

# ポテンシャルネットによる基本表情認識

4P-4

松野 勝弘\* 李 七雨\*  
\*(財)イメージ情報科学研究所

辻 三郎\*\*  
\*\*大阪大学基礎工学部

## 1. はじめに

科学、技術の高度な発達により、優れた処理能力を持つ機械が製造されるようになった。しかし、操作する人間と機械とのギャップはますます大きくなっている。このギャップを埋めるために、近年、インタラクティブなマンマシンインターフェースの研究がさかに行なわれている[1]。このようなインターフェースを構築するには、機械は、ユーザが誰であるかを知り、その感情や意図を正しく理解することが必要である。このことが実現されると、機械はユーザの反応を確かめ、必要なサポートをすることができ、人に優しいマンマシンインターフェースを構築することができる。本稿では、テレビカメラで撮影した1枚の画像から、表情を認識する手法について述べる。

## 2. 顔表情認識

エクマンらの表情と感情を結び付ける研究[2]では、幸福、怒り、驚き、悲しみ、嫌悪、恐怖の6つの基本的な表情を取り上げ、これらの感情と表情のマッピングは万国共通であることを実験的に確かめている。

計算機を用いた表情認識の研究は、特徴点を抽出する方法[3]、筋肉の動きをオプティカルフローより検出する方法[4]が提案されているが、特徴量の自動抽出が前提となっており、安定した抽出法の検討が必要である。このように、計算機による表情自動認識は困難であり、その問題点を挙げてみると以下ようになる。

- (1) 表情の変化の定式化が困難である。
- (2) 特徴量の抽出、追跡が困難である。

これらの問題を解決するには、顔画像より、個人性を抑制し、表情構成要素を個別に捕えるのではなく、顔全体の表情の特徴をパターンとして捕えることが必要である。我々は、顔全体にある表情構成要素(眉、目、口、鼻等)の特徴を総合的に捕えることができるポテンシャルネットというモデルを提案している[5]。本稿では、さらに検討を行なったので、そのことについて述べる。

## 3. ポテンシャルネットによる表情の解析

### 3.1 ポテンシャルネットの概要

ポテンシャルネットは、ネットに各ノードの分布がある系のポテンシャルを表す物理モデルである。その構造を図1に示す。外郭のノードは枠に固定さ

れ、中心部分のノードは4近傍のノードにバネを通して結合されたグリッド構造である。従って、あるノードの位置変化はネット全体の構造に影響を及ぼす。その程度はネットを構成するバネの性質に依存するが、変形力はバネの弾性係数が小さいほど遠くまで伝達される。

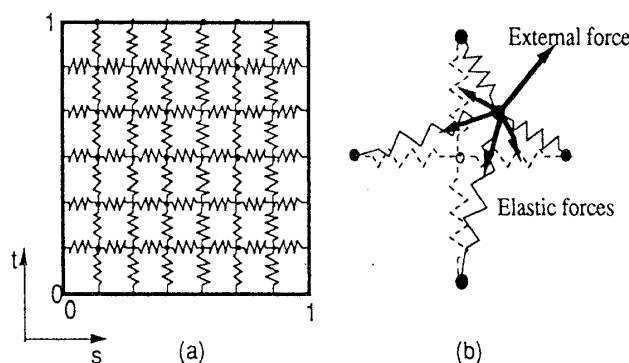


Fig 1. The structure of Potential Net ;  
(a) Potential Net and node array  
(b) Balance of forces at a node

### 3.2 ポテンシャルネットの収束

ネット上に設けられた有限個のノードにノード番号 ( $i = 1 \sim M, j = 1 \sim N$ ) を付け、ノード ( $i, j$ ) の座標を  $\vec{n}_{i,j} = (x_{i,j}, y_{i,j})^T$  とする。適当な外力が加えられた場合の各ノードの運動方程式を表すと

$$m_{i,j} \frac{d^2 \vec{n}_{i,j}}{dt^2} + \gamma \frac{d \vec{n}_{i,j}}{dt} + \vec{F}_{spring} = \vec{F}_{ext} \quad (1)$$

となる。ここで、

$m_{i,j}$ : ノードの質量、 $\gamma$ : 減衰定数、 $\vec{F}_{spring}$ : 弾性力  
 $\vec{F}_{ext}$ : 外力である。

この外力を用いて、運動方程式(1)を解くことになるが、動的問題として解を求めようとすると各項は時間的に変化するので解析が困難になる。そこで、静的問題として考え、運動方程式(1)の平衡状態を考える。

$$\vec{F}_{spring} = \vec{F}_{ext} \quad (2)$$

式(2)は、弾性力と画像からの力(外力)の平衡状態を求めることになる。弾性力は式(3)で表される。

$$\vec{F}_{spring} = k \sum_a \left( \frac{|\vec{l}_{i,j,a}|}{|\vec{l}_{i,j,a}|} - l_0 \right) \frac{\vec{l}_{i,j,a}}{|\vec{l}_{i,j,a}|} \quad (3)$$

$$\bar{l}_{i,j,1} = \bar{n}_{i,j} - \bar{n}_{i+1,j} \quad \bar{l}_{i,j,3} = \bar{n}_{i,j} - \bar{n}_{i,j+1}$$

$$\bar{l}_{i,j,2} = \bar{n}_{i,j} - \bar{n}_{i-1,j} \quad \bar{l}_{i,j,4} = \bar{n}_{i,j} - \bar{n}_{i,j-1}$$

ここで、 $k$  はバネの弾性定数、 $l_0$  はバネの自然長である。画像からの力は式 (4) で表される。

$$\bar{F}_{ext} = b(\nabla(G_\sigma * I(x_{i,j}, y_{i,j}))) \quad (4)$$

ここで、 $I(x_{i,j}, y_{i,j})$  はノード  $(i, j)$  の画素の濃度値、 $b$  は重み係数である。次に、式 (2) を次のように変形して、式 (5) を定義する。

$$\bar{F}_{total} = \bar{F}_{ext} - \bar{F}_{spring} \quad (5)$$

各ノードは2つの力が釣り合う方向、即ち、合力  $\bar{F}_{total}$  が0の方向に移動する。この処理を全ノードに対して繰り返し実行し、各ノードの安定な位置を求める。顔画像の特徴情報がバランスを保ちながら、ネットは、ある形状に変形する。ノードの最終位置の座標を  $\bar{n}^e_{i,j} = (x^e_{i,j}, y^e_{i,j})^T$  とすると、初期位置の座標  $\bar{n}^0_{i,j} = (x^0_{i,j}, y^0_{i,j})^T$  からの変位ベクトルを  $\bar{v}_{i,j}$  で表す。

$$\begin{aligned} \bar{v}_{i,j} &= \bar{n}^e_{i,j} - \bar{n}^0_{i,j} \\ &= (x^e_{i,j}, y^e_{i,j})^T - (x^0_{i,j}, y^0_{i,j})^T \\ &= (\Delta x_{i,j}, \Delta y_{i,j}) \end{aligned} \quad (6)$$

パラメータ ( $k, b$ ) を調整することにより、個性を抑制した表情の特徴情報を持った変位ベクトルを求めることができる。

### 3. 3変位ベクトルによる表情の識別

メディアンフィルタにより、画像の雑音を除去した顔画像からマウスで顔領域を指定し、大きさの正規化を行ない、微分フィルタを用いて各点の微分値を計算した後、図1(a)のネットを顔画像上にはり、前節の手法によりネットを収束させ、変位ベクトルを求める。あらかじめ、作成しておいた表情データベース ( $\bar{v}_{dbase}$ ) と入力顔画像の変位ベクトル ( $\bar{v}$ ) との類似度を比較することにより、表情の識別を行う。

表情データベースは、10人の4表情(幸福、怒り、驚き、悲しみ)の変位ベクトルの平均(平均変位ベクトルと呼ぶ)と平均より少し正及び負の方向にずれた変位ベクトルである。すなわち、1つの表情について9つの変位ベクトル、合計で36の変位ベクトルを作成しておく。表情データベースの顔画像の1例を図2に示す。



Fig. 2 Examples of 4 facial expressions

ここでは、 $K$ -最近傍法を用いる。実験では、 $K=9$ として、9つの最近傍になるクラスから多数決

で識別を行なう。

$$D(\bar{v}_{dbase}, \bar{v}) = \sum_i \sum_j |\bar{v}_{(dbase)_{i,j}} - \bar{v}_{i,j}| \quad (7)$$

予備実験により、パラメータは  $k=2.0$ 、 $b=0.5$  とした。平均変位ベクトルに対応するネットを図3に示す。

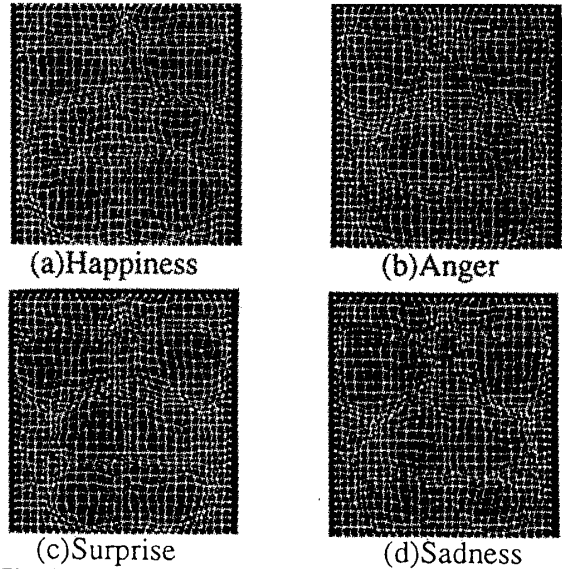


Fig.3 The model nets for happiness, anger, surprise and sadness

### 4. 実験結果

表情顔画像は、CCDカメラにより、テープに取り込み、その後、表情が最も良く表れていると思われる1コマを取り出し、それを用いて実験を行なった。マンマシンインターフェースの入力の1要素として考えると、表情のデータは、幸福、怒り、驚き、悲しみの4種類で十分であると思われる。今回は、表情データベースの作成に用いた人物とは異なる6人の24枚のデータについて識別実験を行ない、識別率は88(21/24)%であった。

### 5. おわりに

本論文では、ポテンシャルネットを用いる表情認識手法に関して述べ、ポテンシャルネットが表情認識に有効であることを示した。

### 参考文献

- [1]重点領域研究「感性情報処理の情報学・心理学研究」第2回全体会議予稿集
- [2]P. Ekman and W.V. Friesen, "Facial Action Coding System", Consulting Psychologist Press(1977)
- [3]小林他「ニューラルネットによる人の基本表情認識」、HC90-29
- [4]間瀬他「表情動画像からの感情の認識の1手法」、PRU91-29
- [5]松野他「画像からの感情の抽出法の検討」、情報処理第46回全国大会