

# 要旨

## 特徴選択を行う遺伝的アルゴリズム における適応度の評価

古井 健裕

本論文では、特徴選択を行う遺伝的アルゴリズム (Genetic Algorithm : GA) における適応度のロバスト性を評価する。類似文字を識別するには類似文字間の差分を抽出することが必要である。特徴には様々な変動が生じると、正しい識別ができなくなるという問題がある。本論文では類似文字識別において特徴変動が生じたときでも正しい認識ができる特徴を GA によって自動的に発見できるであろうかという問題提起を行う。GA においては適応度が重要な役割を担っている。適応度を変えることにより選ばれる特徴も変わってくる。ここでは適応度としてパターン分離の尺度を表す F 比, 類似度, 距離差を用いた。GA 実行時に特徴変動を模擬したノイズを付加する場合と付加しない場合の両方について特徴選択を行った。選択された特徴に対しノイズを付加して認識を行うことによりロバスト性を評価した。類似文字“王”, “玉”に適用した結果, ノイズの無い環境下で GA を実行した場合, 適応度が F 比, 類似度, 距離差の順にロバスト性が高いという結果を得た。逆にノイズ有りの環境下で GA を実行した場合は適応度が距離差, 類似度, F 比の順に高くなる結果を得た。このことより類似文字識別において特徴変動にも強い特徴選択を行うには適応度として距離差を用い, ノイズ環境下で GA を実行することが効果的であると言える。

**キーワード** 文字認識, 類似文字識別, 遺伝的アルゴリズム, 適応度, ロバスト性

# Abstract

## The evaluation of fitness value on GA selecting feature selection

Takehiro Furui

This paper evaluates robustness of fitness value in genetic algorithm(GA) which performs feature selection. Extraction of difference between similar-shaped characters is necessary to distinguish similar-shaped characters. When various variation appear in the feature, serious problem that correct recognition cannot be done is caused. This paper addresses the problem whether GA can automatically discover the feature which can recognize similar-shaped character when feature variation occurred. Fitness value plays an important role in GA. The selected feature changes also when fitness value changes. We select a F ratio, similarity and distance difference, which represent degree of pattern separation, as fitness value. Feature selection was performed by both GA with noise and GA without noise. Robustness is evaluated by the task that recognition is performed with noise added to the selected features. The result that robustness was high in order of the F ratio, the similarity and the distance difference was obtained. On the contrary, on experiments using GA without noise for similar-shaped characters “王” and “玉”, the result that robustness becomes high in order of the distance difference, the similarity, and the F ratio was obtained under the condition using GA with noise. The result shows that GA with noise using the distance difference as fitness value is effective to select feature which is robust to a feature variation for similar-shaped character recognition.

***key words*** Character recognition, Similar-shaped character recognition, Genetic algorithm, Fitness value, Robustness