

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2013-61590
(P2013-61590A)

(43) 公開日 平成25年4月4日(2013.4.4)

(51) Int.Cl. F I テーマコード (参考)
G09B 19/00 (2006.01) G09B 19/00 H
G06Q 50/20 (2012.01) G06F 17/60 128

審査請求 未請求 請求項の数 15 O L (全 28 頁)

(21) 出願番号 特願2011-201451 (P2011-201451)
 (22) 出願日 平成23年9月15日 (2011.9.15)

(特許庁注：以下のものは登録商標)

1. QRコード

(71) 出願人 304025138
 国立大学法人 大阪教育大学
 大阪府柏原市旭ヶ丘4-698-1
 (74) 代理人 100115749
 弁理士 谷川 英和
 (72) 発明者 仲矢 史雄
 大阪府柏原市旭ヶ丘4-698-1 国立
 大学法人大阪教育大学内

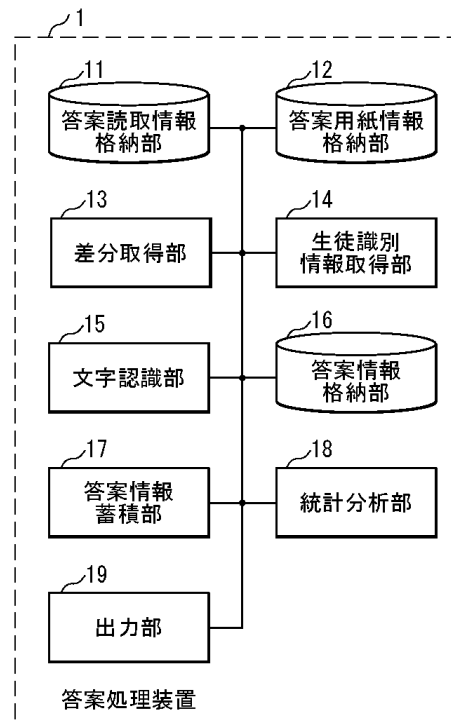
(54) 【発明の名称】 答案処理装置、答案処理方法、プログラム、およびシール

(57) 【要約】

【課題】生徒が記入した答案を適切に処理することが可能な答案処理装置を提供する。

【解決手段】生徒が答案を記入した答案用紙であって、答案の記入を行った各生徒の生徒識別情報と対応付けられた識別画像が配置された答案用紙を読み取った答案読取情報が格納される答案読取情報格納部 11 と、答案読取情報の読み取り元となる答案用紙の、答案が記入される前の状態を示す答案用紙情報が格納される答案用紙情報格納部 12 と、答案読取情報の、答案用紙情報に対する差分に関する情報である差分関連情報を取得する差分取得部 13 と、この差分を取得した答案読取情報の識別画像と対応付けられた生徒識別情報を取得する生徒識別情報取得部 14 と、差分取得部 13 が取得した差分関連情報と、生徒識別情報取得部 14 が取得した生徒識別情報とを有する答案情報を答案情報格納部 16 に蓄積する答案情報蓄積部 17 とを備えた。

【選択図】 図 1



【特許請求の範囲】

【請求項 1】

1 以上の生徒が答案を記入した答案用紙であって、答案の記入を行った各生徒の識別情報である生徒識別情報と対応付けられた画像の情報である識別画像が配置された答案用紙を読み取った情報である答案読取情報が格納される答案読取情報格納部と、前記答案読取情報の読み取り元となる答案用紙の、答案が記入される前の状態を示す情報である答案用紙情報が格納される答案用紙情報格納部と、前記答案読取情報格納部に格納されている答案読取情報の、前記答案用紙情報に対する差分に関する情報である差分関連情報を取得する差分取得部と、前記差分取得部が差分を取得した答案読取情報の識別画像を検出し、当該識別画像と対応付けられた生徒識別情報を取得する生徒識別情報取得部と、前記差分関連情報と、前記生徒識別情報とを有する答案情報が格納される答案情報格納部と、前記差分取得部が取得した差分関連情報と、前記生徒識別情報取得部が取得した生徒識別情報とを有する答案情報を前記答案情報格納部に蓄積する答案情報蓄積部とを備えた答案処理装置。

10

【請求項 2】

前記答案情報格納部に格納された答案情報に含まれる生徒識別情報を用いて、生徒ごとに当該答案情報に関する統計分析を行う統計分析部と、前記統計分析部の統計分析結果を出力する出力部とを備えた請求項 1 記載の答案処理装置。

20

【請求項 3】

前記答案情報には、当該答案情報に対応した日付を示す情報である日付情報を更に有しており、前記統計分析部は、前記答案情報が有する日付情報を用いて、予め指定された期間ごとに、前記答案情報に関する統計分析を行う請求項 2 記載の答案処理装置。

【請求項 4】

前記差分取得部が取得した差分関連情報に対して文字認識処理を行う文字認識部を更に備え、前記統計分析部は、前記答案情報に関する統計分析として、前記答案情報に含まれる差分関連情報に記載されている文字に関する統計分析を、当該答案情報に含まれる差分関連情報に対して前記文字認識部により行われた文字認識処理の結果を示す情報を用いて行う請求項 2 または請求項 3 いずれか記載の答案処理装置。

30

【請求項 5】

前記差分関連情報に記載されている文字に関する統計分析は、前記文字認識部により当該差分関連情報について認識された文字の量に関する統計分析である請求項 4 記載の答案処理装置。

【請求項 6】

前記差分関連情報に記載されている文字に関する統計分析は、前記文字認識部により当該差分関連情報について認識された文字の多様性に関する統計分析である請求項 4 記載の答案処理装置。

40

【請求項 7】

前記差分関連情報に記載されている文字に関する統計分析は、当該記載されている文字の属性に関する統計分析である請求項 4 記載の答案処理装置。

【請求項 8】

前記文字の属性に関する統計分析は、当該差分関連情報に対して前記文字認識部により行われた文字認識結果に関する統計分析である請求項 7 記載の答案処理装置。

【請求項 9】

前記文字の属性に関する統計分析は、当該差分関連情報に対して前記文字認識部により行われた文字認識結果が示す認識された文字のサイズに関する統計分析である請求項 7 記載

50

の答案処理装置。

【請求項 10】

前記差分取得部が取得した差分関連情報に対して文字認識処理を行う文字認識部を更に備え、

前記差分関連情報に記載されている文字に関する統計分析は、前記文字認識部が認識した文字が示す文章の正確度に関する統計分析である請求項 2 または請求項 3 いずれか記載の答案処理装置。

【請求項 11】

前記差分取得部が取得した差分関連情報に対して文字認識処理を行う文字認識部を更に備え、

前記差分関連情報に記載されている文字に関する統計分析は、当該記載されている文字の単位面積当たりの文字数に関する統計分析である請求項 2 または請求項 3 いずれか記載の答案処理装置。

【請求項 12】

1 以上の生徒が答案を記入した答案用紙であって、答案の記入を行った各生徒の識別情報である生徒識別情報と対応付けられた画像の情報である識別画像が配置された答案用紙を読み取った情報である答案読取情報が格納される答案読取情報格納部と、前記答案読取情報の読み取り元となる答案用紙の、答案が記入される前の状態を示す情報である答案用紙情報が格納される答案用紙情報格納部と、差分取得部と、生徒識別情報取得部と、答案情報が格納される答案情報格納部とを用いて行われる答案処理方法であって、

前記差分取得部が、前記答案読取情報格納部に格納されている答案読取情報の、前記答案用紙情報に対する差分に関する情報である差分関連情報を取得する差分取得ステップと、前記生徒識別情報取得部が、前記差分取得ステップにより差分を取得した答案読取情報の識別画像を検出し、当該識別画像と対応付けられた生徒識別情報を取得する生徒識別情報取得ステップと、

前記答案情報蓄積部が、前記差分取得ステップにより取得した差分関連情報と、前記生徒識別情報取得ステップにより取得した生徒識別情報とを有する答案情報を前記答案情報格納部に蓄積する答案情報蓄積部とを備えた答案処理方法。

【請求項 13】

1 以上の生徒が答案を記入した答案用紙であって、答案の記入を行った各生徒の識別情報である生徒識別情報と対応付けられた画像の情報である識別画像が配置された答案用紙を読み取った情報である答案読取情報が格納される答案読取情報格納部と、前記答案読取情報の読み取り元となる答案用紙の、答案が記入される前の状態を示す情報である答案用紙情報が格納される答案用紙情報格納部と、答案情報格納部と、にアクセス可能なコンピュータを、

前記答案読取情報格納部に格納されている答案読取情報の、前記答案用紙情報に対する差分に関する情報である差分関連情報を取得する差分取得部と、

前記差分取得部が差分を取得した答案読取情報の識別画像を検出し、当該識別画像と対応付けられた生徒識別情報を取得する生徒識別情報取得部と、

前記差分取得部が取得した差分関連情報と、前記生徒識別情報取得部が取得した生徒識別情報とを有する答案情報を前記答案情報格納部に蓄積する答案情報蓄積部として機能させるためのプログラム。

【請求項 14】

生徒の識別情報である生徒識別情報と対応付けられた識別画像が表面に配置されたシール。

【請求項 15】

前記識別画像の周辺に、当該識別画像に対応する生徒識別情報を示す文字列が配置された請求項 14 記載のシール。

【発明の詳細な説明】

【技術分野】

10

20

30

40

50

【0001】

本発明は、生徒が答案用紙に記載した答案を処理する装置等に関するものである。

【背景技術】

【0002】

従来、答案を処理するシステムとして、入力手段と表示手段とを備えた答案処理システムであって、(a)試験の未処理の答案が画像データ化された答案データが記憶される第1の記憶手段と、(b)処理が終了した答案が画像データ化された答案データが記憶される第2の記憶手段と、(c)前記試験で問われている問題に関する説明データが予め記憶されているデータベースと、(d)前記入力手段から入力された特定の答案の表示命令にตอบสนองして、前記第1の記憶手段から前記特定の答案に相当する答案データを読み出して前記表示手段に表示させ、表示された前記答案データにตอบสนองして前記入力手段から入力された所定の基準に基づく前記答案の処理データを表示されている前記答案データに付加すると共に、前記答案に含まれている問題に関する説明データを前記データベースから読み出して前記答案データに付加した上で、前記処理データと前記説明データとが付加された答案データを第2の記憶手段に記憶する答案処理手段と、を備えていることを特徴とする答案処理システムが知られていた(例えば、特許文献1参照)。

10

【先行技術文献】

【特許文献】

【0003】

【特許文献1】特開2006-227317号公報(第1頁、第1図等)

20

【発明の概要】

【発明が解決しようとする課題】

【0004】

しかしながら、従来においては、生徒が記入した答案を適切に処理することができないという課題があった。

【0005】

例えば、上記のように、生徒による答案を、画像データ化して管理することは可能であるが、この場合、答案を生徒と対応付けて管理するためには、答案を画像データ化する際に、生徒の識別情報(例えば、名前等)等を手動で入力しなければならず、非常に手間がかかり、容易に処理が行えなかった。

30

【0006】

また、生徒の答案を、統計処理等に使う際には、答案の画像データ中の生徒の記入内容を取り出して、いわゆるOCRで文字の認識を行って、文字の情報を取得する必要がある。しかしながら、答案用紙に元々印刷されている文字等の、生徒が記入した内容以外の文字等も、OCRで文字認識されてしまい、不必要な情報が発生してしまうという問題があった。この結果として、例えば、生徒の記入内容を利用した統計処理等が適切に行えないという問題もあった。

【課題を解決するための手段】

【0007】

本発明の答案処理装置は、1以上の生徒が答案を記入した答案用紙であって、答案の記入を行った各生徒の識別情報である生徒識別情報と対応付けられた画像の情報である識別画像が配置された答案用紙を読み取った情報である答案読取情報が格納される答案読取情報格納部と、答案読取情報の読み取り元となる答案用紙の、答案が記入される前の状態を示す情報である答案用紙情報が格納される答案用紙情報格納部と、答案読取情報格納部に格納されている答案読取情報の、答案用紙情報に対する差分に関する情報である差分関連情報を取得する差分取得部と、差分取得部が差分を取得した答案読取情報の識別画像を検出し、識別画像と対応付けられた生徒識別情報を取得する生徒識別情報取得部と、差分関連情報と、生徒識別情報とを有する答案情報が格納される答案情報格納部と、差分取得部が取得した差分関連情報と、生徒識別情報取得部が取得した生徒識別情報とを有する答案情報を答案情報格納部に蓄積する答案情報蓄積部とを備えた答案処理装置である。

40

50

【0008】

かかる構成により、答案読取情報から、生徒が記入した答案と生徒識別情報とを容易に取得して、答案情報として蓄積することができるため、生徒が記入した答案を適切に処理することができる。

【0009】

また、本発明の答案処理装置は、前記答案処理装置において、答案情報格納部に格納された答案情報に含まれる生徒識別情報を用いて、生徒ごとに答案情報に関する統計分析を行う統計分析部と、統計分析部の統計分析結果を出力する出力部とを備えた答案処理装置である。

【0010】

かかる構成により、答案情報を用いて、適切な統計分析を行うことができる。

【0011】

また、本発明の答案処理装置は、前記答案処理装置において、答案情報には、答案情報に対応した日付を示す情報である日付情報を更に有しており、統計分析部は、答案情報が有する日付情報を用いて、予め指定された期間ごとに、答案情報に関する統計分析を行う答案処理装置である。

【0012】

かかる構成により、予め指定された期間ごとの統計分析結果を容易に得ることができる。

【0013】

また、本発明の答案処理装置は、前記答案処理装置において、差分取得部が取得した差分関連情報に対して文字認識処理を行う文字認識部を更に備え、統計分析部は、答案情報に関する統計分析として、答案情報に含まれる差分関連情報に記載されている文字に関する統計分析を、答案情報に含まれる差分関連情報に対して文字認識部により行われた文字認識処理の結果を示す情報を用いて行う答案処理装置である。

【0014】

かかる構成により、生徒の答案における文字に関する統計分析結果を容易に得ることができる。

【0015】

また、本発明の答案処理装置は、前記答案処理装置において、差分関連情報に記載されている文字に関する統計分析は、文字認識部により差分関連情報について認識された文字の量に関する統計分析である答案処理装置である。

【0016】

かかる構成により、例えば、生徒の答案における文字の記入量に関する統計分析結果を容易に得ることができる。

【0017】

また、本発明の答案処理装置は、前記答案処理装置において、差分関連情報に記載されている文字に関する統計分析は、文字認識部により差分関連情報について認識された文字の多様性に関する統計分析である答案処理装置である。

【0018】

かかる構成により、例えば、各生徒の文字や文字列、単語等に関する知識等についての統計分析結果を容易に得ることができる。

【0019】

また、本発明の答案処理装置は、前記答案処理装置において、差分関連情報に記載されている文字に関する統計分析は、記載されている文字の属性に関する統計分析である答案処理装置である。

【0020】

かかる構成により、生徒の答案における文字の属性に関する統計分析結果を容易に得ることができる。

【0021】

10

20

30

40

50

また、本発明の答案処理装置は、前記答案処理装置において、文字の属性に関する統計分析は、差分関連情報に対して文字認識部により行われた文字認識結果に関する統計分析である答案処理装置である。

【0022】

かかる構成により、例えば、生徒の答案における文字のきれいさ等に関する統計分析結果を容易に得ることができる。

【0023】

また、本発明の答案処理装置は、前記答案処理装置において、文字の属性に関する統計分析は、差分関連情報に対して文字認識部により行われた文字認識結果が示す認識された文字のサイズに関する統計分析である答案処理装置である。

10

【0024】

かかる構成により、例えば、生徒の答案における文字のきれいさ等に関する統計分析結果を容易に得ることができる。

【0025】

また、本発明の答案処理装置は、前記答案処理装置において、差分取得部が取得した差分関連情報に対して文字認識処理を行う文字認識部を更に備え、差分関連情報に記載されている文字に関する統計分析は、文字認識部が認識した文字が示す文章の正確度に関する統計分析である答案処理装置である。

【0026】

かかる構成により、例えば、生徒が正確な文章が作成できるか等に関する統計分析結果を容易に得ることができる。

20

【0027】

また、本発明の答案処理装置は、前記答案処理装置において、差分取得部が取得した差分関連情報に対して文字認識処理を行う文字認識部を更に備え、差分関連情報に記載されている文字に関する統計分析は、記載されている文字の単位面積当たりの文字数に関する統計分析である答案処理装置である。

【0028】

かかる構成により、例えば、答案において生徒が文字を記入する量の度合いについての統計分析結果を容易に得ることができる。

【0029】

本発明のシールは、生徒の識別情報である生徒識別情報と対応付けられた識別画像が表面に配置されたシールである。

30

【0030】

かかる構成により、シールを貼付するだけで、答案用紙に容易に生徒識別情報を配置することができ、生徒の答案に容易に、各生徒に対応した生徒識別情報を配置することができる。

【0031】

また、本発明のシールは、前記シールにおいて、前記識別画像の周辺に、当該識別画像に対応する生徒識別情報を示す文字列が配置されたシールである。

【0032】

かかる構成により、各シールが、どの生徒の識別画像が配置されたものであるかを容易に確認することができる。

40

【発明の効果】

【0033】

本発明による答案処理装置等によれば、生徒が記入した答案を適切に処理することができる。

【図面の簡単な説明】

【0034】

【図1】本発明の実施の形態における答案処理装置のブロック図

【図2】同シールの一例を示す図を示す図

50

- 【図 3】同答案処理装置の動作について説明するフローチャート
- 【図 4】同答案処理装置の動作について説明するフローチャート
- 【図 5】同答案処理装置の答案用紙情報の一例を示す図（図 5（a））、および、答案読取情報の一例を示す図（図 5（b））
- 【図 6】同答案処理装置の答案用紙情報管理表を示す図
- 【図 7】同答案処理装置の答案読取情報管理表を示す図
- 【図 8】同答案処理装置の差分関連情報を示す図
- 【図 9】同答案処理装置により取得されたテキスト情報を示す図（図 9（a））、および、認識管理表を示す図（図 9（b））
- 【図 10】同答案処理装置の答案情報管理表を示す図
- 【図 11】同答案処理装置の年別取得情報を示す図
- 【図 12】同答案処理装置の年別分析結果管理表を示す図
- 【図 13】同答案処理装置の年別分析結果管理表を示す図
- 【図 14】同発明の実施の形態におけるコンピュータシステムの外觀の一例を示す図
- 【図 15】同コンピュータシステムの構成の一例を示す図
- 【発明を実施するための形態】

10

【0035】

以下、答案処理装置等の実施形態について図面を参照して説明する。なお、実施の形態において同じ符号を付した構成要素は同様の動作を行うので、再度の説明を省略する場合がある。

20

【0036】

（実施の形態）

図 1 は、本実施の形態における答案処理装置 1 のブロック図である。

【0037】

答案処理装置 1 は、答案読取情報格納部 1 1、答案用紙情報格納部 1 2、差分取得部 1 3、生徒識別情報取得部 1 4、文字認識部 1 5、答案情報格納部 1 6、答案情報蓄積部 1 7、統計分析部 1 8、出力部 1 9 を備える。

【0038】

答案読取情報格納部 1 1 には、1 以上の答案読取情報が格納される。答案読取情報は、1 以上の生徒が答案を記入した答案用紙であって、答案の記入を行った各生徒の生徒識別情報と対応付けられた識別画像が配置された答案用紙を読み取った情報である。答案読取情報は、通常は、答案用紙をスキャナ等で読み取った情報である。ただし、デジタルカメラ等で読み取った（撮影した）情報であっても良い。答案読取情報は、例えば、答案用紙の読み取り面の画像情報である。答案読取情報は、例えば、J P E G 形式や P N G 形式や、T I F F 形式のラスタ画像である。ただし、ラスタ画像を自動トレース等によりベクトル化した画像であってもよい。答案読取情報の色深度等は問わない。通常、答案においては、色の情報が不要な場合が多いので、白黒等の二値の画像や、グレースケールの画像であっても良い。また、カラーの画像であっても良い。色深度の変更は、答案用紙の読取後に行っても良い。また、答案読取情報の解像度は高い方が生徒の記入した文字を正確に読み取ることができる上で好ましいが、答案読取情報等のデータ量が増えるため、例えば、200 p p i ~ 600 p p i 程度の解像度が利用されることが多い。

30

40

【0039】

答案用紙は、問題に対する答案を記入する用紙である。答案用紙の材質は問わない。答案用紙には、通常、答案を記入するための記入欄や、この記入欄と問題とを対応付けるために用いられる問題番号等の問題の識別情報が配置（例えば、印刷）されている。答案を記入した答案用紙は、例えば、筆記用具等を用いて生徒によって記入された答案を示す文字が配置されている答案用紙である。また、文字以外の答案等が記入されていてもよい。ここでの生徒は、学校や、塾、予備校等の生徒である。答案用紙は、問題も含んでいても良い。答案用紙には、日付の記入欄や、答案用紙や、答案用紙に対応する問題を識別する情報が配置されていても良い。

50

【 0 0 4 0 】

生徒識別情報とは、各生徒の識別情報である。例えば、生徒識別情報は、生徒の名前や、学籍番号、住民コード、メールアドレス等である。

【 0 0 4 1 】

識別画像とは、識別情報（ここでは特に生徒識別情報）と対応付けられた画像の情報である。識別画像とは、図形によって、文字列（具体的にはコード）を表現した画像である。この文字列は、数字等であっても良い。識別画像は、例えば、バーコードである。バーコードとは、例えば、様々な太さや長さを有する複数の矩形と、その並びおよびその間隔等で文字を表現するものである。バーコードは、例えば、様々な太さの複数の線を一方に配列したものであって、その線の並び及び間隔で文字を表現する一次元のバーコードであってもよい。また、バーコードは、例えば、様々な太さや長さを有する複数の矩形を、縦方向及び横方向に配列したものであって、その矩形の並び（位置関係）およびその間隔等で文字を表現する二次元のバーコード（以下、二次元バーコードと称す）であってもよい。二次元バーコードとは、二次元コードとも呼ばれる。二次元バーコードは、例えば、QRコードやPDF417等である。本実施の形態においては、説明の便宜上、識別画像が二次元バーコードである場合を例に挙げて説明する。

10

【 0 0 4 2 】

生徒識別情報と対応付けられた識別画像とは、この識別画像（例えば二次元バーコード）から読み出される情報（文字列）が、生徒識別情報である識別画像、あるいは、読み出される情報（文字列）が、生徒識別情報と対応付けられている識別画像である。識別画像からの文字列等の読み出しは、例えば、いわゆるバーコードリーダ（図示せず）等の、識別画像を読み取って、対応する文字列に変換する装置によって行われる。例えば、二次元バーコードの一つであるQRコード等においては、このQRコードを専用のバーコードリーダを読み出す処理を行うことで、生徒名等の生徒識別情報を取得することが可能である。識別画像は答案用紙にどのように配置されていてもよく、例えば、答案用紙に、最初から印刷されていても良いし、二次元バーコードが表面に配置された（例えば、印刷された）シールが、答案用紙に貼り付けられていても良い。

20

【 0 0 4 3 】

なお、答案用紙に配置される二次元バーコードは、答案に記入を行う生徒の生徒識別情報と対応付けたものとする必要があることから、各答案用紙には、それぞれ、答案記入を行う生徒の生徒識別情報と対応付けられた二次元バーコードが配置されたシールを貼付することが好ましい。

30

【 0 0 4 4 】

図2は、生徒識別情報と対応付けられた識別画像が配置されたシールの一例を示す図である。

【 0 0 4 5 】

シール20の表面には、識別画像が配置されている。ここでは、特に識別画像として、二次元バーコードを用いた場合を例に挙げて示している。ただし、他の識別画像であっても良い。シール20の裏面には、例えば、接着剤や粘着剤等が設けられている。この識別画像21のシール20においては、図2に示すように、識別画像と対応付けられた生徒識別情報を示す文字列22（ここでは、山田A男）が、識別画像の近傍に配置されている（例えば、印刷されている）ようにすることが好ましい。ここでの文字列22は数字等も含む概念である。このようにすることで、どの生徒の生徒識別情報に対応付けられた識別画像が配置されたシール20であるかを、識別画像を読みださなくても容易に判断できる。なお、識別画像の近傍に配置される生徒識別情報は、この識別画像と対応付けられている生徒識別情報（例えば、この識別画像から読み出される生徒識別情報）と対応付けられたものであればよく、必ずしも同一のものでなくてよい。例えば、識別画像の近傍に配置される生徒識別情報が生徒の氏名で、この識別画像と対応付けられている生徒識別情報が生徒の学籍番号等であっても良い。このようなシール20を、各生徒が答案を記入した答案用紙、あるいは各生徒が答案を記入しようとする答案用紙に貼付することで、各生徒の生

40

50

徒識別情報と対応付けられた識別画像を、容易に答案用紙に配置することができる。

【 0 0 4 6 】

なお、答案用紙の、シール 2 0 の貼り付ける位置、または貼り付け可能な範囲等は、答案用紙上に枠線等を用いて指定されていても良い。

【 0 0 4 7 】

答案読取情報は、生徒が答案を記入した答案用紙の日付に関する情報と対応付けられていても良い。生徒が答案を記入した答案用紙の日付に関する情報とは、生徒が答案を記入した日付の情報であっても良いし、答案用紙の読み取りが行われた日付に関する情報であっても良い。日付に関する情報は、通常、年月日の情報である。日付に関する情報は、時刻の情報を含んでもよい。

10

【 0 0 4 8 】

答案読取情報格納部 1 1 は、不揮発性の記録媒体が好適であるが、揮発性の記録媒体でも実現可能である。かかることは、他の格納部等においても同様である。

【 0 0 4 9 】

答案用紙情報格納部 1 2 には、答案読取情報の読み取り元となる答案用紙の、答案が記入される前の状態を示す情報である答案用紙情報が格納される。答案用紙情報は、例えば、生徒が答案を記入する前の答案用紙、または、生徒が答案を記入した答案用紙と同じ未記入の答案用紙を、スキャナ等を用いて読み取った答案用紙の画像情報である。また、答案用紙を作成するための、例えば、ワードプロセッサや DTP ソフトウェア用のデータ、もしくはこれらのデータをラスタライズして得られた情報であっても良い。答案用紙情報は、例えば、答案用紙のラスタ画像を取得可能な情報であればよい。なお、答案用紙情報の解像度と、答案読取情報の解像度は、同じであることが好ましい。一の答案用紙に対応する答案読取情報と答案用紙情報とは、例えば、対応付けて管理される。例えば、各答案読取情報は、対応する答案用紙情報の識別情報と対応付けて管理される。なお、予め答案用紙に、答案用紙の識別情報を表記しておくようにすることで、この答案用紙の識別情報を読み取って得られた識別情報を、答案読取情報と答案用紙情報との両方に対応付けるようにしても良い。例えば、この場合、答案用紙に、この答案用紙の識別情報と対応付けられた識別画像（例えば、バーコード）が、答案用紙に配置しておくようにし、答案読み取り情報および答案用紙情報から、この識別画像を読み取って得られた識別情報を、答案読取情報と答案用紙情報との両方に対応付けるようにしても良い。

20

30

【 0 0 5 0 】

差分取得部 1 3 は、答案読取情報格納部 1 1 に格納されている答案読取情報の、答案用紙情報格納部 1 2 に格納されている答案用紙情報に対する差分に関する情報である差分関連情報を取得する。ここでの差分とは、例えば、差分の画像情報である。答案読取情報の答案用紙情報に対する差分とは、答案読取情報に対する答案用紙情報の差分も含む概念である。差分関連情報とは、差分の情報そのものであっても良いし、差分に対して予め指定された加工を行って得られる情報、例えば、差分の情報を二値化した情報や、いわゆる輪郭補正等を行った情報であっても良い。差分関連情報は、例えば、答案読取情報と答案用紙情報との差分を示す画像情報、または、この画像情報に対して予め指定された加工を行って得られる画像情報である。差分取得部 1 3 は、例えば、差分関連情報の取得対象となる 1 以上の答案読取情報と、各答案読取情報に対応する答案用紙情報とをそれぞれ読み出し、各答案読取情報について、それぞれ、対応する答案用紙情報に対する差分を取得することで差分関連情報を取得する。あるいは、取得した差分に加工等を行って差分関連情報を取得する。なお、答案用紙情報が、ベクタ画像等である場合、適宜ラスタライズを行って差分を取得する。差分取得部 1 3 は、答案読取情報の画素のうちの、答案用紙情報と重なる画素間（例えば、座標が一致する画素）について、画素値が一致する場合に、その画素を削除することで、差分関連情報を取得する。または、画素値が一致する場合に、その画素の色の値を予め指定されている白色等の色の値に変更してもよい。もしくはその画素を不可視としてもよい。これによって、差分の画像を示す差分関連情報を取得することができる。画素を不可視とするためには、例えば、各画素の透過度を指定するチャンネル（

40

50

いわゆるアルファチャンネル)を差分関連情報に設けるようにすればよい。なお、差分関連情報においては、差分の存在する部分を含む最小矩形を検出して、この矩形以外の部分を削除しても良い。

【0051】

生徒識別情報取得部14は、差分取得部13が差分を取得した答案読取情報の識別画像を検出し、識別画像と対応付けられた生徒識別情報を取得する。生徒識別情報取得部14は、答案読取情報内において、識別画像の画像を検出し、検出した識別画像を、いわゆるバーコードリーダ等の識別画像に対応付けられた識別情報を読み出す装置等が利用するアルゴリズムと同様のアルゴリズムで処理することで、例えば、識別画像に対応する識別情報である生徒識別情報を取得する。これにより、生徒識別情報取得部14は、生徒識別情報

10

を取得したことになる。答案読取情報内から識別画像を検出する際には、例えば、識別画像のコーナー等に設けられている予め指定されているドットパターン等や輪郭線等を、答案読取情報内から探索し、検出したコーナー等で囲まれる領域に識別画像が配置されていると判断する。また、例えば、識別画像の配置される場所が座標等で予め決まっている場合、答案読取情報のこの予め決まっている場所の画像を、識別画像の画像として読み出す、あるいはこの場所について、上記と同様の識別画像の探索等を行うようにする。なお、画像内において、識別画像の位置を検出する処理等は、バーコードリーダの技術等として公知であるので、ここでは詳細な説明は省略する。また、二次元バーコード等の識別画像から対応する識別情報を取得する処理等は、バーコードリーダ等の識別画像を読み取る装置の技術等として公知であるので、ここでは詳細な説明は省略する。

20

【0052】

文字認識部15は、差分取得部13が取得した差分関連情報に対して文字認識処理を行う。文字認識処理として、例えば、OCR(光学文字認識:Optical Character Recognition)用いられる。文字認識の結果として、例えば、差分関連情報について認識された文字で構成されるテキスト情報を取得する。このテキスト情報は、文字サイズ等の書式の情報等を有するものであっても良い。OCRの技術については、公知の技術であるので説明は詳細な説明は省略するが、以下に、OCRの処理の一例を簡単に説明する。

【0053】

例えば、文字認識部15は、差分関連情報に対して2値化処理等を行って文字の画素の境界を検出したり、文字と文字以外の部分の色の濃度の変化点等を検出して文字の画素の境界を検出し、生徒の記入した各文字をそれぞれひとつの連続したかたまりとして認識し、その文字のかたまりをそれぞれ個別に囲む矩形のエリアが求められる。これにより、差分関連情報に含まれる各文字を個別に含む画像を検出したこととなる。そして、求めた各矩形のエリアに含まれる文字について、例えばベクトル化等を行って、交点位置情報や端点位置情報等の文字の特徴を示す値を取得して、これらの値等を、予め用意されている辞書に複数の文字と対応付けて格納されている同様の特徴と示す値と比較すること等によって、辞書に含まれる文字と、各矩形エリアの文字との、類似度のスコアを取得する。そして、一番スコアの良い文字を、認識された文字として辞書から取得する。なお、差分関連情報に含まれる何らかの文字を、個別に検出するまでの処理を、文字認識処理と考えても良い。

30

40

【0054】

文字認識部15が取得する文字認識処理の結果は、例えば、上述したような差分関連情報について認識された文字で構成されるテキスト情報であってもよいし、差分関連画像中の、文字であると判断された画像の数(例えば、上述した矩形エリアの数)等であっても良い。あるいは、上述したような文字認識の際に算出された類似度等を表すスコアであっても良い。例えば、文字認識処理の結果は、文字認識の対象となる文字の画像のそれぞれに対する、類似度が最も高いと判断された文字認識処理用の辞書に含まれる文字の、類似度等を表すスコアであってもよい。

【0055】

文字認識部15が文字認識処理を行う対象の差分関連情報は、答案情報の一部として後

50

述する答案情報格納部 16 に蓄積される前の差分関連情報でも良いし、後述する答案情報格納部 16 に格納されている答案情報に含まれる差分関連情報でも良い。例えば、文字認識部 15 は、差分取得部 13 が一の差分関連情報を取得した直後に、この差分関連情報について文字認識処理を行っても良いし、答案情報格納部 16 に格納されている答案情報に含まれる差分関連情報に対して、順次、文字認識処理を行ってもよい。

【0056】

答案情報格納部 16 には、差分関連情報と、生徒識別情報とを有する答案情報が格納される。一の答案情報が有する差分関連情報および生徒識別情報は、例えば、一の答案読取情報について、差分取得部 13 および生徒識別情報取得部 14 がそれぞれ取得した差分関連情報および生徒識別情報である。また、答案情報は、更に、この答案情報に含まれる差分関連情報について文字認識部 15 が文字認識処理を行った結果として取得されたテキスト情報を有しても良い。また、答案情報は、この答案情報に対応した日付を示す情報（以下、日付情報）を更に有していてもよい。日付情報（答案情報に対応した日付を示す情報）とは、上述したような答案読取情報に対応付けられていた日付の情報であっても良いし、答案情報が答案情報格納部 16 に蓄積された日付の情報であっても良い。ここでの格納は、一時記憶も含む概念である。

10

【0057】

答案情報蓄積部 17 は、差分取得部 13 が取得した差分関連情報と、生徒識別情報取得部 14 が取得した生徒識別情報とを有する答案情報を答案情報格納部 16 に蓄積する。答案情報蓄積部 17 が蓄積する一の答案情報は、一の答案読取情報について取得された差分関連情報と生徒識別情報とを有する答案情報である。また、答案情報蓄積部 17 は、答案情報に含まれる差分関連情報について文字認識部 15 が文字認識処理を行った結果として取得されたテキスト情報をさらに有する答案情報を答案情報格納部 16 に蓄積しても良い。また、答案情報蓄積部 17 は、上述したような答案情報に対応した日付を示す日付情報を更に有する答案情報を答案情報格納部 16 に蓄積しても良い。

20

【0058】

統計分析部 18 は、例えば、答案情報格納部 16 に格納された答案情報に含まれる生徒識別情報を用いて、生徒ごとに答案情報に関する統計分析を行う。生徒ごとに統計分析を行うということは、例えば、生徒識別情報別に統計分析を行うことである。例えば、同じ生徒識別情報を有する答案情報を用いて、統計分析を行うことである。

30

【0059】

また、統計分析部 18 は、答案情報が有する日付情報を用いて、予め指定された期間ごとに、答案情報に関する統計分析を行うようにしてもよい。例えば、統計分析部 18 は、答案情報格納部 16 に格納されている答案情報から、予め指定された期間ごと（例えば、年ごと）に、期間内に含まれる日付を示す日付情報を有する答案情報を検出し、検出した期間ごとの答案情報を用いて、統計分析を行う。なお、上述した生徒ごとの答案分析を、期間ごとに行うようにしても良い。

【0060】

なお、ここで述べる統計分析とは、どのような統計分析であっても良く、どのような統計分析が行われるかは、例えば、統計分析の目的や用途等に応じて設定される。例えば、ここでの統計分析は、統計分析の対象となる情報のうちの予め指定された条件を満たすものを条件ごとに集計したり、その集計結果の統計分析の対象となる情報全体に対する比率等を取得することである。また、例えば、統計分析の対象となる情報のうちの、予め指定された条件を満たす値を有するものについて、その値の平均値や、分散や、標準偏差、最大値、最小値、中央値等を取得することであってもよい。統計分析は、クラスター分析等の多変量解析であっても良い。統計分析の際には、統計分析に用いられる情報の値や、情報数や、統計分析の結果等を、適宜、正規化しても良い。なお、かかることは以下においても同様である。

40

【0061】

統計分析部 18 は、答案情報に関する統計分析として、例えば、答案情報に含まれる差

50

分関連情報に記載されている文字に関する統計分析を行う。例えば、統計分析部 18 は、答案情報に含まれる差分関連情報に対して文字認識部 15 により行われた文字認識処理の結果を示す情報を用いて文字に関する統計分析を行う。ここで述べる文字認識処理の結果とは、上述したような各差分関連情報について認識された文字で構成されるテキスト情報であってもよいし、差分関連画像中の、一の文字であると判断された画像の数（即ち、上記の矩形のエリアの数）や、文字認識された画像の数や、文字認識できなかった画像の数等であっても良い。あるいは、各差分関連情報において、文字であると判断された各画像に対して文字認識の処理において得られた、辞書等に含まれる文字に対する類似度等を表すスコアであっても良い。

【0062】

統計分析部 18 が行う差分関連情報に記載されている文字に関する統計分析とは、例えば、文字認識部 15 による文字認識処理において、差分関連情報について認識された文字の量に関する統計分析である。ここで述べる文字の量とは、例えば、文字数である。また、文字の量は、文字認識処理の結果として得られるテキスト情報に含まれる単語数等と考えても良い。テキスト情報に含まれる単語は、例えば、形態素解析を行うこと等により検出することができる。形態素解析については、公知の技術であるので、ここでは詳細な説明を省略する。

【0063】

文字の量に関する統計分析とは、例えば、予め指定された条件を満たす（あるいは満たさない）文字の集計を行うことや、この集計数の全体に対する比率を算出することと等である。予め指定された条件は、例えば、文字種を指定する条件である。文字種とは、例えば、漢字、ひらがな、アルファベット等である。また、予め指定された条件は、文字認識処理で認識された全ての文字という条件であってもよい。また、複数の答案読取情報に含まれる予め指定された条件を満たす（あるいは満たさない）文字の集計数の平均であっても良い。例えば、一のユーザに対応する複数の差分関連情報に対してそれぞれ行われた文字認識処理により得られたテキスト情報に含まれる全ての文字の文字数の集計を各差分関連情報ごとにそれぞれ行う。そして、その集計結果の平均値を統計分析結果として取得する。

【0064】

統計分析部 18 が行う差分関連情報に記載されている文字に関する統計分析とは、文字認識部 15 により差分関連情報について認識された文字（具体的には、文字認識結果として得られたテキスト情報に含まれる文字）の多様性に関する統計分析である。文字の多様性とは、文字で構成される単語の多様性と考えても良い。ここでの多様性とは、例えば、文字認識により認識された文字（または単語）のうちの、ユニークな文字（または単語）の数である。例えば、重複を除外して（例えば、重複するものは一度しかカウントしないようにして）文字や単語をカウントした集計数である。あるいは文字認識により認識された文字（または単語）のうちの、文字種が予め指定された文字種（漢字、ひらがな等）であるユニークな文字または単語の数である。単語の多様性とは、特定のグループに属する単語の使用されている数や比率等を示す情報であっても良い。特定のグループとは、例えば、語彙の難易度で分類された単語のグループである。例えば、どのような難易度の語彙がテキスト情報等に含まれているかを取得する処理等は、以下の非特許文献等の技術を用いることで実現可能である。（非特許文献：川村よし子、"語彙チェッカーを用いた日本語教科書の分析"、[online]、[平成11年9月8日検索]、インターネット<URL: <http://language.tiu.ac.jp/castel99.pdf>>）。

【0065】

統計分析部 18 が行う差分関連情報に記載されている文字に関する統計分析とは、例えば、記載されている文字の属性に関する統計分析である。文字の属性とは、文字のサイズや、文字認識のしやすさ等である。

【0066】

例えば、統計分析部 18 が行う文字の属性に関する統計分析とは、差分関連情報に対し

10

20

30

40

50

て文字認識部 15 により行われた文字認識の結果に関する統計分析である。文字認識の結果に関する統計分析とは、例えば、文字認識処理において認識された各文字と、これらの文字の認識対象となった差分関連情報内の各文字を含む画像との類似度等を示すスコアの、1 以上の差分関連情報についての、平均値や、分散等を取得することであっても良い。スコアの高い文字は、文字認識によって認識しやすい文字、例えば、きれいな文字であると考えられるため、このような統計分析結果は文字のきれいさを示すものとなる。なお、このスコアが予め用意された閾値以下の文字数と、文字認識された全ての文字数との比率等を取得することを、文字認識結果に関する統計分析と考えても良い。

【0067】

また、統計分析部 18 が行う文字の属性に関する統計分析とは、差分関連情報に対して文字認識部 15 により行われた文字認識結果が示す認識された文字のサイズに関する統計分析である。認識された文字のサイズとは、例えば、文字認識処理で認識された文字のサイズ（例えばポイント数）であっても良い。また、文字認識処理の過程で、一の文字を含む画像に対して設定された矩形のエリアの縦や横の長さとしても良い。統計分析部 18 は、例えば、認識された文字の数をサイズ別に集計した値や、集計した値の全体の文字数に対する比率等を統計分析部 18 が統計分析結果として取得する。また、文字サイズの分散等を統計分析結果として取得することで、文字サイズのばらつきを判断できるようにしても良い。例えば、一般的に、文字サイズのばらつきが小さい方が、きれいな文字であると判断可能である。

10

【0068】

統計分析部 18 が行う差分関連情報に記載されている文字に関する統計分析とは、例えば、文字認識部 15 が認識した文字が示す文章の正確度に関する統計分析である。ここでの文字認識部 15 が認識した文字とは、例えば、文字認識結果として出力されるテキスト情報に含まれる文字である。文章の正確度に関する統計分析とは、例えば、文字認識結果として出力された文章、即ち、差分関連情報の文字の画像から判断された特定の文字で構成される文章内における誤字脱字、漢字の間違い、助詞の組合せの間違い、辞書にない単語の使用、同音語誤り、送りがなの間違い、句読点の間違い、慣用表現の間違い等の項目が発生している数についての統計分析である。これらの統計分析は項目別の統計分析であっても、複数の項目をまとめた統計分析であっても良い。なお、文章から上記のような項目の箇所を検出する処理は、文書校正支援ソフトウェアや、ワードプロセッサ等で公知の技術であるので、ここでは詳細な説明は省略する。例えば、以下の非特許文献等に一例が開示されている。（非特許文献：“細やかなチェックで、校正者の負担を軽減 | 文章校正支援ツール Just Right!4 Pro CE”、[online]、[平成 11 年 9 月 8 日検索]、インターネット<URL : <http://www.justsystems.com/jp/products/justright/feature1.html>>）。

なお、文章の正確度についての統計分析とは、結果的に文章の正確度の状況や傾向等が判断可能な統計分析であればよく、文章の不正確さについての統計分析も、文章の正確度についての統計分析の一形態と考えて良い。

20

30

【0069】

また、統計分析部 18 が行う差分関連情報に記載されている文字に関する統計分析とは、記載されている文字の単位面積当たりの文字数に関する統計分析である。単位面積当たりの文字数とは、例えば、一の差分関連情報の面積（例えば読み取られた面積）に対する差分関連情報において認識された文字数の比率である。この認識された文字数は、文字認識部 15 により認識された文字数（例えば、文字認識結果として得られたテキスト情報に含まれる文字数）であってもよいし、文字認識部 15 で文字認識の過程で差分関連情報に検出した一の文字を含む矩形のエリアの数であっても良い。ここでの文字数は、上述したような特定の属性を有する文字の文字数であっても良い。この単位面積当たりの文字数は、例えば、答案用紙に対して文字が記入されている度合いを示す情報となる。

40

【0070】

出力部 19 は、統計分析部 18 の統計分析結果を出力する。ここで述べる出力とは、ディスプレイへの表示、プロジェクターを用いた投影、プリンタへの印字、外部の装置への

50

送信、記録媒体への蓄積、他の処理装置や他のプログラムなどへの処理結果の引渡しなどを含む概念である。例えば、出力部 19 は、統計分析結果を、グラフ等を用いて出力しても良い。

【0071】

出力部 19 は、ディスプレイ等の出力デバイスを含むと考えるとも含まないと考えるとも良い。出力部 19 は、出力デバイスのドライバーソフトまたは、出力デバイスのドライバーソフトと出力デバイス等で実現され得る。

【0072】

次に、答案処理装置 1 の動作について図 3 のフローチャートを用いて説明する。

なお、同じ答案用紙について得られた答案用紙情報と、答案読取情報とには、予め同じ答案用紙の識別情報（以下、答案用紙識別情報と称す）が対応付けられているものとする。また、答案用紙情報と、答案読取情報とには、生徒により答案が記入された日付の情報が対応付けられているものとする。

10

【0073】

（ステップ S 101）差分取得部 13 は、差分関連情報を取得する処理が未処理である答案読取情報が、答案読取情報格納部 11 に蓄積されているか否かを判断する。蓄積されている場合、ステップ S 102 に進み、蓄積されていない場合、ステップ S 111 に進む。例えば、差分取得部 13 が差分関連情報を取得した答案読取情報については、処理済であることを示すフラグを付与することで、このフラグの有無により、未処理の答案読取情報の有無を判断することができる。あるいは、処理済の答案読取情報は、削除するようにしても良い。

20

【0074】

（ステップ S 102）差分取得部 13 は、カウンター m に 1 を代入する。

【0075】

（ステップ S 103）差分取得部 13 は、差分関連情報の取得が未処理の答案読取情報に、m 番目の答案読取情報があるか否かを判断する。有る場合、ステップ S 104 に進み、ない場合、ステップ S 111 に進む。

【0076】

（ステップ S 104）差分取得部 13 は、答案用紙情報格納部 12 から、m 番目の答案読取情報に対応する答案用紙情報を取得する。例えば、m 番目の答案読取情報に対応付けられている答案用紙識別情報と同じ答案用紙識別情報に対応付けられている答案用紙情報を取得する。

30

【0077】

（ステップ S 105）差分取得部 13 は、答案用紙情報に対する m 番目の答案読取情報の差分である差分関連情報を取得する。なお、差分に対して、シャープネス処理等の処理を行ったものを差分関連情報として取得してもよい。また、差分関連情報を取得する際には、識別画像情報が配置される領域近傍等については、差分の取得を行う対象領域から予め除外しておくようにしても良い。なお、m 番目の答案読取情報には、差分関連情報を取得する処理が処理済みであることを示すフラグを付与する。

【0078】

（ステップ S 106）文字認識部 14 は、ステップ S 105 において取得した差分関連情報に対して文字認識処理を行う。そして、例えば文字認識結果を示す情報として、文字認識により特定された文字列で構成されるテキスト情報を取得する。また、ここでは、例えば、文字認識した際の各文字についてのスコアを示す情報を取得する。なお、このテキスト情報は、文字サイズの情報等を有するものであるとする。

40

【0079】

（ステップ S 107）生徒識別情報取得部 14 は、m 番目の答案読取情報から、識別画像を検出する。

【0080】

（ステップ S 108）生徒識別情報取得部 14 は、ステップ S 107 で検出した識別画

50

像情報を読み出して、生徒識別情報を取得する。

【0081】

(ステップS109) 答案情報蓄積部17は、ステップS105で取得した差分関連情報と、ステップS108で取得した生徒識別情報と、ステップS106で取得した文字認識結果(ここでは、テキスト情報と文字ごとのスコア)と、m番目の答案読取情報に対応付けられた日付の情報である日付情報とを有する答案情報を、答案情報格納部16に蓄積する。

【0082】

(ステップS110) 差分取得部13は、カウンターmの値を1インクリメントする。そして、ステップS103に戻る。

10

【0083】

(ステップS111) 統計分析部18は、統計分析を行うか否かを判断する。例えば、ユーザ等により図示しない受付部等を介して統計分析を行う指示を受け付けている場合に、統計分析を行うと判断し、受け付けていない場合、統計分析しないと判断する。統計分析を行う場合、ステップS112に進み、受け付けていない場合、ステップS101に戻る。

【0084】

(ステップS112) 統計分析部18は、図示しない受付部等を介してユーザから分析対象の1以上の期間を指定する情報を受け付けたか否かを判断する。ここでの期間は、例えば年や年度である。受け付けた場合、ステップS113に進み、受け付けていない場合、ステップS112に戻る。

20

【0085】

(ステップS113) 統計分析部18は、カウンターnに1を代入する。

【0086】

(ステップS114) 統計分析部18は、答案情報格納部16に格納されている答案情報の中に、n番目の生徒識別情報があるか否かを判断する。ある場合、ステップS115に進み、ない場合、ステップS120に進む。

【0087】

(ステップS115) 統計分析部18は、カウンターkに1を代入する。

【0088】

(ステップS116) 統計分析部18は、ステップS112において指定された期間の中に、k番目の期間の指定があるか否かを判断する。有る場合、ステップS117に進み、ない場合、ステップS119に進む。

30

【0089】

(ステップS117) 統計分析部18は、統計処理を行う。この処理の詳細については、後述する。

【0090】

(ステップS118) 統計分析部18は、カウンターkの値を1インクリメントする。そして、ステップS116に戻る。

【0091】

(ステップS119) 統計分析部18は、カウンターnの値を1インクリメントする。そして、ステップS114に戻る。

40

【0092】

(ステップS120) 出力部19は、統計分析部18が取得した統計処理の結果を出力する。そして、処理を終了する。

【0093】

なお、統計分析部18は、ステップS112の前後において、統計分析対象となる生徒識別情報を受け付けても良い。そして、統計分析部18は、この受け付けた生徒識別情報について、ステップS114の処理等を行うようにしてもよい。

【0094】

50

次に、答案処理装置 1 による統計処理の詳細な動作について図 4 のフローチャートを用いて説明する。この処理は、上述したステップ S 1 1 7 の詳細な動作の一例である。

【0095】

(ステップ S 2 0 1) 統計分析部 1 8 は、カウンター p の値を 1 インクリメントする。

【0096】

(ステップ S 2 0 2) 統計分析部 1 8 は、n 番目の生徒識別情報と、k 番目の期間に含まれる日付を示す日付情報とを有する p 番目の答案情報があるか否かを判断する。ある場合、ステップ S 2 0 3 に進み、ない場合、ステップ S 2 1 0 に進む。

【0097】

(ステップ S 2 0 3) 統計分析部 1 8 は、p 番目の答案情報に含まれる文字認識処理結果を示す情報(例えば、テキスト情報)を用いて、文字の量を示す情報を取得する。例えば、認識された文字数を取得する。取得した情報は、例えば、p 番目の答案情報と対応付けて、図示しない記憶媒体等に蓄積する。

10

【0098】

(ステップ S 2 0 4) 統計分析部 1 8 は、p 番目の答案情報に含まれる文字認識処理結果を示す情報を用いて、文字の多様性を示す情報を取得する。例えば、文字認識処理結果に含まれているユニークな単語の数を取得する。ユニークな単語の数とは、重複した単語を 1 つの単語としてカウントした場合の単語の数である。取得した情報は、例えば、p 番目の答案情報と対応付けて、図示しない記憶媒体等に蓄積する。

【0099】

(ステップ S 2 0 5) 統計分析部 1 8 は、p 番目の答案情報に含まれる文字認識処理の結果に関する情報を取得する。例えば、統計分析部 1 8 は、文字認識が行われた際の各文字のスコアを用いて、文字認識された文字のスコアの平均値を取得する。取得した情報は、例えば、p 番目の答案情報と対応付けて、図示しない記憶媒体等に蓄積する。

20

【0100】

(ステップ S 2 0 6) 統計分析部 1 8 は、p 番目の答案情報に含まれる文字認識処理結果を示す情報を用いて、文字のサイズに関する値を取得する。例えば、文字認識処理結果に含まれる複数の文字の文字サイズの値を用いて、文字のサイズの平均値を取得する。取得した情報は、例えば、p 番目の答案情報と対応付けて、図示しない記憶媒体等に蓄積する。

30

【0101】

(ステップ S 2 0 7) 統計分析部 1 8 は、p 番目の答案情報に含まれる文字認識処理結果を示す情報を用いて、文字の正確度に関する値を取得する。例えば、文字認識処理結果に含まれる複数の文字を用いて、誤字脱字の数等を取得する。取得した情報は、例えば、p 番目の答案情報と対応付けて、図示しない記憶媒体等に蓄積する。

【0102】

(ステップ S 2 0 8) 統計分析部 1 8 は、p 番目の答案情報に含まれる文字認識処理結果を示す情報を用いて、単位面積当たりの文字数を取得する。例えば、文字認識処理結果が示す文字数を、p 番目の答案情報に対応する答案読取情報のサイズ(面積)や、p 番目の答案情報に含まれる差分関連情報のサイズで除算して単位面積当たりの文字数を取得する。取得した情報は、例えば、p 番目の答案情報と対応付けて、図示しない記憶媒体等に蓄積する。

40

【0103】

(ステップ S 2 0 9) 統計分析部 1 8 は、カウンター p の値を 1 インクリメントする。そして、ステップ S 2 0 2 に戻る。

【0104】

(ステップ S 2 1 0) 統計分析部 1 8 は、p 個の答案情報についてステップ S 2 0 3 からステップ S 2 0 8 においてそれぞれ取得した情報を用いて、各ステップで取得したそれぞれの値についての平均値を統計分析結果として取得する。なお、平均値を取得する代わりに、他の統計処理に関する値、例えば、分散や標準偏差等、を取得しても良い。

50

【 0 1 0 5 】

(ステップ S 2 1 1) 統計分析部 1 8 は、ステップ S 2 1 0 で取得した統計分析結果を、n 番目の生徒識別情報と、k 番目の期間を示す情報と対応付けて図示しない記憶媒体等に蓄積する。そして、上位の処理にリターンする。

【 0 1 0 6 】

以下、本実施の形態における答案処理装置 1 の具体的な動作について説明する。

【 0 1 0 7 】

例えば、ある学校の各生徒に、それぞれ、図 2 に示すような、各生徒の生徒識別情報と対応付けられた識別画像 2 1 が配置されたシール 2 0 が配布されており、各生徒が試験を受ける際に、このシールを答案用紙の予め指定された領域に貼付するものとする。ここでは、生徒識別情報は、一例として、生徒名であるとする。また、識別画像 2 1 は、ここでは一例として、二次元バーコードの一つである QR コードであるとする。このシール 2 0 は、図 2 に示したように、識別画像 2 1 の下辺近傍に、生徒識別情報 2 2 を示す文字が記載されているものとする。

10

【 0 1 0 8 】

ここでは、例えば、生徒に対してテストが行われるごとに、テストに用いられた答案用紙と、テストにおいて生徒が答案を記入した答案用紙とが回収され、教師等によって、これらの答案用紙が、スキャナ等(図示せず)を用いて画像情報に変換されて、答案用紙情報格納部 1 2 および答案読取情報格納部 1 1 に蓄積されているとする。また、ここでは、一例として、答案読取情報を蓄積する際には、テストが行われた日付の情報が対応付けて蓄積されるとともに、一のテストに用いられた答案用紙について得られる答案用紙情報と答案読取情報との両方には、教師等によって図示しない受付部等を介して入力された上記の答案用紙の識別情報(以下、答案用紙識別情報と称す)が共に対応付けて蓄積されるものとする。

20

【 0 1 0 9 】

図 5 は、上述した生徒が受ける一のテストに用いられた答案用紙の答案の記入面をスキャナ等(図示せず)で読み取った答案用紙情報 5 1 を示す図(図 5 (a))、およびこの答案用紙と同じ答案用紙に対して、各生徒が答案を記入した答案用紙の答案面をスキャナ等(図示せず)で読み取った答案読取情報 5 2 の一例を示す図(図 5 (b))である。図 5 (a)の答案用紙情報 5 1 に対応する答案用紙は、答案が未記入の答案用紙である。この答案用紙は、答案用紙のサンプルや原本等と考えても良い。また、図 5 (b)の答案読取情報 5 2 に対応する生徒が答案を記入した答案用紙には、予め指定された領域である枠 5 0 1 内に、答案を記入した生徒のシール 2 0 が、生徒によって貼付されているものとする。

30

【 0 1 1 0 】

図 6 は、答案用紙情報格納部 1 2 に格納されている答案用紙情報を管理する答案用紙情報管理表である。答案用紙情報管理表は、「答案用紙情報」と「答案用紙 ID」という項目を有している。「答案用紙情報」は、答案用紙情報のファイル名を示している。「答案用紙 ID」は、答案用紙情報と対応付けられた答案用紙識別情報である。図 5 (a)に示した答案用紙情報に対応付けられた「答案用紙 ID」は「K 2 0 1 1 - 1 0 1」とであるとする。なお、答案用紙情報のファイル名を答案用紙識別情報としても用いるようにすることで、「答案用紙情報」と「答案用紙 ID」とを一つの項目としても良い。

40

【 0 1 1 1 】

図 7 は、答案読取情報格納部 1 1 に格納されている答案読取情報を管理する答案読取情報管理表を示す図である。答案読取情報管理表は、「答案読取情報」と、「答案用紙 ID」と、「日付」と、「処理済」という項目を有している。「答案読取情報」は答案読取情報のファイル名を示している。「答案用紙 ID」は、答案読取情報に対応付けられた答案用紙識別情報であり、図 6 の「答案用紙 ID」と対応している。「日付」は、答案読取情報に対応する試験が行われた日付である。「処理済」は、差分関連情報を取得する処理が済んでいる答案読取情報であるか否かを示すフラグ情報であり、値「0」は、未処理、値

50

「1」は処理済であることを示す。

【0112】

次に、教師等のユーザが、テストに対する統計分析を行うために、答案処理装置1を起動させたとすると、差分取得部13は、答案読取情報格納部11に、差分関連情報を取得する処理が未処理の答案読取情報があるか否かを判断する。ここでは、図7に示した答案読取情報で「処理済」の値が、「0」であるものがあるか否かを判断する。ここでは、「処理済」の値が「0」であるレコードが存在していたとする。ただし、判断開始のトリガーやタイミング等は問わないものとする。

【0113】

差分取得部13は、未処理の答案読取情報があるため、未処理の答案読取情報を順次読み出し、差分関連情報を取得する処理を行う。例えば、まず、図7に示した答案読取情報管理表の一番目のレコード(行)から、ファイル名が「01000125.tif」である答案読取情報をメモリ等に読み出す。即ち、図5(b)に示した答案読取情報を読み出す。また、この答案読取情報に対応する答案用紙識別情報、具体的には、同じレコードの「答案用紙ID」の値である「K2011-001」を読み出し、図6に示した答案用紙情報管理表から、「答案用紙ID」が「K2011-001」と対応付けられた答案用紙情報、具体的には、ファイル名が「100545.tif」である答案用紙情報を取得する。即ち、図5(a)に示した答案用紙情報を読み出す。

10

【0114】

差分取得部13は、上記で読み出したファイル名が「01000125.tif」である答案読取情報の、ファイル名が「100545.tif」である答案用紙情報に対する差分の画像を取得する。即ち、図5(b)に示した答案読取情報の、図5(a)に示した答案用紙情報に対する差分の画像を取得する。例えば、重複する座標にある画素同士のうち、色情報(あるいは階調情報)が一致するものを削除する。ただし、ここでは予め指定された範囲内の色情報の違いがあるものも一致と考えてよい。また、画素を削除する代わりに、予め指定された色(例えば白色)の画素に変更するようにしても良い。なお、ここでは、一例として、上述したようなシール20を貼付するための枠501内の領域については、予め、この領域を示す座標等を示しておくことで、差分の画像を取得する対象に含めないようにして、この枠501内の画像は、差分の画像から削除する。そして、差分の画像に対して、適宜、シャープネスの処理や、レベル補正の処理等を行って、差分関連情報を得る。これにより、生徒の記入した答案のみを示す画像である差分関連情報を得る。なお、差分の画像をそのまま、差分関連情報として取得してもよい。また、差分取得部13は、答案読取情報「01000125.tif」に対応付けられていた日付の情報である「2011/5/25」を日付情報として取得する。また、差分取得部13は、答案読取情報管理表において、差分関連情報を取得した答案読取情報「01000125.tif」を含むレコードの「取得済」の値を「1」に変更する。

20

30

【0115】

図8は、差分取得部13が取得した差分関連情報の一例を示す図である。取得した差分関連情報には、例えば、予め指定されたルール等に応じてファイル名等が付与され、図示しない記憶媒体等に一時記憶される。ここでは、「500001.tif」というファイル名が付与されたとする。

40

【0116】

次に、生徒識別情報取得部14は、上述したファイル名が「100545.tif」である答案用紙情報の、枠501内において、識別画像21を検出する。識別画像21は、通常は、例えば、識別画像21のコーナー等に設けられた特定の形状の画像をパターンマッチング等で検出することによって検出可能である。ここでは、シール20上の識別画像21であるQRコードを検出する。そして、この識別画像21を読み出して、読み出した識別画像21に対応付けられた生徒識別情報を取得する。ここでは、例えば、「山田A男」という生徒識別情報を取得したとする。

【0117】

50

次に、文字認識部 15 は、図 8 に示した差分関連情報に対して文字認識処理を行う。具体的には OCR の処理を行う。例えば、図 8 に示した差分関連情報において、一文字と考えられる画素のかたまりを順次検出して、それぞれの塊を囲む矩形のエリアを設定する。

【0118】

次に、各矩形のエリアに含まれる文字の画像の特徴を示す情報を取得し、各矩形のエリアに含まれる文字の画像と、予め用意されている文字認識処理用の文字の辞書（図示せず）に含まれる各文字との類似度等を示すスコアを各矩形のエリアごとに取得し、各矩形のエリアにおいて、最もスコアの高かった文字を、矩形のエリアについて認識された文字として取得する。そして、認識された文字を認識された位置の近傍等に配置したテキスト情報を文字認識結果として取得する。このテキスト情報のファイル名は、例えば、予め指定されたルール等により付与される。ここでは、例えば、「500001.rtf」というファイル名が付与されたとする。また、ここでは、認識された文字の文字認識処理時に得られたスコア、および認識された文字のサイズの情報を、認識された文字ごとに対応付けた情報である認識管理表も文字認識結果として取得する。この認識管理表のファイル名は、例えば、予め指定されたルール等により付与される。ここでは、例えば、「500001.csv」というファイル名が付与されたとする。

10

【0119】

図 9 は、文字認識部 15 による文字認識の結果として取得されたテキスト情報を示す図（図 9（a））、および、認識された各文字と、文字認識処理時のスコアと、文字のサイズの情報とを対応付けて管理する認識管理表を示す図（図 9（b））である。認識管理表は、認識された文字を示す「認識文字」、認識された文字のスコアである「スコア」、認識された文字のサイズである「サイズ」という項目を有している。認識管理表は、ここでは、例えば、差分関連情報ごとに作成されるものとする。ただし、複数の差分関連情報に対して一の認識管理表を用意し、各差分関連情報に対応するレコードに対しては、差分関連情報の識別情報等に対応付けるようにしても良い。

20

【0120】

答案情報蓄積部 17 は、差分取得部 13 が取得したファイル名が「500001.tif」である差分関連情報と、生徒識別情報取得部 14 が取得した生徒識別情報「山田 A 男」と、答案用紙識別情報「K2011-001」と、文字認識処理の結果として得られたテキスト情報「500001.rtf」と、文字認識処理の結果として得られた認識管理表「500001.csv」と、差分取得部 13 が取得した日付情報「2011/5/25」とを有する答案情報を、答案情報格納部 16 に蓄積する。

30

【0121】

以下、同様にして、差分関連情報が未処理の答案読取情報について、同様の処理を繰り返して、答案情報を蓄積（追記）していく。なお、蓄積を行う際に、生徒識別情報の値と、答案用紙識別情報の値とが一致する答案情報が既に蓄積されているか否かを判断し、蓄積されていると判断した場合、同じテストに対する同じ生徒の答案情報が、二重登録されることとなるため、二重登録がされる旨の警告や、何らかの不正が行われている恐れがある旨の警告等をユーザ等に出力するようにしても良い。

【0122】

図 10 は、答案情報格納部 16 に格納されている答案情報を管理する答案情報管理表を示す図である。答案情報管理表は、「答案 ID」と、「差分関連情報」と、「生徒 ID」と、「答案用紙 ID」と、「認識テキスト」と、「認識管理表」と、「日付情報」という項目を有している。「答案 ID」は、答案情報の識別情報である。ここでは、一例として、答案情報に含まれる差分関連情報のファイル名から、拡張子を除いたものを「答案 ID」としている。ただし、どのようなルールに従って「答案 ID」を付与するかは問わない。「差分関連情報」は、差分関連情報のファイル名である。「生徒 ID」は生徒識別情報である。「答案用紙 ID」は、答案用紙識別情報であり、図 6 および図 7 の「答案用紙 ID」に対応する。「認識テキスト」は、文字認識処理により得られたテキスト情報である。「認識管理表」は、文字認識処理により得られた認識管理表である。「日付情報」は、

40

50

差分取得部 13 が取得した日付情報である。答案読取情報「01000125.tif」について上記で取得された答案情報は、「差分関連情報」が「500001.tif」であるレコードである。

【0123】

次に、教師等のユーザが、上記で取得した答案情報についての統計分析を行う指示を、図示しない受付部等を介して、答案処理装置 1 に与えたとすると、答案処理装置 1 は、統計分析の対象となる期間を指定する指示を受け付けるためのインターフェース画面（図示せず）を、図示しないモニタ等に表示する。そして、ユーザが、このインターフェース画面を操作して、2008年から2011年までの期間の各年についての統計分析を行う指示を答案処理装置 1 に与えたとする。

10

【0124】

この指示を受け付けると、統計分析部 18 は、まず、図 10 の答案情報管理表の各レコードからそれぞれ「生徒 ID」の値を順次読み出し、図示しない記憶媒体等に一時的に蓄積（追記）していく。このとき、同じ「生徒 ID」が既に蓄積されている場合、重複した蓄積は行わない。このようにして、答案用紙を提出した生徒についてのユニークな生徒識別情報のリストを得る。

【0125】

次に、統計分析部 18 は、取得した生徒識別情報のリストから、一の生徒識別情報を取得する。ここでは、例えば、生徒識別情報「山田 A 男」を取得したとする。また、統計分析部 18 は、上記で指定された期間内の一の年を示す情報、例えば「2011」を取得する。そして、図 10 に示した答案情報管理表において、「生徒 ID」が「山田 A 男」であり、かつ、「日付情報」の値が、「2011/1/1」から「2011/12/31」までのいずれかの日付の値と一致するレコード（答案情報）の一つを検索する。ここでは、例えば、「答案 ID」が「500001」であるレコードを検出したとする。

20

なお、上記の年を示す情報を、年度の情報として扱うようにし、年を示す情報として「2011」を取得した場合に、小学校や中学校等の年度の開始日と終了日とを考慮して、統計分析部 18 は、「日付情報」の値が、この 2011 年度を示す範囲の「2011/4/1」から「2012/3/31」までのいずれかの日付の値と一致するレコード（答案情報）の一つを検索するようにしてもよい。つまり、年度の最初を、指定された年の「4月1日」とし、年度の最終を、指定された年の翌年の「3月31日」としてもよい。

30

【0126】

そして、まず、統計分析部 18 は、検索したレコードで管理される「認識テキスト」の文字の量として、文字数を取得する。例えば、ここでは、認識テキスト「500001.rtf」について文字の数をカウントする。カウント結果が、例えば、「124」であったとする。統計分析部 18 は、この値を、例えば、答案 ID「500001」と対応付けて図示しない記憶媒体等に蓄積する。

【0127】

次に、統計分析部 18 は、上記で検索したレコードで管理される認識テキスト「500001.rtf」について文字の多様性を示す値を取得する。ここでは、例として、文字の多様性を示す値として、ユニークな単語数を取得する。例えば、図 9 (a) に示した認識テキスト「500001.rtf」のテキスト情報に対して、形態素解析を行い、テキスト情報を単語に分割する。そして、分割した単語を、重複したものは一度だけカウントするようにしてカウントすることで、ユニークな単語数を取得する。ここで取得したユニークな単語数は、「53」であったとする。統計分析部 18 は、この値を、例えば、答案 ID「500001」と対応付けて図示しない記憶媒体等に蓄積する。

40

【0128】

次に、統計分析部 18 は、上記で検索したレコードで管理される認識管理表「500001.csv」を用いて、文字認識処理の結果に関する値を取得する。ここでは、例として、文字認識された各文字についての、文字認識処理に取得されたスコアを取得して、これらの平均値を取得する。例えば、図 9 (b) に示した認識管理表「500001.csv

50

v」の全てのレコードの「スコア」の値を加算し、加算した値を全レコード数で除算することで、スコアの平均値を取得する。取得した平均値が、例えば、「71」であったとする。統計分析部18は、この値を、例えば、答案ID「500001」と対応付けて図示しない記憶媒体等に蓄積する。

【0129】

次に、統計分析部18は、上記で検索したレコードで管理される認識管理表「500001.csv」を用いて、文字認識処理で認識された文字のサイズに関する値を取得する。ここでは、例えば、文字認識された各文字についてのサイズを取得して、これらの平均値を取得する。例えば、図9(b)に示した認識管理表「500001.csv」の全てのレコードの「サイズ」の値を加算し、加算した値を全レコード数で除算することで、文字サイズの平均値を取得する。取得した平均値が、例えば、「15.8」(ポイント)であったとする。統計分析部18は、この値を、例えば、答案ID「500001」と対応付けて図示しない記憶媒体等に蓄積する。

10

【0130】

次に、統計分析部18は、上記で検索したレコードで管理される認識テキスト「500001.rtf」について文字の正確度に関する値を取得する。ここでは、例として、文字の正確度に関する値として、誤字脱字の数を取得する。つまり、誤字脱字の数が小さければ生徒が記入した文章の正確度が高いことを示すこととなる。例えば、図9(a)に示した認識テキスト「500001.rtf」のテキスト情報に対して、誤字脱字を検出する処理を行い誤字脱字を検出する。そして、検出した誤字脱字の数をカウントする。そして、カウントした誤字脱字数を、上記で取得したテキスト情報の文字数で除算することで、正規化を行う。これにより取得した正規化された誤字脱字の数は、「5」であったとする。統計分析部18は、この値を、例えば、答案ID「500001」と対応付けて図示しない記憶媒体等に蓄積する。

20

【0131】

次に、統計分析部18は、上記で検索したレコードに含まれる認識テキスト「500001.rtf」について単位面積当たりの文字数を示す値を取得する。例えば、上記で取得した認識テキストの文字数を、差分関連情報の面積で除算することで、単位面積当たりの文字数を取得する。例えば、差分関連情報のサイズが、高さ7インチ、幅12インチであったとすると、単位面積当たり(ここでは1平方インチ当たり)の文字数は、124(文字)÷84 = 「1.48」文字となる。統計分析部18は、この値を、例えば、答案ID「500001」と対応付けて図示しない記憶媒体等に蓄積する。

30

【0132】

これにより、「生徒ID」が「山田A男」であり、かつ、「答案ID」が「500001」である答案情報についての処理が終了する。

【0133】

統計分析部18は、更に、「生徒ID」が「山田A男」であり、かつ、「日付情報」の先頭の四文字が、「2011」と一致する残りのレコード(答案情報)についても、上記と同様の処理を繰り返し行う。

【0134】

これにより、「生徒ID」が「山田A男」である生徒が、「2011」年に提出した答案用紙についての処理が終了する。

40

【0135】

図11は、統計分析部18が、「生徒ID」が「山田A男」である生徒が「2011」年に提出した答案用紙について上記の処理で取得した情報である年別取得情報を示す図である。この情報は、上記の処理により、図示しない記憶媒体値に蓄積された情報である。年別取得情報は、「答案ID」と、「文字数」と、「単語数」と、「平均スコア」と、「平均サイズ」と、「誤字脱字数」と、「単位文字数」という項目を有している。「文字数」は、文字認識結果であるテキスト情報に含まれる文字数、「単語数」は、文字認識結果であるテキスト情報に含まれるユニークな単語数、「平均スコア」は、文字認識処理の際

50

に取得されたスコアの平均値、「平均サイズ」文字認識処理で認識された文字サイズの平均値、「誤字脱字数」は文字認識結果であるテキスト情報に含まれる正規化された誤字脱字数、「単位文字数」は差分関連情報における単位面積当たりの文字数である。

【0136】

次に、統計分析部18は、図11に示した各項目ごとに、値の平均値を取得する。具体的には、全てのレコードの同じ項目の値同士を加算し、加算した値を、レコード数で除算して、各項目の平均値を取得する。この平均値は、「生徒ID」が「山田A男」である生徒が、「2011」年に提出した答案用紙についての各項目の値の平均値である。「生徒ID」が「山田A男」である生徒が、「2011」年に提出した答案用紙についての統計分析結果と考えても良い。そして、取得した各項目の平均値を、統計分析結果として、生徒識別情報「山田A男」と、年を示す値「2011」と対応付けて図示しない記憶媒体に蓄積する。

10

【0137】

同様にして、統計分析部18は、残りの年である「2010」年から「2008」年までの各年について、同様の処理を行う。

【0138】

図12は、上記の処理より取得されて蓄積された各年の統計分析結果を示す年別分析結果管理表である。「年」は、統計分析対象の年を示す。「答案ID」と、「文字数」と、「単語数」と、「平均スコア」と、「平均サイズ」と、「誤字脱字数」と、「単位文字数」という項目は、図11に示す同じ名称の項目の値の平均値である。

20

【0139】

更に、統計分析部18は、上記で取得した生徒識別情報のリストの、「山田A男」以外の生徒識別情報についても、上記と同様の処理を行い、処理により得られた情報を、図12に示す年別分析結果に追記していく。

【0140】

図13は、「山田A男」以外の生徒識別情報についての統計分析結果が追記された年別分析結果管理表を示す。

【0141】

そして、出力部19は、図13に示すような統計分析結果を図示しないモニタ等に表示する。なお、図13に示す統計分析結果の全体や一部等を、グラフ等で表示しても良いことはいうまでもない。

30

【0142】

以上、本実施の形態によれば、生徒が答案を記入した答案用紙を読み取った答案読取情報と、生徒が答案を記入する前の答案用紙の答案用紙情報との差分に関する差分関連情報を取得し、この差分関連情報と、この答案用紙に配置された識別画像から読み出した生徒識別情報とを有する答案情報を蓄積するようにしたので、生徒が記入した答案を適切に管理できる。

【0143】

また、差分関連情報を文字認識処理した情報を用いて様々な統計分析を行うことができ、生徒の答案を、生徒の成長過程の評価等の様々な用途に有効利用することができる。

40

【0144】

なお、上記各実施の形態において、各処理（各機能）は、単一の装置（システム）によって集中処理されることによって実現されてもよく、あるいは、複数の装置によって分散処理されることによって実現されてもよい。

【0145】

また、上記各実施の形態では、答案処理装置がスタンドアロンである場合について説明したが、答案処理装置は、スタンドアロンの装置であってもよく、サーバ・クライアントシステムにおけるサーバ装置であってもよい。後者の場合には、出力部や受付部は、通信回線を介して入力を受け付けたり、画面を出力したりすることになる。

【0146】

50

また、上記各実施の形態において、差分取得部 13、生徒識別情報取得部 14、文字認識部 15、答案情報蓄積部 17、統計分析部 18等の各構成要素は、専用のハードウェアにより構成されてもよく、あるいは、ソフトウェアにより構成され、プログラムを実行することによって実現されてもよい。例えば、ハードディスクや半導体メモリ等の記録媒体に記録されたソフトウェア・プログラムをCPU等のプログラム実行部が読み出して実行することによって、各構成要素が実現され得る。

【0147】

なお、上記各実施の形態における答案処理装置を実現するソフトウェアは、以下のようなプログラムである。つまり、このプログラムは、1以上の生徒が答案を記入した答案用紙であって、答案の記入を行った各生徒の識別情報である生徒識別情報と対応付けられた画像の情報である識別画像が配置された答案用紙を読み取った情報である答案読取情報が格納される答案読取情報格納部と、答案読取情報の読み取り元となる答案用紙の、答案が記入される前の状態を示す情報である答案用紙情報が格納される答案用紙情報格納部と、答案情報格納部と、にアクセス可能なコンピュータを、答案読取情報格納部に格納されている答案読取情報の、答案用紙情報に対する差分に関する情報である差分関連情報を取得する差分取得部と、差分取得部が差分を取得した答案読取情報の識別画像を検出し、識別画像と対応付けられた生徒識別情報を取得する生徒識別情報取得部と、差分取得部が取得した差分関連情報と、生徒識別情報取得部が取得した生徒識別情報とを有する答案情報を答案情報格納部に蓄積する答案情報蓄積部として機能させるためのプログラムである。

10

【0148】

なお、上記プログラムにおいて、上記プログラムが実現する機能には、ハードウェアでしか実現できない機能は含まれない。例えば、情報を取得する取得部や、情報を出力する出力部などにおけるモデムやインターフェースカードなどのハードウェアでしか実現できない機能は、上記プログラムが実現する機能には含まれない。

20

【0149】

また、このプログラムを実行するコンピュータは、単数であってもよく、複数であってもよい。すなわち、集中処理を行ってもよく、あるいは分散処理を行ってもよい。

【0150】

図14は、上記プログラムを実行して、上記実施の形態による答案処理装置を実現するコンピュータの外観の一例を示す模式図である。上記実施の形態は、コンピュータハードウェア及びその上で実行されるコンピュータプログラムによって実現されうる。

30

【0151】

図14において、コンピュータシステム900は、CD-ROM(Compact Disk Read Only Memory)ドライブ905、FD(Floppy(登録商標)Disk)ドライブ906を含むコンピュータ901と、キーボード902と、マウス903と、モニター904とを備える。

【0152】

図15は、コンピュータシステム900の内部構成を示す図である。図15において、コンピュータ901は、CD-ROMドライブ905、FDドライブ906に加えて、MPU(Micro Processing Unit)911と、ブートアッププログラム等のプログラムを記憶するためのROM912と、MPU911に接続され、アプリケーションプログラムの命令を一時的に記憶すると共に、一時記憶空間を提供するRAM(Random Access Memory)913と、アプリケーションプログラム、システムプログラム、及びデータを記憶するハードディスク914と、MPU911、ROM912等を相互に接続するバス915とを備える。なお、コンピュータ901は、LANへの接続を提供する図示しないネットワークカードを含んでいてもよい。

40

【0153】

コンピュータシステム900に、上記実施の形態による答案処理装置等の機能を実行させるプログラムは、CD-ROM921、またはFD922に記憶されて、CD-ROMドライブ905、またはFDドライブ906に挿入され、ハードディスク914に転送さ

50

れてもよい。これに代えて、そのプログラムは、図示しないネットワークを介してコンピュータ 901 に送信され、ハードディスク 914 に記憶されてもよい。プログラムは実行の際に RAM 913 にロードされる。なお、プログラムは、CD-ROM 921 や FD 922、またはネットワークから直接、ロードされてもよい。

【0154】

プログラムは、コンピュータ 901 に、上記実施の形態による答案処理装置の機能を実行させるオペレーティングシステム (OS)、またはサードパーティプログラム等を必ずしも含んでいなくてもよい。プログラムは、制御された態様で適切な機能 (モジュール) を呼び出し、所望の結果が得られるようにする命令の部分のみを含んでいてもよい。コンピュータシステム 900 がどのように動作するのかについては周知であり、詳細な説明は省略する。

10

【0155】

本発明は、以上の実施の形態に限定されることなく、種々の変更が可能であり、それらも本発明の範囲内に包含されるものであることは言うまでもない。

【産業上の利用可能性】

【0156】

以上のように、本発明にかかる答案処理装置等は、答案用紙を処理する装置等として適しており、特に、答案用紙から、生徒が答案を記入した答案用紙を処理する装置等として有用である。

【符号の説明】

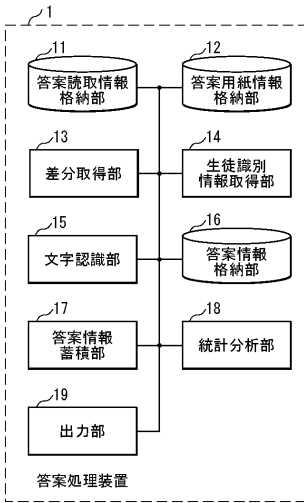
20

【0157】

- 1 答案処理装置
- 11 答案読取情報格納部
- 12 答案用紙情報格納部
- 13 差分取得部
- 14 生徒識別情報取得部
- 15 文字認識部
- 16 答案情報格納部
- 17 答案情報蓄積部
- 18 統計分析部
- 19 出力部
- 20 シール
- 21 識別画像
- 22 文字列
- 51 答案用紙情報
- 52 答案読取情報

30

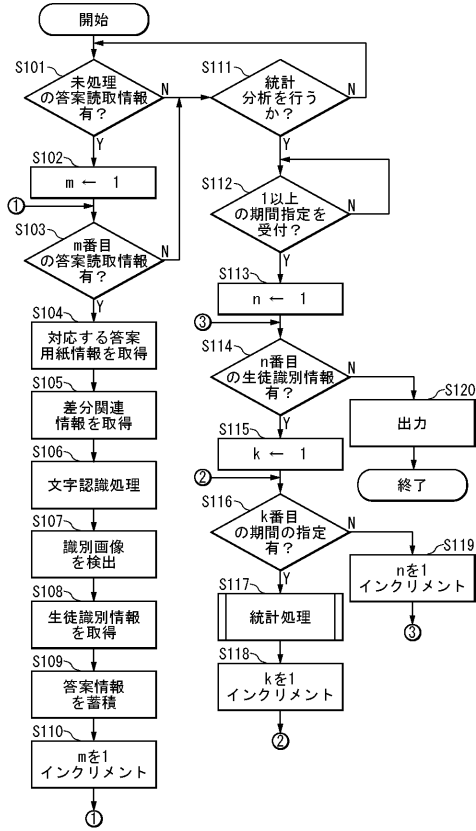
【図1】



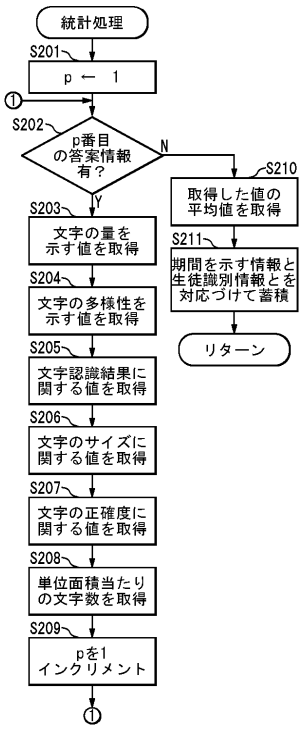
【図2】



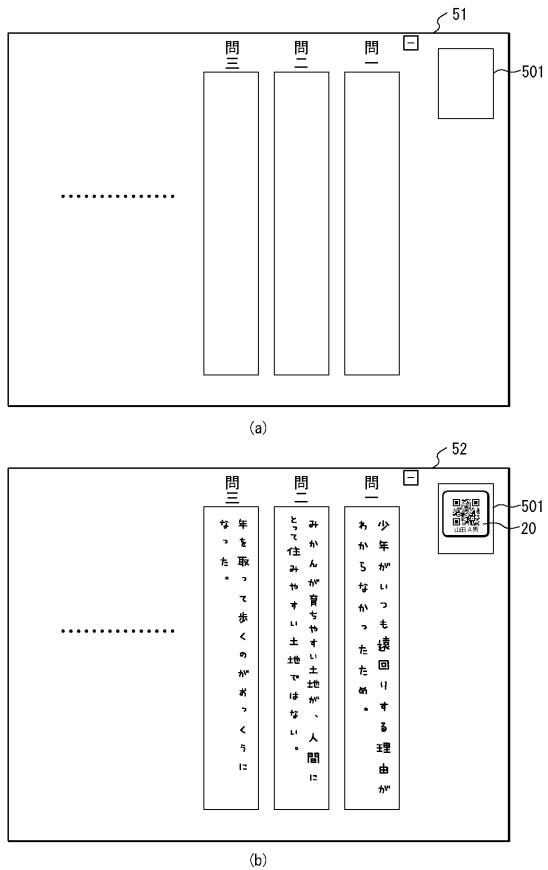
【図3】



【図4】



【図5】



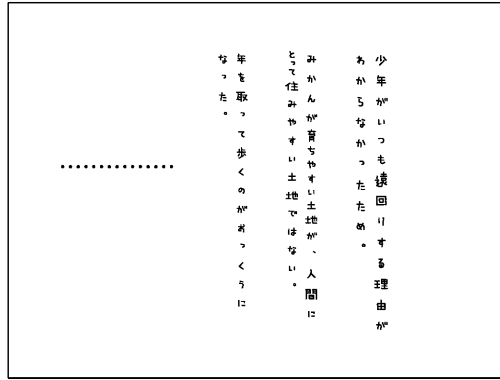
【 図 6 】

答案用紙情報	答案用紙ID
100545.tif	K2011-101
100546.tif	K2011-102
100547.tif	K2011-103
100548.tif	K2011-104
100549.tif	K2011-105
⋮	⋮

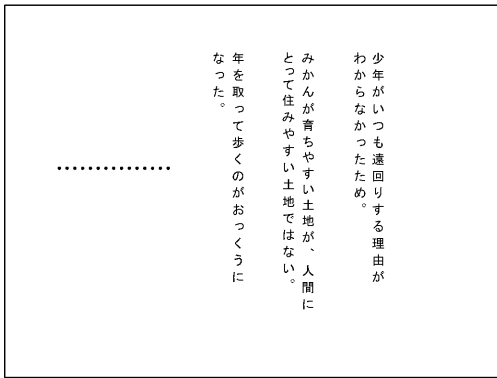
【 図 7 】

答案読取情報	答案用紙ID	日付	処理済
01000125.tif	K2011-101	2011/5/25	0
01000126.tif	K2011-101	2011/5/25	0
01000127.tif	K2011-101	2011/5/25	0
01000128.tif	K2011-101	2011/5/25	0
01000129.tif	K2011-101	2011/5/25	0
⋮	⋮	⋮	⋮
01000165.tif	K2011-102	2011/6/22	0
01000166.tif	K2011-102	2011/6/22	0
01000167.tif	K2011-102	2011/6/22	0
⋮	⋮	⋮	⋮

【 図 8 】



【 図 9 】



(a)

認識文字	スコア	サイズ(ポイント)
少	77	16
年	54	16
が	62	15
い	76	14
つ	61	15
も	88	16
遠	82	16
回	57	16
り	84	16
⋮	⋮	⋮

(b)

【 図 10 】

答案ID	答案読取情報	生徒ID	答案用紙ID	認識テキスト	認識管理系	日付情報
500001	500001.tif	山田A男	K2011-001	500001.tif	500001.csv	2011/5/25
500002	500002.tif	川上B夫	K2011-001	500002.tif	500002.csv	2011/5/25
500003	500003.tif	長嶋C太	K2011-001	500003.tif	500003.csv	2011/5/25
500004	500004.tif	水村D子	K2011-001	500004.tif	500004.csv	2011/5/25
500005	500005.tif	堀内E孝	K2011-001	500005.tif	500005.csv	2011/5/25
500006	500006.tif	佐野F助	K2011-001	500006.tif	500006.csv	2011/5/25
⋮	⋮	⋮	⋮	⋮	⋮	⋮
500041	500041.tif	山田A男	K2011-001	500041.tif	500041.csv	2011/6/22
500042	500042.tif	川上B夫	K2011-001	500042.tif	500042.csv	2011/6/22
500043	500043.tif	長嶋C太	K2011-001	500043.tif	500043.csv	2011/6/22
⋮	⋮	⋮	⋮	⋮	⋮	⋮
500081	500081.tif	山田A男	K2011-001	500081.tif	500081.csv	2011/7/14
500082	500082.tif	川上B夫	K2011-001	500082.tif	500082.csv	2011/7/14
500083	500083.tif	長嶋C太	K2011-001	500083.tif	500083.csv	2011/7/14

【 図 1 1 】

答案ID	文字数	単語数	平均スコア	平均サイズ	読字脱字数	単位文字数
500001	124	53	71	15.8	5	1.48
500041	85	42	78	15.7	4	1.56
500081	151	66	68	15.2	7	1.38
500121	142	68	70	15.2	6	1.42
...

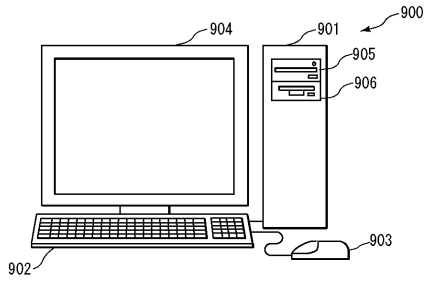
【 図 1 2 】

生徒ID	年	文字数	単語数	平均スコア	平均サイズ	読字脱字数	単位文字数
山田A男	2011	125	58	74	15.5	5.6	1.46
山田A男	2010	110	45	67	16.8	5.8	1.31
山田A男	2009	72	30	55	17.8	6.7	0.86
山田A男	2008	43	21	39	18.3	7.1	0.51
川上B女	2011	112	48	82	16.1	7.2	1.33
川上B女	2010	102	37	74	16.3	8.2	1.21
川上B女	2009	64	24	64	17.5	8.8	0.76
川上B女	2008	38	20	49	17.9	9.1	0.45
長島C女	2011	118	51	76	14.7	6.3	1.4
...

【 図 1 3 】

生徒ID	年	文字数	単語数	平均スコア	平均サイズ	読字脱字数	単位文字数
山田A男	2011	125	58	74	15.5	5.6	1.46
山田A男	2010	110	45	67	16.8	5.8	1.31
山田A男	2009	72	30	55	17.8	6.7	0.86
山田A男	2008	43	21	39	18.3	7.1	0.51
川上B女	2011	112	48	82	16.1	7.2	1.33
川上B女	2010	102	37	74	16.3	8.2	1.21
川上B女	2009	64	24	64	17.5	8.8	0.76
川上B女	2008	38	20	49	17.9	9.1	0.45
長島C女	2011	118	51	76	14.7	6.3	1.4
...

【 図 1 4 】



【図 15】

