

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7329331号
(P7329331)

(45)発行日 令和5年8月18日(2023.8.18)

(24)登録日 令和5年8月9日(2023.8.9)

(51)国際特許分類

F I

G 0 6 V 30/14 (2022.01)

G 0 6 V 30/14 3 4 0 K

G 0 6 V 30/412 (2022.01)

G 0 6 V 30/14 3 6 0 C

G 0 6 V 30/412

請求項の数 8 (全13頁)

(21)出願番号	特願2019-17157(P2019-17157)	(73)特許権者	599001389
(22)出願日	平成31年2月1日(2019.2.1)		株式会社プリマジェスト
(65)公開番号	特開2020-126326(P2020-126326 A)		神奈川県川崎市幸区堀川町 5 8 0 番地
(43)公開日	令和2年8月20日(2020.8.20)	(74)代理人	ソリッドスクエア東館 1 2 階
審査請求日	令和4年1月25日(2022.1.25)	(74)代理人	100079108
			弁理士 稲葉 良幸
		(74)代理人	100109346
			弁理士 大貫 敏史
		(74)代理人	100117189
			弁理士 江口 昭彦
		(74)代理人	100134120
			弁理士 内藤 和彦
		(72)発明者	山井 淳彦
			神奈川県川崎市幸区堀川町 5 8 0 番地
			ソリッドスクエア東館 1 2 階 株式会社
			最終頁に続く

(54)【発明の名称】 帳票に含まれる文字列を認識するための装置、方法、及びプログラム

(57)【特許請求の範囲】

【請求項 1】

帳票に含まれる文字列を認識するための文字列認識装置であって、
前記帳票のカラー画像を取得するように構成されたカラー画像取得部と、
取得した前記カラー画像において、前記帳票に含まれる所定の項目の名称に対応する第 1 文字列を検出するように構成された項目名称検出部であって、前記第 1 文字列は、特定の色により表示される、項目名称検出部と、
前記カラー画像から少なくとも前記特定の色が消去されたドロップアウト画像を生成するように構成されたドロップアウト画像生成部と、
生成した前記ドロップアウト画像において、前記カラー画像において検出された前記第 1 文字列の位置に少なくとも基づいて、前記所定の項目の値に対応する第 2 文字列を検出するように構成された項目値検出部と、を備える、
文字列認識装置。

【請求項 2】

前記ドロップアウト画像生成部は、黒色を除く色が消去されたドロップアウト画像を生成するように構成されている、
請求項 1 の文字列認識装置。

【請求項 3】

前記項目値検出部は、前記カラー画像における前記第 1 文字列との間の相対的な位置関係が所定の関係にある前記ドロップアウト画像における文字列を前記第 2 文字列として検

出するように構成されている、

請求項 1 又は 2 の文字列認識装置。

【請求項 4】

前記項目値検出部は、前記カラー画像における前記第 1 文字列の位置に基づいて設定される前記ドロップアウト画像における所定の領域に含まれる文字列を前記第 2 文字列として検出するように構成されている、

請求項 3 の文字列認識装置。

【請求項 5】

前記所定の領域は、前記カラー画像における前記第 1 文字列の右側及び / 又は下側の領域が少なくとも含まれるように構成されている、

請求項 4 の文字列認識装置。

【請求項 6】

前記所定の項目は、金額に関する項目であり、

前記項目名称検出部は、前記第 1 文字列に加えて、通貨に関する所定の文字を検出するように構成されており、

前記項目値検出部は、前記カラー画像において検出された前記第 1 文字列及び前記所定の文字の位置に少なくとも基づいて前記第 2 文字列を検出するように構成されている、

請求項 1 ないし 5 の何れか一項の文字列認識装置。

【請求項 7】

1 又は複数のコンピュータによって実行され、帳票に含まれる文字列を認識するための文字列認識方法であって、

前記帳票のカラー画像を取得する工程と、

取得した前記カラー画像において、前記帳票に含まれる所定の項目の名称に対応する第 1 文字列を検出する工程であって、前記第 1 文字列は、特定の色により表示される、工程と、

前記カラー画像から少なくとも前記特定の色が消去されたドロップアウト画像を生成する工程と、

生成した前記ドロップアウト画像において、前記カラー画像において検出された前記第 1 文字列の位置に少なくとも基づいて、前記所定の項目の値に対応する第 2 文字列を検出する工程と、を備える、

文字列認識方法。

【請求項 8】

帳票に含まれる文字列を認識するための文字列認識プログラムであって、コンピュータに、

前記帳票のカラー画像を取得する処理と、

取得した前記カラー画像において、前記帳票に含まれる所定の項目の名称に対応する第 1 文字列を検出する処理であって、前記第 1 文字列は、特定の色により表示される、処理と、

前記カラー画像から少なくとも前記特定の色が消去されたドロップアウト画像を生成する処理と、

生成した前記ドロップアウト画像において、前記カラー画像において検出された前記第 1 文字列の位置に少なくとも基づいて、前記所定の項目の値に対応する第 2 文字列を検出する処理と、を実行させる、

文字列認識プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、帳票に含まれる文字列を認識するための装置、方法、及びプログラムに関するものである。

【背景技術】

10

20

30

40

50

【 0 0 0 2 】

従来、例えば、銀行や郵便局等の金融機関に持ち込まれる帳票を処理するために、当該帳票の種類を特定し、さらに、当該帳票に含まれる項目（対応する処理の依頼者、口座番号、及び各種の金額等）の種類及び位置を特定することが行われている。例えば、下記特許文献 1 は、事前に帳票のサイズや罫線情報、項目の位置等を様式情報として登録しておき、帳票の読取画像と登録済みの様式情報とを照合することを開示している。

【先行技術文献】

【特許文献】

【 0 0 0 3 】

【文献】特開 2 0 0 5 - 2 4 2 7 8 6 号公報

10

【発明の概要】

【発明が解決しようとする課題】

【 0 0 0 4 】

しかしながら、帳票の様式は、典型的には、その作成元（例えば、収納企業や地方自治体等）毎に異なるので、様式の数膨大（数千又は数万等）となり、これらの様式情報の全てを登録することが非現実的となってしまう場合がある。したがって、様式情報の事前登録なしで帳票の処理が可能な仕組みの実現が望まれる。

【 0 0 0 5 】

本発明は、様式情報の事前登録なしで行われる帳票の処理を支援することを目的の 1 つとする。本発明の他の目的は、本明細書全体を参照することにより明らかとなる。

20

【課題を解決するための手段】

【 0 0 0 6 】

本発明の実施形態に係る文字列認識装置は、帳票に含まれる文字列を認識するための文字列認識装置であって、前記帳票のカラー画像を取得するように構成されたカラー画像取得部と、取得した前記カラー画像において、前記帳票に含まれる所定の項目の名称に対応する第 1 文字列を検出するように構成された項目名称検出部と、前記カラー画像から少なくとも特定の色が消去されたドロップアウト画像を生成するように構成されたドロップアウト画像生成部と、生成した前記ドロップアウト画像において、前記カラー画像において検出された前記第 1 文字列の位置に少なくとも基づいて、前記所定の項目の値に対応する第 2 文字列を検出するように構成された項目値検出部と、を備える。

30

【 0 0 0 7 】

この構成によれば、帳票に含まれる所定の項目の名称に対応する第 1 文字列を検出し、当該第 1 文字列の位置に基づいて当該所定の項目の値に対応する第 2 文字列を検出するから、帳票の様式情報の事前登録なしに、所定の項目の値（第 2 文字列）の認識が可能となる。さらに、帳票のカラー画像から特定の色が消去されたドロップアウト画像において所定の項目の値（第 2 文字列）を検出するから、例えば、カラー画像において所定の項目の値が特定の色（罫線 / 罫線等）に重なっている場合であっても（こうした罫線 / 罫線等との接触は不読 / 誤読の原因となる。）、ドロップアウト画像においては当該特定の色（罫線 / 罫線）が消去され、所定の項目の値の不読 / 誤読が抑制される。

【 0 0 0 8 】

40

また、上記装置において、前記ドロップアウト画像生成部は、黒色を除く色が消去されたドロップアウト画像を生成するように構成することもできる。こうすれば、黒色以外の様々な色の罫線 / 罫線を有する様々な帳票における所定の項目の値の不読 / 誤読を抑制することができる。

【 0 0 0 9 】

また、上記装置において、前記項目値検出部は、前記カラー画像における前記第 1 文字列との間の相対的な位置関係が所定の関係にある前記ドロップアウト画像における文字列を前記第 2 文字列として検出するように構成することもできる。こうすれば、第 1 文字列との間の相対的な位置関係に基づいて第 2 文字列を検出することが可能となる。

【 0 0 1 0 】

50

また、上記装置において、前記項目値検出部は、前記カラー画像における前記第1文字列の位置に基づいて設定される前記ドロップアウト画像における所定の領域に含まれる文字列を前記第2文字列として検出するように構成することもできる。前記所定の領域は、例えば、前記カラー画像における前記第1文字列の右側及び／又は下側の領域が少なくとも含まれるように構成され得る。こうすれば、第1文字列の位置に基づく所定の領域に含まれる文字列を第2文字列として検出することを可能とする。

【0011】

また、上記装置において、前記所定の項目は、金額に関する項目であり、前記項目名称検出部は、前記第1文字列に加えて、通貨に関する所定の文字を検出するように構成されており、前記項目値検出部は、前記カラー画像において検出された前記第1文字列及び前記所定の文字の位置に少なくとも基づいて前記第2文字列を検出するように構成することもできる。通貨に関する所定の文字は、例えば、「円」又は「¥」を含む。こうすれば、第1文字列の位置に加えて、通貨に関する所定の文字の位置に基づいて第2文字列が検出されるから、当該第2文字列の検出の精度が向上し得る。

【0012】

本発明の実施形態に係る文字列認識方法は、1又は複数のコンピュータによって実行され、帳票に含まれる文字列を認識するための文字列認識方法であって、前記帳票のカラー画像を取得する工程と、取得した前記カラー画像において、前記帳票に含まれる所定の項目の名称に対応する第1文字列を検出する工程と、前記カラー画像から少なくとも特定の色が消去されたドロップアウト画像を生成する工程と、生成した前記ドロップアウト画像において、前記カラー画像において検出された前記第1文字列の位置に少なくとも基づいて、前記所定の項目の値に対応する第2文字列を検出する工程と、を備える。

【0013】

この構成によれば、帳票に含まれる所定の項目の名称に対応する第1文字列を検出し、当該第1文字列の位置に基づいて当該所定の項目の値に対応する第2文字列を検出するから、帳票の様式情報の事前登録なしに、所定の項目の値(第2文字列)の認識が可能となる。さらに、帳票のカラー画像から特定の色が消去されたドロップアウト画像において所定の項目の値(第2文字列)を検出するから、例えば、カラー画像において所定の項目の値が特定の色の枠線／罫線等に重なっている場合であっても(こうした枠線／罫線等との接触は不読／誤読の原因となる。)、ドロップアウト画像においては当該特定の色の枠線／罫線が消去され、所定の項目の値の不読／誤読が抑制される。

【0014】

本発明の実施形態に係る文字列認識プログラムは、帳票に含まれる文字列を認識するための文字列認識プログラムであって、コンピュータに、前記帳票のカラー画像を取得する処理と、取得した前記カラー画像において、前記帳票に含まれる所定の項目の名称に対応する第1文字列を検出する処理と、前記カラー画像から少なくとも特定の色が消去されたドロップアウト画像を生成する処理と、生成した前記ドロップアウト画像において、前記カラー画像において検出された前記第1文字列の位置に少なくとも基づいて、前記所定の項目の値に対応する第2文字列を検出する処理と、を実行させる。

【0015】

この構成によれば、帳票に含まれる所定の項目の名称に対応する第1文字列を検出し、当該第1文字列の位置に基づいて当該所定の項目の値に対応する第2文字列を検出するから、帳票の様式情報の事前登録なしに、所定の項目の値(第2文字列)の認識が可能となる。さらに、帳票のカラー画像から特定の色が消去されたドロップアウト画像において所定の項目の値(第2文字列)を検出するから、例えば、カラー画像において所定の項目の値が特定の色の枠線／罫線等に重なっている場合であっても(こうした枠線／罫線等との接触は不読／誤読の原因となる。)、ドロップアウト画像においては当該特定の色の枠線／罫線が消去され、所定の項目の値の不読／誤読が抑制される。

【発明の効果】

【 0 0 1 6 】

本発明の様々な実施形態は、様式情報の事前登録なしで行われる帳票の処理を支援する。

【図面の簡単な説明】

【 0 0 1 7 】

【図 1】本発明の一実施形態に係る文字列認識装置として機能する画像処理装置 20 を含む帳票読取システム 1 のシステム構成を概略的に例示する構成図である。

【図 2】画像処理装置 20 のハードウェア構成を概略的に例示する構成図である。

【図 3】画像処理装置 20 が有する機能を概略的に例示する構成図である。

【図 4】画像処理装置 20 が実行する処理を例示するフローチャートである。

10

【図 5】帳票の様式を例示する図である。

【図 6】図 5 に例示した帳票に含まれる項目の一部に対して値が印刷されている状態の帳票を例示する図である。

【図 7】対象の項目名称の矩形領域 A 1 及び通貨関連文字の矩形領域 A 2 が検出される様子を例示するための図である。

【図 8】図 6 に例示した帳票のカラー画像に基づいて生成されるドロップアウト画像を例示する図である。

【図 9】対象の項目値が検出される様子を例示するための図である。

【図 10】対象の項目値が、帳票の様式における枠線 / 罫線に重なって印刷されてしまっている場合を説明するための図。

20

【発明を実施するための形態】

【 0 0 1 8 】

図 1 は、本発明の一実施形態に係る文字列認識装置として機能する画像処理装置 20 を含む帳票読取システム 1 のシステム構成を概略的に例示する構成図である。帳票読取システム 1 は、図示するように、処理対象となる帳票を光学的に読み取って画像を生成する帳票読取装置 10 と、当該帳票読取装置 10 によって生成された画像に対する画像処理を行う画像処理装置 20 と、当該画像処理装置 20 と通信ネットワーク N を介して通信可能に接続されたホストサーバ 30 と、を備える。システム 1 は、帳票読取装置 10 によって光学的に読み取られる帳票に含まれる文字列を、画像処理装置 20 によって実行される画像処理を介して認識するように構成されている。本実施形態における帳票読取システム 1 による読取の対象となる帳票は、様々な種類の帳票が含まれ、例えば、口座振替依頼書及び税公金の領収済通知書等を含む任意の帳票である。

30

【 0 0 1 9 】

帳票読取装置 10 は、一般的な業務用スキャナとして構成されており、所定の位置にセットされた帳票を光学的にスキャンし、当該スキャンを介して生成された画像を画像処理装置 20 に入力するように構成されている。

【 0 0 2 0 】

ホストサーバ 30 は、一般的なサーバ装置として構成されており、帳票内の文字列の認識のために必要な様々な情報 / データを記憶し、当該情報 / データを画像処理装置 20 に対して適宜に提供するように構成されている。

40

【 0 0 2 1 】

本実施形態における画像処理装置 20 は、帳票読取装置 10 によって入力される画像を用いて、帳票内に存在する文字列を認識する。図 2 は、画像処理装置 20 のハードウェア構成を概略的に例示する構成図である。画像処理装置 20 は、一般的なコンピュータとしての構成を有しており、図示するように、CPU 21 と、メインメモリ 22 と、ストレージ 23 と、帳票読取装置インターフェイス 24 と、入力インターフェイス 25 と、表示コントローラ 26 と、通信インターフェイス 27 と、リーダ / ライタ 28 とを有する。

【 0 0 2 2 】

画像読取装置インターフェイス 24 は、CPU 21 と帳票読取装置 10 との間のデータ伝送を仲介し、当該装置 10 から入力される画像を一時的に蓄積するための画像バッ

50

ファ２４１を有している。入力インターフェイス２５は、ＣＰＵ２１と各種の入力部との間のデータ伝送を仲介する。入力部には、マウス３１、キーボード、及びタッチパネル等が含まれる。表示コントローラ２６は、液晶モニタ等のディスプレイ３２が接続されており、当該ディスプレイ３２における表示を制御する。通信インターフェイス２７は、ＣＰＵ２１とパーソナルコンピュータ及びホストサーバ３０等との間のデータ伝送を仲介する。リーダ／ライタ２８は、ＣＰＵ２１と各種の記録媒体（メモリカード３３が含まれる）との間のデータ伝送を仲介する。各インターフェイスは、例えばＵＳＢ等の規格に従うように構成される。

【００２３】

画像処理装置２０は、ＣＰＵ２１が、ストレージ２３等に格納されているプログラムをメインメモリ２２に読み込んで実行することにより、各種の処理を実行する。ストレージ２３等に格納されているプログラムは、例えば、メモリカード３３や光ディスク等のコンピュータ読取可能な記録媒体を介して提供され、又は、インターネット等の通信ネットワークＮを介して提供される。なお、本実施形態におけるプログラムは、単体のアプリケーションプログラムとして構成されていてもよいし、他のプログラムの一部として組み込まれるモジュールとして構成されていてもよい。また、こうしたプログラムを介して実行される処理の一部又は全部は、ＡＳＩＣ等の専用回路を介して実行されるように構成され得る。

10

【００２４】

図３は、本実施形態の画像処理装置２０が有する機能を概略的に例示する構成図である。画像処理装置２０は、図示するように、カラー画像取得部２０１と、項目名称検出部２０２と、ドロップアウト画像生成部２０３と、項目値検出部２０４と、様々な情報をストレージ２３等において記憶する記憶部２０５とを有する。

20

【００２５】

カラー画像取得部２０１は、帳票のカラー画像の取得に関する処理を実行する。例えば、カラー画像取得部２０１は、帳票読取装置１０から入力されるカラー画像を取得するように構成されている。

【００２６】

項目名称検出部２０２は、帳票に含まれる所定の項目の名称（以下、「項目名称」と言う場合がある。）に対応する文字列（第１文字列）の検出に関する処理を実行する。例えば、項目名称検出部２０２は、カラー画像取得部２０１によって取得されるカラー画像を解析することにより、当該カラー画像に含まれる、項目名称に対応する文字列を検出するように構成されている。

30

【００２７】

本実施形態において、文字列の認識の対象となる所定の項目は、「納入金額」、「税額」、「納付額」、「収納金額」、「請求金額」等の金額に関する項目として構成される。これらの項目名称に対応する文字列は、例えば、所定の形式のファイル内にリスト化されており、当該ファイルは、記憶部２０５（ストレージ２３）及び／又はホストサーバ３０等に格納される。なお、本発明の他の実施形態においては、文字列の認識の対象となる所定の項目は、金額に関する項目以外の様々な項目が含まれ得る。

40

【００２８】

本実施形態において、項目名称検出部２０２は、対象の項目名称に対応する文字列に加えて、通貨に関する通貨関連文字（所定の文字）を検出するように構成されている。例えば、項目名称検出部２０２は、カラー画像に含まれる「円」及び「¥」という文字を検出するように構成されている。

【００２９】

ドロップアウト画像生成部２０３は、ドロップアウト画像の生成に関する処理を実行する。例えば、ドロップアウト画像生成部２０３は、カラー画像取得部２０１によって取得されたカラー画像に対してドロップアウト処理を行うことにより、ドロップアウト画像を生成するように構成されている。本実施形態におけるドロップアウト処理は、黒

50

色以外の色を除去する処理であり、例えば、カラー画像から、R成分、G成分、及び、B成分を順に除去する処理として構成される。なお、黒色とは、R成分、G成分、及び、B成分を完全に除去したものに限定されず、帳票に含まれる項目値を検出できる程度に各色の成分が除去されればよい。また、本発明の他の実施形態において、ドロップアウト処理においてカラー画像から除去される色は、黒色以外の色に限定されず、例えば、黒色及び他の特定の色以外の色、又は、1又は複数の特定の色であってもよい。

【0030】

項目値検出部204は、帳票に含まれる所定の項目の値（以下、「項目値」と言う場合がある。）に対応する文字列（第2文字列）の検出に関する処理を実行する。例えば、項目値検出部204は、ドロップアウト画像生成部203によって生成されたドロップアウト画像を解析し、当該解析結果と、項目名称検出部202によって検出されたカラー画像における項目名称に対応する文字列の位置と、に少なくとも基づいて、当該ドロップアウト画像における項目値に対応する文字列を検出するように構成されている。

10

【0031】

次に、このように構成された本実施形態の画像処理装置20の動作について説明する。図4は、画像処理装置20が実行する処理を例示するフローチャートである。当該装置20は、まず、図示するように、帳票読取装置10から入力される帳票のカラー画像を取得し（S101）、当該カラー画像に含まれる対象の項目名称及び通貨関連文字を検出する（S102）。具体的には、上述したように、項目名称としての「納入金額」、「税額」、「納付額」、「収納金額」、「請求金額」等の文字列が検出されると共に、関連文字としての「円」及び「¥」という文字が検出される。以下、具体例を用いて説明する。

20

【0032】

図5は、本実施形態における帳票の様式を例示する。図5に例示される様式は、「水道料金等納入済み通知書」の様式であって、様々なテキスト、枠線/罫線、項目名称（例えば、お客様番号、水道料金、開栓手数料、納入期限等）、単位（例えば、「円」等）等の情報が印刷されている。様式として予め印刷されているこれらの情報は、黒色以外の色（例えば、青色）で印刷されている。

【0033】

図6は、図5に例示した帳票に含まれる項目の一部に対して値が印刷されている状態の帳票を例示する。具体的には、図6の帳票においては、お客様番号、使用期間、お客様氏名、水道料金、下水道使用料、収納金額、納入期限、及び、発行日に対する値がそれぞれ対応する枠内に印刷されている。また、図6の帳票においては、領収日付印が押印されている。これらの各項目の値、及び、領収日付印は、黒色で印刷又は押印されている。ステップS101では、このように項目の値が印刷されている状態の帳票のカラー画像が取得されている。

30

【0034】

例えば、図5、6の帳票における項目「収納金額」は、対象の項目名称としてリスト化されているので、当該文字列がカラー画像において検出される。また、図5、6の帳票における「円」という文字が通貨関連文字としてカラー画像において検出される。具体的には、図7に例示するように、「収納金額」と印字されている1つの矩形領域A1、及び、「円」と印字されている6つの矩形領域A2がカラー画像において検出される。

40

【0035】

図4のフローチャートに戻り、続いて、画像処理装置20は、カラー画像に基づいてドロップアウト画像を生成する（S103）。上述したように、本実施形態におけるドロップアウト画像は、黒色以外の色が除去された画像（二値画像）である。図8は、図6に例示した帳票のカラー画像に基づいて生成されるドロップアウト画像を例示する。図示するように、ドロップアウト画像では、帳票において黒色以外の色で印刷されていた様々なテキスト、枠線/罫線、項目名称、単位等の情報が除去され、黒色で印刷/押

50

印されていた項目値及び領収日付印が残存する。

【0036】

図4のフローチャートに戻り、こうしてドロップアウト画像を生成すると、次に、画像処理装置20は、当該ドロップアウト画像に含まれる対象の項目値を検出する(S104)。項目値の検出は、ステップS102において検出された項目名称及び通貨関連文字のカラー画像上での位置に基づいて行われる。

【0037】

図9は、対象の項目値が検出される様子を例示するための図である。本実施形態では、図示するように、対象の項目名称の矩形領域A1を拡張した拡張領域A11に含まれる文字列が対象の項目値の候補となる。当該拡張領域A11は、領域A1を、上方向及び左方向に第1の幅W1だけ拡張し、且つ、下方向に第1の幅W1よりも大きい第2の幅W2だけ拡張すると共に右方向に第1の幅W1よりも大きい第3の幅W3だけ拡張した領域として構成される。図9の例では、拡張領域A11に含まれる「148,004」という文字列、及び、「平成29年3月27日」という文字列が対象の項目値の候補となる。

10

【0038】

また、対象の項目値は、通貨関連文字の矩形領域A2との相対的な位置関係を考慮して検出される。具体的には、「円」という通貨関連文字の左側の文字列が優先的に対象の項目値として検出され、例えば、図9の例では、対象の項目値の候補となる2つの文字列「148,004」及び「平成29年3月27日」のうち、通貨関連文字の領域A2の左側に位置する「148,004」という文字列が対象の項目値として検出される。なお、「¥」という通貨関連文字が検出されている場合には、当該文字の右側の文字列が優先的に対象の項目値として検出される。

20

【0039】

ここで、対象の項目値が、帳票の様式における枠線/罫線に重なって印刷されている場合を考える。例えば、図10の上側に例示するように、対象の項目である「収納金額」の項目値「148,004」がカラー画像において枠の下側の枠線/罫線に重なってしまっている場合、当該カラー画像を用いて項目値を検出しようとする、不読/誤読のおそれがあるが、図10の下側に例示するように、ドロップアウト画像においては黒色以外の色で印刷されていた枠線/罫線が消去されるので、枠線/罫線に重なってしまうことに起因する項目値の不読/誤読のおそれなくなる。

30

【0040】

本実施形態では、対象の項目名称の矩形領域A1を拡張した拡張領域A11に含まれる文字列が対象の項目値として検出されるようにしたが、こうした手法は例示であって、本発明の他の実施形態では、他の手法が適用され、例えば、対象の項目名称の矩形領域A1との間の相対的な位置関係が所定の関係にある文字列が対象の項目値として検出されるような様々な手法が適用される。

【0041】

本実施形態では、カラー画像において、対象の項目名称に対応する文字列に加えて、通貨に関する通貨関連文字を検出するように構成したが、本発明の他の実施形態では、通貨関連文字の検出は行わず、例えばカラー画像に存在する罫線ないし枠線の情報を用いて、項目名称の属する矩形領域を特定し、その右または下の矩形領域を探索対象領域として、項目値の検出を行なう手法が適用されてもよい。

40

【0042】

以上説明した本発明の実施形態に係る画像処理装置20は、帳票に含まれる所定の項目の名称に対応する文字列を検出し、当該名称に対応する文字列の位置に基づいて当該所定の項目の値に対応する文字列を検出するから、帳票の様式情報の事前登録なしに、所定の項目の値に対応する文字列の認識が可能となる。さらに、帳票のカラー画像から特定の色(例えば、黒色以外の色)が消去されたドロップアウト画像において項目値に対応する文字列を検出するから、例えば、カラー画像において項目値が特定の色

50

／罫線に重なってしまっている場合であっても、ドロップアウト画像においては当該特定の色の枠線／罫線が消去され、項目値に対応する文字列の不読／誤読が抑制される。

【符号の説明】

【 0 0 4 3 】

1 ... 帳票読取システム、10 ... 帳票読取装置、20 ... 画像処理装置、21 ... CPU、
22 ... メインメモリ、23 ... ストレージ、24 ... 帳票読取装置インターフェイス、24
1 ... 画像バッファ、25 ... 入力インターフェイス、26 ... 表示コントローラ、27 ... 通
信インターフェイス、28 ... リーダ／ライター、30 ... ホストサーバ、31 ... マウス、3
2 ... ディスプレイ、33 ... メモリカード、201 ... カラー画像取得部、202 ... 項目名
称検出部、203 ... ドロップアウト画像生成部、204 ... 項目値検出部、205 ... 記憶
部。

10

20

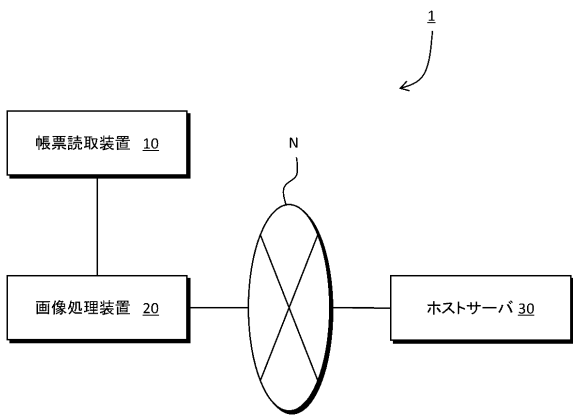
30

40

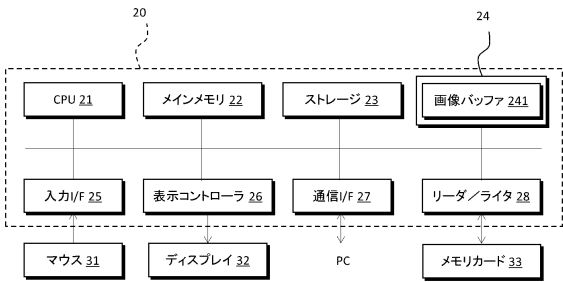
50

【図面】

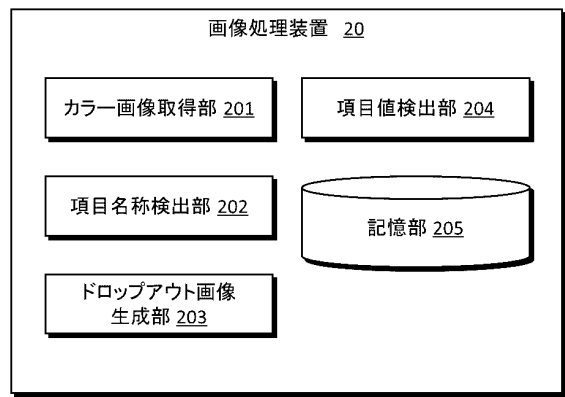
【図 1】



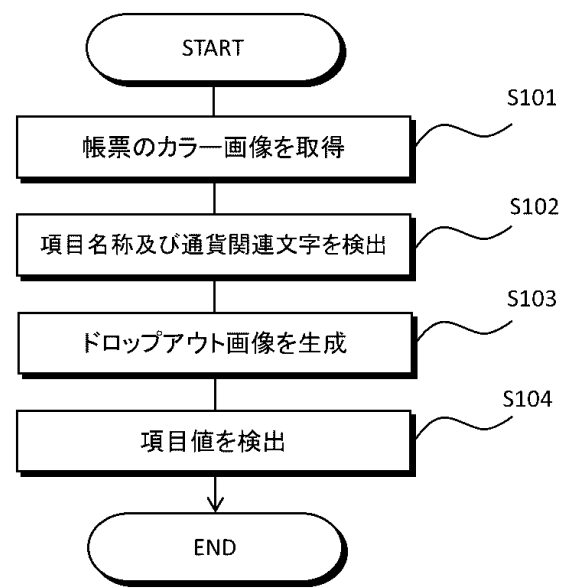
【図 2】



【図 3】



【図 4】



10

20

30

40

50

【 図 5 】

〇〇〇水道料金等納入済通知書

.....

お客様番号

使用期間

お客様氏名

	金額		金額
水道料金	円	開栓手数料	円
下水道使用料	円	修繕工事費	円
下水道使用料 (井戸)	円	収納金額	円

納入期限

発行日

.....

領
収
日
付
印

【 図 6 】

〇〇〇水道料金等納入済通知書

.....

お客様番号 9999999999

使用期間 平成29年2月～
平成29年3月

お客様氏名 △△△△△△△

	金額		金額
水道料金	70,050 円	開栓手数料	円
下水道使用料	77,954 円	修繕工事費	円
下水道使用料 (井戸)	円	収納金額	148,004 円

納入期限

平成29年4月20日

発行日

平成29年3月27日

.....

領
収
日
付
印

領
29. 4. 15
収

10

20

【 図 7 】

〇〇〇水道料金等納入済通知書

.....

お客様番号 9999999999

使用期間 平成29年2月～
平成29年3月

お客様氏名 △△△△△△△

	金額		金額
水道料金	70,050 円	開栓手数料	円
下水道使用料	77,954 円	修繕工事費	円
下水道使用料 (井戸)	円	収納金額	148,004 円

納入期限

平成29年4月20日

発行日

平成29年3月27日

.....

領
収
日
付
印

領
29. 4. 15
収

A2

A2

A2

A1

【 図 8 】

9999999999 平成29年2月～
平成29年3月

△△△△△△△

70,050

77,954

148,004

平成29年4月20日 平成29年3月27日

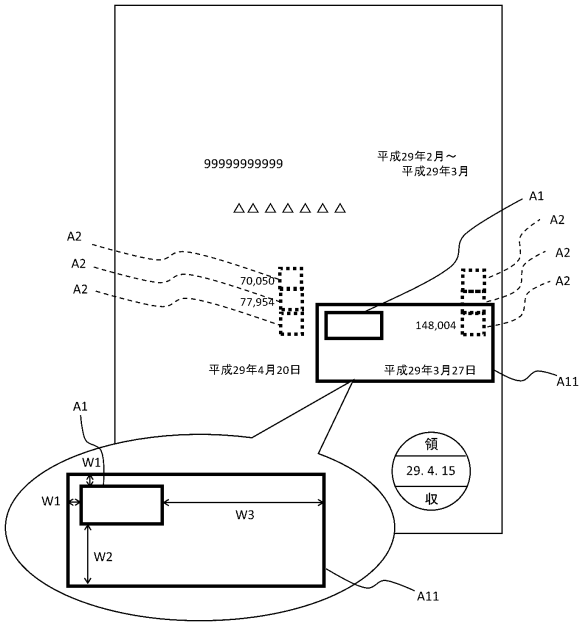
領
29. 4. 15
収

30

40

50

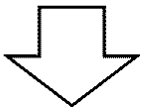
【図 9】



【図 10】

(カラー画像)

収納金額	円
148,004	



(ドロップアウト画像)

148,004

10

20

30

40

50

フロントページの続き

 プリマジェスト内
(72)発明者 加藤 真
 神奈川県川崎市幸区堀川町５８０番地 ソリッドスクエア東館１２階 株式会社プリマジェスト内
 審査官 小池 正彦
(56)参考文献 特開２０１３－０３００４０（ＪＰ，Ａ）
 特開平０６－０６８２９９（ＪＰ，Ａ）
 特開２０１５－１１８４８８（ＪＰ，Ａ）
 特開２００１－１２６０１０（ＪＰ，Ａ）
(58)調査した分野 (Int.Cl.，ＤＢ名)
 Ｇ０６Ｖ ３０／１４
 Ｇ０６Ｖ ３０／４１２