

明 細 書

発明の名称： 記入領域抽出装置および記入領域抽出プログラム

技術分野

[0001] 本発明は、記入領域抽出装置および記入領域抽出プログラムに関するものである。

背景技術

[0002] 特許文献1には、複数の定型文書の各々を複数の小領域に分割し、小領域ごとに、RGB成分の平均濃度および平均色の特徴量を抽出し、特徴量のバラツキ量を基に、記入欄の領域を抽出する技術が記載されている。

[0003] 特許文献2には、ニューラルネットワークで生成された中間画像を加工し、損失関数の変化量を基に、対象物の特徴を示す領域の画像を抽出して合成する技術が記載されている。

[0004] 特許文献3には、低解像度画像に対して局所適応的閾値処理と膨張処理とを行ってカラー成分の局所適応的二値化を行うことにより、文字領域を抽出する技術が記載されている。

先行技術文献

特許文献

[0005] 特許文献1：特開2010-061523号公報

特許文献2：特開2017-059090号公報

特許文献3：特開2007-251518号公報

発明の概要

発明が解決しようとする課題

[0006] 特許文献1に記載の技術は、ズレおよびノイズの影響を受けやすく、抽出精度がパラメータに大きく左右される。小領域の決め方およびバラツキ量の閾値設定が煩雑であるため、特許文献1に記載の技術を適用しても、実際には十分な精度が得られないおそれがある。

[0007] 本発明は、書類サンプルから記入欄の領域を精度良く抽出することを目的

とする。

課題を解決するための手段

[0008] 本発明の一態様に係る記入領域抽出装置は、

書類の種類によって決められた記入欄を含み各記入欄が書類ごとに個別に記入されている複数の書類の画像から書類の種類に応じた特徴を学習して得られる学習モデルを用いて、少なくとも1つの書類サンプルの画像から、前記書類サンプルの種類に応じた特徴を示している領域である特徴領域を抽出する特徴領域抽出部と、

前記書類サンプルの画像における、前記特徴領域抽出部により抽出された特徴領域を除いた残りの領域から、記入欄の領域である記入領域を抽出する記入領域抽出部と

を備える。

[0009] 前記特徴領域抽出部は、前記書類サンプルの画像の画素ごとに前記書類サンプルの種類に応じた特徴の顕著性を数値化し、当該顕著性の数値を閾値によって二値化して前記書類サンプルの画像を前記特徴領域と前記残りの領域とに区分する。

[0010] 前記書類サンプルの画像から、前記書類サンプルにおいてオブジェクトが記載された領域である1つ以上のオブジェクト領域を前記記入領域の候補として抽出するオブジェクト領域抽出部をさらに備え、

前記記入領域抽出部は、前記オブジェクト領域抽出部により抽出されたオブジェクト領域のうち、前記特徴領域に重なるオブジェクト領域を前記記入領域の候補から除外する。

[0011] 前記記入領域抽出部は、前記オブジェクト領域抽出部により抽出されたオブジェクト領域のうち、互いに重なる2つ以上のオブジェクト領域を合成する。

[0012] 前記記入領域抽出部は、前記オブジェクト領域抽出部により抽出されたオブジェクト領域のうち、互いの距離が閾値以下の2つ以上のオブジェクト領

域を合成する。

[0013] 前記記入領域抽出部は、前記オブジェクト領域抽出部により抽出されたオブジェクト領域のうち、面積が閾値以下のオブジェクト領域を前記記入領域の候補から除外する。

[0014] 前記オブジェクト領域抽出部は、少なくとも文字および印を前記オブジェクトとして認識する。

[0015] 前記特徴領域抽出部により複数の書類サンプルの画像から前記特徴領域が抽出され、前記記入領域抽出部により異なる書類サンプルの画像における前記特徴領域を除いた残りの領域から抽出された2つ以上の記入領域が互いに重なる場合、当該2つ以上の記入領域を合成する記入領域合成部をさらに備える。

[0016] 本発明の一態様に係る記入領域抽出プログラムは、
コンピュータに、

書類の種類によって決められた記入欄を含み各記入欄が書類ごとに個別に記入されている複数の書類の画像から書類の種類に応じた特徴を学習して得られる学習モデルを用いて、少なくとも1つの書類サンプルの画像から、前記書類サンプルの種類に応じた特徴を示している領域である特徴領域を抽出する特徴領域抽出処理と、

前記書類サンプルの画像における、前記特徴領域抽出処理により抽出された特徴領域を除いた残りの領域から、記入欄の領域である記入領域を抽出する記入領域抽出処理と

を実行させる。

発明の効果

[0017] 本発明では、書類サンプルの画像を細かく分割することなく、その画像における、学習モデルを用いて抽出された特徴領域を除いた残りの領域から、記入欄の領域を抽出するため、ズレおよびノイズの影響を抑えることができる。したがって、本発明によれば、書類サンプルから記入欄の領域を精度良く抽出することができる。

図面の簡単な説明

- [0018] [図1]実施の形態1に係る記入領域抽出装置の構成を示すブロック図。
- [図2]実施の形態1に係る記入領域抽出装置の動作を示すフローチャート。
- [図3]実施の形態1に係る記入領域抽出装置が行う処理の流れを示す図。
- [図4]実施の形態1に係る記入領域抽出装置が行う二値化処理の例を示す図。
- [図5]実施の形態1に係る記入領域抽出装置が行う特徴領域抽出処理の例を示す図。
- [図6]実施の形態1に係る記入領域抽出装置が行うオブジェクト領域抽出処理の例を示す図。
- [図7]実施の形態1に係る記入領域抽出装置が行う記入領域抽出処理の例を示す図。
- [図8]実施の形態1に係る記入領域抽出装置が行う記入領域マージ処理の例を示す図。
- [図9]実施の形態1に係る記入領域抽出装置が行う記入領域合成処理の例を示す図。

発明を実施するための形態

- [0019] 以下、本発明の実施の形態について、図を用いて説明する。各図中、同一または相当する部分には、同一符号を付している。実施の形態の説明において、同一または相当する部分については、説明を適宜省略または簡略化する。なお、本発明は、以下に説明する実施の形態に限定されるものではなく、必要に応じて種々の変更が可能である。例えば、以下に説明する実施の形態は、部分的に実施されても構わない。

- [0020] 実施の形態1.

本実施の形態について、図1から図9を用いて説明する。

- [0021] ***構成の説明***

図1を参照して、本実施の形態に係る記入領域抽出装置10の構成を説明する。

- [0022] 記入領域抽出装置10は、コンピュータである。記入領域抽出装置10は

、プロセッサ１１を備えるとともに、メモリ１２等の他のハードウェアを備える。プロセッサ１１は、信号線を介して他のハードウェアと接続され、これら他のハードウェアを制御する。

[0023] 記入領域抽出装置１０は、機能要素として、入力部２１と、学習部２２と、特徴領域抽出部２３と、オブジェクト領域抽出部２４と、記入領域抽出部２５と、記入領域合成部２６と、出力部２７とを備える。入力部２１、学習部２２、特徴領域抽出部２３、オブジェクト領域抽出部２４、記入領域抽出部２５、記入領域合成部２６および出力部２７の機能は、ソフトウェアにより実現される。

[0024] プロセッサ１１は、記入領域抽出プログラムを実行する装置である。記入領域抽出プログラムは、入力部２１、学習部２２、特徴領域抽出部２３、オブジェクト領域抽出部２４、記入領域抽出部２５、記入領域合成部２６および出力部２７の機能を実現するプログラムである。プロセッサ１１は、例えば、CPU、GPU、DSP、または、これらのうちいくつか、もしくは、すべての組み合わせである。「CPU」は、Central Processing Unitの略語である。「GPU」は、Graphics Processing Unitの略語である。「DSP」は、Digital Signal Processorの略語である。

[0025] メモリ１２は、記入領域抽出プログラムを記憶する装置である。メモリ１２は、例えば、RAM、フラッシュメモリまたはこれらの組み合わせである。「RAM」は、Random Access Memoryの略語である。

[0026] 記入領域抽出装置１０には、外部のハードウェアとして、ディスプレイ５０が接続される。

[0027] ディスプレイ５０は、記入領域抽出プログラムから出力されるデータを画面に表示する機器である。ディスプレイ５０は、例えば、LCDである。「LCD」は、Liquid Crystal Displayの略語である。なお、ディスプレイ５０

は、内部のハードウェアとして、記入領域抽出装置 10 に備えられていてもよい。

[0028] 記入領域抽出装置 10 には、内部の他のハードウェアとして、通信デバイス、入力機器またはこれらの両方が備えられていてもよい。

[0029] 通信デバイスは、記入領域抽出プログラムに入力されるデータを受信するレシーバと、記入領域抽出プログラムから出力されるデータを送信するトランスミッタとを含む。通信デバイスは、例えば、通信チップまたは NIC である。「NIC」は、Network Interface Card の略語である。

[0030] 入力機器は、記入領域抽出プログラムへのデータの输入のためにユーザにより操作される機器である。入力機器は、例えば、マウス、キーボード、タッチパネル、スキャナ、または、これらのうちいくつか、もしくは、すべての組み合わせである。

[0031] 記入領域抽出装置 10 には、通信デバイス、入力機器またはその他のインタフェースを介して、複数の書類の画像 30 が入力される。それぞれの書類には、書類の種類によって決められた記入欄が含まれている。各記入欄は、書類ごとに個別に記入されている。例えば、書類の種類として、口座開設申込書があり、いずれの口座開設申込書にも、共通して氏名の記入欄が含まれるが、個々の口座開設申込書の記入者が違うため、記入欄に記入される氏名は口座開設申込書によって異なる。

[0032] 記入領域抽出装置 10 には、通信デバイス、入力機器またはその他のインタフェースを介して、少なくとも 1 つの書類サンプルの画像 40 も入力される。本実施の形態では、記入領域抽出装置 10 には、複数の書類サンプルの画像 40 が入力される。

[0033] メモリ 12 には、学習モデル 31 が記憶されるとともに、記入領域データベース 41 が構築される。学習モデル 31 は、複数の書類の画像 30 から書類の種類に応じた特徴を学習して得られるモデルである。記入領域データベース 41 は、複数の書類サンプルそれぞれの画像 40 から抽出される記入欄

の領域を画像データまたはその他の形式で記録するデータベースである。

- [0034] 記入領域抽出プログラムは、メモリ 12 からプロセッサ 11 に読み込まれ、プロセッサ 11 によって実行される。メモリ 12 には、記入領域抽出プログラムだけでなく、OS も記憶されている。「OS」は、Operating System の略語である。プロセッサ 11 は、OS を実行しながら、記入領域抽出プログラムを実行する。なお、記入領域抽出プログラムの一部または全部が OS に組み込まれていてもよい。
- [0035] 記入領域抽出プログラムおよび OS は、補助記憶装置に記憶されていてもよい。補助記憶装置は、例えば、HDD、フラッシュメモリまたはこれらの組み合わせである。「HDD」は、Hard Disk Drive の略語である。記入領域抽出プログラムおよび OS は、補助記憶装置に記憶されている場合、メモリ 12 にロードされ、プロセッサ 11 によって実行される。
- [0036] 記入領域抽出装置 10 は、プロセッサ 11 を代替する複数のプロセッサを備えていてもよい。これら複数のプロセッサは、記入領域抽出プログラムの実行を分担する。それぞれのプロセッサは、例えば、CPU、GPU、DSP、または、これらのうちいくつか、もしくは、すべての組み合わせである。
- [0037] 記入領域抽出プログラムにより利用、処理または出力されるデータ、情報、信号値および変数値は、メモリ 12、補助記憶装置、または、プロセッサ 11 内のレジスタまたはキャッシュメモリに記憶される。
- [0038] 記入領域抽出プログラムは、入力部 21、学習部 22、特徴領域抽出部 23、オブジェクト領域抽出部 24、記入領域抽出部 25、記入領域合成部 26 および出力部 27 により行われる処理をそれぞれ入力処理、学習処理、特徴領域抽出処理、オブジェクト領域抽出処理、記入領域抽出処理、記入領域合成処理および出力処理としてコンピュータに実行させるプログラムである。記入領域抽出プログラムは、コンピュータ読取可能な媒体に記録されて提供されてもよいし、記録媒体に格納されて提供されてもよいし、プログラムプロダクトとして提供されてもよい。

[0039] 記入領域抽出装置 10 は、1 台のコンピュータで構成されていてもよいし、複数台のコンピュータで構成されていてもよい。記入領域抽出装置 10 が複数台のコンピュータで構成されている場合は、入力部 21、学習部 22、特徴領域抽出部 23、オブジェクト領域抽出部 24、記入領域抽出部 25、記入領域合成部 26 および出力部 27 の機能が、各コンピュータに分散されて実現されてもよい。

[0040] ***動作の説明***

図 2 を参照して、本実施の形態に係る記入領域抽出装置 10 の動作を説明する。記入領域抽出装置 10 の動作は、本実施の形態に係る記入領域抽出方法に相当する。

[0041] ステップ S101 において、複数の書類の画像 30 が入力部 21 に入力される。具体的には、入力部 21 が、複数の書類をスキャンする操作をユーザより受け付ける。

[0042] ステップ S102 において、学習部 22 は、入力部 21 に入力された複数の書類の画像 30 から、書類の種類に応じた特徴を学習して学習モデル 31 を得る。具体的には、学習部 22 は、書類を分類できるように CNN を用いて画像 30 の特徴を学習する。すなわち、学習部 22 は、ステップ S101 におけるスキャンの結果である複数枚の画像 30 から、ニューラルネットワークを用いて分類問題を学習する。「CNN」は、Convolutional Neural Network の略語である。

[0043] ステップ S103 において、学習部 22 は、得られた学習モデル 31 をメモリ 12 に書き込む。具体的には、学習部 22 は、ステップ S102 における学習の結果である学習モデル 31 をメモリ 12 に構築されたデータベースに保存する。

[0044] ステップ S104 において、入力部 21 は、学習が十分であるかどうかをユーザに問い合わせる。学習が十分でないとユーザにより判断された場合、ステップ S105 において、書類の数が増やされ、ステップ S101 の処理が再び行われる。学習が十分であるとユーザにより判断された場合、ステッ

プS 1 0 6において、1つの書類サンプルの画像4 0が入力部2 1に入力される。具体的には、入力部2 1が、1つの書類をスキャンする操作をユーザより受け付ける。

[0045] ステップS 1 0 7において、特徴領域抽出部2 3は、図3から図5に示すように、学習モデル3 1を用いて、入力部2 1に入力された書類サンプルの画像4 0から、書類サンプルの種類に応じた特徴を示している領域である特徴領域4 3を抽出する。具体的には、特徴領域抽出部2 3は、書類サンプルの画像4 0の画素ごとに書類サンプルの種類に応じた特徴の顕著性を数値化してマップ4 2を生成する。そして、特徴領域抽出部2 3は、生成したマップ4 2を二値化して特徴領域4 3と残りの領域とに区分する。すなわち、特徴領域抽出部2 3は、計算した顕著性の数値を閾値によって二値化して書類サンプルの画像4 0を特徴領域4 3と残りの領域とに区分する。

[0046] より具体的には、特徴領域抽出部2 3は、図3および図4に示すように、ステップS 1 0 6におけるスキャンの結果である1枚の画像4 0における、分類対象の書類の種類に対応した発火位置をCAMまたはGrad-CAM等の手法によって求めることで、マップ4 2としてヒートマップを生成する。そして、特徴領域抽出部2 3は、発火位置を閾値で分割してヒートマップを二値化する。すなわち、特徴領域抽出部2 3は、ヒートマップを白黒画像に変換する。「CAM」は、Class Activation Mappingの略語である。「Grad-CAM」は、Gradient-weighted Class Activation Mappingの略語である。二値化のための閾値は、固定値でもよいが、GUIによりユーザが調整できるようにしてもよい。「GUI」は、Graphical User Interfaceの略語である。

[0047] 図5の例では、書類サンプルの画像4 0の中で、書類の種類を表す「口座開設申込書」という文字列と、口座開設申込書に共通する記入欄の見出しである「氏名」、「生年月日」、「住所」、「電話番号」および「Eメール」という文字列とが記載された領域が1つの特徴領域4 3として抽出されている。

る。記入欄そのもの、すなわち、口座開設申込書ごとに異なる氏名、生年月日、住所、電話番号およびEメールが記入された領域は、特徴領域43ではないと判定されている。

[0048] ステップS108において、オブジェクト領域抽出部24は、図3および図6に示すように、入力部21に入力された書類サンプルの画像40から、書類サンプルにおいてオブジェクトが記載された領域である1つ以上のオブジェクト領域45を抽出する。本実施の形態では、オブジェクト領域抽出部24は、少なくとも文字および印をオブジェクトとして認識する。なお、オブジェクト領域抽出部24は、矩形の枠等の図形をオブジェクトとして認識してもよい。

[0049] 具体的には、オブジェクト領域抽出部24は、図3に示すように、ステップS106におけるスキャンの結果である1枚の画像40を白黒反転して反転画像44に変換した後、反転画像44の白色箇所を膨張させたり、収縮させたりして、1つの文字または1つの印に相当するオブジェクト同士を一塊として認識する。これにより、文字または印の分裂を防ぐことができる。その後、オブジェクト領域抽出部24は、オブジェクトの輪郭を抽出し、抽出した輪郭を囲む矩形をオブジェクト領域45として抽出する。オブジェクト領域45は、次ステップで切り出される領域の候補となる。なお、オブジェクト領域抽出部24は、画像40を白黒反転せずに、画像40の黒色箇所を膨張させたり、収縮させたりして、1つの文字または1つの印に相当するオブジェクト同士を一塊として認識してもよい。

[0050] 図6の例では、書類の種類を表す「口座開設申込書」という文字列の各文字と、口座開設申込書に共通する記入欄の見出しである「氏名」、「生年月日」、「住所」、「電話番号」および「Eメール」という文字列の各文字とのいずれもがオブジェクトに相当する。記入欄に記入された氏名、生年月日、住所、電話番号およびEメールそのものの文字列の各文字もオブジェクトに相当する。書類サンプルの画像40の中で、これらのオブジェクトが記載された領域がそれぞれオブジェクト領域45として抽出されている。

- [0051] なお、ステップS108の処理は、OCRと併用されるか、あるいは、OCRに置き換えられてもよい。「OCR」は、Optical Character Recognitionの略語である。
- [0052] ステップS109において、記入領域抽出部25は、図3および図7に示すように、書類サンプルの画像40における、特徴領域抽出部23により抽出された特徴領域43を除いた残りの領域から、記入欄の領域である記入領域46を抽出する。具体的には、記入領域抽出部25は、オブジェクト領域抽出部24により記入領域46の候補として抽出されたオブジェクト領域45のうち、特徴領域43に重なるオブジェクト領域45を記入領域46の候補から除外し、残りの候補を記入領域46として抽出する。より具体的には、記入領域抽出部25は、ステップS108で抽出された領域の候補から、ステップS107で求められた発火位置に重なる矩形を除外し、残りの矩形を記入領域46として切り出す。
- [0053] 本実施の形態では、記入領域抽出部25は、オブジェクト領域抽出部24により抽出されたオブジェクト領域45のうち、互いに重なる2つ以上のオブジェクト領域45を合成する。具体的には、記入領域抽出部25は、図8に示すように、ある矩形A101と別の矩形A102とが一部重複している場合、これらを1つの矩形A103としてマージする。また、記入領域抽出部25は、図8に示すように、ある矩形A111が別の矩形A112に内包されている場合、これらを1つの矩形A113としてマージする。
- [0054] また、本実施の形態では、記入領域抽出部25は、オブジェクト領域抽出部24により抽出されたオブジェクト領域45のうち、互いの距離が閾値以下の2つ以上のオブジェクト領域45を合成する。具体的には、記入領域抽出部25は、図8に示すように、ある矩形A121と別の矩形A122とが隣接している場合、これらを1つの矩形A123としてマージする。距離の閾値は、0に設定されてもよいし、0よりも大きい値に設定されてもよい。
- [0055] また、本実施の形態では、記入領域抽出部25は、オブジェクト領域抽出

部24により抽出されたオブジェクト領域45のうち、面積が閾値以下のオブジェクト領域45を記入領域46の候補から除外する。具体的には、記入領域抽出部25は、一定の面積以下の矩形を削除する。

[0056] 記入領域抽出部25によるマージの条件となる距離の閾値、あるいは、削除の条件となる面積の閾値は、GUIによりユーザが調整できるようにしてもよい。マージの条件に面積の閾値が加えられてもよい。

[0057] 図7の例では、「口座開設申込書」、「氏名」、「生年月日」、「住所」、「電話番号」および「Eメール」という文字列の各文字のオブジェクト領域45が特徴領域43に重なっているため、記入領域46の候補から除外されている。そして、記入欄に記入された氏名、生年月日、住所、電話番号およびEメールそのものの文字列の各文字のオブジェクト領域45が文字列ごとに合成された上で記入領域46として抽出されている。すなわち、5個の文字列の記入範囲がそれぞれ記入領域46として抽出されている。なお、氏名の文字列については、第1段階において、姓の各文字のオブジェクト領域45がマージされて姓に対応する1つの矩形が得られ、名の各文字のオブジェクト領域45がマージされて名に対応する1つの矩形が得られる。そして、第2段階において、姓に対応する矩形と名に対応する矩形とが互いに近い距離にあることから、これら2つの矩形がマージされて1つの矩形である記入領域46が得られる。この例では、姓および名に対してまとめて1つの記入領域46が抽出されているが、姓および名に対して別々の記入領域46が得られるように、前述した距離の閾値を小さめに設定してもよい。

[0058] ステップS110において、入力部21は、すべての書類サンプルの画像40について処理が完了したかどうかをユーザに問い合わせる。処理が完了していないとユーザにより判断された場合、ステップS106の処理が再び行われる。処理が完了したとユーザにより判断された場合、ステップS111の処理が行われる。

[0059] ステップS111において、記入領域合成部26は、図3に示すように、記入領域抽出部25により異なる書類サンプルの画像40における特徴領域

4 3を除いた残りの領域から抽出された2つ以上の記入領域4 6が互いに重なる場合、それら2つ以上の記入領域4 6を合成する。具体的には、記入領域合成部2 6は、ある書類サンプルの1つの矩形と別の書類サンプルの1つの矩形とが重複している場合、それら2つの矩形を1つの矩形としてマージする。これにより、例えば、2つの口座開設申込書のうち、一方の口座開設申込

書では住所が住所欄内で左寄りに記入されていて、他方の口座開設申込書では住所が住所欄内で右寄りに記入されていたとすると、住所欄全体を1つの記入領域4 6として認識できる。

[0060] このように、本実施の形態では、記入領域合成部2 6は、複数枚の書類間で重複する矩形をマージする。このマージの条件として、矩形同士の重複部分の面積が設定されてもよい。すなわち、複数枚の書類間で重複する矩形をマージするかどうかは重複部分の面積によって決められてもよい。さらに、記入領域合成部2 6は、複数枚の書類間で位置が近い矩形をマージしてもよい。このマージの条件として、矩形間の距離が設定されてもよい。すなわち、複数枚の書類間で位置が近い矩形をマージするかどうかは距離によって決められてもよい。これらのマージの条件となる距離および面積の閾値は、GUIによりユーザが調整できるようにしてもよい。その場合、ユーザは、距離および面積の条件を緩めながら、矩形を一定の個数に絞り込むことができる。あるいは、距離および面積の閾値の代わりに、矩形の個数をいくつに絞り込むかをユーザが指定できるようにしてもよい。その場合、記入領域合成部2 6は、距離および面積の条件を緩めながら、矩形の個数を指定された個数に自動的に絞り込む。具体的には、記入領域合成部2 6は、図9に示すように、距離の閾値を調整することで、ある書類サンプルの矩形A 1 3 1と別の書類サンプルの矩形A 1 3 2とを1つの矩形A 1 3 3としてマージするかどうかを選択することができる。

[0061] ステップS 1 1 2において、記入領域合成部2 6は、ステップS 1 1 0で抽出され、ステップS 1 1 1で適宜合成された記入領域4 6を記入領域デー

データベース 4 1 に登録する。記入領域 4 6 は、任意の形式で記録されてよいが、本実施の形態では矩形の 1 頂点の座標と縦横の長さによって記録される。

[0062] ステップ S 1 1 3 において、出力部 2 7 は、記入領域データベース 4 1 に登録されたデータをディスプレイ 5 0 の画面に出力する。

[0063] 以上説明したように、本実施の形態において、特徴領域抽出部 2 3 は、ニューラルネットワークによって分類された結果を基に、画像 4 0 の特徴領域 4 3 を抽出する。オブジェクト領域抽出部 2 4 は、画像 4 0 の文字列または印鑑のマーク等のオブジェクト領域 4 5 を特定する。記入領域抽出部 2 5 は、特徴領域 4 3 とオブジェクト領域 4 5 とから、記入領域 4 6 を抽出する。

[0064] オブジェクト領域抽出部 2 4 は、1 枚の書類のオブジェクト同士を膨張処理により一塊とし、オブジェクトの輪郭を矩形で囲むことによって 1 枚の書類のオブジェクト領域 4 5 を抽出する。

[0065] 特徴領域抽出部 2 3 は、ニューラルネットワークの分類結果と、ニューラルネットワークの隠れ層の出力結果とを基に、書類の特徴領域 4 3 を抽出する。

[0066] 記入領域合成部 2 6 は、記入領域抽出部 2 5 から出力された 1 枚の書類に対する記入領域 4 6 を、複数枚重ね合わせて重複する矩形を合成する。

[0067] ***実施の形態の効果の説明***

本実施の形態では、書類サンプルの画像 4 0 を細かく分割することなく、その画像 4 0 における、学習モデル 3 1 を用いて抽出された特徴領域 4 3 を除いた残りの領域から、記入欄の領域を抽出するため、ズレおよびノイズの影響を抑えることができる。したがって、本実施の形態によれば、書類サンプルから記入欄の領域を精度良く抽出することができる。

[0068] 従来は、書類から人が記入した情報を取得する場合、人手で取得位置を指定したり、書類のテンプレートを定義したりする必要があるが、本実施の形態によれば、自動的に人手によって入力された文字または印鑑等の領域を特定することができる。そのために、本実施の形態では、ニューラルネットワ

ークによる書類判別結果を基に、判別の大きな要因となった画素群である特徴領域43を判別結果と中間層のニューロンとから特定し、特徴領域43を除いたオブジェクト領域45を記入領域46とする。すなわち、本実施の形態では、特徴領域43をニューラルネットワークの分類結果から得る。そして、特徴領域43以外の場所を、書類のフォーマット以外の領域とすることで、記入領域46を抽出する。

[0069] 本実施の形態によれば、テンプレートまたは書類の特徴を定義することなく、複数枚の書類サンプルから自動で記入領域46を抽出できる。

[0070] 本実施の形態によれば、書類の分類結果を用いて記入領域46を抽出するため、計算リソースを節約できる。すなわち、記入領域46の抽出のために、新たにCNNを構築したり、記入領域46の抽出用のルールを作成したりする必要がない。

[0071] ***他の構成***

本実施の形態では、入力部21、学習部22、特徴領域抽出部23、オブジェクト領域抽出部24、記入領域抽出部25、記入領域合成部26および出力部27の機能がソフトウェアにより実現されるが、変形例として、入力部21、学習部22、特徴領域抽出部23、オブジェクト領域抽出部24、記入領域抽出部25、記入領域合成部26および出力部27の機能がソフトウェアとハードウェアとの組み合わせにより実現されてもよい。すなわち、入力部21、学習部22、特徴領域抽出部23、オブジェクト領域抽出部24、記入領域抽出部25、記入領域合成部26および出力部27の機能の一部が専用のハードウェアにより実現され、残りがソフトウェアにより実現されてもよい。

[0072] 専用のハードウェアは、例えば、単回路、複合回路、プログラム化したプロセッサ、並列プログラム化したプロセッサ、ロジックIC、GA、FPGA、ASIC、または、これらのうちいくつか、もしくは、すべての組み合わせである。「IC」は、Integrated Circuitの略語である。「GA」は、Gate Arrayの略語である。「FPGA」は

、Field-Programmable Gate Arrayの略語である。「ASIC」は、Application Specific Integrated Circuitの略語である。

- [0073] プロセッサ11および専用のハードウェアは、いずれも処理回路である。すなわち、入力部21、学習部22、特徴領域抽出部23、オブジェクト領域抽出部24、記入領域抽出部25、記入領域合成部26および出力部27の機能がソフトウェアにより実現されるか、ソフトウェアとハードウェアとの組み合わせにより実現されるかに関わらず、入力部21、学習部22、特徴領域抽出部23、オブジェクト領域抽出部24、記入領域抽出部25、記入領域合成部26および出力部27の動作は、処理回路により行われる。

符号の説明

- [0074] 10 記入領域抽出装置、11 プロセッサ、12 メモリ、21 入力部、22 学習部、23 特徴領域抽出部、24 オブジェクト領域抽出部、25 記入領域抽出部、26 記入領域合成部、27 出力部、30 画像、31 学習モデル、40 画像、41 記入領域データベース、42 マップ、43 特徴領域、44 反転画像、45 オブジェクト領域、46 記入領域、50 ディスプレイ。

請求の範囲

- [請求項1] 書類の種類によって決められた記入欄を含み各記入欄が書類ごとに個別に記入されている複数の書類の画像から書類の種類に応じた特徴を学習して得られる学習モデルを用いて、少なくとも1つの書類サンプルの画像から、前記書類サンプルの種類に応じた特徴を示している領域である特徴領域を抽出する特徴領域抽出部と、
- 前記書類サンプルの画像における、前記特徴領域抽出部により抽出された特徴領域を除いた残りの領域から、記入欄の領域である記入領域を抽出する記入領域抽出部と
- を備える記入領域抽出装置。
- [請求項2] 前記特徴領域抽出部は、前記書類サンプルの画像の画素ごとに前記書類サンプルの種類に応じた特徴の顕著性を数値化し、当該顕著性の数値を閾値によって二値化して前記書類サンプルの画像を前記特徴領域と前記残りの領域とに区分する請求項1に記載の記入領域抽出装置。
- [請求項3] 前記書類サンプルの画像から、前記書類サンプルにおいてオブジェクトが記載された領域である1つ以上のオブジェクト領域を前記記入領域の候補として抽出するオブジェクト領域抽出部をさらに備え、
- 前記記入領域抽出部は、前記オブジェクト領域抽出部により抽出されたオブジェクト領域のうち、前記特徴領域に重なるオブジェクト領域を前記記入領域の候補から除外する請求項1または2に記載の記入領域抽出装置。
- [請求項4] 前記記入領域抽出部は、前記オブジェクト領域抽出部により抽出されたオブジェクト領域のうち、互いに重なる2つ以上のオブジェクト領域を合成する請求項3に記載の記入領域抽出装置。
- [請求項5] 前記記入領域抽出部は、前記オブジェクト領域抽出部により抽出されたオブジェクト領域のうち、互いの距離が閾値以下の2つ以上のオブジェクト領域を合成する請求項3または4に記載の記入領域抽出装置。

置。

[請求項6] 前記記入領域抽出部は、前記オブジェクト領域抽出部により抽出されたオブジェクト領域のうち、面積が閾値以下のオブジェクト領域を前記記入領域の候補から除外する請求項3から5のいずれか1項に記載の記入領域抽出装置。

[請求項7] 前記オブジェクト領域抽出部は、少なくとも文字および印を前記オブジェクトとして認識する請求項3から6のいずれか1項に記載の記入領域抽出装置。

[請求項8] 前記特徴領域抽出部により複数の書類サンプルの画像から前記特徴領域が抽出され、前記記入領域抽出部により異なる書類サンプルの画像における前記特徴領域を除いた残りの領域から抽出された2つ以上の記入領域が互いに重なる場合、当該2つ以上の記入領域を合成する記入領域合成部をさらに備える請求項1から7のいずれか1項に記載の記入領域抽出装置。

[請求項9] コンピュータに、
書類の種類によって決められた記入欄を含み各記入欄が書類ごとに個別に記入されている複数の書類の画像から書類の種類に応じた特徴を学習して得られる学習モデルを用いて、少なくとも1つの書類サンプルの画像から、前記書類サンプルの種類に応じた特徴を示している領域である特徴領域を抽出する特徴領域抽出処理と、

前記書類サンプルの画像における、前記特徴領域抽出処理により抽出された特徴領域を除いた残りの領域から、記入欄の領域である記入領域を抽出する記入領域抽出処理と
を実行させる記入領域抽出プログラム。