

## SOMによる顔方向認識を応用した頭部ジェスチャ認識システム

服部 公央亮<sup>†</sup>, 田口 亮<sup>†</sup>, 保黒 政大<sup>‡</sup>, 梅崎 太造<sup>†</sup>名古屋工業大学<sup>†</sup> 中部大学<sup>‡</sup>

## 1. まえがき

画像を用いたセンシングの中で、顔のセンシングは利用時の抵抗感の少なさ、インタフェースの簡便さ等の特徴を有しており、様々な分野での応用が期待されている<sup>[8]-[10]</sup>。特にロボットビジョンなどでの応用を考えた場合、人の視線方向に対応する顔の向きという情報は、自然なインタラクションを返す際に非常に有用な情報として使用できる。

我々は、入力画像に対する次元圧縮を行わずに、連続的な動きをニューラルネットワークの一種である自己組織化マップ(SOM)により補間することで、少ない学習データから高精度に姿勢推定を行う手法を提案している。本稿では、これを連続的に動画に対して適用することで、首振りや傾きなどのジェスチャ認識を行うシステムを提案する。

## 2. 自己組織化マップを用いた姿勢推定

姿勢推定の手法としては、2次元画像、もしくは3次元情報を用いた手法が主に行われている。一般的に、3次元情報を用いた手法は高精度だが、データの取得にコストがかかるという問題がある。本稿では、3次元データ取得のための処理や装置が必要無く、システムが簡易なもので済むため、実用面において優位性が高い2次元画像のみによる手法を検討した。

姿勢変動のような変化が連続的であることを利用して、2次元画像を低次元のベクトルに圧縮した上で、その固有空間上で照合や姿勢推定を行う手法として、パラメトリック固有空間法が提案されている。しかし、主成分分析などにより得られた空間は、画像集合の次元圧縮という意味では最適であるが、パラメータ推定を行う上で最適な空間とは言えないことが知られており、パラメータ推定のための様々な空間生成法が検討されている<sup>[24][25]</sup>。

我々は、入力画像に対する次元圧縮を行わず

一種である自己組織化マップ(SOM)により補間することで、少ない学習データから高精度に姿勢推定を行う手法を提案している。これにより、2次元画像を用いながらも従来研究に比べて分解能の高い推定を実現している。使用するSOMの構成を図1に示す。

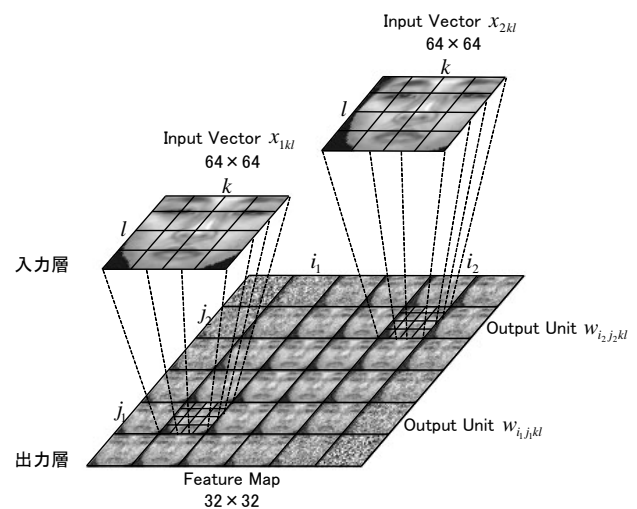


図1 SOMの構成

SOMの学習過程について説明する。まず、各ユニットの結合重み $w_{ijkl}$ に小さな値をランダムに与えて初期化する。次に(1)式より全ユニットと入力画像 $x_{kl}$ との類似度 $u_{ij}$ を求め、 $u_{ij}$ が最大のものを類似度最大ユニットとする。最後に類似度最大ユニットを中心とした学習半径 $d_s$ 内のユニット結合重みを $x_{kl}$ に近づくように(4)式で修正する。ここで、 $d_i$ と $d_j$ は最適整合ユニットから更新する重みまでの $i$ 方向と $j$ 方向の距離( $d_s > d_i$ ,  $d_s > d_j$ )、 $T$ は総学習回数、 $t$ は現在の学習回数、 $\alpha(t)$ は学習係数である。また、 $X$ は任意の変数である。通常SOMでは学習が進むにつれて学習半径を小さくするが、本稿では重み更新時に(2)式を用いる。類似度最大ユニットからの距離が離れるに従い重み更新量を減少させることで、なだらかな変化が生じることが先行研究により確認されている<sup>[15]</sup>。学習係数は、

(3)式を用いて更新することで、学習が進むにつれて小さくなる．重みの更新には(4)式を用いる．

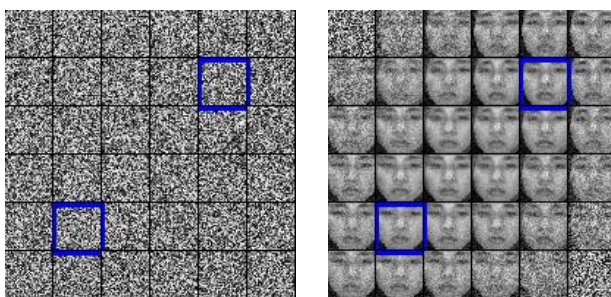
$$u_{ij} = \frac{\sum x_{kl} w_{ijkl}}{\sqrt{\sum x_{kl}^2 \sum w_{ijkl}^2}} \dots\dots\dots (1)$$

$$w'_{ijkl} = w_{ijkl} + f(d_i) f(d_j) \alpha(t) (x_{kl} - w_{ijkl}) \dots\dots\dots (2)$$

$$f(X) = \cos\left(\frac{\pi X}{2d_s}\right) \dots\dots\dots (3)$$

$$\alpha(t+1) = \alpha(t) \left(1 - \frac{t}{T}\right) \dots\dots\dots (4)$$

更新に伴い、初期のランダム値から学習画像に近づいてゆき、さらに学習画像の類似度最大位置間に補間されたデータが生成されてゆく．



(a) 初期化状態 (b) 更新時  
図2 自己組織化マップ学習の様子

なお、実験の簡略化とマップの視認性の確保を目的に、輝度情報を特徴ベクトルに使用している．照明変動や不特定者へ対応するためには、前処理としてPCAなどの検討が必要である．

### 3. 姿勢推定を応用したジェスチャ認識

まず、SOMに2自由度の姿勢変動を学習することで、首振りや傾きに対応した姿勢推定用のマップを作成する．SOMに画像を学習させる際には、マップの縦軸*j*を傾き方向の動きに、横軸*i*を首振り方向の動きにそれぞれ対応させるため、マップ上のセントロイド位置を固定して学習した．中間姿勢についても、学習姿勢の中央位置に中間姿勢のセントロイド配置することで、推定可能であることを確認している．

本稿では、2自由度の姿勢変動を学習したSOMに対して、入力画像の類似度最大位置がどう動くかを判定することで、首振り・傾きなどのジェスチャ認識が可能ではないかと考えた．ジェスチャ認識を行う際には、マップの座標変動を時系列で処理することで、ジェスチャを認識する．図3～5に傾きと首振りの各動作時の座標変動を示す．各ジェスチャの動作に合わせて類似度最大位置が移動していることが分かる．この変移の縦横それぞれの大きさを比較することで、動作の判定ができる．

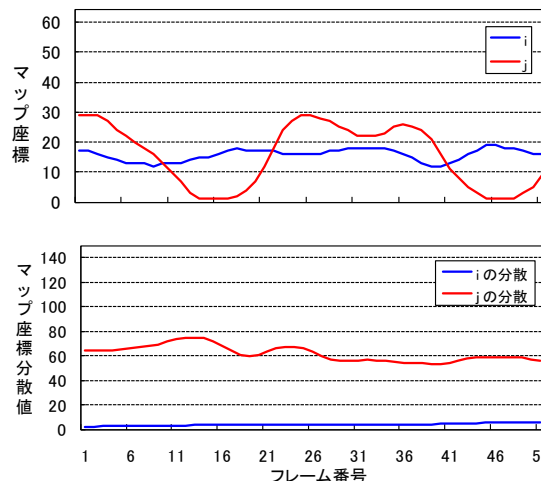


図3 傾き時の変化

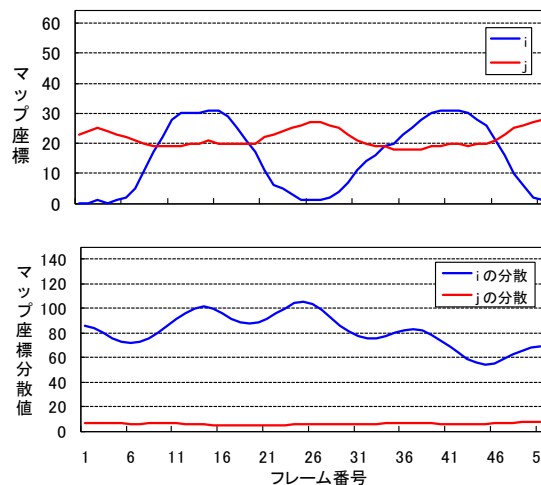


図4 首振り時の変化

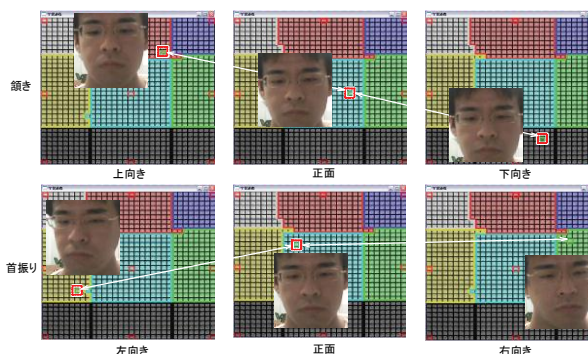


図5 認識時の様子

### 4. むすび

本稿では、「首振り」「傾き」のジェスチャ認識手法を提案した．今後、顔検出器を組み込み、SOMを用いた姿勢推定手法と併用することでより実用的なシステムを構築したい．

#### 参考文献

[1] 服部中央亮, 田口亮, 梅崎太造: 自己組織化マップを用いた顔方向の認識, 日本顔学会誌, Vol.10, No.1, pp.29-36 (2010).