

ニューラルネットワークを用いた音声による男女認識

電子情報学類 情報システム 0852040051 226 3年 澤口光生

1. 研究課題

ニューラルネットワークを用いて、ある音声は男女どちらの性別の音声かを予測する。

2. 研究方法

- 音声データの収集
サークルの男女各4名づつに47都道府県を発音してもらい録音する。
- 録音した音声データをニューラルネットワークに組み込むために前処理を行う
- 変換したデータをニューラルネットワークに組み込んで実験を行う。

3. 実験手法

今回の実験方法では大きく分けて2種類の実験を行った。

- ① 男女各3人づつの女性すべてを入力データとして、使用していない男女各1名の音声データをすべてテストデータとする。
- ② 男女全員の音声の一部を学習に使用して、残りの音声データをテストデータとする。
そして同様な実験を数回行い、正答率の平均値を出してそれを最終正答率とする。

4. 研究結果

- ① 男女各3名づつの音声を使用

男123：女123	男4の正答率：56.5%	女4の正答率：93.6%
男234：女234	男1の正答率：69.6%	女1の正答率：61.7%
男341：女341	男2の正答率：58.7%	女2の正答率：78.7%
男412：女412	男3の正答率：87.0%	女3の正答率：59.6%

- ② 音声の一部をテストに使用

全音声の2/3を学習に使用	男の正答率：73.3%	女の正答率：82.2%
全音声の1/2を学習に使用	男の正答率：74.5%	女の正答率：79.6%
全音声の1/3を学習に使用	男の正答率：71.1%	女の正答率：74.2%

5. 考察

全体の正答率を見てみると、女性のほうが高いことに気がついた。これは録音した音声のデータが高い声・低い声がバランスよく選ばれていたためでないと思われる。

県名の一部を使用した時のほうが比較的高判別率が得られた。これから、発音している言葉が違ったとしても、発音している人物が同じであれば判別は容易であると思われる。

そして以上より、学習データ数を増やすことができれば判別率の向上につながると思われる。