知識工学実験

＜認識手順＞

・512byteのデータを16\*4の64個のデータに分けて，それぞれをユークリッドで距離を計算し，認識した．

＜データ収集＞

・字種ごとのデータ認識率　（図１）

・データ全体に対する平均認識率

(3922）\*(9200)\*100=42.630435

・誤認文字の例

例）

　　「お」と「ま」が認識率が異常に非常に低かった．具体的には，

「お」が，16.5%となっており，2割を切っていた．

そして，「ま」は，9%となっており，1割を切ってしまっていた．

＜考察＞

・自主ごとのデータ認識率が，全体的に思っていた以上に悪かった．サンプルデータで実験をしたときは平均認識率はなんと8割もいっていた．しかし，実験データでは4割というとても低い結果になっていた．

・平均認識率の低い文字に関しては，「お」と「ま」となっていて，そのどちらも，黒画素が多く，また広く分布していて，区別をつけるのが難しいのだと考えた．

・具体的な誤認文字としては，「お」であれば，「あ」や，「め」，「ぬ」などがあると思われる．また，「ま」は「は」，「ほ」などととても似ているので，おそらくはその文字と誤認したのだと思われる

図１：文字の認識率のグラフ