

表情表出における3次元特徴点が生み出す動的変位の分析

IWASA, Kaori / 岩佐, 香織

(出版者 / Publisher)

法政大学大学院理工学・工学研究科

(雑誌名 / Journal or Publication Title)

法政大学大学院紀要. 理工学・工学研究科編

(巻 / Volume)

56

(開始ページ / Start Page)

1

(終了ページ / End Page)

6

(発行年 / Year)

2015-03-24

(URL)

<https://doi.org/10.15002/00011539>

表情表出における 3 次元特徴点が生み出す動的変位の分析

ANALYSIS OF 3D DYNAMIC DISPLACEMENT OF FEATURE POINTS IN FACIAL EXPRESSIONS GENERATION

岩佐香織

Kaori IWASA

指導教員 赤松茂

法政大学大学院理工学研究科システム工学専攻修士課程

My research goal is to achieve accurate and natural automatic facial expression recognition by examining features that reveal the dynamic deformation of three-dimensional (3D) faces. We conducted classification experiments to test whether the geometric features obtained by a motion capture system could be used to discriminate among different categories of facial expressions. Classification experiments use the simple minimum distance classification method based on Euclidean distance, and Leave-one-out method is used to the method of evaluation. We discovered that geometric features representing 3D coordinates of sparse feature points defined on a subject's face could be measured by a motion capture system for each time frame. The measurements of these features are mostly effective at discriminating different facial expressions, especially at time frames that reveal maximum expressiveness. However, the generation of facial expressions includes a series of temporal transitions of feature points from an initial emotionless face to an expression's peak. Therefore, this study examines the effect of the temporal integration of gradual 3D deformation into features used for facial expression discrimination.

Key Words : facial expression recognition, 3D shape deformation, motion capture

1. はじめに

本研究は、表情を表出している時の顔特徴点の位置情報が、表情の違いをどれだけ表現しているかを識別実験によって検証している。まずモーションキャプチャによって得られた「疎な特徴点の位置情報を含んだ測定データ」と、それを主成分分析によって圧縮して求めた「表情の違いを効率的に表す低次元の特徴ベクトル」の二つに対して最近傍法を用いて識別実験を行い、Leave-One-Out法によって交差検証を行った。その結果は「ピーク表情時における識別」、「表情サンプルを時間軸に沿って10分割した区間に属するフレームにおける識別率の平均」といった形で集計した。しかし、これはあくまで動きの中の一瞬の位置情報を用いて識別である。そのため追加実験として動作途中の特徴ベクトルをつなぎ合わせることで連続の動作を表現した「表情表出特徴ベクトル」を用いての表情識別を行った。

本報告の構成は次のとおりである。まず2.で研究背景を述べる。次に3.で実験環境を述べ、続く4.で識別実験の手順を示す。5.では識別実験の結果を示す。さらに6.で結果に対する議論を述べる。最後に7.で本報告の成果をまとめる。

2. 背景

機械による人間の感情や態度の認識が発展すれば、ヒューマンインタフェースの高度化への貢献が期待される。感情を認識できることで、その感情に沿った反応を提示できるからである。例として相手の反応を読み取る擬人化エージェント^[1]などが挙げられる。

感情を読み取る上で非言語コミュニケーションは重要である。それは人間が相手の感情や態度を認識する際、言語コミュニケーションと非言語コミュニケーションに矛盾が生じた際は後者による影響が大きい^[2]事などからも伺える。なかでも顔表情は特に重要であり、顔からは何を言っているかという発話情報や、感情・意図・関心など心理状態の情報なども見て取ることができるのは経験的に実感できるだろう。

本研究は顔の疎な特徴点の動的変異情報について着目した。その理由は二つある。一つは、表情は顔という3次元物体の変形をとまなう複雑な動きによって表出されるものであるため、その動的特性の分析にモーションキャプチャの利用が期待できる事^[3]。もう一方は、顔面上の僅かな個数の点光刺激でも表情判断が成立することが報告されている事^[4]である。

3. 実験環境

本研究では、表情表出時の顔特徴点位置データの計測を行なうために光学式モーションキャプチャシステム OptiTrack（米国 Natural Point 社製）を使用した。本システムは2台以上の赤外線照射機能を搭載した赤外線カメラで赤外線を反射するマーカを撮影し、三角測量の原理によってマーカの三次元座標を取得することができる。

実際の計測環境では図 1 のように被験者を囲うように7台のカメラを設置し、被験者の顔表面に貼付したフェイシャルマーカを常に3台以上のカメラで捕らえられるようにした。

また、顔面の特徴点として被験者の顔表面にはフェイシャルマーカを37個貼付し、さらに頭部の揺らぎ補正を行なうために4つのマーカが固定されたヘッドバンドを装着した（図 2）。



図 1 カメラの設置環境

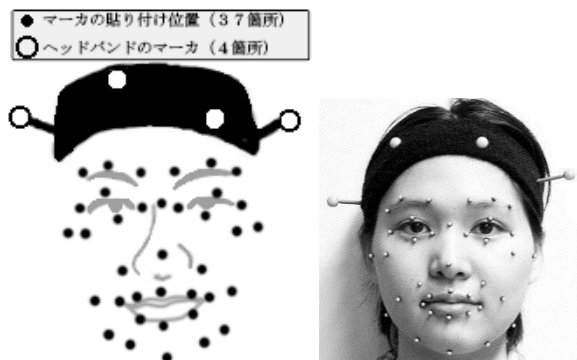


図 2 マーカの貼り付け位置とヘッドバンド

4. 識別実験方法

(1) 計測

大学生および大学院生の男女計 30 名に対して、表 1 に示す9表情を真顔から表出してもらい、それぞれ3回ずつ計測した。合計 810 データを得た。

表 1 計測した表情の種類

1	[a]の発話	6	開口笑顔
2	[i]の発話	7	閉口笑顔
3	[u]の発話	8	唇を尖らせる
4	[e]の発話	9	頬を膨らませる
5	[o]の発話		

(2) データ処理

本研究で必要なデータは開始が真顔、終了がピーク表情時である。その為、計測データから表情表出の開始フレームと表情表出のピークフレームを目視によって決定し、前後の不要なフレームを削除する事で真顔から表情のピークまでのデータを抽出した。（図 3）

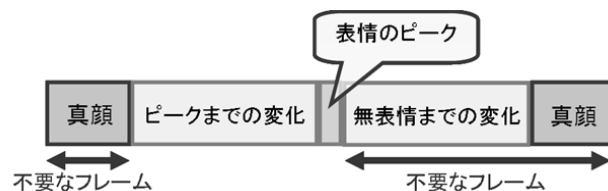


図 3 モーションキャプチャデータの不要フレーム除去

また測定したデータは頭部の姿勢変化に影響され、向いている方向や頭部の位置が同じではない為、正規化を行った。

(3) 特徴ベクトルの定義

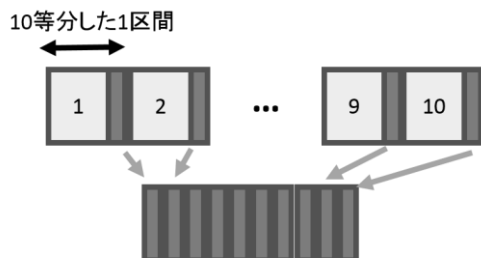
人物 k ($k = 1, 2, \dots, 20$) が表情 c ($c = 1, 2, \dots, 8$) について m ($m = 1, 2, 3$) 番目の試行で行った表情表出動作の計測データに対し、 t ($t = 1, 2, \dots, T$) 番目のフレームの位置座標ベクトルを $X_{k,c,m}^{(t)}$ と表す。ただし、 $t = 1$ は真顔、 $t = T$ は表情表出のピークを表す。なお、表情の表出時間は試行ごとで一定ではない為、 T は一定の値をとらない。真顔から表情表出過程のフレーム t における各マーカ点の三次元変位量を表す特徴ベクトル $Z_{k,c,m}^{(t)}$ は位置座標ベクトルの差分により次式のように求められる。

$$Z_{k,c,m}^{(t)} = X_{k,c,m}^{(t)} - X_{k,c,m}^{(1)} \quad (1)$$

表情表出時間は試行毎に一定ではないため、各表情サンプルをフレーム数によって10等分し、その末尾に当たるフレームを表情表出途中の代表フレームとする。代表フレームを任意につなげることで、任意の表情表出パターンを得ることができる（図 4）。これを、表情表出特徴ベクトルと呼ぶことにする。加えてこの代表フレームを太字の数字で表すこととする。今回、生成したパターンは区間番号でいうところの10=7（最終フレームのみ）10~9, 10~8, ..., 10~2, 10~1（すべての代表フレーム）の

10 種類である．代表フレーム 1 における特徴ベクトルを $Z_{k,c,m}^{(1)}$ とすると，例えば表情表出特徴ベクトル 10～1 は次のようにあらわされる．なお，T は転置を意味する．

$$\left[Z_{k,c,m}^{(1)T} \quad Z_{k,c,m}^{(2)T} \quad \dots \quad Z_{k,c,m}^{(9)T} \quad Z_{k,c,m}^{(10)T} \right]^T \quad (2)$$



区間の末尾の特徴ベクトルをつなぎ合わせる

図 4 表情表出特徴ベクトルのイメージ (1～10)

(4) 主成分分析による次元圧縮

この特徴ベクトル $Z_{k,c,m}^{(i)}$ は 37 個のマーカ点の x, y, z 座標値からなる 111 次元のベクトルである．さらに，表情表出特徴ベクトル 10～1 (式 2) に至っては 1110 次元のベクトルとなる．表情表出特徴ベクトル 10 と表情表出特徴ベクトル 10～1 では特徴ベクトルの長さが異なり，次元数に差があるため，公平に比較するためには主成分分析による次元圧縮が必要となる．

各表情表出特徴ベクトルにおいて，それぞれのパターンに対して主成分分析を適用し，得られた固有ベクトルを用いて各表情表出特徴ベクトルの次元圧縮を行う．

(5) 表情識別及び性能評価

各表情表出特徴ベクトルについてユークリッド距離による最近傍法で表情識別し，その分類性能を Leave-One-Out 法によって交差検証を行う．識別においては表 1 に示される 9 表情に無表情を加えた 10 表情への強制選択を取った．人による表情認知においては未知人物の表情識別に相当する性能を評価するため，Leave-One-Out 法によ

る交差検証では人物単位にテスト用データを選択し，当該人物の全ての表情顔が学習データに含まれることのないようにした．

5. 識別実験の結果

(1) ピーク表情における識別実験

主成分分析を行って得られた累積寄与率を図 5 に表す．次に，次元ごとの平均識別率を図 6 に示す．最後に最も識別率が良かった次元数における混同行列を

表 2 に，次元減少前における混同行列を表 3 に示す．なお，行に示された表情は実際の表情で，列に示された表情は識別された表情である．グレーに塗られたマスの値は正答率を，残りの表情に誤答した率を示している．

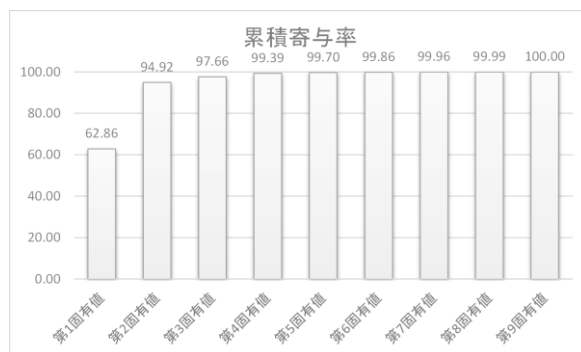


図 5 ピーク表情に主成分分析を行って得られた累積寄与率

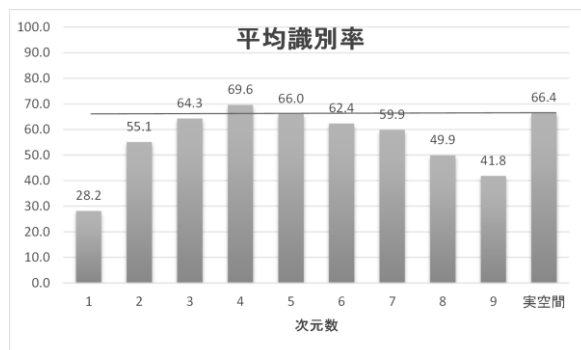


図 6 ピーク表情における次元ごとの平均識別率

表 2 ピーク表情における第 4 主成分まで次元圧縮した場合の混同行列

表情	[a] の 発話	[i] の 発話	[u] の 発話	[e] の 発話	[o] の 発話	開 口 笑顔	閉 口 笑顔	唇を尖 らせる	頬を膨ら ませる	真 顔
[a]の発話	56.7	2.2	3.3	28.9	6.7	0.0	2.2	0.0	0.0	0.0
[i]の発話	6.7	52.2	4.4	13.3	1.1	1.1	18.9	0.0	0.0	0.0
[u]の発話	7.8	4.4	63.3	1.1	8.9	0.0	0.0	13.3	0.0	1.1
[e]の発話	34.4	10.0	4.4	43.3	7.8	0.0	0.0	0.0	0.0	0.0
[o]の発話	12.2	3.3	11.1	5.6	67.8	0.0	0.0	0.0	0.0	0.0
閉口笑顔	3.3	1.1	0.0	0.0	0.0	93.3	2.2	0.0	0.0	0.0
開口笑顔	3.3	20.0	0.0	1.1	0.0	2.2	73.3	0.0	0.0	0.0
唇を尖らせる	3.3	0.0	14.4	0.0	1.1	0.0	0.0	81.1	0.0	0.0
頬を膨らませる	3.3	0.0	1.1	0.0	0.0	0.0	0.0	0.0	95.6	0.0

表 3 ピーク表情における次元圧縮前の混同行列

表情	[a] の 発話	[i] の 発話	[u] の 発話	[e] の 発話	[o] の 発話	開 口 笑顔	閉 口 笑顔	唇を尖 らせる	頬を膨ら ませる	真 顔
[a]の発話	58.9	5.6	0.0	34.4	1.1	0.0	0.0	0.0	0.0	0.0
[i]の発話	8.9	46.7	3.3	20.0	2.2	1.1	14.4	0.0	0.0	0.0
[u]の発話	2.2	2.2	61.1	4.4	8.9	0.0	0.0	20.0	1.1	0.0
[e]の発話	45.6	12.2	1.1	38.9	2.2	0.0	0.0	0.0	0.0	0.0
[o]の発話	8.9	2.2	6.7	13.3	68.9	0.0	0.0	0.0	0.0	0.0
開口笑顔	0.0	1.1	0.0	3.3	0.0	92.2	3.3	0.0	0.0	0.0
開口笑顔	0.0	21.1	0.0	7.8	0.0	4.4	66.7	0.0	0.0	0.0
唇を尖らせる	0.0	0.0	24.4	3.3	0.0	0.0	0.0	71.1	1.1	0.0
頬を膨らませる	0.0	0.0	1.1	3.3	0.0	0.0	0.0	2.2	93.3	0.0

図 5 からはデータセット 2014 において 99%の情報を得るためには第 4 主成分までを利用すればよいことがわかる。図 6 からは第 4 次元まで圧縮した場合が 69.6%の識別率で、次元圧縮された中でもっとも高く、次元圧縮無しの場合の 66.4%も上回ったことが分かる。第 4 次元の前後では識別率のグラフが山形になるように識別率が減少している。

表 2, 表 3 からは, 各表情の正答率が「[e]の発話」の 43.3%から「頬を膨らませる」の 95.6%であることが分かる。加えて, 太字で示したマスから誤答しやすい表情対があることが判明した。その表情対とは, 「[a]の発話」対「[e]の発話」, 「[i]の発話」対「[e]の発話」, 「[i]の発話」対「開口笑顔」, 「[u]の発話」対「[o]の発話」, 「[u]の発話」対「唇を尖らせる」, 「[e]の発話」対「[o]の発話」である。

(2) サンプリングした表情表出過程の結合ベクトルにおける識別実験

主成分分析を行って得られた累積寄与率を表 4 に表す。すべてのパターンにおける表情表出特徴ベクトルの識別率を、次元数ごとに平均したグラフを図 7 に示す。なお、グレーに塗られたマスは累積寄与率 99%以上を満たす最初の累積寄与率である。つぎに次元数ごとに平均した表情表出特徴ベクトルの識別率においてもっとも成績が良かった第 4 主成分までを用いた場合にて、パターン毎の識別率の変化を示したグラフを図 8 に示す。最後に最も識別率が良かった次元数における混同行列を表 5 に示す。なお、行に示された表情は実際の表情で、列に示された表情は識別された表情である。グレーに塗られたマスの値は正答率を、残りのある表情に誤答した率を示している。誤答した率が高いと思われたマスは太字で示す。

表 4 データセット 2014 の全表情表出特徴ベクトルにおけるパターンごとの累積寄与率

[illegible]

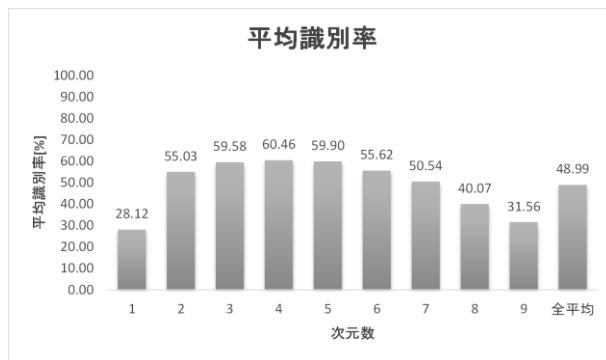


図 7 データセット 2014 の全表情表出特徴ベクトルにおける次元ごとの平均識別

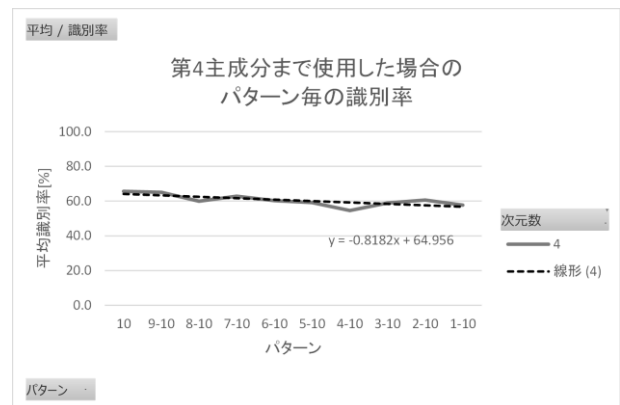


図 8 全表情表出特徴ベクトルにおける第 4 主成分まで用いた場合のパターンごとの平均識別率

表 5 データセット 2014 の全表情表出特徴ベクトルにおけるパターン 1~10 の第 4 主成分まで用いた場合の混同行列

表情	[a] の 発話	[i] の 発話	[u] の 発話	[e] の 発話	[o] の 発話	閉口 笑顔	開口 笑顔	唇を尖 らせる	頬を膨ら ませる	真顔
[a]の発話	45.6	1.1	2.2	27.8	23.3	0	0	0	0	0
[i]の発話	2.2	36.7	12.2	15.6	7.8	0	18.9	0	1.1	0
[u]の発話	11.1	11.1	45.6	5.6	15.6	0	0	10.0	1.1	0
[e]の発話	25.6	11.1	4.4	45.6	11.1	0	1.1	0	0	0
[o]の発話	24.4	4.4	10.0	14.4	46.7	0	0	0	0	0
閉口笑顔	0	0	0	1.1	3.3	90.0	5.6	0	0	0
開口笑顔	0	25.6	0	2.2	3.3	3.3	64.4	0	0	0
唇を尖らせる	0	0	10	0	3.3	0	0	82.2	3.3	1.1
頬を膨らませる	1.1	6.7	11.1	1.1	3.3	2.2	0	3.3	67.8	1.1

表 4 によると多くのパターンにおいて第 4 主成分まで、もしくは第 3 主成分まで用いた場合が良いという結果になった。図 7 によると第 4 主成分まで用いた場合の識別率が最も高いことが分かる。図 8 からは各パターンにおける識別率の変化が、より多くのベクトルを用いた場合になるほどやや右肩下がりに識別率が減少していることが分かる。表 5 からは、パターン 1~10 の第 4 主成分まで用いた場合、各表情の正答率が「[i]の発話」の 36.7%から「閉口笑顔」の 90.0%であることが分かる。加えて、太字で示したマスから誤答しやすい表情対があることが判明した。その表情対とは、「[a]の発話」対「[e]の発話」、「[a]の発話」対「[o]の発話」、「[i]の発話」対「[u]の発話」、「[i]の発話」対「[e]の発話」、「[i]の発話」対「開口笑顔」、「[u]の発話」対「[o]の発話」、「[u]の発話」対「唇を尖らせる」、「[e]の発話」対「[o]の発話」である。

6. 考察

(1) ピーク表情における識別実験

今回累積寄与率の閾値を 99.0%と仮定して、識別実験に使用するために妥当な次元数の目処をつけた。その結果、第 4 主成分が妥当であるという仮説は正しかった。

しかし、参考に 2010, 2011 それぞれに撮影したデータセットで同様の実験を行ったときは閾値 99.0%では不十分であり、より主成分を用いたデータのほうが識別率は高かった。2010・2011 に撮影したデータセットと今回のデータセットではデータ数が 480 と 810 で倍近く異なり、今回のデータセットのほうがより多くのデータ数を保有しているため精度が異なった可能性が考えられる。

続いて次元圧縮により次元圧縮前とどのような変化が起きたかを議論する。まず次元圧縮により平均識別率が 66.4%から 69.6%に向上した。

表 2, 表 3 の混同行列を見ると次元減少により赤字で示した混同された表情への誤識別が概ね減少している。これは、次元圧縮によって表情特徴を良く表す表情ベクトルが得られたことを意味する。おそらく特徴の中でも個人を現す特徴や性差を表す特徴など、表情を識別する上ではノイズになる特徴が次元減少により削減されたと考えられる。

誤答しやすい表情対については概ね予想通りの結果となった。「閉口笑顔」や「頬を膨らませる」といったほかにはない動きをする表情は識別率が高くなり、お互いが似通いやすい発話表情は識別率がやや低くなった。特に「[e]の発話」についてはほかの発話表情に似ている傾向

が強い。

(2) サンプリングした表情表出過程の結合ベクトルにおける識別実験

今回累積寄与率の閾値を 99.0%と仮定して、識別実験に使用するために妥当な次元数の目処をつけた。その結果、第4主成分が妥当であるという仮説は正しかった。

今回注目すべきパターン毎の識別率は予想とは異なり、ベクトル数を増やしたもののほうが悪い結果となった。これは本研究で用いた識別手法では、表情表出過程の情報がかえってノイズとなってしまう、識別率の悪化を招いた可能性がある。逆に言ってしまうと本研究で用いた識別手法ではピーク表情のみであれば十分識別できるということになる。

混同行列を見ると、ピーク表情時のものと比べて全体的に正答率が悪化しており、特に発話表情の混同が増えている。本実験で使用したデータセットは被験者に自然な表情をとってもらったため、発話表情については対象の音を発してもらう以外の制限はつけなかった。そのためはっきりと口を動かしていないデータが多く、どの発話表情も似通ってしまった可能性がある。今回の実験ではベクトルを長く取ることで、それが顕著に出てしまった可能性が考えられる。

7. むすび

本研究ではモーションキャプチャによって得られた疎

な特徴点の動的変異情報から、時系列変化による微細な特徴を含んだ表情表出特徴ベクトルを求め、その識別性能を評価した。その結果、本手法、本データではピーク表情のみを用いて4次元程度に次元減少させた場合に識別性能が高かった。ピーク表情ベクトルと比較して表情表出特徴ベクトルの優位性は見られなかった。つまり最近傍法で識別する場合、ピーク表情のみで十分識別可能であることが判明した。最近傍法では特徴空間の次元数が多くなってしまうと識別性能に限界があることが判明した。

謝辞：本研究の一部は、科学研究費補助金（基盤研究(B)24300087）の助成を得た。ここに記して謝意を示す。

参考文献

- 1) 中村和晃, 角所考, 正司哲郎, 美濃導彦, 澤木美奈子, 南泰浩, 前田英作: 擬人化エージェントとの音声対話時におけるユーザの非言語動作からの難/易及び興味/退屈の推定, 電気情報通信学会論文誌.A, Vol.J95-A, No.1, pp.85-96, Jan.2012.
- 2) Mehrabian, A. : Silent messages, Wadsworth, Belmont, 1971.
- 3) 柳澤博昭, 祖川慎治, 前島謙宣: モーションキャプチャによる顔表情の定量表現, 電子情報通信学会技術研究報告.HCS, Vol.104, No.744, pp.7-12, Mar.2005.
- 4) 松寄直幸, 佐藤隆夫: 表情認知とバイオロジカルモーション, 電子情報通信学会技術研究報告. MVE, Vol.97, No.85, pp.131-136, Jun.1997.