

(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2016-536837

(P2016-536837A)

(43) 公表日 平成28年11月24日 (2016. 11. 24)

(51) Int. Cl.	F I	テーマコード (参考)
H04N 1/387 (2006.01)	H04N 1/387	5B057
G06T 5/00 (2006.01)	G06T 5/00 725	5C076

審査請求 未請求 予備審査請求 有 (全 48 頁)

(21) 出願番号 特願2016-517415 (P2016-517415) (86) (22) 出願日 平成26年9月23日 (2014. 9. 23) (85) 翻訳文提出日 平成28年5月24日 (2016. 5. 24) (86) 国際出願番号 PCT/US2014/057065 (87) 国際公開番号 W02015/048045 (87) 国際公開日 平成27年4月2日 (2015. 4. 2) (31) 優先権主張番号 61/883, 865 (32) 優先日 平成25年9月27日 (2013. 9. 27) (33) 優先権主張国 米国 (US) (31) 優先権主張番号 14/491, 901 (32) 優先日 平成26年9月19日 (2014. 9. 19) (33) 優先権主張国 米国 (US)	(71) 出願人 505149170 コファックス, インコーポレイテッド アメリカ合衆国 カリフォルニア 926 18, アーバイン, ラグーナ キャニ オン ロード 15211 (74) 代理人 110001195 特許業務法人深見特許事務所 (72) 発明者 マッチオーラ, アンソニー アメリカ合衆国、91709 カリフォル ニア州、チノ・ヒルズ、グレン・リッジ・ ドライブ、15283 (72) 発明者 マ, ジーヨン アメリカ合衆国、92127 カリフォル ニア州、サン・ディエゴ、ブルー・グラニ ット・ドライブ、10507 最終頁に続く
---	--

(54) 【発明の名称】 取込んだ画像データを3次元で幾何学的に再構成するためのシステムおよび方法

(57) 【要約】

さまざまな実施形態において、モバイル装置によって取込まれたデジタル画像を処理するための方法、システム、およびコンピュータプログラムプロダクトが開示される。非常に多くの特徴が、そうでなければ技術的に不可能または非現実的な、モバイル装置を用いたこのようなデジタル画像の処理を、可能におよび/または容易にするとともに、従来のフラットベッドスキャナ、給紙スキャナ、または多機能周辺装置ではなくカメラを用いて取込まれた画像が示す固有の課題に対処する。特に、ここに開示されるシステムおよび技術は、モバイル装置のカメラを用いて取込まれた画像に示されているオブジェクトの三次元再構成を可能にする。再構成により、カメラによる取込みによって生じる透視歪みを補正または補償する。

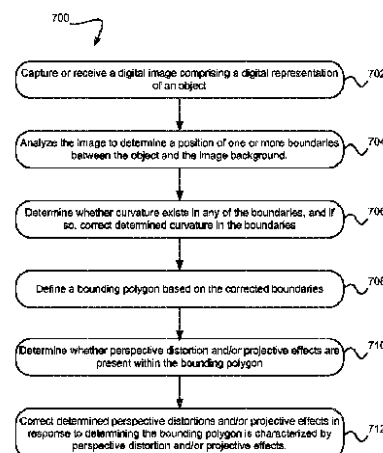


FIG. 7

【特許請求の範囲】

【請求項 1】

デジタル画像を再構成する方法であって、前記方法は、
四角形で囲まれたオブジェクトのデジタル表現を含むデジタル画像を受けるステップと

、
前記四角形の湾曲を補正して四辺形を形成するステップと、
前記四辺形における射影効果を補正して矩形を形成するステップとを含む、方法。

【請求項 2】

前記四辺形は平行四辺形である、請求項 1 に記載の方法。

【請求項 3】

前記湾曲を補正するステップは、前記四角形の 1 つ以上の部分を、固有の四角形座標に少なくとも一部基づいて変換するステップを含む、請求項 1 に記載の方法。

【請求項 4】

前記射影効果を補正するステップは、4 点アルゴリズムを用いて前記四角形を変換するステップを含む、請求項 1 に記載の方法。

【請求項 5】

前記デジタル画像は、前記オブジェクトに対する法線から約 30 度以上ずれた取込み角度によって特徴付けられる、請求項 1 に記載の方法。

【請求項 6】

前記射影効果を補正するステップは特異値分解 (SVD) を含む、請求項 1 に記載の方法。

【請求項 7】

前記射影効果を補正するステップは直接線形変換 (DLT) を含む、請求項 1 に記載の方法。

【請求項 8】

前記射影効果を補正するステップは、ホモグラフィ変換 H を推定するステップを含む、請求項 1 に記載の方法。

【請求項 9】

前記 H を推定するステップは、1 つ以上の点対応

【数 1】

$p_i \leftrightarrow P'_i$ を検出するステップを含み、 $p_i = (x_i, y_i, 1)^T$ である、

請求項 8 に記載の方法。

【請求項 10】

前記 H を推定するステップは、4 つの点対応を検出するステップを含み、

【数 2】

各点対応 $p_i \leftrightarrow P'_i$ は、前記四角形の角位置 p_i と前記矩形の対応する角位置 P'_i に
対応する、

請求項 9 に記載の方法。

【請求項 11】

H は、式 $p_i = H P_i$ を満たし、
= 三次元基準座標位置 (X, Y, Z) と、取込み装置の対応する基準座標位置との間の
焦点距離であり、

$p_i = (x_i, y_i, 1)^T$ であり、

$P_i = (X_i, Y_i, 1)^T$ であり、

i は、対応する画像における点 p または P のそれぞれの指数特徴である、請求項 8 に
記載の方法。

【請求項 1 2】

前記射影効果を補正するステップは、非線形リファインメントを実行するステップを含む、請求項 1 に記載の方法。

【請求項 1 3】

再構成されたデジタル画像を出力するステップをさらに含み、

前記再構成されたデジタル画像は、前記オブジェクトに対して垂直である取込み角度を用いて取込まれた前記オブジェクトのスキャンされた 5 0 0 D P I 画像と比較して約 5 画素以下の画素位置誤差によって特徴付けられる、請求項 1 に記載の方法。

【請求項 1 4】

前記四角形の少なくとも一边は、二次以上の多項式によって特徴付けられる、請求項 1 に記載の方法。

10

【請求項 1 5】

前記四角形の各辺は、二次以上の多項式によって特徴付けられる、請求項 1 4 に記載の方法。

【請求項 1 6】

前記射影効果を補正するステップは、前記デジタル画像と再構成された画像との点対応の座標を正規化するステップを含む、請求項 1 に記載の方法。

【請求項 1 7】

前記オブジェクトは、既知の高さと幅の比率によって特徴付けられる文書である、請求項 1 に記載の方法。

20

【請求項 1 8】

システムであって、

ロジックを実行するように構成されたプロセッサと、

四角形で囲まれたオブジェクトのデジタル表現を含むデジタル画像を受けるように構成されたロジックと、

前記四角形の湾曲を補正して四辺形を形成するように構成されたロジックと、

前記四辺形における射影効果を補正して矩形を形成するように構成されたロジックとを含む、システム。

【請求項 1 9】

コンピュータ読取可能なプログラムコードが格納されたコンピュータ読取可能な記録媒体を含むコンピュータプログラムプロダクトであって、前記コンピュータ読取可能なプログラムコードは、

30

四角形で囲まれたオブジェクトのデジタル表現を含むデジタル画像を受けるように構成されたコンピュータ読取可能なプログラムコードと、

前記四角形の湾曲を補正して四辺形を形成するように構成されたコンピュータ読取可能なプログラムコードと、

前記四辺形における射影効果を補正して矩形を形成するように構成されたコンピュータ読取可能なプログラムコードとを含む、コンピュータプログラムプロダクト。

【発明の詳細な説明】**【技術分野】**

40

【0 0 0 1】

発明の分野

本発明は、デジタル画像データの取込みおよび処理に関し、より具体的には、カメラを用いた画像の取込みに固有の歪み効果を原因とする画像アーティファクトの修正に関する。

【背景技術】**【0 0 0 2】**

発明の背景

書簡、勘定書、請求書、納品伝票等の文書が示されているデジタル画像は、従来、ラップトップまたはデスクトップコンピュータ等のコンピュータワークステーションに接続さ

50

れたスキャナまたは多機能周辺装置を用いて取込まれて処理されてきた。このような取込みおよび処理を実行することが可能な方法およびシステムは、当該技術では周知であり、これら方法およびシステムが採用されるタスクに十分適応している。

【0003】

しかしながら、日々の活動、計算、および仕事をモバイル装置を用いて行なうことが増えている時代においては、スマートフォン、デジタルカメラ、タブレットコンピュータ等のモバイルプラットフォームに導入して使用するための、同様の文書取込みおよび処理システムを提供することが、非常に有益であろう。

【0004】

従来の文書取込みおよび処理技術からの移行における主な課題は、モバイル装置内の現在利用できるハードウェアを用いたときに得られる処理能力および画像解像度には限界があることである。これらの限界は重要な課題を示している。なぜなら、一般的には従来のスキャナで得られる解像度よりも遥かに低い解像度で取込まれた画像を処理することは、不可能または非現実的であるからである。結果として、従来のスキャナに基づく処理アルゴリズムは一般的に、モバイル装置を用いて取込まれたデジタル画像に対しては上手く機能しない。

【0005】

加えて、モバイル装置上で利用できる処理およびメモリには限界があるので、スキャナに対して採用される従来の画像処理アルゴリズムは、計算コストが極端に高くなる。従来のスキャナに基づく画像処理アルゴリズムを処理しようとすると、時間がかかりすぎて、現代のモバイルプラットフォーム上のアプリケーションとしては非実用的である。

【0006】

さらに他の課題が、モバイル取込み部品（たとえば携帯電話、タブレット等のカメラ）の性質によって示される。従来のスキャナは、デジタル画像内の物理的文書を忠実に表現し、このデジタル画像内の物理的文書のアスペクト比、寸法、および形状を正確に維持することが可能であるが、モバイル取込み部品ではこのような結果をもたらすことができないことが多い。

【0007】

具体的には、カメラによって取込まれた文書の画像は、スキャナによって取込まれた画像を扱うときには生じない、新たな種類の処理上の問題を示す。その原因の一部は、文書画像の取得方法にも装置の構成方法にも、本質的な相違があることにある。一部のスキャナは、紙と直線的に配列されたセンサとの間の相対運動を生じさせる搬送機構を使用して機能する。これらセンサは、文書を通過するときこの文書の画素値を生成し、取込まれた一連のこれら画素値が1つの画像を形成する。したがって、一般的にはセンサ自体のノイズまでは水平または垂直方向の一貫性があり、その列のすべての画素を提供するのは同じセンサである。

【0008】

これに対し、カメラには、たとえば典型的には矩形に配置された、非直線的な配列のより多くのセンサがある。このため、これら個々のセンサはすべて独立しており、典型的には水平方向にも垂直方向にも一貫性がない画像データを提供する。加えて、カメラは、撮像角度の関数である射影効果を導入する。たとえば、スキャナのように直線配列の場合は、紙の搬送がセンサの並びに対して完全に垂直でなくいくらかのスキューが導入されたとしても、カメラのような射影効果は生じない。加えて、カメラによる取込みの場合は、カメラの光学部品が原因で非線形歪みが導入されることがある。

【0009】

上記課題に鑑み、モバイル装置を用いた画像の取込みおよび処理に関連する問題を排除くおよび/または修正する一方で、効率的な処理方法によって計算コストを低く保つことが可能な、画像取込みおよび処理アルゴリズムならびにそのアプリケーションを提供することが有益であろう。

【発明の概要】

10

20

30

40

50

【課題を解決するための手段】

【0010】

発明の概要

ある実施形態において、デジタル画像を再構成する方法は、四角形で囲まれたオブジェクトのデジタル表現を含むデジタル画像を受けるステップと、この四角形の湾曲を補正して四辺形を形成するステップと、この四辺形における射影効果を補正して矩形を形成するステップとを含む。

【0011】

システムは、ロジックを実行するように構成されたプロセッサと、四角形で囲まれたオブジェクトのデジタル表現を含むデジタル画像を受けるように構成されたロジックと、この四角形の湾曲を補正して四辺形を形成するように構成されたロジックと、この四辺形における射影効果を補正して矩形を形成するように構成されたロジックとを含む。

10

【0012】

コンピュータプログラムプロダクトは、コンピュータ読取可能なプログラムコードが格納されたコンピュータ読取可能な記録媒体を含む。コンピュータ読取可能なプログラムコードは、四角形で囲まれたオブジェクトのデジタル表現を含むデジタル画像を受けるように構成されたコンピュータ読取可能なプログラムコードと、この四角形の湾曲を補正して四辺形を形成するように構成されたコンピュータ読取可能なプログラムコードと、この四辺形における射影効果を補正して矩形を形成するように構成されたコンピュータ読取可能なプログラムコードとを含む。

20

【図面の簡単な説明】

【0013】

【図1】ある実施形態に従うネットワークアーキテクチャを示す。

【図2】ある実施形態に従う、図1のサーバおよび/またはクライアントに関連付けることができる代表的なハードウェア環境を示す。

【図3A】ある実施形態に従う、文書のデジタル表現を含むデジタル画像の概略図である。

【図3B】ある実施形態に従う、文書のデジタル表現および複数のページ検出解析ウィンドウを含むデジタル画像の概略図である。

【図3C】ある実施形態に従う、複数の候補エッジポイントによって特徴付けられる文書のデジタル表現を含むデジタル画像の概略図である。

30

【図3D】ある実施形態に従う、デジタル画像の複数の画素を含む大解析ウィンドウとこの大解析ウィンドウの中の小解析ウィンドウの概略図である。

【図4】ある実施形態に従う、目標の四角形によって囲まれた文書のデジタル表現を含むデジタル画像の概略図である。

【図5A】ある実施形態に従う、ページ矩形化アルゴリズムの一回目の繰返しの図形表現である。

【図5B】ある実施形態に従う、ページ矩形化アルゴリズムに対する入力 of 図形表現である。

【図6A】ある実施形態に従う、取込み角度を測定するための座標系を示す、単純化した模式図である。

40

【図6B】ある実施形態に従う、オブジェクトに対して垂直な取込み角度を用いて取込まれた矩形のオブジェクトの典型的な模式図を示す。

【図6C】ある実施形態に従う、オブジェクトに対してわずかに傾斜した取込み角度を用いて取込まれた矩形のオブジェクトの典型的な模式図を示す。

【図6D】ある実施形態に従う、オブジェクトに対して大きく傾斜した取込み角度を用いて取込まれた矩形のオブジェクトの典型的な模式図を示す。

【図7】ある実施形態に従う方法のフローチャートである。

【発明を実施するための形態】

【0014】

50

詳細な説明

以下の説明は、本発明の一般的な原理を例示することを目的となされ、本明細書においてクレームされる発明の概念を限定することを意図しているのではない。さらに、本明細書に記載の特定の特徴は、可能なさまざまな組合せおよび置換え各々において、記載されている他の特徴と組合わせて使用することができる。

【0015】

本明細書において特に指定しない限り、すべての用語は、明細書において示される意味および当業者によって理解されるおよび／または辞書、論文等において定義される意味を含む、可能な最も広い解釈が与えられるべきものである。

【0016】

また、本明細書およびそれに続く請求項において使用される単数形「a」、「an」および「the」は、特に指定しない限り複数のものを含むことに注意しなければならない。

【0017】

本願は画像処理に関連する。特に、本願は、デジタル画像に示されたオブジェクトを非矩形形状から実質的に矩形の形状、または好ましくは矩形の形状に変換するように構成された、システム、方法、およびコンピュータプログラムプロダクトを開示する。さらに好ましくは、これは、最初にオブジェクトの湾曲を補正し次に画像の射影効果を補正する二段階のプロセスを採用することによって実現される。

【0018】

以下の定義は、さまざまな実施形態に従う本明細書に記載の発明の概念の理解に役立つであろう。以下の定義は、典型的なものとみなされるべきもので、この開示をより明確にするための説明を目的として提供されるが、本明細書に開示される発明の概念の範囲を限定するものとみなされてはならない。

【0019】

以下における「四辺形」は、(1)各辺が直線であり、(2)隣合う辺がこれら辺の交点で頂点を形成する、四辺の図形である。例示としての2つの実施形態に従う、典型的な四辺形が、以下の図6Cおよび図6Dに示される。

【0020】

「平行四辺形」は、特殊な種類の四辺形である、すなわち、(1)各辺が直線であり、(2)対向する辺が平行で、(3)隣合う辺が必ずしも垂直でなくしたがって隣合う辺の交点における頂点がなす角度の値が必ずしも90度でない、四辺の図形である。

【0021】

「矩形」または「矩形形状」は、特殊な種類の四辺形であり、これは、(1)各辺が直線であり、(2)対向する辺が平行で、(3)隣合う辺が垂直でありしたがって隣合う辺で構成される辺の対各々において、これら隣合う辺が頂点でなす内角が直角すなわち90度の角度である、四辺の図形であると定義される。例示としてのある実施形態に従う典型的な矩形が図6Bに示される。

【0022】

加えて、本明細書における「矩形」および「矩形形状」は、「実質的に矩形の形状」を含むとみなされる。「実質的に矩形の形状」は、(1)各辺の大部分が直線であり(たとえば、さまざまな実施形態において、各辺の長さの少なくとも90%、95%、または99%が($y = mx + b$ 等の)一次多項式で特徴付けられる)、(2)隣合う辺で構成される各対が値 θ の内角をなし、 θ が、(a)隣合う2つの辺の間の頂点、(b)1つの辺の、大部分が直線である部分の延長線と、その隣の辺との間の頂点、または(c)1つの辺の、大部分が直線である部分の延長線と、その隣の辺の、大部分が直線である部分の延長線との間の頂点、のいずれかにおいて、およそ90°である(たとえば θ が $85^\circ < \theta < 95^\circ$ という関係を満たす)、四辺の形状として定義される。典型的な「実質的に矩形の形状」が以下の図7に示されている(なお、示されている運転免許証の角は湾曲しているので、隣合う辺で形成される独立した頂点はないが、隣接する辺の延長線がそれら延長

10

20

30

40

50

線の交点でおおよそ90°である角度を有する頂点をなす)。

【0023】

本明細書における「非矩形形状」は、上記のように定義される「矩形」でも「実質的に矩形形状」でもない、いかなる形状も含む。好ましい実施形態において、「非矩形形状」は、本明細書において四辺の図形と呼ばれ、(1)各辺の全体または一部が、選ばれた関数のクラスから選択された(たとえば、好ましくはゼロ次から五次の範囲の多項式、より好ましくは一次から三次の多項式、さらに好ましくは一次から二次の多項式のクラスから選択された)等式によって特徴付けられ、かつ(2)この図形の隣合う辺がこれら辺の交点で頂点をなす、「四角形」である。ある実施形態に従う、本明細書における代表的な四角形が図4に示されている。

10

【0024】

本明細書に開示される発明の概念の、ある一般的な実施形態において、デジタル画像を再構成する方法は、四角形で囲まれたオブジェクトのデジタル表現を含むデジタル画像を受けるステップと、この四角形の湾曲を補正して四辺形を形成するステップと、この四辺形における射影効果を補正して矩形を形成するステップとを含む。

【0025】

別の一般的な実施形態において、システムは、ロジックを実行するように構成されたプロセッサと、四角形で囲まれたオブジェクトのデジタル表現を含むデジタル画像を受けるように構成されたロジックと、この四角形の湾曲を補正して四辺形を形成するように構成されたロジックと、この四辺形における射影効果を補正して矩形を形成するように構成されたロジックとを含む。

20

【0026】

他の一般的な実施形態において、コンピュータプログラムプロダクトは、コンピュータ読取可能なプログラムコードが格納されたコンピュータ読取可能な記録媒体を含む。コンピュータ読取可能なプログラムコードは、四角形で囲まれたオブジェクトのデジタル表現を含むデジタル画像を受けるように構成されたコンピュータ読取可能なプログラムコードと、この四角形の湾曲を補正して四辺形を形成するように構成されたコンピュータ読取可能なプログラムコードと、この四辺形における射影効果を補正して矩形を形成するように構成されたコンピュータ読取可能なプログラムコードとを含む。

【0027】

米国特許出願第13/740,127号(2013年1月11に出願)に記載されているような以前の矩形化方法は、カメラで取込まれた画像において観察される湾曲効果を上手く補正する。しかしながら、カメラのピッチおよび/またはロールが十分に大きい(たとえば約30度以上)ときは、画像の中の射影効果/アーティファクトが非常に大きいので、この技術の射影効果補正能力は不十分であることが多い。

30

【0028】

以前に説明されている湾曲補正方法と本明細書に記載の射影効果補正方法を組合わせて、2つの目的を兼ねた1つの手順にすると、これら手法それぞれの強みが効果的に組合わされることになる。ある実施形態において、これら手法は次のように組合わせてもよい。第1に、湾曲補正要素が、オブジェクトを取囲む湾曲した四角形を、角は同一であるが辺が直線である(すなわち線形または一次多項式によって特徴付けられる)四角形にマッピングすることにより、湾曲を補正する。第2に、この(辺が直線になった)四角形を、以下で説明するような4ポイント法を用いて、目標とする矩形にマッピングする。

40

【0029】

重要な利点として、この二重目的の手順は、矩形表現をより正確に再構成するが、座標の2段マッピングにしか作用せず、その一方で、画像の実際の変換が一度しか生じないことが、挙げられる。処理時間のほとんどが、メモリにおいて2つの大きな画像を操作することと、何らかの方法で正しい座標を計算するのではなく非整数(x, y)座標を囲む4つの画素を組合わせることに費やされるので、この二重目的の方法は、従来のフラットベッドスキャナまたは同様の装置から得られる対応する画像と比較して再構成画像における

50

相対的な画素位置の精度が改善されているにもかかわらず、計算コストはたがって実行時間が、元来の方法よりもわずかに約 3 % しか高くない。

【 0 0 3 0 】

実験のテストは、湾曲が著しくかつ射影歪みが大きい写真において、上記二重目的の方法が、座標誤差（矩形化された画素から、同一解像度の同一文書のスキャンされた画像における同一画素までの最大距離）を、座標ベースの方法のみの場合の誤差と比較して、約 3 分の 1 にすることを、示している。ある実施形態において、残留誤差は、分解能 5 0 0 D P I で約 5 画素、または約 1 0 0 分の 1 インチであった。

【 0 0 3 1 】

画像（たとえば絵、数字、略図、映画、ビデオ、フィルム、クリップの 1 コマ等）は、好ましくはカメラ、特にモバイル装置のカメラによって取込まれたデジタル画像である。本明細書では、モバイル装置が、物理的な接続（たとえばワイヤ、コード、ケーブル等）を介した電力供給なしでデータを受信することが可能で、かつ、物理的なデータ接続（たとえばワイヤ、コード、ケーブル等）なしでデータを受信することが可能な任意の装置であることが理解される。本開示の範囲に含まれるモバイル装置は、携帯電話、スマートフォン、タブレット、携帯情報端末、i P o d（登録商標）、i P a d（登録商標）、B L A C K B E R R Y（登録商標）装置等の代表的な装置を含む。

【 0 0 3 2 】

しかしながら、さまざまな機能の説明から明らかになるように、ここで開示されているモバイル画像処理アルゴリズムは、時には何らかの修正が施されて、スキャナおよび多機能周辺装置（M F P）からの画像に適用することができる。同様に、いくつかのアプローチでは、ここで開示されている処理アルゴリズムを用いて処理された画像を、従来のスキャナ処理アルゴリズムを用いてさらに処理してもよい。

【 0 0 3 3 】

当然、本明細書に記載のさまざまな実施形態は、ハードウェア、ソフトウェア、またはその所望の組み合わせを用いて実現し得る。そのために、本明細書に記載のさまざまな機能を実現することが可能な何らかの種類のロジックを使用してもよい。

【 0 0 3 4 】

モバイル装置の使用における 1 つの利点は、データプランを用いて、取込まれた画像に基づく画像処理および情報処理を、スキャナの存在に頼る以前の方法よりも、一層便利で能率的で統合されたやり方で、行なうことができる点である。しかしながら、モバイル装置を文書取込みおよび / または処理装置として使用することは、さまざまな理由から今までは不可能であると考えられてきた。

【 0 0 3 5 】

あるアプローチでは、画像をモバイル装置のカメラによって取込んでもよい。「カメラ」という用語は、一枚の紙等の装置の外部にある物理的なオブジェクトの画像を取込むことができるいずれの種類の装置も含むものとして広く解釈されねばならない。「カメラ」という用語は、周辺装置としてのスキャナまたは多機能装置は包含しない。どのような種類のカメラを使用してもよい。好ましい実施形態は、解像度が高い、たとえば 8 M P 以上、理想的には 1 2 M P 以上のカメラを使用することがある。画像の取込みは、カラーで、無彩色で、白黒で、またはそれ以外の光学効果を利用して行なってもよい。本明細書に記載の「画像」という用語は、生のデータ、処理済みのデータ等を含む、カメラの出力に相当するどの種類のデータも包含することを意図している。

【 0 0 3 6 】

本明細書の記載は、当業者が本発明を実施し使用することを可能にするために示され、本発明の特定の用途およびその必要条件という文脈においてなされる。開示されている実施形態のさまざまな変形は、当業者には直ちに明らかになるであろう。また、本明細書において定義される一般的な原理は、本発明の精神および範囲から外れることなく他の実施形態および用途に応用し得る。このように、本発明は、示されている実施形態に限定されることを意図しているのではなく、本明細書に開示される原理および特徴に従う最も広い

10

20

30

40

50

範囲に一致するものである。

【 0 0 3 7 】

特に、本明細書に記載の発明のさまざまな実施形態は、複数のコンピュータシステム間の通信手段としてインターネットを用いて実現される。本発明が通信媒体としてのインターネットの使用に限定されないこと、および、本発明の代替方法が、私用イントラネット、ローカルエリアネットワーク（LAN）、ワイドエリアネットワーク（WAN）、またはそれ以外の通信手段の使用に適応し得ることを、当業者は理解するであろう。加えて、有線、無線（たとえば無線周波数）および光通信リンクをさまざまに組合わせたものを利用してよい。

【 0 0 3 8 】

本発明のある実施形態を実行し得るプログラム環境は、例示として、汎用コンピュータ、またはハンドヘルドコンピュータ等の専用機器を、1つ以上含む。このような機器（たとえばプロセッサ、メモリ、データ記憶装置、入力および出力装置）の詳細は、周知であり、わかりやすくするために、ここでは示さない。

【 0 0 3 9 】

本発明の技術がさまざまな科学技術を用いて実現し得ることも理解されるはずである。たとえば、本明細書に記載の方法は、コンピュータシステムで実行されるソフトウェアにおいて実現してもよく、または、方法の動作を実行するための1つ以上のプロセッサおよびロジック（ハードウェアおよび/またはソフトウェア）、特定用途向け集積回路、フィールドプログラマブルゲートアレイ（FPGA）等のプログラマブルロジックデバイス、および/またはこれらをさまざまに組み合わせたものを利用するハードウェアにおいて実現してもよい。例示としてのあるアプローチにおいて、本明細書に記載の方法は、物理的な（たとえば一時的でない）コンピュータ読取可能な媒体等の記録媒体にある、コンピュータによる実行が可能な一連の命令によって、実現してもよい。加えて、本発明の特定の実施形態は、オブジェクト指向ソフトウェアプログラミングコンセプトを採用し得るが、本発明はそのように限定される訳ではなく、コンピュータの動作の他の指示形態の採用に容易に適応する。

【 0 0 4 0 】

本発明は、計算装置（たとえばプロセッサ）および/またはシステムによって実行し得るコンピュータコードを有するコンピュータ読取可能な記録または信号媒体を含むコンピュータプログラムプロダクトの形態で提供することもできる。コンピュータ読取可能な記録媒体は、読取専用のおよび書込可能なCDおよびDVD等の光媒体、磁気メモリまたは媒体（たとえばハードディスクドライブ、テープ）、半導体メモリ（たとえばフラッシュメモリおよびその他のポータブルメモリカード等）、チップにおいて符号化されたファームウェア等を含む、計算装置またはシステムが使用する、コンピュータ符号を格納することが可能な任意の媒体を含み得る。

【 0 0 4 1 】

コンピュータ読取可能な信号媒体は、上記記録媒体の部類には入らないものである。たとえば、例示としてのコンピュータ読取可能な信号媒体は、システム内で、システム間で、たとえば物理または仮想ネットワーク等を介して一時的な信号を伝達するかそうでなければ転送する。

【 0 0 4 2 】

図1は、ある実施形態に従うアーキテクチャ100を示す。図1に示されるように、第1の遠隔ネットワーク104および第2の遠隔ネットワーク106を含む複数の遠隔ネットワーク102が与えられている。ゲートウェイ101は、遠隔ネットワーク102と近接ネットワーク108との間に接続されていてもよい。このネットワークアーキテクチャ100の文脈において、ネットワーク104、106は各々、インターネット、公衆電話交換網（PSTN）、内部電話網等のLAN、WANを含むがこれに限定されない形態を取り得る。

【 0 0 4 3 】

使用時、ゲートウェイ 101 は、遠隔ネットワーク 102 から近接ネットワーク 108 への入口のポイントとして機能する。このため、ゲートウェイ 101 は、ゲートウェイ 101 に到達した所与のデータパケットを導くことができるルータとして、および、所与のパケットに対してゲートウェイ 101 に出入りする実際の経路を与えるスイッチとして機能し得る。

【0044】

ゲートウェイ 101 を介して遠隔ネットワーク 102 からアクセスすることが可能な、近接ネットワーク 108 に接続された少なくとも 1 つのデータサーバ 114 がさらに含まれる。なお、データサーバ 114 は、いずれかの種類の計算装置 / グループウェアを含み得る。各データサーバ 114 に接続されているのは複数のユーザ装置 116 である。このようなユーザ装置 116 は、デスクトップコンピュータ、ラップトップコンピュータ、ハンドヘルドコンピュータ、プリンタ、またはその他の種類のロジックを含み得る。なお、ある実施形態において、ユーザ装置 111 もネットワークのいずれかに直接接続されていてもよい。

10

【0045】

1 つの周辺装置 120 または一連の周辺装置 120、たとえば、ファクシミリ機、プリンタ、ネットワークされた記憶ユニット等が、ネットワーク 104、106、108 のうちの 1 つ以上に接続されてもよい。なお、データベース、サーバ、および / またはその他構成要素が、ネットワーク 104、106、108 に接続された何らかの種類のネットワーク要素とともに使用されてもよく、またはこれに組み込まれてもよい。本明細書の文脈において、ネットワーク要素は、ネットワークのいずれかの構成要素を意味し得る。

20

【0046】

いくつかのアプローチに従うと、本明細書に記載の方法およびシステムは、仮想システム、および / または 1 つ以上の他のシステムをエミュレートするシステム、たとえば MAC OS 環境をエミュレートする UNIX (登録商標) システム、MICROSOFT WINDOWS (登録商標) 環境を事実上ホストする UNIX システム、MAC OS 環境をエミュレートする MICROSOFT WINDOWS システム等とともに実現されてもよく、および / またはそのシステム上で実現されてもよい。この仮想化および / またはエミュレートは、いくつかの実施形態では VMWARE のソフトウェアの使用を通して向上させてもよい。

30

【0047】

さらに他のアプローチにおいて、1 つ以上のネットワーク 104、106、108 は、一般的に「クラウド」と呼ばれているシステムのクラスタを表わしていてもよい。クラウドコンピューティングでは、処理能力、周辺機器、ソフトウェア、データ処理および / または記憶、サーバ等の共有リソースが、クラウド内の任意のシステムに、好ましくはオンデマンドの関係で与えられることにより、多数の計算システム全体におけるアクセスおよびサービスの分配が行なえるようにしている。クラウドコンピューティングは、典型的にはクラウド内で動作するシステム間のインターネットまたはその他高速接続 (たとえば 4G LTE、光ファイバー等) を必要とするが、システムの他の接続技術を使用してもよい。

40

【0048】

図 1 は、ある実施形態に従うアーキテクチャ 100 を示す。図 1 に示されるように、第 1 の遠隔ネットワーク 104 および第 2 の遠隔ネットワーク 106 を含む複数の遠隔ネットワーク 102 が与えられている。ゲートウェイ 101 は、遠隔ネットワーク 102 と近接ネットワーク 108 との間に接続されていてもよい。このアーキテクチャ 100 の文脈において、ネットワーク 104、106 は各々、インターネット、公衆電話交換網 (PS TN)、内部電話網等の LAN、WAN を含むがこれに限定されない形態を取り得る。

【0049】

使用時、ゲートウェイ 101 は、遠隔ネットワーク 102 から近接ネットワーク 108 への入口のポイントとして機能する。このため、ゲートウェイ 101 は、ゲートウェイ 1

50

01に到達した所与のデータパケットを導くことができるルータとして、および、所与のパケットに対してゲートウェイ101に出入りする実際の経路を与えるスイッチとして機能し得る。

【0050】

ゲートウェイ101を介して遠隔ネットワーク102からアクセスすることが可能な、近接ネットワーク108に接続された少なくとも1つのデータサーバ114がさらに含まれる。なお、データサーバ114は、いずれかの種類の計算装置/グループウェアを含み得る。各データサーバ114に接続されているのは複数のユーザ装置116である。このようなユーザ装置116は、デスクトップコンピュータ、ラップトップコンピュータ、ハンドヘルドコンピュータ、プリンタ、またはその他の種類のロジックを含み得る。なお、ある実施形態において、ユーザ装置111もネットワークのいずれかに直接接続されていてもよい。

10

【0051】

1つの周辺装置120または一連の周辺装置120、たとえば、ファクシミリ機、プリンタ、ネットワークされたおよび/またはローカル記憶ユニットまたはシステム等が、ネットワーク104、106、108のうちの1つ以上に接続されてもよい。なお、データベースおよび/またはその他構成要素が、ネットワーク104、106、108に接続された何らかの種類のネットワーク要素とともに使用されてもよく、またはこれに組み込まれてもよい。本明細書の文脈において、ネットワーク要素は、ネットワークのいずれかの構成要素を意味し得る。

20

【0052】

いくつかのアプローチに従うと、本明細書に記載の方法およびシステムは、仮想システム、および/または1つ以上の他のシステムをエミュレートするシステム、たとえばMAC OS環境をエミュレートするUNIX（登録商標）システム、MICROSOFT WINDOWS（登録商標）環境を事実上ホストするUNIXシステム、MAC OS環境をエミュレートするMICROSOFT WINDOWSシステム等とともに実現されてもよく、および/またはそのシステム上で実現されてもよい。この仮想化および/またはエミュレートは、いくつかの実施形態ではVMWAREのソフトウェアの使用を通して向上させてもよい。

【0053】

さらに他のアプローチにおいて、1つ以上のネットワーク104、106、108は、一般的に「クラウド」と呼ばれているシステムのクラスタを表わしていてもよい。クラウドコンピューティングでは、処理能力、周辺機器、ソフトウェア、データ処理および/または記憶、サーバ等の共有リソースが、クラウド内のシステムに、好ましくはオンデマンドの関係で与えられることにより、多数の計算システム全体におけるアクセスおよびサービスの分配が行なえるようにしている。クラウドコンピューティングは、典型的にはクラウド内で動作するシステム間のインターネットまたはその他高速接続（たとえば4G LTE、光ファイバー等）を必要とするが、システムの他の接続技術を使用してもよい。

30

【0054】

図2は、ある実施形態に従う、図1のユーザ装置116および/またはサーバ114に関連する代表的なハードウェア環境を示す。この図は、システムバス212を介して相互に接続されたマイクロプロセッサ等の中央処理装置210およびそれ以外の多数の装置を有するワークステーションの典型的なハードウェア構成を示す。

40

【0055】

図2に示されるワークステーションは、ランダムアクセスメモリ（RAM）214と、読取専用メモリ（ROM）216と、ディスク記憶ユニット220等の周辺装置をバス212に接続するためのI/Oアダプタ218と、キーボード224、マウス226、スピーカ228、マイク232、および/またはタッチスクリーンおよびデジタルカメラ（図示せず）等の他のユーザインターフェイス装置をバス212に接続するためのユーザインターフェイスアダプタ222と、ワークステーションを通信ネットワーク235（たとえ

50

ばデータ処理ネットワーク)に接続するための通信アダプタ234と、バス212をディスプレイ装置238に接続するためのディスプレイアダプタ236とを含む。

【0056】

ワークステーションには、Microsoft Windows(登録商標)のオペレーティングシステム(OS)、MAC OS、UNIX OS等のオペレーティングシステムが常駐していてもよい。好ましい実施形態は、上記のもの以外のプラットフォームおよびオペレーティングシステム上でも実現し得ることが理解されるであろう。好ましい実施形態は、オブジェクト指向プログラミング技法とともに、JAVA(登録商標)、XML、C、および/またはC++言語、またはその他のプログラミング言語を用いて記述されてもよい。複雑なアプリケーションの開発のために使用されることが増えているオブジェクト指向プログラミング(OOP)を用いてもよい。

10

【0057】

モバイル画像取込みおよび処理アルゴリズムのさまざまな実施形態

モバイル画像取込みおよび処理アルゴリズムのさまざまな実施形態、ならびに本開示の範囲に含まれるこのようなアルゴリズム処理を使用し易くするように構成されたいくつかのモバイルアプリケーションについて、以下で説明する。以下の各セクションに記載される機能は、本明細書に記載の機能のうちのいずれかまたはすべてを含むその他のセクションに開示される機能と組合わせて使用し得ることが、理解されるはずである。加えて、いくつかのアプローチでは、処理アルゴリズムの実施形態の機能とモバイルアプリケーションの実施形態を、組合わせてもよく、および/またはさまざまな計算リソースおよび/またはシステム全体に、何らかのやり方で分散させてもよい。

20

【0058】

アプリケーションは、モバイル装置にインストールされてもよく、たとえばモバイル装置の不揮発性メモリに格納されてもよい。あるアプローチにおいて、このアプリケーションは、モバイル装置上で画像処理を実行することを指示する命令を含む。別のアプローチにおいて、このアプリケーションは、1つ以上の非モバイル装置、たとえばネットワークサーバ等の遠隔サーバ、遠隔ワークステーション、クラウドコンピューティング環境等に画像を送ることを指示する命令を含むが、当業者はこのことを本明細書の記載を読んで理解するであろう。さらに他のアプローチにおいて、このアプリケーションは、モバイル装置上で一部またはすべての処理を実行するか否か、および/または画像を遠隔場所に送るか否かを判断することを指示する命令を含み得る。如何にして画像を処理し得るかについての例を、以下でより詳細に示す。

30

【0059】

ある実施形態において、モバイル装置上で実行し得る処理と、遠隔サーバ上で実行し得る処理との間に、処理速度、利用できるメモリに対する制約等を除いて、違いがない場合がある。加えて、たとえばモバイルアプリケーションの一部としての、モバイル装置上にあるさまざまなユーザインターフェイスと、非モバイル装置と通信するディスプレイ上にある対応するユーザインターフェイスとの間に、何らかの相違がある、または相違がない場合がある。

【0060】

他の実施形態において、遠隔サーバは、より高い処理能力、より多くの機能、より多くの処理アルゴリズム等を有し得る。さらに他の実施形態において、モバイル装置は、画像を遠隔サーバに送信するのに必要なもの以外、このアプリケーションに関連する画像処理機能を有さない場合がある。さらに別の実施形態において、遠隔サーバは、処理された画像を遠隔サーバから受信するのに必要なもの以外、本明細書に示されるプラットフォームに関連する画像処理能力を有さない場合がある。したがって、画像は、部分的にもしくは全体的にモバイル装置上で処理されてもよく、および/または部分的にもしくは全体的に遠隔サーバ上で処理されてもよく、および/または部分的にもしくは全体的にクラウドで処理されてもよく、および/または部分的にもしくは全体的に、間にあるアーキテクチャ全体のうちのいずれかの部分において処理されてもよい。加えて、いくつかの処理工程は

40

50

別々の装置上で繰返されてもよい。

【 0 0 6 1 】

どの装置が処理のどの部分を実行するかについては、たとえば、ユーザによって規定されてもよく、予め定められていてもよく、実行中に決定されてもよい。加えて、いくつかの処理工程は、たとえばユーザからの要求を受けて、再度実行されてもよい。したがって、生の画像データ、部分的に処理された画像データ、または完全に処理された画像データは、モバイル装置から、たとえば無線データネットワークを用いて、遠隔システムに送信されてもよい。遠隔システムで処理された画像データは、出力のためにおよび／またはさらに処理するために、モバイル装置に戻されてもよい。

【 0 0 6 2 】

さらに他のアプローチにおいて、画像は分割されてもよく、画像のさまざまな部分の処理がさまざまな装置に割当てられてもよく、たとえば、1 / 2 がモバイル装置に割当てられ1 / 2 が遠隔サーバに割当てられて、その後、処理されたこれら1 / 2 の部分が組合される。

【 0 0 6 3 】

ある実施形態において、処理を実行する装置の選択は、モバイル装置上のローカルな処理速度と、サーバとの通信速度との比較に、少なくとも一部基づいていてもよい。

【 0 0 6 4 】

あるアプローチにおいて、処理機能のライブラリがあってもよく、モバイル装置上のアプリケーションまたは遠隔サーバ上のアプリケーションは、このライブラリに対する呼出しを行なうだけでよく、実際、この呼出しの意味が、実行する処理の種類を定める。次に、装置はその処理を実行し、処理された画像を、おそらくは何らかの対応するメタデータとともに出力する。

【 0 0 6 5 】

さまざまな実施形態において、当該技術において周知のおよび／または本明細書で新たに示される、どのような種類の画像処理も、任意に組合わせて実行し得る。

【 0 0 6 6 】

次に、例示としての画像処理を説明する。カメラは、画像を取込むエリアセンサとみなすことができ、この画像には、いくつかの射影効果がある場合があり、場合によっては非線形効果がある場合がある。画像を処理してこのような効果を補正することができる。加えて、処理中に、画像内の文書の位置と境界線、たとえば紙の1つ以上の実際のページを囲む背景の中にある、この(これらの)ページの境界線を、検出してもよい。さまざまな実施形態のモバイルの性質のため、この紙は、ほとんどどのようなものの上にも載っている可能性がある。このことは、スキャナを用いて生成される文書の画像の処理と比較して、画像の解析を複雑にする。なぜなら、スキャナの背景特性は一定であって典型的には既知であるのに対し、モバイル取込みの背景は、文書がある場所および画像背景に取込まれた周囲の対応するテクスチャに応じて、また、変動し得る照明条件が原因で、ほぼ無限に変化する可能性があるからである。

【 0 0 6 7 】

したがって、カメラによる取込みのために1枚の紙が置かれるであろう表面の背景が一樣でないことが、1つの課題を示し、非線形効果および射影効果が、さらに他の課題を示す。さまざまな実施形態はこれら課題を克服する。このことはまもなく明らかになるであろう。

【 0 0 6 8 】

ある代表的な動作モードにおいて、モバイル装置上のアプリケーションは、たとえばアプリケーションを開くことを指示するユーザ要求に応じて、開始されてもよい。たとえば、このアプリケーションを表わすアイコンをユーザが選択したことが検出されてもよい。

【 0 0 6 9 】

いくつかのアプローチにおいて、ユーザ認証が要求されおよび／または実行されることがある。たとえば、ユーザIDおよびパスワード、またはその他の認証情報が、要求され

10

20

30

40

50

るおよび／またはユーザから渡される。

【0070】

さらに他のアプローチにおいて、アプリケーションのグラフィカルユーザインターフェイスを介してさまざまなタスクが可能にされてもよい。たとえば、タスクのリストが示されてもよい。このような場合、ユーザによるこのようなタスクのうちの1つの選択が検出されてもよく、その他の選択肢がユーザに示されてもよく、予め定められたタスクが開始されてもよく、カメラが起動されてもよい。

【0071】

画像は、モバイル装置のカメラによって、好ましくは、何らかの種類のユーザ入力を受けたことに応じて、たとえばモバイル装置の画面上のタップ、モバイル装置上のボタンの押下、音声コマンド、ジェスチャー等を受けたことに応じて、取込まれてもよい。可能性のあるもう1つのシナリオは、たとえばビデオストリームからの連続フレームの何らかのレベルの解析を伴ない得る。シーケンシャルフレーム解析の次に、高解像度の1つの画像フレームの取込みに切換えられてもよく、これは、いくつかのアプローチにおいて、自動的にまたはユーザによってトリガされてもよい。加えて、このトリガは、1つ以上のモバイル装置センサから受けた情報に基づいていてもよい。

10

【0072】

たとえば、ある実施形態において、モバイル装置内のまたはモバイル装置に接続された加速度計が、カメラの安定度を示してもよく、アプリケーションは、文書の低解像度のビデオフレームを解析してもよい。文書が検出されると、アプリケーションは、焦点合わせ動作を実行し、検出された文書の高解像度の画像を取得してもよい。

20

【0073】

さらに処理するのは低解像度の画像であっても高解像度の画像であってもよいが、好ましい実施形態は、さらに処理する対象として高解像度の画像を使用する。さらに他のアプローチにおいて、先に述べたようなシングルフレームモードへの切換えは、特に名刺および領収書のようなより小さな文書の場合、不要である場合がある。処理速度を高めるとともに処理リソースの消費を減じるために、文書の種類の識別によって、シングルフレームモードに切換えるか否かおよび／または処理の対象として高解像度の画像を取込むか否かを判断し易くしてもよい。ここでは1つ以上の文書の画像が取り込まれると想定する。

【0074】

一般的にモバイル装置には従来の非モバイル装置の処理能力がないと想定して、あるアプローチでは、モバイル装置に対して何らかの限定された処理を実行することによって、たとえばページが正しく検出されたこと、画像が不鮮明でないこと、および／または照明が十分であること、たとえばある種のプレビューを、ユーザに確認させてもよい。

30

【0075】

あるアプローチにおいて、カメラによって取込まれた画像の中の文書が検出されてもよい。

【0076】

文書の1本以上の境界線を検出する他の方法も本明細書に示される。画像の中の文書が非線形性を有しているかまたは矩形でない場合に、補正処理を適用してもよい。

40

【0077】

画像の中のページが検出されると、ある実施形態は、当然元の一枚の紙が矩形であると想定して、このページを矩形にするために、スムーズな変換を実行する。画像に対するもう1つの有効な補正は、照明のムラの緩和であってもよい。

【0078】

ある代表的なアプローチにおいて、ページ検出および矩形化は、実質的に下記のように実行し得る。

【0079】

モバイルページ検出のさまざまな実施形態

次に、ページ検出を実行するための代表的な方法体系を示す代表的な実施形態を、図3

50

A～図4を参照しながら説明する。これらの説明を参照すると、本明細書に記載のモバイル処理アルゴリズムに対して実現される利点が、如何にして、エリアセンサ（カメラ）によって取込まれた画像を扱いそのために現われる固有の障害を取除くかが、より明らかになるであろう。

【0080】

あるアプローチでは、特に図3A～図3Bを参照すると、エッジ検出アルゴリズムは、デジタル画像300の境界線から、背景の特性に関してわかっているものと十分に異なるポイントを探して、画像300の中心領域に向かって進む。特に、同じモバイル装置で取込まれていても、画像の背景304は毎回異なる場合があるので、画像の中の文書を識別するための新たな技術を提供する。

10

【0081】

本開示に従う、カメラで取込まれた画像の中のページのエッジを発見することにより、たとえばスキャナとは対照的に、モバイル装置を用いて取込まれた画像の特性の重要な相違に対応し易くなる。たとえば、射影効果が原因で、写真の中の矩形の文書の画像は、実際矩形に見えない場合があり、この画像内の文書の対向する辺の長さが同一でない場合がある。第2に、最高のレンズであっても何らかの非線形性があるので、結果として、オブジェクトの中の直線、たとえば実質的に矩形の文書の直線の辺が、このオブジェクトの取込まれた画像ではわずかに湾曲して見えることになる。第3に、カメラを用いて取込んだ画像は、取込んだ画像に、ムラのある照明効果を導入する傾向が極めて高い。この照明ムラが原因で、文書を置くことができる面の背景が完全に均質であっても、画像の中では、輝度にばらつきがあるように見え、特にページが完全に平坦でない場合はページのエッジの周りに影があることが多い。

20

【0082】

代表的なアプローチにおいて、背景の中の変化を間違っページのエッジとみなすのを避けるために、現在のアルゴリズムは、下記機能のうちの1つ以上を利用する。

【0083】

さまざまな実施形態において、画像のフレームは、文書302のデジタル表現を、その周囲の背景304である余白とともに含む。好ましい実装例において、個々のページエッジ306の探索は、画像の行と列を外側から内側に向かって解析するステップオーバー手法で実行してもよい。ある実施形態において、ステップオーバー手法は、図3A～図3Bに示されるような複数の解析ウィンドウ308をデジタル画像300の中に画定してもよい。本明細書では、解析ウィンドウ308は、1つ以上の「背景ウィンドウ」、すなわちデジタル画像300の背景304を示す画素のみを含むウィンドウと、1つ以上の「テストウィンドウ」、すなわちデジタル画像300の背景304、文書302のデジタル表現、または双方を示す画素を含むウィンドウを、含み得ると理解される。

30

【0084】

好ましい実施形態において、文書のデジタル表現を、デジタル画像の中で、文書が置かれた面の、背景304に対応する画像の余白に、第1の解析ウィンドウ308すなわち背景解析ウィンドウを画定することによって、検出してもよい。第1の解析ウィンドウ308では、複数の小解析ウィンドウ（たとえば図3Dに示されるようなテストウィンドウ312）が第1の解析ウィンドウ308の中に画定されていてもよい。複数のテストウィンドウ312を利用して、背景304を記述する1つ以上の統計的特性の1つ以上の分布を推定してもよい。

40

【0085】

引続き上記好ましい実施形態について説明する。文書のデジタル表現の境界線の検出における次のステップは、デジタル画像の中に複数のテストウィンドウ312を画定することと、このデジタル画像の対応する領域を解析することとを含み得る。テストウィンドウ312ごとに、この画像の対応する領域を記述する1つ以上の統計値を計算してもよい。さらに、これら統計値を、背景304を記述する、対応する統計値の分布と比較してもよい。

50

【0086】

好ましいアプローチにおいて、上記複数のテストウィンドウ312は、経路、特に直線の経路に沿って画定してもよい。とりわけ好ましいアプローチにおいて、複数のテストウィンドウ312は、水平方向および／または鉛直方向に、たとえばデジタル画像の行と列に沿って、画定してもよい。加えて、段階的進行を採用して、上記経路に沿い、および／または行および／または列の間に、テストウィンドウ312を画定してもよい。いくつかの実施形態では、段階的進行を利用すると有利に文書検出プロセスの計算効率を高めることができる。当業者はこのことを本明細書の記載を読んで理解するであろう。

【0087】

加えて、いくつかの実施形態では、開始段の大きさを、画像の解像度または画素サイズに基づいて推定してもよいが、この段は、以下でさらに説明するように、文書の辺の確実な検出に好都合であれば、小さくしてもよい。

【0088】

さらに他の実施形態において、上記アルゴリズムは、文書を囲む背景の中に置かれた大解析ウィンドウ308の中において発見された、画像特性を記述する統計値の分布を、推定する。あるアプローチにおいて、複数の小さなウィンドウ312を、大解析ウィンドウ308の中に画定してもよく、小さなテストウィンドウ312を記述する統計値の分布を推定してもよい。ある実施形態において、大解析ウィンドウ308は、デジタル画像の背景領域、たとえば画像の左上の角に画定される。

【0089】

背景画素を記述する統計値は、最小値、最大値、中央値、値の範囲の広がり、分散、標準偏差等の、デジタル画像データから生成することができる統計値を含み得る。当業者はこのことを本明細書の記載を読んで理解するであろう。デジタル画像300を記述するデータから、1つ以上のカラーチャネル、たとえば赤、緑、青すなわちRGB、シアン、マゼンタ、イエロー、ブラックすなわちCMYKにおける輝度値、色相飽和値すなわちHSV等の値を、サンプリングしてもよい。当業者はこのことを本明細書の記載を読んで理解するであろう。

【0090】

図3Dに示されるように、小解析ウィンドウ312は各々、大解析ウィンドウ308の中の複数の画素のサブセットを含み得る。加えて、小解析ウィンドウ312は、大解析ウィンドウ308の境界線の中に嵌めることができるどのようなサイズおよび／または形状であってもよい。好ましい実施形態において、小解析ウィンドウ312は、矩形形状で特徴付けられてもよく、より好ましくは、第1の方向の長さ（たとえば高さ）が3画素で、第2の方向の長さ（たとえば幅）が7画素であることを特徴とする矩形によって特徴付けられてもよい。もちろん、これ以外の小解析ウィンドウのサイズ、形状、および寸法も、本明細書に開示される処理アルゴリズムで実現するのに適している。

【0091】

ある実施形態において、テストウィンドウを用いることにより、画像を解析するとともに、この画像内に示されている文書のデジタル表現の境界線を検出してもよい。背景ウィンドウは、背景の当初の統計的特性の推定、および／または背景の局所的な統計的特性の再推定に使用される。照明ムラおよび／または背景テクスチャのばらつき等のアーティファクトに対処するためには、再推定が必要および／または好都合であろう。

【0092】

好ましくは、いくつかのアプローチにおいて、文書ページの外側にある余白の中の大解析ウィンドウ308の中にある複数の小解析ウィンドウ312のうちのいくつかまたはすべてに対して、統計的推定を行なってもよい。このような推定は、大解析ウィンドウ308の中で小解析ウィンドウ312を段階的に移動させることを利用して実行してもよく、所与の画素に対して取るサンプルの数を変えるために、段階的な移動の増分を適切な増分にしてもよい。たとえば、計算効率を高めるために、解析プロセスは、各画素318が一回サンプリングされることを保証するのに十分な、大解析ウィンドウ308内の小解析ウ

10

20

30

40

50

インドウ 3 1 2 の数を定めてもよい。このように、この高計算効率アプローチにおいて定めた複数の小解析ウィンドウ 3 1 2 は、境界線を共有するが重なり合うことはない。

【 0 0 9 3 】

統計的推定のロバスト性を高めるように設計された別のアプローチにおいて、解析プロセスは、たとえば、段を小さくして、連続して画定される小解析ウィンドウ 3 1 2 間の所与の方向におけるシフトを 1 画素のみにすることにより、各画素 3 1 8 が最大回数サンプリングされることを保証するのに十分な、多数の小解析ウィンドウ 3 1 2 を、大解析ウィンドウ 3 0 8 の中に画定してもよい。もちろん、本明細書に開示される処理アルゴリズムのさまざまな実施形態において採用する段の増分は、どのような増分であってもよい。当業者はこのことを本明細書の記載を読んで理解するであろう。

10

【 0 0 9 4 】

デジタル画像内の局所的背景の統計値を再推定するのに利用される大解析ウィンドウ 3 0 8 と、テストウィンドウは、望ましい任意の方法で、デジタル画像内に配置できることを、当業者は理解するであろう。

【 0 0 9 5 】

たとえば、図 3 A に示されるある実施形態に従うと、所与の行 i における左側のエッジの探索は、所与の行 i を中心とする画像の左辺上のフレーム境界線に隣接する大解析ウィンドウ 3 0 8 の中の上記統計値を計算することから始まる。

【 0 0 9 6 】

さらに他の実施形態において、上記アルゴリズムが画像の外側領域から画像の内側領域に向かって進行するとき起こり得る、背景でないテストウィンドウ（たとえば推定された統計値が、最後にわかった局所背景の統計的特徴の分布と似ていないテストウィンドウ）との遭遇が生じたとき、このアルゴリズムは、いくつかの実施形態において、以前求めた背景領域に後戻りし、新たな大解析ウィンドウ 3 0 8 を形成し、背景の統計値の分布を再推定して、小解析ウィンドウ 3 1 2 の中の選択された統計値と、大解析ウィンドウ 3 0 8 の中の対応する統計値の局所分布との違いの妥当性を再評価してもよい。

20

【 0 0 9 7 】

当業者が本明細書の記載を読んで理解するように、このアルゴリズムは、画像 3 0 0 の外側領域から、画像 3 0 0 の内側領域まで、さまざまな態様で進行し得る。たとえば、あるアプローチにおいて、このアルゴリズムは、実質的に螺旋のパターンでテストウィンドウ 3 1 2 を画定することから始まる。他のアプローチにおいて、このパターンは、鉛直方向または水平方向いずれかに沿って実質的に蛇行してもよい。さらに他のアプローチにおいて、このパターンは実質的に屋根板状のパターンであってもよい。このパターンはまた、デジタル画像 3 0 0 の一部または全面に置かれた、チェッカー盤パターン、鉛直、水平、または対角線方向のストライプのパターン、同心円形状等の、「シーケンスマスク」によって画定されてもよい。当業者はこのことを本明細書の記載を読んで理解するであろう。他の実施形態において、大解析ウィンドウ 3 0 8 および / または小解析ウィンドウ 3 1 2 等の解析ウィンドウを、何らかの規定された手順に従って、ランダム、疑似ランダム、確率論的等のやり方で、デジタル画像 3 0 0 全体に画定してもよい。当業者はこのことを本明細書の記載を読んで理解するであろう。このアルゴリズムは、経路が既知の背景に後戻りできる経路でありかつこの経路が望ましい粒度の画像全体をカバーする限り、何らかの望ましいやり方で一連のテストウィンドウを続けることができる。

30

40

【 0 0 9 8 】

好都合なことに、このように統計値を再計算すると、デジタル画像 3 0 0 および / または背景 3 0 4 に固有の照度のドリフトに対処し易くなる。さもなければ画像内の背景でないポイントであると間違って識別することになり得る（たとえば図 3 C に示されるアウトライナー候補エッジポイント 3 1 6 ）。

【 0 0 9 9 】

さらに他の実施形態において、違いが統計上有効であるとき、このアルゴリズムは、木目、表面上の擦り傷、表面の模様、小さな影等の、背景 3 0 4 のテクスチャの小さな変化

50

を再び確認して迂回するために、その経路に沿って特定距離だけさらに飛越してもよい。当業者はこのことを本明細書の記載を読んで理解するであろう。

【0100】

追加のおよび／または代替の実施形態において、背景でない可能性があるポイントを発見した後に、このアルゴリズムは、このポイントが影のエッジ上にあるか否か判断し（特にページのエッジが背景面よりも高い場合に起こり得る）、実際のページのエッジに到達しようと試みる。このプロセスは、通常影は実際のエッジに向かって濃くなりそれに続いて画像が突然明るくなるという観察に基づいている。

【0101】

上記ページエッジ検出のアプローチが使用されたのは、いくつかの理由から、標準的なエッジ検出器の使用が不要でありそれどころか望ましくないからである。第1に、最も標準的なエッジ検出器は、時間がかかる動作を必要とし、第2に、このアルゴリズムは、エッジがどれほど細いか、どの方向に沿っているかをモニタリングすることのような、その他の要件とは無関係である。より重要なのは、ページのエッジ306を探すときにエッジ検出そのものは必ずしも必要でないことである。すなわち、本開示に従うページエッジ検出は、文書の境界線（たとえばページエッジ306）を探索するのではなく、背景から文書への移行に関連する画像の特徴を探索するやり方で、実行し得るものである。たとえば、この移行は、光沢紙の中のオフホワイトの輝度レベルが平坦になること、すなわち、平均グレーレベルまたはカラーレベルではなくテクスチャの変化によって、特徴付けられる場合がある。

【0102】

その結果、格子上の各行および列における、実質的に最初と最後の非背景画素である、候補エッジポイント（たとえば図3Cに示される候補エッジポイント314）を得ることができる。ランダムなアウトライアー（たとえば図3Cに示されるアウトライアー候補エッジポイント316）を排除し、どの候補エッジポイント314がページの各辺に対応するか判断するためには、あるアプローチにおいて、近隣の候補エッジポイントを解析することが有効である。

【0103】

ある実施形態において、「ポイント」は、画素、画素間の位置（たとえば2画素×2画素の正方形の中心のような、分率座標を有するポイント）、画素の小さなウィンドウ等の、デジタル画像の中のいずれかの領域とみなすことができる。当業者はこのことを本明細書の記載を読んで理解するであろう。好ましい実施形態において、候補エッジポイントは、局所的背景を記述する統計値の分布と異なると判断された統計値によって特徴付けられることがわかっている、テストウィンドウ（たとえば3画素×7画素のウィンドウ）の中心と関連付けられる。

【0104】

本明細書では、「近隣の」候補エッジポイントまたは「近隣の」画素はそれぞれ、対象ポイントまたは画素（たとえば画素318）の近くまたは隣りのポイントまたは画素であり、たとえば、対象ポイントまたは画素の境界線に少なくとも一部に沿って位置するポイントまたは画素、対象ポイントまたは画素からしきい値距離以内（対象ポイントまたは画素の1つの行の中、対象ポイントまたは画素の1つの列の中の、所与の方向における2画素、10画素、64画素以内等）に位置するポイントまたは画素等であるとみなされると理解される。当業者はこのことを本明細書の記載を読んで理解するであろう。好ましいアプローチにおいて、「近隣の」ポイントまたは画素は、特定方向、たとえば水平方向および／または鉛直方向に沿って、対象ポイントに最も近い候補エッジポイントであってもよい。

【0105】

各「優良」エッジポイントは、理想的には少なくとも2つの隣接ポイントが（両側に1つずつ）あり、これら隣接ポイントを結ぶ直線の線分から大きく離れておらず、「優良」エッジポイント、たとえば、候補エッジポイントと、少なくとも2つの隣接ポイントは、

線形回帰に適していてもよく、その結果が、 0.95 以上の決定係数(R^2)によって特徴付けられてもよい。この線分の、デジタル画像の1本以上の境界線に対する角度は、その相対位置とともに、このエッジポイントが割当てられているのが、ページの上の辺なのか、左の辺なのか、右の辺なのか、または下の辺なのかを、決定する。好ましい実施形態において、候補エッジポイントと、2つの近隣エッジポイントを、三角形のそれぞれの角に割当ててもよい。この三角形の、候補エッジポイントにおける角度が、 180 度に近い場合、この候補エッジポイントは「優良な」エッジポイントであるとみなし得る。この三角形の、候補エッジポイントにおける角度が、しきい値を上回る角度(20 度以上等)分、 180 度から大きくずれている場合、この候補エッジポイントは、一組の「優良な」エッジポイントから除外してもよい。この発見的問題解決法の背後にある論理的根拠は、行と列の中の最初と最後の非背景画素の判断におけるランダム誤差を解消したいという希望にある。これら画素が存在する線が不変であることはありえないので、距離と方向という観点から近隣を調べることは、いくつかのアプローチにおいて特に好都合である。

10

【0106】

速度のためには、この格子の段を32等の大きな数から初めてもよいが、これを2分の1に減じ、エッジポイントの探索を、ページの辺の最小二乗平均(LMS)に基づく等式を求めるのに十分なエッジポイントが得られるまで繰返してもよい(下記を参照)。画像内のすべての行と列を使用した後になってもこのプロセスによって確実に辺を求めることができなければ、断念して画像全体をページとして扱う。

20

【0107】

ある実施形態において、ページの辺の等式は次のようにして求められる。最初に、このアルゴリズムは、最良のLMS直線を各辺に適合させる。そのときに使用する手法は、LMS直線からの距離が小さい範囲の中に、残りのサポートエッジがすべて入るまで、最悪のアウトライアーを捨てるという手法である。たとえば、文書の特定の境界線に沿う複数の候補エッジポイントを結ぶ実質的に真直ぐな線から最大の距離のポイントを、「最悪の」アウトライアーと指定してもよい。この手順を繰返すことによって、複数の候補エッジポイントから、1つ以上の「最悪の」アウトライアーを指定および/または削除すればよい。いくつかのアプローチにおいて、候補エッジポイントが、複数のエッジポイントを結ぶ線から離れているであろう距離は、デジタル画像のサイズおよび/または解像度に少なくとも一部に基づく。

30

【0108】

上記直線が、この直線の範囲全体に沿って十分にサポートされていない場合、このアルゴリズムは、最良の二次多項式(放物線)を、同じ当初の候補ポイントに適合させようと試みる場合がある。このアルゴリズムにおける、最良の放物線の発見と、最良の直線の発見の違いは、重要ではない。すなわち、直線の方角とオフセットを決める2つの未知の係数の代わりに、放物線の曲率、方向、およびオフセットと決める3つの係数がある。しかしながら、ある実施形態において、その他の点ではこのプロセスは本質的には同じである。

【0109】

放物線のサポートが直線のサポートよりも強力である場合、特に、候補エッジの範囲の端部により近い場合、結論として、このアルゴリズムでは、放物線の方が、画像の中のページの辺のより優れたモデルとして、好ましいはずである。それ以外の場合は、さまざまなアプローチにおいて、線形モデルが採用される。

40

【0110】

発見された文書の4つの辺の交点を計算することにより、四角形(たとえば図4に示され以下でより詳しく説明される四角形400)の角(わずかに湾曲している可能性がある)を発見してもよい。好ましい実装例では、そうするために、3つのケースを考慮することが必要である。すなわち、2つの直線の交点の計算、直線と放物線の交点の計算、および2つの放物線の交点の計算である。

【0111】

50

第1のケースでは、解は1つであり（なぜなら、上側と下側のページエッジ306は大部分が水平方向に延び、左側と右側のページエッジ306は大部分が鉛直方向に延び、対応するLMS直線は平行になり得ないからである）、この解が、対応するページの角の座標を決定する。

【0112】

第2のケースの、直線と放物線の交点の計算は、もう少し複雑である。すなわち、結果として得られる二次方程式の解は、ゼロ、1つ、または2つの可能性がある。交点がないということは、ページ検出の致命的な問題を示すかもしれないが、その結果は採用されない可能性がある。解が1つということは、どちらかと言えばありえないが、さらに他の問題は示さない。交点が2つの場合は、選択肢を示し、その場合、フレームの対応する角により近い交点が、より優良な候補ということになる。実際、方程式の他方の解は、画像フレームの座標範囲から非常に遠い可能性がある。

10

【0113】

第3のケースの、2つの放物線の交点の計算は、結果として、（原則として）解析的に解き得る四次多項式である。しかしながら、実際は、解を得るのに必要な計算の数が、所望のサブピクセル精度も保証する近似反復アルゴリズムよりも多い。

【0114】

以下では、この目的のために使用される代表的な手順について、あるアプローチに従う、文書302のデジタル表現の矩形化について述べながら、詳細に説明する。

【0115】

20

結果として得られる、目標の四角形（たとえば図4に関連して以下でさらに詳しく説明する四角形400）の妥当性には、いくつかの制約がある。すなわち、この四角形は小さすぎる（たとえば画像の総面積の25%といった所望の値の予め定められたしきい値未満）ことがないのが好ましく、この四角形の角は画像のフレームの外側に大きく離れすぎていない（たとえば100画素以下離れている）ことが好ましく、角そのものが、左上、右上、左下、および右下の角であると解釈でき、対角線が四角形の内側で交差することが好ましい等である。これらの制約が満たされていなければ、いくつかの実施形態において、所与のページ検出結果は採用されない場合がある。

【0116】

文書302のデジタル表現の検出された四角形が妥当である、例示としての実施形態において、このアルゴリズムは、目標の矩形を求めてもよい。目標の矩形の幅と高さはそれぞれ、四角形の上側の辺と下側の辺の平均と、左側の辺と右側の辺の平均に設定してもよい。

30

【0117】

ある実施形態において、スキュー補正を行なう場合、目標の矩形のスキューの角度をゼロに設定して、ページの辺が水平方向および鉛直方向になるようにしてもよい。そうでなければ、スキュー角度を、水平軸に対する上側の辺と下側の辺の角度の平均と、鉛直軸に対する左側の辺と右側の辺の平均に設定してもよい。

【0118】

同様に、クロップ補正を行わない場合、目標の矩形の中心を指定して四角形の4つの角の座標の平均に一致するようにしてもよい。そうでなければ、他の実施形態において、目標の矩形が画像フレームの左上になるように中心を計算してもよい。

40

【0119】

いくつかのアプローチにおいて、ページ検出結果が何らかの理由で採用されない場合、本明細書に記載のプロセスのいくつかまたはすべてのステップを、段の増分を小さくして繰返すことにより、より多くの候補エッジポイントを取得し、好都合に、より妥当な結果を得てもよい。最悪のケースのシナリオでは、許容される最小の段にしても問題が解消せず、検出されたページを画像フレーム全体に設定してもよく、元の画像には触れないままであってもよい。

【0120】

50

次に、本明細書に記載の発明のページ検出の実施形態の代表的な実装例を特に参照すると、あるアプローチにおいて、ページ検出は、下記のような方法を実行することを含む。この方法は、本明細書に記載され本明細書に開示される図面のいずれかに示されるものを含む、任意の環境で実行し得る。当業者はこのことを本明細書を読んで理解するであろう。

【 0 1 2 1 】

ある実施形態において、この方法は、デジタル画像の背景から文書のデジタル表現に移行する部分に対応する複数の候補エッジポイントを定める動作を含む。

【 0 1 2 2 】

さまざまな実施形態において、動作時に複数の候補エッジポイントを定めることは、以下で説明するような動作を含む、1つ以上の他の動作を含み得る。

【 0 1 2 3 】

ある動作では、ある実施形態に従うと、大解析ウィンドウ（たとえば図3A～図3Bおよび図3Dに示される大解析ウィンドウ308）がデジタル画像300の中に画定される。好ましくは、第1の大解析ウィンドウは、非背景部分（たとえば、動作を参照しながら以下でより詳しく説明する背景統計値等の、文書302のデジタル表現）に特有の情報との比較対象のための、デジタル画像の背景304に特有の情報を取得するために、デジタル画像の背景304の複数の画素を示すが非背景部分は示さない領域に、画定される。たとえば、第1の大解析ウィンドウ308は、デジタル画像300の角（左上の角等）に画定されてもよい。当然、第1の大解析ウィンドウは、本開示の範囲から外れることなく、デジタル画像300のいずれの部分にも画定し得る。

【 0 1 2 4 】

加えて、大解析ウィンドウ308は、どのようなサイズであってもよく、および/またはどのような適切な寸法で特徴付けられてもよいが、好ましい実施形態において、大解析ウィンドウ308は、高さが約40画素で幅が約40画素である。当業者はこのことを本明細書を読んで理解するであろう。

【 0 1 2 5 】

特に好ましいアプローチにおいて、大解析ウィンドウ308は、デジタル画像の角領域に画定されてもよい。たとえば、図3Aを参照すると、デジタル画像300が示されており、デジタル画像300は、複数の辺306を有する文書302のデジタル表現302と、背景304とを含む。動作に関して先に述べたように、大解析ウィンドウ308は、複数の背景画素を含むが文書302のデジタル表現に対応する画素は含まない領域に画定されてもよい。加えて、大解析ウィンドウ308は、いくつかのアプローチにおいて、デジタル画像300の角に画定されてもよい。

【 0 1 2 6 】

別の動作において、ある実施形態に従い、複数の小解析ウィンドウ312が、デジタル画像300の中、たとえば、大解析ウィンドウ308の中に、画定されてもよい。小解析ウィンドウ312は、他の1つ以上の小解析ウィンドウ312と、少なくとも一部が重なり合って、図3Dに示されるように1つ以上の重畳領域320を含むことによって特徴付けられてもよい。好ましいアプローチにおいて、可能なすべての小解析ウィンドウ312が、大解析ウィンドウ308の中に画定される。当然、図3Bに示されるように、小解析ウィンドウはデジタル画像のどの部分の中に画定されてもよく、好ましくは、小解析ウィンドウは、各小解析ウィンドウが1つの中心画素によって特徴付けられるように、画定されてもよい。

【 0 1 2 7 】

さらに他の動作において、ある実施形態に従い、1つ以上の小解析ウィンドウ312（たとえば大解析ウィンドウ308内の1つ以上の小解析ウィンドウ312）について1つ以上の統計値が計算され、対応する統計値の1つ以上の分布が推定される（たとえば複数の小解析ウィンドウ312全体について推定される統計値の分布）。別の実施形態において、1つ以上の大解析ウィンドウ308全体について統計値の分布が推定され、任意でマ

ージされてもよい。

【0128】

加えて、値は、デジタル画像の背景に関連する何らかの特徴を記述するものであればよく、たとえば、背景輝度の値、背景カラーチャネルの値、背景テクスチャの値、背景明度の値、背景コントラストの値、背景鮮鋭度の値等である。当業者はこのことを本明細書を読んで理解するであろう。さらにまた、統計値は、大解析ウィンドウ308内の複数の小解析ウィンドウ312全体のデジタル画像背景304を示す複数の画素の、1つ以上のカラーチャネルにおける輝度値の、最小値、最大値、および/または範囲を含み得る。

【0129】

さらに他の動作において、再び代表的なある実施形態に従うと、背景統計値の1つ以上の分布が推定される。統計値の分布を推定することにより、たとえば、大解析ウィンドウ308内のデジタル画像300の背景304の特性を特徴付ける記述的分布を得ることができる。

【0130】

この分布は、好ましくは小解析ウィンドウごとに計算された背景統計値に対応し、たとえば、輝度最小値の分布、輝度最大値の分布等を含み得る。この分布から、最小輝度値の最小値および/または最大値、最小輝度値の最小値および/または最大値、輝度値の最小および/または最大の広がり、最小カラーチャネル値の最小値および/または最大値、最大カラーチャネル値の最小値および/または最大値、カラーチャネルの最小および/または最大の広がりといった、分布の統計的記述子が得られる。当業者はこのことを本明細書を読んで理解するであろう。当然、計算された背景統計値(たとえば輝度の値、カラーチャネルの値、コントラストの値、テクスチャの値、明度の値、鮮鋭度の値等)を集めて1つの分布にしてもよく、分布を記述する任意の値を本開示の範囲から外れることなく採用することができる。

【0131】

さらに他の動作では、ある実施形態に従い、図3A~図3Bに示されるような解析ウィンドウ308等の大解析ウィンドウが、デジタル画像300内に画定される。

【0132】

加えて、ウィンドウの形状は、デジタル画像300の一部としてウィンドウの境界線をポジティブに設定することによって画定されてもよく、たとえばマスクをデジタル画像300に与えマスクされていないデジタル画像300の領域を解析ウィンドウとして画定することによってネガティブに画定されてもよい。加えて、特にマスクをデジタル画像300に与えることによってウィンドウがネガティブに画定される実施形態では、ウィンドウはパターンに従って画定されてもよい。当然、本開示の範囲から外れることなく、ウィンドウを画定する他の方法が採用されてもよい。

【0133】

さらに他の実施形態において、この方法は、解析ウィンドウ312について1つ以上の統計値が算出される動作を実行することを含み得る。加えて、好ましい実施形態において、各解析ウィンドウ統計値は、動作において大解析ウィンドウ308について推定された背景統計値の分布に対応する。たとえば、ある実施形態において、最大輝度は背景輝度最大値の分布に対応し、最小輝度は背景輝度最小値の分布に対応し、輝度の広がりには背景輝度の広がり分布に対応する等である。当業者はこのことを本明細書を読んで理解するであろう。

【0134】

さらに他の実施形態において、この方法は、少なくとも1つの解析ウィンドウ統計値と、対応する背景統計値の分布との間に、統計上重要な相違が存在するか否かを判断することを含む。当業者が本明細書を読んで理解するように、統計上重要な相違が存在するか否かの判断は、p値、z検定、カイ二乗相関等の、周知の統計上重要な評価方法または評価基準を用いて行なわれてもよい。当業者はこのことを本明細書を読んで理解するであろう。

【0135】

10

20

30

40

50

追加のおよび／または代替のアプローチにおいて、この方法は、画素 3 1 8 を記述する値と、対応する背景統計値の分布との間に統計上重要な相違がある解析ウィンドウにおける、1 つ以上のポイント（たとえば最も中心にある画素 3 1 8 またはポイント）を、候補エッジポイントとして指定することを含む。この指定は、この画素に対応するフラグを設定すること、この画素の座標を記憶すること、画素座標のアレイを作成すること、画素 3 1 8 を記述する 1 つ以上の値（輝度、明度、コントラスト等）を変更すること、またはその他の適切な手段といった、当該技術では周知の適切な方法によって、行なってもよい。

【0 1 3 6】

動作時には、ある実施形態に従い、動作のうちの 1 つ以上を一回以上繰返してもよい。好ましい実施形態において、このような繰返しを複数行なってもよく、各繰返しはデジタル画像の異なる部分に対して実行される。好ましくは、繰返しは、文書のデジタル表現の各辺の評価が完了するまで行なわれてもよい。さまざまなアプローチにおいて、解析ウィンドウ 3 0 8、3 1 2 の画定により、結果として、一本以上の境界線を共有する、全体または一部が重なり合う、および／または共有する境界線がなく重なり合わない、複数の解析ウィンドウ 3 0 8、3 1 2 が得られるであろう。当業者はこのことを本明細書を読んで理解するであろう。

【0 1 3 7】

特に好ましい実施形態において、複数の繰返しは、非背景ウィンドウの可能性のあるウィンドウ（たとえば、候補エッジポイントを含むウィンドウ、または、照明ムラ、背景テクスチャのばらつき等のアーティファクトを含むウィンドウ）を検出したときに、局所的な背景統計値を再推定することに関連したやり方で、行なわれてもよい。

【0 1 3 8】

動作時に、ある実施形態に従うと、四角形 4 0 0 の 4 つの辺が、複数の候補エッジポイントに基づいて画定される。好ましくは、四角形 4 0 0 の辺は、デジタル画像 3 0 0 内の文書のデジタル表現 3 0 2 のエッジ 3 0 6 を取囲む。四角形 4 0 0 の辺を画定することは、あるアプローチにおいて、1 以上の最小二乗平均（LMS）近似を実行することを含み得る。

【0 1 3 9】

さらに他のアプローチにおいて、四角形 4 0 0 の辺を画定することは、1 つ以上のアウトライアー候補エッジポイントを識別し、1 つ以上のアウトライアー候補エッジポイントを複数の候補エッジポイントから取除くことを含み得る。さらに、四角形 4 0 0 の辺を画定することは、上記 1 つ以上のアウトライアー候補エッジポイントを除外する少なくとも 1 つの追加の LMS 近似を実行することを含み得る。

【0 1 4 0】

またさらに、ある実施形態において、四角形 4 0 0 の各辺は、関数のクラスから選択された等式によって特徴付けられ、少なくとも 1 つの LMS 近似を実行することは、等式ごとに、好ましい実装例における二次多項式の最良の係数といった、1 つ以上の係数を求めることを含む。これらのアプローチに従うと、四角形 4 0 0 の辺を画定することは、文書のデジタル表現の各辺が、二次多項式または二次多項式の代わりに線形関数等のより単純な関数といった、所与のクラスの関数に含まれるか否か判断することを含み得る。

【0 1 4 1】

好ましいアプローチにおいて、方法を実行することによって、文書の 4 つの主要な辺の周りに四角形を正確に画定する一方で、図 3 A ~ 図 3 C および図 4 に示されるリップ 3 1 0 および／またはタブ 3 2 0 のような、文書の主要な辺からの 1 つ以上のずれは無視する。

【0 1 4 2】

本明細書に開示される四角形 4 0 0 の、他のおよび／または代替の実施形態は、各辺が上記多項式関数のような 1 つ以上の等式によって特徴付けられる、4 つの辺を有することを特徴とし得る。たとえば、四角形 4 0 0 の辺が 2 つ以上の等式によって特徴付けられる実施形態は、1 つ以上の辺を、各線分が上記多項式関数のような等式によって特徴付けら

10

20

30

40

50

れる、複数の線分に分割することを含み得る。

【0143】

さまざまな実施形態において、四角形400を画定することは、代わりにおよび/または加えて、四角形400の1つ以上の角を画定することを含む。たとえば、四角形400の角は、四角形400の隣合う辺と辺の1つ以上の交点を計算し、複数の交点が計算された場合は計算された1つ以上の交点から適切な交点を指定することによって、画定し得る。さらに他の実施形態において、角を画定することは、1つ以上の等式を解くことを含み得る。この場合の各等式は、N次多項式等の選ばれた関数のクラスに属することを特徴とする。当業者はこのことを本明細書を読んで理解するであろう。

【0144】

さまざまな実施形態において、四角形400の角は、四角形400の隣合う2つの湾曲した辺と辺の交点、実質的に直線である2つの線の交点、および、実質的に直線である一本の線と実質的に湾曲している一本の線の交点、のうちの、1つ以上によって画定されてもよい。

【0145】

さらに他の実施形態において、この方法は、文書のデジタル表現302および四角形400がモバイル装置のディスプレイに出力される、追加のおよび/または代替の動作を含み得る。出力は、どのようなやり方で行なわれてもよく、モバイル装置のハードウェアおよび/またはソフトウェアの構成に応じて決まるであろう。

【0146】

加えて、出力は、さまざまなアプローチにおいて、さらに他の処理および/またはこの出力とユーザの対話を容易にするように行なわれてもよい。たとえば、ある実施形態において、四角形400を、たとえば四角形400の辺を、特定の色、パターン、照明モチーフ、アニメーション等で表示することにより、四角形400をデジタル画像300の他の特徴から区別するように設計されたやり方で、表示してもよい。当業者はこのことを本明細書を読んで理解するであろう。

【0147】

またさらに、いくつかの実施形態において、四角形400と文書のデジタル表現302を出力することにより、ユーザが何らかの適切なやり方で四角形400を手動で調整および/または画定することを容易にしてもよい。たとえば、ユーザは、モバイル装置のディスプレイと対話することにより、四角形400を変換してもよい、すなわち四角形400のアスペクト比、形状、エッジの長さ、面積等は維持したままで、四角形400の位置を1つ以上の方向に移動させてもよい。これに加えておよび/またはこれに代えて、ユーザは、モバイル装置のディスプレイと対話することにより、四角形400の角の位置を手動で画定または調整してもよく、たとえば、四角形400の角をタッピングしこの角をデジタル画像300内の所望の位置、たとえば文書のデジタル表現302の角に、ドラッグしてもよい。

【0148】

再び図4を参照すると、ページ検出の理想的な結果の具体的な一例が示されており、デジタル画像300の中にある文書のデジタル表現302が示され、四角形400が文書のデジタル表現302のエッジを取囲んでいる。

【0149】

いくつかのアプローチにおいて、上述のようなページ検出方法は、以下で説明するような1つ以上の追加のおよび/または代替の動作を含み得る。

【0150】

あるアプローチにおいて、ページ検出は、文書のデジタル表現とこの文書のデジタル表現に関連する音声データとを含む画像データのうちの1つ以上を取込むことをさらに含み得る。取込みは、マイク、カメラ、加速度計、センサ等の、モバイル装置に接続された1つ以上の取込み装置を用いて行なえばよい。当業者はこのことを本明細書を読んで理解するであろう。

10

20

30

40

50

【0151】

別のアプローチにおいて、ページ検出は、新たな大解析ウィンドウ309を画定することと、この新たな大解析ウィンドウ309の背景統計値の分布を、統計上重要な相違が存在すると判断したときに、再推定することを含み得る。すなわち、動作1908および/または1910を、非背景点の可能性がある点に近い点、たとえば文書のエッジ306のうちの1つに近い点が識別された場合に、その点に近い、デジタル画像300の異なる領域において、実質的に繰返すことを含み得る。

【0152】

いくつかの代表的な実施形態において、大解析ウィンドウ308を、ある行の左端の非背景画素またはその近くに配置してもよく、ある行の右端の非背景画素またはその近くに配置してもよく、ある列の上端の非背景画素またはその近くに配置してもよく、または、ある列の下端の非背景画素またはその近くに配置してもよい。

10

【0153】

このような再推定を含むアプローチはさらに、少なくとも1つの小解析ウィンドウ（たとえばテストウィンドウ）統計値と、大解析ウィンドウ統計値の対応する再推定された分布との間に、統計上重要な相違が存在するか否か判断することを含み得る。このようにして、統計上重要な相違が存在するか否かについて、より信頼性が高い判断を得ることができ、したがって、たとえばデジタル画像内のテクスチャのばらつき、照明の異常、および/またはその他のアーティファクトの場合とは異なり、デジタル画像の背景から文書のデジタル表現への真の移行をより良く識別することができる。

20

【0154】

加えて、上述の再推定を実行してもしなくても、デジタル画像内の照明および/または背景テクスチャのばらつき等といった1つ以上のアーティファクトを回避する方法を容易にし得る。このアーティファクトは、デジタル画像の背景から文書のデジタル表現への真の移行に対応しないものである。いくつかのアプローチにおいて、アーティファクトの回避は、デジタル画像の1つ以上の領域（たとえばテクスチャ、ばらつき等によって特徴付けられる領域であって、このテクスチャ、ばらつき等は領域を真の背景から区別する）を迂回する形態を取ってもよい。

【0155】

いくつかのアプローチにおいて、大解析ウィンドウ308について推定された統計分布と小解析ウィンドウ312について計算された対応する統計値との間に統計上重要な相違が存在すると判断し、この小解析ウィンドウの近くに新たな大解析ウィンドウを画定し、この新たな大解析ウィンドウについて統計値の分布を再推定し、この再推定された統計分布と小解析ウィンドウ312について計算された対応する統計値との間に統計上重要な相違がないと判断したときは、1つ以上の領域を迂回してもよい。

30

【0156】

他のアプローチにおいて、迂回は、経路に沿ってさらに別の解析ウィンドウ312を検査し、この検査したウィンドウの統計値が、たとえば統計上の重要性のテストによって示されるように、背景の既知の統計特性と大きく変わらないと判断したときには、非背景領域への移行部の探索を再開することによって、実現されてもよい。

40

【0157】

迂回は、経路に沿ってさらに別の解析ウィンドウを検査することによって実現されてもよく、当業者はこのことを本開示を読んで理解するであろう。

【0158】

さらに他のアプローチにおいて、ページ検出は、さらにおよび/または代替的に、四角形400が1つ以上の品質管理メトリクスを満たすか否か判断することと、四角形400が品質管理基準のうちの1つ以上を満たさないと判断されると四角形400を採用しないことを含み得る。加えて、品質管理メトリクスは、LMSサポートメトリクス、最小の四角形400の面積のメトリクス、四角形400の角の位置のメトリクス、および四角形400の対角線の交点の位置のメトリクス等の尺度を含み得る。

50

【0159】

実際、四角形400がこれらメトリクスのうちの1つ以上を満たすか否かの判断は、方法の実施についての検査の役割を果たす。たとえば、検査は、四角形400がデジタル画像全体の面積の少なくともしきい値をカバーするか否か、たとえば、四角形400が画像の総面積の少なくとも25%を含むか否かの判断を含み得る。

【0160】

さらに、検査は、四角形400の対角線が四角形400の境界の内側で交差するか否か判断すること、LMS近似のうちの1つ以上が、そこから得られる統計値のロバストな信頼度を得るのに十分なデータから計算されたか否かを判断すること、すなわち、LMS近似に十分な「サポート」（さまざまなアプローチにおいて少なくとも5つのデータポイント、または、データポイントの総数の少なくとも四分の一から計算された近似等）があるか否か判断すること、および/または（四角形400の各辺を特徴付ける等式によって定められる）四角形400の角の位置がデジタル画像のエッジのしきい値距離内に存在するか否か判断すること、たとえば、四角形400の角が所与の方向においてデジタル画像のエッジから100画素を超える画素離れているか否か判断することを、含み得る。当然ながら、これら開示の範囲から逸脱することなく、他のメトリクスおよび/または検査を使用してもよく、当業者はこのことを本明細書を読んで理解するであろう。

【0161】

あるアプローチにおいて、品質メトリクスおよび/または検査により、最適ではない四角形400の定義を不採用にすることを容易にすることができ、さらに、四角形400の辺の定義を改善することを容易にすることができる。たとえば、あるアプローチは、複数の候補エッジポイントに基いて四角形400の四辺を画定したがこれは有効な四角形400を画定していない、すなわち品質管理メトリクスのうちの1つ以上を満たさないという表示を受けることと、複数の候補エッジポイントを再度画定することを含む。特に、この実施形態において、複数の候補エッジポイントを再度画定することは、デジタル画像内の、失敗した先の試みでサンプリングしたポイントの数よりも多い数のポイントをサンプリングすることを含む。これは、あるアプローチでは、デジタル画像の行または列のうちの1つ以上に対するステップを減じ、より多くの数の候補エッジポイントを解析するためにアルゴリズムのステップすべてを繰返すことによって、可能であろう。このステップは、垂直方向において減じてもよく、水平方向において減じてもよく、または両方の方向において減じてもよい。当然ながら、候補エッジポイントを再度画定するおよび/またはデジタル画像内のポイントを再度サンプリングするための他の方法を、本開示の範囲から逸脱することなく利用してもよい。

【0162】

さらに、特に方法を複数回繰返しデジタル画像解析全体において進行のステップを大幅に小さくしても有効な四角形400が画定されない場合、ページ検出は、デジタル画像全体を文書のデジタル表現として指定することを含み得る。あるアプローチにおいて、デジタル画像全体を文書のデジタル表現として指定することは、画像の角を文書の角として定めること、画像の辺を文書の辺として定めること等を含み得る。当業者はこのことを本明細書を読んで理解するであろう。

【0163】

本明細書に記載の、四角形400の対角線は、四角形400の計算された左上の角と四角形400の計算された右下の角とを結ぶ第1の直線と、四角形400の計算された右上の角と四角形400の計算された左下の角とを結ぶ第2の直線によって特徴付けてもよい。加えて、第1の直線と第2の直線は好ましくは四角形400の内部で交差する。

【0164】

さまざまなアプローチにおいて、上記動作のうちの1つ以上をプロセッサを用いて実行してもよく、プロセッサは、モバイル装置、特に一体化されたカメラを有するモバイル装置の一部であってもよい。

【0165】

湾曲補正

本明細書に記載の発明の概念は、デジタル画像における湾曲を補正すること、より具体的には文書のデジタル表現の辺の湾曲を補正することを含む。以下では、代表的に図5A～図5Bを参照して湾曲の補正についてのさまざまなアプローチを詳細に説明する。

【0166】

ある実施形態において、湾曲補正アルゴリズムの目的は、四角形400を四辺形にスムーズに変換することである。特に、四角形400は複数の等式によって特徴付けられ、各等式は、四角形400の一辺に対応し、選ばれたクラスの関数から選択される。たとえば、四角形400の各辺は、一次多項式、二次多項式、三次多項式等によって特徴付けられてもよく、当業者はこのことを本明細書を読んで理解するであろう。

10

【0167】

あるアプローチにおいて、四角形400の辺は等式によって説明してもよく、好ましい実施形態において、四角形400の左辺は二次多項式 $x = a_2 * y^2 + a_1 * y + a_0$ によって特徴付けられ、四角形400の右辺は二次多項式 $x = b_2 * y^2 + b_1 * y + b_0$ によって特徴付けられ、四角形400の上辺は二次多項式 $y = c_2 * x^2 + c_1 * x + c_0$ によって特徴付けられ、四角形400の下辺は二次多項式 $y = d_2 * x^2 + d_1 * x + d_0$ によって特徴付けられる。

【0168】

本明細書に示す湾曲補正の説明では、四角形の中の、複数の四角形ベースの固有座標対 (p, q) の定義を利用する。各固有座標対 (p, q) は、左辺と右辺の等式から、上下方向の曲線の係数比におけるすべての対応する係数 $p \sim 1 - p$ を組み合わせることによって得られる等式によって特徴付けられる上下方向の曲線と、上辺と下辺の等式から、左右方向の曲線の係数比におけるすべての対応する係数 $q \sim 1 - q$ を組み合わせることによって得られる等式によって特徴付けられる左右方向の曲線との交点に対応し、 $0 \leq p \leq 1$ 、 $0 \leq q \leq 1$ である。

20

【0169】

四角形400の辺が二次多項式によって特徴付けられる好ましい実施形態において、固有座標 p に対応する上下方向の曲線は、等式 $x = ((1 - p) * a_2 + p * b_2) * y^2 + ((1 - p) * a_1 + p * b_1) * y + ((1 - p) * a_0 + p * b_0)$ によって特徴付けられ、固有座標 q に対応する左右方向の曲線は、等式 $y = ((1 - q) * c_2 + q * d_2) * x^2 + ((1 - q) * c_1 + q * d_1) * x + ((1 - q) * c_0 + q * d_0)$ によって特徴付けられる。当然ながら、当業者が本明細書を読めば理解するように、他の等式によって上記辺および/または曲線を特徴付けてもよい。

30

【0170】

平行四辺形の場合、固有座標は特に単純になる。平行四辺形の中で、各固有座標対 (p, q) は、平行四辺形の左辺および平行四辺形の右辺各々に平行な線、たとえば $p \sim 1 - p$ の部分において上辺および下辺双方を分割する線と、平行四辺形の上辺および平行四辺形の下辺各々に平行な辺、たとえば $q \sim 1 - q$ の部分において上辺および下辺双方を分割する辺との交点に対応し、 $0 \leq p \leq 1$ 、 $0 \leq q \leq 1$ である。別の特殊なケースにおいて、四角形が単位正方形、すなわち長さ1の辺を有する正方形のとき、座標は通常のデカルト座標そのものであり、座標 (p, q) のポイントは、鉛直線 $x = p$ と水平線 $y = q$ の交点である。

40

【0171】

下記湾曲補正アルゴリズムの目的は、湾曲補正された画像内の各ポイントを、元の画像内の対応するポイントにマッチングし、このマッチングを、四角形400の四辺各々を、現在の四角形の角を結ぶ実質的に真直ぐな線に変換するように、行なうことである。しかしながら、同じ技術によって、その四辺の等式によって表わされる四角形を、他のこのような四角形にスムーズに変換することができる。

【0172】

下記座標マッピングアルゴリズムの主な意図は、この目的を、第1に、目標とする画像

50

内の各ポイント P (図示せず) の固有座標 (p , q) を計算し、第 2 に、これらを、元の画像の固有座標の同じ対 (p , q) にマッチングし、第 3 に、これら固有座標にそれぞれ対応する左右湾曲および上下湾曲の交点の座標を計算し、最後に、元の画像において発見されたポイントの色またはグレー値をこのポイント P に割当てることによって、達成することである。

【 0 1 7 3 】

次に、ある実施形態に従うページ湾曲補正アルゴリズムの第 1 の反復を図示する図 5 A を参照する。図 5 A に示されるように、デジタル画像 5 0 0 内の各ポイントは、ポイントに関連付けられた (上記のような) 固有座標に対応する、上下曲線 5 0 4 と左右曲線 5 0 6 との交点に対応し得る (曲線は、真直ぐな線、湾曲した線、たとえば放物線等を含み得るものであり、当業者はこのことを本明細書を読んで理解するであろう) 。

10

【 0 1 7 4 】

この説明から明らかになるように、湾曲補正は、複数のこのような左右方向の線 5 0 6 と上下方向の線 5 0 4 を定めることを含み得る。

【 0 1 7 5 】

加えて、湾曲補正は、目標とする固有座標を、文書のデジタル表現 5 0 2 の元の固有座標にマッチングすることを含み得る。

【 0 1 7 6 】

図 5 A に示されるように、このマッチングは、所与の左右曲線 5 0 6 と所与の上下曲線 5 0 4 との交点を繰返し探索することを含み得る。図 5 A は、本開示の範囲に含まれる代表的な反復探索の第 1 の反復を示す。

20

【 0 1 7 7 】

以下でさらに詳細に説明するあるアプローチに従うと、反復探索は、座標 (x_0 , y_0) を有する起点 5 0 8 を指定することを含む。起点 5 0 8 は、文書のデジタル表現 5 0 2 内のどこに位置していてもよいが、好ましくは目標とする四角形の中心または中心近くに位置する。

【 0 1 7 8 】

反復探索は、交差する 2 つの曲線 5 0 4 、 5 0 6 のうちの一方の上に起点 5 0 8 を射影することを含み得る。起点はこれら曲線 5 0 4 、 5 0 6 のうちいずれに射影してもよいが、あるアプローチにおいて、反復探索における第 1 の探索の前半は、起点 5 0 8 を上下曲線上に射影することにより、次の点の x 座標 (x_1) を得ることを含み、この射影の結果を図 5 A における点 5 1 0 で表わすとその座標は (x_1 , y_0) である。同様に、いくつかの実施形態において、反復探索の第 1 の反復の後半は、この点 5 1 0 を左右曲線 5 0 6 上に射影することにより、次の点の y 座標 (y_1) を得ることを含み、この射影の結果を図 5 A における点 5 1 2 で表わすとその座標は (x_1 , y_1) である。

30

【 0 1 7 9 】

図 5 B は、ある実施形態に従い、文書のデジタル表現 5 0 2 を、複数の上下曲線 5 0 4 と複数の左右曲線 5 0 6 によって定められた同じサイズの複数の区画に分割した後の、ページ湾曲補正アルゴリズムの基点を図示する。

【 0 1 8 0 】

40

いくつかの実施形態において、以下でさらに詳細に説明する同様の手法を用いてさらに反復を行なってもよい。

【 0 1 8 1 】

引続き図 5 A ~ 図 5 B を参照して、ある実施形態に従う、デジタル画像内の文書のデジタル表現の 1 つ以上の空間特性を修正する方法は、以下の動作のうちの 1 つ以上を含み得る。当業者が本明細書を読めば理解するように、この方法は、図面および本開示の対応する説明において示されおよび / または記載されるものを含む、何らかの適切な環境で実行し得る。

【 0 1 8 2 】

ある実施形態において、この方法は、四角形 4 0 0 を四辺形に変換する動作を含む。特

50

に、四角形 4 0 0 は複数の等式によって特徴付けられ、各等式は四角形 4 0 0 の一辺に対応し、選ばれたクラスの関数から選択される。たとえば、四角形 4 0 0 の各辺は、当業者が本明細書を読めば理解するように、一次多項式、二次多項式、三次多項式等によって特徴付けられてもよい。

【0183】

ある実施形態において、四角形 4 0 0 の辺は等式によって説明してもよく、好ましい実施形態において、四角形 4 0 0 の左辺は二次多項式 $x = a_2 * y^2 + a_1 * y + a_0$ によって特徴付けられ、四角形 4 0 0 の右辺は二次多項式 $x = b_2 * y^2 + b_1 * y + b_0$ によって特徴付けられ、四角形 4 0 0 の上辺は二次多項式 $y = c_2 * x^2 + c_1 * x + c_0$ によって特徴付けられ、四角形 4 0 0 の下辺は二次多項式 $y = d_2 * x^2 + d_1 * x + d_0$ によって特徴付けられる。加えて、上下曲線の等式は $x = ((1 - p) * a_2 + p * b_2) * y^2 + ((1 - p) * a_1 + p * b_1) * y + ((1 - p) * a_0 + p * b_0)$ であり、左右曲線の等式は $y = ((1 - q) * c_2 + q * d_2) * x^2 + ((1 - q) * c_1 + q * d_1) * x + ((1 - q) * c_0 + q * d_0)$ である。当然ながら、当業者が本明細書を読めば理解するように、他の等式によって上記辺および/または曲線のいずれかを特徴付けてもよい。

10

【0184】

ある実施形態において、曲線 5 0 4、5 0 6 は、以下の一般的な形態のうちの 1 つ以上に適合する代表的な多項式関数によって説明してもよい。

20

【0185】

$$x_1 = u_2 * y_0^2 + u_1 * y_0 + u_0$$

$$y_1 = v_2 * x_1^2 + v_1 * x_1 + v_0$$

式中、 $u_i = (1 - p) * a_i + p * b_i$ 、 $v_i = (1 - q) * c_i + q * d_i$ であり、 a_i は四角形の左辺の等式における係数であり、 b_i は四角形の右辺の等式における係数であり、 c_i は四角形の上辺の等式における係数であり、 d_i は四角形の下辺の等式における係数であり、 p および q は曲線 5 0 4、5 0 6 に対応する四角形ベースの固有座標である。いくつかのアプローチにおいて、 a_i 、 b_i 、 c_i 、 d_i 等の係数は、ページ検出に関して先に述べたようなページ検出法等のページ検出を実行する過程で得られる計算、推定、および/または判断から導出されるものであってもよい。

30

【0186】

当然ながら、当業者が本明細書を読めば理解するように、四角形 4 0 0 を四辺形に変換することは、以下でより詳細に説明するように 1 以上の加算演算を含み得る。

【0187】

ある実施形態において、方法は、さらにおよび/または代替的に、四角形 4 0 0 の 1 つ以上の領域を、さらに他のアーティファクト（四角形の内部領域の歪み等）を結果として得られる四辺形に導入しないよう十分スムーズなやり方で、伸張することを含み得る。

【0188】

いくつかのアプローチにおいて、四角形 4 0 0 を矩形に変換することは、矩形の高さ、矩形の幅、矩形のスキュー角、および/または矩形の中心位置を求めることを含み得る。たとえば、このような変換は、目標矩形の幅を、四角形 4 0 0 の上辺の幅と下辺の幅の平均として定めること、目標矩形の高さを、四角形 4 0 0 の左辺の高さと右辺の高さの平均として定めること、目標矩形の中心を、画像内における矩形の所望の配置に応じて定めること、および、目標矩形のスキューの角度を、たとえば、ユーザからの文書のデジタル表現のスキュー除去要求に応じて定めることを含み得る。

40

【0189】

（辺が直線の）四辺形を得ると、本明細書に開示するアルゴリズムは、補正後の画像におけるさまざまなオブジェクト要素の画素位置をスキャンされた画像のそれぞれのオブジェクト要素の対応する画素位置と比較することにより測定される誤差が好都合に非常に少ない、四辺形を目標の矩形に変換する射影補正要素を開始することができる。好ましい実施形態において、残余誤差は、1 インチ当たりのドット数 (DPI) が約 5 0 0 の解像度

50

の画像において、約 5 画素以下である。これは、オブジェクトの「補正後の画像」上において、オブジェクトのスキャン画像内の同じ画素の対応する位置からいずれの方向にも 6 画素以上離れている画素がないことに相当する。

【0190】

射影補正

あるアプローチにおいて、本明細書に記載の射影補正は、本質的に、上記湾曲補正アルゴリズムによって生成された四辺形を、真の矩形に変換することを含む。好ましくは、この矩形は、（当業者が本明細書を本で理解するように、たとえば従来のフラットベッドスキャナ、多機能装置でオブジェクトをスキャンする場合の取込み角度のような、オブジェクトに対して垂直な角度の）特定の視点から見たときの、デジタル画像に取込まれたオブジェクトの実際の寸法、アスペクト比等を実質的に表わす。

10

【0191】

さまざまな取込み角度および関連する射影効果が図 6 A ~ 図 6 D に概略的に示される。

いくつかのアプローチにおいて、射影補正は、4 点アルゴリズム等のアルゴリズムを画像データに適用することを含み得る。さまざまな実施形態において、このようなアルゴリズムは、以下の仮定 1) ~ 5) のうちの 1 つ以上に依拠するまたはしない場合がある。1) 取込み角度の視点から、3 D オブジェクトの厚さはゼロであり、取込まれた 3 D オブジェクトのサイズは幅寸法および高さ寸法各々に沿ってゼロでない。2) 幅寸法と高さ寸法のアスペクト比がわかっている。アスペクト比の値は、正確にわかっている必要はなく、3 D 再構成誤差に影響するかもしれない小さな測定誤差は許容し得る。好ましい実施形態において、約 10 % 未満の誤差によって特徴付けられる測定値は許容可能である（たとえば、書簡、法律文書、A 4、A 5、運転免許証、クレジットカード、領収書、名刺等の所定の種類の文書に対応するアスペクト比であり、当業者はこのことを本明細書を読んで理解するであろう）。3) 取込まれた画像における 4 つの角の画素の 2 D 画素位置は、本明細書に記載のページ検出で使用されるような画像分割技術を用いて推定可能である。4) 取込まれた画像および基準画像（オブジェクトの「リアルワールド座標」に対応する）におけるオブジェクトの角の位置は、基準画像および取込まれた画像における 4 つの画素 / 角の予め確立された対応関係によって説明される。5) 3 D 再構成により、たとえば同じ視点からオブジェクトの 2 D 表現をスキャンすることによって生成されるオブジェクトに対して垂直な取込み角度を用いて取込まれた 500 DPI 画像によって観察されたような、実際のオブジェクトの特定の視点から観察されたのと実質的に同一の、再構成画像の画素位置が得られる。

20

30

【0192】

ある実施形態において、本明細書に記載のページ検出アルゴリズムを利用して、文書の角の 2 D 画素位置を推定してもよい。（わずかに湾曲しているかもしれない）ページの四角形（たとえば図 4 に示される四角形 400）の角を見出すために、発見された文書の 4 つの辺の交点を計算してもよい。

【0193】

別の実施形態において、特に基準画像と取込まれた画像の画素座標 / 位置（特に角の座標 / 位置）の対応関係について、基準画像におけるオブジェクト（たとえば文書）の左上の角、左下の角、右下の角、および右上の角の座標は、取込まれた画像のオブジェクトの角の座標 / 位置にそれぞれ対応することが好ましい。対応する角と角の正確な関係の判断において、テキストおよび / または画像の特徴を、判断における基準点として用いてもよい。たとえば、ある実施形態において、テキストの向きおよび文書のアスペクト比を用いて、取込まれた文書の向きを判断してもよい。

40

【0194】

専ら例示を目的として、以下の説明は、運転免許証のデジタル画像に対して実施される透視補正のある実施形態を示す。運転免許証の基準画像は、好ましくは、スキャナ、多機能プリンタ、または当該技術で周知の他の装置を用いて、取込んだ画像にスキューまたは歪みを導入しないように、取込めばよい。基準画像は、好ましくは、スキャナを用いて、

50

既知の解像度、最も好ましくはおよそ500DPIの解像度で、既知の取込み角度、最も好ましくは文書に対して垂直の角度で（たとえば図6Bの先に示した90度の取込み角度で）取込まれている。

【0195】

基準画像を用い、好ましくはオブジェクトの「真の」形状、サイズ、寸法、テクスチャ等に関する事前の知識で補って、このオブジェクトの再構成を、再構成された画像において表現されたオブジェクトが基準画像と同一またはほぼ同一の特徴を有するように、行なうことが可能である。いくつかの実施形態において、再構成された画像の縮尺と基準画像の縮尺は、異なっているが比例する場合がある。このような場合、再構成された画像に対してスケーリング動作を適用することにより、縮尺の違いをなくして、再構成されたオブジェクトが、基準画像に示されるオブジェクトと同一またはほぼ同一の特徴を有するようにしてもよい。

【0196】

ユーザは、運転免許証の画像を、モバイル装置を用いて取込むときに、急な取込み角度（たとえば法線から約30度以上ずれた取込み角度）で取込む可能性がある。結果として、取込まれた画像における運転免許証の表現は、3D透視歪みによって特徴付けられた画像となり、その結果、実質的に矩形の文書の形状が台形に見えることになる。法線から30度を超える角度ずれた取込み角度を用いて観察される場合のように、極端なケースでは、たとえば取込み装置から最も遠いエッジの長さが取込み装置に最も近いエッジの長さよりも短く見える場合がある。当業者はこのことを本明細書を読んで理解するであろう。ある実施形態において、遠い側のエッジの長さは著しく短く見える場合があり、特に急な取込み角度（たとえば法線から30度以上ずれた角度）を用いたときは、近い側のエッジの長さのわずか50%に見える場合がある。

【0197】

さまざまな実施形態において、3D再構成は、好ましくは、元の画像を再構成するときに導入される歪みを最小にして、取込まれたオブジェクト／文書の矩形表現を生成する。

【0198】

ある実施形態において、透視補正は、画像をモバイル装置を用いて取込むことと、取込んだ画像の中の4つの点、画素等を識別することを含み得るものであり、この点／画素は各々、検出された運転免許証（またはその他の四角形の文書）の、角の可能性がある部分に対応し、透視補正はさらに、4つの角の画素に少なくとも一部基づいて3D変換を構成することを含み得る。これら4つの画素の位置は、画像の境界線を形成する四辺の多角形を分割しようとするときに推定することができる。

【0199】

平面ホモグラフィ／射影変換は、2つの平面間の非特異線形関係である。この場合、ホモグラフィ変換は、取込まれた画像と画像平面との間の、4つの角の画素／位置の線形マッピングを規定する。いくつかのアプローチにおいて、カメラパラメータの計算は、等式（1）に示されるようなホモグラフィ（homography）変換Hの推定を利用し得る。

【0200】

【数1】

$$\lambda \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \underbrace{\begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix}}_{\text{homography } H} \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix} \quad (1)$$

【0201】

上記式（1）において、

・ は、「基準」または「リアルワールド」座標系（たとえば上記図6Bまたは図7に示されるような基準画像から得られた座標系）における位置（X，Y，Z）の焦点深度である。言い換えると、は、基準座標系における点（X，Y，Z）と取込み装置との間の直

10

20

30

40

50

線距離と考えることができる。

- ・ (x, y, z) は、取込まれた画像における所与の画素位置の座標である。
- ・ H は、要素 h_{ij} を有する (3×3) 行列であり、 i および j はそれぞれ対応する行と列の指数を定める。

【0202】

あるアプローチにおいて、等式 1 に示される (x, y) 座標および (X, Y) 座標は、取込まれた画像平面および基準画像（たとえば図 6 B に示される）におけるそれぞれの点の座標に対応する。 z 座標は 0 に設定され、これは、各々に示されるオブジェクトが厚さゼロの 1 つの（たとえば $X - Y$ ）平面に沿っているという仮定に対応する。ある実施形態において、上記計算から等式 1 の z 値を省略してもよい。なぜなら、 z 値はホモグラフィ

10

【0203】

よって、ホモグラフィ H は、4 点の対応

【0204】

【数 2】

$$p_i \leftrightarrow P'_i$$

【0205】

を検出することによって推定することができる。 $p_i = (x_i, y_i, 1)^T$ は、取込まれた画像平面における 4 つの角の位置であり、 $P'_i = (X_i, Y_i, 1)^T$ は、対応する 4 つの角の点の座標であり、 i は、以下の説明における 1 から n の範囲の点指数値である。先に述べた表記を用いて、等式 (1) を下記の等式 (2) に示すように記述することができる。

20

【0206】

【数 3】

$$\lambda p_i = H P'_i. (2)$$

【0207】

ある実施形態において、スケーリングファクタを除くために、等式 (3) に示されるように等式 (2) の各項のクロス積を計算することが可能である。

30

【0208】

【数 4】

$$p_i \times (\lambda p_i) = p_i \times (H P'_i), (3)$$

【0209】

$p_i \times p_i = 0_3$ なので、等式 (3) は以下の等式 (4) に示すように記述することができる。

【0210】

【数 5】

$$p_i \times H P'_i = 0_3. (4)$$

40

【0211】

したがって、行列の積 $H P'_i$ は、等式 (5) のように表わすことができる。

【0212】

【数 6】

$$H P'_i = \begin{bmatrix} h^{1T} P'_i \\ h^{2T} P'_i \\ h^{3T} P'_i \end{bmatrix}, (5)$$

【0213】

50

式中、 h^{mT} は、 H の m 番目の行の転置である（たとえば、 h^{1T} は H の 1 番目の行の転置、 h^{2T} は H の 2 番目の行の転置等）。したがって、等式（４）は次のように書き直すことが可能である。

【 0 2 1 4 】

【 数 7 】

$$p_i \times H P_i' = \begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} \times \begin{bmatrix} h^{1T} P_i' \\ h^{2T} P_i' \\ h^{3T} P_i' \end{bmatrix} = \begin{bmatrix} y_i h^{3T} P_i' - h^{2T} P_i' \\ h^{1T} P_i' - x_i h^{3T} P_i' \\ x_i h^{2T} P_i' - y_i h^{1T} P_i' \end{bmatrix} = 0_3. (6)$$

【 0 2 1 5 】

特に、等式（６）は h^{mT} において線形であり、 $h^{mT} P_i = P_i^T h^m$ である。したがって、等式（６）は以下の等式（７）に示すように書き直すことができる。

【 0 2 1 6 】

【 数 8 】

$$\begin{bmatrix} 0_3^T & -P_i'^T & y_i P_i'^T \\ P_i'^T & 0_3^T & -x_i P_i'^T \\ -y_i P_i'^T & x_i P_i'^T & 0_3^T \end{bmatrix} \begin{bmatrix} h^1 \\ h^2 \\ h^3 \end{bmatrix} = 0_9. (7)$$

【 0 2 1 7 】

なお、等式（７）に示される行列の行は線形的に独立していない。たとえば、ある実施形態において、第 3 行は、 $-x_i \times$ 第 1 行と $-y_i \times$ 第 2 行の和である。したがって、各点对応において、等式（７）は線形的に独立した 2 つの等式を提供する。好ましくは、2 つの第 1 行を H を解くために使用する。ホモグラフィ変換は同種の座標を用いて記述されるので、ある実施形態では、ホモグラフィ H を 8 つのパラメータに加えて同種のスケールファクタ（9 番目のパラメータとみなし得る）を用いて規定することができる。このような実施形態において、8 つの等式を提供する少なくとも 4 つの点对応を用いて、ホモグラフィを計算することができる。実際、ある代表的な実施形態に従うと、よりロバストな結果（たとえば相対画素位置の誤差が少ない）が得られる優決定系線形系が得られるように、対応関係の数は多い方が好ましい。 H をベクトル形態 $h = [h_{11}, h_{12}, h_{13}, h_{21}, h_{22}, h_{23}, h_{31}, h_{32}, h_{33}]^T$ に書き換えることにより、 n 対の点对応によって、等式（８）によって表わされる $2n \times 9$ の線形系を構成することができる。

【 0 2 1 8 】

【 数 9 】

$$\underbrace{\begin{pmatrix} 0 & 0 & 0 & -X_1 & -Y_1 & -1 & y_1 X_1 & y_1 X_1 & y_1 \\ X_1 & Y_1 & 1 & 0 & 0 & 0 & -x_1 X_1 & -x_1 Y_1 & -x_1 \\ 0 & 0 & 0 & -X_2 & -Y_2 & -1 & y_2 X_2 & y_2 X_2 & y_2 \\ X_2 & Y_2 & 1 & 0 & 0 & 0 & -x_2 X_2 & -x_2 Y_2 & -x_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & -X_n & -Y_n & -1 & y_n X_n & y_n X_n & y_n \\ X_n & Y_n & 1 & 0 & 0 & 0 & -x_n X_n & -x_n Y_n & -x_n \end{pmatrix}}_C \begin{pmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \\ h_{33} \end{pmatrix} = 0_9. (8)$$

【 0 2 1 9 】

等式 8 に示されるように、最初の 2 行は、座標 X 、 Y 、 x 、 y の添え字の値によって示される、第 1 の特徴点に対応し、この場合の添え字の値は 1 である。2 番目の 2 行は、添え字の値 2 によって示される、第 2 の特徴点に対応し、最後の 2 行は、第 n の特徴点に対応する。4 点アルゴリズムにおいて、 n は 4 であり、特徴点は文書ページの 4 つの角である。

10

20

30

40

50

【 0 2 2 0 】

この線形系を解くことは、特異値分解 (Singular Value Decomposition) (SVD) の計算を含む。このような SVD は、行列を、行列積 $C = U D V^T$ の形に書き直すことに対応し、解 h は、行列 C の最小固有値の固有ベクトルに対応し、これは、ある実施形態において、固有値が降順で分類されたときの行列 V の最終列に位置し得る。

【 0 2 2 1 】

立体マシビジョンのために従来行なわれているような、2 つ以上のカメラを使用するときに本質的マトリクスを推定するための 8 点アルゴリズムで使用される典型的な行列と、行列 C とは異なることは、注目に値する。より具体的には、8 点アルゴリズムにおいて従来使用されている要素は 2 つのカメラ平面上に射影された特徴点からなるのに対し、本明細書に記載の行列 C の要素は、1 つのカメラ平面のみの上に射影された特徴点と 3 D オブジェクト上の対応する特徴点とからなる。

10

【 0 2 2 2 】

ある実施形態において、数値の不安定性を回避するために、好ましくは点对応の座標を正規化してもよい。これは、たとえば、正規化された直接線形変換 (Direct Linear Transformation) (DLT) アルゴリズムとして知られている技術を用いて実現し得る。たとえば、ある実施形態において、ホモグラフィ行列の推定後に、等式 1 を用いて、 (X, Y) の所与の値に対する各画素位置 (x, y) を計算してもよい。実際の応用における課題は、 (x, y) の値が与えられているまたは事前にわかっているときに (X, Y) を計算することを含む。等式 1 に示されるように、好ましい実施形態において、 (x, y) と (X, Y) は、対称である (すなわち、 (x, y) と (X, Y) の値を入れ替えたとき、等式 1 は有効であると言える)。この場合、「逆」ホモグラフィ行列を推定してもよく、この「逆」ホモグラフィ行列を用いて、取込まれた画像において、たとえばカメラの視野において示されるオブジェクトの対応する 2 D 座標を与えられて、3 D (すなわち「基準」または「リアルワールド」) 座標を再構成してもよい。

20

【 0 2 2 3 】

これまでの説明に基づいて、本明細書に記載の 4 点アルゴリズム (ならびに当業者が本明細書を読んで理解するであろう等価の変形および / または修正形) を実現することが可能であり、この 4 点アルゴリズムは、さまざまな実施形態で用いて、少なくとも何らかの透視歪みを特徴とするデジタル画像を、効率的かつ効果的に、このような透視歪みがない、補正されたデジタル画像に再構成することができ、補正された画像は、約 5 画素以下の画素位置誤差によって特徴付けられる。

30

【 0 2 2 4 】

さまざまな実施形態は、これに加えておよび / またはこれに代えて、上記データ、計算、結果、および / または概念を用いて、取込まれた画像、オブジェクト等に関する有益なその他の情報を得ることを含み得る。たとえば、さまざまな実施形態において、取込まれるオブジェクトと取込み装置との間の距離、取込み装置のピッチおよび / またはロール角等を求めることが可能であり、当業者はこのことを本明細書を読んで理解するであろう。

【 0 2 2 5 】

(X, Y) 値の推定後、等式 1 の式は次のように書くことができる。

40

【 0 2 2 6 】

【 数 1 0 】

$$\lambda = h_{31}X + h_{32}Y + h_{33} \quad (9)$$

【 0 2 2 7 】

したがって、ある実施形態において、3 D (すなわち「基準」または「リアルワールド」) 座標系における各点 (X, Y, Z) と取込み装置の間の距離として知られている焦点深度を、上記等式 9 を用いて計算することができる。

【 0 2 2 8 】

オブジェクトの回転行列の決定

50

3Dオブジェクトの位置の推定後、取込まれた画像の画素ごとに (X, Y) と Z_c が得られる。なお、 (X, Y) はワールド座標系における座標であり、 Z_c はカメラ座標系における点 (X, Y) までの距離である。3Dオブジェクトが剛体であると仮定する場合、ワールド座標系からカメラ座標系への回転行列を推定するためのアルゴリズムを提示する。以下の式は点 $(X, Y, 0)$ の回転および並進に適用できる。

【0229】

【数11】

$$\begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix} = R \begin{pmatrix} X \\ Y \\ 0 \end{pmatrix} + t \quad (10)$$

10

【0230】

式中、 (X_c, Y_c, Z_c) は、カメラ座標系に対する座標であり、これらは、ワールド座標系の点 (X, Y, Z) を、回転行列 R および並進ベクトル t で回転させることによって得られ、 t は (X, Y) から独立した定数である。なお、等式9を用いて先に推定したように、 Z_c の値は Z の値と同一である。

【0231】

ホモグラフィ行列 H と、固有カメラパラメータ行列 A と、それぞれ第1列ベクトルおよび第2列ベクトルである r_1 、 r_2 の関係を考慮すると、次の関係が明らかになる。

【0232】

【数12】

$$H = \sigma A(r_1, r_2, t) \quad (11)$$

20

【0233】

式中、 σ は定数であり、 A は次のように定義される固有カメラパラメータ行列である。

【0234】

【数13】

$$A = \begin{pmatrix} a & c & d \\ & b & e \\ & & 1 \end{pmatrix} \quad (12)$$

30

【0235】

式中、 a および b は、カメラの焦点距離情報からなるスケーリングファクタ、 $a = f / d_x$ 、 $b = f / d_y$ 、 f は焦点距離、 d_x 、 d_y は画像のスケーリングファクタ、 c は2つの画像軸に関するスキューパラメータ、 (d, e) は対応する主点の座標である。

【0236】

ホモグラフィ行列 H の推定後、行列 A を次のように推定することができる。

【0237】

【数14】

$$a = \sqrt{w/B_{11}}; \quad (12.1)$$

40

$$b = \sqrt{wB_{11}(B_{11}B_{22} - B_{12}^2)}; \quad (12.2)$$

$$c = -B_{12}a^2b/w; d = \frac{v v_0}{b} - B_{13}a^2/w; \quad (12.3)$$

$$v = -B_{12}a^2b/w; \quad (12.4)$$

$$e = (B_{12}B_{13} - B_{11}B_{23})/(B_{11}B_{22} - B_{12}^2); \quad (12.5)$$

$$w = B_{33} - (B_{13}^2 + e(B_{12}B_{13} - B_{11}B_{23}))/B_{11}. \quad (12.6)$$

【0238】

50

上記関係において、未知のパラメータは B_{ij} である。これらの値は以下の等式によって推定される。

【 0 2 3 9 】

【 数 1 5 】

$$\begin{pmatrix} v_{12}^t \\ (v_{11} - v_{22})^t \end{pmatrix} G = 0, \quad (12.7)$$

【 0 2 4 0 】

式中、 G は上記等式の解であり、代わりに次のように表わされる。

【 0 2 4 1 】

【 数 1 6 】

$$G = (B_{11}, B_{12}, B_{22}, B_{13}, B_{23}, B_{33})^t, \quad (12.8)$$

$$v_{ij} = (h_{i1}h_{j1}, h_{i1}h_{j2} + h_{i2}h_{j1}, h_{i2}h_{j2}, h_{i3}h_{j1} + h_{i1}h_{j3}, h_{i3}h_{j2} + h_{i2}h_{j3}, h_{i3}h_{j3})^t \quad (12.9)$$

【 0 2 4 2 】

なお、従来の 4 点アルゴリズムでは、スケーリングファクタ a 、 b を正確に推定することが可能なので、スキューファクタ c はゼロであると想定され、このことは、カメラのスキュー歪みは無視し得ることを意味する。ある実施形態では、 d および e の値がゼロである ($d = 0$, $e = 0$) と想定することがさらに有益である。

【 0 2 4 3 】

等式 (11) から、 $B = (r_1 \ r_2 \ t)$ であり、 $A^{-1} A^{-1} H = B$ である。この関係を用いると、等式 $C = (r_1 \ r_2 \ 0)$ から r_1 、 r_2 を推定する新たなアプローチが可能になり、 C の第 1 列ベクトルおよび第 2 列ベクトルは B の第 1 列ベクトルおよび第 2 列ベクトルであり、 C の第 3 列ベクトルは 0 である。

【 0 2 4 4 】

まず、行列 C を SVD (特異値分解) 法で分解すると、 $C = U V^t$ となり、 U は 3×3 の直交行列、 V は 3×3 の直交行列である。そうすると、 r_1 および r_2 は以下の等式によって推定される。

【 0 2 4 5 】

【 数 1 7 】

$$(r_1 \ r_2 \ 0) = U \begin{pmatrix} W \\ 0 \end{pmatrix} \quad (13)$$

【 0 2 4 6 】

式中、 W は、第 1 行ベクトルおよび第 2 行ベクトルがそれぞれ V^t の第 1 行ベクトルおよび第 2 行ベクトルである 2×3 行列である。上記計算において、 W は 1 であると想定する。このスケーリングファクタは、 U および W の値に影響しないので、 r_1 および r_2 の推定に影響しない。(たとえば等式 13 を用いて) r_1 および r_2 を推定した後、 R は回転行列であるという事実を利用して、 r_3 を推定することが有益であり、 r_3 は、求める符号 (1 または -1) を有する r_1 または r_2 のクロス積である。 R には 2 つの解が可能である。右手座標系を用いる一例において、 r_3 の値は r_1 および r_2 のクロス積の値である。

【 0 2 4 7 】

回転マトリクスのヨー、ピッチ、およびロールの決定

ヨー、ピッチ、およびロール (それぞれ α 、 β 、および γ で示される) は、オイラー角としても知られており、これらは、ある実施形態において、それぞれ z 、 y 、および x 軸を中心とする回転角度として定義される。このアプローチに従うと、等式 10 の回転行列 R は次のように定義できる。

【 0 2 4 8 】

10

20

30

40

50

【数 1 8】

$$R = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \quad (14)$$

【0 2 4 9】

式中、各 r は行列 R の要素である。

所与の回転行列 R から、およびパラメータを直接求めることが好都合であることが多い。ある実施形態において、ロールは、以下の等式によって推定し得る（たとえば r_{33} がゼロに等しくないとき）。

【0 2 5 0】

10

【数 1 9】

$$\gamma = \text{atan2}(r_{32}, r_{33}) \quad (15)$$

【0 2 5 1】

同様に、別のアプローチにおいて、ピッチは以下の等式によって推定し得る。

【0 2 5 2】

【数 2 0】

$$\beta = \text{atan2}(-r_{31}, \sqrt{r_{11}^2 + r_{21}^2}) \quad (16)$$

【0 2 5 3】

20

さらに他のアプローチにおいて、ヨーは以下の等式によって推定し得る（ R_{11} がゼロでないとき）。

【0 2 5 4】

【数 2 1】

$$\alpha = \text{atan2}(r_{21}, r_{11}) \quad (17)$$

【0 2 5 5】

特に、いくつかのアプローチにおいて、 r_{11} 、 r_{33} 、または

【0 2 5 6】

【数 2 2】

30

$$\sqrt{r_{11}^2 + r_{21}^2}$$

【0 2 5 7】

がゼロに近い値のとき（たとえば $0 < r_{11} < \epsilon$ 、 $0 < r_{33} < \epsilon$ 、または

【0 2 5 8】

【数 2 3】

$$0 < \sqrt{r_{11}^2 + r_{21}^2} < \epsilon$$

【0 2 5 9】

40

であり、値 ϵ が、ある実施形態では数値安定性を考慮して $0 < \epsilon < 0.01$ 等の妥当な値に設定され、特に好ましい実施形態では $\epsilon = 0.0001$ に設定される。一般的に、 ϵ の値は、全体的または部分的に、限られたコンピュータワード長等に基づいて求めることができ、当業者はこのことを本明細書を読んで理解するであろう）、これは、回転行列 R の縮退に対応し、特殊な式を用いてヨー、ピッチ、およびロールの値を推定する。

【0 2 6 0】

オブジェクトと取込み装置の間の距離の推定

さらに他の実施形態において、オブジェクトと取込み装置の間の距離を、オブジェクトサイズの知識がなくても、カメラの固有パラメータ（たとえば焦点距離、画像平面内の（ u ， v ）のスケールファクタ等）といった情報を用いて推定することが可能である。

50

【 0 2 6 1 】

あるアプローチにおいて、このアルゴリズムの必要条件を次のように要約することができる。１）装置のＡＰＩコールによって、取込まれた画像についてのカメラの焦点距離を与えこの焦点距離にアクセスすることができる（たとえば、アンドロイド装置は、ＡＰＩコールを与えて、取込まれた画像についての焦点距離情報を得る）。２） d_x および d_y のスケールファクタは等式 12 . 1 および 12 . 2 のアルゴリズムによって推定される。

【 0 2 6 2 】

このことは、ある種の装置についてのスケールファクタ d_x 、 d_y を推定することができる、個々の装置について個別にこれらを推定する必要はないことを意味する。たとえば、アップル社のiPhone（登録商標）４スマートフォンを使用する代表的なある実施形態において、上記アルゴリズムを用いることで、既知のサイズのオブジェクトを用いてスケールファクタを推定することが可能である。２つのスケーリングファクタはその後、同一の装置種類については同一であると想定すればよい。

10

【 0 2 6 3 】

例示としてのあるアプローチに従う、オブジェクトからカメラまでの距離を推定するアルゴリズムは次の通りである。以下の式において (u, v) 、 (X, Y) を正規化する。

【 0 2 6 4 】

【 数 2 4 】

$$\lambda \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = H \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix} \quad (18)$$

20

【 0 2 6 5 】

なお、等式 18 は、等式 1 の (x, y) の代わりに等式 18 では (u, v) を用いることを除いて、等式 1 と等価である。

【 0 2 6 6 】

【 数 2 5 】

$$\tilde{u} = u/L_u, \tilde{v} = v/L_v; \tilde{x} = X/L_X; \tilde{y} = Y/L_Y$$

30

【 0 2 6 7 】

であり、 L_u 、 L_v は座標 u および v の画像サイズであり、 L_X 、 L_Y は求めるべきオブジェクトサイズであると想定する。

【 0 2 6 8 】

そうすると等式 18 は次のように表わすことができる。

【 0 2 6 9 】

【数 2 6】

$$\lambda \begin{pmatrix} \tilde{u} \\ \tilde{v} \\ 1 \end{pmatrix} = \tilde{H} \begin{pmatrix} \tilde{x} \\ \tilde{y} \\ 1 \end{pmatrix} \quad (19)$$

$$\tilde{H} = \begin{pmatrix} 1/L_u & & \\ & 1/L_v & \\ & & 1 \end{pmatrix} H \begin{pmatrix} L_x & & \\ & L_y & \\ & & 1 \end{pmatrix} \quad (20)$$

10

正規化されたホモグラフィ行列 \tilde{H} は、等式 (20) によって推定することができる。

等式 11 から、

$$H = \sigma A \begin{pmatrix} r_1 & r_2 & t \end{pmatrix} \quad (21)$$

が得られる。

【0 2 7 0】

カメラの固有パラメータ行列は以下の単純な形態で想定される。

【0 2 7 1】

20

【数 2 7】

$$A = \begin{pmatrix} f/dx & c & d \\ & f/dy & e \\ & & 1 \end{pmatrix} \quad (22)$$

【0 2 7 2】

式中、f はカメラの焦点距離、dx、dy はカメラのスケーリングファクタであり、セクション [0 0 1 8 2] ~ [0 0 2 0 9] に示されるアルゴリズムによって推定される。

【0 2 7 3】

30

等式 (19)、(20)、および (21) から、次のようになる。

【0 2 7 4】

【数 2 8】

$$\sigma A \begin{pmatrix} r_1 & r_2 & t \end{pmatrix} \begin{pmatrix} L_x & & \\ & L_y & \\ & & 1 \end{pmatrix} = \tilde{\tilde{H}} \quad (23)$$

$$\tilde{\tilde{H}} = \begin{pmatrix} L_u & & \\ & L_v & \\ & & 1 \end{pmatrix} \tilde{H}$$

40

【0 2 7 5】

A がわかっているので、等式 (23) から、次のようになる。

【0 2 7 6】

【数 2 9】

$$\sigma \begin{pmatrix} r_1 & r_2 & t \end{pmatrix} \begin{pmatrix} L_x & & \\ & L_y & \\ & & 1 \end{pmatrix} = A^{-1} \tilde{\tilde{H}} \quad (24)$$

$K = A^{-1} \tilde{\tilde{H}}$ 、 $K = (k_1, k_2, k_3)$ とすると、式 (24) から、次のようになる。

$$\sigma r_1 L_x = k_1 \quad (25)$$

$$\sigma r_2 L_y = k_2 \quad (26)$$

$$\sigma t = k_3 \quad (27)$$

10

【0 2 7 7】

等式 (27) の t は、カメラに対するオブジェクトの並進ベクトルであるためである。 t の L_2 ノルム (ユークリッドノルム) は次の通りである。

【0 2 7 8】

【数 3 0】

$$\|t\| = \|k_3\| / \sigma \quad (28)$$

20

は、オブジェクトの左上の角からカメラまでの距離である。

$\|r_1\| = \|r_2\| = 1$ であるため、等式 (8) および (9) から、次のようになる。

$$L_x = \|k_1\| / \sigma \quad (29)$$

$$L_y = \|k_2\| / \sigma \quad (30)$$

【0 2 7 9】

30

等式 (29) および (30) を用いて、 X および Y 座標に沿う文書のサイズを推定することができる。このアプローチを用いると、スケーリングファクタは未知のままとなり得る。

【0 2 8 0】

上記回転行列を推定するアルゴリズムでは、スケーリングファクタは不要である。むしろ、いくつかのアプローチでは、 $\sigma = 1$ と想定するのが適切である。上記アルゴリズムを用いてロール、ピッチ、およびヨーを推定することができる。等式 (29) および (30) から、オブジェクトのアスペクト比を次のように推定することができる。

【0 2 8 1】

【数 3 1】

40

$$\text{アスペクト比} = L_x / L_y = \|k_1\| / \|k_2\| \quad (31)$$

【0 2 8 2】

想定した矩形からのピッチおよびロールの推定

実際、最も一般的なケースは、標準サイズの手紙、名刺、運転免許証およびその他の免許証等の、矩形の文書のカメラによる取込みである。カメラの焦点距離は変わらずヨーに関する知識は上記種類の文書の画像処理とは無関係なので、文書の対応する画像を矩形化するためには、文書の面に対するカメラのロールとピッチを求めるだけでよい。

【0 2 8 3】

このアルゴリズムの概念は、単純に、何らかの相対的なピッチ - ロールの組合わせに対

50

し、画像において検出された四角形に対応する文書のオブジェクト座標を、(縮尺、回転、およびシフトに応じて)計算できるということである。この計算されたオブジェクト座標の四角形は、ピッチおよびロールの正確な値が使用されたときは90度の角度を特徴とし、ずれは、4つの角度差の二乗の和によって特徴付けることができる。この基準は有用である。なぜなら、これはスムーズであり個々の大きなずれに対して効果的にペナルティを課すからである。

【0284】

この基準に基づく傾斜降下処置により、わずかミリ秒単位で適切なピッチ・ロール対を見出すことができる。これは、画像内の四角形が正確に求められた場合について実験で確認されている。このアプローチは、ゼロに等しいヨーとオブジェクトまでの距離として任意の固定値を使用する。なぜなら、これら値の変化によって、オブジェクト座標のさらに他の直交変換が加わるだけであるからである。このアプローチはまた、座標変換の計算においてカメラの既知の焦点距離を使用するが、4つの角すべてが発見されており独立した3つの角度がある場合は、同じ基準と、やや複雑度が増した傾斜降下処置を用いて、ピッチおよびロールに加えて焦点距離を推定することができる。このことは、入力された画像に、どのようなカメラで撮影されたかに関する情報があってもなくても、サーバベースの処理には有益であろう。

10

【0285】

興味深いことに、ページ検出が誤りであったとき、最適なピッチ・ロール対であっても相当大きな残余の角度誤差(1度以上)が残る、または、少なくとも、ページが自身に平行にクロップされたばかりであれば、見出されたオブジェクト座標から得られるアスペクト比は、実際のアスペクト比と一致しない。

20

【0286】

加えて、このアルゴリズムは、文書の検出された辺のうちの1つの位置が疑わしいかまたは全体的に欠落しているときであっても(たとえば、文書のその辺が部分的にまたは完全に遮断されている、示されていない、または認識できないほど不鮮明である等)、適用することが可能である。所望の結果を成し遂げるには、上述のように規定された基準を修正して、2つの角度、たとえば、下辺に隣接する角度のみを、傾斜降下処置で用いることが、有益である。このようにして、このアルゴリズムをなおも用いて、偽のおよび/または検出不可能な左上の角および右上の角を有するが総の四角形から、ピッチとロールを推定し得る。

30

【0287】

一例において、画像フレームの上部に近い左辺および右辺上の任意のポイントを、左上の角と右上の角として指定することができる。最適に推定されたピッチ・ロールでは、オブジェクト座標において偽の左上の角および右上の角を等しく生成することになるが、文書はそれでも正しく矩形化されるであろう。文書の欠落している(たとえば上の)辺は再構成できる。なぜなら、この上辺は向い側の(たとえば下の)辺と実質的に平行であり隣接する(たとえば左および/または右の)辺に直交するはずだからである。

【0288】

残っている問題は、画像のコンテキスト全体において欠落している辺をどこに配置するかであり、アスペクト比がわかっているならば、欠落している辺のオフセットを上手く推定することができる、わかっているなければ、データを失うのではなく、フレームのエッジまで押し出せばよい。このアルゴリズムの変形によって、画像が、その辺に沿って文書の一部、たとえば、預金伝票を含むインボイスの下側しか含まないときに、ユーザの重要なケースを解決することができる。このような状況においては、文書の下辺、左辺、および右辺を、正しく判断し、ピッチおよびロールの推定に用いることができる。これらの角度を焦点距離とともに用いて、文書の見えている部分を矩形化することができる。

40

【0289】

よって、図7に示される方法700によって例示される、ある一般的なアプローチにおいて、再構成は、動作702において、オブジェクト、好ましくは特徴(サイズ、テクス

50

チャ、カラープロファイル等)がわかっている文書またはその他のオブジェクトの、デジタル表現を含む、デジタル画像を取込むまたは受けることを含む。

【0290】

動作704において、取込んだまたは受けた画像を解析して、オブジェクトのデジタル表現を画像背景または画像内の他のオブジェクトから分離する1本以上の境界線の位置を求める。

【0291】

動作706において、境界線を解析して、境界線のうちの1本以上に湾曲(すなわち非線形領域)があるか否か判断する。たとえば、各々が、異なるクラスの関数であり得る、一次、二次、三次、四次等の多項式などの、特定のクラスの関数に、境界線の特徴付ける多項式が適合すると判断することにより、境界線のうちの1つに湾曲があると判断してもよい。湾曲があると判断された場合、これを補正して、その長さ全体に沿って実質的に線形特徴を有する境界線を生成することが好ましい。

【0292】

判断された湾曲が補正されたら、動作708において、線形のエッジの境界線を解析および/または外装することにより、取囲む多角形、好ましくは取囲む四角形、より好ましくは取囲む平行四辺形、台形、または矩形を定める。

【0293】

動作710において、デジタル画像および/または取囲む多角形を解析することにより、透視歪みおよび/または射影効果が、取囲む多角形の中にあるか否か判断する。

【0294】

動作712において、取囲む多角形、デジタル画像等の中に透視歪みおよび/または射影効果があるという判断に応じて、透視歪みおよび/または射影効果を補正して、再構成された多角形を生成する。好ましくは、取囲む多角形は四辺形であり、再構成された多角形は矩形である。

【0295】

当然ながら、代表的な方法700の上記開示は、ロジックを実行するように構成されたシステムとして、および/または上記機能のうちのいずれかと実質的に同様の機能を実行するように構成されたコンピュータ読取可能なプログラムコード含むコンピュータプログラムプロダクトとして、実現し得る。

【0296】

同様に、本明細書に記載の、発明の概念、特徴、技術、構成要素、システム、製品等すべては、モジュール方式とみなされるべきものであって、当業者が本明細書を読んで理解するであろう適切なやり方で組合わせてもよい。特に、これら組合わせは、本明細書を読んだ当業者が相互に排他的または両立しないと認識するもののみを除いて、開示された特徴のいずれの可能な組も含む。したがって、この発明の概念のさまざまな実施形態は、個々の特徴を含めてまたは除いて、いずれかの数の開示された特徴を記載することによって、定義し得る。

【0297】

たとえば、あるアプローチにおいて、デジタル画像を再構成する方法は、四角形によって取囲まれたオブジェクトのデジタル表現を含むデジタル画像を受けるステップと、この四角形の湾曲を補正して四辺形を形成するステップと、この四辺形における射影効果を補正して矩形を形成するステップとを含む。この四辺形は平行四辺形であってもよい。湾曲を補正するステップは、四角形の1つ以上の部分を、固有の四角形座標に少なくとも一部基づいて変換するステップを含み得る。射影効果を補正するステップは、4点アルゴリズムを用いて四角形を変換するステップを含み得る。より具体的には、射影効果の補正は、特異値分解(SVD)、直接線形変換(DLT)、非線形リファインメント(non-linear refinement)(NLR)、デジタル画像と再構成された画像との点对応の座標を再構成すること、および/または、ホモグラフィ変換(H)の推定のうちの、1つ以上を含み得る。Hを推定することは、

1 つ以上の点对応

【 0 2 9 8 】

【 数 3 2 】

$p_i \leftrightarrow P'_i$ を検出するステップを含み、 $p_i = (x_i, y_i, 1)^T$ であり、

【 0 2 9 9 】

および / または、

4 つの点对応を検出するステップを含み、

【 0 3 0 0 】

【 数 3 3 】

10

各点对応 $p_i \leftrightarrow P'_i$ は、四角形の角位置 p_i と前記矩形の対応する角位置 P'_i に対応する。

【 0 3 0 1 】

任意で、デジタル画像は、オブジェクトに対する法線から約 30 度以上ずれた取込み角度によって特徴付けられる。好ましくは、 H は、式 $p_i = H P_i$ を満たし、 H = 三次元基準座標位置 (X, Y, Z) と、取込み装置の対応する基準座標位置との間の焦点距離であり、 $p_i = (x_i, y_i, 1)^T$ であり、 $P_i = (X_i, Y_i, 1)^T$ であり、 i は、対応する画像における点 p または P のそれぞれの指数特徴である。この方法はまた、再構成されたデジタル画像を出力するステップを含み得る。この再構成されたデジタル画像は、オブジェクトに対して垂直である取込み角度を用いて取込まれたオブジェクトのスキャンされた 500 DPI 画像と比較して約 5 画素以下の画素位置誤差によって特徴付けられる。四角形の辺は、二次多項式、または二次よりも高次の多項式にであってもよい多項式によって特徴付けることができる。最も好ましくは、オブジェクトは、既知の高さと幅の比率によって特徴付けられる文書である。

20

【 0 3 0 2 】

代表的なシステムは、1 つ以上のプロセッサと、画像取込み装置と、上記段落に記載の動作のうちの 1 つ以上を実行するように、かつ、任意の数の具体的な特徴を利用するように構成された、対応する機能的論理とを含み得る。上記具体的な特徴は、オブジェクトが矩形の文書で既知の高さと幅の比率、 p_i 、 P'_i 等を有することであり、当業者はこのことを本明細書を読んで理解するであろう。

30

【 0 3 0 3 】

代表的なコンピュータプログラムプロダクトは、上記段落に記載の動作のうちの 1 つ以上を実行し、かつ、任意の数の具体的な特徴を利用する、コンピュータが実装された、コンピュータ読取可能な記録媒体を含み得る。上記具体的な特徴は、オブジェクトが矩形の文書で既知の高さと幅の比率、 p_i 、 P'_i 等を有することであり、当業者はこのことを本明細書を読んで理解するであろう。

【 0 3 0 4 】

以上さまざまな実施形態が説明されているが、これら実施形態は限定ではなく専ら例示のために示されていることが理解されるはずである。よって、本発明の実施形態の広さおよび範囲は、上記代表的な実施形態のいずれによっても限定されてはならず、以下の請求項およびこれらの均等物に従ってのみ定義されるべきものである。

40

【図 3 A】

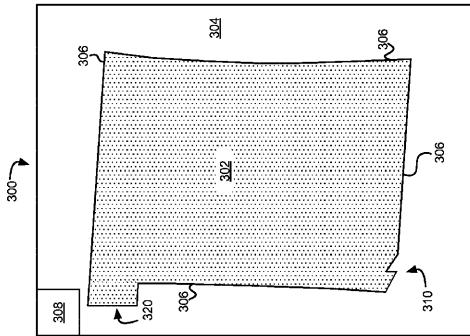


FIG. 3A

【図 3 B】

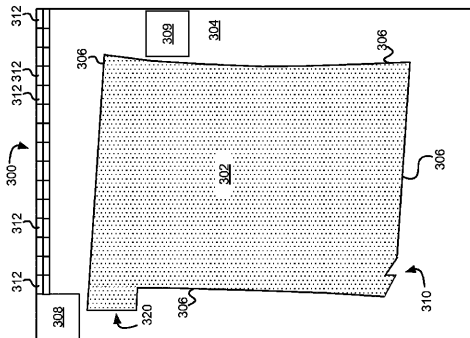


FIG. 3B

【図 3 C】

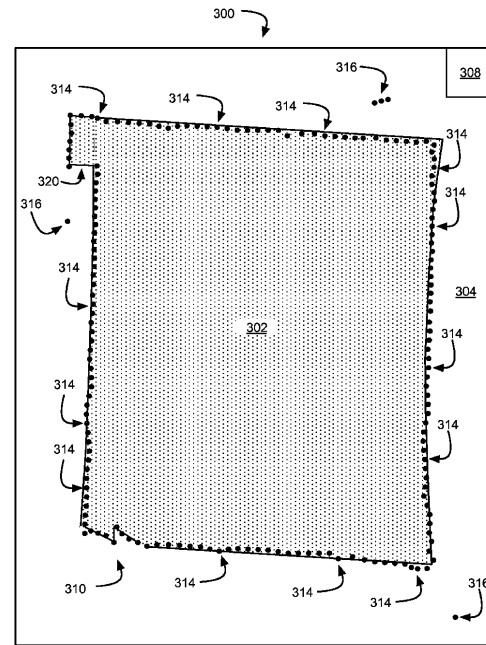


FIG. 3C

【図 3 D】

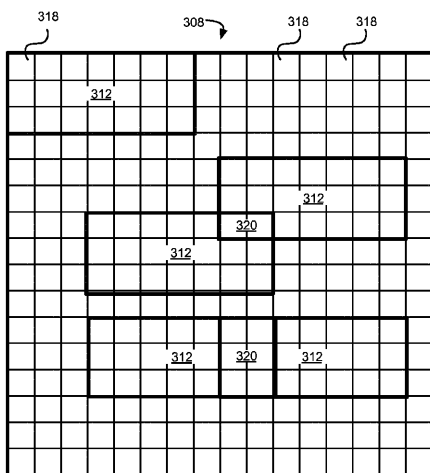


FIG. 3D

【図 4】

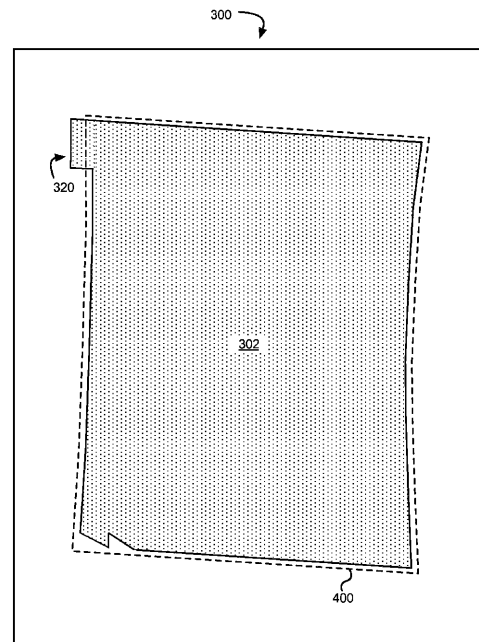


FIG. 4

【図 5 A】

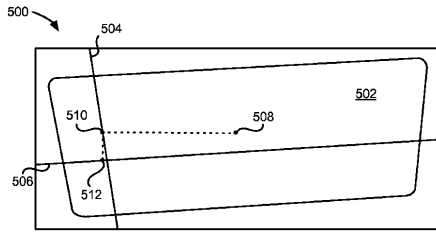


FIG. 5A

【図 5 B】

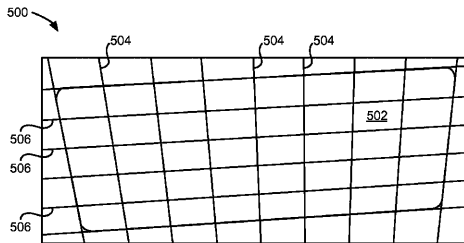


FIG. 5B

【図 6 A】

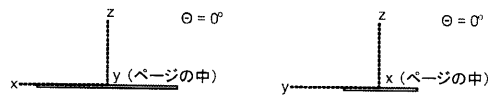


FIG. 6A

【図 6 B】

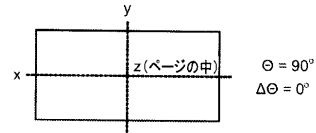


FIG. 6B

【図 6 C】

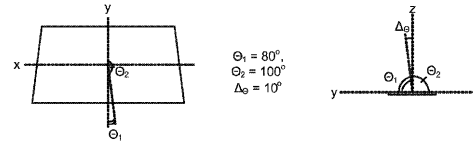


FIG. 6C

【図 6 D】

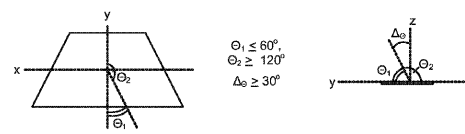


FIG. 6D

【図 7】

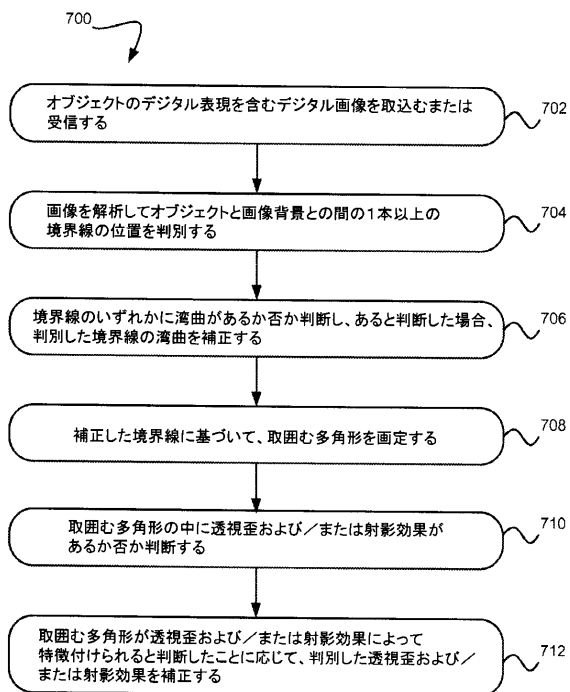


FIG. 7

【図 1】

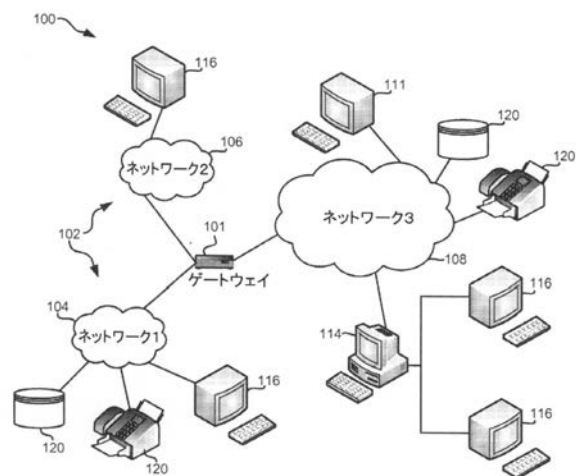


FIG. 1

【図 2】

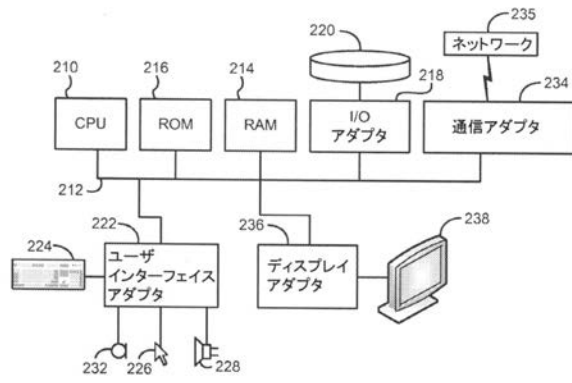


FIG. 2

【国際調査報告】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US2014/057065																																										
A. CLASSIFICATION OF SUBJECT MATTER IPC(8) - G06F 15/00 (2014.01) CPC - G06K 15/02 (2014.09) 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) - G06F 15/00; G06K 1/00; G01T 5/00 (2014.01) USPC - 358/1.9, 3.26, 3.27, 520, 522, 537, 538 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched CPC - G06K 15/02; H04N 1/00002, 1/4092 (2014.09) (keyword delimited) Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) Orbit, Google Patents, ProQuest Search terms used: digital image, curvature, tetragon, rectangle, correction, parallelogram																																												
C. DOCUMENTS CONSIDERED TO BE RELEVANT <table border="1"> <thead> <tr> <th>Category*</th> <th>Citation of document, with indication, where appropriate, of the relevant passages</th> <th>Relevant to claim No.</th> </tr> </thead> <tbody> <tr> <td>Y</td> <td>US 2013/0182973 A1 (MACCIOLA et al) 18 July 2013 (18.07.2013) entire document</td> <td>1-19</td> </tr> <tr> <td>Y</td> <td>US 2011/0279456 A1 (HIRANUMA et al) 17 November 2011 (17.11.2011) entire document</td> <td>1-19</td> </tr> <tr> <td>Y</td> <td>US 2009/0225180 A1 (MARUYAMA et al) 10 September 2009 (10.09.2009) entire document</td> <td>2</td> </tr> <tr> <td>Y</td> <td>US 2011/0090337 A1 (KLOMP et al) 21 April 2011 (21.04.2011) entire document</td> <td>4</td> </tr> <tr> <td>Y</td> <td>US 2008/0068452 A1 (NAKAO et al) 20 March 2008 (20.03.2008) entire document</td> <td>5</td> </tr> <tr> <td>Y</td> <td>US 6,729,733 B1 (RASKAR et al) 04 May 2004 (04.05.2004) entire document</td> <td>6</td> </tr> <tr> <td>Y</td> <td>US 2009/0324062 A1 (LIM et al) 31 December 2009 (31.12.2009) entire document</td> <td>7</td> </tr> <tr> <td>Y</td> <td>US 2012/0236019 A1 (OH et al) 20 September 2012 (20.09.2012) entire document</td> <td>8-11</td> </tr> <tr> <td>Y</td> <td>US 2012/0301011 A1 (GRZECHNIK) 29 November 2012 (29.11.2012) entire document</td> <td>11</td> </tr> <tr> <td>Y</td> <td>US 2012/0300020 A1 (ARTH et al) 29 November 2012 (29.11.2012) entire document</td> <td>12</td> </tr> <tr> <td>Y</td> <td>US 4,651,287 A (TSAO) 17 March 1987 (17.03.1987) entire document</td> <td>13</td> </tr> <tr> <td>A</td> <td>US 2004/0019274 A1 (GALLOWAY et al) 29 January 2004 (29.01.2004) entire document</td> <td>1-19</td> </tr> <tr> <td>A</td> <td>US 2008/0298718 A1 (LIU et al) 04 December 2008 (04.12.2008) entire document</td> <td>1-19</td> </tr> </tbody> </table>			Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	Y	US 2013/0182973 A1 (MACCIOLA et al) 18 July 2013 (18.07.2013) entire document	1-19	Y	US 2011/0279456 A1 (HIRANUMA et al) 17 November 2011 (17.11.2011) entire document	1-19	Y	US 2009/0225180 A1 (MARUYAMA et al) 10 September 2009 (10.09.2009) entire document	2	Y	US 2011/0090337 A1 (KLOMP et al) 21 April 2011 (21.04.2011) entire document	4	Y	US 2008/0068452 A1 (NAKAO et al) 20 March 2008 (20.03.2008) entire document	5	Y	US 6,729,733 B1 (RASKAR et al) 04 May 2004 (04.05.2004) entire document	6	Y	US 2009/0324062 A1 (LIM et al) 31 December 2009 (31.12.2009) entire document	7	Y	US 2012/0236019 A1 (OH et al) 20 September 2012 (20.09.2012) entire document	8-11	Y	US 2012/0301011 A1 (GRZECHNIK) 29 November 2012 (29.11.2012) entire document	11	Y	US 2012/0300020 A1 (ARTH et al) 29 November 2012 (29.11.2012) entire document	12	Y	US 4,651,287 A (TSAO) 17 March 1987 (17.03.1987) entire document	13	A	US 2004/0019274 A1 (GALLOWAY et al) 29 January 2004 (29.01.2004) entire document	1-19	A	US 2008/0298718 A1 (LIU et al) 04 December 2008 (04.12.2008) entire document	1-19
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.																																										
Y	US 2013/0182973 A1 (MACCIOLA et al) 18 July 2013 (18.07.2013) entire document	1-19																																										
Y	US 2011/0279456 A1 (HIRANUMA et al) 17 November 2011 (17.11.2011) entire document	1-19																																										
Y	US 2009/0225180 A1 (MARUYAMA et al) 10 September 2009 (10.09.2009) entire document	2																																										
Y	US 2011/0090337 A1 (KLOMP et al) 21 April 2011 (21.04.2011) entire document	4																																										
Y	US 2008/0068452 A1 (NAKAO et al) 20 March 2008 (20.03.2008) entire document	5																																										
Y	US 6,729,733 B1 (RASKAR et al) 04 May 2004 (04.05.2004) entire document	6																																										
Y	US 2009/0324062 A1 (LIM et al) 31 December 2009 (31.12.2009) entire document	7																																										
Y	US 2012/0236019 A1 (OH et al) 20 September 2012 (20.09.2012) entire document	8-11																																										
Y	US 2012/0301011 A1 (GRZECHNIK) 29 November 2012 (29.11.2012) entire document	11																																										
Y	US 2012/0300020 A1 (ARTH et al) 29 November 2012 (29.11.2012) entire document	12																																										
Y	US 4,651,287 A (TSAO) 17 March 1987 (17.03.1987) entire document	13																																										
A	US 2004/0019274 A1 (GALLOWAY et al) 29 January 2004 (29.01.2004) entire document	1-19																																										
A	US 2008/0298718 A1 (LIU et al) 04 December 2008 (04.12.2008) entire document	1-19																																										
<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 "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family																																												
Date of the actual completion of the international search 18 November 2014		Date of mailing of the international search report 30 DEC 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, ST, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, RU, TJ, TM), EP(AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US

(72)発明者 シュストロヴィッチ, アレクサンダー

アメリカ合衆国、 1 4 5 3 4 ニュー・ヨーク州、ピッツフォード、ストーントン・ドライブ、
3 9

(72)発明者 スラッシャー, クリストファー・ダブリュ

アメリカ合衆国、 1 4 6 2 4 ニュー・ヨーク州、ロチェスター、ウッドラフ・グレン、 1 3

(72)発明者 アントラップ, ジャン

アメリカ合衆国、 2 0 8 1 5 メリーランド州、チェビー・チェース、ノース・パーク・アベニュー、
4 6 1 5、ナンバー・ 1 2 1 1

Fターム(参考) 5B057 AA11 CA08 CA12 CA16 CB08 CB12 CB16 CD01 DC08 DC09
5C076 AA40