

文字認識技術

文字認識技術は身近なところで活躍しています。例えば膨大な量のハガキや小包の仕分けを、人が手作業で行うのではなく、7桁の郵便番号を文字認識して、地域ごとに機械が自動的に仕分けしてくれています。

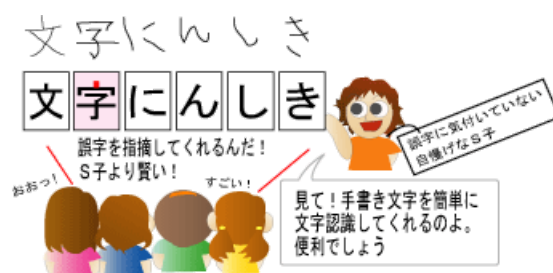
ご利用にあたっての注意

この講座の内容は、2008年当時の情報です。予告なしに更新、あるいは掲載を終了することがあります。あらかじめご了承ください。

最終更新日 2008年8月4日

もくじ

- ↓ 文字認識ってなんだろう
- ↓ オフライン文字認識（紙に書かれている文字を判別する技術）
 - 原理
 - 手書きの住所の文字認識
 - 新聞や雑誌、カラー画像の文字認識
 - 応用
- ↓ オンライン文字認識（画面に入力した手書き文字を判別する技術）
 - 原理




- 文字評価技術を使った自動採点（富士通オリジナル）
- オンライン手書き学習ソフトを使いやすくするための工夫

↓ 小話

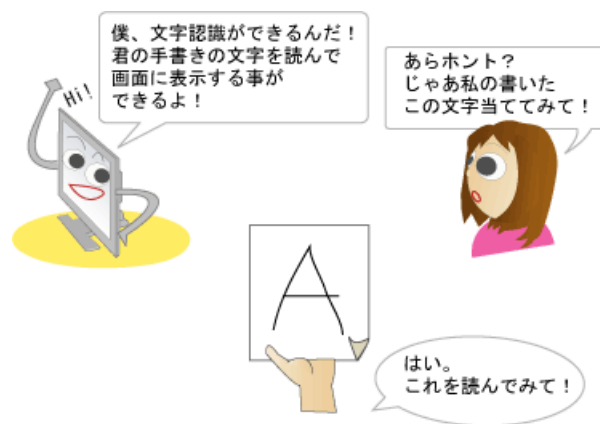
↓ 関連ページへのリンク

（講座メモ）

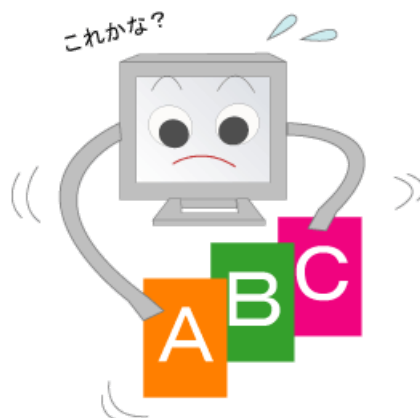
 ボタンを押すとサンプル音声聞けます。

文字認識ってなんだろう

文字認識とは、紙に書かれた手書きの文字や、印刷文字などを判別する技術の事です。



僕はテンプレート（形板）を持っているから、まず合いそうな物を探すんだ。

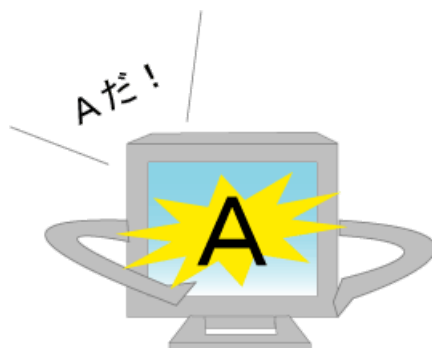


そしてテンプレートと文字を重ねて、後ろから光を当ててるんだ。透かして見て、選んだ物が正しければ光が遮られて真っ暗に見えるから、この文字は『A』って事がわかるんだよ。光が漏れて見えると違う字だな

って判断して、また他の文字を探すんだ。これはパターンマッチングって言って、1928年に開発された最も典型的な技術なんだよ。意外と歴史があるでしょ。



こうやって調べた結果はAと判定。正解しました。
他にも色々な方法や技術があるんだ。現在はすごく進歩してるしね。

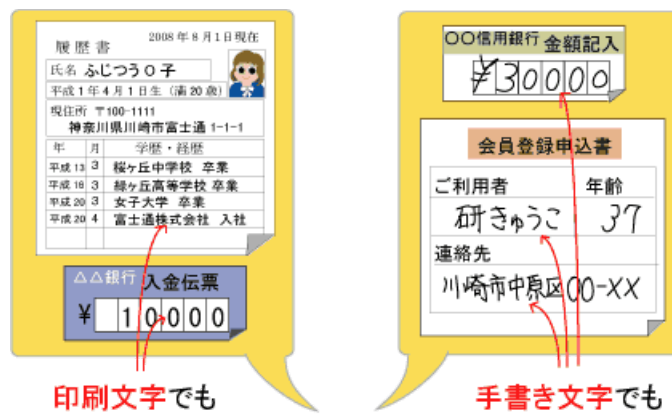


文字認識には、オフライン文字認識（紙に書かれている文字を判別する技術）とオンライン文字認識（画面に入力した手書き文字を判別する技術）があります。それぞれ紹介します。

オフライン文字認識（紙に書かれている文字を判別する技術）

オフライン文字認識

紙に印字された印刷文字や、伝票に書いた手書き文字などを、その場で読み取って、キーボードで入力しなくても自動的にパソコンに表示できます。



紙の情報を読み取って



原理

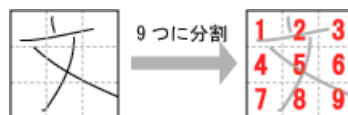
どうして文字がわかっちゃうんだろう

1. 文章を1文字ずつに区切ります。

文/を/作/る/。

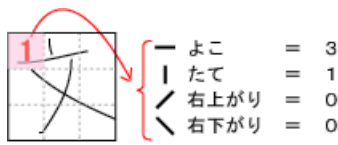
2. 1文字を分割します。

分割する数はひらがな、カタカナ、漢字等で異なりますが、ここでは仮に9つに分割します。



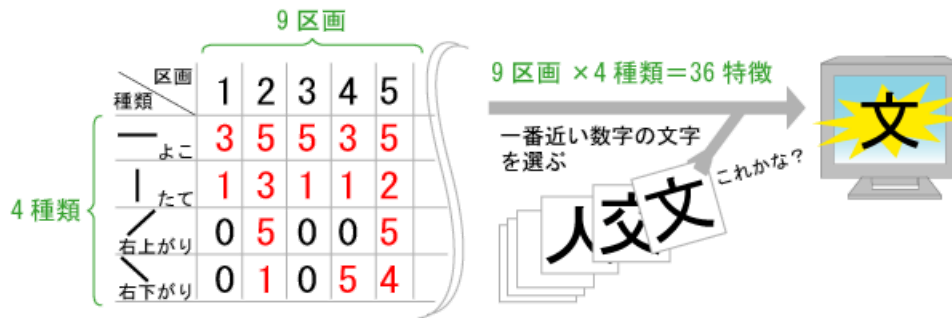
3. 1区画内の文字がどんな方向を向いているか特徴を調べます。

ここでは主な特徴として「横、縦、右上がり、右下がり」の4種類を数字で表します。数字が大きいほど、その方向を向いていることを示します。



4. 1区画4種類の特徴で表し、それが9区画あるので、全部で36コの特徴を表す数字が表示され、そこから文字を断定します。

手書き文字には個人差があるため、辞書に記録してある数字と必ずしも同じ数字ではないため、一番近い数字の文字を選び、知識（場所、名前など）をプラスして、文字を決めます。

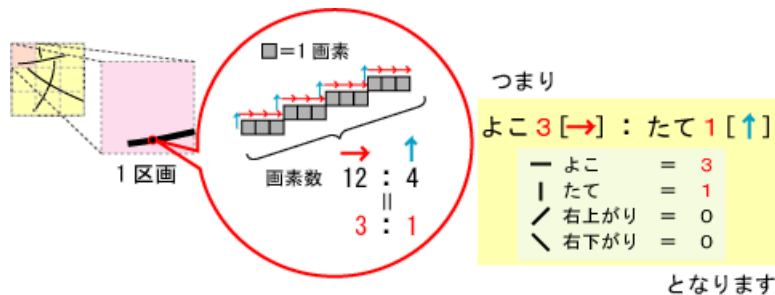


- 通常は1文字を、100～400（文字によって異なります）の特徴で表します。
- この技術によって、1秒間に認識できる文字は、（3.2GHzのパソコンを利用した場合）3080文字です。これは世界トップレベルの速さです。
- この技術を使うと、中国語や韓国語の文字も同じように認識できます。

特徴を表す数字はどうやって決まるのかな

手書き文字の線を拡大してみると、横方向に3画素並んでいて、そして上方向に1画素分上がって、また横に3画素、上方向に1画素という具合に画素が並んでいます。1区画内の横方向の画素を数えると12、縦方向には4ありますので、横縦の比率が3対1となります。この1区画の特徴は横が3、縦が1となるわけです。

（どうして「右上がり」の特徴はゼロなのかといいますと、この「文」という手書き文字の右上がり具合が、X軸から右上22.5度よりも低いので横の線と判断されたためです。）

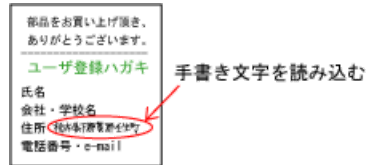


手書きの住所の文字認識

大量のユーザ登録ハガキやアンケート結果等の手書き住所をひとつひとつ入力していると大変です。そこで、当社は枠のない住所欄も認識できる技術を、世界で初めて実用レベルに達しました。従来の認識技術の場合、住所単位（町域部）の認識率は55%、当社の開発した認識技術は94%です。

どうやって手書き文字を認識しているのかな

1. 原画像を読み込みます。



2. 区切り文字を抽出します。

住所の中の都道府県や市郡町村を探し出します。

栃木県下都賀郡壬生町
県 都 郡 町

3. 区切り文字組みを決めます。

この住所の場合、3通り考えられます。例えば「県」と「府」は一つの住所に同時に使われません。また、「町」が一番最初にくることもありません。このような知識データの下に、可能性のある区切り文字組みを決定します。

4. 単語を一括認識します。

考えられる3通りの組合せ

1. 県→郡→町 栃木県下都賀郡壬生町 → ○全て存在する地名なので正解
栃木 下都賀 壬生
2. 都→郡→町 栃木県下都賀郡壬生町 → ×存在しない地名があるのでバツ
栃木県下 賀 壬生
3. 郡→町 栃木県下都賀郡壬生町 → ×存在しない地名があるのでバツ
栃木県下都賀 壬生

5. 認識結果を表示します。

結果→ 栃木県 下都賀郡 壬生町

従来技術と開発技術の違い

・従来技術

「接触文字」や「分離文字」に弱く、一文字ずつ切出して認識できない場合があります。

「栃木県下都賀郡壬生町」

県



接触文字

下都賀



分離文字「都」

切出し失敗

・開発技術

単語単位でまとめるので、地名データを利用し認識することができます。また、一文字切出しによるエラーがありません。

栃木県下都賀郡壬生町

住所の区切りとなる文字の抽出

下都賀

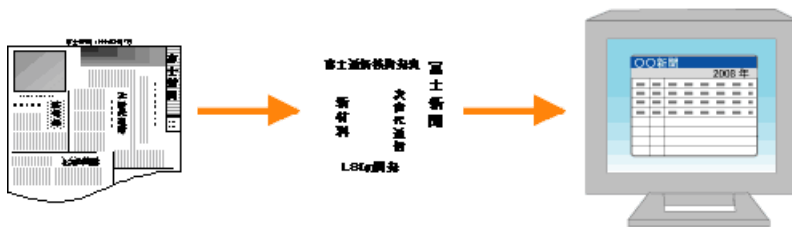
区切りの間の文字を一括認識

新聞や雑誌、カラー画像の文字認識

新聞や雑誌の文字認識

・タイトルを抽出し、文字の背景を除去

新聞の管理情報を自動登録でき、入力の簡易化を実現しました。



新聞の情報を自動登録、検索 OK!

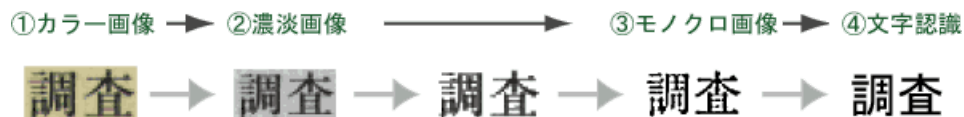
A3 200dpi新聞紙紙面 (Pentium166Mhz)		
タイトル領域抽出	94%	5.8秒
背景模様の除去	96%	0.5秒
トータル	90%	6.3秒

カラー画像の文字認識

・認識方法

カラー画像は、低解像度画像のため色ずれがあるので、色情報を使うより、濃淡情報を使います。

1. 原画像を読み込みます。
2. 文字を粗く抽出し、背景を除去します。
3. 低解像のため元々無い画素を、隣同士の画素から予測した新しい画素を加えて高解像にします。1文字単位で「白・黒」に判別する値(しきい値)を決めて、モノクロにします。



・従来方法

モノクロにした時、字がつぶれてしまいます。



・文字正読率

文書画像の解像度	開発技術	従来法
150dpi	97.5%	94.3%
100dpi	91.3%	80.2%

処理時間：1秒/A4サイズ（150dpi文書画像、500MHzのパソコンを使用した場合）

応用

文字認識は色々な場所で使われています。

銀行で

種類の多い手書きの伝票や印刷された伝票の文字をスキャナで読み取って、そのままデータとして取り扱えます。

また、押したばかりの印鑑をスキャナで読み取ると、朱肉が機械についてしまいます。そこで、デジタルカメラで撮影した画像データから文字認識して、データとして扱えます。



オフィスや家庭で

手書きのメモから書類まで様々な紙に書かれた文字、そのままではどこに何が書いてあるか探すのは大変です。

そこで、スキャナで読み取って文字認識し、テキストデータとして管理していれば、そのメモをパソコンで検索できます。また、キーワードを入力すれば目的の情報を見つけられます。自分で作るMyデータベースですね。



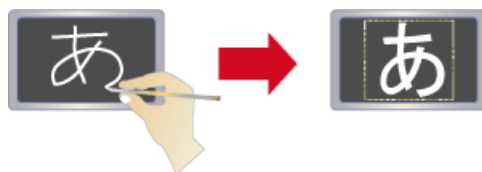
オンライン文字認識（画面に入力した手書き文字を判別する技術）

パソコンやワープロ、携帯ゲーム機等に、キーボード入力ではなく、専用のペンなどを使って直接書き込んだ文字を認識し、表示できます。

パソコンの例で紹介します。

表示入力一体型

画面に直接ペンをあてて、書き込みができるタイプです。

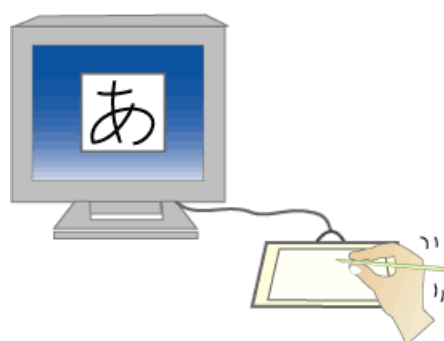


タブレット方式	電磁誘導式	タッチパネル式
ペン	専用ペンのみ使用可能	先端が尖っていれば使用可能
画面に手をつくこと	可能	画面接触面積が大きければ可能
書きやすさ	良好	パネル特性により筆跡途切れあり

表示入力分離型

パソコンに接続したタブレット上に書き込みをするタイプです。

パソコンにUSBなどで接続して、マウスや専用のペンを使って書き込みます。



オンライン文字認識の原理

原理

「研究所」という手書き文字を認識する例をご紹介します。

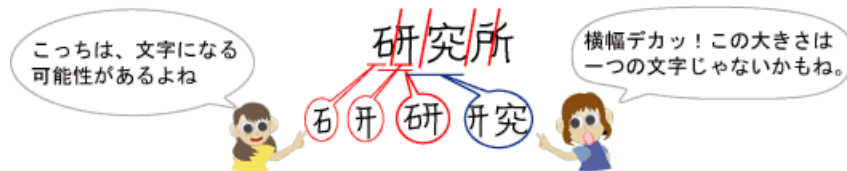
- ↓ 1. 「研究所」という手書き文字を入力します。
- ↓ 2. 手書き文字を1文字になる可能性のあるパーツに分けます。
- ↓ 3. パーツごとに辞書から可能性のある候補をオフライン文字認識とオンライン文字認識から選び出します。
- ↓ 4. 文字同士の「つながりやすさ」を考えます。
- ↓ 5. 一番可能性の高い文字の並びを選びます。
- ↓ 6. 「研究所」という文字の並びが選ばれます。

1. 「研究所」という手書き文字を入力します。



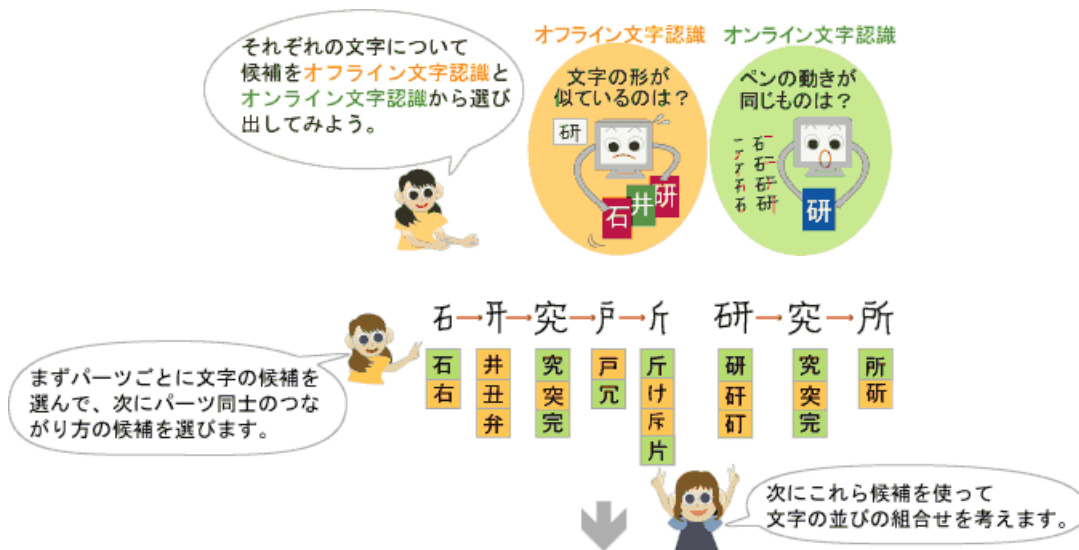
2. 手書き文字を1文字になる可能性のあるパーツに分けます。

例えば、研究所の「研」の漢字は左右に分かれて1文字となる可能性もあるので、左右のパーツにも分けられます。

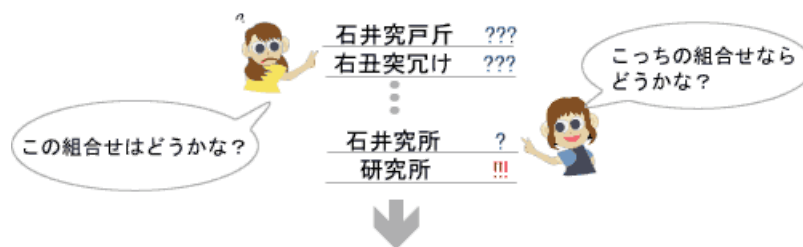


3. パーツごとに辞書から可能性のある候補をオフライン文字認識とオンライン文字認識から選び出します。

- オフライン文字認識は、書き順に関係なく、文字の形だけを見て判断します。
- オンライン文字認識は、書き順も記録してある辞書を使って、文字を書くときのペンの動きを見て判断します。
(手書き入力の際に、書き順を意識する必要があります)



4. 文字同士の「つながりやすさ」を考えます。記録してある辞書と比較します。



5. 一番可能性の高い文字の並びを選びます。



6. その結果「研究所」という文字の並びが選ばれます。

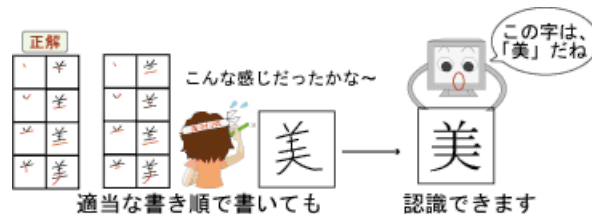


富士通研究所のオンラインの文字認識率が高い理由

文字認識はパソコンに記録されているの辞書の中から、可能性のある候補を提案して、その中から一番可能性が高いものを選びます。その時、オフライン・オンラインのどちらかだけで候補を出すよりも、両方から候補をだし、その中で一番可能性が高いものを選ぶと、より認識率が高くなります。

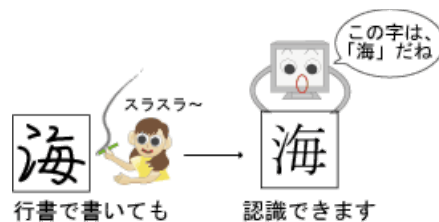
<オフライン文字認識の良いところ>

書き順が違っていても、書いた文字の形があっていれば認識できます。



<オンライン文字認識の良いところ>

行書などの崩し文字でも、ペンの運びで判断しているので認識できます。

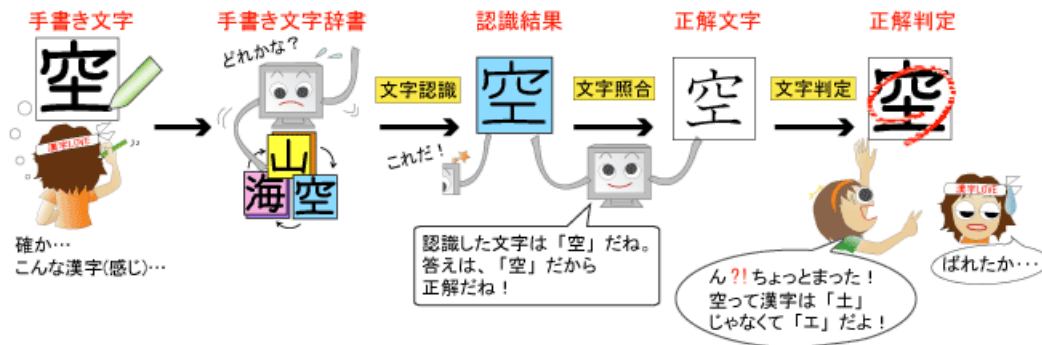


文字評価技術を使った自動採点（富士通オリジナル）

文字の正しさを評価する技術で、誤字の指摘を含めた自動採点ができます。文字認識技術を使った従来の一般的な自動採点と比較説明をします。

文字認識を使った一般的な自動採点

手書き文字を記録してある辞書と照合して一番類似した文字を文字認識により選出します。この時、手書き文字の一部が誤っていた場合でも、辞書から選出した文字が正しければ、正解として採点されます。

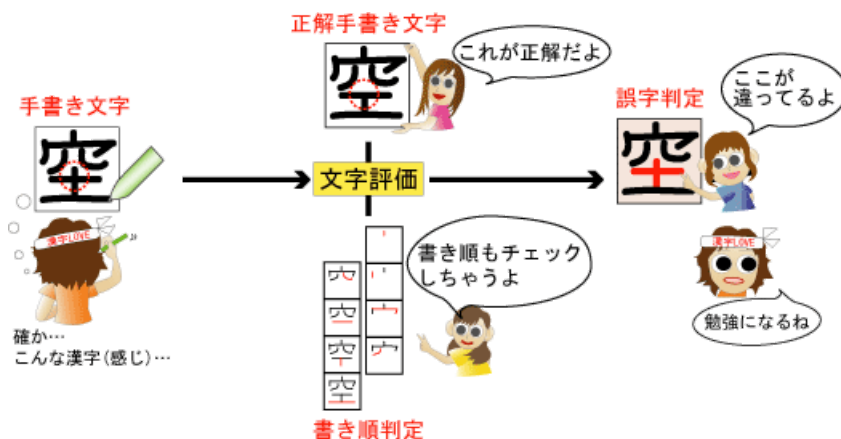


富士通オリジナルの文字評価技術を使った自動採点

読み取った手書き文字を正しい手書き文字と照合して評価します。正しい書き順、形になっていなければ誤字と判定します。

書き順も記録してあるオンライン文字認識の辞書を使うので、文字の形だけでなく、書き順なども正誤チェックができます。

記入後、すぐに正誤判定がフィードバックされるので、達成感があり、やる気に繋がります。



オンライン手書き学習ソフトを使いやすくするための工夫

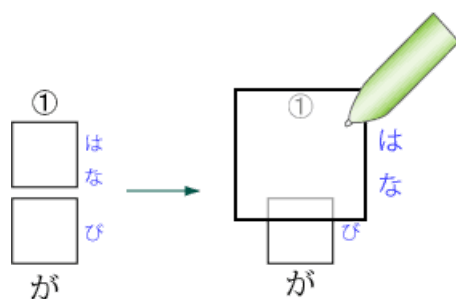
文字評価技術を使ってオンライン手書き学習ソフトを作ることができます。オンライン手書き学習ソフトを使いやすくするために、様々な工夫をしました。

- ↓ 枠を大きく表示
- ↓ スクラッチアウト (消しゴム機能)

- ↓ 間違った理由ごとに色分け
- ↓ 正解表示
- ↓ 書き順練習
- ↓ 枠外への記入

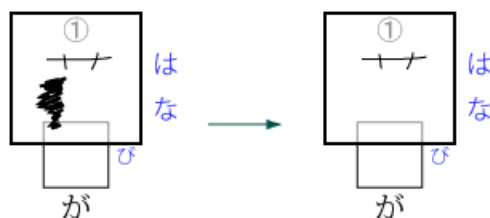
枠を大きく表示

ペンを近づけると枠が拡大表示されます。元のままのサイズでは、記入しづらいのでペンを記入欄に持つていくと枠が自動的に大きくなって記入し易くなります。
(画面に近づけるだけでペンの位置を認識するタイプのみ)



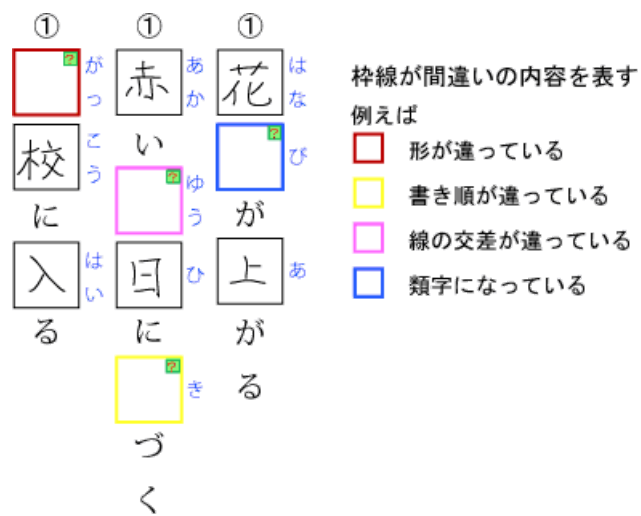
スクラッチアウト

間違った部分を塗りつぶすことで削除できます。通常は、消しゴムツールを選んで消したい場所を指定しますが、消しゴムツールを選ばずとも消したい部分を塗りつぶすだけで消えます。



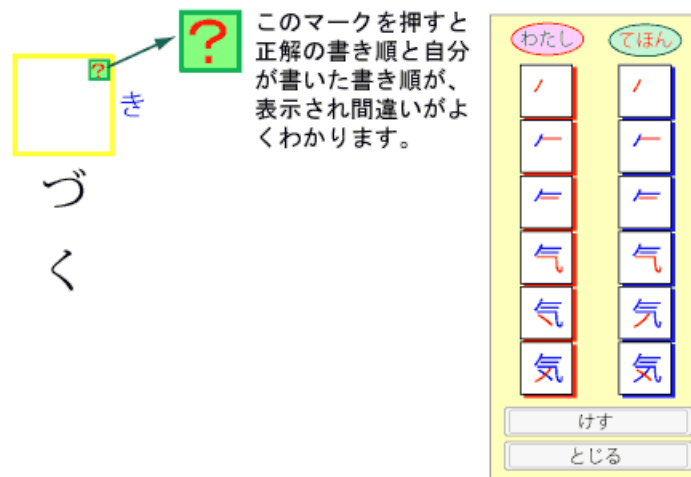
間違った理由ごとに色分け

評価判定した後、間違った理由に合わせて枠の色を変えています。間違い直しをする際にどんな間違いをしていたのか、一目瞭然でわかりやすくなっています。



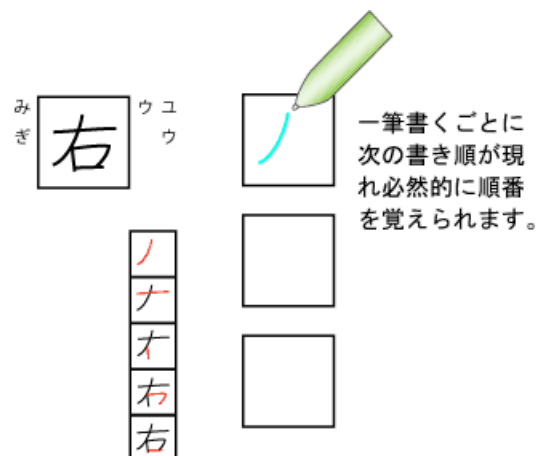
正解表示

間違っただけには、「？」マークが表示されます。例えば、マークを押すと正しい書き順が表示されます。



書き順練習

一筆ごとに書き順が表示されるので、必然的に書き順を覚えられます。



枠外への記入

枠以外の場所へのメモが可能です。

例えば、筆算など計算をする際に、繰り上がりの数字を欄外に記入しておくことができます。記入欄もカーソルなどで指定する必要はなく、ペンでそのまま入力できます。

$$\begin{array}{r} 78 \\ + 64 \\ \hline \end{array}$$

筆算で計算をする際に繰り上がりの数字を欄外に記入できます。

1 4 2

小話

世界の数字のお話

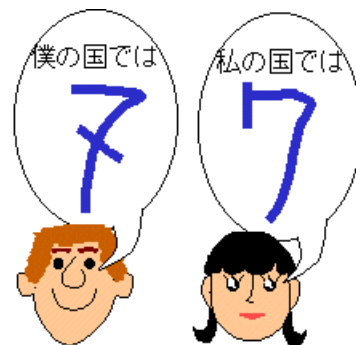
『数字』というのは、一見同じようであっても、国によって微妙に書き方が違ったりします。

特に「7」は最も違いがでるようです。

例えばこんな実話があります。

日本在住のアメリカ人が、日本で銀行口座を作ろうとしたところ、不備があるという事で受け付けてもらえませんでした。原因は、暗証番号記入欄に数字「7」を書いた際に、縦の長い線のところに横棒をちょこっと入れて記入した事が原因でした。日本ではこのような横棒が入る「7」は認められないため、横棒が入らない「7」に書き直して、無事受け付けてもらえました。

このように同じ文字でも、国による特長、個人のクセ字などにより微妙に変化がでます。しかし、それらを判別し、世界レベルで文字認識の技術が使えるよう、研究員は各国の『数字サンプル』収集をしています。



1 2 3 4 5 6 7 8 9 10 フランス

1 2 3 4 5 6 7 8 9 10 ドイツ

1 2 3 4 5 6 7 8 9 10 ウクライナ

1 2 3 4 5 6 7 8 9 10 オランダ

1 2 3 4 5 6 7 8 9 10 日本

1 2 3 4 5 6 7 8 9 10 日本

この人の家は「5丁目」だそうです、ダイレクトメールはいつも「5丁目」で来るそうです

1 2 3 4 5 6 7 8 9 10

日本
(弓道用)

関連ページへのリンク

プレスリリース

- ＞ 業界初、レイアウト定義不要な非定型帳票の認識技術（2006年10月16日）
- ＞ 世界初、HSV空間を利用した「リアルカラー認識技術」を開発（2006年2月21日）

製品

- ＞ カラーイメージスキャナScanSnap(スキャンスナップ)

その他

- ＞ 富士通の研究開発