

顔ロボット開発に向けた口唇部周辺の複雑で広範な皮膚の流れ場のクラスタ分析

Cluster analysis of complex and various skin flow fields around human lips for enriching robot facial expressions

太田 信行 (阪大) 石原 尚 (阪大)
正 浅田 稔 (阪大)

Nobuyuki OTA, Osaka University, nobuyuki.ota@ams.eng.osaka-u.ac.jp
Hisashi ISHIHARA, Osaka University, ishihara@ams.eng.osaka-u.ac.jp
Minoru ASADA, Osaka University, asada@ams.eng.osaka-u.ac.jp

To enrich human-like robot facial expressions, robot designers should know how much and to which directions each human facial part moves for various facial expressions. However, combinations of facial part movements are too complex to determine appropriate design parameters of robot facial deformation mechanisms such as the number, locations, and orientations of actuators for realizing intended reproducibility of each human facial expression. This paper proposes an analysis method of various human facial expressions to determine the design parameters for human-like face robots. The proposed method superimposes several measured human skin flow fields, and extracts the direction of the first principal component of the motion vectors for each facial marker. Its contribution rate was 86%, which indicates each facial part moves almost only in one direction even when a human exhibits several facial expressions. Then, the proposed method estimates which face area moves similarly through cluster analysis of a vector field of the first principal component vector of each facial marker.

Key Words: Face robot, Facial expression analysis, Cluster analysis, Flow field analysis

1 緒言

人の表情は柔らかい皮膚の複雑な流れの組み合わせによって表現され、表情表出時に顔の表面形状は多様に変化する。そして、これまで人の豊かな表情をロボットで模倣し、役立てようとする試みが様々な分野でなされてきた [1, 2, 3, 4]。しかし、これまでの顔ロボットの開発では、定石の機構が未だ確立されておらず、設計者によって試行錯誤で駆動部の数・位置・移動方向が設定されていた。また、顔ロボットが再現できる表情の種類は限定的であった。多様な表情を 1 つのロボットで実現することが困難な原因として、まず人の口唇部が複雑な構造になっていることが挙げられる。口唇部は 13 の筋肉が密に配置された構造をしており [5]、筋肉配置をそのまま 13 個のアクチュエータに置き換えることは顔ロボットのスペースの制約上、非常に困難である。また、仮に全ての筋肉をアクチュエータに置き換えられた場合でも、その上層にある皮膚の変形を再現できるとは限らない。なぜなら、多様な表情を表出した際、皮膚の流れは非常に複雑で広範かつ、顔の動きごとに大きく変化するためである。よって、表情表出時の多様な皮膚変形を再現するためには、表情表出に伴い、皮膚にどのような流れが起こるのかを把握した上で機構を設計することが望ましい。

そこで、本研究では、顔の中でも特によく動く口唇部周辺に注目し、まず多様な表情表出時の皮膚の流れ場の計測を行い、そのバリエーションを明らかにする。そして、その多様な流れをロボットで再現する上で適切な機構配置を決定するための皮膚の流れ場の分析手法を提案する。そのうえで、顔ロボットの自由度数が筋肉の数と同数の 13 以下の場合の表情の再現度を評価する。

2 顔の動きの観察

観察対象とする顔の動きは FACS(Facial Action Coding System) という表情分析手法 [6] の分類に基づき、決定した。FACS では様々な顔の動きが AU(Action Unit) として分類されている。顔下部で分類されている 16 種類の AU を観察対象として採用し

た。それぞれの AU と顔の動きは、AU9:鼻に皺を寄せる、AU10:上唇を上げる、AU11:鼻唇溝を深める、AU12:唇両端を引っ張り上げる、AU13:唇両端を鋭く上げる、AU14:笑窪を作る、AU15:唇両端を下げる、AU16:下唇を下げる、AU17:顎の先を上げる、AU18:唇をすぼめる、AU20:唇を引き伸ばす、AU22:唇をじょうご型にする、AU23:唇を締める、AU24:唇を押さえつける、AU25:唇を開く、AU28:唇を吸い込む、である。流れの観察は顔に直接マーカを配置することで行い、マーカ数は眼の下から顎にかけて 96 点となった。そして、FACS を基にした顔の動きに伴うマーカの移動をビデオカメラで撮影した。例として、AU12 の動きを図 1 に示す。顔の変形前後のマーカの移動をベクトルで表した。また、流れの計測結果は、画像から求められるピクセルの値を式 (1) を用いて顔上の実際の移動長さに対応する距離 [mm] に変換した。ただし、この変換は便宜上採用したものであり、精度は十分でない可能性がある。

$$\text{距離 [mm]} = \text{pixel 数} \times 0.46 \quad (1)$$



Fig.1 Example of a measured skin flow field with Lip Corner Puller action, which is defined as AU12 in FACS.

3 AU 表出時の頭部の移動の補正

AU を表出した際の各マーカの移動ベクトルを分析する上で、撮影時の頭部全体の動きが大きな誤差の問題となる。そこで、画像上で頭部の動きの補正を試みる。まず各 AU 内で無表情時と AU 表出時の頭部位置補正を実施し、その上で全ての AU のマーカの移動ベクトルの始点が揃うように画像補正を行った。各 AU 内での補正は図 2 のように AU 表出時に動きにくいと考えられる部分に基準点を設け、それらの位置が画像上で一致するよう画像のアスペクト比を一定にし、拡大及び移動を行った。補正の結果、基準点の誤差は平均約 1mm となった。

AU ごとのベクトルの始点のズレの補正は各 AU に対する無表情時の全マーカに対して、アフィン変換を行い、基準の無表情顔のマーカ位置に揃えた。全ての AU のベクトルの重ねあわせたものについて補正前後の結果を図 3 に示す。なお図 3 ではベクトルの分布の様子が把握しやすいよう、角度に基づいて色付けしている。ベクトルの起点のズレは補正前が全マーカ平均 3.0mm であったのに対し、補正後は 0.9mm まで改善された。

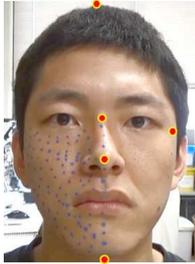
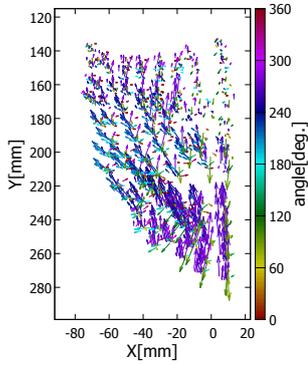
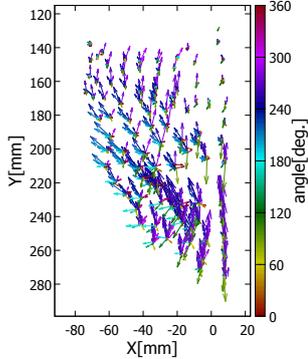


Fig.2 Reference points adopted in order to compensate differences of the size, location, and angle of the head among facial images.



(a) Without compensation.



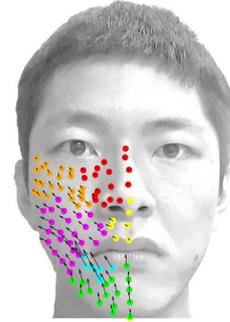
(b) With compensation.

Fig.3 Superposed skin flow vector fields of 16 AUs with/without the face movement compensation.

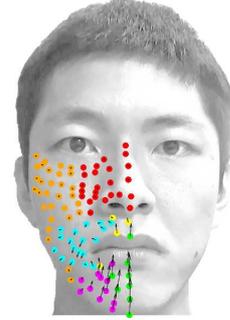
4 AU ごとの皮膚の流れ場の分析

ロボットのアクチュエータ数が少なくかつ再現度が高くなるよう顔ロボットの機構を配置するために、動きの似た部位同士を 1 つのクラスタとしてまとめ、そのクラスタ単位で皮膚を駆動することが望ましい。AU ごとに流れ分布のクラスタリングを k-means 法で実施した。クラスタリングに用いた変数はマーカの移動ベクトルの始点位置・x 変位・y 変位である。それぞれの値は式 (2) を用いて正規化した。式 (2) において、各変数の値を x_i 、各変数の平均を \bar{x}_i 、各変数の標準偏差を σ_i とする。なお本研究では標準偏差は全て不偏標準偏差を用いた。クラスタリング結果の例として、図 4a に唇両端を引っ張り上げる AU12、図 4b に唇両端を下げる AU15 のクラスタリング結果を示す。皮膚の流れ場のクラスタリングの結果、AU ごとにクラスタリング結果が異なり、多様な表情を 1 つの顔ロボットで同時に実現することの困難さが確認された。この原因として、皮膚の各点において AU ごとに流れの向きと大きさが異なることが挙げられる。

$$z_i = \frac{x_i - \bar{x}_i}{\sigma_i} \quad (2)$$



(a) AU12:Lip Corner Puller



(b) AU15:Lip Corner Depressor

Fig.4 Example of an estimated clusters of the measured skin flow field with Lip Corner Puller action, which is defined as AU12 and Lip Corner Depressor action, which is defined as AU15 in FACS. Markers represented with the same color were estimated as members of the same cluster.

5 全 AU の重ね合わせの分析

ロボット開発において、図 3b のような様々な皮膚の流れ全てに対応した機構を配置することは、スペースの制約や自由度数が膨大になることにより、現実的ではない。そこで、これら全ての流れの特徴を上手く捉えた機構配置を行うため、各マーカの移動量に対し、主成分分析を行い、皮膚の流れに「共通する流れ成分」を抽出する。この流れ成分が元の多様な流れの情報を保持してい

る場合、その流れ成分のみをロボットに実装した場合でも再現度は損なわれないと考えられる。

図 5 に主成分分析結果を示す。この図では、ベクトルの向きが主成分方向を表し、各マーカに関して、口唇部の動きに伴って最もよく現れる皮膚の移動方向と対応する。そして、ベクトルの長さが主成分得点の分散に対応する。なおこの図では、主成分得点の分散の値を 2 で割っている。

主成分分析の結果、第 1 主成分の寄与率が 86% となり、皮膚の各点の多様な動きは第 1 主成分方向の 1 方向にほぼ集約されることが示された。

また、図 6 のように、主成分分析結果を筋肉と重ね合わせると、第 1 主成分方向と筋繊維方向が概ね一致していることが分かる。故に主成分分析結果はこの観点からも、合理的であり妥当であると言える。

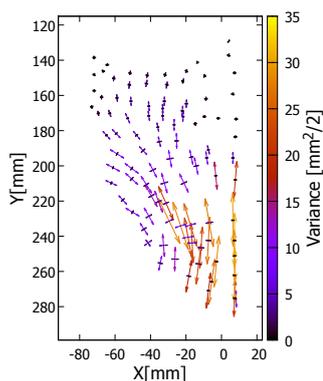


Fig.5 Vector field of the each marker's representative vector whose direction and magnitude are the direction of the first principal components of the superposed skin flows and the score variance of the first principal component scores of them. The score variance is divided by 2.

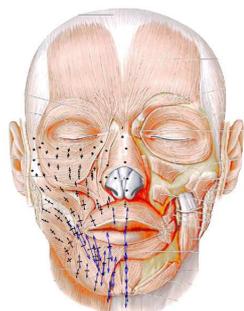


Fig.6 Position correspondence between the vector field shown in Fig. 5 and facial muscles. The facial anatomical chart is reprinted from a reference [5].

6 主成分分析結果のクラスタリング

主成分分析を行うことで、各マーカの移動を第 1 主成分方向の流れで表したが、皮膚上の全てのマーカそれぞれにアクチュエータを割り当てることは現実的ではない。そこで、類似した流れが起こる部位を 1 つの駆動機構で実現するため、共通する流れ成分である第 1 主成分に基づいてクラスタリングを行った。クラスタリングには k-means 法を用いた。クラスタリングに用いた変数は各マーカの位置ベクトル・主成分方向の角度・主成分得点の分散をそれぞれ式 (2) で正規化したものである。

このクラスタリングではクラスタ数と各 AU の再現度の関係を把握するため、クラスタ数を 2 から筋肉の数である 13 まで変化させた。k-means 法のクラスタ重心の初期位置は筋肉の配置を基に決定した。図 7 に、用意するクラスタ数を増やした際に追加したクラスタ重心の初期配置を示す。このクラスタリングにより、多様な顔の動きに対して類似した流れが起こる部位をクラスタとして分類する。

クラスタ数が 7 の場合を例として、第 1 主成分のクラスタリング結果を図 8 に示す。この図より顎部の鉛直方向の流れが緑、頬部の斜め方向の流れが紫といったように類似した流れがクラスタとしてまとめられていることが分かる。

ここでクラスタ内の主成分方向のばらつきをクラスタ数ごとに比較する。クラスタ数と第 1 主成分方向角度の標準偏差のクラスタ平均の関係を図 9 に示す。この図において横軸のクラスタ数が顔ロボットの口唇部の自由度数に該当する。そして、縦軸の第 1 主成分方向角度の標準偏差のクラスタ平均がロボットの表情と人の表情の誤差に対応し、この値が小さいほどそのクラスタ数での顔の動きの再現度が高いと言える。このグラフからロボットの自由度数とその自由度数での顔の動きの再現度が推定できる。クラスタ数 3 まで急激に減少し、それ以上のクラスタ数になると、減少傾向は緩やかになっており、最低でも口唇部の自由度は 3 必要であると考えられる。

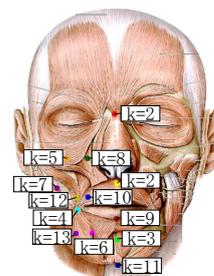


Fig.7 Initial position setting of the cluster centers for k-means clustering of the vector field shown in Fig. 5. The clustering was carried out with several numbers of clusters and the numbers in this figure show which positions are adopted as initial positions in each setting. The facial anatomical chart is reprinted from a reference [5].

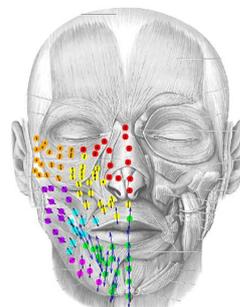


Fig.8 Example of an estimated clusters of the vector field shown in Fig. 5 with 7 clusters. Markers represented with the same color were estimated as members of the same cluster. The facial anatomical chart is reprinted from a reference [5].

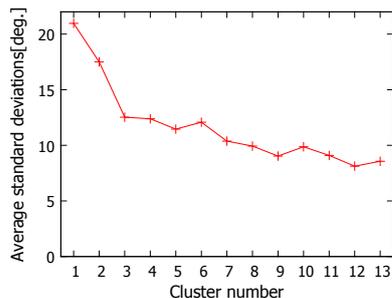


Fig.9 Averaged standard deviations of the first principal component directions in each cluster, which was estimated with different numbers of clusters.

7 第1主成分方向と各顔の動きのズレ

第1主成分方向のみの流れをアクチュエータで実装した場合、顔の本来の動きから皮膚の流れがどの程度ずれるのかを評価する。図10に口角のマーカーを例に、第1主成分方向と各AU表出時のマーカーの移動ベクトルの終点の関係を表している。原点は無表情時のマーカー位置である。この図において各マーカーの移動ベクトルの終点から第1主成分方向に下ろした垂線の長さ、つまり第2主成分方向の値が第1主成分方向のみに皮膚を動かした際の本来の皮膚の流れからのズレを表す。そこで、次に各顔の動きに関して第2主成分得点のクラスタ内の平均を求める。図11にクラスタ数7の場合を例に、AU12(唇両端を引っ張り上げる)のクラスタ内の第2主成分得点の値を示す。第2主成分得点の値を16種類のAUについて求めた結果、16種中、14種(AU10-18, AU22-25, AU28)は第2主成分得点のクラスタ平均が1-3mmであり、第1主成分方向に皮膚を動かすようにアクチュエータを実装することで、効率的に多くの表情を再現することが可能であることが分かった。

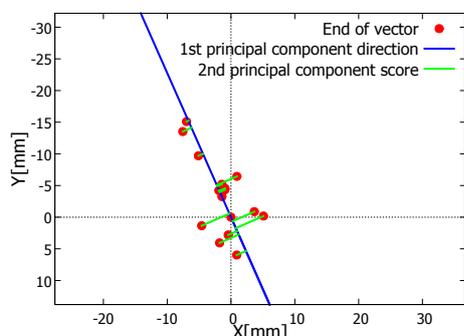


Fig.10 Example of the relationship between the measured flow vectors and the direction of their first principal component. Red dots represent the end point of each flow vector and foot lengths of a green perpendicular line from each red dot to the straight line indicate how much reproduction errors occur when this facial part is moved only in the direction of the first principal component.

8 結言

本研究では、多様な顔の動きを再現可能な顔ロボットの機構設計を推定するための分析手法として、複数の表情に伴う各皮膚部位の動きを重畳させて共通する流れ成分である第一主成分を抽出した上で、位置と動きの近いクラスタに分類する手法を提案した。そして、分析結果の評価を行ったところ、第1主成分方向のみの

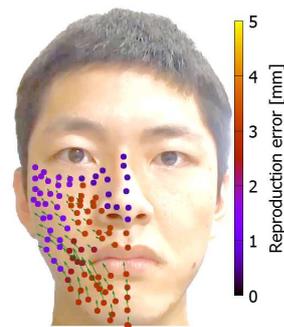


Fig.11 Deviation of the first principal component directions and the skin flows of AU12. Marker colors indicate each marker's reproduction error, which is an average of the 2nd principal component scores in each cluster.

方向のアクチュエータを実装しても、多様な顔の表情を再現可能であることが示唆された。

本研究では、2次元の画像データから求めたマーカーの初期位置と移動後の位置のデータを用いて顔の観察・分析を行った。しかし、本来顔は3次元に変形し、変形のタイミングも顔の部位ごとに異なる。よって、今後の展開として、本分析手法を顔の変形の3次元の時系列データに拡張することが挙げられる。また本研究で用いたデータのサンプル数は1人であったが、顔の変形は人によって異なる可能性があり、それらも考慮して分析を行うことが望まれる。また、本分析から得られた知見を基に顔口ポットを作成することも課題として挙げられる。

9 謝辞

本研究において、クラスタリング手法や分析手法に関し、多くの助言、助力を頂いた横山裕樹特任研究員、博士後期課程2年河合祐司氏に深く感謝の意を表します。本研究は、科学研究費補助金 特別推進研究(24000012)及び科学技術人材育成費補助金 テニユアトラック普及・定着事業の補助を受けた。

文献

- [1] 権泰錫. 顔表情ロボットにおける駆動ユニットの分類に関する研究:- アニマトロニクスのためのムービングユニットの提案-. PhD thesis, 九州大学, 2008.
- [2] H Shimauchi, Y Takeuchi, T Tenkumo, and K Sasaki. Development of clinical training program for sophisticated dental education. In *Hiroshima Conference on Education and Science in Dentistry*, p. 45, 2013.
- [3] 坂本大介, 神田崇行, 小野哲雄, 石黒浩, 萩田紀博. 遠隔存在感メディアとしてのアンドロイド・ロボットの可能性. *情報処理学会論文誌*, Vol. 48, No. 12, pp. 3729-3738, 2007.
- [4] Hisashi Ishihara, Yuichiro Yoshikawa, and Minoru Asada. Realistic child robot "affetto" for understanding the caregiver-child attachment relationship that guides the child development. In *Development and Learning (ICDL), 2011 IEEE International Conference on*, 第2巻, pp. 1-5. IEEE, 2011.
- [5] Baker and Eric W. プロメテウス解剖学アトラス 口腔・頭頸部. 医学書院, 2012. 坂井 建雄, 天野 修 監訳.
- [6] P Ekman, WV Friesen, and JC Hager. *Facs manual. A Human Face*, 2002.