

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2019年2月7日(07.02.2019)



(10) 国際公開番号

WO 2019/026147 A1

(51) 国際特許分類:
G06K 9/00 (2006.01)

(21) 国際出願番号: PCT/JP2017/027758

(22) 国際出願日: 2017年7月31日(31.07.2017)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(71) 出願人: 株式会社 P F U (PFU LIMITED) [JP/JP]; 〒9291192 石川県かほく市宇野気ヌ98番地の2 Ishikawa (JP).

(72) 発明者: 谷崎 裕紀 (TANIZAKI, Hiroki); 〒9291192 石川県かほく市宇野気ヌ98番地の2 株式会社 P F U 内 Ishikawa (JP). 西川 満

(NISHIKAWA, Mitsuru); 〒9291192 石川県かほく市宇野気ヌ98番地の2 株式会社 P F U 内 Ishikawa (JP). 小坂 清人 (KOSAKA, Kiyoto); 〒9291192 石川県かほく市宇野気ヌ98番地の2 株式会社 P F U 内 Ishikawa (JP).

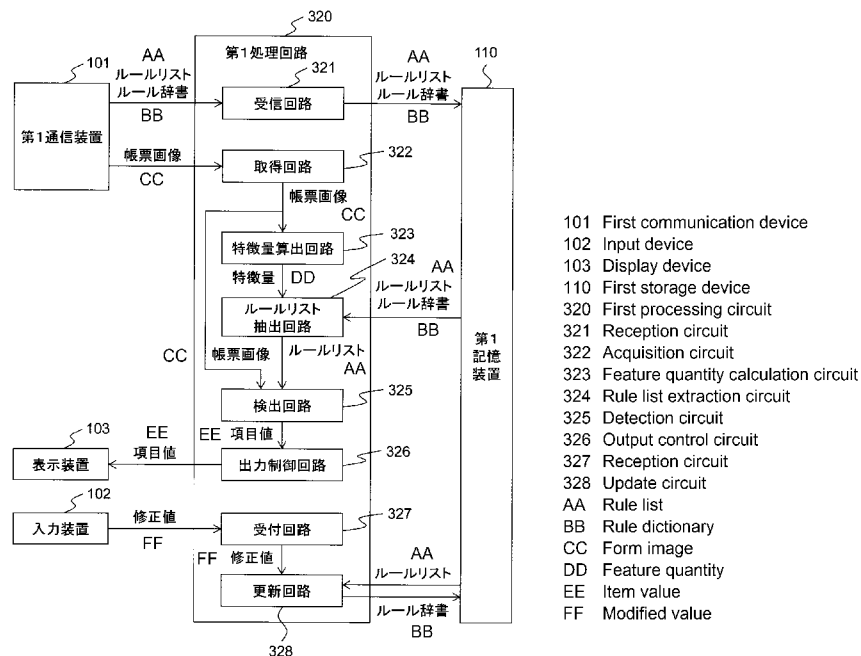
(74) 代理人: 青木 篤, 外 (AOKI, Atsushi et al.); 〒1058423 東京都港区虎ノ門三丁目5番1号 虎ノ門37森ビル青和特許法律事務所 Tokyo (JP).

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,

(54) Title: INFORMATION PROCESSING DEVICE, INFORMATION PROCESSING SYSTEM, CONTROL METHOD, AND CONTROL PROGRAM

(54) 発明の名称: 情報処理装置、情報処理システム、制御方法及び制御プログラム

[図17]



(57) Abstract: The present invention makes it possible to more accurately detect an item value for a desired item from a form image that indicates an item name and an item value for each of a plurality of items. This information processing device processes a form image indicating an item name and an item value for each of a plurality of items, and comprises: a storage unit which stores a rule dictionary associating a feature quantity of each of a plurality of sample form images with one of a plurality of rule lists that each specify a plurality of item names and a priority assigned to each item name;

HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH,
KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY,
MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ,
NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT,
QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA,
UG, US, UZ, VC, VN, ZA, ZM, ZW.

- (84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (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), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類：

- 一 国際調査報告（条約第21条(3)）

an acquisition unit which acquires a form image; a feature quantity calculation unit which calculates a feature quantity from the acquired form image; a rule list extraction unit which extracts the rule list associated by the rule dictionary with the calculated feature quantity; a detection unit which sequentially determines whether or not each item name specified in the extracted rule list is included in the form image, in the priority order specified in the rule list, and detects the item value associated with the first item name that was determined to be included in the form image; and an output unit which outputs the detected item value.

(57) 要約：複数の項目毎に項目名及び項目値が表された帳票画像から、所望の項目の項目値をより精度良く検出することを可能とする。情報処理装置は、複数の項目毎に項目名及び項目値が表された帳票画像を処理し、複数のサンプル帳票画像毎の特徴量のそれぞれと、複数の項目名及び各項目名の優先順位が規定された複数のルールリストのそれぞれとが関連付けられたルール辞書を記憶する記憶部と、帳票画像を取得する取得部と、取得された帳票画像から特徴量を算出する特徴量算出部と、ルール辞書において、算出された特徴量に関連付けられたルールリストを抽出するルールリスト抽出部と、抽出されたルールリストに規定された優先順位の順に、そのルールリストに規定された各項目名が帳票画像に含まれるか否かを順次判定し、最初に含まれると判定された項目名に対応する項目値を検出する検出部と、検出された項目値を出力する出力部とを有する。

明 細 書

発明の名称：

情報処理装置、情報処理システム、制御方法及び制御プログラム

技術分野

[0001] 本開示は、情報処理装置、情報処理システム、制御方法及び制御プログラムに関し、特に、複数の項目毎に項目名及び項目値が表された帳票画像を処理する情報処理装置、情報処理システム、制御方法及び制御プログラムに関する。

背景技術

[0002] 生命保険会社では、保険金支払いの際、顧客から送付された医療領収書等の帳票を、担当者が手作業によりデータ化している。生命保険会社によっては、顧客から膨大な数の帳票が送付されるため、担当者の業務負担は大きく、帳票のデータ化作業のアシストに対する要望が高まっている。

[0003] 任意のフォーマットで情報が記入された媒体から読み取ったイメージに基づいて、その情報の認識を行う媒体処理装置が開示されている（特許文献1を参照）。この媒体処理装置は、読み取ったイメージのデータから、学習更新が可能な解析用辞書を参照してそのイメージのレイアウトの特徴を抽出し、その特徴から認識すべき情報の存在する位置を特定するとともに、そのレイアウトの特徴を学習によって更新していく。

[0004] また、入力された帳票の画像情報の枠及び文字列を検出し、その文字列が項目名である可能性が高い枠を項目名領域候補として検出し、項目名領域候補に含まれる文字列が項目名であるか否かを判定する帳票認識装置が開示されている（特許文献2を参照）。

[0005] また、帳票原稿から取得した画像データについて帳票処理を行う帳票処理装置が開示されている（特許文献3を参照）。この帳票処理装置は、各画像データから抽出した罫線データの種に基づいて、各画像データを複数の画像グループに分類し、各画像グループに割り当てられた帳票テンプレートデ

ータに基づいて、各画像グループの画像データについて帳票処理を行う。

[0006] また、帳票上に任意の表構造で存在する文字列の認識を行う帳票認識装置が開示されている（特許文献4を参照）。この帳票認識装置は、帳票が電子化された帳票画像上に存在する文字列を認識し、認識した文字列のなかから予め定めた文字列である見出し文言を抽出する。そして、この帳票認識装置は、抽出した見出し文言の帳票画像上の配置に基づいて、帳票画像上に存在する表構造を判定し、その判定結果を用いて、見出し文言とデータの間の対応関係を特定する。

先行技術文献

特許文献

[0007] 特許文献1：WO 97／05561

特許文献2：特開2009－93305号公報

特許文献3：特開2014－89575号公報

特許文献4：特開2010－3155号公報

発明の概要

[0008] 複数の項目毎に項目名及び項目値が表された帳票画像を処理する情報処理装置において、所望の項目の項目値をより精度良く検出できることが望まれている。

[0009] 情報処理装置、情報処理システム、制御方法及び制御プログラムの目的は、複数の項目毎に項目名及び項目値が表された帳票画像から、所望の項目の項目値をより精度良く検出することを可能とすることにある。

[0010] 本発明の一側面に係る情報処理装置は、複数の項目毎に項目名及び項目値が表された帳票画像を処理する情報処理装置であって、複数のサンプル帳票画像毎の特徴量のそれぞれと、複数の項目名及び各項目名の優先順位が規定された複数のルールリストのそれぞれとが関連付けられたルール辞書を記憶する記憶部と、帳票画像を取得する取得部と、取得された帳票画像から特徴量を算出する特徴量算出部と、ルール辞書において、算出された特徴量に関連付けられたルールリストを抽出するルールリスト抽出部と、抽出されたル

ールリストに規定された優先順位の順に、そのルールリストに規定された各項目名が帳票画像に含まれるか否かを順次判定し、最初に含まれると判定された項目名に対応する項目値を検出する検出部と、検出された項目値を出力する出力部とを有する。

[0011] 本発明の一側面に係る情報処理システムは、生成装置と、複数の項目毎に項目名及び項目値が表された帳票画像を処理する情報処理装置とを有する情報処理システムであって、生成装置は、複数のサンプル帳票画像と、複数の項目名及び各項目名の優先順位が規定された複数のルールリストとを記憶する記憶部と、複数のルールリストのそれぞれに適合するサンプル帳票画像の特微量と、各ルールリストとを関連付けたルール辞書を生成するルール辞書生成部と、ルール辞書を情報処理装置に送信する送信部とを有し、情報処理装置は、生成装置からルール辞書を受信する受信部と、帳票画像を取得する取得部と、取得された帳票画像から特微量を算出する特微量算出部と、受信されたルール辞書において、算出された特微量に関連付けられたルールリストを抽出するルールリスト抽出部と、抽出されたルールリストに規定された優先順位の順に、そのルールリストに規定された各項目名が帳票画像に含まれるか否かを順次判定し、最初に含まれると判定された項目名に対応する項目値を検出する検出部と、検出された項目値を出力する出力部とを有する。

[0012] また、本発明の一側面に係る制御方法は、記憶部と、出力部とを有し、複数の項目毎に項目名及び項目値が表された帳票画像を処理する情報処理装置の制御方法であって、複数のサンプル帳票画像毎の特微量のそれぞれと、複数の項目名及び各項目名の優先順位が規定された複数のルールリストのそれぞれとが関連付けられたルール辞書を記憶部に記憶し、帳票画像を取得し、取得された帳票画像から特微量を算出し、ルール辞書において、算出された特微量に関連付けられたルールリストを抽出し、抽出されたルールリストに規定された優先順位の順に、そのルールリストに規定された各項目名が帳票画像に含まれるか否かを順次判定し、最初に含まれると判定された項目名に対応する項目値を検出し、検出された項目値を出力部に出力することを含む

。

[0013] また、本発明の一側面に係る制御プログラムは、記憶部と、出力部とを有し、複数の項目毎に項目名及び項目値が表された帳票画像を処理する情報処理装置の制御プログラムであって、複数のサンプル帳票画像毎の特徴量のそれぞれと、複数の項目名及び各項目名の優先順位が規定された複数のルールリストのそれぞれとが関連付けられたルール辞書を記憶部に記憶し、帳票画像を取得し、取得された帳票画像から特徴量を算出し、ルール辞書において、算出された特徴量に関連付けられたルールリストを抽出し、抽出されたルールリストに規定された優先順位の順に、そのルールリストに規定された各項目名が帳票画像に含まれるか否かを順次判定し、最初に含まれると判定された項目名に対応する項目値を検出し、検出された項目値を出力部に出力することを情報処理装置に実行させる。

[0014] 本実施形態によれば、情報処理装置、情報処理システム、制御方法及び制御プログラムは、複数の項目毎に項目名及び項目値が表された帳票画像から、所望の項目の項目値をより精度良く検出することが可能となる。

[0015] 本発明の目的及び効果は、特に請求項において指摘される構成要素及び組み合わせを用いることによって認識され且つ得られるだろう。前述の一般的な説明及び後述の詳細な説明の両方は、例示的及び説明的なものであり、特許請求の範囲に記載されている本発明を制限するものではない。

図面の簡単な説明

[0016] [図1]実施形態に従った情報処理システム10の概略構成の一例を示す図である。

[図2]第1記憶装置110及び第1CPU120の概略構成を示す図である。

[図3]第2記憶装置210及び第2CPU220の概略構成を示す図である。

[図4]複数のサンプル帳票画像400、401、402の一例を示す図である。

。

[図5]複数のルールリスト500、501、502の一例を示す図である。

[図6]ルール辞書600の一例を示す図である。

[図7]ルール辞書の生成処理の動作の例を示すフローチャートである。

[図8]罫線の形状に関する特徴量について説明するための図である。

[図9]ロゴに関する特徴量について説明するための図である。

[図10]文字認識結果に関する特徴量について説明するための図である。

[図11]生成処理により生成されるルール辞書の一例を示す図である。

[図12]生成処理により生成されるルール辞書の一例を示す図である。

[図13]生成処理により生成されるルール辞書の一例を示す図である。

[図14]項目値の検出処理の動作の例を示すフローチャートである。

[図15]認識結果画面の一例を示す図である。

[図16]更新処理の動作の例を示すフローチャートである。

[図17]他の情報処理装置の第1処理回路320の概略構成を示す図である。

[図18]他の生成装置の第2処理回路420の概略構成を示す図である。

発明を実施するための形態

[0017] 以下、本開示の一側面に係る情報処理装置について図を参照しつつ説明する。但し、本開示の技術的範囲はそれらの実施の形態に限定されず、特許請求の範囲に記載された発明とその均等物に及ぶ点に留意されたい。

[0018] 図1は、実施形態に従った情報処理システム10の概略構成の一例を示す図である。

[0019] 情報処理システム10は、情報処理装置100と、生成装置200とを有する。

[0020] 情報処理装置100及び生成装置200は、それぞれ、有線又は無線による通信機能を有し、ネットワーク300に接続され、ネットワーク300を介して相互に通信する。

[0021] 情報処理装置100は、PC (Personal Computer)、ノートPC等の情報処理装置であり、そのユーザである作業者により使用され、複数の項目毎に項目名及び項目値が表された帳票画像を処理する。なお、情報処理装置100は、タブレットPC、多機能携帯電話（いわゆるスマートフォン）、携帯情報端末等の携帯可能な装置でもよい。情報処理装置100は、第1通信装

置 1 0 1 と、入力装置 1 0 2 と、表示装置 1 0 3 と、第 1 記憶装置 1 1 0 と、第 1 C P U (Central Processing Unit) 1 2 0 とを有する。以下、情報処理装置 1 0 0 の各部について詳細に説明する。

[0022] 第 1 通信装置 1 0 1 は、T C P / I P (Transmission Control Protocol/Internet Protocol) 等の有線の通信インタフェース回路を有する。第 1 通信装置 1 0 1 は、イーサネット (登録商標) 等の通信方式に従って、ネットワーク 3 0 0 と通信接続する。第 1 通信装置 1 0 1 は、ネットワーク 3 0 0 を介して生成装置 2 0 0 から受信したデータを第 1 C P U 1 2 0 に供給し、第 1 C P U 1 2 0 から供給されたデータをネットワーク 3 0 0 を介して生成装置 2 0 0 に送信する。なお、第 1 通信装置 1 0 1 は、外部の装置と通信できるものであればどのようなものでもよい。例えば、第 1 通信装置 1 0 1 は、無線 L A N (Local Area Network) 通信方式に従って不図示のアクセスポイントを介して生成装置 2 0 0 と通信してもよい。または、第 1 通信装置 1 0 1 は、携帯電話通信方式に従って不図示の基地局装置を介して生成装置 2 0 0 と通信してもよい。

[0023] 入力装置 1 0 2 は、入力部の一例であり、タッチパネル式の入力装置、キーボード、マウス等の入力デバイス及び入力デバイスから信号を取得するインタフェース回路を有する。入力装置 1 0 2 は、ユーザにより入力された入力データを受け付け、受け付けた入力データに応じた信号を第 1 C P U 1 2 0 に対して出力する。

[0024] 表示装置 1 0 3 は、出力部の一例であり、液晶、有機 E L (Electro-Luminescence) 等から構成されるディスプレイ及びディスプレイに画像データ又は各種の情報を出力するインタフェース回路を有する。表示装置 1 0 3 は、第 1 C P U 1 2 0 と接続されて、第 1 C P U 1 2 0 から出力された画像データをディスプレイに表示する。なお、タッチパネルディスプレイを用いて、入力装置 1 0 2 と表示装置 1 0 3 を一体に構成してもよい。

[0025] 第 1 記憶装置 1 1 0 は、記憶部の一例である。第 1 記憶装置 1 1 0 は、R A M (Random Access Memory)、R O M (Read Only Memory) 等のメモリ装

置、ハードディスク等の固定ディスク装置、又はフレキシブルディスク、光ディスク等の可搬用の記憶装置等を有する。また、第1記憶装置110には、情報処理装置100の各種処理に用いられるコンピュータプログラム、データベース、テーブル等が格納される。コンピュータプログラムは、例えばCD-ROM (compact disk read only memory)、DVD-ROM (digital versatile disk read only memory) 等のコンピュータ読み取り可能な可搬型記録媒体からインストールされてもよい。コンピュータプログラムは、公知のセットアッププログラム等を用いて第1記憶装置110にインストールされる。また、第1記憶装置110には、データとして、ルール辞書及び複数のルールリスト等が記憶される。ルール辞書及びルールリストの詳細については後述する。

[0026] 第1CPU120は、予め第1記憶装置110に記憶されているプログラムに基づいて動作する。第1CPU120は、汎用プロセッサであってもよい。なお、第1CPU120に代えて、DSP (digital signal processor)、LSI (large scale integration) 等が用いられてよい。また、CPU160に代えて、ASIC (Application Specific Integrated Circuit)、FPGA (Field-Programmable Gate Array) 等が用いられてもよい。

[0027] 第1CPU120は、第1通信装置101、入力装置102、表示装置103及び第1記憶装置110と接続され、これらの各部を制御する。第1CPU120は、第1通信装置101を介したデータ送受信制御、入力装置102の入力制御、表示装置103の出力制御及び第1記憶装置110の制御等を行う。また、第1CPU120は、第1通信装置101等から取得した帳票画像を処理する。

[0028] 生成装置200は、情報処理装置100が使用するルール辞書を生成するサーバである。ルール辞書は、帳票画像を処理するために、特に帳票画像から注目する項目の項目値を検出するために、使用されるデータである。帳票画像は、帳票を表す画像データであり、帳票は、複数の項目毎に項目名及び項目値が表された文書である。本実施形態では、帳票として多種レイアウト

の帳票が使用される。多種レイアウトの帳票とは、複数の異なる機関によって個別に作成され、帳票毎に項目名の記載位置が異なり且つ項目名の類義語が複数存在する帳票のことをいう。注目する項目とは、電子データとして演算又は出力等の処理を行う必要がある項目である。生成装置200は、第2通信装置201と、第2記憶装置210と、第2CPU220とを有する。以下、生成装置200の各部について詳細に説明する。

[0029] 第2通信装置201は、第1通信装置101と同様の通信インタフェース回路を有し、第1通信装置101と同様の通信方式に従って、ネットワーク300と通信接続する。第2通信装置201は、ネットワーク300を介して情報処理装置100から受信したデータを第2CPU220に供給し、第2CPU220から供給されたデータをネットワーク300を介して情報処理装置100に送信する。

[0030] 第2記憶装置210は、記憶部の一例である。第2記憶装置210は、第1記憶装置と同様のメモリ装置、固定ディスク装置又は記憶装置等を有する。また、第2記憶装置210には、生成装置200の各種処理に用いられるコンピュータプログラム、データベース、テーブル等が格納される。コンピュータプログラムは、例えばCD-ROM、DVD-ROM等のコンピュータ読み取り可能な可搬型記録媒体からインストールされてもよい。コンピュータプログラムは、公知のセットアッププログラム等を用いて第2記憶装置210にインストールされる。また、第2記憶装置210には、データとして、複数のサンプル帳票画像、ルール辞書及び複数のルールリスト等が記憶される。サンプル帳票画像、ルール辞書及びルールリストの詳細については後述する。

[0031] 第2CPU220は、予め第2記憶装置210に記憶されているプログラムに基づいて動作する。第2CPU220は、汎用プロセッサであってもよい。なお、第2CPU220に代えて、DSP、LSI、ASIC、FPGA等が用いられてもよい。

[0032] 第2CPU220は、第2通信装置201及び第2記憶装置210と接続

され、これらの各部を制御する。第2CPU220は、第2通信装置201を介したデータ送受信制御及び第2記憶装置210の制御等を行う。また、第2CPU220は、ルール辞書を生成する。

[0033] 図2は、情報処理装置100の第1記憶装置110及び第1CPU120の概略構成を示す図である。

[0034] 図2に示すように、第1記憶装置110には、受信プログラム111、取得プログラム112、特徴量算出プログラム113、ルールリスト抽出プログラム114、検出プログラム115、出力制御プログラム116、受付プログラム117及び更新プログラム118等の各プログラムが記憶される。これらの各プログラムは、プロセッサ上で動作するソフトウェアにより実装される機能モジュールである。第1CPU120は、第1記憶装置110に記憶された各プログラムを読み取り、読み取った各プログラムに従って動作する。これにより、第1CPU120は、受信部121、取得部122、特徴量算出部123、ルールリスト抽出部124、検出部125、出力制御部126、受付部127及び更新部128として機能する。

[0035] 図3は、生成装置200の第2記憶装置210及び第2CPU220の概略構成を示す図である。

[0036] 図3に示すように、第2記憶装置210には、ルールリスト生成プログラム211、ルール辞書生成プログラム212及び送信プログラム213等の各プログラムが記憶される。これらの各プログラムは、プロセッサ上で動作するソフトウェアにより実装される機能モジュールである。第2CPU220は、第2記憶装置210に記憶された各プログラムを読み取り、読み取った各プログラムに従って動作する。これにより、第2CPU220は、ルールリスト生成部221、ルール辞書生成部222及び送信部223として機能する。

[0037] 図4は、生成装置200の第2記憶装置210に記憶される複数のサンプル帳票画像400、401、402の一例を示す図である。

[0038] 各サンプル帳票画像400、401、402は、ルール辞書を生成するた

めに事前に登録された帳票画像である。図4に示すように、各サンプル帳票画像400、401、402には、複数の項目毎に、各項目に対応する項目名と、各項目名に対応する項目値とが対応付けて表示されている。また、各サンプル帳票画像400、401、402には、注目する項目の項目値が、正解値として関連付けて記憶される。

[0039] 図4に示す例では、帳票として医療領収書が用いられている。この例では、医療に係る各費用が複数の項目に分類され、項目名として各費用の費用名目が記載され、項目値として各費用の金額が記載されている。また、この例では、注目する項目として「患者の負担額」が指定されている。しかしながら、医療領収書のような多種レイアウトの帳票では、同一の項目に対する項目名が帳票毎に異なる可能性がある。図4に示す例では、項目「患者の負担額」に対して、サンプル帳票画像400では項目名「患者負担額」403が用いられ、サンプル帳票画像401では項目名「一部負担金」404が用いられ、サンプル帳票画像402では項目名「合計」405が用いられている。各サンプル帳票画像400、401、402が登録される際に、サンプル帳票画像400、401、402のそれぞれにおける項目「患者の負担額」に対する正解値「15000」が登録者により登録される。

[0040] 図5は、複数のルールリスト500、501、502の一例を示す図である。

[0041] 各ルールリスト500、501、502は、生成装置200により生成されて第2記憶装置210に記憶されるとともに、情報処理装置100に送信されて第1記憶装置110に記憶される。図5に示すように、各ルールリスト500、501、502には、各ルールリストを識別するためのルールリストIDが割り当てられる。また、各ルールリスト500、501、502には、複数のルール（図5に示す表の中の各行に対応する情報）が関連付けて規定され、各ルール毎に、ルール番号、優先順位、項目名及び位置情報等が関連付けて規定される。即ち、一つのルールは、一つの項目名に対応する。

[0042] ルール番号は、各ルールを識別するための番号である。優先順位は、各ルールリストを用いて帳票画像から注目する項目の項目値を検出する際に、各ルールで規定された項目名が検査される順序を規定するための順位である。この例では、数値「1」の優先順位が最も高く、数値が大きくなる程、優先順位が低くなるように、各ルールに優先順位が割り当てられている。項目名は、各ルールリストを用いて帳票画像から注目する項目の項目値を検出する際に、各ルールにおいて検査される名称である。位置情報は、各サンプル帳票画像における各項目名と各項目名に対応する項目値の表示位置関係を規定するための情報である。位置情報として、例えば項目名に対する項目値の相対位置（例えば、上、下、右、左、右上、右下、左上、左下）が規定される。位置情報は、各ルールリストを用いて帳票画像から注目する項目の項目値を検出する際に、各ルールで規定された項目名に対応する項目値を特定するために使用される。

[0043] 図6は、ルール辞書600の一例を示す図である。

[0044] ルール辞書600は、生成装置200により生成されて第2記憶装置210に記憶されるとともに、情報処理装置100に送信されて第1記憶装置110に記憶される。図6に示すように、ルール辞書600には、複数の特徴量のそれぞれと、複数のルールリストIDのそれぞれとが関連付けて規定される。

[0045] 特徴量は、各サンプル帳票画像における特徴量である。特徴量の詳細については後述する。ルール辞書600では、各ルールリストに適合するサンプル帳票画像の特徴量と、各ルールリストとが関連付けられている。

[0046] 図7は、生成装置200によるルール辞書の生成処理の動作の例を示すフローチャートである。

[0047] 以下、図7に示したフローチャートを参照しつつ、生成装置200による生成処理の動作の例を説明する。なお、以下に説明する動作のフローは、予め第2記憶装置210に記憶されているプログラムに基づき主に第2CPU220により生成装置200の各要素と協働して実行される。

- [0048] 最初に、ルールリスト生成部２２１はルールリストを初期化し、ルール辞書生成部２２２はルール辞書を初期化する（ステップＳ１０１）。ルールリスト生成部２２１は、第２記憶装置２１０に記憶された全てのサンプル帳票画像から、注目する項目に対応し得る全ての項目名を抽出する。なお、注目する項目に対応し得る項目名は、事前に登録されていてもよい。ルールリスト生成部２２１は、抽出した各項目名に対して、想定し得る全ての位置に対応する位置情報のそれぞれが関連付けられたルールを生成し、生成した全てのルールを含むルールリストを初期ルールリストとして生成する。初期ルールリストでは、各ルールに、任意の順序でルール番号が割り当てられ、優先順位として「不定」が設定される。また、ルール辞書生成部２２２は、特徴量及びルールリストＩＤとして「不定」が設定されたルール辞書を初期ルール辞書として生成する。
- [0049] 次に、ルールリスト生成部２２１は、初期ルールリストから一つのルールを選択する（ステップＳ１０２）。
- [0050] 次に、ルールリスト生成部２２１は、選択したルールに基づいて、生成済みのルールリストの何れにも適合していない全てのサンプル帳票画像のそれぞれから項目値を検出する（ステップＳ１０３）。ステップＳ１０３の処理が最初に実行される場合、ルールリスト生成部２２１は、第２記憶装置２１０に記憶された全てのサンプル帳票画像から項目値を検出する。ルールリスト生成部２２１は、公知のＯＣＲ（Optical Character Recognition）技術を利用して、各サンプル帳票画像から、選択したルールにおける項目名を検出する。さらに、ルールリスト生成部２２１は、公知のＯＣＲ技術を利用して、検出した項目名に対して、選択したルールにおける位置情報に規定された表示位置関係に位置する所定サイズの領域から項目値を検出する。このように、ルールリスト生成部２２１は、ルールリストにおいて規定された表示位置関係に基づいて、各項目名に対応する項目値を検出する。
- [0051] 次に、ルールリスト生成部２２１は、各サンプル帳票画像から検出した各項目値が、各項目値を検出した各サンプル帳票画像に関連付けられた正解値

と一致している割合を示す一致率を算出する（ステップS 1 0 4）。ルールリスト生成部2 2 1は、各サンプル帳票画像から検出した各項目値と、各項目値を検出した各サンプル帳票画像に関連付けられた正解値が一致するか否かを判定し、検出した項目値の数に対する、各正解値と一致する項目値の割合を一致率として算出する。

[0052] 次に、ルールリスト生成部2 2 1は、初期ルールリストに含まれる全てのルールについて処理が完了したか否かを判定する（ステップS 1 0 5）。まだ処理が完了していないルールが存在する場合、ルールリスト生成部2 2 1は、処理をステップS 1 0 2に戻し、未選択のルールに対して、ステップS 1 0 2～S 1 0 4の処理を実行する。

[0053] 一方、初期ルールリストに含まれる全てのルールについて処理が完了した場合、ルールリスト生成部2 2 1は、ステップS 1 0 4で算出した一致率が高い順に優先順位が高くなるように各ルールの優先順位を割り当てたルールリストを生成する（ステップS 1 0 6）。このように、ルールリスト生成部2 2 1は、複数のルールリストの何れにも適合しない各サンプル帳票画像において、正解値に対応している割合が高い項目名ほど、優先順位が高くなるようにルールリストを生成する。また、ルールリスト生成部2 2 1は、生成したルールリストにルールリストIDを割り当てる。

[0054] 次に、ルール辞書生成部2 2 2は、生成済みのルールリストの何れにも適合していないサンプル帳票画像の中から一つのサンプル帳票画像を選択する（ステップS 1 0 7）。

[0055] 次に、ルール辞書生成部2 2 2は、ステップS 1 0 6で生成されたルールリストに基づいて、選択したサンプル帳票画像から項目値を抽出する（ステップS 1 0 8）。ルール辞書生成部2 2 2は、公知のOCR技術を利用して、ルールリストに規定された優先順位の順（優先順位が高い順）に、ルールリストに規定された各項目名がサンプル帳票画像に含まれるか否かを順次判定していく。ルール辞書生成部2 2 2は、最初に含まれると判定された項目名に対して、その項目名に関連付けられた位置情報に規定された表示位置関

係に位置する所定サイズの領域から項目値を検出し、検出した項目値をその項目名に対応する項目値として抽出する。

[0056] 次に、ルール辞書生成部222は、抽出した項目値と、選択したサンプル帳票画像に関連付けられた正解値とが一致するか否かを判定する（ステップS109）。

[0057] 項目値と正解値とが一致しない場合、ルール辞書生成部222は、ステップS106で生成されたルールリストと、ステップS107で選択されたサンプル帳票画像が適合しないと判定し、処理をステップS113に移行する。

[0058] 一方、項目値と正解値とが一致する場合、ルール辞書生成部222は、ステップS106で生成されたルールリストと、ステップS107で選択されたサンプル帳票画像が適合すると判定する（ステップS110）。

[0059] 次に、ルール辞書生成部222は、ルールリストに適合すると判定されたサンプル帳票画像から特徴量を算出する（ステップS111）。ルール辞書生成部222は、サンプル帳票画像の特徴量として、罫線の形状に関する特徴量、もしくは、ロゴに関する特徴量、又は、文字認識結果に関する特徴量を算出する。

[0060] 図8は、罫線の形状に関する特徴量について説明するための図である。

[0061] 図8に示すように、帳票画像800に表801が含まれる場合、表801を構成する各罫線802によって形状の異なる複数の頂点803又は交点804が形成される。ルール辞書生成部222は、形状の異なる頂点803及び交点804毎に、各形状を有する頂点803及び交点804の数を計数する。ルール辞書生成部222は、各形状を有する頂点803及び交点804の数を要素とするヒストグラム（特徴ベクトル）810を帳票画像800の特徴量として算出する。ルール辞書生成部222は、罫線の形状に関する特徴量を用いることにより、表が含まれる可能性の高い多種レイアウトの各帳票を高精度に識別することができる。

[0062] 図9は、ロゴに関する特徴量について説明するための図である。

[0063] 図9に示すように、一般に、帳票画像900には、その帳票を発行する機関を表す、模様のロゴ901又は文字のロゴ902等の様々なロゴ画像が含まれる。ルール辞書生成部222は、公知のパターンマッチング技術を利用して、各形状を有するロゴ901、902を検出し、各形状を有するロゴ901、902の数を計数する。例えば、生成装置200は、各ロゴ901、902の画像を二値化した正解画像を予め記憶しておく。ルール辞書生成部222は、帳票画像900内で切り出し範囲を移動させながら、画像を順次切り出し、切り出した画像を二値化した二値化画像と正解画像との画素の一致率が閾値以上である場合に、その切り出した範囲をロゴとして検出する。ルール辞書生成部222は、各形状を有するロゴ901、902の数を要素とするヒストグラム（特徴ベクトル）910を帳票画像900の特徴量として算出する。ルール辞書生成部222は、各機関のロゴに関する特徴量を用いることにより、各機関によって個別に作成される各帳票を高精度に識別することができる。

[0064] 図10は、文字認識結果に関する特徴量について説明するための図である。

[0065] 図10に示すように、一般に、帳票画像1000には、様々な文字列1001～1005が含まれる。ルール辞書生成部222は、公知のOCR技術を利用して、帳票画像1000から文字列1001～1005を検出し、各文字列1001～1005の数を計数する。ルール辞書生成部222は、各文字列1001～1005の数を要素とするヒストグラム（特徴ベクトル）1010を帳票画像1000の特徴量として算出する。ルール辞書生成部222は、文字認識結果に関する特徴量を用いることにより、多種レイアウトの各帳票をより高精度に識別することができる。

[0066] なお、ルール辞書生成部222は、罫線の形状に関する特徴量、ロゴに関する特徴量、又は、文字認識結果に関する特徴量の内の二つ以上を組み合わせ、サンプル帳票画像の特徴量として用いてもよい。これにより、ルール辞書生成部222は、各帳票をより高精度に識別することが可能となる。ま

た、生成装置 200 は、各サンプル帳票画像の特徴量を事前に算出しておき、第 2 記憶装置 210 に記憶しておいてもよい。その場合、ルール辞書生成部 222 は、第 2 記憶装置 210 から各サンプル帳票画像の特徴量を取得する。これにより、ルール辞書生成部 222 は、生成処理にかかる処理負荷を低減し、より短時間でルール辞書を生成することが可能となる。

[0067] 次に、ルール辞書生成部 222 は、算出した特徴量を、ステップ S106 で生成されたルールリストのルールリスト ID と関連付けてルール辞書に記憶（追加）する（ステップ S112）。

[0068] なお、ルール辞書生成部 222 は、算出した特徴量が過去に算出した特徴量と一致する場合、ルール辞書の更新を省略する。また、ルール辞書生成部 222 は、算出した特徴量が過去に算出した特徴量と類似する場合も、ルール辞書の更新を省略してもよい。その場合、ルール辞書生成部 222 は、算出した特徴量と、過去に算出した特徴量の類似度が閾値以上である場合に、二つの特徴量が類似すると判定する。類似度は、例えば各特徴量（特徴ベクトル）のユークリッド距離の逆数、又は、コサイン類似度等である。また、ルール辞書生成部 222 は、算出した特徴量が過去に算出した特徴量と一致又は類似する場合、ルール辞書に既に記憶されている、その特徴量に関連付けられたルールリスト ID を新たに生成したルールリストのルールリスト ID に更新（上書き）してもよい。

[0069] また、ルール辞書生成部 222 は、算出した特徴量に特徴量 ID を割り当てて、割り当てた特徴量 ID をルールリスト ID と関連付けてルール辞書に記憶してもよい。例えば、ルール辞書生成部 222 は、算出した特徴量（特徴ベクトル）の各要素を所定の有効桁数に丸めた値（ベクトル）を特徴量 ID として使用する。なお、ルール辞書生成部 222 は、算出した特徴量（特徴ベクトル）を正規化した値（ベクトル）を特徴量 ID として使用してもよい。

[0070] 次に、ルール辞書生成部 222 は、生成済みのルールリストの何れにも適合していない全てのサンプル帳票画像についてステップ S107～S112

の処理が完了したか否かを判定する（ステップS 1 1 3）。まだ処理が完了していないサンプル帳票画像が存在する場合、ルール辞書生成部2 2 2は、処理をステップS 1 0 7に戻し、未選択のサンプル帳票画像に対して、ステップS 1 0 7～S 1 1 2の処理を実行する。

[0071] 一方、各サンプル帳票画像について処理が完了した場合、ルール辞書生成部2 2 2は、全てのサンプル帳票画像が生成済みのルールリストの何れかと適合したか否かを判定する（ステップS 1 1 4）。生成済みのルールリストの何れにも適合しないサンプル帳票画像が存在する場合、ルール辞書生成部2 2 2は、処理をステップS 1 0 2に戻す。この場合、ルールリスト生成部2 2 1は、ステップS 1 0 2～S 1 0 6の処理により、生成済みのルールリストの何れにも適合していないサンプル帳票画像に基づいて、新たなルールリストを生成する。一方、ルール辞書生成部2 2 2は、ステップS 1 0 7～S 1 1 2の処理により、新たに生成されたルールリストを用いて、ルール辞書を更新する。

[0072] 一方、全てのサンプル帳票画像が生成済みのルールリストの何れかと適合した場合、ルール辞書生成部2 2 2は、一連のステップを終了する。生成処理が終了した時点で第2記憶装置2 1 0に記憶されているルール辞書が、ルール辞書生成部2 2 2により生成されたルール辞書となる。

[0073] ステップS 1 0 6のように、ルールリスト生成部2 2 1は、生成済みのルールリストの何れにも適合していない各サンプル帳票画像における項目値と正解値の一致率に基づいて、次に各サンプル帳票画像との適合検査を行うルールリストにおける優先順位を決定する。これにより、ルール辞書生成部2 2 2は、各サンプル帳票画像と各ルールリストが適合するかを効率良く検査することが可能となり、ルール辞書をよりシンプルにし且つルール辞書の生成時間を低減することが可能となる。

[0074] なお、ルール辞書が生成された後にサンプル帳票画像が新たに追加された場合、ルールリスト生成部2 2 1は、新たに追加されたサンプル帳票画像に基づいて、新たなルールリストを生成する。その場合、ルールリスト生成部

221は、新たに追加されたサンプル帳票画像を、生成済みのルールリストの何れにも適合していないサンプル帳票画像として、ステップS102～S106の処理を実行し、新たなルールリストを生成する。一方、ルール辞書生成部222は、ステップS107～S112の処理により、新たに生成されたルールリストを用いて、ルール辞書を更新する。但し、新たに生成されたルールリストが生成済みのルールリストの何れかと同一である場合、ルール辞書生成部222は、その生成済みのルールリストを用いて、ルール辞書を更新する。

[0075] これにより、生成装置200は、一旦生成されたルール辞書に対して、新たにサンプル帳票画像を追加して、ルール辞書を更新することができる。特に生成済みのルールリストの何れにも適合しないサンプル帳票画像が新たに追加された場合、新たに追加された各サンプル帳票画像における項目値と正解値の一致率に基づいて、各サンプル帳票画像との適合検査を行うルールリストにおける優先順位が決定される。これにより、ルール辞書生成部222は、新たに追加された各サンプル帳票画像と各ルールリストが適合するかを効率良く検査することが可能となり、ルール辞書をよりシンプルにし且つルール辞書の生成時間を低減することが可能となる。

[0076] 図11～図13は、図7に示した生成処理により生成されるルール辞書の一例を示す図である。

[0077] 図11～図13では、図11に示す9個のサンプル帳票画像1～9に対して、生成処理が実行されている。まず、図11に示すように、各サンプル帳票画像1～9から項目名「患者負担額」に対応する各項目値が検出され（ステップS103）、検出された各項目値と、各サンプル帳票画像1～9に関連付けられた正解値との一致率が算出される（ステップS104）。同様に、項目名「一部負担金」に対応する各項目値、項目名「合計」に対応する各項目値についても一致率が算出され、一致率が高い順に優先順位が割り当てられたルールリスト1101が生成される（ステップS106）。この例では、ルールリスト1101において、患者負担額、一部負担金、合計の順に

優先順位が割り当てられている。

[0078] 次に、サンプル帳票画像 1 に対して、ルールリスト 1 1 0 1 に規定された優先順位が高い順（患者負担額、一部負担金、合計の順）に各項目名が含まれるか否かが判定され、最初に含まれると判定された項目名に対応する項目値が抽出される（ステップ S 1 0 8）。そして、抽出された項目値と、サンプル帳票画像 1 に関連付けられた正解値とが一致する場合、サンプル帳票画像 1 の特徴量 A とルールリスト 1 1 0 1 のルールリスト I D とが関連付けてルール辞書 1 1 0 2 に記憶される（ステップ S 1 0 9）。同様に、サンプル帳票画像 2 ～ 9 から抽出された項目値と、サンプル帳票画像 2 ～ 9 に関連付けられた正解値とが一致する場合、サンプル帳票画像 2 ～ 9 の特徴量 B ～ I とルールリスト 1 1 0 1 のルールリスト I D とが関連付けてルール辞書 1 1 0 2 に記憶される。

[0079] この例ではサンプル帳票画像 1 ～ 6 から抽出された項目値と、サンプル帳票画像 1 ～ 6 に関連付けられた正解値とが一致して、サンプル帳票画像 1 ～ 6 の特徴量 A ～ F とルールリスト 1 1 0 1 のルールリスト I D とが関連付けてルール辞書 1 1 0 2 に記憶されている。一方、サンプル帳票画像 7 ～ 9 から抽出された項目値と、サンプル帳票画像 7 ～ 9 に関連付けられた正解値とは一致せず、サンプル帳票画像 7 ～ 9 の特徴量 G ～ I は、まだルール辞書 1 1 0 2 に記憶されていない。

[0080] 次に、図 1 2 に示すように、ルールリスト 1 1 0 1 と適合しなかったサンプル帳票画像 7 ～ 9 からルールリスト 1 2 0 1 が生成される。この例では、ルールリスト 1 2 0 1 において、一部負担金、患者負担額、合計の順に優先順位が割り当てられている。

[0081] 次に、サンプル帳票画像 7 ～ 9 に対して、ルールリスト 1 2 0 1 を用いて、サンプル帳票画像 7 ～ 9 から抽出した項目値と、サンプル帳票画像 7 ～ 9 に関連付けられた正解値とが一致するか否かが判定され、ルール辞書 1 2 0 2 が更新される。この例ではサンプル帳票画像 7 ～ 8 から抽出された項目値と、サンプル帳票画像 7 ～ 8 に関連付けられた正解値とが一致して、サン

ル帳票画像 7～8 の特徴量 G～H とルールリスト 1201 のルールリスト ID とが関連付けてルール辞書 1202 に記憶されている。一方、サンプル帳票画像 9 から抽出された項目値と、サンプル帳票画像 9 に関連付けられた正解値とは一致せず、サンプル帳票画像 9 の特徴量 I は、まだルール辞書 1202 に記憶されていない。

[0082] 次に、図 13 に示すように、ルールリスト 1201 と適合しなかったサンプル帳票画像 9 からルールリスト 1301 が生成される。この例では、合計、患者負担額、一部負担金の順に優先順位が割り当てられている。

[0083] 次に、サンプル帳票画像 9 に対して、ルールリスト 1301 を用いて、サンプル帳票画像 9 から抽出した項目値と、サンプル帳票画像 9 に関連付けられた正解値とが一致するか否かが判定され、ルール辞書 1302 が更新される。この例では、サンプル帳票画像 9 から抽出された項目値と、サンプル帳票画像 9 に関連付けられた正解値とが一致して、サンプル帳票画像 9 の特徴量 I とルールリスト 1301 のルールリスト ID とが関連付けてルール辞書 1302 に記憶されている。以上により、全てのサンプル帳票画像 1～9 に対する処理が完了し、ルール辞書 1302 が完成している。

[0084] 図 14 は、情報処理装置 100 による帳票画像からの項目値の検出処理の動作の例を示すフローチャートである。

[0085] 以下、図 14 に示したフローチャートを参照しつつ、情報処理装置 100 による検出処理の動作の例を説明する。なお、以下に説明する動作のフローは、予め第 1 記憶装置 110 に記憶されているプログラムに基づき主に第 1 CPU 120 により情報処理装置 100 の各要素と協働して実行される。

[0086] 最初に、受信部 121 は、第 1 通信装置 101 を介して生成装置 200 に、ルール辞書及びルールリストの取得を要求する取得要求を送信し、第 1 通信装置 101 を介して生成装置 200 からルール辞書及びルールリストを受信する（ステップ S201）。なお、生成装置 200 の送信部 223 は、第 2 通信装置 201 を介して情報処理装置 100 から取得要求を受信すると、事前に生成されたルール辞書及びルールリストを第 2 記憶装置 210 から読

出し、第2通信装置201を介して情報処理装置100に送信する。受信部121は、受信したルール辞書及びルールリストを第1記憶装置110に記憶する。なお、情報処理装置100は、事前にルール辞書及びルールリストを取得して第1記憶装置110に記憶しておいてもよい。

[0087] 次に、取得部122は、第1通信装置101を介して不図示の画像読取装置から、項目値を検出する対象となる帳票画像を取得する（ステップS202）。なお、取得部122は、USB（Universal Serial Bus）等の不図示のインタフェース回路を介して、画像読取装置から帳票画像を取得してもよい。または、情報処理装置100が撮像装置を有し、取得部122は、撮像装置により帳票を撮像して帳票画像を取得してもよい。

[0088] 次に、特徴量算出部123は、図7のステップS111の処理と同様にし、取得部122により取得された帳票画像から特徴量を算出する（ステップS203）。特徴量算出部123は、生成装置200のルール辞書生成部222が算出する特徴量と同じ特徴量を算出する。なお、ルール辞書生成部222が、特徴量IDをルールリストIDと関連付けてルール辞書に記憶している場合、特徴量算出部123は、ルール辞書生成部222と同じ方式により、算出した特徴量に特徴量IDを割り当てる。その場合、情報処理装置100は、特徴量の代わりに特徴量IDを用いて以下の処理を実行する。

[0089] 次に、ルールリスト抽出部124は、第1記憶装置110に記憶されたルール辞書に、特徴量算出部123により算出された特徴量が記憶されているか否かを判定する（ステップS204）。なお、ルールリスト抽出部124は、ルール辞書に、特徴量算出部123により算出された特徴量と同一の特徴量だけでなく、類似する特徴量が記憶されている場合も、特徴量算出部123により算出された特徴量が記憶されていると判定してもよい。その場合、ルールリスト抽出部124は、特徴量算出部123により算出された特徴量と、ルール辞書に記憶されている特徴量の類似度が閾値以上である場合に、二つの特徴量が類似すると判定する。類似度は、例えば各特徴量（特徴ベクトル）のユークリッド距離の逆数、又は、コサイン類似度等である。

[0090] ルール辞書に、特徴量算出部 1 2 3 により算出された特徴量が記憶されている場合、ルールリスト抽出部 1 2 4 は、ルール辞書において、その特徴量に関連付けられたルールリスト I D に対応するルールリストを抽出する（ステップ S 2 0 5）。

[0091] 一方、ルール辞書に、特徴量算出部 1 2 3 により算出された特徴量が記憶されていない場合、即ち、その特徴量に関連付けられたルールリストが存在しない場合、ルールリスト抽出部 1 2 4 は、ルール辞書から、代替ルールリストを抽出する（ステップ S 2 0 6）。ルールリスト抽出部 1 2 4 は、例えば、ルール辞書において最も多くの特徴量に関連付けられたルールリストを代替ルールリストとして抽出する。これにより、ルールリスト抽出部 1 2 4 は、ルール辞書に特徴量が登録されていないような帳票画像からも、注目する項目の項目値を精度良く検出することが可能となる。なお、ルールリスト抽出部 1 2 4 は、他の条件によって代替ルールリストを抽出してもよい。

[0092] 次に、検出部 1 2 5 は、ルールリスト抽出部 1 2 4 により抽出されたルールリストに基づいて、帳票画像から項目値を検出する（ステップ S 2 0 7）。検出部 1 2 5 は、公知の O C R 技術を利用して、ルールリストに規定された優先順位の順（優先順位が高い順）に、ルールリストに規定された各項目名が帳票画像に含まれるか否かを順次判定していく。検出部 1 2 5 は、最初に含まれると判定された項目名に対して、その項目名に関連付けられた位置情報に規定された表示位置関係に位置する所定サイズの領域から項目値を検出し、その項目名に対応する項目値として検出する。このように、検出部 1 2 5 は、ルールリストにおいて規定された表示位置関係に基づいて、各項目名に対応する項目値を検出する。

[0093] 次に、出力制御部 1 2 6 は、検出部 1 2 5 により検出された項目値を含む認識結果画面を表示装置 1 0 3 に表示することにより、その項目値を出力する（ステップ S 2 0 8）。

[0094] 図 1 5 は、認識結果画面の一例を示す図である。

[0095] 図 1 5 に示すように、認識結果画面 1 5 0 0 には、注目する項目毎に、項

目 1 5 0 1、認識結果 1 5 0 2、修正ボタン 1 5 0 3 及び確定ボタン 1 5 0 4 等が表示される。帳票画像から検出された各項目の項目値は、各項目 1 5 0 1 と対応付けて認識結果 1 5 0 2 として表示される。ユーザにより入力装置 1 0 2 を用いて修正ボタン 1 5 0 3 が押下されると、認識結果 1 5 0 2 が作業者により修正（編集）可能となる。また、ユーザにより入力装置 1 0 2 を用いて確定ボタン 1 5 0 4 が押下されると、帳票画像から検出された各項目の項目値が、認識結果 1 5 0 2 において指定された修正値に修正される。なお、項目毎に表示される修正ボタン 1 5 0 3 及び確定ボタン 1 5 0 4 の代わりに、全ての項目に共通に用いられる一つの修正ボタン及び確定ボタンが表示されてもよい。その場合、修正ボタンが押下されると、全ての認識結果 1 5 0 2 が作業者により修正（編集）可能となり、確定ボタンが押下されると、各認識結果 1 5 0 2 において指定された全ての修正値が項目値に反映される。

[0096] なお、出力制御部 1 2 6 は、項目値を表示装置 1 0 3 に表示する代わりに、第 1 通信装置 1 0 1 を介して、帳票に係る各情報を集約するサーバ等に送信することにより、項目値を出力してもよい。

[0097] 次に、受付部 1 2 7 は、ユーザにより入力装置 1 0 2 を用いて確定ボタン 1 5 0 4 が押下されることにより、ユーザによる、認識結果画面 1 5 0 0 に表示された項目値の修正指示を受け付けたか否かを判定する（ステップ S 2 0 9）。ユーザによる修正指示が受け付けられなかった場合、情報処理装置 1 0 0 は、一連のステップを終了する。

[0098] 一方、ユーザによる修正指示が受け付けられた場合、更新部 1 2 8 は、更新処理を実行し（ステップ S 2 1 0）、一連のステップを終了する。更新部 1 2 8 は、更新処理で、ルール辞書において、特徴量算出部 1 2 3 により算出された特徴量に関連付けられたルールリストを更新する。更新処理の詳細については後述する。

[0099] 図 1 6 は、更新処理の動作の例を示すフローチャートである。図 1 6 に示す更新処理は、図 1 4 のステップ S 2 1 0 において実行される。

- [0100] 最初に、更新部１２８は、第１記憶装置１１０に記憶されたルールリストの中から一つのルールリストを選択する（ステップＳ３０１）。
- [0101] 次に、更新部１２８は、選択したルールリストに基づいて、帳票画像から項目値を抽出する（ステップＳ３０２）。更新部１２８は、公知のＯＣＲ技術を利用して、ルールリストに規定された優先順位の順（優先順位が高い順）に、ルールリストに規定された各項目名がサンプル帳票画像に含まれるか否かを順次判定していく。更新部１２８は、最初に含まれると判定された項目名に対して、その項目名に関連付けられた位置情報に規定された表示位置関係に位置する所定サイズの領域から項目値を検出し、その項目名に対応する項目値として抽出する。
- [0102] 次に、更新部１２８は、抽出した項目値と、ユーザにより入力装置１０２を用いて指定された修正値とが一致するか否かを判定する（ステップＳ３０３）。
- [0103] 項目値と正解値とが一致する場合、更新部１２８は、ステップＳ３０１で選択されたルールリストと、帳票画像が適合すると判定する。その場合、更新部１２８は、ステップＳ２０３で特徴量算出部１２３により算出された特徴量を、ステップＳ３０１で選択されたルールリストのルールリストＩＤと関連付けてルール辞書に記憶し（ステップＳ３０４）、一連のステップを終了する。このように、更新部１２８は、第１記憶装置１１０に記憶されたルール辞書を更新する。なお、更新部１２８は、更新したルール辞書を第１通信装置１０１を介して生成装置２００に送信し、第２記憶装置２１０に記憶されたルール辞書を更新してもよい。これにより、更新部１２８は、他の情報処理装置が使用するルール辞書についても適切に更新することが可能となる。
- [0104] また、ルール辞書に、特徴量算出部１２３により算出された特徴量が既に記憶されている場合、更新部１２８は、その特徴量に関連付けられたルールリストＩＤを更新（上書き）する。これにより、更新部１２８は、ルール辞書において不適切に登録されたルールリストを適宜更新して、ルール辞書の

品質を効率良く向上させることが可能となる。一方、ルール辞書に、特徴量算出部 123 により算出された特徴量がまだ記憶されていない場合、更新部 128 は、その特徴量を、そのルールリスト ID と関連付けてルール辞書に新たに追加する。これにより、更新部 128 は、まだ登録されていない新しい情報をルール辞書に追加して、ルール辞書の品質を効率良く向上させることが可能となる。

[0105] 一方、項目値と修正値とが一致しない場合、更新部 128 は、第 1 記憶装置 110 に記憶された全てのルールリストについて処理が完了したか否かを判定する（ステップ S305）。まだ処理が完了していないルールリストが存在する場合、更新部 128 は、処理をステップ S301 に戻し、未選択のルールリストに対して、ステップ S301～S305 の処理を実行する。一方、全てのルールリストについて処理が完了した場合、更新部 128 は、一連のステップを終了する。

[0106] 以上詳述したように、情報処理装置 100 は、帳票画像の特徴量に関連付けられたルールリストにおいて規定された優先順位に従って、帳票画像から所望の項目の項目値を検出する。多種レイアウトの帳票では、帳票毎に項目名の記載位置が異なり且つ一つの帳票の中に項目名の類義語が複数存在する可能性がある。一方、同様の特徴量を有する帳票画像では、同一の項目に対して同一の項目名が使用されている可能性が高い。情報処理装置 100 は、複数の項目毎に項目名及び項目値が表された帳票画像から、所望の項目に対応する可能性が高い順に、各項目名を検査していくことが可能となり、所望の項目の項目値をより精度良く検出することが可能となった。

[0107] また、情報処理装置 100 は、各項目名に対応する項目値の表示位置関係が規定されたルールリストを用いて、帳票画像から所望の項目の項目値を検出する。これにより、情報処理装置 100 は、帳票毎に項目名に対する項目値の記載位置が異なる多種レイアウトの帳票に係る帳票画像から、所望の項目の項目値を高精度に検出することが可能となった。

[0108] また、生成装置 200 は、ルールリストに規定された優先順位の順に各項

目名を検査していき、最初にサンプル帳票画像に含まれると判定された項目名に対応する項目値とその正解値とが一致する場合に、そのルールリストとサンプル帳票画像を関連付ける。これにより、生成装置200は、各項目名を優先順位の順に検査していくことにより所望の項目の項目値を正しく検出できるルールリストを生成することが可能となった。

[0109] また、生成装置200は、複数のサンプル帳票画像を用いて事前学習することにより、ルール辞書を生成する。これにより、生成装置200は、ルール辞書に含まれるルールリストの数が膨大な量になることを抑制しつつ、レイアウトが未知である帳票画像に対しても、所望の項目の項目値を高精度に検出できるルール辞書を生成することが可能となった。

[0110] 図17は、他の実施形態に係る情報処理装置における第1処理回路320の概略構成を示すブロック図である。

[0111] 第1処理回路320は、第1CPU320の代わりに用いられ、第1CPU320の代わりに検出処理を実行する。第1処理回路320は、受信回路321、取得回路322、特徴量算出回路323、ルールリスト抽出回路324、検出回路325、出力制御回路326、受付回路327及び更新回路328等を有する。

[0112] 受信回路321は、受信部の一例であり、受信部121と同様の機能を有する。受信回路321は、第1通信装置101を介して生成装置からルール辞書及びルールリストを受信し、受信したルール辞書及びルールリストを第1記憶装置110に記憶する。

[0113] 取得回路322は、取得部の一例であり、取得部122と同様の機能を有する。取得回路322は、第1通信装置101を介して不図示の画像読取装置から帳票画像を受信し、受信した帳票画像を特徴量算出回路323及び検出回路325に出力する。

[0114] 特徴量算出回路323は、特徴量算出部の一例であり、特徴量算出部123と同様の機能を有する。特徴量算出回路323は、取得回路322から帳票画像を受信し、受信した帳票画像から特徴量を算出し、ルールリスト抽出

回路 3 2 4 に出力する。

[0115] ルールリスト抽出回路 3 2 4 は、ルールリスト抽出部の一例であり、ルールリスト抽出部 1 2 4 と同様の機能を有する。ルールリスト抽出回路 3 2 4 は、第 1 記憶装置 1 1 0 からルール辞書を読み出すとともに、特徴量算出回路 3 2 3 から特徴量を受信し、ルール辞書及び特徴量からルールリストを抽出して検出回路 3 2 5 に出力する。

[0116] 検出回路 3 2 5 は、検出部の一例であり、検出部 1 2 5 と同様の機能を有する。検出回路 3 2 5 は、取得回路 3 2 2 から帳票画像を受信するとともに、ルールリスト抽出回路 3 2 4 からルールリストを受信し、帳票画像及びルールリストから注目する項目の項目値を検出し、出力制御回路 3 2 6 に出力する。

[0117] 出力制御回路 3 2 6 は、出力制御部の一例であり、出力制御部 1 2 6 と同様の機能を有する。出力制御回路 3 2 6 は、検出回路 3 2 5 から項目値を受信し、受信した項目値を表示装置 1 0 3 に出力する。

[0118] 受付回路 3 2 7 は、受付部の一例であり、受付部 1 2 7 と同様の機能を有する。受付回路 3 2 7 は、入力装置 1 0 2 から修正値を受信し、受信した修正値を更新回路 3 2 8 に出力する。

[0119] 更新回路 3 2 8 は、更新部の一例であり、更新部 1 2 8 と同様の機能を有する。更新回路 3 2 8 は、第 1 記憶装置 1 1 0 からルールリストを読み出すとともに、受付回路 3 2 7 から修正値を受信し、ルールリスト及び修正値からルール辞書を更新する。

[0120] 図 1 8 は、他の実施形態に係る生成装置における第 2 処理回路 4 2 0 の概略構成を示すブロック図である。

[0121] 第 2 処理回路 4 2 0 は、第 2 C P U 2 2 0 の代わりに用いられ、第 2 C P U 2 2 0 の代わりに生成処理を実行する。第 2 処理回路 4 2 0 は、ルールリスト生成回路 4 2 1、ルール辞書生成回路 4 2 2 及び送信回路 4 2 3 等を有する。

[0122] ルールリスト生成回路 4 2 1 は、ルールリスト生成部の一例であり、ルー

ルリスト生成部 221 と同様の機能を有する。ルールリスト生成回路 421 は、第 2 記憶装置 210 からサンプル帳票画像を読み出し、読み出したサンプル帳票画像に基づいてルールリストを生成して第 2 記憶装置 210 に記憶する。

[0123] ルール辞書生成回路 422 は、ルール辞書生成部の一例であり、ルール辞書生成部 222 と同様の機能を有する。ルール辞書生成回路 422 は、第 2 記憶装置 210 からサンプル帳票画像及びルールリストを読み出し、読み出したサンプル帳票画像及びルールリストに基づいてルール辞書を生成して第 2 記憶装置 210 に記憶する。

[0124] 送信回路 423 は、送信部の一例であり、送信部 223 と同様の機能を有する。送信回路 423 は、第 2 記憶装置 210 からルール辞書を読み出し、読み出したルール辞書を第 2 通信装置 201 を介して情報処理装置に送信する。

[0125] 以上詳述したように、情報処理装置が第 1 処理回路 320 を用い、生成装置が第 2 処理回路 420 を用いる場合においても、情報処理装置は、所望の項目の項目値をより精度良く検出することが可能となった。

[0126] 以上、本発明の好適な実施形態について説明してきたが、本発明はこれらの実施形態に限定されるものではない。例えば、生成処理は、生成装置 200 により実行されるのではなく、情報処理装置 100 により実行されてもよい。その場合、第 1 記憶装置 110 が、第 2 記憶装置 210 が記憶する各情報を記憶し、第 1 CPU 120 が、図 2 に示した各部に加えて、第 2 CPU 220 の各部と同様の機能を持つ各部を有する。

[0127] また、各帳票画像から検出される注目する項目の項目値の数は一つに限定されない。生成装置 200 は二つ以上の項目のそれぞれについてルール辞書及びルールリストを生成し、情報処理装置 100 は各帳票画像から二つ以上の項目の項目値を検出してもよい。

符号の説明

[0128] 100 情報処理装置

- 1 0 2 入力装置
- 1 0 3 表示装置
- 1 1 0 第 1 記憶装置
- 1 2 1 受信部
- 1 2 2 取得部
- 1 2 3 特徴量算出部
- 1 2 4 ルールリスト抽出部
- 1 2 5 検出部
- 1 2 7 受付部
- 1 2 8 更新部
- 2 0 0 生成装置
- 2 1 0 第 2 記憶装置
- 2 2 1 ルールリスト生成部
- 2 2 2 ルール辞書生成部
- 2 2 3 送信部

請求の範囲

- [請求項1] 複数の項目毎に項目名及び項目値が表された帳票画像を処理する情報処理装置であって、
- 複数のサンプル帳票画像毎の特徴量のそれぞれと、複数の項目名及び各項目名の優先順位が規定された複数のルールリストのそれぞれとが関連付けられたルール辞書を記憶する記憶部と、
- 帳票画像を取得する取得部と、
- 前記取得された帳票画像から特徴量を算出する特徴量算出部と、
- 前記ルール辞書において、前記算出された特徴量に関連付けられたルールリストを抽出するルールリスト抽出部と、
- 前記抽出されたルールリストに規定された前記優先順位の順に、当該ルールリストに規定された各項目名が前記帳票画像に含まれるか否かを順次判定し、最初に含まれると判定された前記項目名に対応する項目値を検出する検出部と、
- 前記検出された項目値を出力する出力部と、
- を有することを特徴とする情報処理装置。
- [請求項2] 前記ルールリストには、前記複数の項目名、各項目名の前記優先順位及び前記サンプル帳票画像における各項目名と各項目名に対応する項目値の表示位置関係が規定され、
- 前記検出部は、前記表示位置関係に基づいて、前記最初に含まれると判定された前記項目名に対応する項目値を検出する、請求項1に記載の情報処理装置。
- [請求項3] 前記帳票画像毎の特徴量は、罫線の形状に関する特徴量、もしくは、ロゴに関する特徴量、又は、文字認識結果に関する特徴量である、請求項1または2に記載の情報処理装置。
- [請求項4] 前記ルールリスト抽出部は、前記ルール辞書に、前記算出された特徴量に関連付けられたルールリストが存在しない場合、最も多くの特徴量に関連付けられたルールリストを抽出する、請求項1～3の何れ

か一項に記載の情報処理装置。

[請求項5] ユーザによる前記出力された項目値の修正指示を受け付ける受付部と、

前記修正指示が受け付けられた場合、前記ルール辞書において、前記算出された特徴量に関連付けられたルールリストを更新する更新部と、をさらに有する、請求項1～4の何れか一項に記載の情報処理装置。

[請求項6] 生成装置と、複数の項目毎に項目名及び項目値が表された帳票画像を処理する情報処理装置とを有する情報処理システムであって、

前記生成装置は、

複数のサンプル帳票画像と、複数の項目名及び各項目名の優先順位が規定された複数のルールリストとを記憶する記憶部と、

前記複数のルールリストのそれぞれに適合する前記サンプル帳票画像の特徴量と、各ルールリストとを関連付けたルール辞書を生成するルール辞書生成部と、

前記ルール辞書を前記情報処理装置に送信する送信部と、を有し、

前記情報処理装置は、

前記生成装置から前記ルール辞書を受信する受信部と、

帳票画像を取得する取得部と、

前記取得された帳票画像から特徴量を算出する特徴量算出部と、

前記受信されたルール辞書において、前記算出された特徴量に関連付けられたルールリストを抽出するルールリスト抽出部と、

前記抽出されたルールリストに規定された前記優先順位の順に、当該ルールリストに規定された各項目名が前記帳票画像に含まれるか否かを順次判定し、最初に含まれると判定された前記項目名に対応する項目値を検出する検出部と、

前記検出された項目値を出力する出力部と、を有する、

ことを特徴とする情報処理システム。

[請求項7] 前記複数のルールリストの何れにも適合しないサンプル帳票画像が存在する場合、前記複数のルールリストの何れにも適合しないサンプル帳票画像に基づいて、新たなルールリストを生成するルールリスト生成部をさらに有する、請求項6に記載の情報処理システム。

[請求項8] 前記記憶部は、前記複数のサンプル帳票画像のそれぞれを正解値と関連付けて記憶し、

前記ルールリスト生成部は、前記複数のルールリストの何れにも適合しない各サンプル帳票画像において、前記正解値に対応している割合が高い項目名ほど、前記優先順位が高くなるように前記新たなルールリストを生成する、請求項7に記載の情報処理システム。

[請求項9] 前記記憶部は、前記複数のサンプル帳票画像のそれぞれを正解値と関連付けて記憶し、

前記ルール辞書生成部は、前記ルールリストに規定された前記優先順位の順に、前記ルールリストに規定された各項目名が前記サンプル帳票画像に含まれるか否かを順次判定し、最初に含まれると判定された前記項目名に対応する項目値と、前記サンプル帳票画像に関連付けられた正解値とが一致する場合に、前記ルールリストと前記サンプル帳票画像が適合すると判定する、請求項6に記載の情報処理システム。

[請求項10] 記憶部と、出力部とを有し、複数の項目毎に項目名及び項目値が表された帳票画像を処理する情報処理装置の制御方法であって、

複数のサンプル帳票画像毎の特徴量のそれぞれと、複数の項目名及び各項目名の優先順位が規定された複数のルールリストのそれぞれとが関連付けられたルール辞書を前記記憶部に記憶し、

帳票画像を取得し、

前記取得された帳票画像から特徴量を算出し、

前記ルール辞書において、前記算出された特徴量に関連付けられた

ルールリストを抽出し、

前記抽出されたルールリストに規定された前記優先順位の順に、当該ルールリストに規定された各項目名が前記帳票画像に含まれるか否かを順次判定し、最初に含まれると判定された前記項目名に対応する項目値を検出し、

前記検出された項目値を前記出力部に出力する、

ことを含むことを特徴とする制御方法。

[請求項11]

記憶部と、出力部とを有し、複数の項目毎に項目名及び項目値が表された帳票画像を処理する情報処理装置の制御プログラムであって、

複数のサンプル帳票画像毎の特徴量のそれぞれと、複数の項目名及び各項目名の優先順位が規定された複数のルールリストのそれぞれとが関連付けられたルール辞書を前記記憶部に記憶し、

帳票画像を取得し、

前記取得された帳票画像から特徴量を算出し、

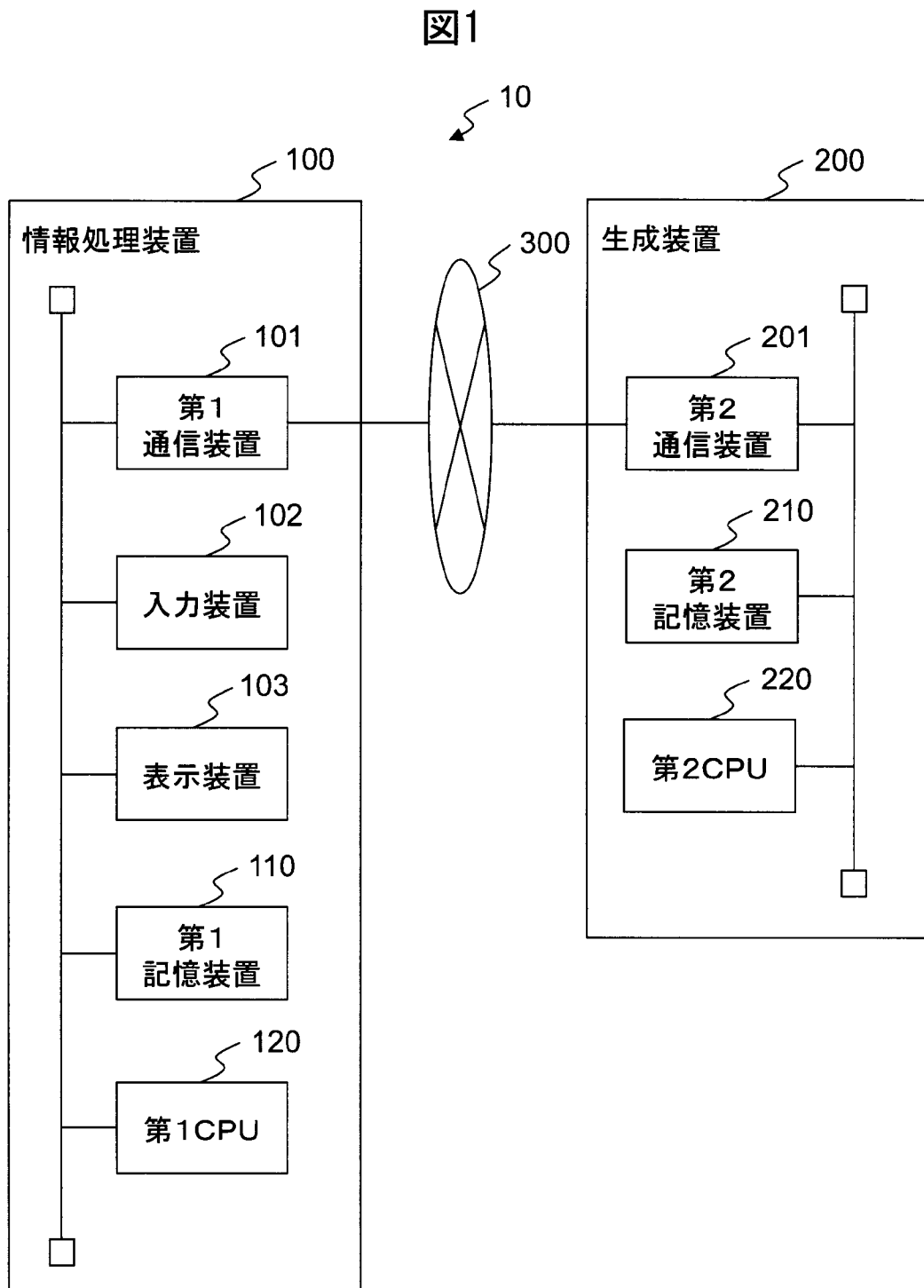
前記ルール辞書において、前記算出された特徴量に関連付けられたルールリストを抽出し、

前記抽出されたルールリストに規定された前記優先順位の順に、当該ルールリストに規定された各項目名が前記帳票画像に含まれるか否かを順次判定し、最初に含まれると判定された前記項目名に対応する項目値を検出し、

前記検出された項目値を前記出力部に出力する、

ことを前記情報処理装置に実行させることを特徴とする制御プログラム。

[図1]



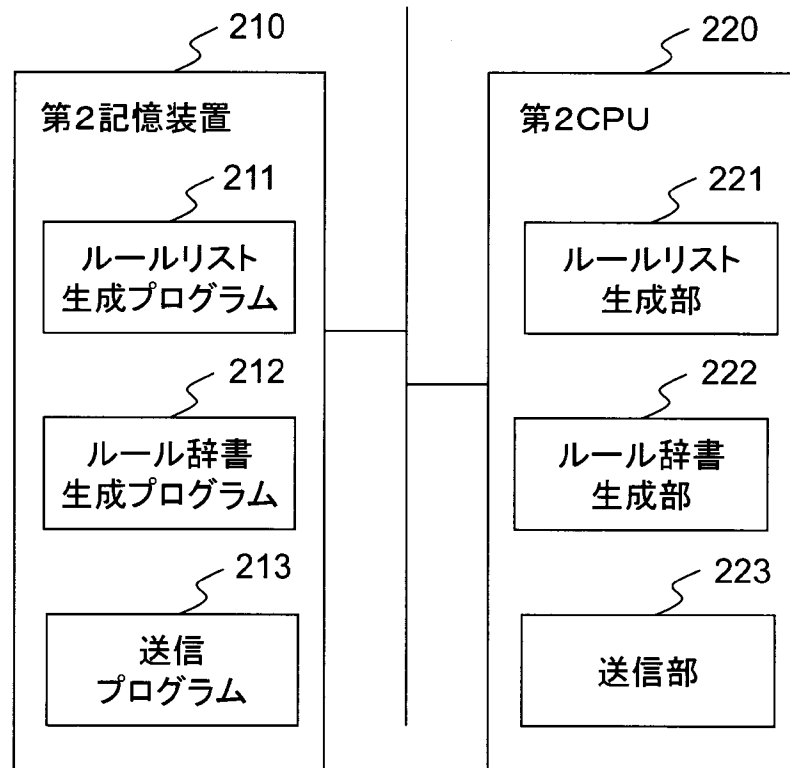
[図2]

図2



[図3]

図3



[図4]

図4

正解値=15000

入院料	30,000
手数料	
麻酔料	
放射線治療	5,000
...	...
合計	50,000
患者負担額	15,000
一部負担金	0

400

403

正解値=15000

401

入院料	手数料	麻酔料	放射線治療	...
30,000			5,000	...
		合計	患者負担額	一部負担金
		50,000	0	15,000

404

正解値=15000

入院料	30,000	患者負担額	0
手数料		一部負担金	0
麻酔料			
放射線治療	5,000		
...	...		
合計	15,000		

402

405

[図5]

図5

500

ルールリストID=1

ルール番号	優先順位	項目名	位置情報
1	1	患者負担額	右
2	2	一部負担金	下
3	3	合計	右
...	...	...	...

501

ルールリストID=2

ルール番号	優先順位	項目名	位置情報
1	2	患者負担額	右
2	1	一部負担金	下
3	3	合計	右
...	...	...	...

502

ルールリストID=3

ルール番号	優先順位	項目名	位置情報
1	2	患者負担額	右
2	3	一部負担金	下
3	1	合計	右
...	...	...	...

[図6]

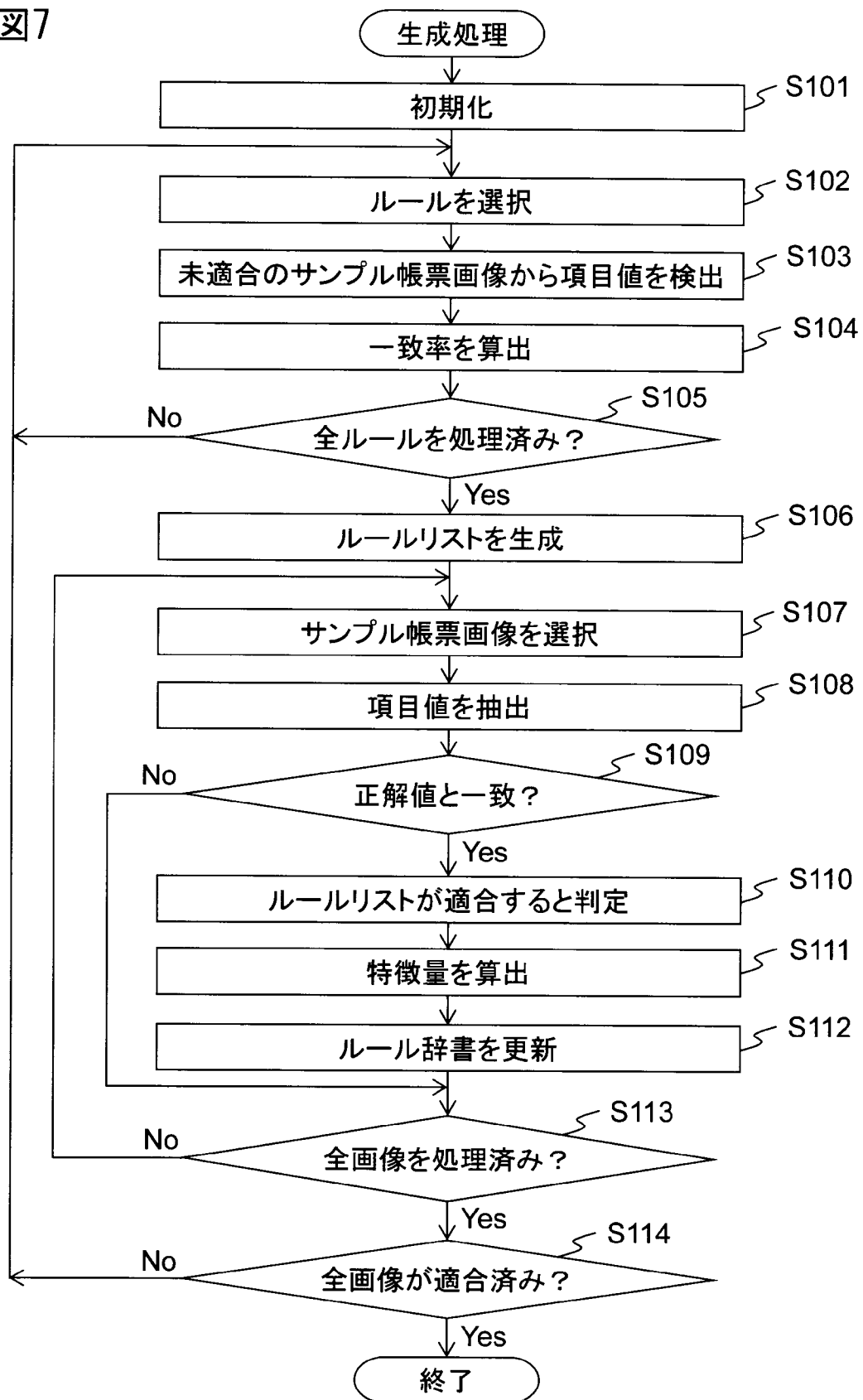
図6

 600

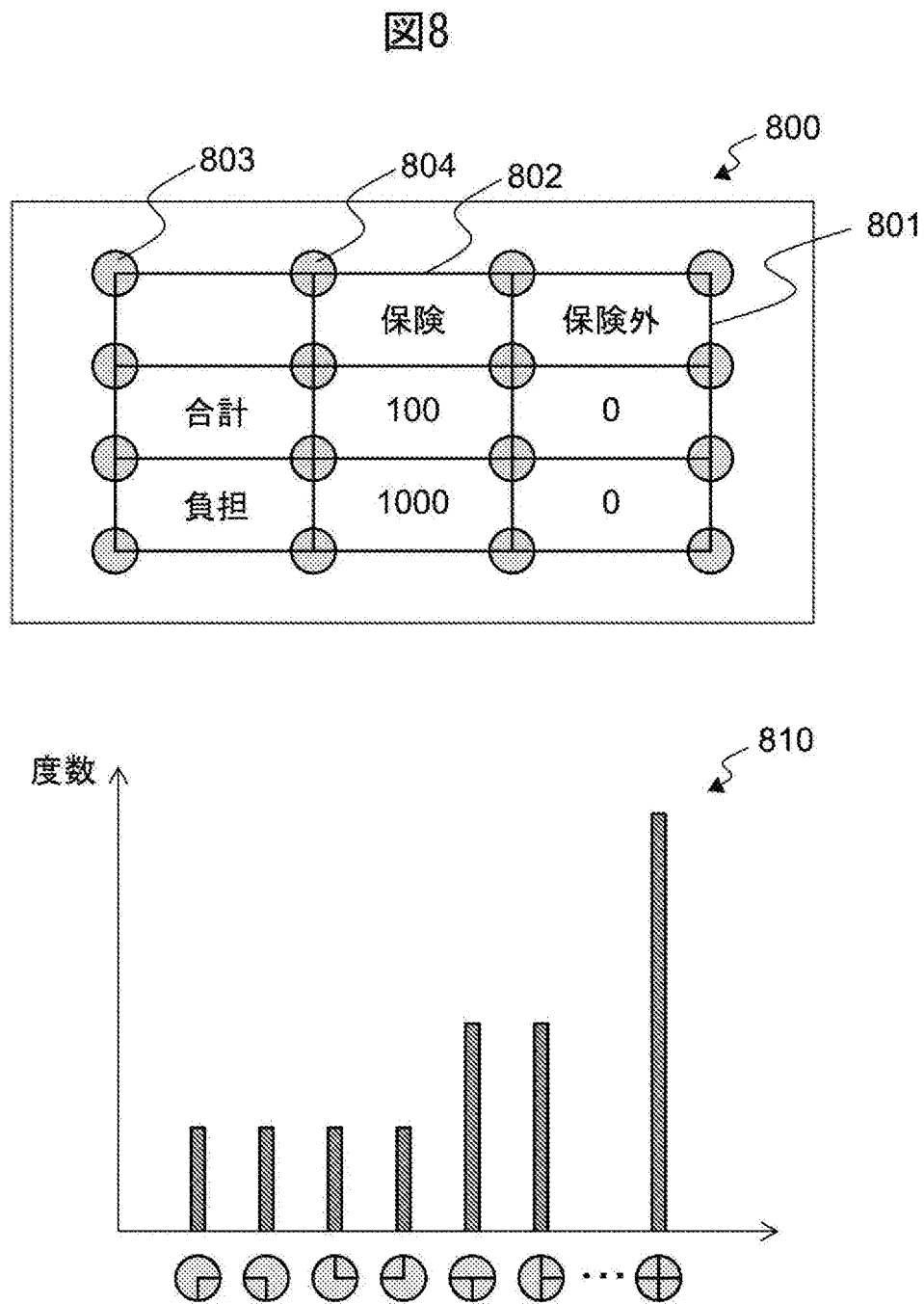
特徴量	ルールリストID
A	1
B	1
C	1
D	1
E	1
F	1
G	2
H	2
I	3
...	...

[図7]

図7

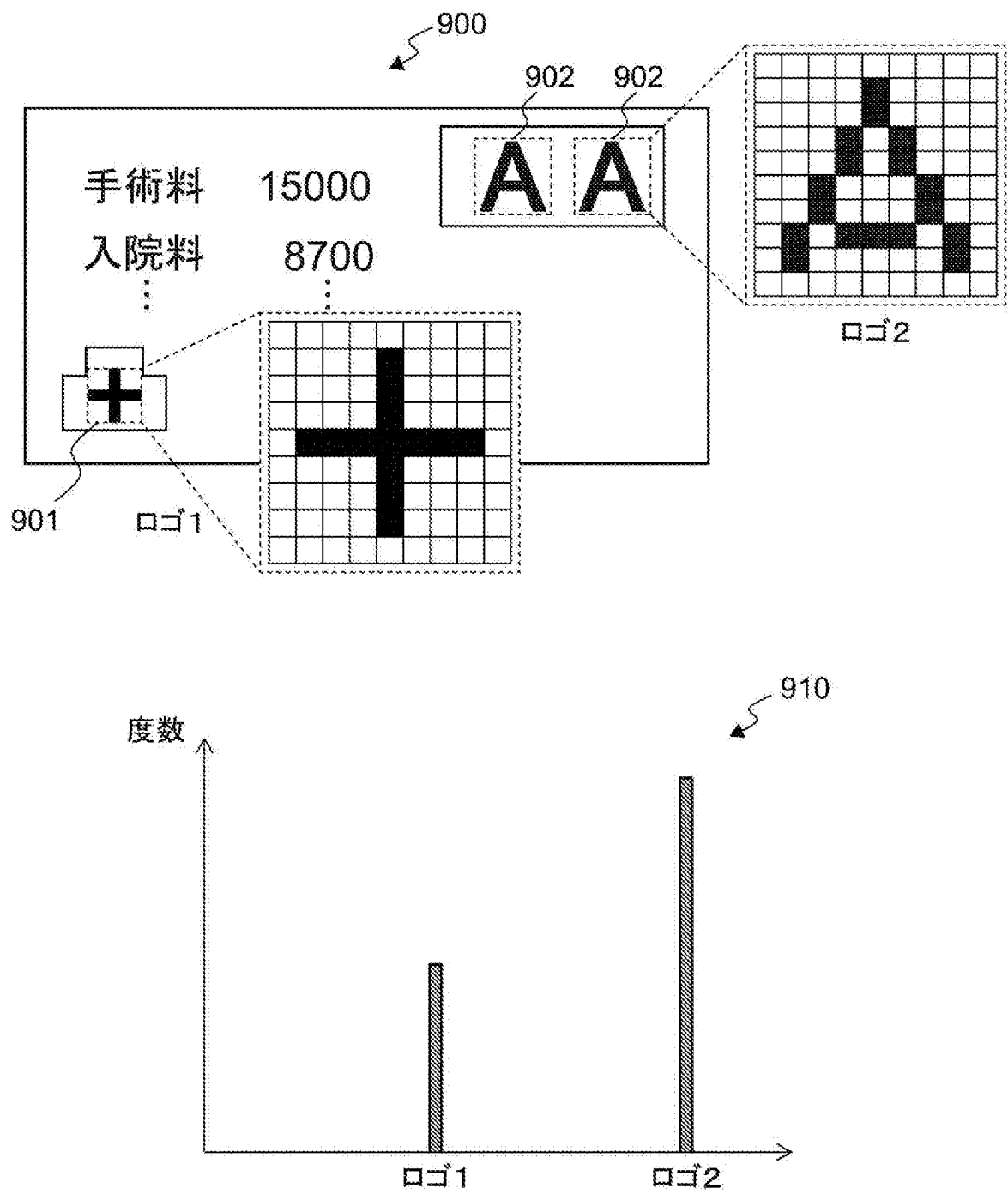


[図8]



[図9]

図9



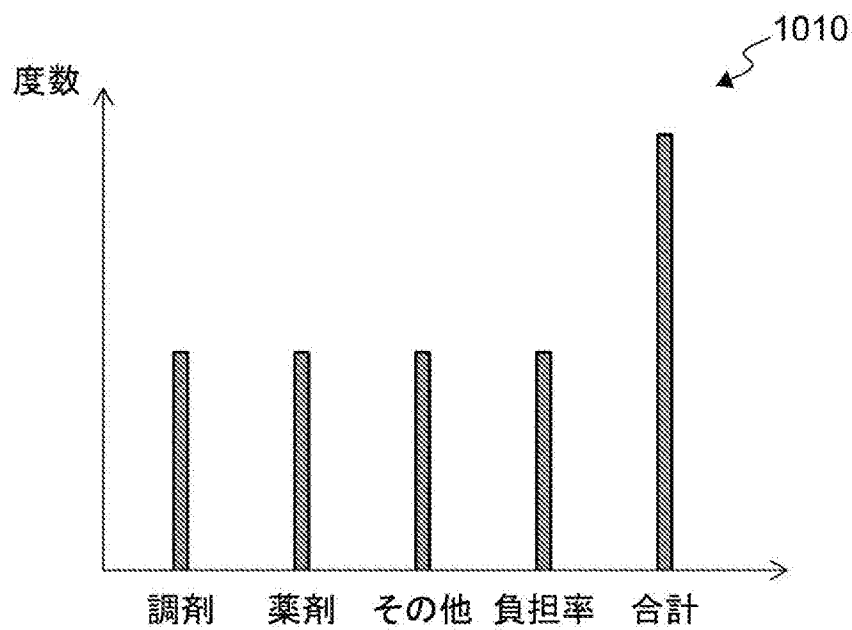
[図10]

図10

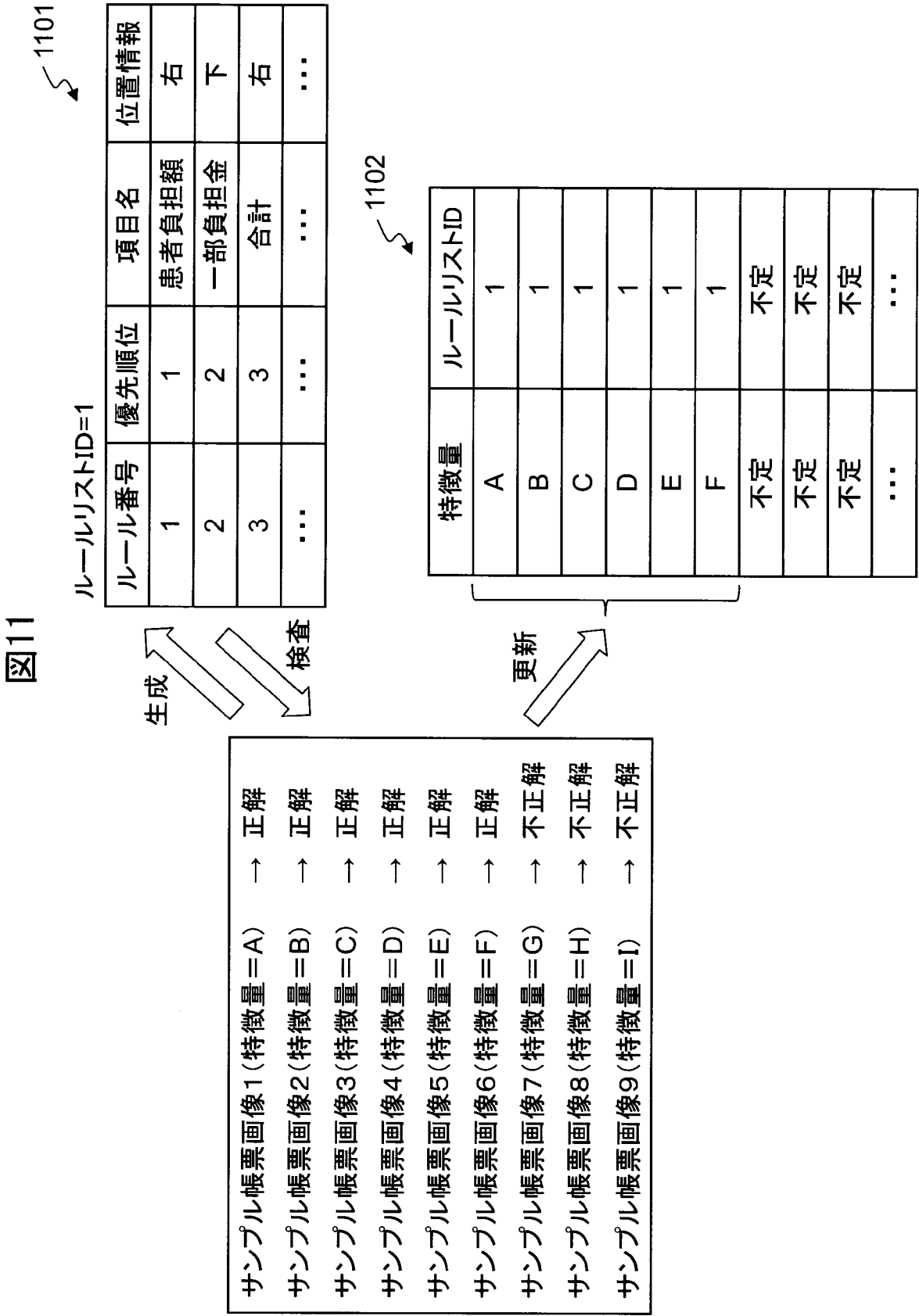
1000

1001	1002	1003	1005
調剤	薬剤	その他	合計
100	0	20	120
負担率			合計
10%			1080

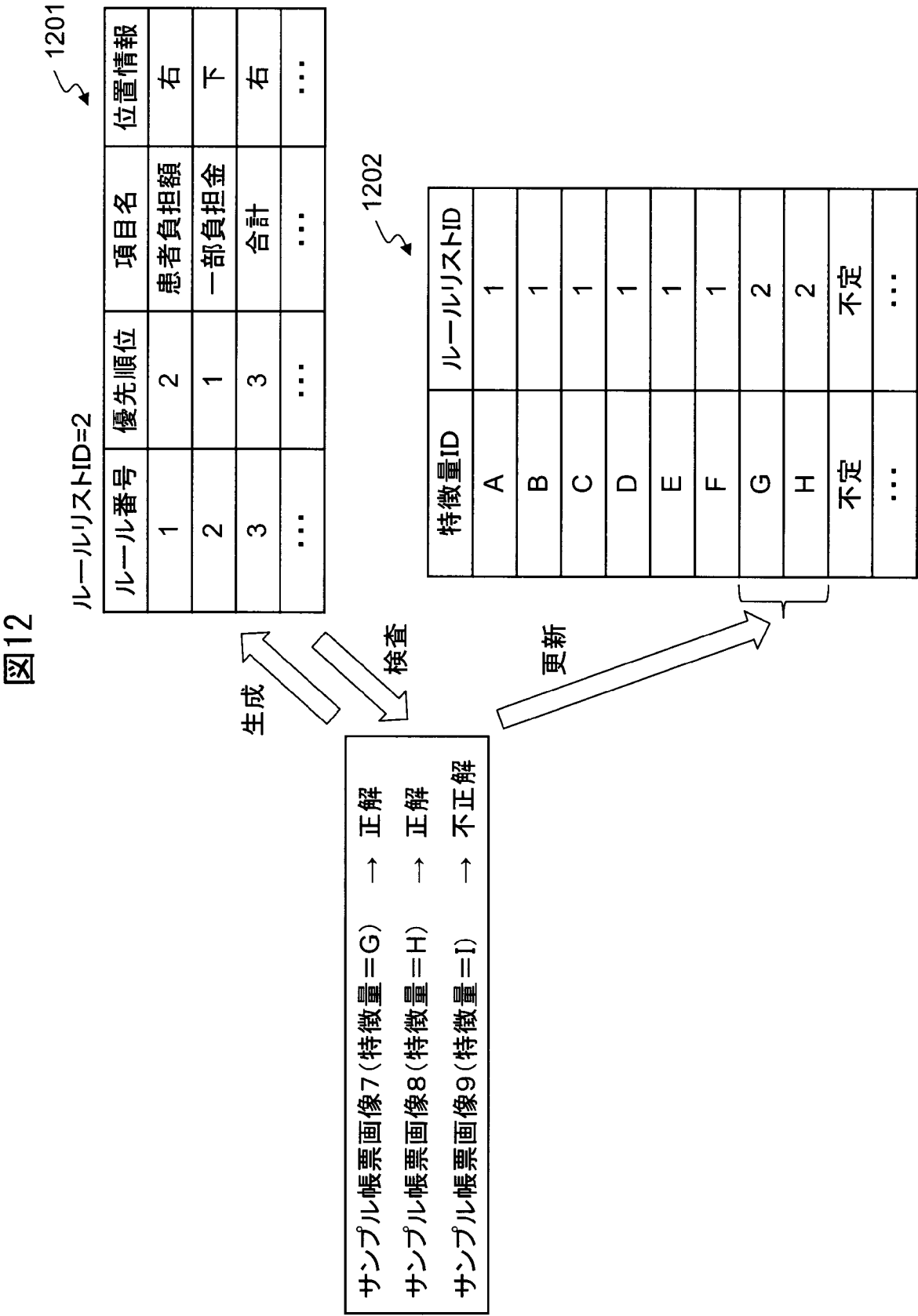
1004 1005



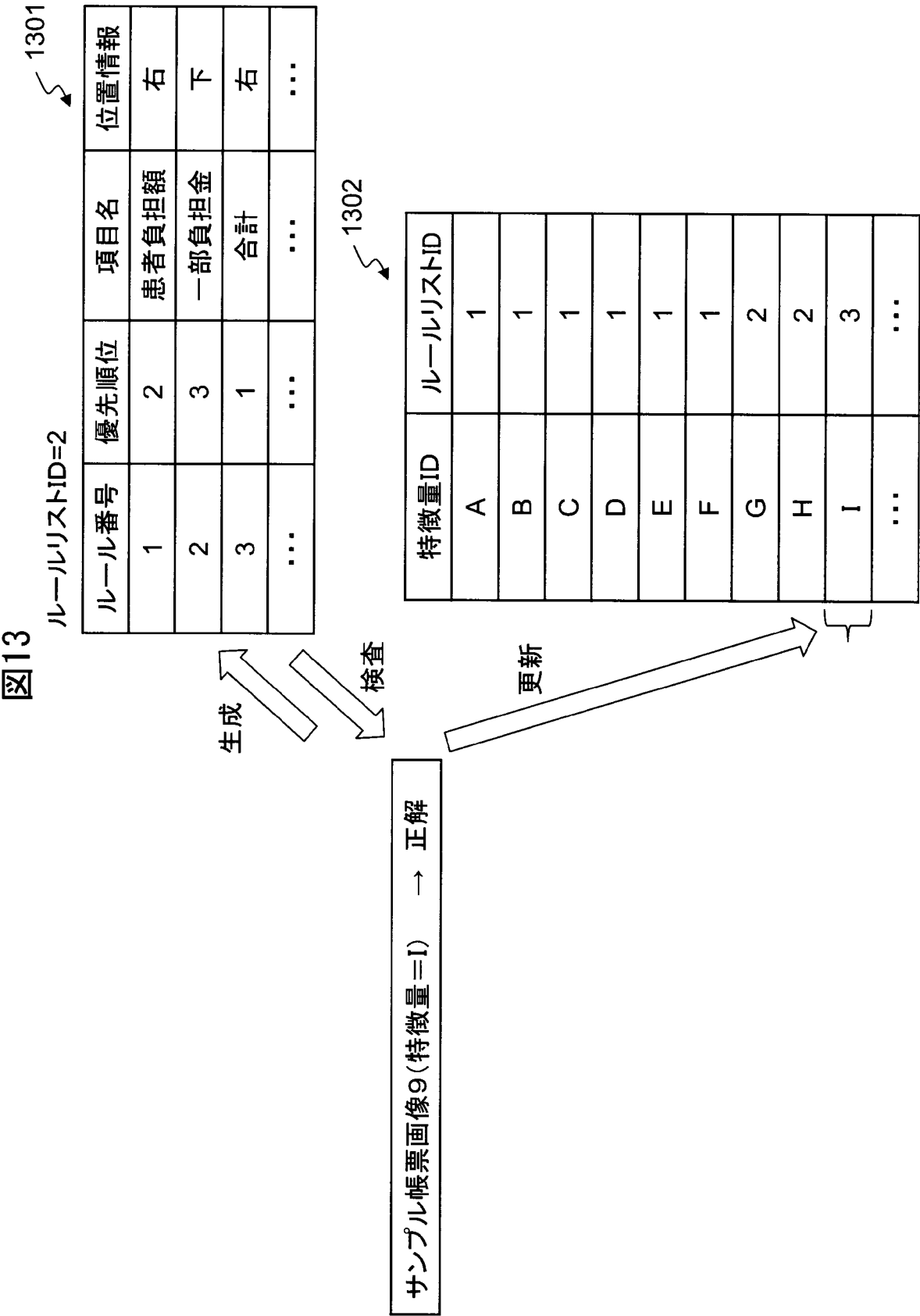
[図11]



[図12]

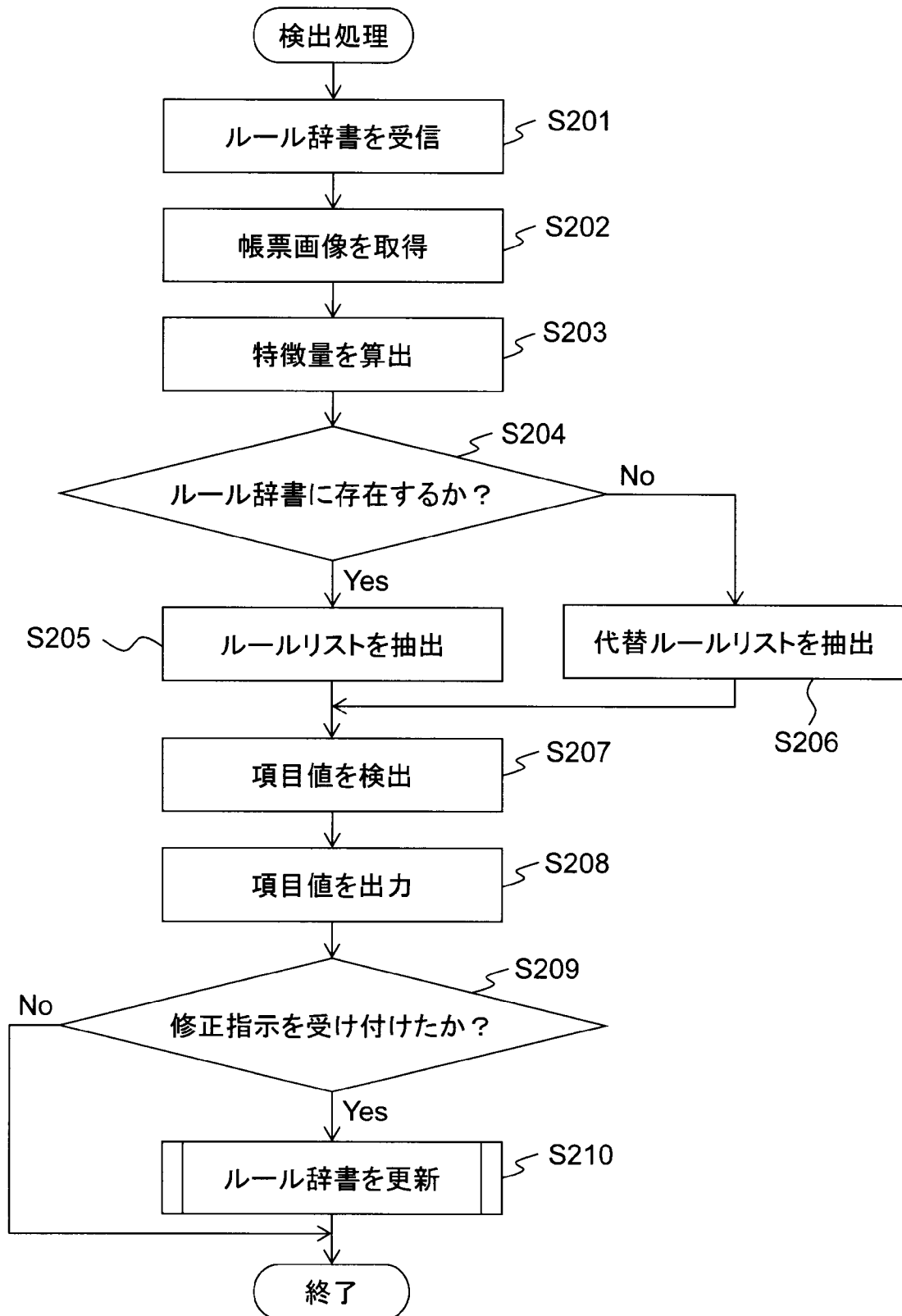


[図13]



[図14]

図14



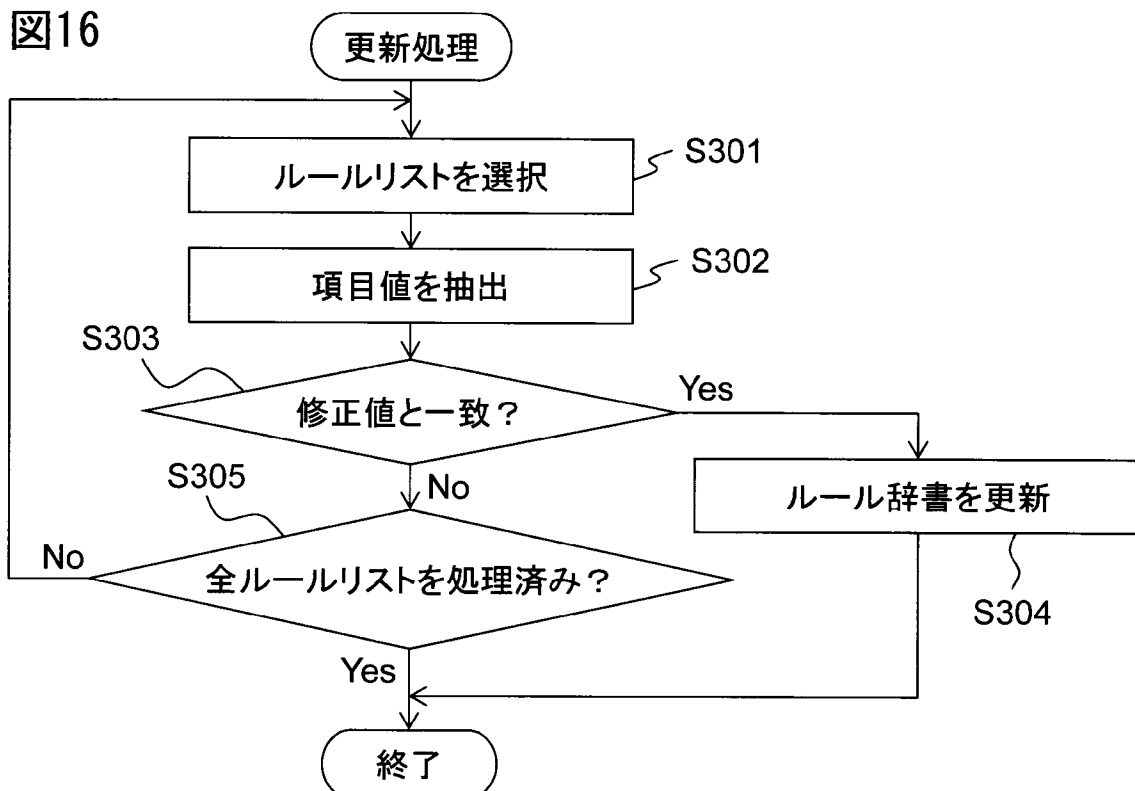
[図15]

図15

1500

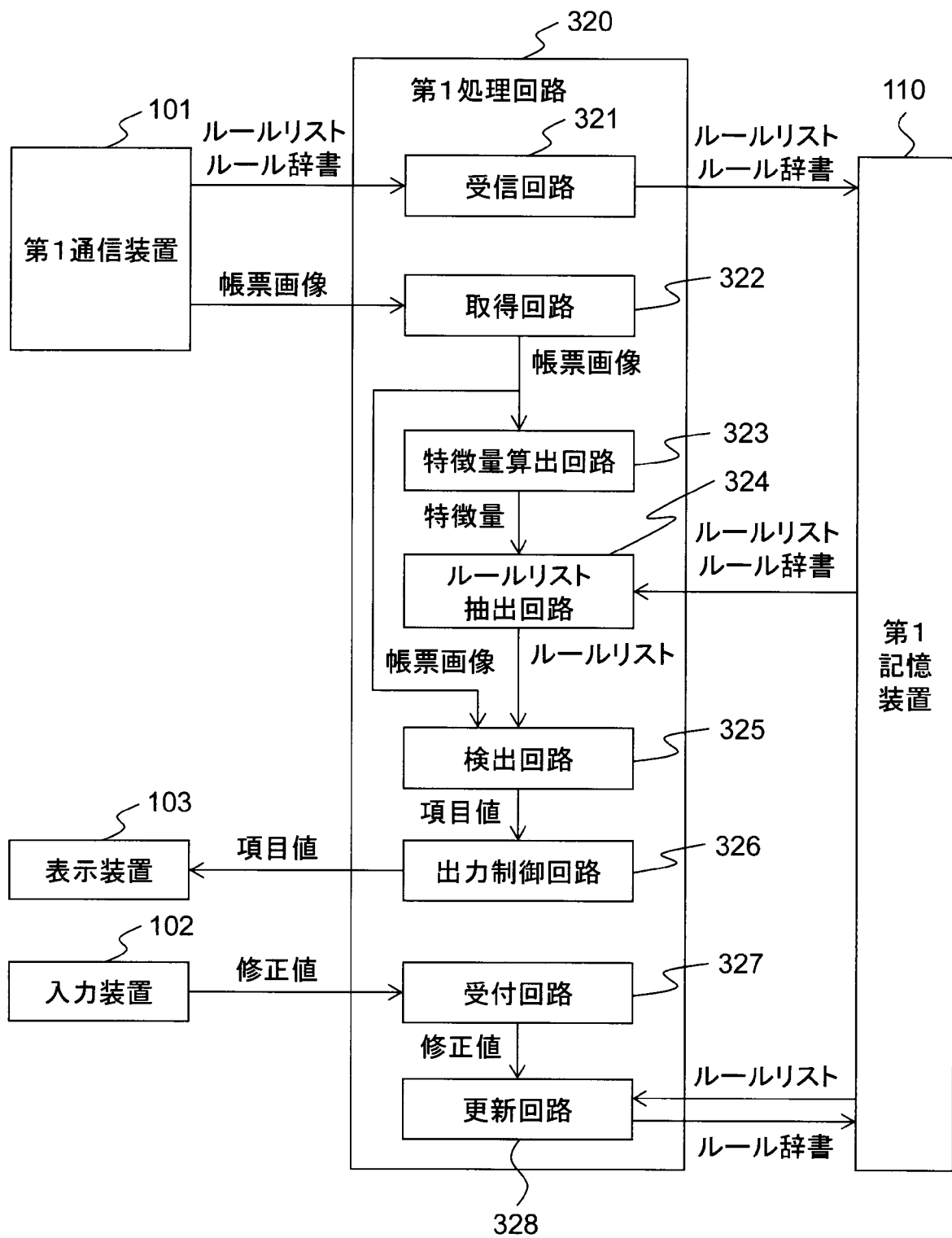
認識結果			
1501 ↓ 項目	1502 ↓ 認識結果	1503 ↓	1504 ↓
負担額	6500円	修正	確定
入院日	2017年5月1日	修正	確定
退院日	2017年6月1日	修正	確定

[図16]



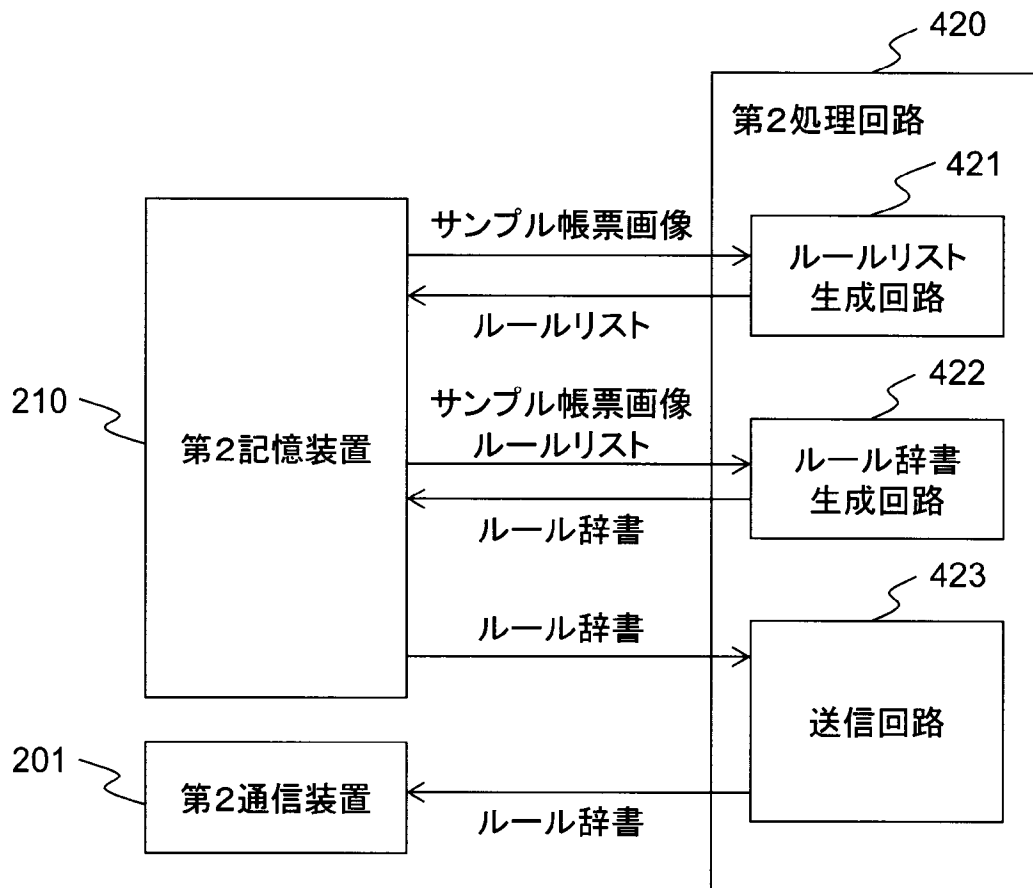
[図17]

図17



[図18]

図18



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2017/027758

A. CLASSIFICATION OF SUBJECT MATTER

G06K9/00(2006.01) i

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)

G06K9/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2017
Kokai Jitsuyo Shinan Koho	1971-2017	Toroku Jitsuyo Shinan Koho	1994-2017

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

JSTPlus (JDreamIII)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2006-134106 A (Hammock Corp.), 25 May 2006 (25.05.2006), entire text; all drawings (Family: none)	1-11
A	JP 2016-051339 A (Hitachi-Omron Terminal Solutions, Corp.), 11 April 2016 (11.04.2016), entire text; all drawings (Family: none)	1-11
A	JP 11-296676 A (Oki Electric Industry Co., Ltd.), 29 October 1999 (29.10.1999), entire text; all drawings (Family: none)	1-11

☒ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
17 August 2017 (17.08.17)

Date of mailing of the international search report
29 August 2017 (29.08.17)

Name and mailing address of the ISA/
Japan Patent Office
3-4-3, Kasumigaseki, Chiyoda-ku,
Tokyo 100-8915, Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2017/027758

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Jun'ichi HIRAYAMA et al., "Kasetu Kenshogata Approach o Mochiita Teigi Less Hiteikei Chohyo Ninshiki Gijutsu", The Transactions of the Institute of Electronics, Information and Communication Engineers D, 01 December 2014 (01.12.2014), vol.J97-D, no.12, pages 1797 to 1808, particularly, '3.6 Mojiretsu Ninshiki Yusendo Settei ni yoru Shori no Kosokuka', table 5	1-11

A. 発明の属する分野の分類（国際特許分類（IPC））

Int.Cl. G06K9/00(2006.01)i

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（IPC））

Int.Cl. G06K9/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2017年
日本国実用新案登録公報	1996-2017年
日本国登録実用新案公報	1994-2017年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

JSTPlus (JDreamIII)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	JP 2006-134106 A（株式会社 ハンモック）2006.05.25, 全文、全図（ファミリーなし）	1-11
A	JP 2016-051339 A（日立オムロンターミナルソリューションズ株式会社）2016.04.11, 全文、全図（ファミリーなし）	1-11
A	JP 11-296676 A（沖電気工業株式会社）1999.10.29, 全文、全図（ファミリーなし）	1-11

☒ C欄の続きにも文献が列挙されている。

☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの
「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
「O」口頭による開示、使用、展示等に言及する文献
「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「&」同一パテントファミリー文献

国際調査を完了した日

17.08.2017

国際調査報告の発送日

29.08.2017

国際調査機関の名称及びあて先

日本国特許庁（ISA/J P）

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官（権限のある職員）

佐田 宏史

5H

4189

電話番号 03-3581-1101 内線 3531

C（続き）． 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	平山 淳一、外 3 名，“仮説検証型アプローチを用いた定義レス非定型帳票認識技術”，電子情報通信学会論文誌D，2014. 12. 01，Vol. J97-D，No. 12，pp. 1797-1808，特に「3. 6 文字列認識優先度設定による処理の高速化」、表 5	1-11