

# 要 旨

## OLCR 辞書による筆順復元過程を用いた詳細識別

小松 圭祐

仮説検証型の構造解析を用いた類似文字における詳細識別の手法を提案する。本手法はリジェクト文字を対象とし、候補文字の OLCR 辞書を用いて入力文字パターンの筆順復元を行い、OLCR 辞書のストロークと入力文字パターンのストロークとの間でストローク対応をとることにより識別を行う。これは候補文字を正解カテゴリと仮定し、ストローク対応によりそれを検証しようとする仮説検証型の識別手法となっている。それにより従来手法では困難であった不良設定問題を解決しようとしている。本手法ではまず、入力パターンに細線化を施し、端点、分岐点、交点、屈曲点といった特徴点を抽出する。次に各特徴点に接続しているストロークの角度や長さを用いてストローク探索し筆順復元を行う。識別部においては入力パターンと候補文字との間で近さの尺度を求め、近い方を識別結果とする。類似文字対「句, 旬」「問, 間」「束, 東」のリジェクトパターンを対象とし、認識率 74.9%を得ることができ、類似文字に対する本手法の有効性を確認できた。

キーワード 文字認識, 詳細識別, 構造解析, 筆順復元, OLCR 辞書, 類似文字

# Abstract

## Detail classification using stroke restoration process on OLCR dictionary

Keisuke Komatsu

This paper proposes detail classification method using structure analysis based on the hypothesis-verification type. At first, the method performs the stroke restoration of an input character pattern using OLCR dictionary of a candidate obtained as reject. Second, detail classification is performed by taking stroke correspondence between OLCR dictionary and input pattern. The method uses technique of the hypothesis-verification type which assumes the candidate is correct category and verifies the hypothesis by stroke correspondence. Using the technique, the method tries to solve ill-posed problem which has been assumed to be difficult to solve until now. At first, the method manipulates thinning to an input pattern and extracts feature point: that is, edge point, bifurcation point, crosspoint and bending point. Next, restoration of stroke is performed by tracing stroke trajectory using an angle and length of a stroke connected to each feature point. This method calculates the measure between the input pattern and the candidate character then output the candidate whose measure is small as final result. On the experiment of similar characters (句, 旬), (問, 問) and (束, 束) for reject samples, the recognition rate of 74.9% were obtained. These results shows the proposed method is effective for similar character recognition.

**key words** character recognition, detail classification, structure analysis, stroke restoration, OLCR, similar characters