

1. 目的

ベイズ識別関数を用いて認識実験をする。

2. 実験方法

字種ごとの平均特徴量，固有値・固有ベクトルを用いて，認識対象文字の特徴量と各字種に対するベイズ識別関数を計算するプログラムを作成し，認識実験を行った。

ベイズ識別関数は，

$$G_k(\vec{x}) = \alpha \sum_{k=1}^n \frac{\{(\vec{x} - \vec{m})^t \vec{e}_k\}^2}{\lambda_k + b} + (1 - \alpha) \log \left(\prod_{k=1}^d \lambda_k \right)$$

の式で求めた。ただし，本実験では

$b = 1 \times 10^3$ (assets/javascript/Assignments.js | 242行目)

$\alpha = 0.9$ (assets/javascript/Assignments.js | 232行目)

$d = 140$ (assets/javascript/Assignments.js | 235行目)

と設定した。括弧内にはプログラム中でこの値が指定されている位置を記した。

3. 実験結果

“う”の文字の認識率が90%，他は100%だった。そのため，全体の認識率は99.78%である。

4. 考察

“う”を“ら”・“つ”と誤認識していた。“ら”は一つ棒が多く，“つ”は一つ点が少ないだけで，見た目が“う”に非常に似ているため，誤認識を起こしてしまった。しかし，前回マハラノビス距離で認識実験を行った際は，認識率100%だったため，パラメータの設定に不十分があると考えられる。

そこで， $b = 10$ としたところ，“う”での誤認識がなくなり，認識率が100%となった。このとき，認識対象文字とそれぞれの文字との間のベイズ識別関数距離は，認識対象文字自身とは100前後，その他の文字とはその1桁程度大きい値となっていた。