

高精度オンライン枠なし文字認識技術

1文字単位の認識結果の組合せで文字列の認識性能を向上

オンライン文字認識技術とは、コンピュータやスマートフォンの画面上に指やペンで直接筆記した文字を機械が認識する技術です。これまでコンピュータの画面に原稿用紙のような文字枠を表示し、文字枠の中に1文字ずつ筆記する枠あり方式が一般的でした。近年、文字枠をなくし、白紙に筆記するように文字列を自由に書ける枠なし方式が広まってきました。しかし、1文字ずつ認識できる枠あり方式と比べ、枠なし方式は文字間の切出し位置がわからないため、認識性能が悪くなります。

今回、東芝は、枠なし方式の認識性能の向上を目指し、新しい方式を開発しました。

オンライン文字認識技術

文字認識技術には、紙に筆記された手書きや印刷の文字を、紙をスキャンすることで画像データとして認識するOCR (Optical Character Recognition) 技術と、コンピュータやスマートフォンの画面上に指やペンで直接筆記された文字を、筆跡の時系列データとして認識するオンライン文字認識技術の2種類があります。

OCR技術は、現在までに、ドキュメントリーダーや郵便の宛名読取り装置など、社会システムへの搭載を中心に広く実用化が進んでいます。

一方、ここで述べるオンライン文字認識技術は、パソコン、スマートフォン、携帯ゲーム機、電子辞書など、パーソナル機器への搭載を中心に実用化されてきました。また、紙への筆記と

同時に筆跡の座標情報をリアルタイムで取得できるデジタルペンが開発され、これにより、紙に筆記された筆跡に対しても紙をスキャンすることなくオンライン文字認識を行えるようになりました。オンライン文字認識技術の応用分野は、今後も更に広まっていくことが予想されます。

オンライン枠なし文字認識技術

オンライン文字認識技術は、これまで文字枠の中の1文字ずつを筆記する枠あり方式が一般的でしたが、近年、文字枠をなくし、文字列を自由に筆記できる枠なし方式が広まってきました。枠なし方式では、筆記者は白紙上に文字を書くように、制約なく自由に文字列を筆記できます。

オンライン枠なし文字認識技術では、文字間の切出し位置がわからない

ため、以下のような手順で認識処理を行います (図1)。

- (1) プレセグメンテーション 入力筆跡を実際の文字よりも細かい基本セグメントに分割する。
- (2) 候補ラティス生成 認識の候補文字列をグラフ構造化した候補ラティスを生成する。候補文字列中の各候補文字は基本セグメントを組み合わせる構成される。
- (3) 経路探索 候補ラティスから、1文字単位の文字認識の評価値などに基づいて最適な経路を探索する。今回、東芝は、前述の手順のうち経路探索処理に着目し、枠なし文字認識技術の性能向上を目指しました。

候補ラティスでの経路探索の問題点

文字認識の評価値を用いて候補ラティ

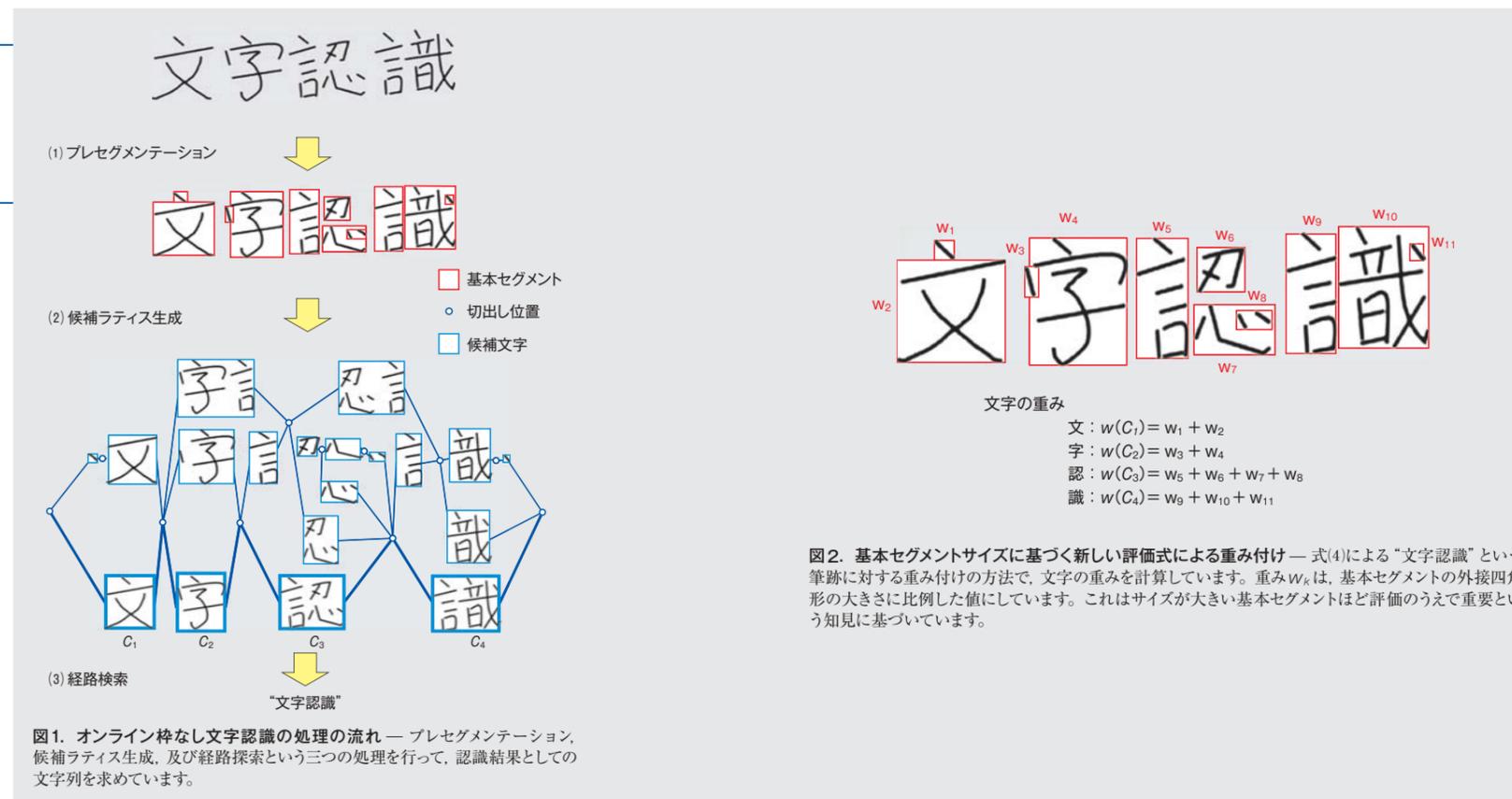


図2. 基本セグメントサイズに基づく新しい評価式による重み付け — 式(4)による“文字認識”という筆跡に対する重み付けの方法で、文字の重みを計算しています。重み w_k は、基本セグメントの外接四角形の大きさに比例した値にしています。これはサイズが大きい基本セグメントほど評価のうえで重要という知見に基づいています。

ス中の経路を評価するための評価式を定義し、その評価式の値が最大となる経路に対応する文字列を認識文字列として決定しています。

これまで、1文字ごとの文字認識の評価値を足し合わせた値を用いた評価式が提案されていました (式(1))。

$$f(C) = \sum_{i=1}^n S(C_i) \quad (1)$$

ここで、 $S(C_i)$ は経路中の文字 C_i に対する文字認識の評価値、 n は経路中の文字数、 C は経路中の文字列 C_1, C_2, \dots, C_n を表しています。

式(1)は加法性を持っているため、動的計画法を用いることで、基本セグメントの個数に対して線形時間で最適経路を探索できますが、文字数が多い経路ほど評価値が高くなるという問題がありました。

式(1)の問題に対して、経路ごとの文字数 n で正規化する評価式が提案され

ています (式(2))。

$$f(C) = \frac{1}{n} \sum_{i=1}^n S(C_i) \quad (2)$$

式(2)では、評価値は文字数に依存しないものの、加法性を持っていないため、動的計画法を用いて経路探索できないという問題がありました。

基本セグメントのサイズで重み付けされた評価式

当社は、これまでの評価式における二つの問題、すなわち評価式が文字数に依存すること、及び動的計画法による最適解探索の適用ができないことを同時に解決するために、文字認識の評価値に対して基本セグメントのサイズに比例した重み付けを行う評価式を考案しました (式(3))。

$$f(C) = \sum_{i=1}^n w(C_i) S(C_i) \quad (3)$$

ここで、 $w(C_i)$ は文字 C_i に対する重みで、文字 C_i を構成する k_i 個の基本セグメントに割り当てられた重み $w_{j_1}, \dots, w_{j_1+k_i-1}$ の和で求めることができます (式(4))。

$$w(C_i) = \sum_{k=j_i}^{j_i+k_i-1} w_k \quad (4)$$

図2では、“文字認識”の文字の重みを計算しています。式(3)では、評価式での重みの合計は経路によらず一定になるため、経路中の文字数に依存しなくなります。また、評価式は加法性を持っているため、動的計画法を用いて最適な経路を探索できます。

この評価式を用いることで、文字数で正規化を行うこれまでの評価式 (式(2)) に比べて、認識性能を約6.4%向上させることができました。

今後の展望

1960年代に始まった文字認識の実用化の歴史は、現在に至るまで着実に続いています。その中で、オンライン文字認識は、携帯ゲーム機、スマートフォン、デジタルペンなど対象となる機器が近年増加しています。今後も、人が機械を意識せずに、白紙の上のように制約なく自由に文字列を筆記したいというニーズは、増え続けることが予想されます。

今回は、文字認識技術の評価値に基づく評価式を工夫することで、認識性能を向上させることに成功しました。今後は、文字間の位置関係や、文字と文字のつながりやすさといった言語的な知識も活用することで、認識性能の更なる向上を目指していきます。

登内 洋次郎

研究開発センター
マルチメディアラボラトリー主任研究員