JSiSE Research Report
vol.34,no.4(2019-11)

人手による採点を完全に代替する手書き入力漢字自動採点システムにおける画数フリーの実現方法

井戸　伸彦
岐阜協立大学　経営情報学科

Stroke-Number Free Method for Kanji Character Automated Scoring which can Completely Replace Manual Scoring

Nobuhiko Ido
Faculty of Business Administration, Gifu Kyoritsu University

I have been developing an automated scoring system for on-line hand-written kanji characters and the system is practically used in a class at the university where I work. In this paper I report a stoke-number free method for the scoring system, which is one of the basic techniques to completely replace manual scoring with automated scoring. A stroke-number free method is necessary to score hand-written input characters in which one stroke is divided into plural strokes or plural strokes are united into one stroke. In the proposed method, strokes of a hand-written input character are converted into the same configuration as those of the correct character corresponded to it, so that the scoring information of the correct characters can be used for various input characters.

キーワード: 画数フリー，手書き入力文字，自動採点，タブレット端末，学習支援

1. はじめに

タブレット端末等で手書き入力した文字や記号を自動採点する試みは各所で行われている(1)(2)．しかしながら，実用化されている漢字自動採点システムには制限もある(3)．報告者は人が行う採点を完全に代替する機能を持つ手書き入力漢字の自動採点システムを開発中であり，これを所属する大学の実際の授業で2019年度 4 月より運用している(4)．開発したシステムを用いれば，単語で解答する試験の自動採点も可能となる．

人が行う場合と同等の採点を行うための要件の一つとして，画数フリーである必要がある．画数フリーとは，正答での一画を複数画に分割した入力，正答での複数画を一画に統合した入力を採点可能である性質を指す．手書き入力文字認識の分野では画数フリーについて古くから検討されている(5)が，正誤の判定結果の理由を明示する自動採点での検討はほとんど例を見ない．

本稿では，手書き漢字自動採点における画数フリーを実現する方法を提案し，実際の授業で適用した結果についても報告する．提案方法では，画数が異なることなどの画構成の差異を仮定し，これに基づく変換／逆変換を施すことで，手書き入力字形を正答の字形と同じ画構成として採点を行う．この方法により，正答の字形に基づく採点情報だけを作成すれば，様々な画構成の字形の手書き入力を採点することが出来る．実際に授業で運用した際の手書き入力を試料とした採点では，画構成に違いのある入力に対して的確にそれを指摘して正しい採点結果を出力することを確認した．

2. 背景となるシステムと技術の概要

開発を行っている手書き入力漢字の自動採点システムについて，本稿で提案する画数フリーの技術の前提となるものを中心にその概要を説明する．

2.1　システムの構成

図 1 に手書き入力漢字自動採点システムの構成と動作の概要を示す．システムは web ページにアクセスすることにより利用する．すなわち，(1)教員が漢字の単語で解答する問題を記述したファイルをアップロードすると，(2)問題の正答に含まれる漢字の漢字採点情報（字形と採点基準）が付加されて，学生が利用するタブレット端末からアクセス出来るようになる．(3)学生はタブレット用のペンを用いて解答を手書き入力し，web ページ上で動作する採点機能(JavaScript により実装)による採点結果を閲覧する．(4)web サーバ上に蓄積された採点結果は，教員が成績評価等のために閲覧できる．なお，採点基準等はシステム管理者により更新・管理されている．

2019 年 4 月から所属する大学の授業において行っている運用では，システムを利用して行う小テストを宿題として課し，学生は情報施設利用窓口でタブレット端末を一時借用して自習室で解答することとしている．図 2 に，学生向けの web ページでの採点結果表示画面の例と学生の利用の様子を示す．図中の画面のように，漢字は一つの入力領域に一文字ずつ手書きする．

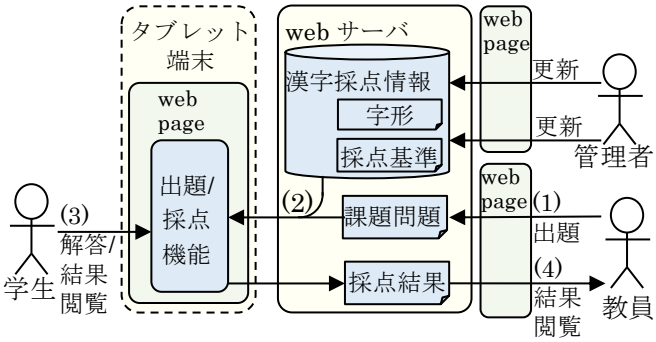

図 1 システム構成と動作の概要

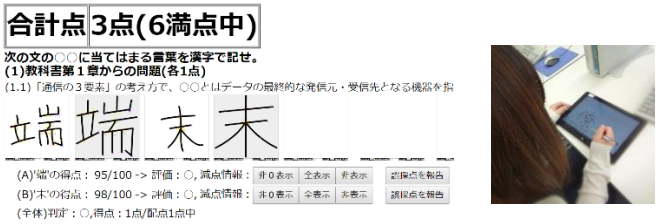

図 2 採点結果表示画面と利用の様子

2.2　自動採点処理

図 3 に自動採点処理の概要を示す．本システムでは漢字の字形をベジェ曲線で表現しており，これにより与えられた正答の字形と被験者による入力の字形との間で「(a)画の対応付け」を行う[6]．次に，対応付けられた画同士で「(b)辺の対応付け」を行う[7]．最後に，対応づけられた画や辺ごとの図形的特徴を用いて，「(c)点数評価」を行う[8]．

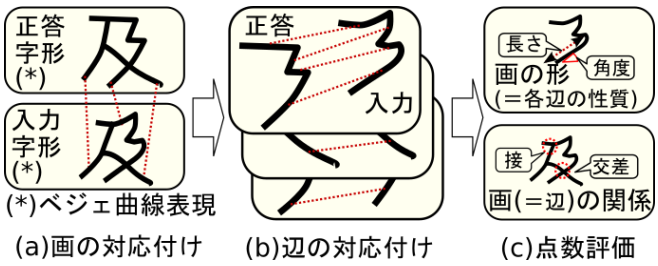

図 3 自動採点処理の概要

2.3　採点情報

本漢字自動採点システムでは，採点において減点方式を取っている[8]．すなわち，様々な減点項目での減点関数により決まる減点値を積算し，満点の 100 点から差し引いた点数が 80 点以上であれば正解としている（このような条件は漢字採点情報中のパラメーター設定により変更出来る）．例えば図 4(a)中に示すように，正答の字形で曲線になっている画がある場合，これに対応する入力の字形の画の様々な曲がり具合の測定値を求め，これに減点関数を適用して減点値を決める．減点関数は図 4(b)のような数値列で表現されており，これは(c)のグラフに示すような減点を表している．このような減点関数を含む採点情報は，報告者が開発した採点情報作成システムにより系統的に生成される．

本システムの採点情報は，平成 28 年に文化庁により示された漢字の正誤基準に関する詳細な指針の文書[9]に全面的に基づいている，

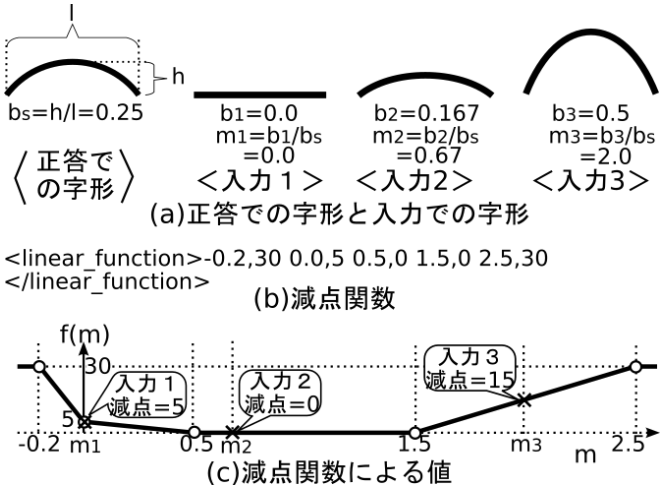

図 4 減点関数

3. 画数フリー化における課題

3.1 画数誤りの独立採点

自動採点における画数フリーにおいては，画数誤りを指摘した場合でも，画数以外の正誤は独立してチェックして採点出来ることが必要である．この独立採点について，図 5 に示した“世”の字の例を用いて説明する．図 5 中，左側の字形は，自動採点時に参照する正答の字形であり，右側の字形が学習者により入力された字形であるが，字形は説明のための模式的なものである．また，画の構成を分かりやすくするためにその終端を矢印で表示している．両者を比較すると，(A)正答側での第 1 画と第 2 画とが交差しているのに対し入力側では接している，(B)第 5 画が 2 つの画に分割されている，(C)第 2 画と第 4 画とがひと続きとなっているという 3 つの誤りを含んでいる．自動採点システムでは(A)(B)(C)のすべての誤りを同定して学習者に指摘し，予め定められた基準に従って(A)(B)(C)を個別に減点して採点結果を算出することが出来る必要がある．画数に誤りがあること ((B)(C)) だけを指摘して(A)を指摘しないと，学習者が誤りに気付かない場合が発生し，学習効果は限定的となる．

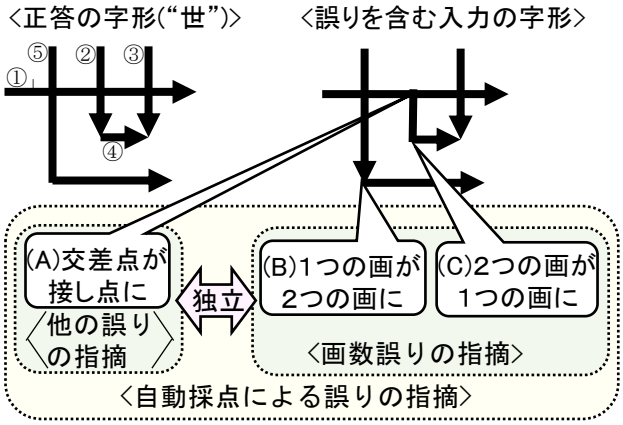

図 5 画数誤りの独立採点

3.2 採点情報の合理的利用方法

前記 2.3 項に示した採点情報は，画数，すなわち，画構成が変わると，いずれの画に対する記述であるかを書き換える必要がある．このため，図 3 に示したような枠組みを単純に適用して採点を行うとすると，図 6(a)に示すように数多くの画の構成に合わせて採点情報の準備が必要となる．想定される画構成は組合せ的に増加するので，図 6(b)に示すように画数(画の構成)に誤りがある字形を，誤りの無い字形の採点情報で採点出来るようにすることが望ましい．

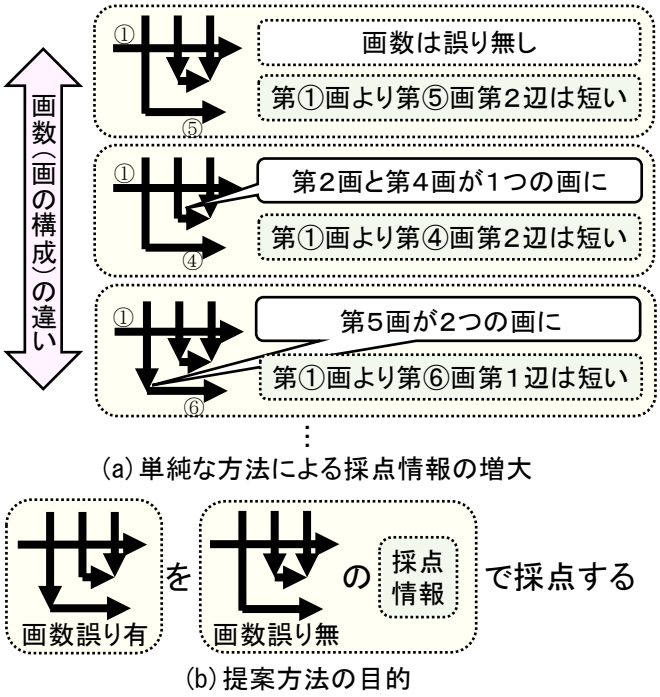

図 6 採点情報に関わる要件

4. 提案方法

4.1 概要

図 7 を用いて，画構成に誤りがある入力字形（図 7 中では(C)）を，正答の字形による採点情報(A)で採点

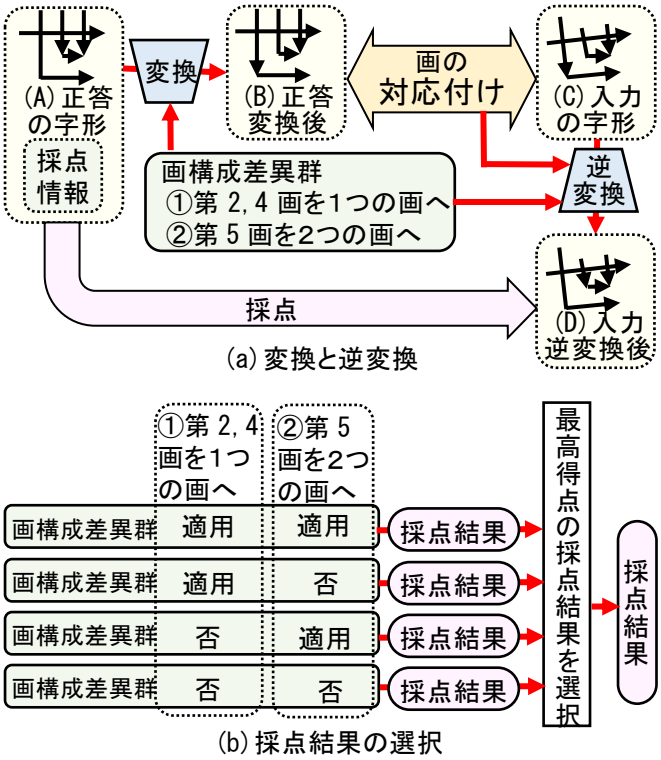

図 7 提案方法の概要

する提案方法の概要を説明する．採点の際には，まず前提とする画構成差異群を仮定する．図 7(a)にて仮定する画構成差異群は，「①第 2,4 画を 1 つの画へ」「②第 5 画を 2 つの画へ」の 2 つから成る．この画構成差異群に基づいて(A)の正答の字形を変換し，(B)の正答変換後の字形を得る．次に，この(B)の字形と(C)の入力の字形との間で画の対応付けを行う．得られた画の対応付けを参照すれば，(C)の入力の字形に対して画構成差異群に基づく逆変換が可能となり，これにより(D)の入力逆変換後の字形を得る．仮定した画構成差異群が入力字形に当てはまるものであれば，(D)は正答の字形と同じ画構成となっているので，(A)正答の字形の採点情報により採点が可能となる．

図 7(a)には特定の画構成差異群を仮定した変換／逆変換による採点を示しているが，図 7(b)に示すように各々の画構成差異を適用するか否かによる様々な変換／逆変換について採点すれば，その中で最高得点となった採点結果を選択することにより，様々な画構成を取る入力の字形を採点することが出来る．

4.2　画構成差異

前項で示した方法により採点を行う際，画構成差異群の要素となる差異にどのような差異を含めるかを決める必要がある．

開発したシステムでは，画分割（1 つの画を複数画に）については，複数の辺を持つすべての画について，辺の境目で分割される差異を画構成差異に含めている．

画統合（複数の画を 1 つの画に）については，報告者が開発した採点情報作成システム（図 8，一部拡大表示）により．どのような差異を画構成差異に含めるかを指定する．このような画構成差異の指定は，他の減点項目（ex.ある画が別の画に比べて長すぎる）と同じように減点項目として扱われており，減点幅を決める減点関数なども指定できる．画を逆向きに書く場合も画構成差異としており，また，画分割において減点幅をデフォルトから変更する場合についても，採点情報作成システムにより減点関数を指定する．

4.3　分割された画の関係

正答では 1 つの画を手書き入力で 2 つの画に統合して書いた場合には，その 2 つの画の関係が問題となる場合がある．具体的には，図 9(a)に示したように，“号”や“糸”の字については明朝体（図中②）の字形に倣った折り方として書いた字形（図中③）がこれに該当する．このような書き方に対して，文化庁指針文書[9]では「1 画が 2 画に見える明朝体のとおりに手書きされても，字形を見誤ることはないので，誤った字であ

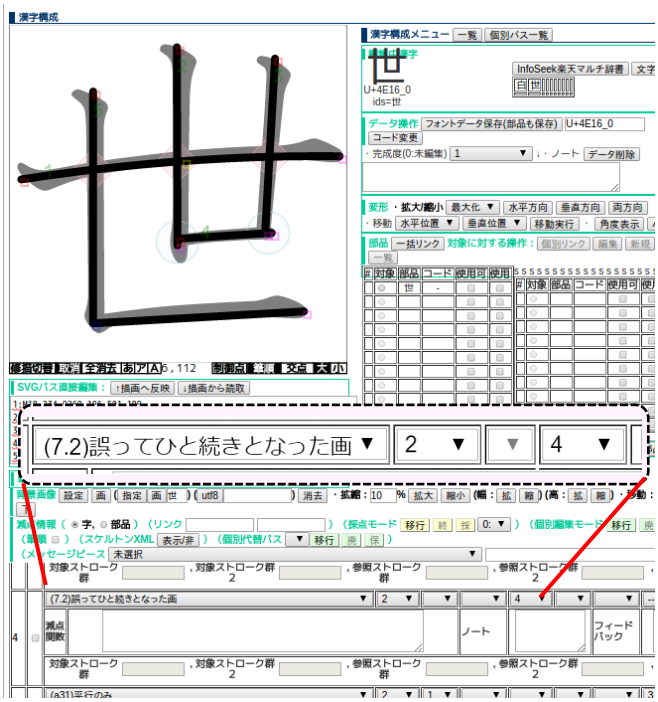

図 8 採点情報作成システムでの設定

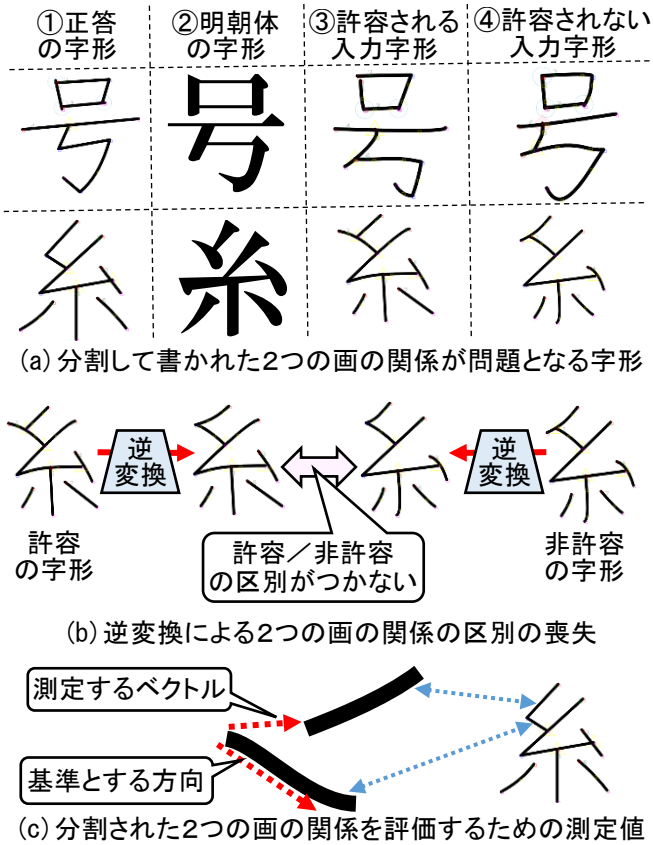

図 9 分割された画の関係

るとまでは言えない」としている．しかしながら，明朝体とは逆の接し方となった図 9(a)中④のような字形については，③よりも厳しく減点して採点すべきである．この「③許容される入力字形」と「④許容されない入力字形」とは，4.1 項で説明した逆変換を施した際に区別が付かなくなってしまう（図 9(b)）．このため，開発したシステムでは，図 9(c)に示すように，分割された２つの画の関係を測定値として評価するようにしている．すなわち，２つの画となって離れた 2 点を結ぶベクトルを，一方の画の方向を基準として測定し，これを減点関数で評価することによって「④許容されない入力字形」の減点を大きくしている．なお，“号”の字の場合は“糸”の字の場合と分割された画の関係が逆になっているため，採点情報作成システムで減点項目のパラメーターを指定することにより，“糸”のような場合と“号”のような場合とを区別している．

4.4 統合された画の逆変換

統合された画を逆変換する際には，その画を分割する位置を決める必要がある．一般的には図 10(a)に示す“こざとへん”の例のように，辺の対応付けにより分割する位置を決める．すなわち，<A>の[正答の字形]を「第１画と第２画を続けて１つの画に統合」という仮定を基に変換した字形<B>と，<C>の[画の統合のある入力字形]との画の対応付けを行った後，辺の対応づけも行って第１画と第２画とを変換で統合した位置に対応する<C>の字形上の位置を求め，ここで分割することで逆変換を行い，<D>の字形を得る．なお，図 10 中では，画の構成を分かりやすくするために，画の端点には四角の印をつけて，２つの画が別であることを示すためにその間を空けて強調している（実際のシステムにおいても，端点に対しては半透明の薄い四角の印を表示するようにしている）．

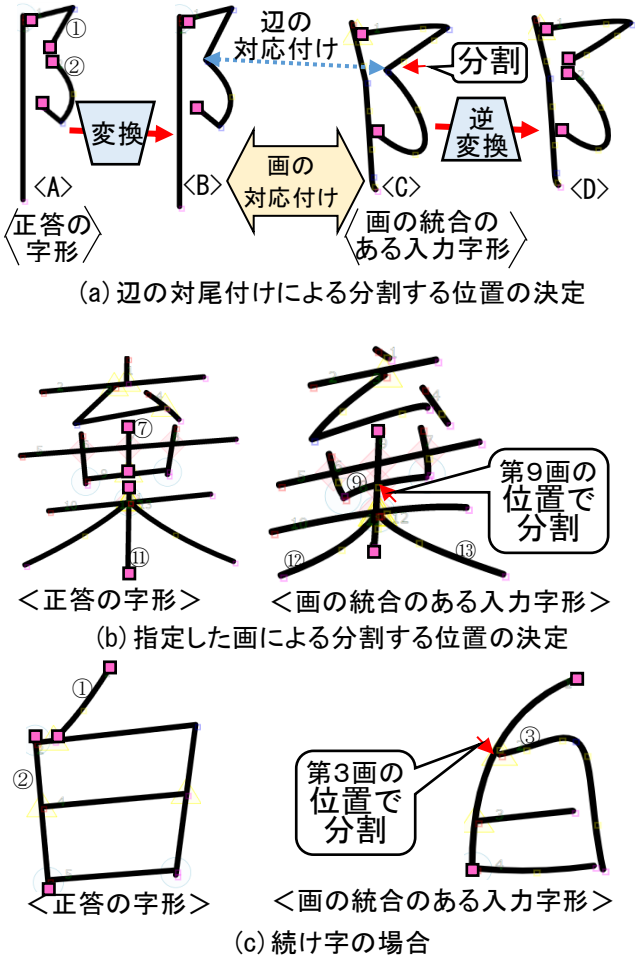

図 10 統合された画の逆変換

図 10(b)に示す“棄”の字の場合，続けて書いて統合するという誤りをしがちな第７画と第１１画とは一直線上にあるので，辺の対応づけにより分割する位置を決めることは出来ない．このため，減点項目として第 1 画と第 2 画との統合を採点情報作成システムで指定する際に，第 9 画と交差する点で分割することも指定しておく．この“棄”の字のように統合される２つの画が一直線上にあることは稀であるが，図 8(c)に示した“白”の字の第１画と第２画とをいわゆる“続け字”で書く場合についても，その分割する位置（“白”の場合は第３画）を指定している．なお，開発したシステムでは，現状，続け字は認めない方針に基づき大きく減点している．もし方針として続け字を認めたいのであれば，続け字の減点関数を定義した１行の記述を修正するだけで変更することが出来る．

また，図 10(b)に記した“棄”の字の入力字形は，正答の字形の第１１画に対応する入力側の画の終点が，第１２画および第１３画を基準として位置が上側過ぎるが，開発したシステムではこれも的確に指摘して減点する．すなわち，図 5 (b)に記した変換／逆変換により，画の統合の無い場合とまったく同様の処理により第１１画の終点の位置の異常が検出されるので，上記 3.1 項に示した“画数誤りの独立採点”の要件が満たされている例となっている．

5. システムの運用での採点例

5.1 対象となる字

開発したシステムは現在も運用しているため，自動採点システムの評価を行う試料としての入力字形は増えつつあり，全体の整理／分析は終わっていない．本稿では，次の単語が正解となる出題に対する解答の手書き入力を試料とした，執筆時点で整理／分析済みの採点結果について報告する．

端末 階層 基本参照 量子化 回線 同僚
符号化 搬送波 誤り 輻輳 全二重通信 送達確認
廃棄 迂回 擬似 乱数

課題の出題時期等により試料数(すなわち解答者数)には，１６～３５個のばらつきがある．

5.2 画の分割

以下では画が分割された例を分類して記す．

5.2.1 明朝体の折り方

上記 4.3 項に記した明朝体の折り方に該当する採点例として，自動採点システムは次の 8 例を検出した．

(1)“階”第５画　35 試料中 5 例（図 11(a)など）

(2)“号”第５画　22 資料中 3 例（図 11(b)など）

このうち(1)“階”については明朝体風であることが明瞭であるが，(2)“号”については，２つに分割された画の終点と始点とが繋がった字形であった．これは明朝体の折り方とは言えないかも知れないが，画を分割していること自体は明朝体の影響であると考えている．上記 4.3 項に記した“分割された画の関係”の観点から減点される例は，残念ながら無かった．

明朝体の折り方については，文化庁指針に“手書きにはよりふさわしい書き方がある”とあるため，開発したシステムでは減点３点を科しているが，決定的な誤りとはしていない．

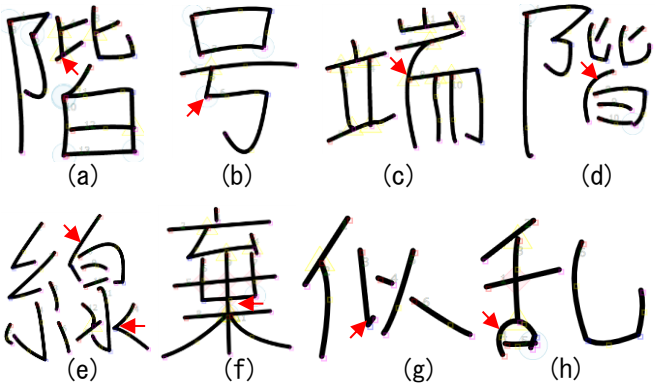

図 11 画数フリーによる採点が適用された入力例

5.2.2 画構成評価に不慣れなための誤り

タブレット端末による入力では画の構成についても正誤の評価対象としているが，被験者である学生は必ずしもこれに慣れている訳では無い．また，タブレット用のペンの扱いについても同様である．見た目にはほとんど違いの無い画構成の間違いを気にせず，ペンの扱いを誤った後に追記するような形で入力したような字形が全漢字試料中に 13 例ほど見受けられたが，自動採点ではこれらすべてを的確に指摘した．特に図 12 中の“階”のこざとへんの例では，「第１画が誤って分割されている」ことと，「第１画と第２画とが誤って統合されている」ことの両方を指摘しており，提案方法が高機能であることの証左であると考えている．

このような誤りは，単に不慣れであるのか，本当に間違えて記憶しているのかの区別が付かないため，どの程度の減点とするかは議論の分かれるところである．現状は不慣れなことを重視して１０点の減点と設定しており，決定的な誤りとはしていない．将来的にはタブレット端末による漢字入力が一般化して画構成の評価に関する理解や慣れが進めば，減点幅を大きくして決定的な誤りと判定することも一案と考えている．

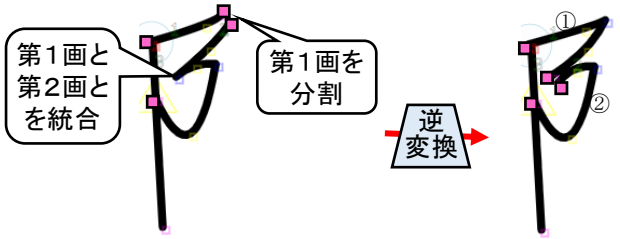

図 12 画の統合と分割の両方が検出された入力例

5.3 画の統合

以下では画が統合された例を分類して記すが，その分類が日本語学的な見地から妥当であるか否かについては，今後専門家の助言を得て変更する必要がある部分も含んでいる．そのような変更は．採点情報の編集等で簡易に行うことが出来る．

5.3.1 続け字

報告者が続け字として採点情報作成システムで画の統合を登録しているものでは，次の 7 例が検出された．

(1)“端”第 10,11 画　34 試料中 1 例（図 11(c)）

(2)“階”第 8,9 画　34 試料中 2 例（図 11(d)など）

(3)“線”第 7,8 画　32 試料中 2 例（図 11(e)など）

(4)“線”第 14,15 画　32 試料中 2 例（図 11(e)など）

このうち(4)“線”については続け字とは考えず，誤字と分類すべきかも知れない．

続け字は，“くずし字”と呼べば美醜の観点から良いものという意味合いが強くなることもあって，誤りと判定することには抵抗があるかも知れない．しかしながら，草書体に連なる様々な字形は教育現場で扱う範疇を超えていると考え，現状システムでは許容しない立場で大きな減点を設定している．

5.3.2 終点と始点とがほぼ繋がっている場合

正答の字形で２つの画の終点と始点がほぼ繋がっている場合には，習慣的に続けて書く人が多いと思われる．そのような字形には画の統合を採点情報作成システムで登録しており，次の例が検出された．

(1)“階”第 1,2 画　35 試料中 15 例

(2)“子”第 1,2 画　34 試料中 7 例

(3)“棄”第 6,9 画　17 試料中 5 例

(4)しんにょう第 2,3 画　100 試料中 20 例
（“送”，“通”，“達”，“迂”）

これらの誤りは軽微と考え，現状システムでは減点５点と設定している．

5.3.3 誤字

報告者が誤字として採点情報作成システムに画の統合を登録しているものでは，次の 7 例が検出された．

(1)“棄”第 7,11 画　17 試料中 3 例（図 11(f)など）

(2)“似”第 3,4 画　17 試料中 4 例（図 11(g)など）

5.3.4 画の統合を検出しない場合

図 11 中(h)は，第４，５画を統合して書いたと見なすことが出来る“乱”の入力例であり，現状ではシステムは画の統合を指摘しないで，“第４画が無い”と指摘する．試みに採点情報作成システムに第４，５画の統合を登録すると指摘を行うようになるが，そのような登録は行っていない．このような字形は，“正しいと採点される漢字を入力する”と意志がある字形の範疇から外れており，“第４画が無い”という指摘で十分であると報告者は考えている．画の統合を登録してもしなくても採点結果は当然不正解となり，試験の採点としても問題は生じない．採点処理に要する時間も増えるため，現状の指摘しない仕様のままとしたいと考えている．

5.4 逆向きの画

5.4.1 どちらの方向でも良いもの

報告者が“どちらの方向でも良い”として採点情報作成システムで逆向きの画を登録しているものでは，次の 9 例が検出された．

(1)“階”第６画　34 試料中 5 例

(2)“化”第３画　34 試料中 1 例

(3)“輳”第１３画　22 試料中 1 例

(4)“擬”第４画　18 試料中 2 例

これらについては，文化庁指針文書にて“右から左にはらって書くことも，横画として左から右に書くこともある”としており，減点は０として表示のみ行っている．

5.4.2 誤字

報告者が誤字として採点情報作成システムで逆向きの画を登録しているものでは，次の 17 例が検出された．

(1)“階”第４画　34 試料中 16 例

(2)“乱”第１画　17 試料中 1 例

これらについては，文化庁指針文書には画の方向について記述は無く，減点５点としている．紙に書いた文字にては，逆向きの画は正しい向きの画と見分け難いものであり，不正解とすることは厳しすぎると考え，減点を軽微にしている．

5.5 まとめ

提案方法は原理としてシンプルであり，画数フリーの採点が難しくなるような入力事例は現状見つかっていない．前述した 5.3.4 項の“乱”の場合のように画数誤りの指摘対象とするか否かを選択する必要はあるものの，実際のシステム運用にても採点上特に問題となるような入力は無かった．今後試料数を増やした検証は必要であるが，人手によるものを完全に代替する自動採点技術として十分機能するとの結果を得たと考えている．

6. おわりに

開発する漢字自動採点システムは既に運用中であり，今後 JIS 第一水準を目指した採点情報の整備，ひらがなのサポート，ユーザインターフェースの改善を行っていく予定である．

参 考 文 献

(1) 中川正樹, 平井佑樹: “記述式解答の採点支援・自動採点に向けてー手書き認識からの挑戦ー”, 情報処理 57.9 , pp. 920-924 (2016)

(2) 石岡恒憲: “コンピュータ上で実施する記述式試験: エッセイタイプ, 短答式, マルチメディア利用について”, 電子情報通信学会誌, 99(10), pp. 1005-1011 (2016)

(3) Benesse, “保護者サポート” , https://faq.benesse.co.jp/faq/show/6646?category_id=855&site_domain=sho(2019/10/01 確認)

(4) 井戸伸彦: “文化庁指針に基づく漢字の字体字形の基準を網羅的に実装する手書き入力漢字の自動採点システムの運用”, JSiSE2019 第 44 回全国大会, pp.389-390 (2019)

(5) 槇重弼, 迫江博昭: “筆順・画数自由オンライン文字認識のための画対応決定法―多層キューブサーチ”, 信学論（D-Ⅱ）, vol.J87-D-Ⅱ, no.5, pp.1112-1119, (2004)

(6) 井戸伸彦: “多軸順序距離を用いた手書き漢字の画の対応付け”, 信学技法(パターン認識・メディア理解),Vol.114, No.42, pp.85-90（2014）

(7) 井戸伸彦: “漢字書き取りの自動採点の画の評価手法ーアグレッシブ DP マッチングによる辺の同定と評価ー”, 信学技報(パターン認識・メディア理解), Vol.114, No.527, pp.217-222（2015）

(8) 井戸伸彦: “画と辺とを同定した後の手書き漢字入力の採点評価方法”,信学技報(パターン認識・メディア理解), Vol.115, No.25,pp.133-138(2015)

(9) 文化庁: “常用漢字表の字体・字形に関する指針(報告)” http://www.bunka.go.jp/seisaku/bunkashingikai/kokugo/hokoku/pdf/jitai_jikei_shishin.pdf(2019/10/01 確認)