ 4 3 1 b を含むことができる。いくつかの実施形態では、ユーザ 4 3 1 b は、任意のモバイルコンピューティング 4 3 1 c を使用して、1 つ以上のパーソナルデジタルアシスタント、少なくとも 1 つのセルラー電話、少なくとも 1 つのモバイル電話、少なくとも 1 つのスマートフォン、少なくとも 1 つのページャ、少なくとも 1 つのデジタルタブレット、及び/または少なくとも 1 つの固定またはモバイルインターネットアライアンスを含むが、これらに限定されない、コンピュータシステム 4 1 0 に結合された無線に接続することができる。

10

20

30

【 0 0 6 8 】

本明細書に記載の主題は、新規の人工知能学習技術の使用による人工知能によるアイテム認識の分野への技術的改善を対象とする。本開示は、1 つ以上のプロセッサ及び 1 つ以上の非一時的コンピュータを含む 1 つ以上のコンピュータを含むマシンが、先行技術に対するシステム及びその改善をどのように実装するかの詳細を説明する。機械によって実行される命令は、人間の心で実行されたり、ピンと紙を使用して人間によって導き出されたりすることはできず、機械はプロセス入力データを有用な出力データに変換する必要がある。さらに、本明細書に提示される特許請求の範囲は、汎用コンピュータによって実装される既知の従来ステップと司法例外を結びつけることを試みるものではなく、単にそれを技術分野にリンクすることによって司法例外を結びつけることを試みるものでもない。実際、本明細書に記載のシステム及び方法は、出願時点で未知であり、及び/またはパブリックドメインに存在せず、従来技術では知られていなかった技術的改善の利点を提供する。さらに、システムは、特許請求の範囲を有用なアプリケーションに限定する非従来型のステップを含む。

40

【 0 0 6 9 】

システムは、その適用において、前述の説明に記載された、または図面に示された構成の詳細及び構成要素の配置に限定されないことが理解される。本明細書に開示されるシステム及び方法は、多数の実施形態の範囲内にある。前述の説明は、当業者がシステムの実施形態を作成及び使用することを可能にするために提示される。いくつかの実施形態に含まれる構造及び/または原理の任意の部分は、いずれか及び/または全ての実施形態に適用することができる：本明細書に提示されるいくつかの実施形態からの特徴は、いくつか

50

他の実施形態による他の特徴と組み合わせられることが理解される。従って、システムのいくつかの実施形態は、図示されるものに限定されることを意図するものではなく、本明細書に開示される全ての原理及び特徴と一致する最も広範囲を付与されるべきである。

【0070】

システムのいくつかの実施形態は、特定の値及び/または設定値で提示される。これらの値及び設定値は、限定することを意図するものではなく、より低い構成と比較してより高い構成の単なる例であり、システムを作成及び使用する当業者のための補助として意図されている。

【0071】

さらに、出願人自身の辞書編纂者として、出願人は、特許請求の範囲の明示的な意味及び/または否認を以下の用語に付与する。 10

【0072】

出願人は、例えば、「A及び/またはB」、または「A及び/またはBの少なくとも1つ」などの「及び/または」の任意の使用を、要素A単独、要素B単独、または要素A及びBを一緒に意味するように定義する。加えて、「A、B、及びCのうちの少なくとも1つ」の記載、「A、B、またはCのうちの少なくとも1つ」の記載、または「A、B、またはCのうちの少なくとも1つ、またはそれらの任意の組み合わせ」の記載はそれぞれ、要素A単独、要素B単独、要素C単独、または要素A、B、及びCの任意の組み合わせ、例えば、AB、AC、BC、またはABCを意味するように定義される。

【0073】

値と併せて使用される場合、「実質的に」及び「約」は、測定される同じ単位及び/またはスケールの5%以下の差を包含する。 20

【0074】

本明細書で使用される「同時に」は、複数のタイプのデータを同時に処理しようとする本明細書に記載のプロセッサ及び/またはネットワークなどの従来の及び/または独自のコンピュータに関連付けられたラグ及び/またはレイテンシ時間を含む。「同時に」は、無線及び/または有線ネットワークを介して、及び/またはプロセッサ回路内で、デジタル信号が1つの物理的な位置から別の物理的な位置に転送されるのにかかる時間も含む。

【0075】

本明細書で使用される場合、「can」または「may」またはそれらの誘導体（例えば、システムディスプレイは、Xを示すことができる）は、説明目的にのみ使用され、システムの測定値及び境界を定義するとき、「構成された」（例えば、コンピュータは、命令Xを実行するように構成される）と同義及び/または交換可能であると理解される。 30

【0076】

加えて、「するように構成される」という用語は、明細書及び/または特許請求の範囲に記載された制限が、記載された機能を実行するためにそのような方法で配置されなければならないことを意味する：「するように構成される」は、記載された機能を実行するように修正されることが「可能である」当該技術分野の構造を除外するが、当該技術分野に関連する開示は、そうするための明確な教示を有していない。例えば、「構造Xから上部に流体を受け入れ、下部から構造Yに流体を送るよう構成された容器」の記載は、構造X、構造Y、及び容器が全て、列挙された機能を実行するように配置されているように開示されているシステムに限定される。記載された文「するように構成された」は、単にそれらの構成によって列挙された機能を実行することが「可能であり得る」要素を除外するが、関連する開示（またはその欠如）は、列挙された全ての構造間の機能的制限を満たすようにそのような修正を行うための教示を提供しない。別の例は、「一連の命令X、Y、及びZを実行するように構成された、またはプログラムされたコンピュータシステム」である。この例では、命令は、コンピュータシステムが、列挙された命令を実行するように「構成される」及び/または「プログラムされる」ように非一時的なコンピュータ可読媒体上に存在しなければならない：「構成される」及び/または「プログラムされる」は、列挙された命令がその上に格納されているが、命令X、Y、及びZの教示がプログラムさ 40 50

れて格納されていない非一時的なコンピュータ可読媒体を有するコンピュータシステムを示唆する技術を除く。記載された文「するように構成される」はまた、物理的構造と併せて使用される場合、動作可能に接続されたものと同義であると解釈され得る。

【0077】

本明細書にて使用される表現及び用語は、説明を目的とするものであり、限定するものとみなすべきではないことが理解されよう。本明細書における「含む (including)」、「備える (comprising)」、または「有する (having)」及びこれらの変形の使用は、それ以降に掲載されるアイテム及びその同等物並びに付加的なアイテムを包含することが意味される。特に指示又は限定がない限り、「搭載 (mounted)」、「接続 (connected)」、「サポート (supported)」及び「結合 (coupled)」という用語及びこれらの変形は広義に使用されるものであり、直接的及び間接的な搭載、接続、サポート及び結合を包含する。更に、「接続」及び「結合」は、物理的又は機械的な接続又は結合に制限されるものでない。

10

【0078】

前述の詳細な説明は、図面を参照して読まれるべきであり、異なる図面の同様の要素は、同様の参照番号を有する。必ずしも一定のスケールではない図面は、いくつかの実施形態を示し、システムの実施形態の範囲を限定することを意図していない。

【0079】

本発明の一部を形成する本明細書に記載の動作のいずれかは、有用な機械動作である。本発明はまた、これらの動作を実行するためのデバイスまたは装置に関する。装置は、専用コンピュータなど、必要な目的のために特別に構築することができる。専用コンピュータとして定義される場合、コンピュータはまた、特殊目的の一部ではない他の処理、プログラム実行、またはルーチンを実行することができ、同時に特殊目的のために動作することができる。あるいは、動作は、コンピュータメモリに格納された、キャッシュに格納された、またはネットワークを介して取得された1つまたは複数のコンピュータプログラムによって選択的にアクティブ化または構成された汎用コンピュータによって処理され得る。データがネットワークを介して取得されるとき、データは、ネットワーク上の他のコンピュータ、例えば、コンピューティングリソースのクラウドによって処理され得る。

20

【0080】

本発明の実施形態はまた、ある状態から別の状態にデータを変換する機械として定義することができる。データは、電子信号として表され、データを電子的に操作することができる物品を表すことができる。変換されたデータは、場合によっては、ディスプレイ上に視覚的に描写され、データの変換から生じる物理的オブジェクトを表し得る。変換されたデータは、一般的に、または物理的及び有形のオブジェクトの構築または描写を可能にする特定のフォーマットでストレージに保存することができる。いくつかの実施形態では、操作は、プロセッサによって実行することができる。このような例では、プロセッサは、データのあるものから別のものに変換する。さらに、いくつかの実施形態は、ネットワークを介して接続することができる1つまたは複数のマシンまたはプロセッサによって処理され得る方法を含む。各マシンは、ある状態または事物から別の状態または事物にデータを変換することができ、また、データを処理すること、データをストレージに保存すること、ネットワークを介してデータを送信すること、結果を表示すること、または結果を別のマシンに通信することもできる。本明細書で使用されるコンピュータ可読記憶媒体は、(信号とは対照的に)物理的または有形の記憶装置を指し、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、または他のデータなどの情報の有形記憶のための任意の方法または技術で実装される揮発性及び不揮発性、リムーバブル及びリムーバブルな記憶媒体を含むが、これらに限定されない。

30

40

【0081】

いくつかの実施形態によれば、方法動作は特定の順序で提示されるが、それらのステップの実行は、明示的に指定されない限り、必ずしもリストされた順序で発生するわけではない。他のハウスキーピング動作は、動作の間に行うことができ、動作は、それらが

50

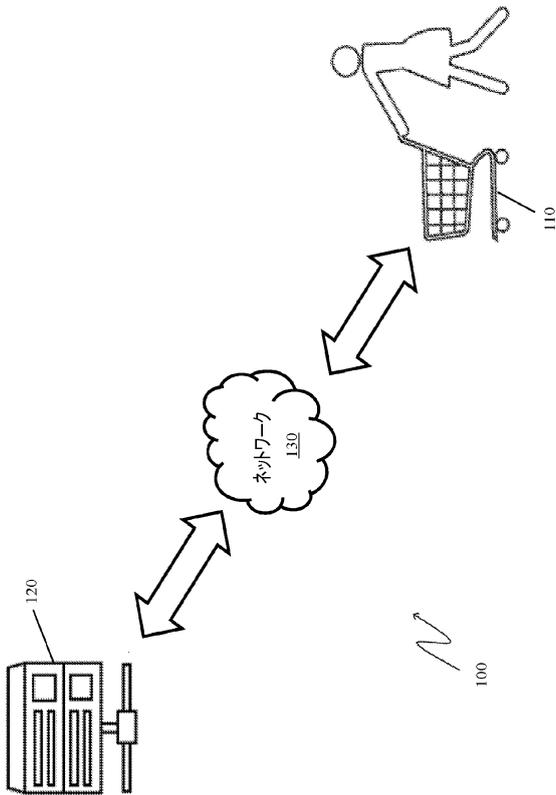
わずかに異なる時間に発生するように調整することができ、及び/または動作は、オーバレイ動作の処理が所望の方法で実行され、所望のシステム出力をもたらす限り、処理に関連付けられた様々な間隔で処理動作の発生を可能にするシステムに分散することができる。

【0082】

本開示及びその利点は詳細に説明されているが、添付の特許請求の範囲によって定義される本開示の精神及び範囲から逸脱することなく、本明細書で様々な変更、置換及び代替を代替を行えることを理解されたい。さらに、本出願の範囲は、本明細書に記載のプロセス、機械、製造、物質の組成、手段、方法及びステップの特定の実施形態に限定されることを意図するものではない。当業者が本開示から容易に理解するように、本明細書に記載の対応する実施形態と実質的に同じ機能を実行する、または実質的に同じ結果を達成する、現在存在する、または後で開発されるプロセス、機械、製造、物質の組成物、手段、方法、またはステップは、本開示に従って利用され得る。従って、添付の特許請求の範囲は、それらの範囲内に、そのようなプロセス、機械、製造、物質の組成物、手段、方法またはステップを含むことを意図している。

【図面】

【図1】



【図2】

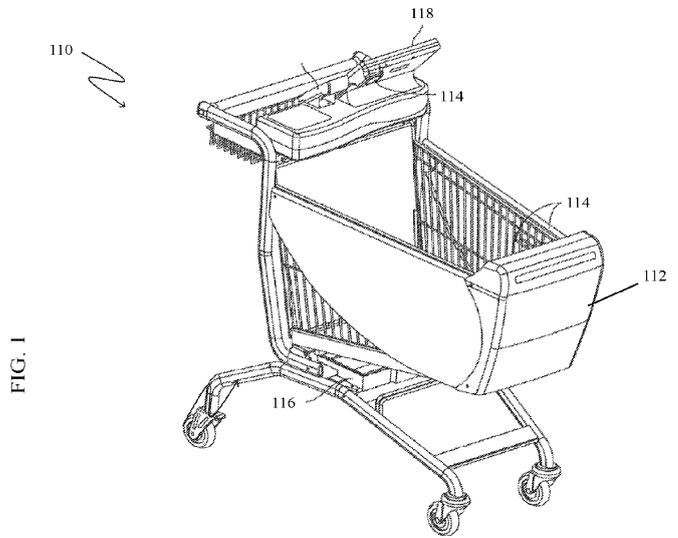


FIG. 1

FIG. 2

10

20

30

40

50

【 図 3 】

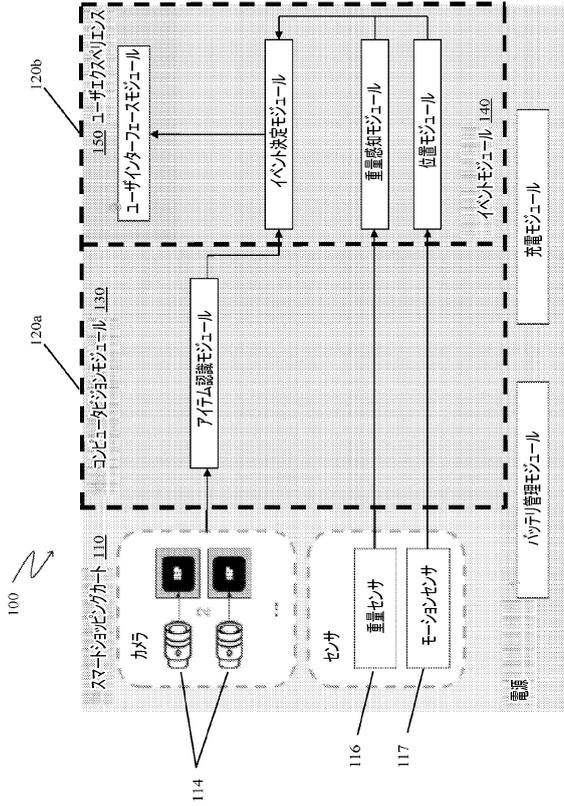


FIG. 3

【 図 4 】

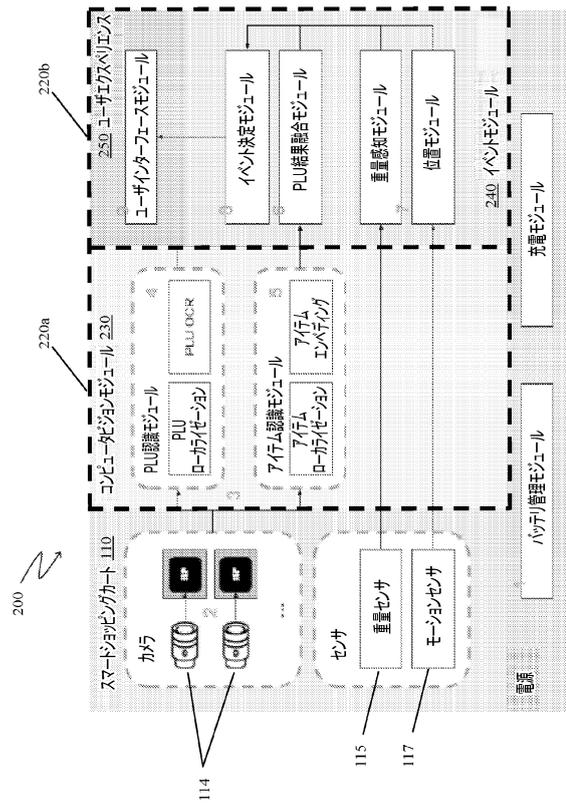


FIG. 4

【 図 5 】

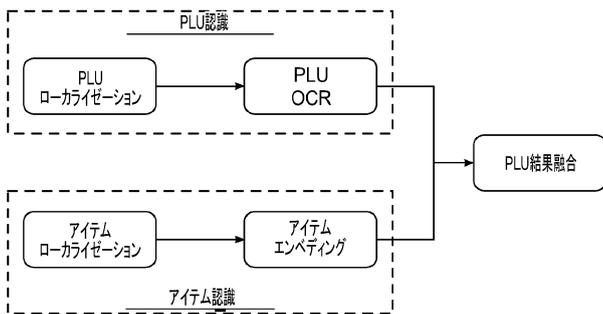


FIG. 5

【 図 6 】

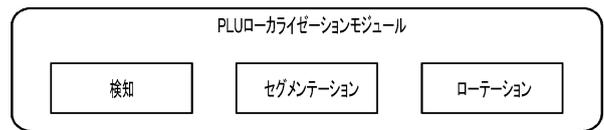


FIG. 6

10

20

30

40

50

【 図 7 】

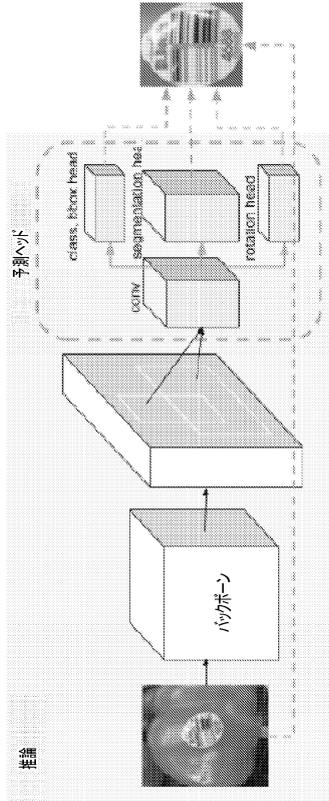


FIG. 7

【 図 8 】

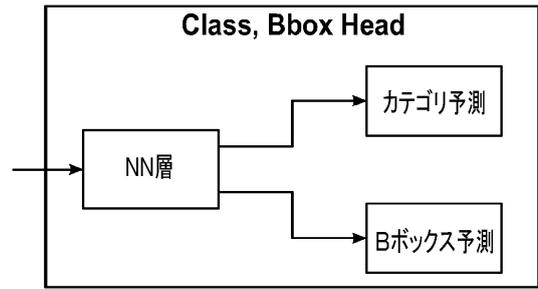


FIG. 8

10

20

【 図 9 】

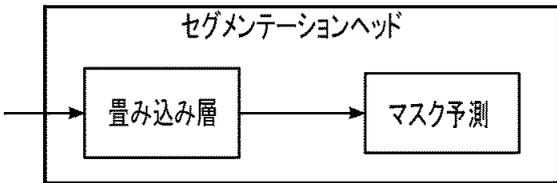


FIG. 9

【 図 10 】

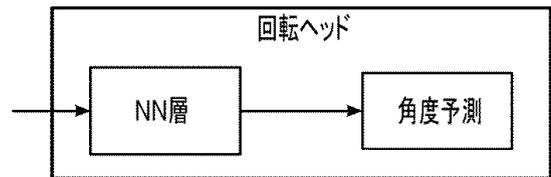


FIG. 10

30

40

50

【 図 1 1 】

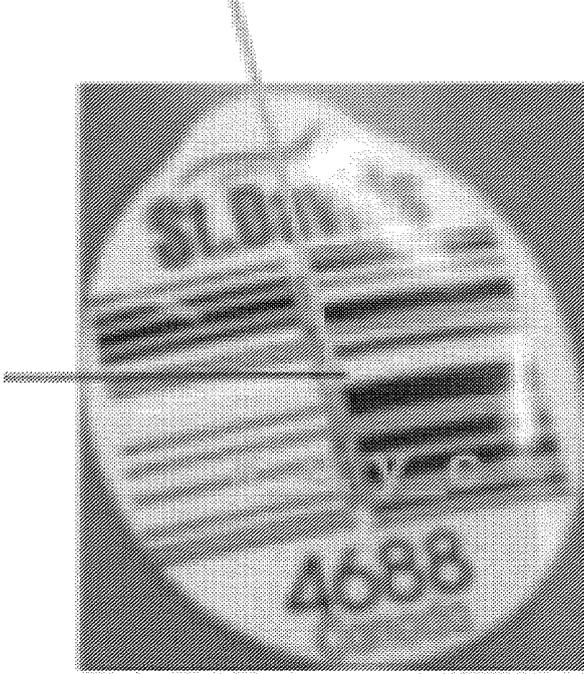


FIG. 11

【 図 1 2 】

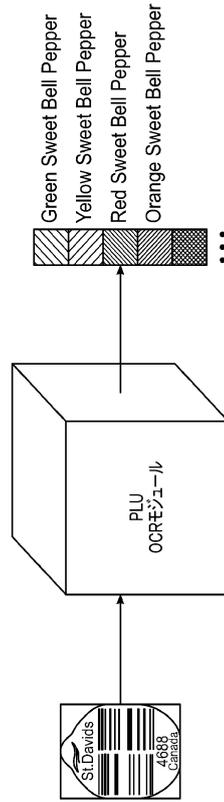


FIG. 12

10

20

【 図 1 3 】

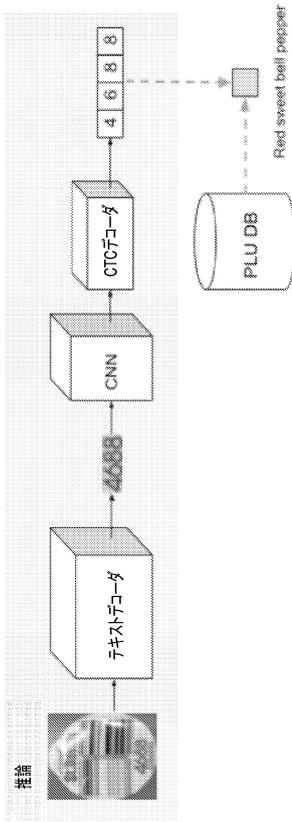


FIG. 13

【 図 1 4 】

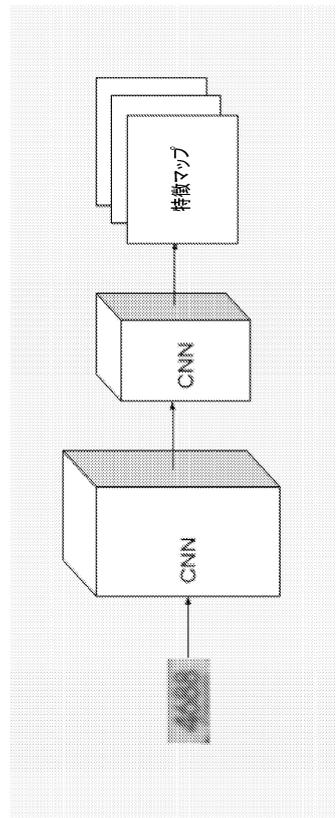


FIG. 14

30

40

50

【 図 1 5 】

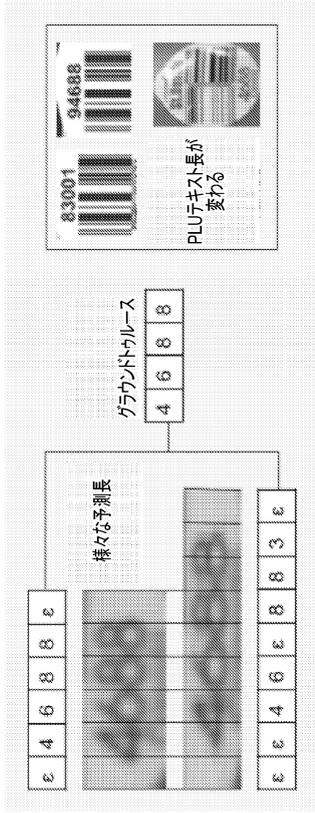


FIG. 15

【 図 1 6 】

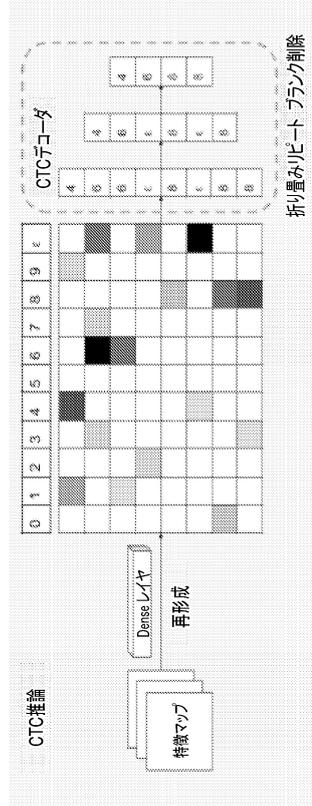


FIG. 16

【 図 1 7 】

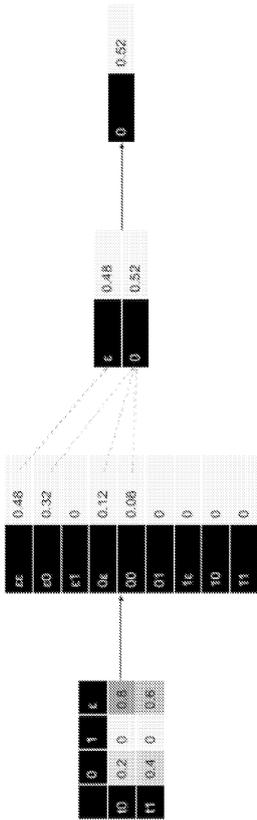


FIG. 17

【 図 1 8 】

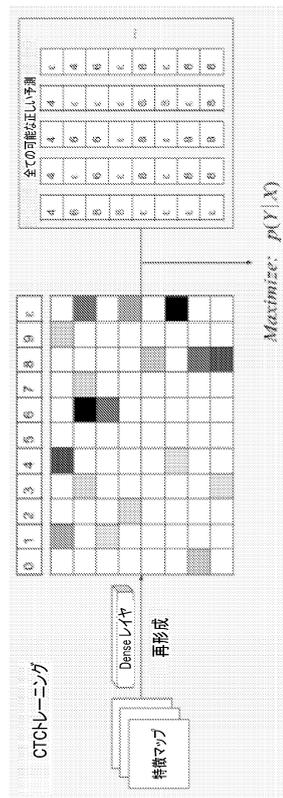


FIG. 18

【 図 19 】

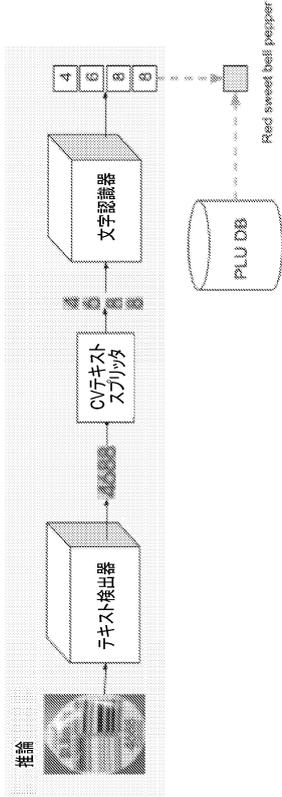


FIG. 19

【 図 20 】

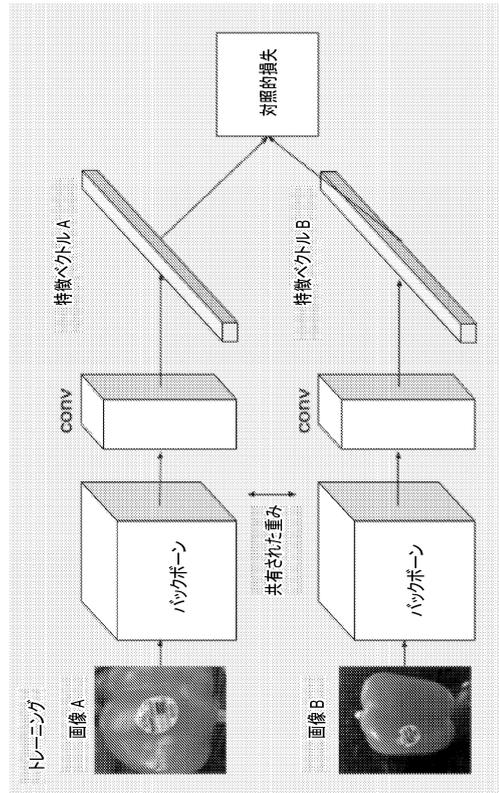


FIG. 20

【 図 21 】

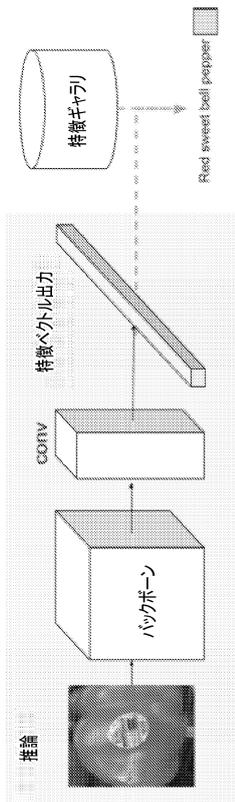


FIG. 21

【 図 22 】

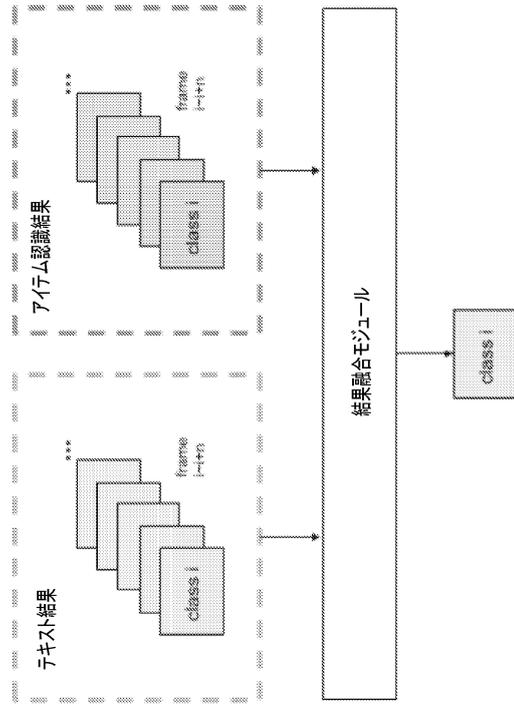


FIG. 22

10

20

30

40

50

【 図 2 3 】

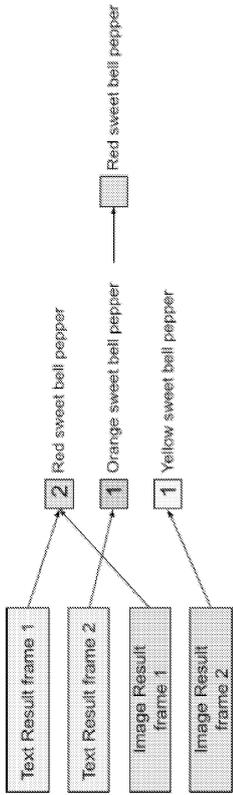


FIG. 23

【 図 2 4 】

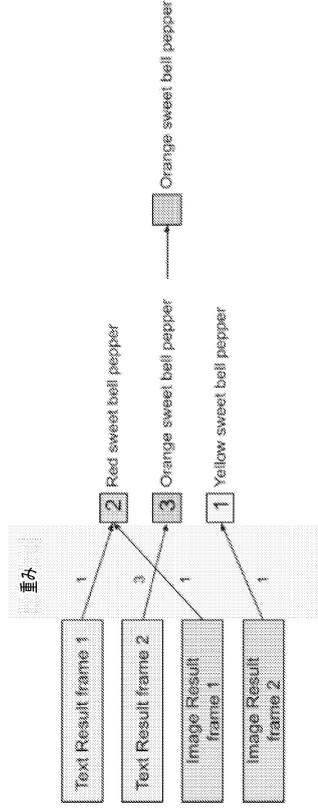


FIG. 24

【 図 2 5 】

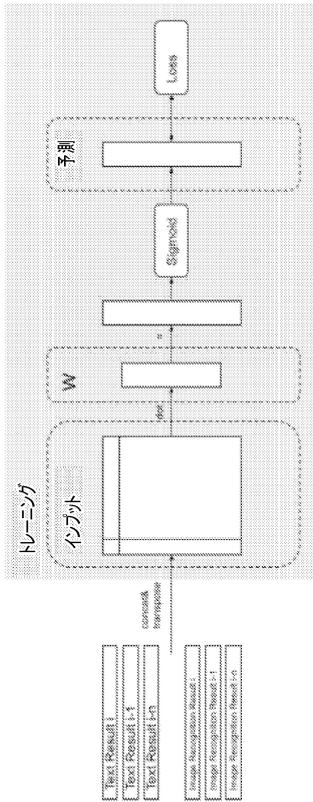


FIG. 25

【 図 2 6 】

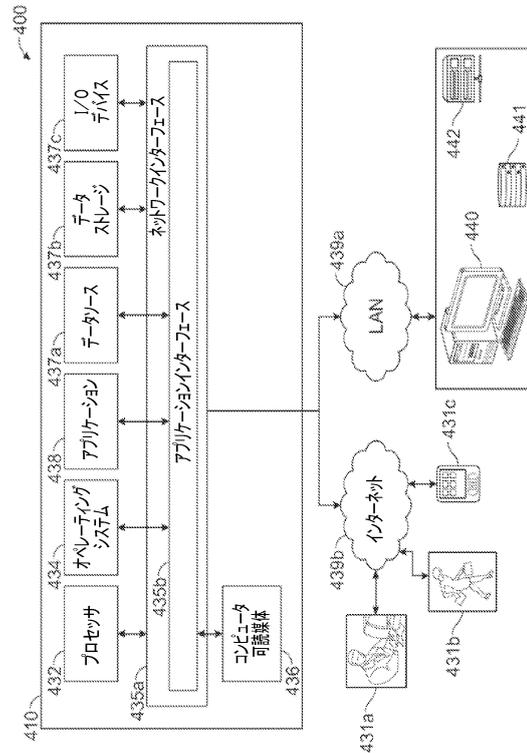


FIG. 26

10

20

30

40

50

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US22/44340

A. CLASSIFICATION OF SUBJECT MATTER		
IPC - INV. G06Q 20/20; B62B 3/14; B62B 5/00; G01G 19/41; G06Q 30/06 (2022.01) ADD. G06K 19/06; G06K 19/07; G06V 10/82; G07G 1/00 (2022.01)		
CPC - INV. G06Q 20/208; B62B 3/14; B62B 5/0096; G01G 19/4144; G06Q 20/201; G06Q 20/203; G06Q 30/0633 ADD. B62B 2203/50; G06K 19/06037; G06K 19/0723; G06T 2207/20081; G06T 2207/20084; G06V 10/454; G06V 10/82		
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) See Search History document		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched See Search History document		
Electronic database consulted during the international search (name of database and, where practicable, search terms used) See Search History document		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2020/0151692 A1 (SBOT TECHNOLOGIES, INC. D/B/A CAPER INC.) 14 May 2020; figures 1-2; paragraphs [0049], [0058]-[0059]	1-2, 5-12, 30 ---
Y		3-4 ---
A		13-14
Y	CN 109741552 B (CHINA UNITED NETWORK COMMUNICATION GROUP CO LTD) 15 June 2021; see machine translation: abstract; page 7, lines 10-15	3-4
A	US 2005/0178840 A1 (CHANG KUANG YUAN) 18 August 2005; paragraphs [0024]-[0025]	1-14, 30
A	US 2018/0218351 A1 (FOCAL SYSTEMS, INC.) 02 August 2018; entire document	1-14, 30
A	US 10,953,906 B1 (AMAZON TECHNOLOGIES, INC.) 23 March 2021; entire document	1-14, 30
<input type="checkbox"/> Further documents are listed in the continuation of Box C.		<input type="checkbox"/> See patent family annex.
<p>* Special categories of cited documents:</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"D" document cited by the applicant in the international application</p> <p>"E" earlier application or patent but published on or after the international filing date</p> <p>"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)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> <p>"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</p> <p>"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</p> <p>"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</p> <p>"&" document member of the same patent family</p>		
Date of the actual completion of the international search 05 December 2022 (05.12.2022)		Date of mailing of the international search report FEB 27 2023
Name and mailing address of the ISA/ Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-8300		Authorized officer Shane Thomas Telephone No. PCT Helpdesk: 571-272-4300

10

20

30

40

50

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US22/44340

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

- 1. Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely: 10

- 2. Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

- 3. Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a). 20

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
Please See Supplemental Page-

- 1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims. 30
- 2. As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
- 3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
- 4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
Group I: Claims 1-14, 30 40

- Remark on Protest**
- The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
 - The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
 - No protest accompanied the payment of additional search fees.

10

20

30

40

50

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US22/44340

---Continued From Box No. III: Observations where unity of invention is lacking---

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I: Claims 1-14, 30 are directed toward a self-checkout vehicle system, comprising: at least one weight sensor.

10

Group II: Claims 15-29 are directed toward a computer vision system for identifying merchandise, comprising: identify, based on at least one of the image(s), an item category of the selected merchandise; and compare the identifier and item category identified in each respective image to determine the most likely identification of the merchandise.

The inventions listed as Groups I-II do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Group I include a self-checkout vehicle system, comprising: a self-checkout vehicle having: at least one weight sensor mounted in the self-checkout vehicle and configured to measure a weight of a merchandise item selected by a shopper operating the self-checkout vehicle when the selected merchandise item is placed in the self-checkout vehicle; and one or more cameras positioned and oriented on the self-checkout vehicle such that its field of view comprises at least one of the weight sensor(s), wherein the one or more cameras are configured to capture one or more images of the selected merchandise item on the at least one weight sensor, which are not present in Group II.

Group II include a computer vision system for identifying merchandise, comprising: identify, based on at least one of the image(s), an item category of the selected merchandise; and compare the identifier and item category identified in each respective image to determine the most likely identification of the merchandise, which are not present in Group I.

20

The common technical features of Groups I-II are at least one processor adapted to obtain computer-executable instructions stored on a non-transitory medium that, when executed by the at least one processor, cause the at least one processor to: identify, in one or more images of merchandise selected by a shopper, an identifier affixed to the selected merchandise using a computer vision technique.

The common technical features of Groups I-II are disclosed by US 2020/0151692 A1 (SBOT). Sbot discloses at least one processor adapted to obtain computer-executable instructions stored on a non-transitory medium that, when executed by the at least one processor, cause the at least one processor to (processor 104 configured to execute instructions stored in memory 116; paragraph [0042]): identify, in one or more images of merchandise selected by a shopper, an identifier affixed to the selected merchandise using a computer vision technique (the one or more cameras 204 obtain images to identify the merchandise identifier such as a barcode to determine the merchandise item that is placed into the cart 102 using computer vision; paragraphs [0058]-[0059]).

Since the common technical features are previously disclosed by the Sbot reference, the common features are not special and so Groups I and II lack unity.

30

40

フロントページの続き

(51)国際特許分類

G 0 6 V 20/60 (2022.01)
G 0 7 G 1/00 (2006.01)

F I

G 0 6 T 7/00 3 0 0 F
G 0 6 V 20/60
G 0 7 G 1/00 3 3 1 C
G 0 7 G 1/00 3 1 1 D
G 0 7 G 1/00 3 1 1 E

テーマコード (参考)

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,N
E,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,
CV,CZ,DE,DJ,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IQ,IR,IS,IT,J
M,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,ST,SV,SY,T
H,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,WS,ZA,ZM,ZW

(72)発明者

ユーフォン ホー

中華人民共和国 2 0 1 8 2 4 シャンハイ チャンニン ディストリクト トンシエ ロード 2 8 8
1 ナンバー 4 1 5

(72)発明者

シアオ ジョウ

中華人民共和国 2 0 1 8 2 4 シャンハイ チャンニン ディストリクト トンシエ ロード 2 8 8
1 ナンバー 4 1 5

(72)発明者

イーリン ホアン

中華人民共和国 2 0 1 8 2 4 シャンハイ チャンニン ディストリクト トンシエ ロード 2 8 8
1 ナンバー 4 1 5

(72)発明者

グリフィン ケリー

アメリカ合衆国 6 0 6 1 4 イリノイ州 シカゴ ウェスト ドラモンド プレイス 6 4 2

(72)発明者

イサベル サイ

アメリカ合衆国 1 0 0 1 0 ニューヨーク州 ニューヨーク ピーター クーパー ロード 8 アパ
ートメント 2 エフ

(72)発明者

アフマド ベシュリー

アメリカ合衆国 1 1 1 0 9 ニューヨーク州 クイーンズ センター ブールバード 4 5 4 5 ユニ
ット 2 9 0 5

F ターム (参考)

3E142 AA01 BA01 EA04 FA41 GA02 GA22 GA32 GA41 JA01
5L030 BB21
5L096 AA06 BA17 CA02 DA01 EA03 EA05 EA39 FA16 FA67 FA69
GA10 GA30 GA51 HA11 JA03 JA11 KA04