

## 画像からの感情の抽出法の検討

## 9C-6 - ポテンシャルネットによる顔表情の認識 -

松野 勝弘\*      李 七雨\*      辻 三郎\*\*  
 (財) イメージ情報科学研究所\*      大阪大学\*\*

## 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もつ機械が簡単に製造できるようになった。しかし、それを操作する人間の能力には限界があり、機械と人間能力のギャップはますます大きくなりつつある。従って人間と機械との距離感を克服するための研究は今後情報工学において大きな課題である。そこで最近、人間の感情または感性を工学的に分析し、その情報を情報工学の一要素として取り入れる動きが始まっている[1][2]。我々はこのような背景に基づいて動画より人間の表情、行動などを理解し、それに反応するビジュアルエージェントの構築を目指して研究を進めている。本稿ではパターン認識の一手法として使えるポテンシャルネットというモデルを提案し、それによる表情認識手法に関して記述する。

## 2. 顔表情認識

人間は他人と対話をするとき多数の情報獲得手段を用いる。その中でも相手の顔表情は大きな情報を伝えてくれる。もし可能であれば人間のこの能力を機械に与えることこそ究極マン・マシン・インターフェースを作ることになるだろう。しかし、顔画像から表情を認識することは非常に難しいことである。それは表情情報自体が定量的なものよりも定性的な側面をもつものであり、それに関わる情報源が何かはまだはっきりしてないからである。また顔画像から顔の各部品を取り出すことが非常に難しいことも理由のひとつである。表情の生成においてよく使われるアクションユニット[3]のような手がかりが画像から得られれば安定な認識結果が出ると思われるが、顔画像は他の画像に比べて一般的に特徴抽出が難しい。そこで、我々は、以下のような方針に基づいてアルゴリズムの開発を行なう。

- ・ 2次元情報のみを用いる。
- ・ 予め登録されたモデルデータを知識として使う。
- ・ 統計的手法を用いる。
- ・ 静的情報と動的情報を融合する。

## 3. ポテンシャルネットによる表情認識

ポテンシャルネットは、ネットの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。これの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を持つ。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。もしこのノードに画像からの変形力を求めて加えることが可能であれば、変形された各ノードの位置は画像を識別する情報源になりえる。

## 3.2. 画像力とポテンシャルネットの変形

Kassらは画像の濃淡分布よりポテンシャルを求めて、画像中の対象物から輪郭線を抽出する方法を提案した[4]。それによるとモデルが本来持っている

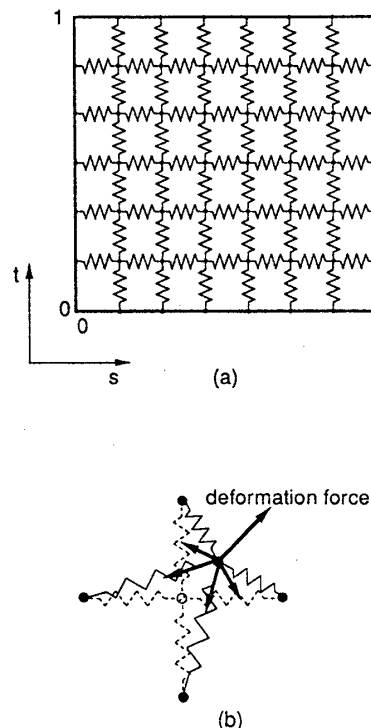


fig 1. The structure of Potential Net ;  
 (a) The Potential Net space and node array  
 (b) Deformation force on a node of potential net

A Study on Emotion Extraction from Facial Images

-Recognition of Facial Expression by Potential Net-

Katsuhiro MATSUNO\*, Chil-Woo LEE\*, Saburo TSUJI\*\*

\*Laboratories of Image Information Science and Technology

\*\*Osaka University

特性（弾性、接続構造のような物理的性質）をもちながらモデルを変形させることができる。本論文では、エネルギー関数を式(1)で定義し、エネルギーが最小になるように各ノードの位置を求める。

$$E = E_{int} + E_{image} \quad (1)$$

式(1)は、内部エネルギーと画像からのエネルギーで構成されている。

内部エネルギーは式(2)で表される。

$$E_{int} = k(l - l_0) \quad (2)$$

式(2)は、バネに撻る力のベクトル量を求めている。 $k$ はバネの弾性係数、 $l$ はある位置でのバネの長さ、 $l_0$ はバネの自然長である。

画像からのエネルギーは式(3)で表される。

$$E_{image} = -\alpha |\nabla(G_0 * I(x,y))| \quad (3)$$

式(3)は、ボケ変換を行い、その濃度勾配を求めている。 $\alpha$ は重み定数である。

ノードエネルギーの最小値は、8近傍探索法により求める。すなわち、各ノードは、式(1)のエネルギーを計算して、エネルギーが一番最小となる方向へ移動する。一つのノードに関して考慮すると画像から得られる変形力とネットが持つ反発力の和が最小になる方向にノードは変形される。

ノードの位置はそのノードが安定するまで繰り返して更新される。処理中で安定したノードの数がある閾値以上になるとネットは安定したと判断し更新処理を終了する。

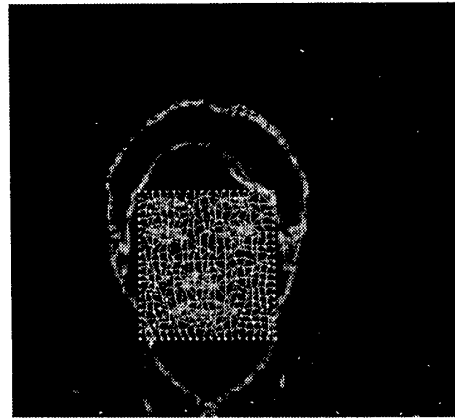
### 3.3. ネットの方向ベクトルと認識

変形されたネットのノードの形は一つのパターンとして扱うことができる。そのパターンの形は局所的画像情報をネット全体に広まった意味を持っている。従って、登録された表情ポテンシャルネットの方向ベクトルと入力された画像の方向ベクトルとのユークリッド距離によって認識を行なう。具体的には、式(4)を最小にするクラス（表情）を求める。

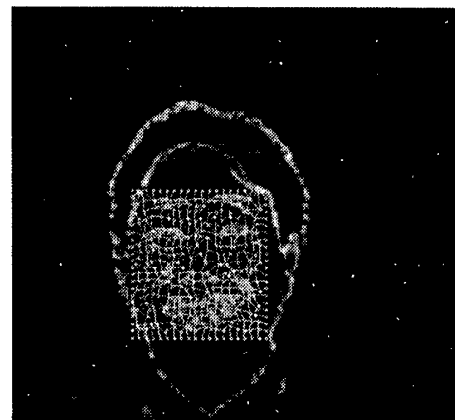
$$D = \vec{V}_{new} - \vec{V}_{class} \quad (4)$$

## 4. 実験結果

今回の実験では正面を向いている「笑う」、「平常」の2種類の顔画像を用いた。動画画像から2表情に相当する画像をデータ画像として登録し、3.3節で記述した処理を行なう。イメージから求めた変形力によってノードは次第にその位置が変わる。図2からわかるように画像のエッジ付近でノードは固定されその影響が他のノードまで及ぶ。これは画像の特徴情報が全体的にバランスを保ちながらある形に固定されたためである。



(a)



(b)

fig 2. The State of Potential Net

(a) Normal (b) Laugh

## 5. おわりに

本論文ではポテンシャルネットというモデル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。今後は、この手法を動画画像に適用し、更に多くの表情について検証を行なう予定である。

## 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1回公開シンポジウム資料
- [2]間瀬他「表情動画画像からの感情の認識の1手法」、PRU91-29
- [3]P. Ekman and W.V. Friesen, "Facial Action Coding System", Consulting Psychologist Press(1977)
- [4]M.Kass 他"Snakes: Active contour models", IJCV,1,4,1987,321-331