

顔特徴点の位置変化による特徴量を用いた表情強度の推定

Estimation of facial expression intensity by using facial features based on the positions of facial feature points

阪上 翔太†
Shota Sakaue

野宮 浩揮†
Hiroki Nomiya

宝珍 輝尚†
Teruhisa Hochin

1. はじめに

近年、マルチメディアデータ記録機器の高性能化やストレージの大容量化、価格の低下に伴い、誰でも簡単に大量のマルチメディアデータを作成できるようになった。このことから、日常生活上の様々なデータを取得・蓄積し、のちに有効活用しようというライフログが提案されている[1]。特にその中でも映像データは、手軽に作成することができ、様々な情報を内包しているため、ライフログデータとして有用であるといえる。

しかし、ライフログ映像はデータの取得が容易である反面、あまりにも多量の映像が蓄積されるため、その中から必要なデータを即座に取り出すことが難しく、有効活用できていないというのが現状である。このような背景から、ライフログ映像の中から有用な映像を簡単に検索したいという要求が高まってきている。

筆者らはこれまでに、映像中の人物の表情を用いた、ライフログ映像における印象的なシーンの検索に関する研究を行ってきた[2]。これらの研究の手法では、映像内の人物に特定の表情が表出しているシーンの検索は可能であるが、表情の程度を推定することは不可能である。例えば、「笑顔」が表出しているシーンは検索可能であるが、「微笑み」や「高笑い」というような感情の程度が異なる表情が考慮されていない。

そこで、森國らは、ライフログ映像中には様々な強さの表情が表出しているシーンが含まれると考え、このような表情の程度を数値化し、シーン検索に活用することを目的として、顔の特徴点を用いた「表情強度」という尺度を提案している[3]。表情強度は、顔特徴点と呼ばれる顔面上のいくつかの点を用いて計算される特徴量から求められる。

森國らによる先行研究では表情強度の算出に使用されている特徴量の評価を行っておらず、また表情強度の推定精度に課題が残っていた。さらに、表情強度を求める際に各特徴量の単位の違いを考慮せず、単純に各特徴量の値を足し合わせて表情強度を算出していた。

そこで本論文では、表情強度を算出するのに使用されている特徴量の評価を行い、表情強度の推定を行う。また、先行研究では「笑顔」と「驚き」のみに焦点を当て実験を行っていたが、本論文では基本6表情(喜び・驚き・嫌悪・怒り・悲しみ・恐怖)に含まれる表情を取り扱う。各特徴量の単位の違いについては、統計学分野で用いられる標準化を行うことで単位の異なる数値データの大きさの違いをなくす。

無表情時・表情表出時の2種の表情データを用いて、先行研究と提案手法のそれぞれの表情強度の平均値に基づいて、表情強度の推定精度を比較する。その結果、全ての表情で提案手法の推定精度が優れていることを示す。

本論文の構成は次の通りである。第2章では森國らによる先行研究について紹介する。第3章では提案手法について説明する。第4章では特徴量の評価と表情強度の推定実験について説明する。第5章では実験の考察を行う。第6章でまとめと今後の課題について述べる。

2. 先行研究

森國らは、表情の程度を表すことを目的として表情強度を提案している[3]。本節では、まず表情強度の導出に使用する顔特徴量、ならびに顔特徴量を求める際に使用する顔特徴点について述べる。その後、顔特徴量を用いた表情強度の導出法について述べる。

2.1 顔特徴点

表情強度は、以下に示す54点の顔特徴点の位置関係に基づいて求められる。図1に54点で構成される顔特徴点の詳細を示す。

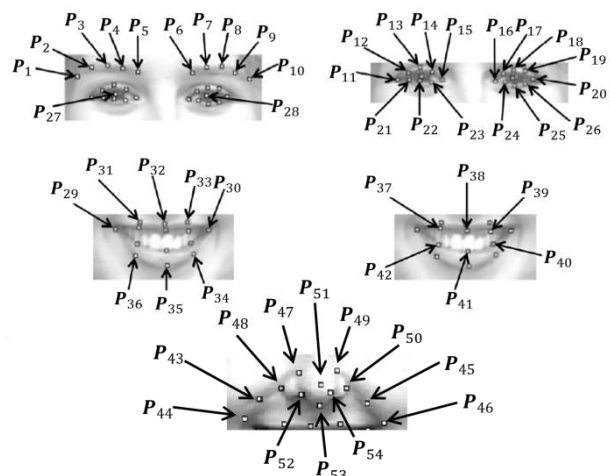


図1 顔特徴点

- 左右の眉:10点(P_1, \dots, P_{10})
- 左右の目:18点(P_{11}, \dots, P_{28})
- 口:14点(P_{29}, \dots, P_{42})
- ほうれい線:4点(P_{43}, \dots, P_{46})
- 鼻:8点(P_{47}, \dots, P_{54})

これらの顔特徴点は FaceSDK4.0[4]を用いて取得する。

2.2 顔特徴量

2.1で示した顔特徴点を組み合わせて、以下に示す11種類の特徴量が定められており、表情強度の推定に用いられている。各特徴量は、表情表出時に人間から見て大きく位置関係が変化すると考えられる特徴点の組み合わせが選択されている。

- 左眉と右眉の傾き f_1

左眉(右眉)の傾きは、5つの特徴点 $P_6, P_7, P_8, P_9, P_{10}, (P_1, P_2, P_3, P_4, P_5)$ から最小二乗法(線形)により算出し、左右の眉がほぼ左右対称と考え、式(1)により求められる。ただし、 a_l は左眉の傾き、 a_r は右眉の傾きである。

$$f_1 = \frac{a_l - a_r}{2} \quad (1)$$

- 眉と目の間の距離 f_2

両眉の10点と上瞼の10点を両方共左から順に1対1対応で結んだ10本の線分の長さの平均値。ただし特徴量が顔画像の大きさに影響を受けないようにするために、左目と右目の距離 L で割って正規化を行う。 f_2 は式(2)により求められる。

$$f_2 = \frac{\sum_{i=1}^{10} \|P_i - P_{i+10}\|}{10 \cdot L} \quad (2)$$

- 眉間の面積 f_3

左眉・右眉・右目・左目の特徴点のうち、それぞれ最も顔の中央に近い点を結んでできる四角形の面積。ただし、 L^2 で割って正規化を行う。 f_3 は式(3)により求められる。

$$f_3 = \frac{S(P_5, P_6, P_{16}, P_{15})}{L^2} \quad (3)$$

ここで、 $S(P_1, \dots, P_n)$ は P_1, \dots, P_n をつないで構成される n 角形の面積である。

- 目の面積 f_4

左目(右目)の外周上の8つの特徴点を結んでできる八角形の面積をそれぞれ算出し、この平均値を目の面積とする。ただし、 L^2 で割って正規化を行う。 f_4 は式(4)により求められる。

$$f_4 = \frac{1}{2 \cdot L^2} \{ S(P_{11}, P_{12}, P_{13}, P_{14}, P_{15}, P_{23}, P_{22}, P_{21}) \} \quad (4)$$

- 目の縦横の比 f_5

左目(右目)の一番上にある点と一番下にある点を結んだ線分の長さを左目(右目)の縦の長さ $l_{vl}(l_{vr})$ 、最も左と右にある2点を結んだ線分の長さを左目(右目)の横の長さ $l_{hl}(l_{hr})$ として縦横の比を算出し、長さ l_{hl}, l_{hr} が0である場合を考慮して、縦横の比の逆正接を取り、それらを平均して式(5)のように求められる。

$$f_5 = \frac{1}{2} \left(\tan^{-1} \frac{l_{vl}}{l_{hl}} + \tan^{-1} \frac{l_{vr}}{l_{hr}} \right) \quad (5)$$

ただし $l_{vl}, l_{vr}, l_{hl}, l_{hr}$ は式(6)、(7)で求められる。

$$\begin{cases} l_{vl} = \|P_{25} - P_{18}\| \\ l_{vr} = \|P_{22} - P_{13}\| \end{cases} \quad (6)$$

$$\begin{cases} l_{hl} = \|P_{15} - P_{11}\| \\ l_{hr} = \|P_{20} - P_{16}\| \end{cases} \quad (7)$$

- 口の内側の面積 f_6

口の内周上の8つの特徴点を結んでできる八角形の面積。ただし、 L^2 で割って正規化を行う。 f_6 は式(8)により求められる。

$$f_6 = \frac{S(P_{29}, P_{31}, P_{32}, P_{33}, P_{30}, P_{34}, P_{35}, P_{36})}{L^2} \quad (8)$$

- 口の外側の面積 f_7

口の外周上の点を用いて f_6 と同様にして求めた面積。 f_7 は式(9)により求められる。

$$f_7 = \frac{S(P_{29}, P_{37}, P_{38}, P_{39}, P_{30}, P_{40}, P_{41}, P_{42})}{L^2} \quad (9)$$

- 口の内側の縦横の比 f_8

口の内周上の点で一番上にある点と一番下にある点を結んだ線分の長さを口の縦の長さ m_v 、最も左と右にある2点を結んだ線分の長さを口の横の長さ m_h として縦横の比を算出し、長さ m_h が0である場合を考慮して、縦横の比の逆正接を取り、式(10)のように求められる。

$$f_8 = \tan^{-1} \frac{m_v}{m_h} \quad (10)$$

ただし m_v, m_h は式(11)で求められる。

$$\begin{cases} m_v = \|P_{41} - P_{38}\| \\ m_h = \|P_{30} - P_{29}\| \end{cases} \quad (11)$$

- 口の外側の縦横の比 f_9

口の外周上の点を用いて f_8 と同様にして求めた縦横比。 f_9 は式(12)により求められる。

$$f_9 = \tan^{-1} \frac{m_v}{m_h} \quad (12)$$

ただし m_v, m_h は式(13)で求められる。

$$\begin{cases} m_v = \|P_{35} - P_{32}\| \\ m_h = \|P_{30} - P_{29}\| \end{cases} \quad (13)$$

- 口角の上がっている度合 f_{10}

口の両端の点 P_{29} と P_{30} と、口の最も上の点と下の点 P_{32} と P_{35} から、式(14)によって求められる。

$$f_{10} = \frac{(y(P_{29}) + y(P_{30})) - (y(P_{32}) + y(P_{35}))}{|y(P_{32}) + y(P_{35})|} \quad (14)$$

ここで、 $y(P)$ は特徴点 P の y 座標を表す。 P_{29} と P_{30} の y 座標の平均値が P_{32} と P_{35} の y 座標の平均値より大きければ正の値を取るため、この値が大きいくほど口角が上がっているといえる。分母は正規化のための項である。

- 口角の角度 f_{11}

口の外周上の最も左側(右側)の点と1つ内側の上下の点1点ずつを結んで構成される2本の線分のなす角の平均値を取ったもので、式(15)により求められる。

$$f_{11} = \frac{A(P_{29}, P_{31}, P_{36}) + A(P_{30}, P_{33}, P_{34})}{2} \quad (15)$$

ここで、 $A(p, q, r)$ は顔特徴点 p, q と p, r をそれぞれつないで構成される2本の線分のなす角で、式(16)によって求められる。

$$A(p, q, r) = \cos^{-1} \frac{(p - q) \cdot (p - r)}{\|p - q\| \cdot \|p - r\|} \quad (16)$$

$$(0 \leq A(p, q, r) \leq \pi)$$

以上、11個の特徴量をライフログ映像中の各フレーム画像に対して求め、フレームごとに11次元ベクトル (f_{m1}, \dots, f_{m11}) を構成している。ここで、 m はフレーム番号を表す。

2.3 表情強度

2.2で示した特徴量11個を利用し表情強度を算出する。まず映像ごとに、学習用データとして表情が表出しているフレームを12フレームと無表情のフレームを12フレームの計24フレームを選択する。

特徴量毎に12フレームの表情強度の平均値を g_1, g_2, \dots, g_{11} (表情表出フレーム) と G_1, G_2, \dots, G_{11} (無表情フレ

ーム)として算出する。次に、式(17)により、特徴量毎に閾値 S_1, S_2, \dots, S_{11} を算出する。

$$S_j = \frac{g_j + G_j}{2} \quad (j = 1, \dots, 11) \quad (17)$$

映像の m フレーム目の各特徴量を $f_{m1}, \dots, f_{mx}, \dots, f_{m11}$ とすると、 m フレーム目の表情強度 E_m は式(18)のように求められる。

$$E_m = \sum_{j=1}^{11} (f_{mj} - S_j) \quad (18)$$

3. 提案手法

先行研究では、特徴量の評価を行っておらず、表情強度の推定精度に課題が残るという点や、「喜び」と「驚き」しか取り扱っていないという点が問題であった。

そこで今回、表情強度の推定精度を向上し、対象とする表情の範囲を広げるために、新たな特徴量の提案を行い、既存の特徴量と共に評価を行った。まず、新たな特徴量について説明する。

3.1 新たな特徴量

以下に新たに提案した特徴量を示す。またこれらの特徴量は 2.2 に示した特徴量と同様、表情表出時に人間の目で見えて大きく位置関係が変化していると考えられる点を組み合わせて作成した。

- ほうれい線の傾き f_{12}

左(右)のほうれい線と鼻の点 $P_{45}, P_{46}, P_{49}, P_{50}(P_{43}, P_{44}, P_{47}, P_{48})$ から、最小二乗法(線形)により算出し、左右のほうれい線がほぼ左右対称と考え、式(19)のように定める。ただし、 b_l は左のほうれい線の傾き、 b_r は右のほうれい線の傾きである。

$$f_{12} = \frac{b_l - b_r}{2} \quad (19)$$

- 口の両端と鼻の頂点で結んだ三角形の面積 f_{13}

口の最も左と右にある点と鼻の頂点を結んでできる三角形の面積。ただし、 L^2 で割って正規化を行う。 f_{13} は式(20)のように定める。

$$f_{13} = \frac{S(P_{29}, P_{30}, P_{51})}{L^2} \quad (20)$$

- 口と鼻の間の距離 f_{14}

上唇の外側の3点と鼻の下部の3点を左から順に1対1対応で結んでできる3本の線分の長さの平均値。ただし、 L で割って正規化を行う。 f_{14} は式(21)のように定める。

$$f_{14} = \frac{\sum_{i=31}^{33} \|P_i - P_{i+21}\|}{3 \cdot L} \quad (21)$$

- ほうれい線と口の両端の点を結んだ四角形の面積 f_{15}

ほうれい線の最下部の点と口の両端の点で結んでできる四角形の面積。ただし、 L^2 で割って正規化を行う。 f_{15} は式(22)のように定める。

$$f_{15} = \frac{S(P_{29}, P_{30}, P_{46}, P_{43})}{L^2} \quad (22)$$

- 目と眉の面積 f_{16}

左目(右目)の最も左と右にある点と左眉(右眉)の最も左と右にある点を結んでできる四角形の面積の平均値。ただし、 L^2 で割って正規化を行う。 f_{16} は式(23)のように定める。

$$f_{16} = \frac{1}{2 \cdot L^2} \left\{ S(P_1, P_{11}, P_{15}, P_5) \right\} \quad (23)$$

3.2 標準化

統計学分野に標準化という手法がある。これは、ある数値データを標準化することで平均が 0、分散が 1 の正規分布になるように値を補正することである。こうすることで、単位の異なるデータの分布の比較などが行えるようになる。

先行研究では、11 種類の特徴量を単純に足し合わせて表情強度を算出しており、単位の違いを考慮していなかった。そこで本研究では、特徴量を標準化することで単位が異なる数値データの大きさの違いをなくす。以下に標準化の手法を述べる。

標準化したい数値データ群を $x_i (i = 1, \dots, n)$ 、標準化したい数値データ群の平均値を μ 、その標準偏差を σ とすると、 μ と σ はそれぞれ式(24)、(25)により算出できる。ここで、 n は標本数である。

$$\mu = \sum_{i=1}^n \frac{x_i}{n} \quad (24)$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n}} \quad (25)$$

x_i を標準化した数値データの値を Z_i とすると式(26)によって算出できる。

$$Z_i = \frac{x_i - \mu}{\sigma} \quad (26)$$

3.3 表情強度

提案手法では特徴量の評価を行い、各表情で使用する特徴量を個別に定める。使用する特徴量については 4.3.1 で述べる。算出された特徴量を 3.2 の方法を用いて標準化を行い、その値を用いて先行研究における式(17)、式(18)を用いて表情強度を算出する。

4. 実験

先行研究で使用されていた特徴量と新たに提案した特徴量の有効性を検証する実験を行う。また、有効だと判断した特徴量を用いて表情強度の推定を行い、先行研究での手法との比較を行う。

4.1 実験の目的

実験の主な目的は以下の4つである。

- 表情強度を算出するための特徴量の有効性を調べる。
- 有効であると判断した特徴量を使用し表情強度の推定を行う。
- 先行研究での手法との比較を行う。
- 先行研究で扱われていない表情に対する表情強度推定精度の評価を行う。

4.2 実験方法

MMI Facial Expression Database[5]に含まれる動画データセットを用いて表情強度の推定を行った。このデータセットには基本 6 表情のデータが存在する。本実験では喜び 8 人、驚き 11 人、嫌悪 8 人、怒り 6 人、悲しみ 5 人、恐怖 6 人分の表情強度を求めた。これらのデータは、筆者の一人が目視により主観的にその表情が表出していると判断した動画を選択した。

本データセットは、各被験者の表情を撮影した動画で、無表情から始まり、だんだんと表情が変化していき表情が最も強く表出し、再び無表情に変化していく動画の構成になっている。図2に表情の変化の例を示す。

本実験では、動画から抽出したフレーム画像のうち、24フレームを学習用データとし、その動画内の全てのフレームをテスト用データとして用いた。各フレーム画像に表情が表出しているかどうかと、特徴点抽出が正しく行えているかどうかは、筆者の一人が主観的に目視によって判断した。

学習用データとして用いる24フレームは、上記と同様にして、筆者の一人の主観的な判断によって決定した。なお、先行研究においても、同様の方法で学習用データとテスト用データが定められている。

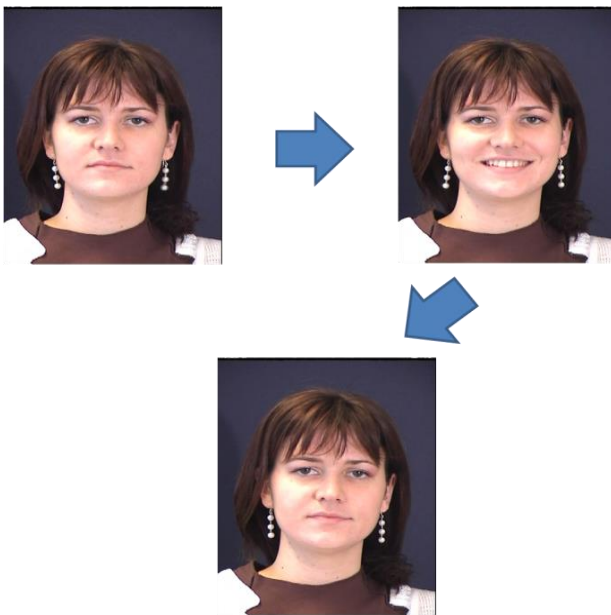


図2 表情の変化の例

4.3 実験結果

4.3.1 特徴量の評価

各表情で2.2、3.1における16種類の特徴量を算出し、データセット中の各動画について、顔特徴点の変化と特徴量の変化を比較しながら表情の変化と特徴量の変化が一致しているかどうかを筆者の一人が判断した。表1に表情強度の算出に使用すべきと判断した特徴量を示す。

表1 使用する特徴量

表情	特徴量
喜び	$f_2, f_4, f_5, f_6, f_7, f_8, f_9, f_{10}, f_{11}, f_{12}, f_{13}, f_{14}$
驚き	$f_1, f_2, f_3, f_4, f_5, f_6, f_7, f_8, f_9, f_{11}, f_{16}$
嫌悪	$f_1, f_2, f_3, f_4, f_5, f_6, f_7, f_8, f_9, f_{10}, f_{11}, f_{14}, f_{16}$
怒り	$f_1, f_2, f_3, f_4, f_5, f_7, f_9, f_{11}, f_{14}$
悲しみ	$f_2, f_3, f_4, f_5, f_7, f_9, f_{10}, f_{12}, f_{13}, f_{14}, f_{15}, f_{16}$
恐怖	$f_1, f_2, f_3, f_4, f_5, f_7, f_9, f_{11}, f_{12}, f_{13}, f_{14}, f_{15}, f_{16}$

4.3.2 表情強度

表1に示した特徴量を用いて各表情で表情強度を算出した。また、先行研究における手法での表情強度も比較のために算出した。その結果を表2~7に示す。ここで、「id~」は動画のファイル名である。各表には、表情表出フレームと無表情フレームにおける、提案手法と先行研究での表情強度の平均値とその差分を示している。比較のため、提案手法と先行研究の手法で得られた表情強度の値を標準化している。また、表情表出フレームと無表情フレームは、筆者の一人が主観的に定めている。各表について、「無」、「有」の列は、それぞれ無表情・表情表出フレームの表情強度を示している。

表2 喜びの表情強度

動画	提案手法			先行研究		
	無	有	差分	無	有	差分
id1795	-1.37	0.59	1.96	-1.36	0.59	1.95
id1825	-0.83	0.39	1.22	-0.04	0.002	0.042
id1856	-1.14	0.86	2.00	-1.12	0.95	2.07
id1877	-1.22	0.66	1.88	-1.32	0.77	2.09
id1924	0.38	-0.02	0.40	1.19	-0.33	1.52
id1980	-1.56	0.51	2.07	-1.27	0.55	1.82
id2007	-0.71	0.20	0.91	-0.63	0.18	0.81

表3 驚きの表情強度

動画	提案手法			先行研究		
	無	有	差分	無	有	差分
id1771	-1.25	0.86	2.11	-1.27	0.80	2.07
id1806	-1.33	0.67	2.00	-1.21	0.76	1.97
id1818	-1.26	1.03	2.29	-1.26	0.94	2.20
id1829	-1.29	0.84	2.13	-1.39	0.86	2.25
id1843	-1.45	0.61	2.06	-1.46	0.59	2.05
id1869	-1.31	0.80	2.11	-1.45	0.68	2.13
id1914	-0.45	0.53	0.98	-0.62	0.92	1.54
id1928	-1.07	0.58	1.65	-0.56	0.029	0.589
id1948	-1.70	0.61	2.31	-1.59	0.64	2.23
id1970	-1.36	0.55	1.91	-1.48	0.60	2.08
id2004	-0.96	0.75	1.71	-1.01	0.86	1.87

表4 嫌悪の表情強度

動画	提案手法			先行研究		
	無	有	差分	無	有	差分
id1821	-1.21	0.83	2.04	-0.97	0.72	1.69
id1832	-0.87	0.64	1.51	-0.25	0.60	0.85
id1875	-1.31	0.85	2.16	-1.17	0.79	1.96
id1882	-0.91	0.48	1.39	-0.62	0.18	0.80
id1952	-0.65	0.60	1.25	-1.15	0.91	2.06
id1964	-1.43	0.58	2.01	-0.77	0.30	1.07
id1991	-0.15	-0.13	0.28	0.56	-0.46	1.02

表5 怒りの表情強度

動画	提案手法			先行研究		
	無	有	差分	無	有	差分
id1809	-1.36	1.14	2.50	-1.46	1.24	2.70
id1866	-0.49	0.26	0.75	0.80	-0.45	1.25
id1874	-0.63	0.23	0.86	0.50	-0.016	0.516
id1900	-0.61	0.25	0.86	-0.20	0.028	0.228
id1918	0.10	0.36	0.26	-0.54	0.71	1.25
id1973	-1.25	0.26	1.51	1.07	-0.43	1.50

表6 悲しみの表情強度

動画	提案手法			先行研究		
	無	有	差分	無	有	差分
id1804	-1.17	0.87	2.04	0.22	-0.0004	0.2204
id1827	-0.47	-0.011	0.481	0.69	-0.38	1.07
id1863	0.80	-0.28	1.08	0.33	-0.0062	0.3362
id1926	-1.26	0.71	1.97	0.079	0.079	0
id1943	0.16	-0.35	0.51	0.36	-0.64	1.00

表7 恐怖の表情強度

動画	提案手法			先行研究		
	無	有	差分	無	有	差分
id1802	-0.98	1.11	2.09	-0.77	0.70	1.47
id1814	0.31	-0.45	0.76	0.55	-0.43	0.98
id1876	-1.08	0.70	1.78	-0.50	-0.082	0.582
id1884	-1.27	0.74	2.01	-0.62	0.58	1.20
id1922	-1.05	1.06	2.11	-0.88	1.02	1.90
id1934	0.15	-0.38	0.53	0.72	-0.24	0.96

5. 考察

5.1 特徴量の評価

以下に各表情での特徴量の評価と選別理由について述べる。

● 喜び

喜びでは口周りの変化が大きく、口に関する特徴量の値は、喜びの表情の強さに比例する傾向が見られた。そのため、口に関する特徴量である $f_6, f_7, f_8, f_9, f_{10}, f_{11}$ に関しては使用できると判断した。一方、目の周りの変化は少ないものが多く、特に f_1, f_3 は特徴量にばらつきがみられたので使用できないと判断した。その他の f_2, f_4, f_5 に関してはそれほどばらつきもなく、笑顔になった際に目を細める人物もいることも考慮し省くべきではないと判断した。

また、新たに提案した特徴量に関しては、 f_{12}, f_{13}, f_{14} は表情の変化とともに特徴量に変化していたので使用できると判断した。しかし f_{15} に関しては、ほうれい線の特徴点のばらつきが大きく、また f_{16} も f_1, f_3 同様にばらつきが大きかったため使用できないと判断した。

● 驚き

驚きも“喜び”と同様に口周りの変化が大きいが、 f_{10} に関しては、特徴点のばらつきが大きかったため省いた。また、“驚き”では眉や目の周りに大きな変化が現れており f_1, f_2, f_3, f_4, f_5 に関しても使用できると判断した。しかし、 f_1 と f_3 に関しては表情強度と負の相関を示していたため、“-1”をかけて使用している。

$f_{12}, f_{13}, f_{14}, f_{15}$ は表情が変化してもあまり変化がないものや顔特徴点の抽出精度が良くないためにばらつきが大きくなっているものであったため使用できないと判断した。 f_{16} は上述の通り変化が大きいため使用できると判断した。

● 嫌悪

眉間にしわを寄せて嫌悪感を示すような表情が多く、目の周りの特徴量がおおむね表情の強さに比例していた。また、 f_2, f_3, f_4, f_5 に関しては目を細めるものが多く、表情強度と負の相関を示していたため“-1”をかけて使用している。 $f_6, f_7, f_8, f_9, f_{11}$ に関しては、多少のばらつきはあるものの省くほどではないと判断し使用している。また、 f_{10} に関しては、ばらつきが大きかったが口をへの字にして嫌悪感を表す場合もあると考えたため省くべきではないと判断し使用している。 f_{12}, f_{13}, f_{15} は特徴点の抽出精度が良くないため使用できないと判断した。 f_{14} と f_{16} は表情の強さに比例しているが f_{16} は表情強度と負の相関を示していたため、“-1”をかけて使用している。

● 怒り

f_1, f_2, f_3, f_4, f_5 は嫌悪同様に顔をしかめる人物が多く変化が現れていた。また、 f_2, f_3, f_4, f_5 は、“-1”をかけて使用している。口に関しては閉じている者が多く、そのため口の内側の特徴点にばらつきが見られたため f_6, f_8 は使用できないと判断した。口の外側の特徴点にはばらつきはあまり見られなかったため f_7, f_9, f_{11} は使用できると判断した。 f_{10} は特徴点抽出の精度が良くないため使用できないと判断した。 $f_{12}, f_{13}, f_{15}, f_{16}$ はほぼ一定の値を示していたが、ノイズになる可能性があるため判断し省いた。 f_{14} はほぼ一定の値であったが、少量の変化がみられる動画もあったため、省くべきではないと判断し使用している。

“怒り”は人物によって表現の仕方が異なり、表情の変化の仕方が異なるものが多く、特徴量を選別するのが難しいものが多かった。

● 悲しみ

うつむいて下を見るような表情が多かったため f_2, f_3 は表情の強さと同様の変化を示すものが多かった。また、 f_4, f_5 は表情の強さとは逆の変化を示していたので“-1”をかけて使用している。眉の変化はほとんどなくばらつきもあったため f_1 は使用できないと判断した。口の変化に関しては“怒り”と同様で f_6, f_8 は使用できない、 f_7, f_9 は使用できると判断した。しかし f_{10} は、口角を下げるものが多く表情の強さとは逆の変化を示していたので“-1”をかけて使用している。 f_{11} は口を閉じている者が多く、また一般的にもそれほど変化がないと考えたので省いた。 $f_{12}, f_{13}, f_{15}, f_{16}$ はほぼ一定の値を示すものが多かったが少量の変化を示すものがあったので使用している。 f_{14} は表情の強さとは逆の変化を示していたので“-1”をかけて使用している。

● 恐怖

目の周りの変化が大きく、 f_2, f_3, f_4, f_5 は使用できると判断した。また、 f_1 は表情の強さとは逆の変化を示していたので“-1”をかけて使用している。口に関しては内側の点の抽出精度がよくなかったため f_6, f_8 は使用できないと判断した。口の外側に関しては誤検出も少なく、 f_7, f_9, f_{11} は使用できると判断した。 f_{10} はノイズになっていたため使用できないと判断した。 $f_{12}, f_{13}, f_{14}, f_{15}, f_{16}$ はほぼ一定の値を示すものが多かったが、変化を示しているものもあったので使用している。 f_{12}, f_{13}, f_{14} は“-1”をかけて使用している。

5.2 表情強度

各表情での比較結果に対する考察を行う。

● 喜び

id1825 については、先行研究の手法では、無表情時と表情表出時の表情強度の値にほとんど差がないのに対し、提案手法では、両者の表情強度の差分は十分に大きくなっており、適切に表情強度が推定できていると言える。図 3、図 4 にこの人物における表情強度の変化を示したグラフを示す。ここで横軸はフレーム番号、縦軸は表情強度である。

この人物は、1~9 フレーム目まで無表情であり、10 フレーム目から表情が変化しはじめる。14 フレーム目で表情が表出し、55 フレーム目から再び表情が変化し 64 フレーム目で無表情となる。37 フレーム目付近と 61 フレーム目付近で表情強度の値が高くなってしまっているのは、まばたきによって顔特徴点が誤検出してしまったために目と眉に関する特徴量に影響が出たためである。

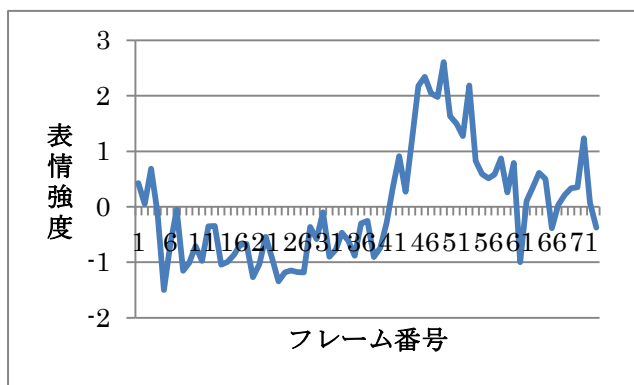


図 3 先行研究における id1825 の表情強度の変化

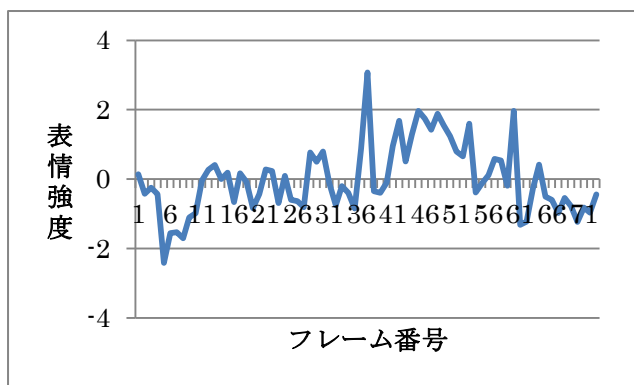


図 4 提案手法における id1825 の表情強度の変化

id1924 では、提案手法と先行研究のいずれにおいても、無表情時の方が表情表出時よりも表情強度が高くなってしまっている。これは、この人物が表情表出の際に口を閉じたまま横に伸ばして笑顔を表しているため、表情の変化による特徴点の位置変化よりも特徴点のばらつきが特徴量に出てしまったことが原因と考えられる。この事例については、いずれの手法も正しく表情強度を推定することはできないが、無表情時においては、提案手法の表情強度の方が低く、表情表出時では、提案手法の表情強度の方が高くなっていることから、提案手法の推定結果の方がより適切であると言える。

● 驚き

多くの事例について、提案手法と先行研究の手法の間に大きな違いは見られなかった。また、いずれの手法も、すべての事例において、無表情時の方が表情表出時よりも表情強度が小さくなっており、ある程度適切な推定が出来ていると言える。

id1914 については、人物の表情変化がそれほど大きくなく、表情強度の推定値のばらつきが大きくなっており、提案手法の推定結果は良好であるとは言えない。

しかし、同様に人物の表情変化が比較的小さい id1928 では、先行研究における表情表出時の表情強度が非常に小さい値であるのに対し、提案手法の表情強度は十分大きい値を取っており、より適切な推定が出来ていると言える。

● 嫌悪

id1952 では、人物が表情の変化と同時に顔を傾けており、それが提案手法における標準化をすることで、複数の特徴量に影響したため、先行研究の手法と比べて表情強度の推定精度がやや低下したと考えられる。

一方、id1991 については、先行研究の手法では無表情時の方が表情表出時よりも表情強度が高くなっているのに対し、提案手法では、無表情時の表情強度の方が低くなっており、より正確な表情強度の推定が出来ていると言える。さらに他の人物でも先行研究の手法より、提案手法の表情強度が十分に大きい値を取っているため、標準化を行うことは有効であると考えられる。

● 怒り

先行研究の手法では、半数の事例(id1866,id1874,id1973)について、無表情時の方が表情表出時よりも表情強度が高くなっているが、提案手法では、すべての事例について、無表情時の表情強度の値の方が小さく、適切な推定が出来ていると言える。

怒りの表情は、人物ごとに多様であることや、比較的表情の変化が乏しく、特徴点のばらつきが大きいことから、正確な表情強度の推定が難しい表情であると言える。この実験結果から、提案手法は、先行研究の手法と比べて、表情強度の推定が難しい表情に対しても有効性が高いと言える。

● 悲しみ

id1863 は人物の表情変化がそれほど大きくない。そのため、特徴点のばらつきが特徴量に出てしまい、標準化したことでそれが大きく表情強度に影響したために提案手法の表情強度推定精度が先行研究と比べ低い。しかし、その他の事例においては、おおむね先行研究よりも推定精度が良いと言える。

悲しみも怒りと同様に、表情の変化が比較的乏しく、特徴点のばらつきが大きい表情である。この結果からも、提案手法は、先行研究の手法と比べて、表情強度の推定が難しい表情に対して有効であると考えられる。

● 恐怖

id1814 と id1934 については、表情の変化が比較的小さく、特徴点のばらつきが大きいこと、提案手法と先行研究のいずれにおいても、表情表出時よりも無表情時の方が、表情強度が大きくなっており、提案手法による改善は見られなかった。

しかし、id1876 については、追加した特徴量や標準化を行うことで、表情の変化と比例した変化を示す特徴量の多くが表情強度に影響を与えたため、提案手法は先行研究と比べて、より適切に表情強度の推定が出来ていると言える。

図 5、図 6 に、この人物における表情強度の変化を示したグラフを示す。

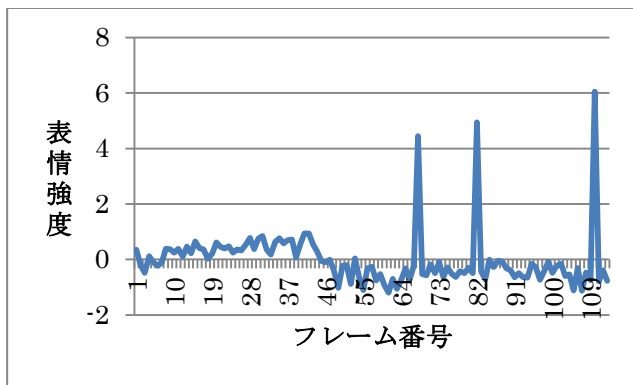


図 5 先行研究における id1876 の表情強度の変化

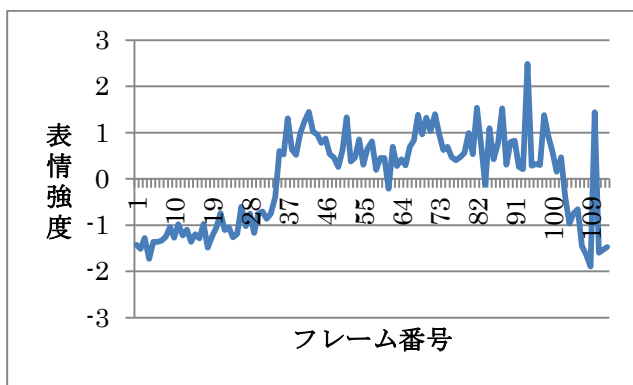


図 6 提案手法における id1876 の表情強度の変化

6. まとめ

本研究では、ライフログ映像におけるシーン検索時に表情表出の程度を考慮することを目的として提案された表情強度の算出に用いられている特徴量の評価と選別を行った。また、表情強度の推定精度の向上を目的として、新たな特徴量を提案した。選別した特徴量から表情強度の推定を行い、先行研究と比較実験を行った。また、先行研究では扱っていなかった表情に対しての実験も行った。その結果、どの表情でも先行研究よりも適切な表情強度推定が出来ていた。また、提案手法は怒りや悲しみのように正確な強度推定が難しい表情に対して有効であることを示した。

今後の課題としては、提案手法でも無表情時の方が、表情強度が大きくなってしまっている動画が見られたため、表情強度推定の精度向上があげられる。

また本研究では、学習用データとテスト用データを分けて使用したが、学習用データとテスト用データを分けて再実験を行う必要がある。さらに、今回使用した動画は意図的に表出された表情からなるデータセットで実験したため、自然に表出した表情が含まれるライフログ映像においても、今回提案した手法が適用できるか確認する必要がある。

7. 謝辞

本研究の一部は、科学研究費補助金（課題番号：15K15993）による。ここに記して謝意を表す。

8. 参考文献

- [1]相澤清晴, ”体験の情報処理-ライフログの取得と処理-“, 電子情報通信学会技術研究報告, パターン認識・メディア理解, 103(738), pp. 284-290, 2004
- [2]野宮浩揮, 森國淳司, 宝珍輝尚, ”ライフログ映像検索に向けた顔特徴点の位置関係に基づく効率的な表情表出シーン検出”, 電子情報通信学会論文誌, J96-D(5), pp. 1279-1289, 2013
- [3]Atsushi Morikuni, Hiroki Nomiya, Teruhisa Hochin, ”Expression Strength for the Emotional Scene Detection from Lifelog Videos”, International Journal of Computer & Information Science, 16 (1), pp. 32-39, 2015
- [4]Luxand, Inc., ”Luxand FaceSDK4.0”, <http://www.luxand.com/facesdk>
- [5]MMI Facial Expression Database, <http://mmifacedb.eu/>