

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4873138号
(P4873138)

(45) 発行日 平成24年2月8日(2012.2.8)

(24) 登録日 平成23年12月2日(2011.12.2)

(51) Int.Cl. F I
G06K 9/03 (2006.01) G06K 9/03 B
 G06K 9/03 J

請求項の数 11 (全 47 頁)

(21) 出願番号 特願2006-170877 (P2006-170877)
 (22) 出願日 平成18年6月21日(2006.6.21)
 (65) 公開番号 特開2008-3740 (P2008-3740A)
 (43) 公開日 平成20年1月10日(2008.1.10)
 審査請求日 平成21年2月10日(2009.2.10)

(73) 特許権者 000005496
 富士ゼロックス株式会社
 東京都港区赤坂九丁目7番3号
 (74) 代理人 110000039
 特許業務法人アイ・ピー・エス
 (72) 発明者 斎藤 照花
 神奈川県足柄上郡中井町境430 グリー
 ンテクなかい 富士ゼロックス株式会社内
 審査官 新井 則和

最終頁に続く

(54) 【発明の名称】 情報処理装置及びプログラム

(57) 【特許請求の範囲】

【請求項1】

原稿の画像データと、前記原稿に対して記入がなされた記入済み原稿の画像データとの差分を抽出する差分抽出手段と、

前記差分抽出手段によって抽出された差分から、予め定められた色成分の画像データを抽出する色成分抽出手段と、

前記色成分抽出手段によって抽出された画像データに含まれる線分のうち、途切れた線分同士が接続されるよう、前記色成分抽出手段によって抽出された画像データを補正する途切れ線分補正手段と、

前記途切れ線分補正手段によって補正された画像データに含まれる図形の形状を認識する図形形状認識手段と、

前記図形形状認識手段によって認識された図形が記入された位置を認識する記入位置認識手段と、

前記図形形状認識手段によって形状が認識され、前記記入位置認識手段によって記入位置が認識された図形それぞれについて、前記差分抽出手段における抽出誤りの信頼度に関する特徴量、前記色成分抽出手段における抽出誤りの信頼度に関する特徴量、前記途切れ線補正手段における補正誤りの信頼度に関する特徴量、前記図形形状認識手段における認識誤りの信頼度に関する特徴量及び前記記入位置認識手段における認識誤りの信頼度に関する特徴量のうち少なくとも複数の特徴量が予め定められた基準を満たすか否かを示すサブ信頼度を算出し、算出されたサブ信頼度が予め定められた水準よりも低い図形を特定す

10

20

る難認識図形特定手段と、

前記難認識図形特定手段によって特定された図形が修正を要することを表示する表示手段と

を備える情報処理装置。

【請求項 2】

前記難認識図形特定手段は、算出されたサブ信頼度のうち最も小さいサブ信頼度を最終的な信頼度として特定し、特定された最終的な信頼度が予め定められた水準よりも低い図形を特定する

請求項 1 に記載の情報処理装置。

【請求項 3】

前記難認識図形特定手段は、算出されたサブ信頼度に重み付け値を乗じて合計した値のうち最も小さい値を最終的な信頼度として特定し、特定された最終的な信頼度が予め定められた水準よりも低い図形を特定する

請求項 1 に記載の情報処理装置。

【請求項 4】

前記難認識図形特定手段は、算出されたサブ信頼度に重み付け値を乗じて合計した値が最も小さい値又は予め定められた水準よりも低い値を算出する基となった特徴量に関する誤りの種別を特定する

請求項 1 に記載の情報処理装置。

【請求項 5】

前記表示手段は、前記難認識図形特定手段において特定された最終的な信頼度又は誤りの種別に応じて、表示方法を変える

請求項 1 乃至 4 のいずれかに記載の情報処理装置。

【請求項 6】

前記表示手段は、前記図形形状認識手段によって形状が認識され、前記記入位置認識手段によって記入位置が認識された図形のうち、前記難認識図形特定手段において特定された図形と、当該図形以外の図形とが区別されるように表示する

請求項 1 乃至 4 のいずれかに記載の情報処理装置。

【請求項 7】

前記表示手段は、前記難認識図形特定手段によって特定された図形が修正を要することを表示するにあたり、当該図形の最終的な信頼度及び当該図形を修正すべき理由を表示する

請求項 1 乃至 4 のいずれかに記載の情報処理装置。

【請求項 8】

前記図形形状認識手段によって形状が認識され、前記記入位置認識手段によって記入位置が認識された図形それぞれに基づいて採点集計処理を実行するデータ処理手段

をさらに有する請求項 1 乃至 7 のいずれかに記載の情報処理装置。

【請求項 9】

前記難認識図形特定手段によって特定された図形の修正を受け付ける受付手段

をさらに有し、

前記データ処理手段は、前記受付手段によって受け付けられた修正後の図形に基づいて自動採点処理を実行する

請求項 8 に記載の情報処理装置。

【請求項 10】

前記差分抽出手段における抽出誤りの信頼度に関する特徴量は、前記差分抽出手段による抽出結果の品質並びに前記原稿の画像データと前記原稿に対して記入がなされた記入済み原稿の画像データとが重複する量及び接触する量であり、

前記色成分抽出手段における抽出誤りの信頼度に関する特徴量は、前記色成分抽出手段による抽出結果の品質並びに前記原稿の画像データと前記色成分抽出手段によって抽出される画像データとが重複する量及び接触する量であり、

10

20

30

40

50

前記途切れ線補正手段における補正誤りの信頼度に関する特徴量は、前記途切れ線補正手段によって接続される途切れた線分の端点及び当該端点間の距離であり、

前記図形形状認識手段における認識誤りの信頼度に関する特徴量は、前記途切れ線分補正手段によって補正された画像データにおいて図形同士が接触するか否か及び前記図形形状認識手段による図形形状認識結果の信頼度であり、

前記記入位置認識手段における認識誤りの信頼度に関する特徴量は、前記図形形状認識手段によって認識された図形が記入された位置が複数認識される可能性及び前記原稿に設けられた記入欄と前記図形形状認識手段によって認識された図形との距離である

請求項 1 乃至 9 のいずれかに記載の情報処理装置。

【請求項 1 1】

原稿の画像データと、前記原稿に対して記入がなされた記入済み原稿の画像データとの差分を抽出する差分抽出手段、

前記差分抽出手段によって抽出された差分から、予め定められた色成分の画像データを抽出する色成分抽出手段、

前記色成分抽出手段によって抽出された画像データに含まれる線分のうち、途切れた線分同士が接続されるよう、前記色成分抽出手段によって抽出された画像データを補正する途切れ線分補正手段、

前記途切れ線分補正手段によって補正された画像データに含まれる図形の形状を認識する図形形状認識手段、

前記図形形状認識手段によって認識された図形が記入された位置を認識する記入位置認識手段、

前記図形形状認識手段によって形状が認識され、前記記入位置認識手段によって記入位置が認識された図形それぞれについて、前記差分抽出手段における抽出誤りの信頼度に関する特徴量、前記色成分抽出手段における抽出誤りの信頼度に関する特徴量、前記途切れ線補正手段における補正誤りの信頼度に関する特徴量、前記図形形状認識手段における認識誤りの信頼度に関する特徴量及び前記記入位置認識手段における認識誤りの信頼度に関する特徴量のうち少なくとも複数の特徴量が予め定められた基準を満たすか否かを示すサブ信頼度を算出し、算出されたサブ信頼度が予め定められた水準よりも低い図形を特定する難認識図形特定手段、及び、

前記難認識図形特定手段によって特定された図形が修正を要することを表示する表示手段

としてコンピュータを機能させるプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、入力修正方法、並びにこの入力修正方法を適用した追記情報処理方法、追記情報処理装置、およびプログラムに関する。より詳細には、たとえば、文書に手書きで追加記入された付加情報（追記情報やアノテーションとも称する）を、文書本体から分離し、文書本体と関わりを持つ様々な情報処理に使用する際に利用される、入力された付加情報に対する修正の仕組みに関する。

【背景技術】

【0002】

情報処理技術の進歩の著しい今日、文書に関する様々な処理を自動処理する仕組みがある。たとえば、予め所定の情報が記載されている文書原本にさらに別の付加情報を手書きで追記し、その付加情報の追記された追記済文書を処理対象として、手書きで追記された付加情報に基づいて所定のデータ処理を自動的に実行する仕組みがある。

【0003】

たとえば、定型伝票（いわゆる帳票）に手書きで情報を記入して、その記入された情報を処理対象とする自動帳票処理の仕組み（たとえば特許文献 1, 2 を参照）や、日付や予定記入欄が用意された主に手帳やメモなどに手書きで予定を記入し、その記入された予定

10

20

30

40

50

を電子データ化する個人情報管理（特にスケジュール管理ともいう）の仕組み（特許文献3を参照）もある。

【0004】

また、学校や学習塾などの教育現場においては、生徒や受験者による解答が記入された答案用紙に採点官が採点した結果の用紙（教育用教材）を処理対象として、自動採点集計処理を行なう仕組みも考えられている（特許文献4を参照）もある。

【0005】

【特許文献1】特開平5 - 342239号公報

【特許文献2】特開平6 - 274156号公報

【特許文献3】特開平5 - 216932号公報

【特許文献4】特開平11 - 31046号公報

10

【0006】

たとえば、特許文献1には、表の種別および表中の各データの行方向の位置並びに列方向の位置を指定するための識別マークを付与して表データを識別マーク付きデータ表帳票として印刷するようにし、データ表帳票に付されている認識マークをマーク読取手段によって読み取らせることで、識別データから帳票を特定して該当する表データを表示手段に表示するようにし、かつ該当する箇所を修正可能な状態に表示することで、少ない作業量で効率よくデータ修正および照合を行なうことができるようにした仕組みが提案されている。

【0007】

20

また、特許文献2には、出力原稿をイメージデータに変換し、変換されたイメージデータを表示し、表示されたイメージ上にフォーム規定位置座標を指示し、指示されたフォーム規定位置座標に基づいてフォーム図形データを導出し、導出されたフォーム図形データに基づいてフォームを表示部に表示されているイメージデータに重ね合わせ表示する構成とすることで、所望の出力原稿に対するフォームを容易に作成できるようにした仕組みが提案されている。

【0008】

また、特許文献3には、個人情報管理用の仕組みとして、書き込みの施された手帳などの紙面をスキャナより入力し、既知の方法により文字や枠線を検出し、検出した文字や枠線の配置を予め記憶された手帳の書式の知識と照合し、各記載項目の属性を表すデータを生成する仕組みが提案されている。

30

【0009】

手帳やメモなどに手書きで記入した情報（本例では手帳に手書きで入力した予定の記載内容）が自動的に電子データに変換され、各記載項目を生成された属性に対応付けて切り換えてデータベースに登録することができるので、利用者がある書式の出力を要求した場合には、予め記憶された出力の書式を参照することにより、データベースの内容を紙上の各内容に対応した場所に印刷するなどができ便利である。

【0010】

また、特許文献4には、答案用紙の手書き記入欄をタブレット上の入力域として定義するとともに、集計対象の入力域を識別可能に定義する書式データを記憶装置に記憶しておき、タブレットに答案用紙が重ね合せられた状態で、手書き記入欄に記入された情報をタブレットの入力域に入力された筆記データとして取り込んで文字認識することにより、採点などの記入作業を行なった後のキー入力などによるデータ入力を省力化する仕組みが提案されている。

40

【0011】

一方、手書きで入力された付加情報に基づいてデータ処理を実行するに当たっては、入力された手書き情報（文字や図形）を如何様にして認識するかや、自動認識結果が信頼できない場合に自動認識処理した認識結果に対して如何様にして効率的に修正するかが問題となり、このような手書き情報の自動認識・修正の仕組みとしては、様々なものが考えられている（たとえば特許文献5を参照）。

50

【0012】

【特許文献5】特開2004-152115号公報

【0013】

ここで、特許文献5に記載の仕組みでは、取得したデータに含まれる複数の項目について項目ごとに内容の確信度を計算し、計算された確信度を用いて提示の方法を動的に変化させるようにしている。入力されたデータに含まれる各項目の確信度を用いて、データの修正方法を動的に変化させることで、郵便番号から住所を修正するなど関連の高い項目の内容を修正したり、常に疑わしい、つまり確信度の低い項目の内容を優先的に修正したりすることができるようにし、データの入力方法に因らずオペレータが介在する入力修正作業を効率よく迅速かつ正確に行なうことができるようにしている。

10

【発明の開示】

【発明が解決しようとする課題】

【0014】

しかしながら、特許文献5に記載の仕組みでは、各項目の認識時にパターン認識の類似度を使用して「確信度」を算出するのみであるため、認識処理よりも前段の各種の前処理などで生じ得る、認識処理に影響を与える事象を考慮することができない。そのため、たとえば、局所的なノイズの有無が「確信度」に影響を与えてしまい、「確信度」の精度が問題となる。前処理の段階で問題があると、認識処理への入力が既に変質しており、「信頼度が高いが間違い」の認識結果を出す可能性があるからである。

【0015】

20

本発明は、上記事情に鑑みてなされたものであり、自動データ処理に供される追記情報について、修正作業を効率よく迅速かつ正確に行なうことができるとともに、高精度に修正を要する追記情報を特定することのできる仕組みを提供することを目的とする。

【課題を解決するための手段】

【0016】

本発明に係る仕組みにおいては、注目する付加情報に関して、認識処理の信頼度に関する特徴量を取得し、各特徴量に基づいて各処理の認識処理に関わるサブ信頼度をそれぞれ算出し、この算出したサブ信頼度に基づいて注目する付加情報に関しての最終的な信頼度を特定し、この最終的な信頼度が一定水準よりも低いかなかを判定することで認識処理の信頼度が一定水準よりも低い難認識情報を特定し、この特定した難認識情報の認識性能を向上させるための認識性能情報を提示するようにした。

30

【0017】

その後には、認識性能情報の提示に対応して記入された修正後の付加情報をデータ処理に反映させる。その修正の反映方法としては、たとえば、認識後に保存されたデータを直接変更することで実現する手法を採ることができる。たとえば、修正結果は、認識後にデータベース操作などで直接データを変更することで反映するとよい。

【0018】

ここで、「認識処理の信頼度に関する特徴量」は、認識処理そのものにおける信頼度に関する特徴量だけでなく、この認識処理よりも前段の認識処理に関わる各種の処理における信頼度に関する特徴量も含むものである。

40

【0019】

認識処理そのものだけでなく、認識処理よりも前段の各種の前処理などで生じ得る、認識処理に影響を与える事象を考慮して、認識処理の信頼度が一定水準よりも低い難認識情報を特定するのである。

【0020】

なお、本発明に係る仕組みは、電子計算機（コンピュータ）を用いてソフトウェアで実現することもでき、このためのプログラムやこのプログラムを格納した記録媒体を発明として抽出することも可能である。プログラムは、コンピュータ読取り可能な記憶媒体に格納されて提供されてもよいし、有線あるいは無線による通信手段を介した配信により提供されてもよい。

50

【発明の効果】

【0021】

本発明によれば、認識処理そのものにおける信頼度に関する特徴量だけでなく、この認識処理よりも前段の認識処理に関わる各種の処理における信頼度に関する特徴量も取得し、認識処理の信頼度が一定水準よりも低い難認識情報を特定するようにした。

【0022】

これにより、認識処理よりも前段の各種の前処理などで生じ得る認識処理に影響を与える事象を考慮することができるようになるので、全ての追記情報をチェックする必要がなく、誤認識を起し得る追記情報の修正作業を効率よく迅速かつ正確に行なうことができるだけでなく、高精度に難認識情報を特定することができる。修正対象の追記情報を高精度に特定して、効率よく迅速かつ正確に修正できるようになる。

10

【発明を実施するための最良の形態】

【0023】

以下、図面を参照して本発明の実施形態について詳細に説明する。

【0024】

< 処理対象文書の例 >

図1は、本発明に係る追記情報処理装置を備えてなる情報処理システムにおいて処理対象とする文書の一例を示す図である。

【0025】

図1に示す文書は、教育現場で紙媒体として使用される教育用教材80であって、典型例として、問題文82およびその解答欄84（本例では括弧で示された部分）を有したものであり、図1（A）はその文書原本8Aを示し、図1（B）は、追記情報記入済の状態（追記済文書8B）を模式的に示している。また、データベースへの登録情報の一例を（C）に示す。

20

【0026】

このような教育用教材80は、具体的には教育機関で用いられるペーパーテストや練習問題シートなどが該当する。なお、教育用教材80は、少なくとも解答欄84を有していればよく、たとえば採点官が読み上げた問題について解答欄84に解答を記入する場合などがあり、問題文82については必ずしも記載されていなくともよい。

【0027】

また、図1（A）に示すように、教育用教材80は、問題文82および第1種の付加情報の一例である解答が記入される解答欄84の他に、配点欄83（問題別の項目点欄83aや部分点欄や合計点欄でなる集計欄83b）と、第1種の付加情報の一例である教育用教材80を識別特定するための情報を記入する識別情報欄85と、第1種の付加情報の一例である解答者情報であって解答欄84への解答記入者に関する情報を記入する解答者情報欄86とを有している。解答欄84、識別情報欄85、解答者情報欄86は、何れも第1種の付加情報を記入する記入欄の一例である。

30

【0028】

識別情報欄85には、たとえば教育用教材80の科目、タイトル、あるいは適用学年などが予め記載されるものとする。ただし、これらの記載に加えて、またはこれらの記載とは別に、教育用教材80を識別するためのコード情報が埋め込まれていてもよい。

40

【0029】

コード情報の埋込みは、公知技術を利用して実現すればよいが、その一つの実例として、たとえば「i T o n e（登録商標）」と呼ばれるもののよう、階調表現としての万線スクリーンまたはドットスクリーンを構成する画素の形態（位置、形状など）を変化させることで、ハーフトーン画像の中にデジタル情報を埋め込むようにする、といった技術を用いることが考えられる。一方、解答者情報欄86には、解答記入者の学級86a、出席番号86b、あるいは氏名86cなどが記入され得ようになっている。

【0030】

配点欄83（特に項目点欄83a）には、各解答欄84についての配点情報が記入され

50

る。配点情報とは、教育用教材 80 における各解答欄 84 について、各位置の解答欄 84 への配点が何点であるかを特定するための情報である。なお、配点は、解答欄 84 ごとに異なっているもよいし、あるいは一律であってもよい。

【0031】

このような教育用教材 80 は、対応する原本（教材原本）の電子データに基づき印刷装置で印刷出力することで得ることができる。教材原本の電子データは、たとえばパーソナルコンピュータなどの電子計算装置を利用してワープロソフトなどのアプリケーションソフトウェアを用いて生成することができ、予め所定のデータベースなどに保存される。

【0032】

なお、教材原本の電子データは、その教育用教材 80 における解答欄 84 や識別情報欄 85 などのレイアウトを特定し得るものであり、かつ、所定のデータベースにて保持蓄積可能なものであれば、そのデータ形式を問わない。たとえば、文書作成ソフトウェアで作成したアプリケーション文書データに限らず、画像データであってもよい。

10

【0033】

教育現場では、図 1 (A) に示した文書原本 8A の一例である教育用教材 80 が生徒や受験者などに配布され、先ず、生徒などによって解答者情報欄 86 への氏名や解答欄 84 への解答などの第 1 種の付加情報が所定欄に記入された後に回収される。この後さらに、図 1 (B) に示す追記済文書 8B の一例である付加情報記入済教材 81 のように、教師などの採点官によって各解答欄 84 に記入された解答に対する採点記号（正誤判定図形）87 や採点記号 87 と関連するその他の図形や文章で示されたコメント 88 などの第 2 種の付加情報が記入される。なお、追記情報処理装置 10 で自動採点処理を行なうので、この時点では、採点官による配点欄 83C への記入はない。

20

【0034】

この際には、通常、生徒などによる第 1 種の付加情報の記入に使用されるペン色と、採点官などによる第 2 種の付加情報の記入に使用されるペン色とは、異なるものが使用されるし、教育用教材 80 として予め記入されている色とも異なるものが使用される。

【0035】

また、採点記号 87 としては、たとえば、正解を示す「」やその他の図形（たとえば楕円図形）、不正解を示す「x」やその他の図形（たとえば「レ点」などのチェックマーク）、あるいは一部正解を示す「」やその他の図形がある。

30

【0036】

コメント 88 は、採点記号 87 を元にした第 1 のデータ処理には直接的な関係を有しない情報であるが、第 1 のデータ処理の結果をサポート（補強）するあるいは全く関係のない第 2 のデータ処理に利用されるものである。

【0037】

教材自動採点システムでは、この追記済文書 8B の一例である付加情報記入済教材 81 を対象として所定のデータ処理を行なうことになる。この際、第 2 種の付加情報（本例では採点記号 87 とコメント 88）の別に、それぞれに応じた個別のデータ処理を行なうようにする。この例では、第 1 のデータ処理として採点記号 87 を元にした自動採点処理を行なう。この自動採点処理結果として、たとえば、配点欄 83C へ記入する得点情報が取得されることになる。また、コメント 88 の追記内容を元にした自動採点処理結果とは別の第 2 のデータ処理を行なう。

40

【0038】

自動採点処理とは別の第 2 のデータ処理としては、たとえば、コメント 88 の追記内容を所定の条件に基づいて分類する処理や、分類したコメント（全ての追記内容には限らず一部の追記内容でもよい）を対応する解答欄 84 についての採点処理結果と関連付けて生徒指導用データベースに登録・蓄積しておく処理などを行なう。

【0039】

こうすることで、自動採点処理とは別の第 2 のデータ処理の利用形態としては、生徒指導用データベースから情報を取り出して、後の生徒指導に利用することができるようにな

50

る。たとえば、全ての採点結果と対応するコメントを表示するようにすれば、優秀、優、並、劣、などの評価の細分類と合わせて採点結果を確認することができる。また、結果が不正解のコメントのみを表示するようにすれば、問題の読み間違いが多い、解答の書き間違いが多い、あるいは計算ミスが多いなどの生徒の能力分析に利用することもできる。あるいは、コメント 88 で示されている注意文やミス修正に基づき生徒指導に利用することも考えられる。

【 0 0 4 0 】

なお、採点記号 87 やコメント 88 を対象としたデータ処理を実行するための教育用教材 80 (原本画像) のデータベースへの登録に当たっては、通常であれば、採点記号 87 についてのデータ処理時に必要となる解答欄 84 の位置情報や問題番号や配点情報などを示す記入欄位置領域情報 38 と、コメント 88 についてのデータ処理時に必要となる分類基準情報とを登録しておく。なお、教育用教材 80 における文字部分 (たとえば問題番号を含む問題文や配点) をテキスト情報として参照され得るように、教育用教材 80 そのものの情報は、テキストデータもしくはテキストデータ相応の文書ファイルデータで保存しておくことが好ましい。

10

【 0 0 4 1 】

たとえば、図 1 (C) に示すように、記入欄位置領域情報 38 の一例である解答欄位置領域情報を、教育用教材 80 上に存在する問題の番号 (第 1 の属性情報の一例) と、その問題の解答に対する配点 (第 2 の属性情報の一例) と、その問題の解答を記入する解答欄 84 として扱われる領域の所定点 (たとえば左上頂点) の $x y$ 座標、並びに、その外接矩形の幅 (W) および高さ (h) とからなる情報で、これらを互いに関連付けるテーブル形式で、所定の格納領域に保持蓄積しておく。

20

【 0 0 4 2 】

付加情報記入済教材 81 に対応する元の教育用教材 80 が文書管理サーバに登録されていないときには、無記入の教育用教材 80 を文書入力装置で読み取り、問題文の位置や解答欄 84 の位置や配点などを特定することにより対処する。

【 0 0 4 3 】

なお、追記情報について自動データ処理を実行する際に、追記情報を複数種類のものに分類し、それぞれの分類ごとに個別のデータ処理を実行しようとする場合には、各追記情報を他方のものと分離して認識・特定することが必要となる。両者の分離認識が適正になされないと、それぞれのデータ処理を適正に実行することができなくなる。

30

【 0 0 4 4 】

たとえば、図 1 に示した付加情報記入済教材 81 の場合には、採点記号 87 に基づく自動採点処理と、コメント 88 に基づく生徒指導用データベースの構築処理があり、このような個別処理を実現するには、採点記号 87 とコメント 88 とを区別して認識処理などを行ってから最終的なデータ処理を行なう必要がある。

【 0 0 4 5 】

一方、答案の採点においては、採点記号 87 以外のコメント 88 を、採点記号 87 と同じペンで記載することがあり、たとえば差分抽出部 132 での抽出結果に対する色成分認識処理を通じて、ペン色と対応する所定色成分についてのものを抽出するだけでは、両者を適切に分離した認識と記載内容の特定ができない。採点記号 87 とコメント 88 とを適切に分離できず、自動採点処理に悪影響を及ぼす。また、追記情報としては、自動採点処理用の採点記号 87 の他に、生徒指導などにも利用し得るコメント 88 が存在するにも関わらず、付加情報記入済教材 81 を生徒に返却した後は、生徒がその記載内容を確認する以外には活用できない事態となる。

40

【 0 0 4 6 】

このような事態を避けるには、たとえば、追記済文書 8B 中に存在する複数の付加情報を、記載位置、画像特徴量、あるいは認識処理時の信頼度などの付加情報が持つ様々な特徴に基づいて、複数の種類に分類 (分離) するようにするのがよい。この際、付加情報が持つ単一の特徴に基づくだけでなく複数の特徴を参照することで、より正確な分離認識を

50

行なうようにするのがよい。なお、これらについては詳細な説明を割愛する。

【 0 0 4 7 】

<システム構成>

図 2 は、本発明に係る追記情報処理装置を備えてなる情報処理システムの一実施形態の構成例を示す図である。なお、この情報処理システムは、答案用紙などの教育用教材 8 0 を処理対象とする教材自動採点システムへの適用例で示す。

【 0 0 4 8 】

図示のように、教材自動採点システム 1 は、システムを中心をなす追記情報処理装置 1 0 と、追記情報処理装置 1 0 に処理対象文書である付加情報記入済教材 8 1 を電子化して入力する文書入力装置 2 0 と、処理対象文書である付加情報記入済教材 8 1 に対応するテンプレート 6 や教育用教材 8 0 (詳しくはその原本画像)の電子データを記憶する文書管理サーバ 3 0 と、情報処理(本例では自動採点処理など)の結果を保存しておく処理結果保存サーバ 4 0 とが、有線や無線を利用してネットワーク接続されて構成されている。

10

【 0 0 4 9 】

文書入力装置 2 0 は、教育用教材 8 0 における解答欄 8 4 への解答記入、解答者情報欄 8 6 への氏名などの記入および解答欄 8 4 に記入された解答に対する採点官による採点記号 8 7 (具体的には、たとえば「」や「」の図形)の記入がされた付加情報記入済教材 8 1 に対して、公知の光学的画像読取技術を用いた画像読取りを行ない、その付加情報記入済教材 8 1 から画像データを得るものである。

【 0 0 5 0 】

文書入力装置 2 0 は、処理対象文書である付加情報記入済教材 8 1 を電子データにする機能を備えたものであればよく、たとえば、画像読取装置としての機能を有した複写機、複合機、またはスキャナ装置を利用して実現することが考えられる。その場合に、自動原稿搬送装置(Automatic Document Feeder ; ADF)が付設されていると、複数の教育用教材に対する画像読取りを連続的に行なうことができ便利である。

20

【 0 0 5 1 】

なお、教育用教材 8 0 を利用した試験などは、紙媒体を用いることに限定されない。たとえばタブレット型の PC を利用して試験を行ない採点するケースでは、最初から付加情報記入済教材 8 1 を電子データの形式で入手することができ、この場合には、システム構成上、文書入力装置 2 0 が不要となる。

30

【 0 0 5 2 】

文書管理サーバ 3 0 は、付加情報記入済教材 8 1 に対応する教育用教材 8 0 の原本画像と、この原本画像を特定するための、たとえば、科目、タイトル、適用学年などの識別情報や識別コードとを対応付けて、ハードディスク装置や光ディスク装置などの所定の記憶媒体に文書原本情報データベース DB 2 として登録している。

【 0 0 5 3 】

また、文書管理サーバ 3 0 は、付加情報記入済教材 8 1 に対応する元の教育用教材 8 0 (原本画像)を保存するとともに、採点記号 8 7 やコメント 8 8 についてのデータ処理時に必要となる問題番号や配点情報などを示す記入欄位置領域情報 3 8 を、図 1 (C) に示したように、テーブル形式で、所定の格納領域に文書原本情報データベース DB 2 として保持蓄積している。

40

【 0 0 5 4 】

処理結果保存サーバ 4 0 としては、追記情報処理装置 1 0 とネットワーク接続され、付加情報記入済教材 8 1 についての自動採点集計結果を管理することができるものであればよく、たとえば、処理結果データベース装置や処理結果ファイルサーバ装置などが該当する。

【 0 0 5 5 】

教材自動採点システム 1 の中心部をなす追記情報処理装置 1 0 は、文書入力装置 2 0 から入力された付加情報記入済教材 8 1 の画像データに基づき所定の信号処理を行なう読取画像処理部 1 1 0 と、読取画像処理部 1 1 0 による処理に基づいて文書入力装置 2 0 から

50

入力された読取画像の元となった文書原本 8 A を識別特定する文書原本特定部 1 2 0 とを備えている。教材自動採点システム 1 においては、文書原本特定部 1 2 0 は、文書原本 8 A の一例である教育用教材 8 0 を特定する教材特定部 1 2 2 として機能する。

【 0 0 5 6 】

また、追記情報処理装置 1 0 は、読取画像処理部 1 1 0 による処理が施された画像データから追記情報（アノテーション）を抽出する追記情報抽出部 1 3 0 と、追記情報抽出部 1 3 0 により抽出されたデータ処理対象の追記情報に基づき記載内容や記入位置を認識し特定するデータ処理対象追記情報特定処理部 1 5 0 と、データ処理対象追記情報特定処理部 1 5 0 により特定された追記情報の記載内容に基づいてデータ処理を行なうデータ処理部 1 7 0 とを備えている。

10

【 0 0 5 7 】

また、追記情報処理装置 1 0 は、本実施形態特有の構成要素として、手書きで記入された文字や図形などの手書き入力情報の認識性能を示す情報（具体的には認識率）に基づいて特定される認識処理の信頼度が一定水準よりも低い難認識情報について、その難認識情報の認識性能を向上させるための認識性能情報をユーザに提示する認識性能情報提示処理部 1 9 0 を備えている。

【 0 0 5 8 】

認識性能情報提示処理部 1 9 0 は、追記入力された採点記号 8 7 やコメント 8 8 の内、認識処理の信頼度が一定水準よりも低い難認識情報を特定し、この特定した難認識情報の認識性能を向上させるための認識性能情報をユーザ端末 1 7 1 にて提示する。データ処理部 1 7 0 は、認識性能情報提示処理部 1 9 0 による認識性能情報の提示に対応してユーザにより記入された修正後の採点記号 8 7 やコメント 8 8 を用いて自動採点処理や自動コメント分類処理などのデータ処理を実行する。

20

【 0 0 5 9 】

ここで、本実施形態の認識性能情報提示処理部 1 9 0 は、認識処理対象の追記情報の効率的な修正を可能とするべく、追記情報処理装置 1 0 の各機能部における複数の処理から、それぞれ認識率に関わる情報を収集し、記入内容を自動認識した追記情報の中で、特に信頼度の低いもの、つまり一定の度合いよりも認識の困難であった難認識情報を抽出して、この難認識情報についての認識性能情報を提示して修正を促す点に特徴を有する。全ての追記情報について、修正の要否をチェックする必要を無くすことで、誤認識の追記情報の修正作業を効率化するのである。

30

【 0 0 6 0 】

たとえば、「全図形をチェックする必要が無い」という効果を得るためには、信頼度による警告が高精度であることが必要になる。何故なら、抜け漏れがあっては警告が信用できないからである。前処理の段階で問題があると、認識処理への入力が既に変質しており、「信頼度が高いが間違い」の認識結果を出す可能性がある。

【 0 0 6 1 】

そこで、認識処理だけでなく複数の処理から情報を収集することで、高精度な信頼度を算出し、そこから高精度な警告を出すことで、はじめて「全図形をチェックする必要がない効率化」の効果が得られるようになるのである。なお、この点についての詳細は後述する。

40

【 0 0 6 2 】

読取画像処理部 1 1 0 は、図示を割愛するが、文書入力装置 2 0 から入力された画像データについて、レイアウト解析、文字図形分離、文字認識、コード情報認識、図形処理、色成分認識などの公知の画像処理技術（それぞれの詳細説明は割愛する）を利用して解析処理を行なう画像データ解析部と、文書入力装置 2 0 から入力された画像データの傾きや主走査方向または副走査方向の拡縮率などの画像歪みを補正する歪み補正部とを有している。なお、歪み補正部は、文書入力装置 2 0 から入力された画像データと、比較対象となる文書管理サーバ 3 0 内の対応する原本画像とを比較照合し、その画像歪み（傾き、拡縮など）を補正してもよい。

50

【 0 0 6 3 】

教材特定部 1 2 2 は、図示を割愛するが、たとえば、画像データ解析部によるデータ解析結果に基づいて、識別情報欄 8 5 に記入されている科目、タイトル、あるいは適用学年などの識別情報を解析する識別情報解析部と、同じく識別情報欄 8 5 に埋め込まれている教育用教材 8 0 を特定するコード情報を解析するコード情報解析部とを有している。

【 0 0 6 4 】

教材特定部 1 2 2 は、画像データ解析部での解析結果に基づいて特定した、たとえば科目、タイトル、適用学年などの識別情報や識別コードと、文書管理サーバ 3 0 に保持蓄積されている教育用教材 8 0 の原本画像の情報（たとえば科目、タイトル、適用学年などの識別情報や識別コード）とを照らし合わせ、該当する原本画像が文書管理サーバ 3 0 に保持蓄積されていない場合は、文書入力装置 2 0 で得られた画像データとの比較対象となる電子データを特定できないと判定して、識別特定エラー信号を出力するようになっている。

10

【 0 0 6 5 】

なお、教材特定部 1 2 2 は、文書入力装置 2 0 から入力された画像データ（付加情報記入済教材 8 1 に相当）と対応する元の教育用教材 8 0 を識別特定できればよく、識別情報解析部とコード情報解析部とは、付加情報記入済教材 8 1 の識別情報欄 8 5 に記載もしくは埋め込まれている識別情報の形式に対応する適正な方を備えていればよく、必ずしも、両者を備えている必要はない。

【 0 0 6 6 】

追記情報抽出部 1 3 0 は、歪み補正部にて画像歪みが補正された画像データと、教材特定部 1 2 2 により特定された、文書入力装置 2 0 から入力された画像データ（付加情報記入済教材 8 1 に相当）に対応する原本画像（教育用教材 8 0 に相当）とを公知の画像処理技術を利用して比較しそれぞれの間の差分を抽出する差分抽出部 1 3 2 を有する。

20

【 0 0 6 7 】

また、追記情報抽出部 1 3 0 は、差分抽出部 1 3 2 による抽出結果に基づき文書入力装置 2 0 で読取り対象となった付加情報記入済教材 8 1 における解答者情報（第 1 種の付加情報の一例）を抽出する解答者抽出部 1 3 4 と、同じく差分抽出部 1 3 2 による抽出結果に基づき文書入力装置 2 0 で読取り対象となった付加情報記入済教材 8 1 における採点記号 8 7 やコメント 8 8 などのデータ処理に供される第 2 種の付加情報を抽出するデータ処理用追記情報抽出処理部 1 4 0 とを有する。

30

【 0 0 6 8 】

解答者抽出部 1 3 4 は、差分抽出部 1 3 2 による抽出結果に基づき、解答者情報欄 8 6 の学級 8 8 a や出席番号 8 8 b や氏名 8 8 c の欄に記入された解答記入者の手書きによる番号や文字の画像をそのまま文字情報として切り出す手書き情報切出部 1 3 6 と、差分抽出部 1 3 2 による抽出結果に基づき（好ましくは、手書き情報切出部 1 3 6 により切り出された手書き情報について）、解答者情報欄 8 6 の手書きによる記入情報を、追記情報処理装置 1 0 上で加工編集が可能な文字データに変換する文字認識処理（OCR ; Optical Character Reader）部 1 3 8 とを有する。

【 0 0 6 9 】

なお、解答者抽出部 1 3 4 は、文書入力装置 2 0 で読取り対象となった付加情報記入済教材 8 1 における解答者情報欄 8 6 に記入された解答者情報を抽出できればよく、手書き情報切出部 1 3 6 と文字認識処理部 1 3 8 の何れか一方を備えていればよい。また、文字認識処理部 1 3 8 を設けない場合や文字認識処理部 1 3 8 で文字認識できなかったコメント 8 8 の部分に関しては、抽出された解答者情報をそのまま画像として取り扱うことにする。

40

【 0 0 7 0 】

データ処理用追記情報抽出処理部 1 4 0 は、注目する追記色の情報を参照して、差分抽出部 1 3 2 による抽出結果に基づき、追記情報抽出部 1 3 0 により抽出された追記情報の内、文書入力装置 2 0 で読取り対象となった追記済文書 8 B（本例では付加情報記入済教材 8 1）におけるデータ処理対象追記情報 9 a（本例では採点記号 8 7 やコメント 8 8）

50

を抽出するデータ処理対象追記情報抽出部 1 4 2 と、データ処理対象追記情報抽出部 1 4 2 で抽出されたデータ処理対象の追記情報をデータ処理に耐え得るように整形する追記情報整形部 1 4 6 とを有する。本実施形態においては、データ処理対象追記情報抽出部 1 4 2 は、採点記号 8 7 を抽出する採点記号抽出部およびコメント 8 8 を抽出するコメント抽出部の機能を備える。

【 0 0 7 1 】

なお、データ処理対象追記情報抽出部 1 4 2 は、データ処理対象追記情報 9 a を抽出できればよく、色を指標に追記情報をさらに分類するとよい。たとえば差分抽出部 1 3 2 での抽出結果に対する色成分認識処理を通じて、採点官が採点記号 8 7 やコメント 8 8 の記入に使用したペン色と対応する所定色成分についてのものを抽出することで行なえばよい。たとえば、付加情報記入済教材 8 1 における採点記号 8 7 やコメント 8 8 の記入は、一般に赤色ペン（赤の筆記具）で行なわれることが多く、この場合には、赤色成分に注目した抽出を行なえばよい。

10

【 0 0 7 2 】

ただし、赤色ペンとはいってもピンク系からオレンジ系というように似通った色気のものがあるし、必ずしも採点記号 8 7 やコメント 8 8 の記入に赤色ペンを使用するとは限らないし、採点記号 8 7 とコメント 8 8 とを別のペン色で記入することもあるので、採点記号 8 7 やコメント 8 8 の記入に使用するペン色の情報を採点記号抽出部やコメント抽出部として機能するデータ処理対象追記情報抽出部 1 4 2 に設定可能に構成しておくことで、抽出性能を向上させるようにするとよい。

20

【 0 0 7 3 】

このため、実際に使用した追記色を特定し、その特定した追記色の情報を参照して、追記色に注目した抽出を行なう。実際に使用されたペン色が特定されていれば、データ処理対象追記情報抽出部 1 4 2 は、抽出許容範囲を狭くすることができる。これにより、採点記号 8 7 やコメント 8 8 をその他の追記情報と高精度に区別して抽出することができる。

【 0 0 7 4 】

追記情報整形部 1 4 6 としては、データ処理対象追記情報抽出部 1 4 2 で抽出されたデータ処理対象の追記情報について、線分同士を接続してその抽出線分間の途切れを解消するように補正処理を行なう抽出線分途切れ補正部 1 4 8 を有する構成とするのがよい。

【 0 0 7 5 】

一般に、追記済文書 8 B 上での図形記入や「2重線」や「波線」や「花丸」や「矢印」などの図形記入やコメント文などは、記入済のものに対して重ねて行なわれることもある。たとえば、付加情報記入済教材 8 1 の場合には、各問題文 8 2 や各解答欄 8 4 を特定する枠や各解答欄 8 4 への解答記入内容などに重ねて「 」や「 x 」などの採点記号 8 7 が記入され、あるいは図形や文字でコメント 8 8 の追記が行なわれることもある。そのため、データ処理対象追記情報抽出部 1 4 2 による所定色成分の抽出結果は、その重なり部分が除かれる結果、図形や文字に途切れ部分が生じたものとなる虞れがある。

30

【 0 0 7 6 】

このことから、抽出線分途切れ補正部 1 4 8 は、「 」や「 x 」や「線」やその他の印（マーク）などの図形や文字であるはずの抽出結果に対して、細線化処理、端点抽出処理、端点間接続処理（いわゆる連結処理）、あるいは線図形の直線近似などを適宜実行する。なお、このときに行なう細線化処理、端点抽出処理、あるいは端点間接続処理や線図形の直線近似などは、公知技術を利用して行なえばよいため、ここではその詳細な説明を割愛する（たとえば、「画像の処理と認識」，安居院猛著，昭晃堂発行などを参照）。

40

【 0 0 7 7 】

データ処理対象追記情報特定処理部 1 5 0 は、差分抽出部 1 3 2 による差分抽出結果に基づいて、具体的には、抽出線分途切れ補正部 1 4 8 で補正されたデータ処理対象追記情報 9 a に基づいて、追記済文書 8 B における第 1 のデータ処理用の追記情報の記入内容を第 2 のデータ処理用の追記情報と分離して認識する第 1 データ処理用追記情報認識部 1 5 4 と第 2 のデータ処理用の追記情報の記入内容を第 1 のデータ処理用の追記情報と分離し

50

て認識する第2データ処理用追記情報認識部164とを有している。

【0078】

データ処理対象追記情報特定処理部150としては、第1データ処理用追記情報認識部154は、付加情報記入済教材81における採点記号87の記入内容をコメント88と分離して認識し、また、第2データ処理用追記情報認識部164は、付加情報記入済教材81におけるコメント88の記入内容を採点記号87と分離して認識する。

【0079】

各データ処理用追記情報認識部154, 164は、抽出線分途切れ補正部で補正されたデータ処理用追記情報の記入内容に対して形状認識処理を行なうことでデータ処理用追記情報の記入内容を認識する図形形状認識部156, 166と、抽出線分途切れ補正部で補正されたデータ処理用追記情報の記入内容に対して文字認識処理を行なうことでデータ処理用追記情報の記入内容を認識する文字認識部157, 167と、図形形状認識部156, 166や文字認識処理部157, 167により認識された各データ処理用追記情報の記入内容の、文書原本8A(追記済文書8B)上における記入位置を認識する記入位置認識部158, 168とを有している。図形形状認識部156, 166および文字認識処理部157, 167により、採点記号87とコメント88とを分離して認識する分離認識処理部155が構成される。



【0080】

文字認識処理部157, 167を設けない場合や文字認識処理部157, 167で文字認識できなかったデータ処理用追記情報の部分に関しては、抽出されたデータ処理用追記情報をそのまま画像として取り扱うことにする。

【0081】

なお、図示のように、図形形状認識部156, 166、文字認識処理部157, 167、並びに記入位置認識部158, 168とは、それぞれ1つの機能部が双方の機能を実現する構成としてもよいし、それぞれを個別の機能部として独立に設けてもよい。

【0082】

たとえば、付加情報記入済教材81を処理対象とする構成においては、第1データ処理用追記情報認識部154は、採点記号87を第1のデータ処理用の追記情報とするものであり採点記号認識部として機能する。この場合、採点記号87についての図形形状認識部156は、採点記号87の記入内容が「正解()」または「不正解(x)」または「一部正解()」であるかなどを図形の側面から認識することができればよく、たとえば「」、「x」、「」の図形形状とのパターンマッチングによって形状認識を行えばよい。あるいは、認識対象図形の特徴量を算出し、その特徴量から形状を認識してもよい。特徴量としては、たとえば、穴の個数や外接矩形に占める対象図形の面積率などを使用することができる。

【0083】

また、採点記号87について文字認識処理部157は、採点記号87の記入内容が「正解()」または「不正解(x)」または「一部正解()」であるかなどを文字の側面から認識することができればよい。なお、採点記号87は図形のみであるとする場合には、文字認識処理部157を割愛することもできる。

【0084】

また、採点記号87についての記入位置認識部158は、たとえば、教育用教材80上における座標解析によって、付加情報記入済教材81上の採点記号87の記入内容の記入位置を認識すればよい。

【0085】

ここで、本実施形態の構成においては、認識性能情報提示処理部190を設けて、追記情報処理装置10の各機能部における複数の処理から、それぞれ認識率に関わる情報を収集し、記入内容を自動認識した追記情報の中で、特に信頼度の低い難認識情報のみを提示して修正を促すようにしているため、全ての追記情報について、修正の可否をチェックする必要が無く、誤認識の追記情報の修正作業を効率的に実行することができる。もちろん

10

20





30

40

50

、その結果として、各データ処理用追記情報認識部 154, 164 における認識率を向上させることができる。



【0086】

なお、図形状認識部 156 は、採点記号 87 に関する形状認識の際には、「」や「」などの採点記号 87 を示す図形を構成する連続画素群を 1 つに纏めて取り扱うために、その連続画素群に対して識別子を付与すべく、一般的な画像処理技術であるラベリング処理を行なう。このことから、記入位置認識部 158 による位置認識の際にも、そのラベリング処理の結果を利用して、「」や「」などの採点記号 87 を示す図形を構成する連続画素群を 1 つの纏まりとして取り扱う。

【0087】

また、記入位置認識部 158 は、採点記号 87 の記入位置の認識処理に当たって、付加情報記入済教材 81 上に複数の採点記号 87 が記入されていることが一般的であるから、その複数の採点記号 87 のそれぞれについて順次予め定められた走査順で検出される採点記号 87 について、順にその位置を認識していく。

【0088】

各採点記号 87 に関する位置認識は、たとえば「」や「」などの採点記号 87 を示す図形（あるいは文字）の外接矩形情報を算出し、さらにその外接矩形の中心座標を算出することによって行なうことが考えられる。具体的には、認識対象となる図形もしくは文字（連続画素群）に対して外接矩形を抽出するとともに、その外接矩形の所定点（たとえば左上頂点）の x y 座標、並びに、その外接矩形の幅（ W ）および高さ（ h ）を算出する。そして、これらの算出結果から、中心 x 座標 = $x + w / 2$ 、中心 y 座標 = $y + h / 2$ を算出し、その算出結果を連続画素群の位置、すなわち採点記号 87 の記入位置の認識結果とする。

【0089】

一方、第 2 データ処理用追記情報認識部 164 は、コメント 88 を第 2 のデータ処理用の追記情報とするコメント認識部として機能する。この場合、コメント 88 についての図形状認識部 166 は、コメント 88 の追記内容を図形の側面から認識することができればよく、たとえば「1 重線」や「2 重線」や「（1 重または 2 重の）波線」などの線を示す図形状とパターンのマッチングによって線に関する形状認識を行えばよい。あるいは、認識対象図形の特徴量を算出し、その特徴量から線の形状を認識してもよい。特徴量としては、たとえば、線数や外接矩形に対する画素密度などを使用することができる。

【0090】

また、コメント 88 について文字認識処理部 167 は、付加情報記入済教材 81 におけるコメント 88 の記入内容を文字の側面から認識することができればよい。また、コメント 88 についての記入位置認識部 168 は、たとえば、教育用教材 80 上における座標解析によって、付加情報記入済教材 81 上のコメント 88 の追記内容の記入位置を認識すればよい。

【0091】

なお、図形状認識部 166 は、コメント 88 に関する形状認識の際には、「2 重線」や「花丸」などのコメント 88 を示す図形を構成する連続画素群を 1 つに纏めて取り扱うために、その連続画素群に対して識別子を付与すべく、一般的な画像処理技術であるラベリング処理を行なう。このことから、記入位置認識部 168 による位置認識の際にも、そのラベリング処理の結果を利用して、「2 重線」や「花丸」などのコメント 88 を示す図形を構成する連続画素群を 1 つの纏まりとして取り扱う。

【0092】

また、記入位置認識部 168 は、コメント 88 の記入位置の認識処理に当たって、付加情報記入済教材 81 上に複数のコメント 88 が記入されていることが一般的であるから、その複数のコメント 88 のそれぞれについて順次予め定められた走査順で検出されるコメント 88 について、順にその位置を認識していく。

【0093】

10

20

30

40

50

各コメント 88 に関する位置認識は、たとえばコメント文や「2重線」や「花丸」などのコメント 88 の文字や図形の外接矩形情報を算出し、さらにその外接矩形の中心座標を算出することによって行なうことが考えられる。具体的には、認識対象となる文字や図形（連続画素群）に対して外接矩形を抽出するとともに、その外接矩形の所定点（たとえば左上頂点）の x y 座標、並びに、その外接矩形の幅（ W ）および高さ（ h ）を算出する。そして、これらの算出結果から、中心 x 座標 = $x + w / 2$ 、中心 y 座標 = $y + h / 2$ を算出し、その算出結果を連続画素群の位置、すなわちコメント 88 の記入位置の認識結果とする。

【0094】

また、この位置認識の際には、各コメント 88 は、ある位置の解答欄 84 への採点記号 87 と対応して、その近傍に記入されることが多いので、記入位置認識部 158 による採点記号 87 についての位置認識と協働して処理を行なうのがよい。こうすることで、双方の位置情報の各解答欄 84 との対応付け、結果としては、採点記号 87 とコメント 88 との関連付けが容易になる。

10

【0095】

データ処理部 170 は、文書入力装置 20 から入力された追記済文書 8B の画像データについて、その追記済文書 8B に記入された第 1 のデータ処理対象追記情報に関する第 1 のデータ処理を実行する第 1 データ処理部 170_1 と、第 2 のデータ処理対象追記情報に関する第 2 のデータ処理を実行する第 2 データ処理部 170_2 を有する。

【0096】

20

各データ処理部 170_1, 170_2 は、データ処理対象の追記情報の記入位置を、当該記入欄の位置情報を保存している装置（テンプレート情報データベース DB1 や文書原本情報データベース DB2 として機能する文書管理サーバ 30）にアクセスして、記入欄位置領域情報 38、記入欄位置領域情報 68、あるいはテンプレート関連付け情報 69 を参照して取得しつつ、追記情報の記入位置と追記情報とを対応付けながらデータ処理を実行する。

【0097】

第 1 データ処理部 170_1 は、追記済文書 8B の一例である付加情報記入済教材 81 の画像データについて、その付加情報記入済教材 81 に記入された採点記号 87 を元に採点集計を行なう採点集計部 172 と、採点集計部 172 による採点集計の結果を、解答者抽出部 134 が抽出した解答者情報と関連付けて出力する集計結果出力部 174 とを備えている。採点集計結果と解答者情報とが関連付けられた状態の処理結果を特に採点認識結果と称する。

30

【0098】

採点集計部 172 は、図形状認識部 156 による採点記号 87 の追記内容の図形の側面からの認識結果や文字認識処理部 157 による採点記号 87 の追記内容の文字情報の側面からの認識結果と、記入位置認識部 158 による採点記号 87 の記入位置の認識結果と、文書入力装置 20 が保持蓄積している付加情報記入済教材 81 に対応する教育用教材 80 の電子データ（原本画像）に含まれる教育用教材 80（付加情報記入済教材 81）の各解答欄 84 についての配点欄 83 で規定されている配点情報とに基づいて、文書入力装置 20 が読み取った付加情報記入済教材 81 について、付加情報記入済教材 81 に記入された採点記号 87 に関する採点処理および集計処理（纏めて採点集計という）を行なう。

40

【0099】

集計結果出力部 174 は、採点集計部 172 により集計された採点集計結果と解答者抽出部 134 が抽出した解答者情報と関連付けて、処理結果保存サーバ 40（処理結果データベース装置や処理結果ファイルサーバ装置など）に登録する。あるいは、採点結果の点数を付加情報記入済教材 81 の集計欄 83b に記入し用紙上に返却答案 81b として出力して生徒などに返却できるようにする。

【0100】

また、第 2 データ処理部 170_2 は、追記済文書 8B の一例である付加情報記入済教材

50

81の画像データについて、その付加情報記入済教材81に記入されたコメント88を元に分類処理を行なうコメント分類処理部176と、コメント分類処理部176による分類結果を集計結果出力部174が出力した採点認識結果や各解答に関連付けて出力するコメント処理結果出力部178とを備えている。

【0101】

コメント分類処理部176は、図形形状認識部166によるコメント88の追記内容の図形の側面からの認識結果や文字認識処理部167によるコメント88の追記内容の文字情報の側面からの認識結果と、記入位置認識部168によるコメント88の記入位置の認識結果と、コメント88の追記内容と対応するように予め規定されている分類情報とに基づいて、文書入力装置20が読み取った付加情報記入済教材81について、その付加情報記入済教材81に記入されたコメント88の分類処理を行なう。

10

【0102】

コメント処理結果出力部178は、コメント分類処理部176による分類結果を各解答欄や集計結果出力部174が出力した採点認識結果と関連付けて、処理結果保存サーバ40(処理結果データベース装置や処理結果ファイルサーバ装置など)に登録する。

【0103】

なお、採点記号87の記入は、一般に教育用教材80上の複数の解答欄84のそれぞれに対応して行なわれ、またコメント88の記入は、採点記号87の記入に付随してその採点記号87の近傍に必要な応じて記入され、かつ採点記号87、コメント88は教師などの採点官によって手書きでされるため、各解答欄84に対する記入位置が必ずしも一義的に定まっている訳ではない。

20

【0104】

その一方で、採点記号87の採点集計に当たっては、各解答欄84と採点記号87の記入位置との対応を明確にする必要がある。採点記号87に関する採点集計は、各解答欄84に対応する採点記号87の記入結果を明確にした上で、採点記号87の内容(正解か不正解か一部正解かなど)および各解答欄84についての配点に基づいて行なわれるからである。同様に、コメント88についての分類処理に当たっては、各解答欄84(つまり採点記号87)とコメント88の記入位置との対応を明確にする必要がある。コメント88に関する分類処理は、各解答欄84に対応するコメント88の記入結果を明確にした上で、コメント88の内容に基づいて行なわれるからである。

30

【0105】

このことから、採点集計部172やコメント分類処理部176は、以下に述べるような手順で、採点記号87の採点集計やコメント88の分類処理を行なう。たとえば、採点集計部172やコメント分類処理部176は、記入位置認識部158,168で特定される「 \square 」や「 \times 」などの採点記号87やコメント88の外接矩形が、付加情報記入済教材81上で解答欄84となる領域との重なるものがあるか否かを判定し、重なる解答欄84と採点記号87やコメント88とを互いに対応付け、その採点記号87やコメント88を解答欄84に対して記入された採点記号87やコメント88の判定結果とする第1の対応付け手法を採用することができる。

【0106】

40

ただし、1つの採点記号87やコメント88が複数の解答欄84の領域に重なる場合には、何れに対応させるべきかを特定することはできないので、第1の対応付け手法による対応付けについての判定が不能であると判断する。また、注目する採点記号87やコメント88の外接矩形が、何れの解答欄84の領域にも重ならない場合にも、何れに対応させるべきかを特定することはできないので、第1の対応付け手法による対応付けについての判定が不能であると判断する。

【0107】

また、採点集計部172やコメント分類処理部176は、記入位置認識部158,168で特定される採点記号87やコメント88の外接矩形と、付加情報記入済教材81上で解答欄84となる領域との重なり面積を求め、その面積(外接矩形に対する面積比でも同

50

様)が最も大きくなる採点記号87やコメント88と解答欄84とを互いに対応付け、その採点記号87やコメント88を解答欄84に対して記入された採点記号87やコメント88の判定結果とする第2の対応付け手法を採用することができる。

【0108】

この第2の対応付け手法を採用すると、1つの採点記号87やコメント88が複数の解答欄84の領域に重なる場合に第1の対応付け手法にては対応付けの特定ができない場合でも、重なり面積の大小に基づいて、何れに対応させるべきかを判定することができる。ただし、重なり面積の外接矩形に対する比が所定閾値未満の場合には、重なる部分が小さいことから、対応付けについての判定が不能であると判断する。

【0109】

あるいは、採点集計部172やコメント分類処理部176は、記入位置認識部158, 168で特定される各採点記号87やコメント88の中心座標位置と各解答欄84の中心座標位置の距離を求め、その距離が最も小さくなる採点記号87やコメント88と解答欄84とを互いに対応付け、その採点記号87やコメント88を解答欄84に対して記入された採点記号87やコメント88の判定結果とする第3の対応付け手法を採用することができる。

【0110】

この第3の対応付け手法を採用すると、注目する採点記号87やコメント88の外接矩形が何れの解答欄84の領域にも重ならない場合に第1の対応付け手法にては対応付けの特定ができない場合や、採点記号87やコメント88が解答欄84からずれて記入されて重なる部分が小さく、重なり面積の外接矩形に対する比が所定閾値未満の場合に第2の対応付け手法にては対応付けの特定ができない場合でも、何れに対応させるべきかを判定することができる。ただし、各解答欄84との間の各距離の差が所定閾値未満の場合には、距離差が小さいことから、対応付けについての判定が不能であると判断する。

【0111】

そして、各採点記号87の解答欄84への対応付けを行なった後は、採点記号87が「」であれば、これに対応する解答欄84についての配点情報から特定される配点を加算し、また採点記号87が「×」であれば、これに対応する解答欄84についての配点加算を行わず、このような採点集計を付加情報記入済教材81上の全ての解答欄84について行なう。

【0112】

なお、付加情報記入済教材81上で解答欄84となる領域は、各解答欄84についての配点情報として、または当該配点情報と同様に、付加情報記入済教材81に対応する文書管理サーバ30に登録されている原本画像に含まれる記入欄位置領域情報38によって特定されるものとする。

【0113】

また、各コメント88の解答欄84への対応付けを行なった後は、コメント88に対応する分類基準から特定される分類先を特定し、このようなコメント分類処理を付加情報記入済教材81上の全てのコメント88について行なう。

【0114】

なお、採点集計部172での採点集計処理やコメント分類処理部176での分類処理に当たっては、完全なる自動処理にしてもよいが、ユーザ端末171のCRT(Cathode Ray Tube)やLCD(Liquid Crystal Display)などで構成された表示部に処理過程や処理結果を表示して、適宜、操作者が処理過程や処理結果をキーボードやマウスなどの指示入力部を介して訂正できるようにしてもよい。

【0115】

また、本実施形態の特徴部分である認識性能情報提示処理部190は、手書き入力情報の履歴を収集しデータベースとしての処理結果保存サーバ40に保存・蓄積しておく追記情報認識履歴保持部192と、一定の度合いよりも認識の困難であった特に信頼度の低い難認識情報を抽出する難認識追記情報抽出部194と、所定の表示態様に従って認識性

10

20

30

40

50

能情報をユーザ端末171上で提示する認識性能情報提示部198とを有している。

【0116】

追記情報認識履歴保持部192は、後述する第1～第3の履歴収集保存手法の何れかもしくはその任意の組合せの履歴収集保存手法を採用して、手書き入力情報の履歴を収集し、この収集した手書き入力情報の履歴を処理結果保存サーバ40に保存する。

【0117】

手書き入力情報の履歴を収集しデータベースに保存しておく際には、どの時点で記入されたものを履歴として残しておくかによって、様々な履歴収集保存手法を採ることができる。たとえば、通常の自動データ処理の過程で認識した手書き入力情報のデータを蓄積し保存していく第1の履歴収集保存手法を採ることが考えられる。第1の履歴収集保存手法を採る際には、手書き入力情報のみを元の文書原本8Aから分離して保存することも考えられるし、手書き入力情報を含む所定範囲の画像情報をも一緒にして、つまり周囲の画像ごと

10

【0118】

第1の履歴収集保存手法を採ると、通常の処理過程で履歴を取って保存していくことができるので、形状や位置や修正状況などのデータを利用することができる。また、特別な手間を掛けずに、採点法などの記入態様をチェックすることができる、しかも、認識の正誤も反映させることができる。ただし、当初は履歴が少なく、また、ユーザによる入力態様の改善も期待できないので認識率が低く、修正の手間が掛かる可能性はある。

【0119】

20

また、通常の自動データ処理の過程ではなく、その処理開始前に、練習用答案などの練習用の文書原本8Aに入力してもらい、その入力情報のデータを蓄積し保存しておく第2の履歴収集保存手法を採ることが考えられる。この第2の履歴収集保存手法を採ると、練習用の記入情報を指定することができるから、修正を考慮しなくてよい。用意された練習用答案などに追記し、通常の自動採点処理を行なう過程で、記入形状も指定されるからである。また、図形形状や文字形状と位置を判定し、認識の正誤も含めて結果を提示できる利点もある。

【0120】

加えて、事前に採点法などの記入態様を練習することができるので、正しい形と位置とを確認でき、認識性能のよい状態での記入が期待でき、その結果として、実際の自動データ処理時には修正を要しない利点が得られる。また、白紙部分に記入させることもできるので、指定された記入情報に対する手書き入力情報のみを保存すればよいので、蓄積容量を低減できる。

30

【0121】

また、自動データ処理には使用しなかった過去の追記済文書8B（たとえば過去の採点答案）における入力情報のデータを蓄積し保存する第3の履歴収集保存手法を採ることが考えられる。なお、自動データ処理を行っていないので、白紙の文書原本8Aを付加して読み取ることで差分を抽出して、手書き入力情報のみを抽出するようにする。また、自動データ処理の全行程は行わずに、形状認識や文字認識のみを行ない、結果を提示するだけにすればよい。

40

【0122】

この第3の履歴収集保存手法を採ると、過去の追記文書を利用できるため、はじめて自動データ処理を使用する場合でも、手書き入力情報に関して認識性能を改善させるための修正手法を提示することができ、実際の自動データ処理時には修正が少なくて済む利点が得られる。また、練習用の文書原本8Aを作成する手間を省くことができる。ただし、結果の正誤は不明のため、正誤を踏まえた判定はできない。

【0123】

難認識追記情報抽出部194は、後述する各種の手法の何れかもしくはその任意の組合せの難認識追記情報抽出手法を採用して、各種の手書き入力情報の内、信頼度が一定レベルよりも低い自動認識が困難であった難認識情報を抽出する。

50

【 0 1 2 4 】

認識性能情報提示部 1 9 8 は、後述する各種の手法の何れかの認識性能情報提示手法を採用して、認識性能情報をユーザ端末 1 7 1 上でユーザに提示する。

【 0 1 2 5 】

< 全体の処理手順 >

図 3 および図 4 は、情報処理システムの一実施形態である教材自動採点システム 1 における教材処理方法の処理動作の手順を説明する図である。ここで、図 3 は、その全体概要をシステム構成図と対応付けて示しており、また図 4 は、教材処理手順を示すフローチャートである。

【 0 1 2 6 】

まず、教育用教材 8 0 を完成させ文書原本情報データベース DB 2 に登録する (S 1 0 4)。なお、各記入欄 6 a に配点が設定されていない形態のテンプレート 6 を使用する場合には、解答欄 8 4 として設定した記入欄 6 a に関しては配点の情報も設定する。

【 0 1 2 7 】

この後、試験を実行する際には、文書原本情報データベース DB 2 から教育用教材 8 0 を読み出して印刷し生徒や受験者に配布する (S 1 0 6)。そして、試験終了後に、採点官は採点記号 8 7 やコメント 8 8 を生徒の解答に対して追記する (S 1 0 8)。

【 0 1 2 8 】

追記情報処理装置 1 0 (特に教材自動採点システム 1 においては教材処理装置に該当する) を利用する場合には、まず、生徒などによって解答者情報欄 8 6 への氏名などの記入および解答欄 8 4 への解答記入、つまり生徒による第 1 種の付加情報の記入がされ、さらに教師などによって各解答欄 8 4 に記入された解答に対する「 」や「 × 」などの採点記号 8 7 やコメント 8 8 などの第 2 種の付加情報の記入がされた付加情報記入済教材 8 1 について、文書入力装置 2 0 は、その付加情報記入済教材 8 1 を読み取り (S 1 1 0)、その付加情報記入済教材 8 1 を表わす画像データを追記情報処理装置 1 0 に入力する (S 1 1 2)。文書入力装置 2 0 は、この文書入力装置 2 0 による画像読取りによって得られた画像データについて、一旦ワークエリアとして用いられるメモリなどに保持しておく。

【 0 1 2 9 】

なお、このとき、文書入力装置 2 0 にて A D F 装置を用いれば、たとえば同一学級のような 1 つのグループに纏めて処理すべき複数の付加情報記入済教材 8 1 について、一括して読み取り (一括スキャン)、各付加情報記入済教材 8 1 に対応する画像データを連続的に追記情報処理装置 1 0 に入力することができる。

【 0 1 3 0 】

追記情報処理装置 1 0 は、文書入力装置 2 0 から取り込んだ各付加情報記入済教材 8 1 の画像データに対して、順次、次のような付加情報抽出・分離処理、付加情報特定処理、および自動採点処理並びに自動コメント処理といった付加情報終末処理を実行する。

【 0 1 3 1 】

たとえば、データ処理部 1 7 0 における自動採点処理並びに自動コメント処理に先立って、ある 1 つの付加情報記入済教材 8 1 から得られた画像データについて、読取画像処理部 1 1 0 の画像データ解析部は解析処理を行ない (S 1 2 2)、教材特定部 1 2 2 は、その解析処理の結果に基づいて付加情報記入済教材 8 1 に対応する元の教育用教材 8 0 の識別特定を行なう (S 1 2 4)。

【 0 1 3 2 】

この識別特定 (S 1 2 4) は、たとえば「理科」「5年」「1. 天気と気温の変化」といった識別情報解析部によるタイトル解析または識別情報欄 8 5 に埋め込まれたコード情報についてのコード情報解析部によるコード解析を通じて行なえばよい。この識別特定を経ることで、教材特定部 1 2 2 では、文書入力装置 2 0 により得られた付加情報記入済教材 8 1 の画像データとの比較対象となる教育用教材 8 0 の電子データ (原本画像) を特定することができる。

【 0 1 3 3 】

10

20

30

40

50

なお、この識別特定は、文書入力装置 20 が画像読取りを行なった複数の付加情報記入済教材 81 のそれぞれについて順次行なうことも考えられるが、一般に 1 つのグループに纏めて処理される付加情報記入済教材 81 は全て同一のものであるため、その纏めて処理される中で最初に処理される付加情報記入済教材 81 についてのみ行なえばよい。

【0134】

教材特定部 122 により各付加情報記入済教材 81 に対応する教育用教材 80 の特定が完了すると、文書管理サーバ 30 は、その特定結果に従いつつ、保持蓄積している中から該当する教育用教材 80 の原本画像（電子データ）を取り出して、これを差分抽出部 132 へ受け渡す（S126）。

【0135】

また、歪み補正部は、ある 1 つの付加情報記入済教材 81 から得られた画像データの歪みを補正する（S128）。この画像歪み補正は、文書入力装置 20 での画像読取りの際に生じ得る画像歪みを補正するために行なうものであり、その後差分抽出部 132 にて行なう原本画像との比較や差分抽出などの精度向上を図るためのものである。

【0136】

差分抽出部 132 は、文書管理サーバ 30 から受け渡された原本画像（教育用教材 80）と、文書入力装置 20 から入力され、歪み補正部により画像歪みが補正された後の画像データ（付加情報記入済教材 81）とを、それぞれ比較して、その差分を抽出する（S130）。差分抽出部 132 は、抽出した差分情報 9 を解答者抽出部 134 やデータ処理用追記情報抽出処理部 140 に渡す。

【0137】

差分抽出部 132 による差分抽出によって、たとえば図 3 中の中央部分に示すように、解答者情報欄 86 および各解答欄 84 への解答者による第 1 種の付加情報の記入内容、並びに各解答欄 84 に対する採点記号 87 やコメント 88 などの採点官による第 2 種の付加情報の記入内容のみで表わされる差分情報 9 が抽出されることになる。

【0138】

解答者抽出部 134 は、差分情報 9 に対する文字認識処理部による文字認識処理などを通じて、文書入力装置 20 で読取り対象となった付加情報記入済教材 81 における解答者情報を抽出する（S132）。これにより、ある 1 つの付加情報記入済教材 81 に解答を記入した解答記入者の学級、出席番号、氏名などを特定できる。

【0139】

また、データ処理用追記情報抽出処理部 140 において、先ず追記部材特定部 141 は、データ処理の対象となる追記情報の記入に使用されたペン色である追記色を特定し（S141）、データ処理対象追記情報抽出部は、追記部材特定部 141 にて特定された追記色に基づき、差分抽出部 132 による差分抽出結果に対して、データ処理用の追記情報を抽出する（S142）。

【0140】

本例の場合、各解答欄 84 への採点記号 87 やコメント 88 の追記内容を抽出するために、その差分情報 9 からさらに所定色成分についてのもの、具体的にはたとえば赤色成分のものを抽出する。所定色成分の抽出は、たとえば差分抽出結果が画素データからなる場合であれば、その画素データを構成する色成分データに着目することで行なうことができる。

【0141】

抽出線分途切れ補正部は、データ処理対象追記情報抽出部による抽出結果に対して、細線化処理、端点抽出処理、端点間接続処理、あるいは線図形の直線近似などの追記情報整形処理を適宜実行する（S146）。抽出線分途切れ補正部は、途切れ補正処理済の採点記号 87 の抽出結果を採点記号認識部として機能する第 1 データ処理用追記情報認識部 154 に渡し、また途切れ補正処理済のコメント 88 の抽出結果をコメント認識部として機能する第 2 データ処理用追記情報認識部 164 に渡す。

【0142】

10

20

30

40

50

データ処理対象追記情報特定処理部 150 は、先ず、分離認識処理部 155 を構成する図形形状認識部 156 , 166 および文字認識処理部 157 , 167 が協働して、文書管理サーバ 30 に保存されている解答欄 84 の位置情報を参照して、採点記号 87 とコメント 88 とを分離し (S 162)、この後、分離した採点記号 87 とコメント 88 の別に、記入内容の特定処理や記入位置の特定処理を実行する。

【 0 1 4 3 】

図形形状認識部 156 や文字認識処理部 157 は、コメント 88 と分離した採点記号 87 について (S 163 - 採点記号)、採点記号 87 の記入内容に対する形状認識あるいは文字認識により、その採点記号 87 の記入内容が「正解」であるかあるいは「不正解」であるかなど、採点記号 87 で示された採点官の採点結果を特定する (S 164)。このとき、ユーザによる修正指示を受け付ける (S 165)。続いて、記入位置認識部 158 は、採点記号 87 の記入内容について、その付加情報記入済教材 81 上における記入位置を認識する (S 166)。

10

【 0 1 4 4 】

このようにして、記入位置認識部 158 が採点記号 87 の記入位置を認識した後は、採点集計部 172 は、図形形状認識部 156 や文字認識処理部 157 による採点記号 87 の記入内容の認識結果と、記入位置認識部 158 による採点記号 87 の記入位置の認識結果と、文書管理サーバ 30 が保持蓄積している付加情報記入済教材 81 に対応する原本画像 (教育用教材 80) に含まれる教育用教材 80 の各解答欄 84 についての配点情報とに基づいて、採点および集計を行なう (S 168)。

20

【 0 1 4 5 】

集計結果出力部 174 は、その採点・集計の結果を処理結果保存サーバ 40 に保存する (S 169)。あるいは採点結果の点数が付加情報記入済教材 81 の集計欄 83 b に記入されて返却答案 81 b として生徒などに返却される。

【 0 1 4 6 】

各付加情報記入済教材 81 についての採点結果 (問題別採点結果) のファイル形式としては、たとえば、図 3 に示すように、付加情報記入済教材 81 上に存在する問題の番号と、その問題の解答に対する正誤判定と、その正誤判定に基づく得点とからなる情報で、これらを互いに関連付けるテーブル形式である。また、集計結果のファイル形式としては、たとえば、図 3 に示すように、出席番号および解答者情報と、得点情報 (集計欄 83 b に記入される項目点や合計点) とからなる情報で、これらを互いに関連付けるテーブル形式である。

30

【 0 1 4 7 】

各付加情報記入済教材 81 上に記入される正誤判定の採点集計の結果が問題別採点結果としてファイル出力されるし、また、問題別の採点集計の結果がファイル出力されるので、処理結果保存サーバ 40 では、付加情報記入済教材 81 についての採点集計結果を、たとえば一覧形式で、管理または利用することが可能となる。

【 0 1 4 8 】

図形形状認識部 166 や文字認識処理部 167 は、採点記号 87 と分離したコメント 88 について (S 163 - コメント)、コメント 88 の記入内容に対する形状認識あるいは文字認識により、採点官により追記されたコメント 88 の記入内容を特定する (S 170)。このとき、ユーザによる修正指示を受け付ける (S 171)。続いて、記入位置認識部 168 は、コメント 88 の記入内容について、その付加情報記入済教材 81 上における記入位置を認識する (S 172)。

40

【 0 1 4 9 】

このようにして、記入位置認識部 168 がコメント 88 の記入位置を認識した後は、コメント分類処理部 176 は、図形形状認識部 166 や文字認識処理部 167 によるコメント 88 の記入内容の認識結果と、記入位置認識部 168 によるコメント 88 の記入位置の認識結果と、文書管理サーバ 30 が保持蓄積している分類基準情報 39 とに基づいて、コメント 88 を分類する (S 174)。

50

【 0 1 5 0 】

コメント処理結果出力部 1 7 8 は、コメント分類処理部 1 7 6 で分類された各コメント 8 8 を、位置が近い採点記号 8 7 の採点結果と関連付けて (S 1 7 8)、処理結果保存サーバ 4 0 に保存する (S 1 7 9)。各コメント 8 8 についての分類結果のファイル形式としては、たとえば、図 3 に示すように、各コメントと近傍の採点記号 8 7 とからなる情報で、これらを互いに関連付けるテーブル形式である。なお、実際にテーブル形式で保存することに限らず、各コメント 8 8 と、このコメント 8 8 と対応する採点集計結果の両者を関連付ける関連付け情報とを対応付けて保存してもよい。

【 0 1 5 1 】

各付加情報記入済教材 8 1 上に記入される採点記号 8 7 の採点集計結果とコメント 8 8 とが対応するようにデータ保存されるので、処理結果保存サーバ 4 0 では、付加情報記入済教材 8 1 についての採点集計結果とコメント 8 8 とを、たとえば一覧形式で、管理または利用することが可能となるし、生徒の能力分析や生徒指導にコメント 8 8 を利用できるようになる。

10

【 0 1 5 2 】

このように、情報処理システムの一実施形態として示した教材自動採点システム 1 では、採点記号 8 7 やコメント 8 8 の記入がされた付加情報記入済教材 8 1 から読み取った画像データと、その付加情報記入済教材 8 1 についての元の教育用教材 8 0、すなわち解答欄 8 4 への解答記入などの生徒などによる第 1 種の付加情報および解答に対する採点官による採点記号 8 7 やコメント 8 8 などの第 2 種の付加情報の記入がされていないものについてのデータとを比較し、互いの差分から採点記号 8 7 やコメント 8 8 の記入内容を分離してその記入内容を特定し、その採点記号 8 7 についての採点集計とコメント 8 8 についての分類処理を実行するようになっている。

20

【 0 1 5 3 】

したがって、採点記号 8 7 に関するデータ処理については、差分抽出部 1 3 2 で抽出される差分結果から、コメント 8 8 の記入内容を排除して採点記号 8 7 のみを分離してその記入内容を特定するようにしているので、同じペンで両者が追記されていても、自動採点に悪影響を及ぼすことがない。

【 0 1 5 4 】

また、採点結果の自動集計を実行できるので、結果として付加情報記入済教材 8 1 についての採点処理が省力化される。付加情報記入済教材 8 1 を紙媒体で入手するケースでは、付加情報記入済教材 8 1 を文書入力装置 2 0 で読み取った画像データを基にすればよく、たとえば、複写機、複合機、またはスキャナ装置などによって実現されるスキャン機能と、パーソナルコンピュータ (P C) などのコンピュータ機器が有する情報記憶処理機能、画像処理機能および演算処理機能とがあれば、システム構成を簡単に実現することができ、専用の機器を必要とすることもない。

30

【 0 1 5 5 】

さらには、付加情報記入済教材 8 1 の画像データを、文書管理サーバ 3 0 が保持する電子データと比較するため、その文書管理サーバ 3 0 に各種の教育用教材 8 0 についての電子データを保持蓄積しておけば、対応可能な付加情報記入済教材 8 1 についての汎用性を十分に確保し得る。さらには、文書管理サーバ 3 0 に予め電子データを保持蓄積しておくことで、文書管理サーバ 3 0 から取り込んだ画像データとの比較を行なう場合において、比較対象となる電子データの入力などを行なう手間を省くことができ、結果として迅速な採点処理を実現することができる。

40

【 0 1 5 6 】

また、コメント 8 8 に関するデータ処理については、差分抽出部 1 3 2 で抽出される差分結果から、採点記号 8 7 の記入内容を排除してコメント 8 8 のみを分離してその記入内容を特定するようにしているので、同じペンで両者が追記されていても、コメント 8 8 に関する分類処理に悪影響を及ぼすことがない。また、各採点記号 8 7 と対応付けて処理結果保存サーバ 4 0 に各コメント 8 8 を保存するようにしたので、コメント 8 8 を生徒など

50

が確認する用途以外に、採点官自らが、能力分析や生徒指導に活用できるようになる。

【 0 1 5 7 】

< < 認識性能情報提示処理 > >

図 5 ~ 図 1 6 は、認識性能情報提示処理部 1 9 0 の処理の詳細、すなわち手書きで記入された文字や図形などの手書き入力情報の履歴を取る手法や、認識性能を向上させるための認識性能情報をユーザに提示する手法を説明する図である。

【 0 1 5 8 】

本実施形態においては、文書原本 8 A へ手書きで記入された手書き入力情報が自動認識に適しているか否かを判定し、認識性能上問題のあるものについては、どのように記入すれば認識率が向上するかをユーザ自身が判断することのできるようなサポート情報（認識性能情報）を提示する。

10

【 0 1 5 9 】

ここで、認識性能情報としては、認識性能を向上させるための修正手法をユーザ自身が具体的にまた容易に判断することのできるものであればよく、様々な提示態様を採ることができる。たとえば、何処をどう直せばいいかを示唆する情報を提示する態様を採ることができる。こうすることで、どのように記載すると認識性能がよくなるのかを簡単に判断することができる。

【 0 1 6 0 】

あるいは、過去の記載の認識信頼度を提示する態様を採ることができる。こうすることで、どのように記載すると認識性能がどの程度になるかを容易に判断することができ、これを踏まえて、ユーザは、より認識性能がよくなる記載をすることができるようになる。

20

【 0 1 6 1 】

何れの提示態様を採っても、認識に適した手書き入力情報の記入を促すことができ、結果的に、認識率を向上させることができる。

【 0 1 6 2 】

特に、本実施形態においては、難認識追記情報抽出部 1 9 4 は、追記情報処理装置 1 0 の各機能部における複数の処理から、それぞれ認識率に関わる情報を収集し、記入内容を自動認識した追記情報の中で、特に信頼度の低い難認識情報のみを特定し、認識性能情報提示部 1 9 8 は、難認識追記情報抽出部 1 9 4 にて特定された難認識情報について、その難認識情報の認識性能を向上させるための認識性能情報を提示して修正を促す。

30

【 0 1 6 3 】

自動認識性能との関係で、特に認識性能の劣る信頼度の低い難認識情報のみを提示することで、チェック頻度を少なくするのである。これにより、全ての追記情報について、修正の可否をチェックする必要が無く、誤認識の追記情報の修正作業を効率的に実行することができる。

【 0 1 6 4 】

加えて、認識処理そのものにおける信頼度に関する特徴量だけでなく、この認識処理よりも前段の認識処理に関わる各種の処理における信頼度に関する特徴量も取得して認識情報を特定するので、修正を要する追記情報（難認識情報）を高精度に特定することができる。

40

【 0 1 6 5 】

なお、手書き入力情報の認識性能情報をユーザに提示するに当たっては、先ず、手書き入力情報の履歴を取っておき、それらの認識時の信頼度を参考にして、認識性能向上のために改善要求度合いのより強いものに関して認識性能情報を提示するようにする。換言すれば、一定の度合いよりも認識の困難であった手書き入力情報（難認識情報）についてのみ、認識性能情報を提示する。こうすることで、認識性能を一定の水準以上にできるようにするのである。認識性能が一定の水準以上にできればよく、必ずしも、手書き入力情報が理想的な基準情報と完全に一致していることは必要ないのである。

【 0 1 6 6 】

手書き入力情報は、ほぼ確実に基準情報とのずれがあるから、過去に記入された全ての

50

手書き入力情報について認識性能情報を提示していたのでは、認識性能との関係で、必要以上の情報を提示してしまう虞れがあるが、提示範囲を難認識情報の範囲に絞ることで、必要以上の情報提示を防止するのである。こうすることで、修正作業を効率的に実行することができるし、必要以上に矯正してしまう虞れを防止することもできる。

【0167】

なお、一般的には、ユーザによって記入癖が異なるので、提示すべき改善態様もユーザごとに異なるから、手書き入力情報の履歴を取る際には、ユーザごとに履歴を取っておくのが好ましい。

【0168】

<信頼度情報に基づく難認識情報抽出>

また、難認識追記情報抽出部194において、一定の度合いよりも認識の困難であった難認識情報についてのみ認識性能情報を提示するに当たっては、分離認識処理部155にて通常の認識処理を行ない、難認識追記情報抽出部194は、その認識処理時の信頼度、換言すれば候補情報（基準情報）に対する類似度に基づいて難認識情報を抽出する。

【0169】

また、本実施形態において、難認識追記情報抽出部194は、注目する追記情報に関して、データ処理対象追記情報特定処理部150における認識処理そのものを含む当該認識処理に関わる各種の処理（たとえば差分抽出部132からデータ処理対象追記情報特定処理部150までの各機能部の各処理）における、認識処理の信頼度に関する特徴量をそれぞれ取得し、この各特徴量に基づいて各処理の認識処理に関わるサブ信頼度をそれぞれ算出し、この算出したサブ信頼度に基づいて注目する付加情報に関しての最終的な信頼度を特定し、この最終的な信頼度が一定水準よりも低いかなかを判定することで難認識情報を特定する。

【0170】

つまり、分離認識処理部155や記入位置認識部158、168における認識処理時の信頼度だけでなく、その前段の各機能部における各処理時の認識率や信頼度に関わる情報（特徴量）をも収集し、収集した各特徴量に基づいて各処理別のサブ信頼度を算出し、この算出した処理別のサブ信頼度に基づいて最終的な信頼度を求める。そして、この最終的な信頼度を元にして修正を要する難認識情報を特定するのである。

【0171】

特開2004-152115号公報に記載の仕組みでは、取得したデータに含まれる複数の項目について項目ごとに内容の確信度を計算し、計算された確信度を用いて修正提示の方法を動的に変化させるようにしているが、この場合、各項目の認識時にパターン認識の類似度を使用して「確信度」を算出するのみであるため、認識処理よりも前段の各種の前処理などで生じ得る、たとえば局所的なノイズの有無など、認識処理に影響を与える事象を考慮することができない。

【0172】

これに対して、本実施形態の難認識追記情報抽出部194では、データ処理対象追記情報特定処理部150よりも前段の各機能部における認識処理と関連する特徴量をも取得して総合的に類似度を算出することで難認識情報を特定するので、データ処理対象追記情報特定処理部150における認識の類似度やこの類似度に基づく確信度に基づいて難認識情報を特定する場合よりも、高精度に難認識情報を特定することができる。

【0173】

すなわち、データ処理対象追記情報特定処理部150における認識処理の類似度だけでなく、データ処理対象追記情報特定処理部150よりも前段の処理をも含む各種の処理における、認識処理の信頼度に関する特徴量をそれぞれ取得し、この各特徴量に基づいて各処理の認識処理に関わるサブ信頼度をそれぞれ算出し、この算出したサブ信頼度に基づいて注目する付加情報に関しての最終的な信頼度を特定し、この最終的な信頼度が一定水準よりも低いかなかを判定することで難認識情報を特定するようにしているので、データ処理対象追記情報特定処理部150での認識処理よりも前段の各種の前処理などで生じ得る

10

20

30

40

50

、局所的なノイズの有無などをも考慮して難認識情報を精度よく特定することができるのである。

【 0 1 7 4 】

< 履歴収集処理 >



なお、各機能部における各処理別のサブ信頼度を算出する際には、分離認識処理部 1 5 5 における認識処理結果のデータベース登録に関わる全行程から、その前段の各機能部における各処理時の認識率や信頼度に関わる特徴量を収集する。

【 0 1 7 5 】

また、認識率や信頼度は、追記情報そのものの種類（文字なのか図形なのか線なのか）や形状によって傾向が異なることもあり、また、処理対象の追記済文書 8 B の状態とも関わりを持つこともあるので、これらの点を考慮して最終的な信頼度の取得ができるように、過去の統計情報を取っておき、サブ信頼度と過去の統計情報とに基づいて信頼度を算出する。

10

【 0 1 7 6 】

たとえば、過去の自動データ処理としては、概ね、「」の認識率（信頼度）は 0 . 9 5 で、「」の認識率（信頼度）は 0 . 9 であつたというように、採点記号 8 7 の別に、認識率（信頼度）が異なることがある。

【 0 1 7 7 】

また、追記済文書 8 B が写真や図形（クリップアート）を含むか否かで認識率（信頼度）は影響を受ける。特に、写真や図形（クリップアート）と接触する部分では間違い易く、写真や図形と接触する部分とそれ以外の部分とで認識率（信頼度）に大きな違いが生じ得る。このため、採点記号 8 7 やコメント 8 8 などの追記情報が、追記済文書 8 B（文書原本 8 A）としての付加情報記入済教材 8 1（教育用教材 8 0）の何処の部分に記入されるものであるかを区別できるようにする。

20

【 0 1 7 8 】

たとえば、図 5 は、各処理時の認識率や信頼度に関わる特徴量を収集する手法の一例を示す図である。この図 5 に示す例では、写真を用いた問題（問 4）が混在しており、その解答欄 8 4 が写真部分に近接して配置されているので、たとえば、問 4 の採点記号 8 7 の信頼度は「0 . 7」で、その他の解答欄 8 4 の採点記号 8 7 の信頼度は「0 . 9 ~ 0 . 9 5」であるなど、追記済文書 8 B（付加情報記入済教材 8 1）の問題ごとに認識率（信頼度）の履歴を取っておく。

30

【 0 1 7 9 】

< 信頼度情報収集処理と信頼度統合処理 >

また、難認識追記情報抽出部 1 9 4 は、各処理別のサブ信頼度に基づいて最終的な信頼度を求めるに当たっては、つまり、処理別のサブ信頼度を統合するに当たっては、たとえば、注目する追記情報（採点記号 8 7 やコメント 8 8）について、処理間で最も信頼度が悪いものを、最終的な信頼度とする第 1 の統合手法を採用することが考えられる。全ての追記情報についてこの第 1 の統合手法を繰り返すことで、サブ信頼度の値が一定水準よりも低い追記情報が抽出される。なお、各処理別のサブ信頼度を算出する手法は、後述する第 3 の統合手法にて説明する。

40

【 0 1 8 0 】

認識性能情報提示部 1 9 8 は、難認識追記情報抽出部 1 9 4 により抽出されたサブ信頼度の値が一定水準よりも低い追記情報について修正を促すように認識性能情報を記入者に提示する。認識性能が一定水準に満たない追記情報のみについて記入者に修正が促されるので、全ての追記情報について修正の可否を判断（チェック）する必要がなく、誤認識を起し得る追記情報の修正作業を効率化することができる。

【 0 1 8 1 】

この第 1 の統合手法を採用すれば、1 つの処理で認識結果に重大な影響を与える誤りを犯してしまう場合に、対応できる。たとえば、三角図形の下半分が写真と重なってしまい、差分抽出処理によって下半分が抽出できなかった場合、記入の仕方によってはバツ図形に見

50

えるため、後段の処理ではバツと「確信を持って」判定してしまう可能性がある。しかし、この第1の統合手法を採れば、差分抽出処理でのサブ信頼度を利用できるため、このような誤りを回避できる効果がある。

【0182】

また、難認識追記情報抽出部194は、注目する追記情報（採点記号87やコメント88）について、処理別のサブ信頼度（複数の処理結果）に重付けをして合計し、最終的な信頼度を算出する第2の統合手法を採ることも考えられる。全ての追記情報についてこの第2の統合手法を繰り返すことで、最終的な信頼度の値が一定水準よりも低い追記情報が抽出される。なお、各処理別のサブ信頼度を「重付け」を考慮して算出する手法は、後述する第3の統合手法にて説明する。

10

【0183】

重付けの一例としては、処理順に拘わらず所定の値をランダムに設定することもできるし、上流に行くほど重付けを重くする、あるいは逆に下流に行くほど重付けを重くするというように、処理順に応じて重付け値を漸次変化させることもできる。

【0184】

上流に行くほど重付けを重くする手法では、初期の処理を重視するため、図5のような処理の上流段階に認識に大きな影響を与えることが想定される場合に対応できる効果があり、下流に行くほど重付けを重くする手法では、歪みなどの補正後の処理を重視するため、元々の記入図形の形状が主因の場合（丸みのある三角など）に対応できる効果がある。

【0185】

20

重付けの別の例としては、特徴量と正しい認識結果のデータ群が有る場合に、重回帰分析によりそのデータ群に最適な重み付けを算出することができる。この手法では、データ群とこれから入力される追記情報が似ている場合（記入者、文書原本が同じなど）には、精度が高い利点がある。

【0186】

認識性能情報提示部198は、難認識追記情報抽出部194により重付けを考慮して抽出された最終的な信頼度の値が一定水準よりも低い追記情報について修正を促すように認識性能情報を記入者に提示する。認識性能が一定水準に満たない追記情報のみについて記入者に修正が促されるので、全ての追記情報について修正の要否を判断（チェック）する必要がなく、誤認識を起し得る追記情報の修正作業を効率化することができる。

30

【0187】

この第2の統合手法を採れば、重付けを自由に変更出来るため、記入者や文書原本などの特徴に合わせて精度が高くなるように調整できる効果がある。

【0188】

また、難認識追記情報抽出部194は、信頼度が低くなる誤認識のタイプ（ミスのタイプ）を予め設定しておき、データ処理時に、この設定した「誤認識のタイプ」に該当するものの内、信頼度が最も低いタイプのものを最終的な信頼度とする第3の統合手法を採ることも考えられる。この際には、誤認識のタイプごとに、判定時に使用する信頼度の組合せと重付けを設定しておく。そして、推定される誤認識のタイプを元にして、処理や特徴量の組合せを選び、重付け加算して信頼度を算出する。そして、加算処理結果に基づいて、信頼度が所定水準よりも低い、あるいは最も低い誤認識のタイプを選ぶ。これにより、難認識追記情報抽出部194は、選択した誤認識のタイプを持つ、認識処理の信頼度が一定水準よりも低い難認識情報を特定することができる。

40

【0189】

ここで、「推定される誤認識のタイプを元にして、処理や特徴量の組合せを選び、重付けして信頼度を算出する」には、あるタイプに特有に現れる特徴量の組合せ（特徴量の特徴）を予め規定しておき、それに当てはまるかどうかを調べる。単純には、各処理で、あるタイプに特有の特徴量に当てはまれば1、当てはまらなければ0として加算し、ある水準を満たすかどうかを見る。

【0190】

50

図6～図10は、認識性能情報提示処理部190の処理であって、各処理と関わる誤認識のタイプを説明する図である。「誤認識のタイプ」の判定としては、たとえば、文書入力装置20と関わりを持つスキャンミス、差分抽出部132やデータ処理対象追記情報抽出部142と関わりを持つ抽出ミス、抽出線分途切れ補正部148と関わりを持つ欠損補間ミス、変形処理部156a, 166a, 157a, 167aと関わりを持つ形状変形、記入位置認識部158, 168と関わりを持つ位置ずれといった各機能部の処理そのものと密接に関係したもののや、複数追記や追記訂正、ペン色・太さ・かすれといった追記情報そのものに起因するものや、紙汚れなどの追記済文書8Bそのものに起因するもの、などを判定するのがよい。

【0191】

たとえば、「抽出ミス」の誤認識のタイプは、形状認識や文字認識の信頼度が低く、かつ濃い背景への重複量が大いとき起こり易い。これは、差分抽出部132での差分抽出処理時に、追記済文書8Bと文書原本8Aとの各画像の位置関係などから追記情報と原本画像との接触の有無によって影響を受け得ることがあるし、また追記情報が写真や図形などの濃い背景に重複しているときにも影響を受け得るからである。また、データ処理対象追記情報抽出部142でのデータ処理用の追記情報の抽出時に、追記に使用されたペン色(たとえば赤ペン)の成分を抽出しようとするときに、注目する色と実際のペン色との相違の影響を受け得るし、赤抽出画像と原本画像との位置関係などから原本画像との接触の有無によっても影響を受け得るからである。

【0192】

たとえば、図6(A)に示すように、追記情報抽出処理の対象画像が、濃い背景に重なって採点記号87が記入されていると、差分抽出部132やデータ処理対象追記情報抽出部142での追記情報の抽出結果は、図6(B)に示すように、重なり部分が欠損して抽出される。抽出線分途切れ補正部148は、その欠損部分が繋がるように補正する。

【0193】

すなわち、抽出線分途切れ補正部148は、データ処理対象追記情報抽出部142による所定色成分の抽出結果、すなわち「」や「x」などの図形であるはずの抽出結果に対して、細線化処理を実行し、さらに端点抽出処理を実行する。これにより、「」や「x」などの図形に途切れ部分が生じている場合に、その途切れ部分における端点が抽出されることになる。そして、端点を抽出したら、その抽出した全ての端点の内、所定のものを接続する。

【0194】

これにより、たとえば図6(B)に示す図形が抽出された場合には、端点Aに対して、所定距離内に端点B, Cが存在していても、その中で最近傍の端点Bを端点Aと接続することで、図6(C)に示すように、「」の画像における途切れ部分を補正する。その結果、記入者は、本来、概ね「」を記入しているにも拘わらず、濃い背景と重なる部分が欠落された歪んだ状態の「」の画像が分離認識処理部155に入力されることになる。

【0195】

その結果、分離認識処理部155は、歪んだ状態の「」の画像に基づいて図形認識処理もしくは文字認識処理によって「」と特定すべきであるが、その際には、重なり部分の欠落によって図形の特徴が大幅に変わってしまうため、認識率(信頼度)が低下してしまい、また認識ミスが生じる可能性が高くなる。

【0196】

このため、図6(D)に示すように、難認識追記情報抽出部194にて、このような「抽出ミス」の誤認識のタイプを特定することができるように、差分抽出部132での差分抽出処理やデータ処理対象追記情報抽出部142での特定色成分抽出処理においては、抽出量などから求められる抽出時の各品質情報J10, J11、あるいは、採点記号87やコメント88の濃い背景への重複量J12、赤抽出画像と原本画像との位置関係などから求められる原本画像と赤抽出画像の接触量J13などを、認識率や信頼度に関わる特徴量として求め、これらの情報を難認識追記情報抽出部194に通知する。

10

20

30

40

50

【 0 1 9 7 】

難認識追記情報抽出部 1 9 4 は、通知された品質情報 J 1 0 , J 1 1、重複量 J 1 2、接触量 J 1 3 に基づいて、たとえば、それぞれ基準を満たせば 1、満たさなければ 0 として加算することで、差分抽出処理のサブ信頼度 T 1 3 2 や特定色成分抽出処理のサブ信頼度 T 1 4 2 を算出する。

【 0 1 9 8 】

また、「欠損補間ミス」の誤認識のタイプは、原本画像との接触量が多く、かつ複数の対応付けがなされたときに起こり易い。抽出線分途切れ補正部 1 4 8 における途切れ補正処理時に、端点同士の接続処理時に、端点間の距離の影響を受け得るからである。なお、複数の対応付けがなされるケースとしては、元々 1 つの画像が原本画像との重なり部分によって分断されるケースと、元々 2 つの画像が原本画像との重なり部分を途切れ補正することによって接続されるケースとがある。

10

【 0 1 9 9 】

元々 1 つの画像が原本画像との重なり部分によって分断されるケースとしては、たとえば、図 7 に示す事例がある。図 7 (A) に示すように、追記情報抽出処理の対象画像が、線分などの比較的小さな濃い背景に重なって採点記号 8 7 が記入されていると、差分抽出部 1 3 2 やデータ処理対象追記情報抽出部 1 4 2 での追記情報の抽出結果は、図 7 (B) に示すように、その重なり部分が欠損して抽出される。抽出線分途切れ補正部 1 4 8 は、その欠損部分が繋がるように補正する。

【 0 2 0 0 】

すなわち、抽出線分途切れ補正部 1 4 8 は、データ処理対象追記情報抽出部 1 4 2 による所定色成分の抽出結果、すなわち「 」や「 x 」などの図形であるはずの抽出結果に対して、細線化処理を実行し、さらに端点抽出処理を実行する。これにより、「 」や「 x 」などの図形に途切れ部分が生じている場合に、その途切れ部分における端点が抽出されることになる。そして、端点を抽出したら、その抽出した全ての端点の内、所定のものを接続する。

20

【 0 2 0 1 】

たとえば図 7 (B) に示す図形が抽出された場合には、「 」の画像が、線分などの比較的小さな濃い背景の重なり部分で分断され、4 つの端点 A ~ D が現われる。抽出線分途切れ補正部 1 4 8 は、それらの端点を所定の条件に基づいて接続しようとする。この際、最近傍の端点同士を接続することを基本条件としていても、何らかの原因で、必ずしもそのように端点が接続されないことも起こり得る。この現象は、特に、文字や図形を示す 1 つの画像が、重なり部分によって分断されることで、複数の対応付けがなされた場合に起こり易い。元々 1 つの画像が分断されたものであるのか、元々別のものではあったのかを切り分けることが困難であり、分断されている画像ごとに端点同士を接続しようとするからである。

30

【 0 2 0 2 】

よって、たとえば、図 7 (C 1) に示すように、分断された上半分の端点 A と端点 C とを接続する一方で、分断された下半分の端点 B と端点 D とは接続しない場合や、図 7 (C 2) に示すように、分断された上半分の端点 A と端点 C とを接続する一方で、分断された下半分の端点 B と端点 D も接続する場合もある。

40

【 0 2 0 3 】

その結果、何れにしても、記入者は、本来、概ね「 」を記入しているにも拘わらず、濃い背景と重なる部分が欠落された歪んだ状態の（概ね半円に近い）「 」の画像が分離認識処理部 1 5 5 に入力されることになる。

【 0 2 0 4 】





その結果、分離認識処理部 1 5 5 は、歪んだ状態の「 」の画像に基づいて図形認識処理もしくは文字認識処理によって「 」と特定すべきであるが、その際には、原本画像との重なり部分の欠落によって図形の特徴が大幅に変わってしまうため、認識率（信頼度）が低下してしまい、また認識ミスが生じる可能性が高くなる。

50


【0205】

また、元々2つの画像が原本画像との重なり部分を途切れ補正することによって接続されるケースとしては、たとえば、図8に示す事例がある。図8(A)に示すように、追記情報抽出処理の対象画像が、解答欄84の枠などの濃い背景に重なって複数の採点記号87が記入されていると、差分抽出部132やデータ処理対象追記情報抽出部142での追記情報の抽出結果は、図8(B)に示すように、それぞれの採点記号87は重なり部分が欠損して抽出される。抽出線分途切れ補正部148は、その欠損部分が繋がるように補正する。



【0206】

すなわち、抽出線分途切れ補正部148は、データ処理対象追記情報抽出部142による所定色成分の抽出結果、すなわち「」や「」などの図形であるはずの抽出結果に対して、細線化処理を実行し、さらに端点抽出処理を実行する。これにより、「」や「」などの図形に途切れ部分が生じている場合に、その途切れ部分における端点が抽出されることになる。そして、端点を抽出したら、その抽出した全ての端点の内、所定のものに接続する。


【0207】

たとえば図8(B)に示す図形が抽出された場合には、2つの「」の画像が、それぞれ枠線などの比較的小さな濃い背景の重なり部分で分断され、それぞれ4つの端点A1~D1, A2~D2が現われる。抽出線分途切れ補正部148は、それらの端点を所定の条件に基づいて接続しようとする。


【0208】

ここで、複数の画像成分(分断されたものであるか否かを問わない)に発生する端点間の距離が、所定値以下であるか否かに基づいて、元々1つの画像が分断されたものであるのか、元々別のものではあったのかを切り分けるようにし、かつ、最近傍の端点同士を接続することを基本条件としていると、先ず、それぞれの「」の画像の抽出部分は、図8(C)に示すように、最近傍の端点同士を接続することで、元の「」がほぼ再現される。



【0209】

しかしながら、2つの「」の画像成分に関しては、何らかの原因で、さらに端点同士の接続がなされることがある。この現象は、元々複数の画像が、それぞれ同じ背景によって分断されることで複数の対応付けがなされた場合に起こり易い。元々複数の画像であっても、元々1つの画像が分断されたものであるのか、元々別のものではあったのかを切り分けることが困難であり、分断されている画像との間で端点同士を接続しようとすることが起こるからである。



【0210】

よって、たとえば、図8(C)に示すように、右側と左側の2つの「」の画像成分が、枠との重なり部分を補正しようとする線分によって接続され、全体として1つの画像になってしまう。

【0211】

その結果、記入者は、本来、それぞれを概ね「」で区別して記入しているにも拘わらず、濃い背景と重なる部分が線分で接続され、歪んだ状態の「」の(2つが1つに繋がった)画像が分離認識処理部155に入力されることになる。

【0212】

その結果、分離認識処理部155は、歪んだ状態の「」の画像に基づいて図形認識処理もしくは文字認識処理によってそれぞれを「」と区別して特定すべきであるが、その際には、原本画像との重なり部分を補正する線分によって図形の特徴が大幅に変わってしまうため、認識率(信頼度)が低下してしまい、また、認識ミスが生じる可能性が高くなる。

【0213】

このため、図7(D), 図8(D)に示すように、難認識追記情報抽出部194にて、このような「欠損補間ミス」の誤認識のタイプを特定することができるよう、抽出線分

10

20

30

40

50

途切れ補正部 148 での途切れ補正処理においては、途切れ補正処理時の補間長さを示す補間長情報 J20 や接続の候補となる端点数 J21 などを、認識率や信頼度に関わる特徴量として求め、補間長情報 J20 などを難認識追記情報抽出部 194 に通知する。

【0214】

難認識追記情報抽出部 194 は、通知された補間長情報 J20 などに基づいて、たとえば、それぞれ基準を満たせば 1、満たさなければ 0 として加算することで、途切れ補正処理のサブ信頼度 T148 を算出する。

【0215】

また、「複数追記」の誤認識のタイプは、形状認識や文字認識の信頼度が低く、隣接する他の記入欄との距離が近くて複数の採点記号 87 やコメント 88 の対応付けがあるときに起こり易い。これは、分離認識処理部 155 の図形形状認識部 156, 166 における図形認識処理時や文字認識処理部 157, 167 における文字認識処理時に、複数の追記情報同士の接触の有無が認識の信頼度に影響を与えることに基づくものである。

10

【0216】

たとえば、図 9 (A) に示すように、2つの解答欄 84 が 2行で記載され、その間の距離が十分でないときに、各解答欄 84 に記入された生徒解答に対してそれぞれ採点記号 87 が採点官によって記入されると、十分な間隔がないために、2つの採点記号 87 が接触してしまうことが典型例である。

【0217】

このため、図 9 (B) に示すように、難認識追記情報抽出部 194 にて、このような「複数追記」の誤認識のタイプを特定することができるように、分離認識処理部 155 においては、先ず採点記号 87 同士の接触の有無を示す接触情報 J30 とコメント 88 同士の接触の有無を示す接触情報 J31 を求め、さらに、図形形状認識部 156 での図形認識処理においては、採点記号 87 やコメント 88 についての図形認識処理結果の認識の信頼度情報 J32, J33 を、採点記号 87 やコメント 88 についての図形認識率や図形の信頼度に関わる特徴量として求める。また、文字認識処理部 157, 167 での文字認識処理においては、採点記号 87 やコメント 88 についての文字認識処理結果の認識の信頼度情報 J34, J35 を、採点記号 87 やコメント 88 についての文字認識率や文字の信頼度に関わる特徴量として求め、これらの情報を難認識追記情報抽出部 194 に通知する。

20

【0218】

難認識追記情報抽出部 194 は、通知された採点記号 87 についての接触情報 J30, 信頼度情報 J32, 信頼度情報 J34 に基づいて、たとえば、それぞれ基準を満たせば 1、満たさなければ 0 として加算することで、図形認識処理のサブ信頼度 T156 と文字認識処理のサブ信頼度 T157 とを算出する。また、コメント 88 についての接触情報 J31, 信頼度情報 J33, 信頼度情報 J35 に基づいて、たとえば、それぞれ基準を満たせば 1、満たさなければ 0 として加算することで、図形認識処理のサブ信頼度 T168 と文字認識処理のサブ信頼度 T167 とを算出する。

30

【0219】

また、「追記訂正」の誤認識のタイプは、形状認識や文字認識の信頼度は高いが、記入欄との距離が遠くて、1つの記入欄に対して複数の採点記号 87 やコメント 88 の対応付けがあるときに起こり易い。これは、データ処理部 170 にて、注目する追記情報について、記入位置認識部 158, 168 により特定された記入位置の特定処理結果に基づいて、何れの記入欄に対応するものであるかを特定してデータ処理を実行していく際に、1つの追記情報が複数の記入位置に関係するか否かがデータ処理結果に影響を与えることに基づくものである。

40

【0220】

たとえば、図 10 (A), (B) に示すように、2つの解答欄 84 が 2行で記載され、その間の距離がある程度広がっているときに、各解答欄 84 に記入された生徒解答に対してそれぞれ採点記号 87 が採点官によって記入されると、間隔が広いために、2つの解答欄 84 の間に、何れか一方用の 1つの採点記号 87 が記入されることが典型例である。

50

【 0 2 2 1 】

ここで、図 1 0 (A) に示すように、追記情報が大きく、上下の解答欄 8 4 に跨がって記入されるケースでは、追記情報の外接矩形とそれぞれの解答欄 8 4 となる領域との重なり面積が概ね同じであるし、各解答欄 8 4 との間の各距離差も概ね同じであるので、2 つの解答欄 8 4 の間に記入された追記情報が何れの解答欄 8 4 に対応させるべきかを特定することが困難になる。換言すれば、1 つの解答欄 8 4 に対して、複数の採点記号 8 7 が対応付けされる可能性が生じる。

【 0 2 2 2 】

また、図 1 0 (B) に示すように、追記情報が小さく、上下の解答欄 8 4 の間に収まって記入されるケースでは、追記情報の外接矩形とそれぞれの解答欄 8 4 となる領域との重なり面積が概ねゼロであるし、各解答欄 8 4 との間の各距離差も概ね同じであるので、2 つの解答欄 8 4 の間に記入された追記情報が何れの解答欄 8 4 に対応させるべきかを特定することが困難になる。換言すれば、1 つの解答欄 8 4 に対して、複数の採点記号 8 7 が対応付けされる可能性が生じる。

10

【 0 2 2 3 】

このため、図 1 0 (C) に示すように、難認識追記情報抽出部 1 9 4 にて、このような「追記訂正」の誤認識のタイプを特定することができるように、記入位置認識部 1 5 8 , 1 6 8 で記入位置認識処理においては、1 つの追記情報の複数の記入欄への対応付けの可能性あるいは1 つの記入欄に対しての複数の追記情報が対応付けされる可能性の有無を示す複数追記情報 J 4 0 , J 4 1 や、各追記情報と各記入欄との距離を示す距離情報 J 4 2 , J 4 3 などを、認識率や信頼度に関わる特徴量として求め、これらの情報を難認識追記情報抽出部 1 9 4 に通知する。

20

【 0 2 2 4 】

難認識追記情報抽出部 1 9 4 は、通知された情報 J 4 0 , J 4 2 に基づいて、たとえば、それぞれ基準を満たせば 1、満たさなければ 0 として加算することで、記入位置認識部 1 5 8 における採点記号 8 7 についての記入位置認識処理のサブ信頼度 T 1 5 8 を算出し、また通知された情報 J 4 1 , J 4 3 に基づいて、たとえば、それぞれ基準を満たせば 1、満たさなければ 0 として加算することで、記入位置認識部 1 6 8 におけるコメント 8 8 についての記入位置認識処理のサブ信頼度 T 1 6 8 を算出する。

【 0 2 2 5 】

難認識追記情報抽出部 1 9 4 は、注目する追記情報（採点記号 8 7 やコメント 8 8）について、それぞれ求めた各処理のサブ信頼度の内で、サブ信頼度の値が一定水準よりも低い誤認識のタイプを抽出する。この際には、サブ信頼度が最も低い誤認識のタイプのみを抽出してもよい。全ての追記情報についてこの第 3 の統合手法を繰り返すことで、サブ信頼度の値が一定水準よりも低い誤認識のタイプを持つ追記情報が抽出される。

30

【 0 2 2 6 】

認識性能情報提示部 1 9 8 は、難認識追記情報抽出部 1 9 4 により抽出されたサブ信頼度の値が一定水準よりも低い誤認識のタイプを持つ追記情報について修正を促すように認識性能情報を記入者に提示する。

【 0 2 2 7 】

認識性能が一定水準に満たない誤認識のタイプを持つ追記情報のみについて記入者に修正が促されるので、全ての追記情報について修正の要否を判断（チェック）する必要がなく、誤認識を起し得る追記情報の修正作業を効率化することができる。信頼度が所定水準よりも低い誤認識のタイプが提示されると、記入者は、その誤認識のタイプの側面から現状の記入状態を改善することができる。1 つの追記情報について、信頼度が所定水準よりも低い誤認識のタイプが 2 つ提示されたときには、記入者は、それぞれの誤認識のタイプの側面から現状の記入状態を改善することができるし、信頼度が最も低い誤認識のタイプのみが提示されたときには、記入者は、その最も信頼度の低い誤認識のタイプの側面から現状の記入状態を改善することができる。

40

【 0 2 2 8 】

50

このような第3の統合手法を採れば、タイプごとに具体的な改善方法を準備しておけるので、記入者が記入方法を注意し易くなり改善効果が向上する効果がある。

【0229】

< 認識性能情報提示手法 >

図11～図16は、認識性能情報提示部198における、認識性能を向上させるための認識性能情報をユーザに提示する手法を説明する図である。

【0230】

認識性能情報提示部198が認識性能情報をユーザ（記入者）に提示する際には、記入者のそれぞれに応じた認識性能情報を提示するのがよい。また、その際の提示手法としては、たとえば、図11に示すように、認識性能情報提示部198は、信頼度が一定水準以下のものについて、さらに信頼度によって認識性能情報の提示方法を変えて提示する第1の提示手法を採ることが考えられる。なお、ここでの「信頼度」は、最終的な信頼度を意味する。

10

【0231】

ここで「信頼度によって表示方法を変える」という第1の提示手法は、最終的な信頼度のレベルや誤認識のタイプを目視で容易に区別することができるようにすることを意図したものであり、その限りにおいて、様々な提示手法による区別ができる。たとえば、最終的な信頼度や誤認識のタイプに応じて、それぞれ異なるグラディエーションを掛けて表示することができる。またたとえば、信頼度が一定水準以下のものについて枠を付ける、さらに信頼度の低いものほど枠線を太くする手法を採ることができる。

20

【0232】

表示方法の違いによって、認識特定した全ての追記情報の中で、信頼度が一定水準よりも低いものがどれであるかを閲覧者（記入者）は判断することができる。よって、この第1の提示手法を採用した場合でも、事実上、信頼度の値が一定水準よりも低い追記情報についてのみ修正を促すように認識性能情報を記入者に提示することになる。記入者は、その信頼度の値が一定水準よりも低い追記情報について記入態様を改善することができる。

【0233】

この第1の提示手法を採れば、修正が必要なものほど目に付き易くすることで、重大な見落としを防止できる。また、信頼度の違いを見せることで、どのような図形が認識に適さないのかを比較して学習することができるため、記入方法の改善に繋がる効果がある。

30

【0234】

また、認識性能情報提示部198は、認識処理された全ての追記情報について、信頼度が一定水準よりも低い難認識情報のみに関して認識性能を向上させるための認識性能情報を提示する第2の提示手法を採ることが考えられる。たとえば、難認識追記情報抽出部194によって第1や第2の統合手法に基づいて特定される最終的な信頼度が所定水準よりも低い追記情報や、第3の統合手法に基づいて特定される信頼度が所定水準よりも低いあるいは最も低い誤認識のタイプを持つ追記情報のみに関して認識性能情報を提示する。

【0235】

ここで、「信頼度が一定水準よりも低い難認識情報のみに関して認識性能を向上させるための認識性能情報を提示する」とは、要するに、認識処理された全ての付加情報について、「信頼度が低い難認識情報」と「その他の追記情報」とを区別して提示することで、「信頼度が低い難認識情報」を確実に特定できるように認識性能情報を提示することを意味する。たとえば、「信頼度が低い難認識情報」以外の情報は全て表示しないようにすることもできる。

40

【0236】

信頼度の値が一定水準よりも低い難認識情報についてのみ修正を促すように認識性能情報を記入者に提示することができ、記入者は、その信頼度の値が一定水準よりも低い難認識情報について記入態様を改善することができる。

【0237】

この第2の提示手法を採れば、信頼度が一定水準よりも低い難認識情報のみに関して認

50

識性能を向上させるための認識性能情報を提示するため、修正の必要性が高い情報だけを素早く確認でき、効率的に修正作業ができる効果がある。

【 0 2 3 8 】

なお、第 2 の提示手法を採る場合において、認識性能が悪く修正を要するものとして認識性能情報の提示対象として抽出した追記情報に関しては、この追記情報が一目で何処の部分であるかを認識できるような状態で強調表示するのがよい。強調表示によって、信頼度が一定水準よりも低い難認識情報とその他の情報とを区別して提示することができる。

【 0 2 3 9 】

たとえば、図 1 2 の各図に示すように、追記情報を枠で囲む（図 1 2（A））、追記情報に矢印（あるいはその他のマーク）を付ける（図 1 2（B））、追記情報に電子付箋を付ける（図 1 3）、あるいは、図示を割愛するが、信頼度によって追記情報の太さを変えて表示する、信頼度によって追記情報の色を変えて表示する、などして、信頼度が一定水準よりも低い難認識情報とその他の情報とを区別して表示するとよい。

10

【 0 2 4 0 】

枠で囲む表示態様では範囲が特定し易いため、特に複数の図形が図 8（C）のように接続してしまった場合に視認し易い効果がある。また、矢印やその他のマークを付ける表示態様では大きさを自由に設定できるため文書原本の記載などに被ってしまう可能性が低く、記入が密な場合でもあまり視認性が低下しない効果がある。電子付箋を付ける表示態様では複数文書を重ねた場合でも最上段以外の付箋を視認できるため、信頼度が低い追記情報の全体量や位置などが確認し易い効果がある。

20

【 0 2 4 1 】

あるいは、認識性能が悪く修正を要するものとして認識性能情報の提示対象として抽出した追記情報に関して、図 1 4 に示すように、リンクの一覧を作成して表示するとともに、この一覧中の何れかがクリックされたときには、付加情報記入済教材 8 1 を表示し、かつ該当箇所の追記情報を表示するよう表示してもよい。この際には、該当箇所の追記情報を図 1 2 に示した強調表示を採用して表示するとよい。

【 0 2 4 2 】

リンクの一覧では、信頼度が一定水準以下の採点記号 8 7 やコメント 8 8 が記入される解答欄 8 4 と対応する問題文のように、難認識情報に辿り着くことのできる情報のみが提示される。このリンクの一覧をベースとして、所要の難認識情報を表示させてから修正作業時に取り掛かることができる。

30

【 0 2 4 3 】

付加情報記入済教材 8 1 上では難認識情報だけでなく全ての追記情報が提示されているが、リンクの一覧では、難認識情報に辿り着くことのできる情報のみを提示しており、このような態様も、「信頼度が一定水準よりも低い難認識情報のみに関して認識性能を向上させるための認識性能情報を提示する」に含むものとする。

【 0 2 4 4 】

あるいは、認識性能が悪く修正を要するものとして認識性能情報の提示対象として抽出した追記情報に関して、図 1 5 に示すように、付加情報記入済教材 8 1 を表示し、かつ、次にジャンプするボタンのクリック、「次に送る」ボタンのクリック、ダブルクリックなどの予め決められたジェスチャ、あるいは予め決められたキー操作（矢印キーなど）、などを契機とすることで、認識性能が悪く修正を要する複数の追記情報を、同一の付加情報記入済教材 8 1 内であるいは別の付加情報記入済教材 8 1 へと、次々とジャンプできるようにして表示してもよい。この際には、該当箇所（ジャンプ先）の追記情報を、図 1 2 に示した強調表示を採用して表示するとよい。

40

【 0 2 4 5 】

また、認識性能情報提示部 1 9 8 は、図 1 6 に示すように、信頼度が一定水準に満たない修正を要する追記情報（難認識情報）について、信頼度とともにそのように判定した判定理由を認識性能情報として提示する第 3 の提示手法を採ることが考えられる。

【 0 2 4 6 】

50

なお、ここでの「信頼度」は、最終的な信頼度を意味する。また、「判定理由」は、たとえば、難認識追記情報抽出部 194 に通知される各種のサブ信頼度を利用して、サブ信頼度の悪い処理を判定理由として特定するのがよい。

【0247】

この第3の提示手法を採れば、信頼度の値が一定水準よりも低い追記情報を通知するだけでなく、そのように判定した理由も認識性能情報として記入者に提示することができ、記入者は、その信頼度の値が一定水準よりも低い追記情報について、提示された理由に応じた記入態様に改善することができる。どの点に注意して記入すればよいかを具体的に把握することができるので、記入方法の習得に利用することができる。手書き入力情報の形状などの不適切な部分を確実に把握し、正確な図形や文字の形状などを確実に学習することができるのである。

10

【0248】

なお、上述の第1～第3の提示手法の何れにおいても、各種の処理過程における認識率に関わる情報を収集し、収集した情報に基づいて信頼度の値が一定水準よりも低い追記情報を特定し、その信頼度の値が一定水準よりも低い追記情報に関して、手書き入力情報の認識性能を向上させるための認識性能情報、つまり現状の追記情報について修正を促す認識性能情報を提示するようにしたので、全ての追記情報をチェックする必要がなく、誤認識をもたらすような追記情報の記載を修正する作業を効率的に実行することができる。

【0249】

また、付加情報記入済教材81の表示中に、認識の信頼度が一致水準より低い現状の追記情報（難認識情報）について修正を促す認識性能情報を提示するようにしているので、特開2000-105796号公報に記載の仕組みとは異なり、訂正専用の画面を作成する必要がない利点もある。

20

【0250】

<信頼度情報収集提示処理；具体例>

図17および図18は、文書原本8Aの一例である教育用教材80を処理対象とする具体的な信頼度情報収集提示処理を示した図である。ここで、図17は、その全体概要をシステム構成図と対応付けて示しており、また図18は、信頼度情報収集提示処理手順を示すフローチャートである。

【0251】

ここでは、教育用教材80としての答案を作成し、答案を用いた試験後に、採点記号87に基づく自動採点処理やコメント88に基づくコメント分類処理を行なう例で示す。

30

【0252】

試験を実行する際には、文書原本情報データベースDB2から教育用教材80を読み出してプリンタで印刷し生徒や受験者に配布する(S206)。そして、試験終了後、採点官は採点記号87やコメント88を生徒の解答に対して追記する(S208)。

【0253】

自動データ処理時には、まず、文書入力装置20は、採点官により採点記号87やコメント88が記入された各解答者の付加情報記入済教材81を読み取り(S210)、その付加情報記入済教材81を表わす画像データを追記情報処理装置10に入力する(S212)。文書入力装置20は、この文書入力装置20による画像読取りによって得られた画像データについて、一旦ワークエリアとして用いられるメモリなどに保持しておく。

40

【0254】

追記情報処理装置10(教材処理装置)は、付加情報記入済教材81の読取画像データを受け取り、ハーフトーン画像中に埋め込まれている文書原本特定コードを特定し(S224)、対応する教育用教材80の原本画像と記入欄位置領域情報38などの原本情報を文書原本情報データベースDB2として機能する文書管理サーバ30から入手する(S225)。

【0255】

そして、自動データ処理として、付加情報記入済教材81に記入されている採点記号8

50

7やコメント88を差分抽出部132やデータ処理用追記情報抽出処理部140にて抽出し(S242)、分離認識処理部155や記入位置認識部158,168にてその記入内容や記入位置を特定した上で(S262)、データ処理部170にて採点記号87に関する自動採点集計処理やコメント88に関する自動コメント分類処理などを実行する(S266)。

【0256】

この際、分離認識処理部155や記入位置認識部158,168あるいはデータ処理部170は、ユーザによる修正指示を受け付ける(S310)。この後、自動データ処理が完了すると(S264-YES)、データ処理部170は、採点・集計の結果やコメント分類結果を処理結果保存サーバ40に登録する(S268)。

10

【0257】

ここで、このような処理過程において、認識性能情報提示処理部190の難認識追記情報抽出部194は、各機能部から、それぞれ処理での認識率に関わる各特徴量を収集して(S300)、収集した各特徴量に基づいて、各処理についてのサブ信頼度を算出する(S302)。

【0258】

たとえば、差分抽出部132は、文書原本8Aの一例である教育用教材80と追記済文書8Bの一例である付加情報記入済教材81との間での差分抽出処理における品質情報J10と、採点記号87やコメント88の濃い背景への重複量J12を抽出して難認識追記情報抽出部194に通知する。また、データ処理用追記情報抽出処理部140のデータ処理対象追記情報抽出部142は、差分抽出部132で抽出された差分情報9の内データ処理対象の追記情報として採点記号87やコメント88を抽出する際の品質情報J11と、赤抽出画像と原本画像との位置関係などから求められる原本画像と赤抽出画像の接触量J13を抽出して難認識追記情報抽出部194に通知する。難認識追記情報抽出部194は、通知された品質情報J10,J11、重複量J12、接触量J13に基づいて、差分抽出処理のサブ信頼度T132と特定色成分抽出処理のサブ信頼度T142とを算出する。

20

【0259】

また、抽出線分途切れ補正部148は、途切れ補正処理時の補間長さを示す補間長情報J20や接続の候補となる端点数J21などを抽出し、難認識追記情報抽出部194に通知する。難認識追記情報抽出部194は、通知された補間長情報J20などに基づいて、途切れ補正処理(欠損補間)のサブ信頼度T148を算出する。

30

【0260】

また、分離認識処理部155は、先ず、採点記号87同士の接触の有無を示す接触情報J30とコメント88同士の接触の有無を示す接触情報J31を抽出し、難認識追記情報抽出部194に通知する。

【0261】

また、分離認識処理部155の図形形状認識部156は、採点記号87についての図形認識処理結果の認識の信頼度情報J32を抽出し、難認識追記情報抽出部194に通知するし、図形形状認識部166は、コメント88についての図形認識処理結果の認識の信頼度情報J33を抽出し、難認識追記情報抽出部194に通知する。

40

【0262】

また、分離認識処理部155の文字認識処理部157は、採点記号87についての文字認識処理結果の認識の信頼度情報J34を抽出し、難認識追記情報抽出部194に通知するし、文字認識処理部167は、コメント88についての文字認識処理結果の認識の信頼度情報J35を抽出し、難認識追記情報抽出部194に通知する。

【0263】

難認識追記情報抽出部194は、通知された採点記号87についての接触情報J30,信頼度情報J32,信頼度情報J34に基づいて、図形認識処理のサブ信頼度T156と文字認識処理のサブ信頼度T157とを算出し、また、コメント88についての接触情報J31,信頼度情報J33,信頼度情報J35に基づいて、図形認識処理のサブ信頼度T

50

168と文字認識処理のサブ信頼度T167とを算出する。

【0264】

また、記入位置認識部158は、採点記号87についての複数追記情報J40と距離情報J42とを抽出して難認識追記情報抽出部194に通知するし、記入位置認識部168は、コメント88についての複数追記情報J41と距離情報J43とを抽出して難認識追記情報抽出部194に通知する。

【0265】

難認識追記情報抽出部194は、通知された情報J40, J42に基づいて採点記号87についての記入位置認識処理のサブ信頼度T158を算出し、また通知された情報J41, J43に基づいてコメント88についての記入位置認識処理のサブ信頼度T168を算出する。

10

【0266】

難認識追記情報抽出部194は、さらに、前述のようにして求めた各処理のサブ信頼度と過去の統計情報とに基づいて、たとえば下記式(1)のようにして、処理ごとに重付けを変えて、最終的な信頼度Tfinalを算出する(S304)。

【0267】

【数1】

$$\begin{aligned}
 \text{信頼度}T_{\text{final}} &= (\text{差分抽出のサブ信頼度}) * \alpha 1 + \\
 &\quad (\text{欠損補間のサブ信頼度}) * \alpha 2 + \\
 &\quad (\text{認識処理のサブ信頼度}) * \alpha 3 + \\
 &\quad (\text{記入位置認識処理のサブ信頼度}) * \alpha 4 \\
 \\
 \text{採点記号の信頼度}T_{\text{final}} &= (T132+T142) * 0.3 + T148 * 0.2 + \\
 &\quad (T156+T157) * 0.4 + T158 * 0.1 \\
 \\
 \text{コメントの信頼度}T_{\text{final}} &= (T132+T142) * 0.3 + T148 * 0.2 + \\
 &\quad (T166+T167) * 0.4 + T168 * 0.1
 \end{aligned}
 \left. \vphantom{\begin{aligned} \dots \end{aligned}} \right\} \dots (1)$$

20

【0268】

認識性能情報提示部198は、自動採点処理用のソフトウェアに組み込まれた自動採点結果の修正ソフト上で、信頼度Tfinalが、予め定めてある閾値以下の低信頼度の追記情報を特定し(S306)、その信頼度の値が閾値よりも低い追記情報に関して、手書き入力情報の認識性能を向上させるための認識性能情報として、特定した現状の追記情報について修正を促す認識性能情報をユーザ端末171上に提示する(S308)。この際には、前述のように、たとえば、その追記情報を強調表示する、あるいは、判定理由を示すなど、修正が必要な追記情報であることや修正方法を容易に判断することができるような状態で表示するのがよい。

30

【0269】

採点記号87やコメント88の記入者は、ユーザ端末171に提示された認識性能情報を確認しながら、信頼度の値が閾値よりも低く、誤認識をもたらすような採点記号87(を表わす文字や図形)やコメント88(を表わす文字や図形)に関してのみ、ユーザ端末171上で修正処理を実行する(S310)。

40

【0270】

そして、全て修正作業が完了すると(S264 - YES)、修正された追記情報をデータ処理に反映させる。たとえば、採点・集計結果の確定データやコメント分類結果の確定データを処理結果保存サーバ40に登録する(S264)。利用者による修正結果を受け付け、認識後にデータベース操作などで直接データを変更することで反映する(S268)。つまり、データベースに保存後のデータを直接に変更すればよい。

【0271】

50

この事例の処理手順では、通常の自動データ処理の過程で認識の正誤を反映させて、認識した手書き入力情報の内、認識性能の劣るものについてのみ修正を促すようにしているので、修正を要するものについてのみ特別な手間を掛けずに、採点記号 87 やコメント 88 などの記入態様を随時チェックしつつ修正を行ない、修正後の追記情報を用いて自動採点処理や自動コメント分類処理の結果を適切に変更することができる。

【0272】

< 追記情報処理装置 ; 計算機構成 >

図 19 は、追記情報処理装置 10 の他の構成例を示すブロック図である。ここでは、パーソナルコンピュータなどの電子計算機を利用して、追記情報処理をソフトウェアで実行するマイクロプロセッサなどから構築されるより現実的なハードウェア構成を示している

10

【0273】

すなわち、本実施形態において、追記情報に関するデータ処理を実行する仕組みは、ハードウェア処理回路により構成することに限らず、その機能を実現するプログラムコードに基づき電子計算機（コンピュータ）を用いてソフトウェア的に実現することも可能である。

【0274】

よって、本発明に係る仕組みを、電子計算機（コンピュータ）を用いてソフトウェアで実現するために好適なプログラムあるいはこのプログラムを格納したコンピュータ読取可能な記憶媒体を発明として抽出することもできる。ソフトウェアにより実行させる仕組みとすることで、ハードウェアの変更を伴うことなく、処理手順などを容易に変更できる利点を享受できるようになる。

20

【0275】

電子計算機に、追記情報に関するデータ処理機能をソフトウェアにより実行させる場合には、そのソフトウェアを構成するプログラムが、専用のハードウェアに組み込まれているコンピュータ（組込マイコンなど）、あるいは、CPU（Central Processing Unit）、論理回路、記憶装置などの機能を 1 つのチップ上に搭載して所望のシステムを実現する SOC（System On a Chip：システムオンチップ）、または、各種のプログラムをインストールすることで各種の機能を実行することが可能な汎用のパーソナルコンピュータなどに、記録媒体からインストールされる。

30

【0276】

記録媒体は、コンピュータのハードウェア資源に備えられている読取装置に対して、プログラムの記述内容に応じて、磁気、光、電気などのエネルギーの状態変化を引き起こして、それに対応する信号の形式で、読取装置にプログラムの記述内容を伝達できるものである。

【0277】

たとえば、コンピュータとは別に、ユーザにプログラムを提供するために配布される、プログラムが記録されている磁気ディスク（フレキシブルディスク FD を含む）、光ディスク（CD-ROM（Compact Disc-Read Only Memory）、DVD（Digital Versatile Disc）を含む）、光磁気ディスク（MD（Mini Disc）を含む）、または半導体メモリなどよりなるパッケージメディア（可搬型の記憶媒体）により構成されるだけでなく、コンピュータに予め組み込まれた状態でユーザに提供される、プログラムが記録されている ROM やハードディスクなどで構成されてもよい。

40

【0278】

また、ソフトウェアを構成するプログラムは、記録媒体を用いずに、記録媒体を介して提供されることに限らず、有線あるいは無線などの通信網を介して提供されてもよい。

【0279】

たとえば、追記情報に関するデータ処理機能を実現するソフトウェアのプログラムコードを記録した記憶媒体を、システムあるいは装置に供給し、そのシステムあるいは装置のコンピュータ（または CPU や MPU）が記憶媒体に格納されたプログラムコードを読み

50

出し実行することによっても、ハードウェア処理回路にて構成する場合と同様の効果は達成される。この場合、記憶媒体から読み出されたプログラムコード自体が追記情報に関するデータ処理機能を実現する。

【0280】

また、コンピュータが読み出したプログラムコードを実行することで、追記情報に関するデータ処理機能を実現されるだけでなく、プログラムコードの指示に基づき、コンピュータ上で稼働しているOS (Operating Systems ; 基本ソフト) などが実際の処理の一部または全部を行ない、その処理により追記情報に関するデータ処理機能を実現される場合であってもよい。

【0281】

さらに、記憶媒体から読み出されたプログラムコードが、コンピュータに挿入された機能拡張カードやコンピュータに接続された機能拡張ユニットに備わるメモリに書き込まれた後、そのプログラムコードの指示に基づき、その機能拡張カードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行ない、その処理によって追記情報に関するデータ処理機能を実現される場合であってもよい。

【0282】

なお、追記情報に関するデータ処理機能を実現するプログラムコードを記述したファイルとしてプログラムが提供されるが、この場合、一括のプログラムファイルとして提供されることに限らず、コンピュータで構成されるシステムのハードウェア構成に応じて、個別のプログラムモジュールとして提供されてもよい。

【0283】

たとえば、コンピュータシステム900は、コントローラ部901と、ハードディスク装置、フレキシブルディスク(FD)ドライブ、あるいはCD-ROM (Compact Disk ROM) ドライブ、半導体メモリコントローラなどの、所定の記憶媒体からデータを読み出したり記録したりするための記録・読取制御部902とを有する。

【0284】

コントローラ部901は、CPU (Central Processing Unit) 912、読出専用の記憶部であるROM (Read Only Memory) 913、随時書込みおよび読出しが可能であるとともに揮発性の記憶部の一例であるRAM (Random Access Memory) 915、および不揮発性の記憶部の一例であるRAM (NVRAMと記述する) 916を有している。

【0285】

なお、上記において“揮発性の記憶部”とは、装置の電源がオフされた場合には、記憶内容を消滅してしまう形態の記憶部を意味する。一方、“不揮発性の記憶部”とは、装置のメイン電源がオフされた場合でも、記憶内容を保持し続ける形態の記憶部を意味する。記憶内容を保持し続けることができるものであればよく、半導体製のメモリ素子自体が不揮発性を有するものに限らず、バックアップ電源を備えることで、揮発性のメモリ素子を“不揮発性”を呈するように構成するものであってもよい。

【0286】

また、半導体製のメモリ素子により構成することに限らず、磁気ディスクや光ディスクなどの媒体を利用して構成してもよい。たとえば、ハードディスク装置を不揮発性の記憶部として利用できる。また、CD-ROMなどの記録媒体から情報を読み出す構成を採ることで不揮発性の記憶部として利用できる。

【0287】

また、コンピュータシステム900は、ユーザインタフェースをなす機能部としての指示入力部903と、操作時のガイダンス画面や処理結果などの所定の情報をユーザに提示する表示出力部904と、各機能部との間のインタフェース機能をなすインタフェース部(IF部)909とを有する。

【0288】

なお、データ処理結果を印刷出力してユーザに提示する構成とするべく、処理結果を所定の出力媒体(たとえば印刷用紙)に出力する画像形成部906を設けることもできる。

10

20

30

40

50

【 0 2 8 9 】

指示入力部 9 0 3 としては、たとえば、ユーザインタフェース部 9 8 5 の操作キー部 9 8 5 b を利用することができる。あるいは、キーボードやマウスなどを利用することもできる。

【 0 2 9 0 】

表示出力部 9 0 4 は、表示制御部 9 1 9 と表示装置とを備える。表示装置としては、たとえば、ユーザインタフェース部 9 8 5 の操作パネル部 9 8 5 a を利用することができる。あるいは、C R T (Cathode Ray Tube ; 陰極線管) や L C D (Liquid Crystal Display ; 液晶) などとなるその他のディスプレイ部を利用することもできる。

【 0 2 9 1 】

たとえば、表示制御部 9 1 9 が、操作パネル部 9 8 5 a やディスプレイ部上に、ガイダンス情報や文書入力装置 2 0 が取り込んだ全体画像などを表示させる。また、各種の情報をユーザに通知する際の表示デバイスとしても利用される。なお、表示面上にタッチパネルを有するディスプレイ部とすることで、指先やペンなどで所定の情報を入力する指示入力部 9 0 3 を構成することもできる。

【 0 2 9 2 】

インタフェース部 9 0 9 としては、処理データ (画像データを含む) や制御データの転送経路であるシステムバス 9 9 1 の他、たとえば、画像形成部 9 0 6 や他のプリンタとのインタフェース機能をなすプリンタ I F 部 9 9 6、およびネットワークとの間の通信データの受け渡しを仲介する通信 I F 部 9 9 9 を有している。

【 0 2 9 3 】

このような構成において、C P U 9 1 2 は、システムバス 9 9 1 を介してシステム全体の制御を行なう。R O M 9 1 3 は、C P U 9 1 2 の制御プログラムなどを格納する。R A M 9 1 5 は、S R A M (Static Random Access Memory) などで構成され、プログラム制御変数や各種処理のためのデータなどを格納する。また、R A M 9 1 5 は、所定のアプリケーションプログラムに従って演算して得たデータや外部から取得したデータなどを一時的に格納する領域を含んでいる。

【 0 2 9 4 】

たとえば、追記情報に関するデータ処理機能をコンピュータに実行させるプログラムは、C D - R O M などの記録媒体を通じて配布される。あるいは、このプログラムは、C D - R O M ではなく F D に格納されてもよい。また、M O ドライブを設け、M O に前記プログラムを格納してもよく、またフラッシュメモリなどの不揮発性の半導体メモリカードなど、その他の記録媒体にプログラムを格納してもよい。さらに、他のサーバなどからインターネットなどのネットワークを経由してプログラムをダウンロードして取得したり、あるいは更新したりしてもよい。

【 0 2 9 5 】

なおプログラムを提供するための記録媒体としては、F D や C D - R O M などの他にも、D V D などの光学記録媒体、M D などの磁気記録媒体、P D などの光磁気記録媒体、テープ媒体、磁気記録媒体、I C カードやミニチュアカードなどの半導体メモリを用いることができる。記録媒体の一例としての F D や C D - R O M などには、追記情報に関するデータ処理機能を実現する際の、一部または全ての機能を格納することができる。

【 0 2 9 6 】

また、ハードディスク装置は、制御プログラムによる各種処理のためのデータを格納したり、自装置で取得したデータや外部から取得したデータなどを大量に一時的に格納したりする領域を含んでいる。

【 0 2 9 7 】

このような構成により、操作キー部 9 8 5 b を介した操作者による指令にて、前述の追記情報処理方法を実行するプログラムが記憶されている C D - R O M などの読取可能な記録媒体から R A M 9 1 5 に追記情報処理プログラムがインストールされ、また操作キー部 9 8 5 b を介した操作者による指令や自動処理にて追記情報処理プログラムが起動される

10

20

30

40

50

。たとえば、教材自動採点システム1を実現する場合であれば、教材処理プログラムでは、所定色成分、具体的にはたとえば赤色成分の差分抽出結果を、採点記号87やコメント88の記入内容として認識し、かつ両者を分離するようにする処理ステップなどが記述されプログラムが起動される。

【0298】

CPU912は、この追記情報処理プログラムに従って前述の追記情報処理方法に伴う計算処理を施し、処理結果をRAM915やハードディスクなどの記憶装置に格納し、必要により操作パネル部985a、あるいはCRTやLCDなどの表示装置に出力する。追記情報処理方法を実行するプログラムが記録した記録媒体を用いることにより、既存のシステムを変えることなく、追記情報処理システムを汎用的に構築することができる。

10

【0299】

なお、このようなコンピュータを用いた構成に限らず、図2を用いて示した各機能部の処理をなす専用のハードウェアの組合せにより、追記情報に関するデータ処理機能を行なう追記情報処理装置10を構成することもできる。

【0300】

また、たとえば、追記情報に関するデータ処理機能のための各機能部分の全ての処理をソフトウェアで行なうのではなく、これら機能部分の一部を専用のハードウェアにて行なう処理回路908を設けてもよい。ソフトウェアで行なう仕組みは、並列処理や連続処理に柔軟に対処し得るものの、その処理が複雑になるに連れ、処理時間が長くなるため、処理速度の低下が問題となる。

20

【0301】

これに対して、ハードウェア処理回路で行なうことで、高速化を図ったアクセラレータシステムを構築することができるようになる。アクセラレータシステムは、処理が複雑であっても、処理速度の低下を防ぐことができ、高いスループットを得ることができる。

【0302】

たとえば、追記情報に関するデータ処理機能を実現する場合であれば、処理回路908としては、図2に示した読取画像処理部110に相当する読取画像処理部908a、文書原本特定部120に相当する文書原本特定部908b、追記情報抽出部130に相当する追記情報抽出部908c、データ処理対象追記情報特定処理部150に相当するデータ処理対象追記情報特定処理部908d、あるいはデータ処理部170に相当するデータ処理部908eなどをハードウェアで構成するとよい。

30

【産業上の利用可能性】

【0303】

以上、本発明について実施形態を用いて説明したが、本発明の技術的範囲は上記実施形態に記載の範囲には限定されない。発明の要旨を逸脱しない範囲で上記実施形態に多様な変更または改良を加えることができ、そのような変更または改良を加えた形態も本発明の技術的範囲に含まれる。

【0304】

また、上記の実施形態は、クレーム(請求項)にかかる発明を限定するものではなく、また実施形態の中で説明されている特徴の組合せの全てが発明の解決手段に必須であるとは限らない。前述した実施形態には種々の段階の発明が含まれており、開示される複数の構成要件における適宜の組合せにより種々の発明を抽出できる。実施形態に示される全構成要件から幾つかの構成要件が削除されても、効果が得られる限りにおいて、この幾つかの構成要件が削除された構成が発明として抽出され得る。

40

【0305】

たとえば、歪み補正処理や途切れ補正処理などは、必ずしも必須ではない。

【0306】

また、処理対象の追記済文書8Bの一例として、自動採点集計処理の仕組みに利用される付加情報記入済教材81を例に説明したが、これは一例に過ぎず、データ処理対象となる手書きの追記情報が含まれているものであれば、その文書の種別は問わない。たとえば

50

、自動帳票処理の仕組みに利用される記入済帳票（特開平5 - 342239号公報や平6 - 274156号公報を参照）や個人情報管理（特にスケジュール管理ともいう）の仕組みに利用される付加情報記入済手帳（特開平5 - 216932号公報を参照）などであってもよい。

【0307】

また、先にも述べたが、処理対象の追記済文書8Bは、紙媒体を用いることに限定されず、最初から追記済文書8Bを電子データの形式で入手する態様の場合であっても、データ処理対象となる追記情報が含まれているものであれば、上述した仕組みを同様に適用することができる。

【図面の簡単な説明】

10

【0308】

【図1】本発明に係る追記情報処理において処理対象とする文書の一例である教育用教材を示す図である。

【図2】本発明に係る追記情報処理装置を備えてなる情報処理システムの一実施形態の構成例を示す図である。

【図3】教材自動採点システムにおける教材処理方法の全体概要をシステム構成図と対応付けて示した図である。

【図4】教材自動採点システムにおける全体処理手順を示すフローチャートである。

【図5】認識性能情報提示処理部の処理であって、各処理時の認識率や信頼度に関わる特徴量を収集する手法の一例を示す図である。

20

【図6】認識性能情報提示処理部の処理であって、「抽出ミス」の誤認識のタイプを説明する図である。

【図7】認識性能情報提示処理部の処理であって、「欠損補間ミス」の誤認識のタイプを説明する図（その1）である。

【図8】認識性能情報提示処理部の処理であって、「欠損補間ミス」の誤認識のタイプを説明する図（その2）である。

【図9】認識性能情報提示処理部の処理であって、「複数追記」の誤認識のタイプを説明する図である。

【図10】認識性能情報提示処理部の処理であって、「追記訂正」の誤認識のタイプを説明する図である。

30

【図11】認識性能情報提示部における、認識性能を向上させるための認識性能情報をユーザに提示する第1の提示手法を説明する図である。

【図12】認識性能情報提示部における、認識性能を向上させるための認識性能情報をユーザに提示する第2の提示手法を説明する図（その1）である。

【図13】認識性能情報提示部における、認識性能を向上させるための認識性能情報をユーザに提示する第2の提示手法を説明する図（その2）である。

【図14】認識性能情報提示部における、認識性能を向上させるための認識性能情報をユーザに提示する第2の提示手法を説明する図（その3）である。

【図15】認識性能情報提示部における、認識性能を向上させるための認識性能情報をユーザに提示する第2の提示手法を説明する図（その4）である。

40

【図16】認識性能情報提示部における、認識性能を向上させるための認識性能情報をユーザに提示する第3の提示手法を説明する図である。

【図17】教育用教材を処理対象とする具体的な信頼度情報収集提示処理をシステム構成図と対応付けて示した図である。

【図18】信頼度情報収集提示処理手順を示すフローチャートである。

【図19】追記情報処理装置を、電子計算機を用いて構成する場合のハードウェア構成の一例を示した図である。

【符号の説明】

【0309】

1 ...教材自動採点システム、8A ...文書原本、8B ...追記済文書、9 ...差分情報、10

50

...追記情報処理装置、20...文書入力装置、30...文書管理サーバ、40...処理結果保存サーバ、80...教育用教材、81...付加情報記入済教材、87...採点記号、88...コメント、110...読取画像処理部、120...文書原本特定部、122...教材特定部、130...追記情報抽出部、132...差分抽出部、134...解答者抽出部、136...手書き情報切出部、138...文字認識処理部、140...データ処理用追記情報抽出処理部、142...データ処理対象追記情報抽出部、146...追記情報整形部、148...抽出線分途切れ補正部、150...データ処理対象追記情報特定処理部、154...第1データ処理用追記情報認識部、155...分離認識処理部、156...図形形状認識部、157...文字認識処理部、158...記入位置認識部、164...第2データ処理用追記情報認識部、166...図形形状認識部、167...文字認識処理部、156a, 166a, 157a, 167a...変形処理部、168...記入位置認識部、170...データ処理部、170_1...第1データ処理部、170_2...第2データ処理部、171...ユーザ端末、172...採点集計部、174...集計結果出力部、176...コメント分類処理部、178...コメント処理結果出力、190...認識性能情報提示処理部、192...追記情報認識履歴保持部、194...難認識追記情報抽出部、198...認識性能情報提示部

10

【図1】

(A) 文書原本8A 80教育用教材 85識別情報欄 86解答者情報欄 83b

理科 5年	1. 天気と気温の変化	86a 年	86b 組	86c 番	30/30/40	合計点
1-	各5点	2-	各5点	3-	各10点	
82 問題文	() ()	83a	84 解答欄	() ()	() ()	() ()

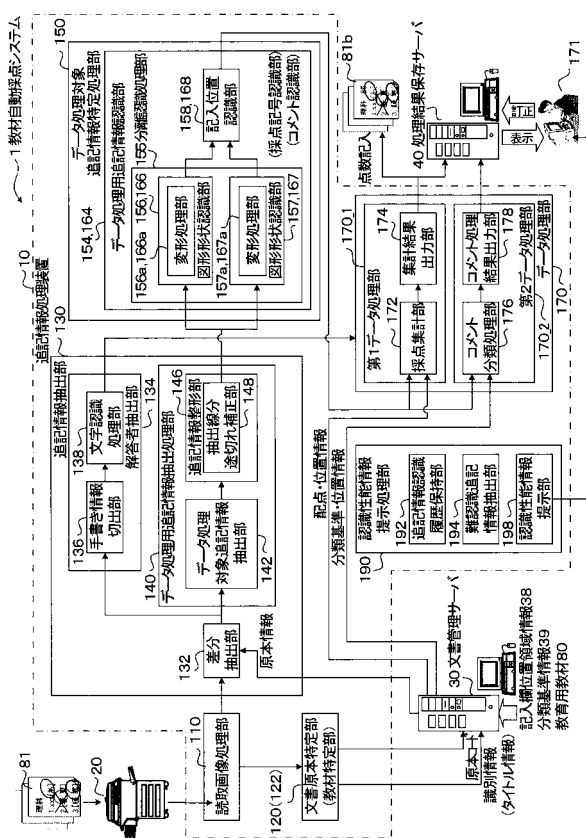
(B) 追記済文書8B 81付加情報記入済教材

理科 5年	1. 天気と気温の変化	年	組	番	30/30/40	合計点
1-	各5点	2-	各5点	3-	各10点	
87 採点記号	() ()	88 コメント	good	() ()	() ()	() ()
88 コメント	() ()	() ()	() ()	() ()	() ()	() ()

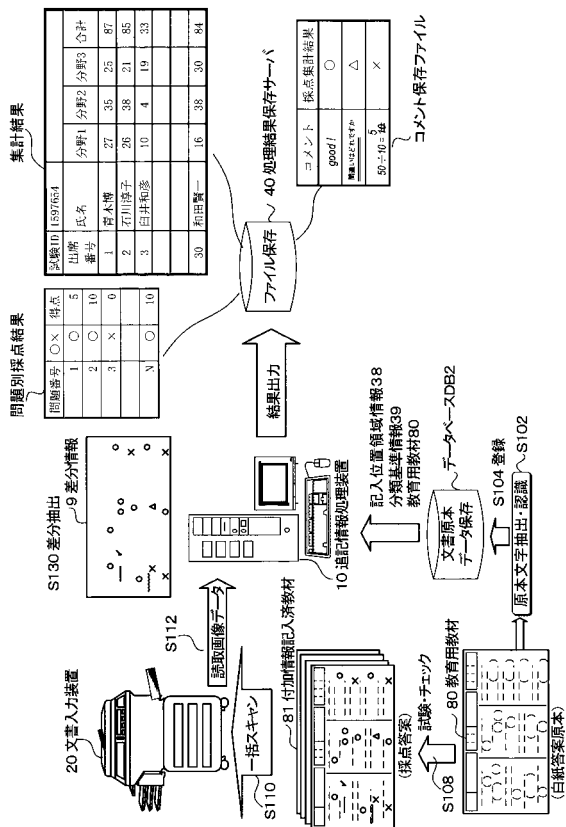
(C) 記入位置領域情報38

問題番号	配点	x	y	W	h
1	5	200	500	240	120
2	10	450	500	300	120
3	5	800	300	240	120
N	10	1000	1700	360	240

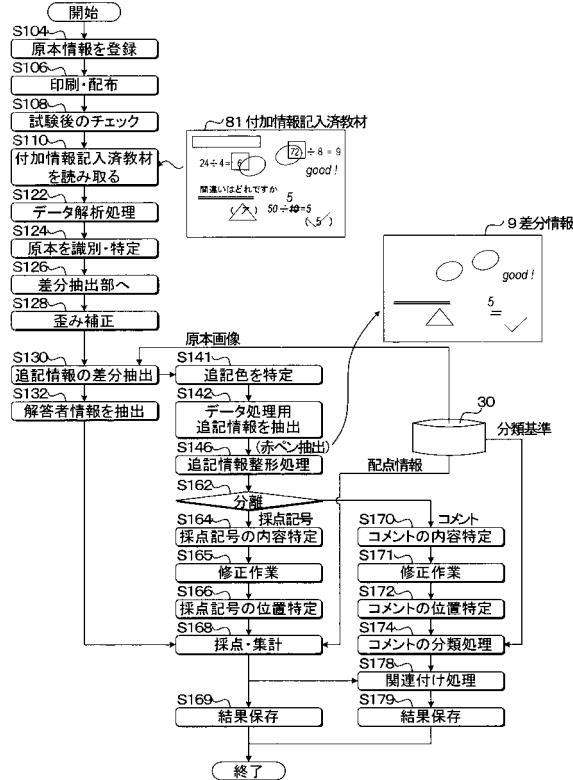
【図2】



【図3】

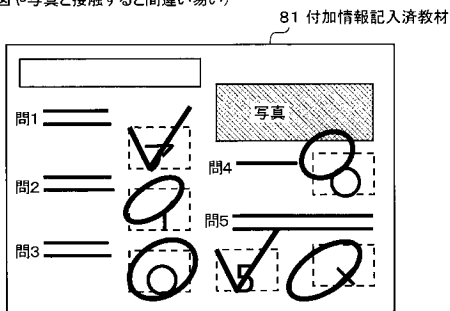


【図4】



【図5】

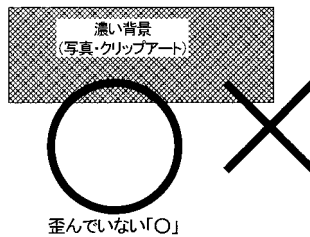
問題ごとに認識率が異なる例 (図や写真と接触すると間違えやすい)



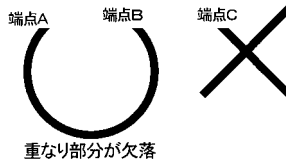
【図6】

<「抽出ミス」の誤認識のタイプ>

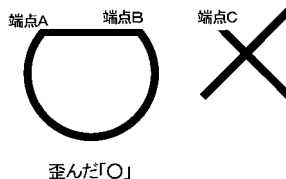
(A) 追記情報抽出処理の対象画像



(B) 追記情報抽出処理後の画像(差分情報9)



(C) 抽出線分途切れ補正後の画像 一図形認識や文字認識の対象画像



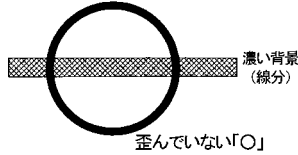
(D) 難認識追記情報抽出部への通知情報

- 抽出量などから求められる抽出時の各品質情報J10, J11
- 採点記号87やコメント88の濃い背景への重複量J12
- 赤抽出画像と原本画像との位置関係などから求められる原本画像と赤抽出画像の接触量J13

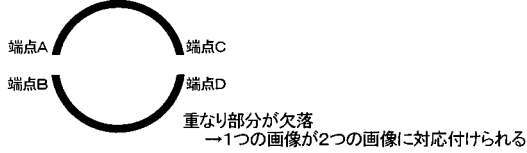
【図7】

<「欠損補間ミス」の誤認識のタイプ;第1例>

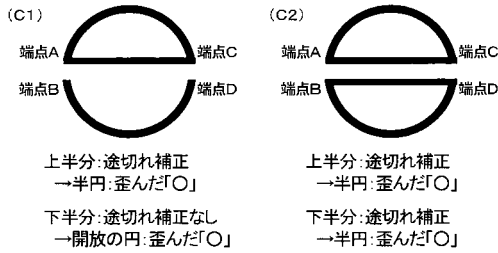
(A) 追記情報抽出処理の対象画像



(B) 追記情報抽出処理後の画像(差分情報9)



(C) 抽出線分途切れ補正後の画像 → 図形認識や文字認識の対象画像

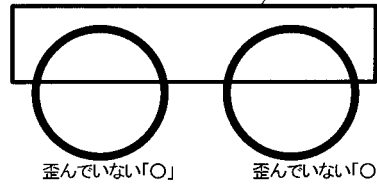


(D) 難認識追記情報抽出部への通知情報
・途切れ補正処理時の補間長さを示す補間長情報J20

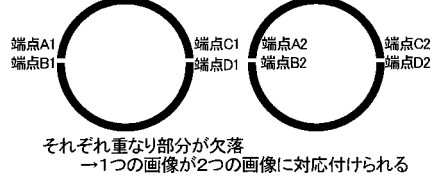
【図8】

<「欠損補間ミス」の誤認識のタイプ;第2例>

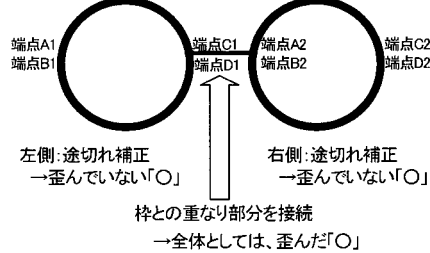
(A) 追記情報抽出処理の対象画像



(B) 追記情報抽出処理後の画像(差分情報9)



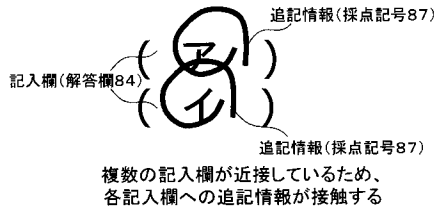
(C) 抽出線分途切れ補正後の画像 → 図形認識や文字認識の対象画像



(D) 難認識追記情報抽出部への通知情報
・途切れ補正処理時の補間長さを示す補間長情報J20

【図9】

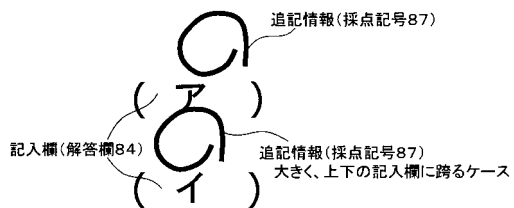
(A) <「複数追記」の誤認識のタイプ>



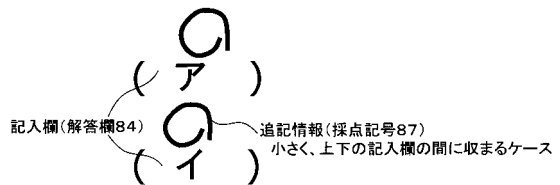
(B) 難認識追記情報抽出部への通知情報
・採点記号87同士の接触の有無を示す情報J30
・コメント88同士の接触の有無を示す情報J31
・各図形認識処理結果の認識の信頼度情報J32, J33
・各文字認識処理結果の認識の信頼度情報J34, J35

【図10】

(A) <「追記訂正」の誤認識のタイプ;第1例>

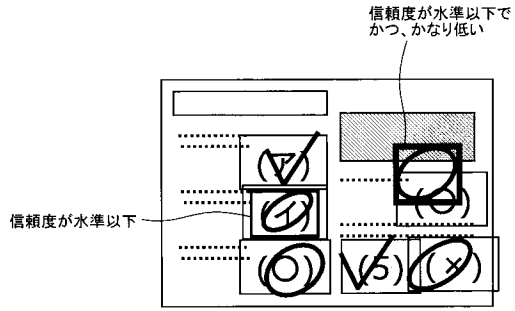


(B) <「追記訂正」の誤認識のタイプ;第2例>



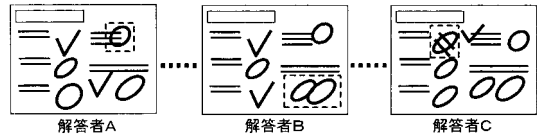
(C) 難認識追記情報抽出部への通知情報
・1つの採点記号87の複数の解答欄84への対応付けの可能性の有無を示す情報J40
・1つのコメント88の複数の解答欄84への対応付けの可能性の有無を示す情報J41
・各採点記号87と各解答欄84との距離を示す情報J42
・各コメント88と各解答欄84との距離を示す情報J43

【図11】

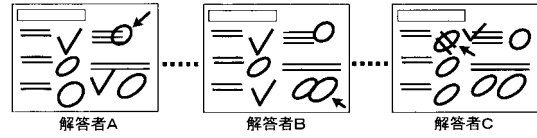


【図12】

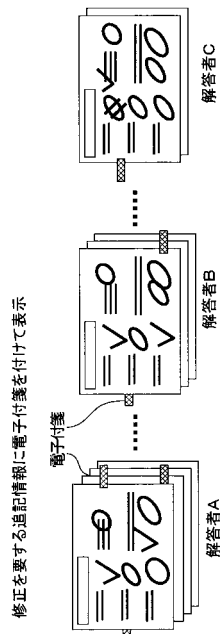
(A) 修正を要する追記情報を枠で囲んで表示



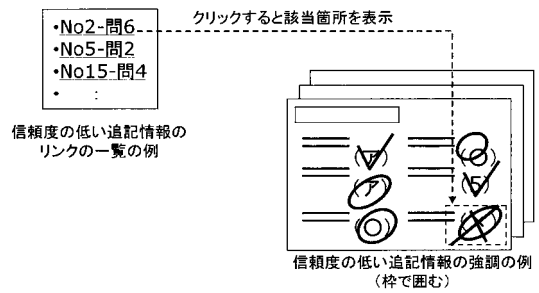
(B) 修正を要する追記情報に矢印を付けて表示



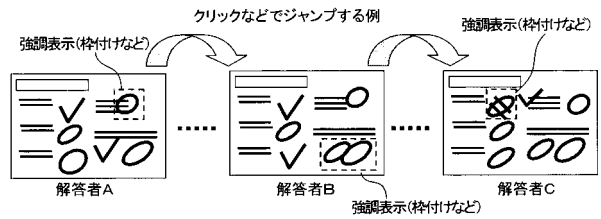
【図13】



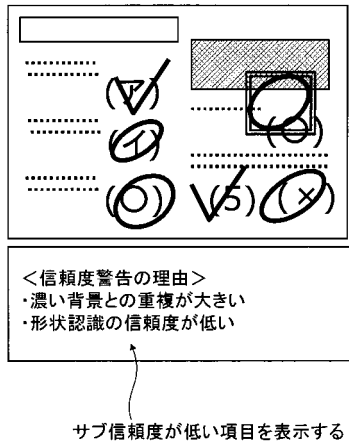
【図14】



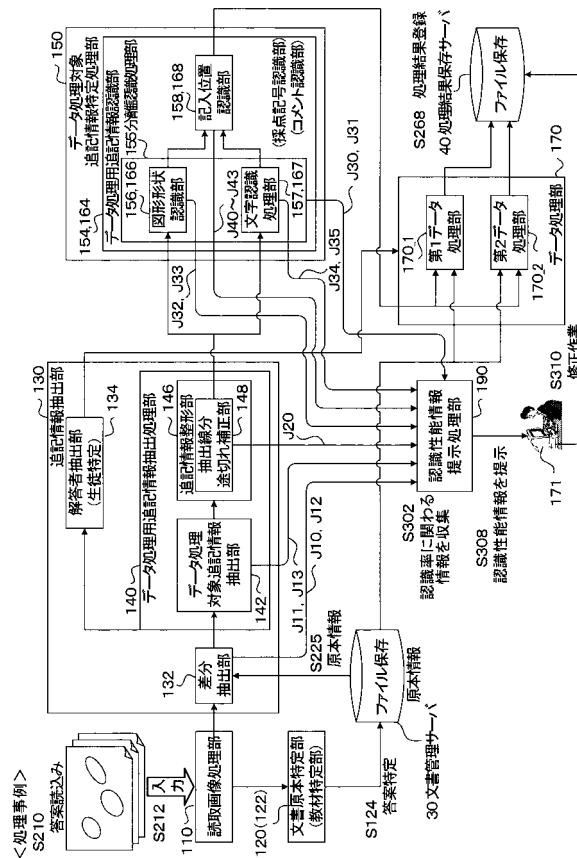
【図15】



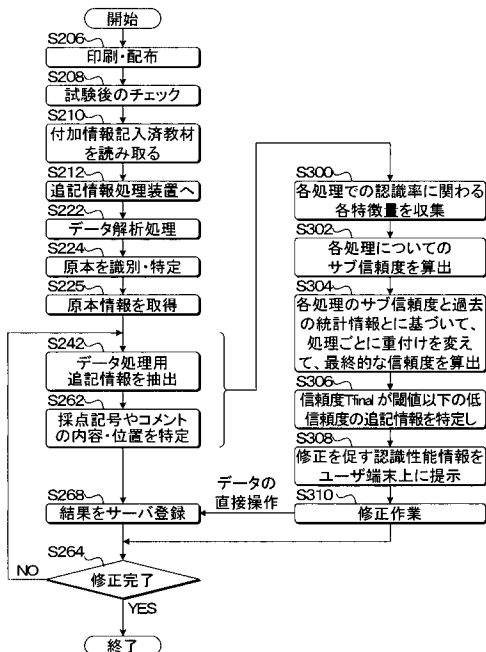
【図 16】



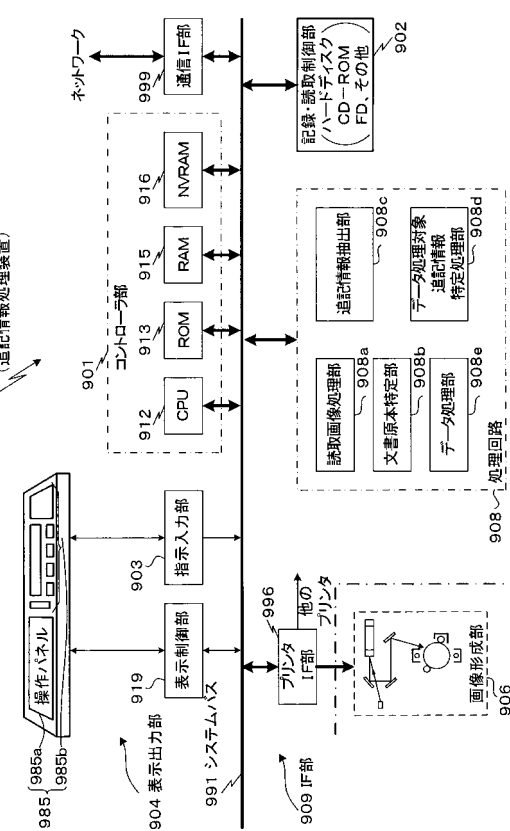
【図 17】



【図 18】



【図 19】



フロントページの続き

(56)参考文献 特開2004-152115(JP,A)
特開平09-161013(JP,A)
特開平10-021326(JP,A)
特開平10-063773(JP,A)

(58)調査した分野(Int.Cl., DB名)
G06K 9/03