

ニューラルネットによる表情からの情動認識

理工学部 数理情報学科

T980048 小西 徹朗

指導教員 佐野 彰

概要

人が表情のみからその内側にある情動を判断する際、過去の経験に基づいて表情を切り分け、その中から当てはまる情動を選択していると考えられる。この切り分けには微妙な個人差が存在するものの、人によって大きな違いはないであろう。その結果、今日では「表情の基本形」と呼ばれるものや、視覚的に識別可能なすべての顔の動きを記述するシステム(顔面動作符号化システム(FACS))などが考案された。これは、表情をパターン化して扱えることを意味している。

本研究では、「怒り」・「悲しみ」・「喜び」・「驚き」の4つの標準的な表情パターンをニューラルネットワークに学習させ、その後、ある未知の表情を与えることによって、ニューラルネットワークと人の表情認識における違い、認識に至るまでの過程の違いを明確にするための実験を行った。

本実験では視覚的情報が重要であるため Java による表情認識アプレットを作成した。アプレットに提示される顔画像は線画であり、表情の認識に重要な「目」・「眉」・「口」は可変とし、認識に大きな影響を及ぼさない「顔の輪郭」「鼻」は固定されたものとした。学習に用いるニューラルネットワークの構造は入力層3、中間層4、出力層4の3層構造で、学習法として慣性項付き誤差逆伝播法(バックプロパゲーション)を用いた。

実験を行った結果、人がそれぞれの情動を識別するときの、「目」・「眉」・「口」の3次元空間上で表情の分布、学習データの数により異なる収束までに要する学習回数と収束の様子、未知データに対するニューラルネットワーク、及び人の識別結果を得ることができた。

以上の結果から、情動状態により生じる「表情の基本形」、誤差逆伝播法による結合荷重の変化の有効性、ネットワークと人との情動認識の違いと認識に至る過程の違いを明確にすることができた。本実験によって、ニューラルネットワークにおいても数多くの学習データを与えてやれば表情による情動の認識は高い確率で可能である一方、人とは認識の手段が大きく異なることが分かった。同時に、表情をパターン化するということの困難さも明らかとなった。