

# 表情表出パターンのセグメント化による動的表情空間の可視化

## Visualization of Dynamic Space for Facial Expression by Segmentation of Time-series Expression Patterns

○大西康平\*, 佐藤和人\*, 間所洋和\*, 門脇さくら\*\*

○Kohei Onishi\*, Kazuhito Sato\*, Hirokazu Madokoro\*, Sakura Kadowaki\*\*

\*秋田県立大学, \*\*Smart Design 株式会社

\*Akita Prefectural University, \*\*Smart Design Ltd.

キーワード： 人間工学 (Ergonomics), 自己組織化マップ (Self-organizing map), Viola-Jones 法 (Viola-Jones method), ガボールウェーブレット変換 (Gabor wavelet transform),

連絡先： 〒 015-0055 秋田県由利本荘市土谷海老ノ口 84-4 秋田県立大学 システム科学技術研究科 機械知能システム学専攻 脳情報工学研究室

大西康平, Tel.: 0184-27-2000, E-mail: M20A004@akita-pu.ac.jp

### 1. はじめに

我々の日常生活においてコミュニケーションは欠かせないものである。対人コミュニケーションにおいて視線、姿勢、ジェスチャー、更に音声や言語、表情などから、個人の特徴を表す情報が得られる<sup>1)</sup>。特に、顔の表情からは人種や性別、年齢、人柄、健康状態などの様々な情報を把握することができ、人の感情や精神状態を読み取ることが可能である<sup>2)</sup>。親しい友人や家族などの間柄なら、体の調子やその時の機嫌なども表情から感じ取ることができる。対人コミュニケーションにおいて表情の果たす役割は大きい。次世代のMMI(Man-Machine Interface)として期待されている<sup>3)</sup>。

我々は、これまでに表情と心理状態との関連性を明らかにするために、ストレス刺激後の意図的表情を繰り返し表出するプロセスに着目した表出テンポとリズムという枠組みを提案した<sup>4)</sup>。表出テンポとリズムは、表情の表出程度をラベル化した時系列変化<sup>5)</sup>において、無表情から表情表出を経て無表情に戻る区間を1テンポ、テンポを複数回繰り返したものを1リズムと位置付け、快刺激時と比較して不

快刺激時の方が、「喜び」表情の1テンポを構成するフレーム数のばらつきが大きいことを示した。また、BNs(Bayesian Networks)を用いて、心理的ストレスの蓄積程度と表情表出の相互関係をグラフィカルに表現したストレスモデルを構築し、心理的ストレスの影響が現れやすい表情が「喜び」と「悲しみ」表情であることを明らかにすると共に、表情の種類によって現れやすい顔部位(目元、口元など)が異なる可能性を示した<sup>6)</sup>。更に、情動喚起ビデオによる快・不快刺激後の「喜び」表情の表出過程に着目し、予め決められた写像空間の中で相対的にカテゴリ分類を行うSOM(Self-Organizing Map)を用いて表情パターンの分類を行い、人間の心理状態に起因する表情表出時の複雑性や曖昧性を客観的に表現することを試み、表出リズムの相互情報量の有効性を示した<sup>7)</sup>。ここでは、意図的に表出した表情を使用しており、予めカテゴリ数を決定した上で分類している。しかし、メンタルヘルスの状態を日常生活の延長線上でチェックするという観点では、自然に表出する表情の変化から、ストレスの蓄積程度を定量化したストレスレベルにより可視化することが重要となる。

人の表情は、時系列で変化し、顔の形状が人物それぞれ異なることから、その特性に応じて修正していく学習メカニズムが必要となる。

本研究では、対人コミュニケーションにおける人の自然な表情変化に着目する。自然な表情は個人固有なカテゴリ数が存在するため、事前に決定できない問題点がある。そこで、従来の SOM に時系列データを含む学習を加えた手法である RSOM(Recurrent SOM)、ノードの成長と階層化の機能を加え、自動で最適なマップサイズを決定する GHSOM(Growing-Hierarchical SOM) を用いることで、時系列変化に応じた新しい表情の追加が可能な動的表情空間を定義し、可視化を試みる。

## 2. 関連研究

表情に関する研究は、画像認識を用いることが多く、加藤<sup>8)</sup>らは、女子学生の顔画像から平均顔を作成し、眉の位置のみを上下に操作した顔画像に対する印象評価を試みた。平均顔は嗜好性の評価が高く、目と眉の間隔を狭めた顔は活動性の印象が強まり、間隔を広げた顔は洗練性の印象が低下することが明らかになった。益子<sup>9)</sup>らは、様々な笑いの表情について、表情の変化の度合い、快・不快感情の度合いの2軸から分類し、パターンごとに平均顔を作成し、変化量を詳細に測定することにより、分類された笑顔間の、物理的な相違を明らかにすることを試みた。眉、目、鼻、口の各移動距離、指示角度の変化から、笑いの違いによる表情変化、部位によって物理的な変化量が異なることが明らかになった。しかしながら加藤、益子らとともに、被験者の故意による表情表出であり、対人コミュニケーションによる自然な表情表出を扱うまでに至っていない。

小原らは、質問応答システムのログを対象に、GHSOM による可視化を行い、システム保守者の負担軽減及び解析の助けとなるツールの作成を試みた。二次元マップ上にログデータを配置することで、その内容の把握が容易となる。また階層化により、各マップは適切なサイズにとどまり、データ量に関わらずコンパクトにまとまったマップが形成できた。表示部分に HTML(Hyper-Text Markup Language) を使用することで、従来の SOM マップよりも自由度の高いマップとなり、システム保守者の解析効率を上

げる効果が見込まれる。

予め決められた写像空間中で相互的にカテゴリ分類を行い、客観的に類似性を検証することができる SOM を扱い、データ間のクラスタリング、エンジン噴射スケジュール問題の解消に試みた従来研究は存在するが、表情表出パターンに着目し、個々人に最適な2次元空間の定義を試みるといった、表情画像に着目した従来研究は比較的少ない。吉見<sup>10)</sup>らは、平面 SOM において中央付近に配置されたデータ周辺の類似度は判定が容易だが、マップの端付近においてはその類似度の判定が難しいため、球面 SOM を用いたパレート解集合の可視化手法を提案した。球面 SOM は、平面 SOM と比較してマップの端でゆがみがないことが報告されているだけでなく、マップの端が存在しないという特性を有するために、各データ間の合理的な位置関係を表示でき、クラスタリングを行う際にも高い精度が得られることが報告されている。多目的最適化の実問題であるディーゼルエンジン噴射スケジュール問題で得られたパレード解集合を利用して、球面 SOM のパレード解表示における有用性を検討した。その結果から、平面 SOM では正常に配置できないデータが球面 SOM では配置可能であることを示した。小笠原<sup>11)</sup>らは、多重解像度解析と SOM を組み合わせた多重解像度木構造形 SOM により、階層型画像特徴の学習アルゴリズムと回路化について提案した。木構造形 SOM 特徴学習部と木構造形テンプレートをを用いる画像認識部を回路化時に統合することで回路規模の削減を実現した。多重解像度木構造型 SOM による特徴学習回路は画像特徴を階層的に学習できたが、ビット数を増やすといった学習係数の実装方法の検討が必要である。

## 3. 実験環境と手順

図1に実験環境を示す。実験は被験者(図1の奥側男性)と対話者(図1の手前側女性)による3分間のコミュニケーションにより行う。実験室は、本大学学部棟 G-514 である。実験3分間で被験者の表情変化を、Transcend 社製の DrivePro520 を用いて、左右2方向から撮影する。画像データは30fps で保存し、解像度は前方:縦 1080×横 1440、後方:縦 720×横 1280 である。会話の話題は自由とし、本研究で

扱う被験者の自然な表情変化を捉えることに努めた。また、被験者、対話者にストレスがなるべくかからないよう、実験室内の空調を調整した。被験者として、男子大学生5名、女子大学生5名の計10名に協力してもらい、11月7日から12月25日の約2ヶ月、週1回ペースの計8回実験を行った。今回、対話者と面識がある被験者Aと、対話者とほとんど面識がない被験者Eの実験1回目の11月14日に行った実験データを対象に解析を行う。なお、本実験は、本大学研究倫理審査委員会の承認を得て行った。



Fig. 1 実験環境

## 4. 提案手法

### 4.1 分類手法

実験により取得した動画データを、Viola-Jones法<sup>12)</sup>らを用いて顔オブジェクト検出し、探索窓の領域ごとに、Cascade構造の識別器でもって顔領域を判定し、画像対象領域である顔画像を生成する。顔画像の解像度は $160 \times 160$ pixel、フレームレートは15fpsである。奇数枚目の顔画像をトレーニングデータ、偶数枚目をテストデータに分別する。トレーニングデータを教師データとし、テストデータと比較し学習させる。学習により識別された表情画像をガボールウェーブレット変換し、解像度を $80 \times 80$ pixel、画像サイズを1/4にする。その後、RSOMによって学習データをユニット数25でセグメント化し、GHSOMにより動的表情マップを生成する。図2に提案手法の処理フローを示す。

### 4.2 RSOM

SOM<sup>13)</sup>とは、Kohonen氏が提案した教師なし学習によるNN(Neural Network)であり、複数の因子を持つ高次元データの集合に対して、データ同士の近似を表す2次元空間を写像する。これによりSOMで複数の多次元データを可視化することが可能となる。RSOMは、時系列データを扱う上で適するよう、SOMを拡張した手法である<sup>14)</sup>。今回、時系列の表情表出パターンを分類するために用いる。

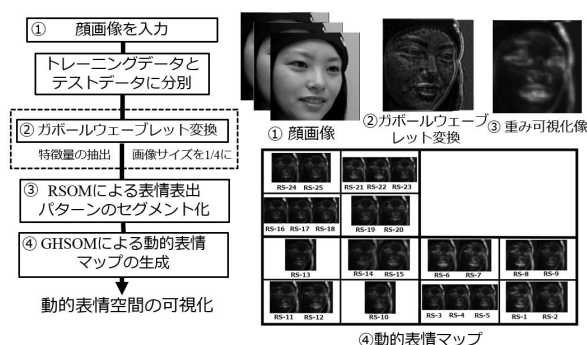


Fig. 2 提案手法の処理

### 4.3 GHSOM

GHSOMとは、SOMにマップ成長と階層化を付与した手法であり、学習の過程で自動的にマップの構造を決定する。マップ成長の調整パラメータ (breadth 値) と階層化の調整パラメータ (depth 値) を制御することにより、類似する表情表出パターンが近傍に写像される。今回、RSOMによって分類された表情表出パターンを動的表情空間で可視化し、個人固有の動的表情空間を定義するためにGHSOMを用いる<sup>15)</sup>。

## 5. 実験結果と考察

### 5.1 被験者Aの考察

図3に動的表情マップの階層化、図4にRSOM、GHSOMによって生成された被験者Aの動的表情マップを示す。図4は、セグメント化された25の重み可視化像をGHSOMによって、特徴が類似している画像同士を近傍に写像した2次元空間であるマップ1層目のマップサイズは $2 \times 2$ のサイズ4であり、

1層目の各ユニットが階層化された2層目の各マップサイズも $2 \times 2$ のサイズ4を表す。2層目は、1層目の近似表情画像から、更に異なる表情を分類した層である。つまり、笑顔と判別された画像から、笑顔の種類をさらに分類したものとイえる。図3ように1, 2層目とも左上から左下, 右上, 右下までをUnit1~Unit4とする。1, 2層目の各マップのいずれも、近傍に分布したユニットどうしは類似している表情画像である。ただし、動的表情マップの1層目Unit3は未発火のため、マップ成長も階層化も行われていない。



Fig. 3 GHSOMによる階層化

図5に被験者Aの時系列表情画像を示す。画像左上が実験開始時の表情画像である。右下画像が実験終了時の被験者Aの表情画像である。表情画像の1枚に着目すると、正常のRGB値の画像と、RGB値が逆転した画像が図5中に存在する。本論文では、それぞれポジティブ画像、ネガティブ画像と定義する。時系列表情画像中のポジティブ画像は、図4の動的表情マップの1層目Unit4重み可視化像の時系列画像を表す。ネガティブ画像は、1層目のUnit1,2重み可視化像の時系列画像を表す。図4の枠線で囲まれた画像は図4のRS-3からRS-7の重み可視化像の時系列表情画像を表している。ここで図5のRS-6, RS-7の時系列表情画像1枚に着目する。またRS-3, RS-4, RS-5の時系列表情画像1枚にも着目する。図4の上部にRS-6, RS-7とRS-3, RS-4, RS-5の時系列表情画像1枚を示す。まずRS-6, RS-7の画像に着目すると、「笑顔」と判別できる。また、RS-3, RS-4, RS-5の画像に着目すると、図4のように「笑顔」とは異なる表情（驚いた表情）と判別できる。また、図5の時系列表情画像中のポジティブ画像に着目すると、連続しているポジティブ画像の集団が確認できる。これは、対話者とのコミュニケーションにおける被験者Aの表情変化が大きく、異なる表情が明確に分別されたためと考える。



Fig. 4 被験者Aの動的表情マップ

図3に被験者Aの動的表情マップ(RS-3,RS-6)を示す。動的表情マップのRS-3, RS-6の重み可視化像に着目すると、RS-6の重み可視化像は、時系列表情画像と同様に「笑顔」と判別できる。また、RS-3の重み可視化像についても「笑顔」と異なる表情（驚いた表情）と判別できる。また、時系列表情画像から、異なる表情と視覚的に判別できる画像は、RSOM, GHSOMによって生成された図3の動的表情マップ

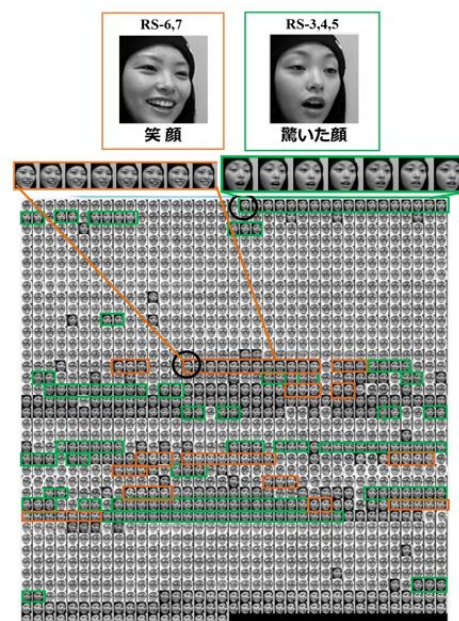


Fig. 5 被験者Aの時系列表情画像





Fig. 6 被験者 E の動的表情マップ

の重み可視化像においても当てはまる。

## 5.2 被験者 E の考察

図6にRSOM, GHSOMによって生成された被験者Eの動的表情マップを示す。被験者Eの動的表情マップと同様に、セグメント化された25の重み可視化像をGHSOMによって、特徴が類似しているもの同士を近傍に写像した。マップ1層目のマップサイズは $2 \times 2$ のサイズ4であり、各ユニットが階層化された2層目の各マップサイズも $2 \times 2$ のサイズ4である。図3と同様に1, 2層目とも左上から右下, 右上, 右下までをUnit1からUnit4とする。動的表情マップの1層目Unit1は未発火のため、マップ成長も階層化も行われていない。

図7に被験者Eの時系列表情画像を示す。画像左上が実験開始時の表情画像である。同様に左上から右に進み、一段落下の右下表情が実験終了時の被験者Eの表情画像である。

図6のポジティブ画像は被験者Eの動的表情マップの1層目Unit2重み可視化像の時系列画像を表す。図7の枠線で囲まれた画像は図6のRS-6からRS-9の重み可視化像の時系列表情画像を表している。ここで図7のRS-8, RS-9の時系列表情画像1枚に着目する。またRS-6, RS-7の時系列表情画像1枚にも着目する。図7上部にRS-8, RS-9とRS-6, RS-7の時系列表情画像1枚を示す。まずRS-8, RS-9の画像に着目すると、「笑顔」と判別することができる。また、RS-6, RS-7の画像に着目すると「笑顔」とは異なる表情（無表情）と判別することができる。また、時系列表情画像のポジティブ画像に着目すると、連続的なポジティブ画像は、被験者Aの時系列表情

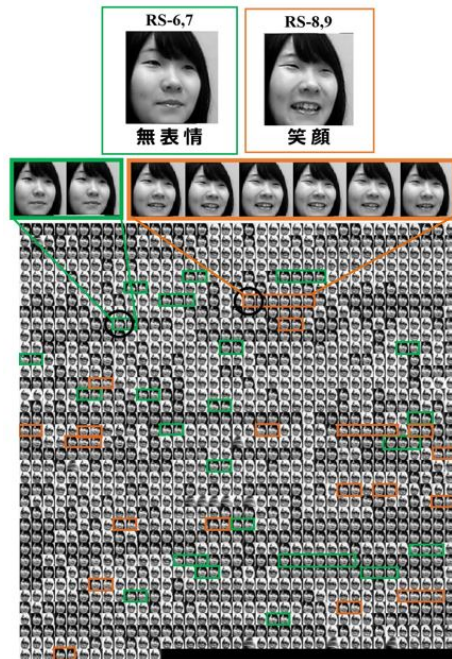


Fig. 7 被験者 E の時系列表情画像

画像と比較して少ない。これは、被験者Aが対話者と面識があり（親友）、被験者Aの表情が表出されやすい環境あるのに対して、被験者Eはほとんど面識がなく、「ぎこちない」コミュニケーションとなり、3分間の会話中の表情の変化が小さかったためと考えられる。また、元々被験者Eが被験者Aに比べ、表情変化とみなされない、わずかな顔部位の変化と判定された可能性がある。

図6の被験者Eの動的表情マップ中RS-6, RS-9の重み可視化像に着目すると、RS-9の重み可視化像は、先述と同様に「笑顔」と判別できる。また、RS-6の重み可視化像についても「笑顔」と異なる表情（無表情）と判別することができる。

図7の被験者Eの時系列表情画像から、異なる表情と視覚的に判別できる画像は、RSOM, GHSOMによって生成された図6の動的表情マップの重み可視化像においても当てはまる。

## 6. おわりに

本研究では、表情表出パターンのセグメント化による動的表情間の可視化と題して、対人コミュニケーションにおける人の自然な表情に着目し、RSOM, GHSOMを用いて、長期的な時系列データにも対応して表情変化をカテゴリ分類し、動的表情空間の可

視化を試みた。評価実験として、被験者と対話者に3分間自由な話題で会話をしてもらい、自然な表情変化を撮影し、被験者として男子大学生5名、女子大学生5名の計10名に協力してもらい、2017年11月7日から12月25日までの2ヶ月間に週1回のペースで、それぞれ計8回実験データを取得した。

本論文では、対話者との関係性が異なる被験者Aと被験者Eを対象に、表情変化の類似性や表出程度の観点から動的表情マップを形成し、各ユニット数による表情表出パターンのセグメント化を行い、動的表情空間の可視化を試みた。

- 得られた動的表情マップの重み可視化像に対応するユニットの時系列表情画像は画像の特徴が類似している
- 被験者と対話者の関係性によって、コミュニケーションにおける表情の変化が異なるため、RSOM, GHSOMによって生成される動的表情マップが変化し、結果として動的表情空間の可視化に影響を及ぼす

今後は、他実験日、他被験者の動的表情マップ、時系列表情画像を比較し、実験日が異なる被験者A, Eの動的表情空間、他実験者の動的表情空間の可視化への影響を明らかにしたい。また、10人の被験者と対話者との関係性の差異、被験者の心理状態による動的表空間の可視化への影響を明らかにする予定である。被験者の心理状態を推定する手段として、POMS試験<sup>16)</sup>が挙げられる。心理状態を把握することによって、被験者と対話者との関係性による心理状態の変化を捉え、動的表情空間の可視化への影響を明らかにしたい。

## 参考文献

- 1) 赤松 茂, "人間とコンピュータによる顔表情の認識の展望: 人間による表情認知に学ぶ", 電子情報通信学会誌, vol.86, no.1, pp.54-61, (2003)
- 2) 永田明徳, 岡崎透, 崔昌石, 原島博, "主成分分析による顔画像の基底生成と空間記述", 電子情報通信学会論文誌, D-II, Vol.J79-D-II, No.7, pp.1230-1235, (1996)
- 3) M.Pantic, L.J.M. Rothkrantz, "Automatic Analysis of Facial Expressions: The State of the Art," IEEE Trans.PAMI, Vol.22, No.12, pp.1424-1445, (2000)

- 4) 佐藤和人, 間所洋和, 門脇さくら, "一過性ストレス刺激が意図的な表情表出に与える影響", 第11回情報科学技術フォーラム, RJ-005, pp29-36, (2012)
- 5) 延谷直哉, 仲谷善雄, "パーソナルテンポを基とした音響リズム支援による会話支援システム", 情報処理学会第71回全国大会, pp.4-227 4-228, (2009)
- 6) 佐藤和人, 間所洋和, 門脇さくら, "意図的な表情表出に及ぼす心理的なストレス要因の分析", 第12回情報科学技術フォーラム, RJ-002, pp21-28, (2013)
- 7) 佐藤和人, 門脇さくら, 間所洋和, "笑顔の表出過程に着目した情動喚起ビデオによる顔部位効果の分析", 第13回情報科学技術フォーラム, RJ-002, pp13-22, (2014)
- 8) 加藤徹也, 青木滉一郎, 菅原 徹, 村上智加, 宮崎正己, "目と眉の間隔の違いが顔の印象に及ぼす影響", 日本感性工学会論文誌, vol.14, No.3, pp.419-424, (2015)
- 9) 益子行弘, 菅場奈津美, 齋藤美穂, "表情の変化量と笑いの分類の検討", 知能と情報(日本知能情報フェジィ学会誌), vol.23, No.2, pp.186-197, (2011)
- 10) 吉見真聡, 西本 要, 廣安知之, 三木光範, "球面SOMによるパレード解集合の可視化の検討-ディーゼルエンジン設計問題における事例の検討", 情報処理学会論文誌, 数理モデル化と応用, vol.3, no.3, pp166-177, (2010)
- 11) 小笠原麦, 横田 怜, 田向 権, 関根優年, "階層型特徴学習機構を有する画像認識システムの回路化", 電子情報通信学会技術研究 報告, RECONF, リンフィギュラブルシステム, no.112(325), pp 77-82, (2012)
- 12) Viola, P. and Jones, M., "Rapid Object Detection Using Boosted Cascade of Simple Features", Proc. Computer Vision and Pattern Recognition, pp.511-518, (2001)
- 13) 徳高平蔵, 藤村喜久郎, "自己組織化マップ(SOM)とその応用", 日本フェジィ学会誌, Vol.13, No.4, pp.345-355(2001).
- 14) O.Yeloglu and A.Nur Zincir, Malcolm I., "Growing Recurrent Self Organizing Map", IEEE.
- 15) Rauber, A., Merkl, D., Dittenbach, M., The growing hierarchical self-organizing map: exploratory Analysis of high-dimensional data, IEEE Transactions on Neural Networks, vol.13, pp.1331-1341, (2002)
- 16) 村松成司, 近藤健吾, 岸 恵美, 広田悠子, 齋藤初恵, "POMS テストからみた大学生の朝の心理特性と生活習慣との関連性について", 千葉大学教育学部研究紀要 50, 503-515, (2002)