

変化点検知を活用した顔特徴量時系列データからの注目点の抽出

平山 孝輔[†] 佐伯 幸郎[†] 中村 匡秀^{†,††}

[†] 神戸大学 〒657-8501 神戸市灘区六甲台町 1-1

^{††} 理化学研究所・革新知能統合研究センター 〒103-0027 東京都中央区日本橋 1-4-1

E-mail: [†]hirayama@ws.cs.kobe-u.ac.jp, ^{††}sachio@carp.kobe-u.ac.jp, ^{†††}masa-n@cs.kobe-u.ac.jp

あらまし 我々は以前より、顔の微細な動きを特徴量として捉え、ケアの定量的評価に寄与するための「表情センシングサービス」の開発を行っている。効率的なケア評価を行うためには、膨大な特徴量時系列データの中から「注目点」を自動的に見つけ出す手法が必要になる。そこで本稿では、被験者の表情に大きな変化がみられるタイミングを注目点とし、変化点検知手法を用いて特徴量データからの注目点の抽出を試みる。これにより、特徴量が大きく変動したり、傾向が変わった箇所を自動的に検知することで、注目点を効率的に探し出せるようにする。また、複数の特徴量を組み合わせた「複合特徴量」を考案し、より実際の表情の変動に即した情報を得ることを目指す。

キーワード 科学的介護、表情分析、ケア効果、変化点検知

Extracting Points of Interest from Facial Feature Time-Series Data by Changing Point Detection

Kosuke HIRAYAMA[†], Sachio SAIKI[†], and Masahide NAKAMURA^{†,††}

[†] Kobe University Rokkodai-cho 1-1, Nada-ku, Kobe, Hyogo, 657-8501 Japan

^{††} Riken AIP 1-4-1 Nihon-bashi, Chuo-ku, Tokyo 103-0027

E-mail: [†]hirayama@ws.cs.kobe-u.ac.jp, ^{††}sachio@carp.kobe-u.ac.jp, ^{†††}masa-n@cs.kobe-u.ac.jp

Abstract We have been developed “*Facial Expression Sensing Service*” for quantitative care evaluation by capturing subtle facial movements as feature values. For efficient care evaluation, we need a method to automatically find “*point of interest*” in a huge amount of time-series data of feature value. In this study, we define points of interest as the moments when a subject’s facial expression changed significantly, and attempt to extract them from feature value data using *changing point detection* method. For this, we can detect when feature values changed significantly or their trend changed, so we can efficiently find points of interest. We also devise “*composite feature value*” that combines multiple feature values to obtain information that is more close to actual facial expression changes.

Key words Scientific long-term care, Facial expression analysis, Care effect, Changing point detection

1. はじめに

近年、日本は超高齢化社会に直面しており、介護を必要とする高齢者の数が増加したことで、慢性的な介護リソースの不足に苛まれている。こうした中、日本は「科学的介護」を国策として実践することを宣言し、介護の効果を科学的エビデンスによって検証することで、介護リソースの適切な配分を目指している [1]。科学的介護の実現のためには、ケアの質や効果をデータを用いて定量的、客観的に評価することが重要である。

以前より我々は、客観的なケアの効果を示す指標の一つとして、顔の表情に着目し、微細な表情の変化を数値データとして取得するためのサービスである「表情センシングサービス」の

検討、開発を行っている [2]。実際に得られたデータを活用した分析を行うに当たっては、ケアの全体を通して対象者の感情がどう変わったかというマクロな視点と、ケアの最中のどのタイミングで対象者の感情に変化が生じたかのミクロな視点からの分析が考えられる。マクロな視点からは、そのケアの満足度や、対象者に対する適性が考察でき、ミクロな視点からは、対象者がどのような会話やコンテンツに対してポジティブな反応を示したかを推測することができる。こういった分析は、ケアの効果を評価するとともに、対象者ひとりひとりの好みに応じた最適なケアを提供するために有用である。しかし、表情センシングサービスは秒単位で特徴量の測定を行うため、得られる時系列データは膨大になる。そのため、特にミクロな視点での

分析において、注目すべき表情の変化点を探し出すことが困難となるおそれがある。

そこで本稿では、表情センシングサービスで得られる特徴量時系列データの中からケア効果を評価するうえで重要な点を**注目点**とし、これを効率的に抽出することを考える。本稿では特に、被験者の表情に大きな変化がみられるタイミングを注目点と定め、これを見つけ出すために**変化点検知**を用いる。そして、特徴量が大きく変動したり、値の傾向が変化する時点を示す**変化点**を自動的に検知し、被験者の表情が大きく変わった点を探し出す。これにより、膨大なデータの中から注目点を容易に見つけ出すことができ、先述のミクロな視点からの分析への寄与が期待できる。

本稿では、本稿筆者の顔から作成した評価用のデータや、ケアを受ける被験者の映像から作成したデータを **ChangeFinder** という変化点検知手法を用いて分析する実験を行い、実際の様子と比較することで、その有用性を考察した。また、単一の特徴量だけではなく、複数の特徴量を組み合わせた**複合特徴量**を作成し、これに対しても変化点検知を適用することで、より実際の表情の変動に即した情報を得るための実験を行った。実験の結果、本稿の手法によって表情の変化や顔の大きな動きを注目点として抽出できることを確認した。

2. 準備

2.1 表情センシングサービス

1.で述べた通り、我々は以前より「表情センシングサービス」の検討・開発を行っている。表情の分析タスクには、機械学習を用いて構築されたコグニティブ API を使用することも考えられる。実際に、過去に我々の研究グループではコグニティブ API を利用したリアルタイム感情分析サービス「Face Emotion Tracker (FET)」を開発し、ケアの定量的評価を試みている [3]。しかし、既存の API は汎用目的であり、例えばケアの対象となる高齢者の表情の変化が機能性あるいは認知機能の障害により弱まっているような場合、その微妙な変化を捉えきれない可能性がある。また、既存の API による感情分析はブラックボックス化されており、なぜその結果が出力されたかに至るまでの結果を追うことができないため、得られたデータを介護効果のエビデンスなどに用いようとする場合に困難が生じる懸念がある。

表情センシングサービスでは、眉、目、口など、顔の特徴的なパーツの座標（＝特徴点）を顔画像から抽出し、その2点を結ぶ長さを**特徴量**として計測し続け、時系列データとして記録する。この特徴量の変化をトラッキングすることで顔の動き＝表情の変化の度合いを測定する。これにより、表情の変化に乏しい人を対象に測定を行う際にも、顔の微妙な変化を見逃さずに記録することが期待できる。また、表情の動きに伴う感情の変化を推測する際に、その表情変化の度合いを明確な数値として把握することで、説明可能な感情分析やケアの評価に寄与することを目指している。

2.2 変化点検知

変化点検知とは、時系列データに対する変化点を検出する技術である。変化点検知の種類には、今までに蓄積されたデータ

の中でバッチ的に変化点を探す**オフライン検知**と、新しいデータが得られるたびに検知を行い得られたデータが変化点であるかどうかを判断する**オンライン検知**がある。本稿では特に、オンライン変化点検知手法の1つである ChangeFinder [4] を特徴量時系列データの分析に採用する。ChangeFinder の計算量は、計算に用いるデータ数に関しては線形のオーダーとなるため、変化点検知を素早く行うことが可能である [5]。

ChangeFinder は、Auto Regression Model (AR モデル) と平滑化を利用した2段階学習を特徴とする変化点検知手法である。時系列が従うモデルが変化した点を検知する仕組みとなっており、変化の程度の大きさをスコア（＝変化スコア）として計算する。内部に AR モデルを2つ持っており、1つ目は元のデータの時系列モデルを学習し、2つ目は1つ目の AR モデルを用いて得られた変化の程度を時系列データとして扱い学習する。これらにより、小さなノイズの影響による変化を除去する仕組みとなっている。また、AR モデルの学習に Sequentially Discounting AR model learning (SDAR) という忘却型学習アルゴリズムを用いており、これによりオンライン処理が可能な処理スピードを実現している。さらに、一般に AR モデルは時系列の定常性を仮定したモデルであるが、ChangeFinder は忘却型学習を取り入れることで、非定常なデータにも対応する。

表情センシングサービスによって得られるような、1次元の時系列データに対する ChangeFinder が行う処理のステップを以下に示す。

Step 1: 入力されたデータに対して、1段階目の AR モデルの学習を行う。この処理ステップでは、SDAR アルゴリズムを用いて、AR モデルが持つ平均・共分散・分散・自己回帰係数の更新を行う。AR モデルの次数を k 、時刻 t ($= 1, 2, \dots$) に対する入力データを x_t 、今回算出される平均を μ_t 、自己回帰係数を $a_{t,i}$ 、共分散を $C_{t,i}$ 、分散を σ_t 、忘却率を r ($0 < r < 1$) としたとき、次の式で更新が行われる。

$$\mu_t = (1-r)\mu_{t-1} + rx_t \quad (1)$$

$$C_{t,i} = (1-r)C_{t-1,i} + r(x_t - \mu_t)(x_{t-i} - \mu_t) \quad (i = 0, \dots, k) \quad (2)$$

$$C_{t,i} = \sum_{j=1}^k a_{t,j} C_{t,i-j} \quad (i = 1, \dots, k) \quad (3)$$

(3) は Yule-Walker 方程式と呼ばれる。ただし、 $C_{t,-i} = C_{t,i}$ とする。これを解くことによって $a_{t,j}$ を計算し、次に以下を計算する。

$$\hat{x}_t = \sum_{i=1}^k a_{t,i}(x_{t-i} - \mu_t) + \mu_t \quad (4)$$

$$\sigma_t = (1-r)\sigma_{t-1} + r(x_t - \hat{x}_t)(x_t - \hat{x}_t) \quad (5)$$

Step 2: 確率密度分布を用いてスコアを算出する。この処理ステップでは、Step 1. の結果得られた平均・分散をもとに、正規分布の確率密度関数を利用することで、スコアの算出を行う。このときの確率密度を $p_t(x_t)$ としたとき、 $p_t(x_t)$ およびスコア y_t は以下のように計算される。

$$p_t(x_t) = \frac{1}{\sqrt{2\pi\sigma_t}} \exp\left(-\frac{1}{2\sigma_t}(x_t - \hat{x}_t)^2\right) \quad (6)$$

$$y_t = -\log p_{t-1}(x_t) \quad (7)$$

Step 3: 得られたスコアを平滑化する．この処理ステップでは，Step 2. で得られたデータを平滑化することでノイズの除去を行う．平滑化するサイズを w としたとき，平滑化後の結果 $Score_t$ は以下のように求まる．

$$Score_t = \frac{\sum_{i=0}^{w-1} y_{t-i}}{w} \quad (8)$$

Step 4: Step 3. で得られたデータ $Score_t$ を新たな時系列データとし，2段階目の学習・スコア算出・平滑化を Step1～3 と同様に行う．そして，2段階目の学習の平滑化後のスコアが，変化スコアとして得られる．

3. 特徴量データを用いた注目点抽出実験

3.1 目的とアプローチ

本稿では1. で述べた「ミクロな分析」を表情センシングサービスによる特徴量時系列データを用いて行うことに焦点を当て，ケアが行われている最中，どの時点で被験者に反応があったか，つまり表情に変化が生じたかを効率的に調べることを目的とする．そのため，特徴量時系列データにおいて，大きな変動が認められる時点を「注目点」とし，これを抽出することを考える．本研究では，2.2 でも述べた通り，ChangeFinder という変化点検知アルゴリズムを用い，特徴量の変化箇所を的確に捉えた上での注目点抽出を試みる．

本研究のキーアイデアとして，「口幅」や「目の開き具合」といった単一の特徴量だけではなく，**複合特徴量**という概念を導入し，これに対しても変化点検知を適用することを考える．複合特徴量は，分析の目的に応じて複数の特徴量から作成された尺度であり，各々の特徴量の変化について考慮するよりも効率的に分析が行える．さらに，表情が変化する際には複数の顔のパーツが動くため，複数の特徴量を組み合わせた尺度の導入によってより実際の顔の動きを反映した妥当性の高い注目点抽出が可能になると考えられる．具体的な手法としては以下のようになる．

Step 1: 複数の特徴量 $x_1(t), x_2(t), \dots$ を組み合わせた複合特徴量 $X(t) = f(x_1(t), x_2(t), \dots)$ を作成する．

Step 2: $X(t)$ を ChangeFinder に入力し，変化スコア $S(t)$ を求める．

Step 3: ある閾値 γ を定め， $\gamma < S(t')$ となる点 t' を注目点の候補とする．

3.2 複合特徴量

本研究ではケアの効果測定という観点から，対象者がどれくらい楽しそうにしているかを示す「**笑顔スケール**」を複合特徴量として定める．笑顔スケールを作成するにあたって，どのような顔の動きが楽しそう・幸福の感情を示唆するかを考察する．

まず，様々な感情に対応する表情の特徴を記述した表情分析に関する書籍[6]や，これを検証した文献[7]を参照しまとめると，幸福を示す表情について以下のような特徴が挙げられる．

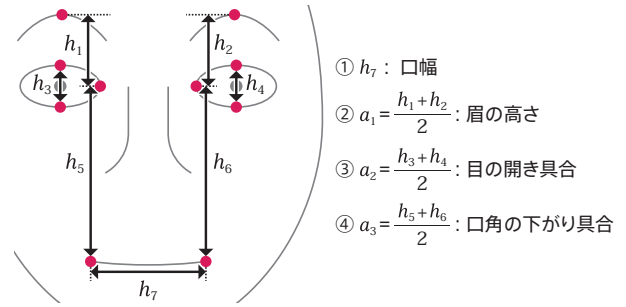


図1 笑顔スケールの作成に用いた特徴量

表1 本稿筆者の顔画像10枚から測定した特徴量の平均値

	口幅 h_7	眉の高さ a_1	目の開き具合 a_2	口角の下がり具合 a_3
平常時平均	0.8199	0.4270	0.0901	1.0431
笑顔時平均	0.8914	0.3981	0.0503	0.9817
差分	0.0715	-0.0289	-0.0398	-0.0614
差分比率	1.0000	0.4035	0.5561	0.8578



図2 平常時の顔のサンプル



図3 笑顔時の顔のサンプル

- ・ 眉は全体的に下に下がる
- ・ 細目になる
- ・ 口角は後ろに引かれ，また位置が上がる
- ・ 頬が上がる

こういった特徴から，笑顔スケールの作成にあたり，図1で示す4つの特徴量に着目する．これらの特徴量のうち，先に述べた表情の特徴と照らし合わせると，「口幅」 h_7 を除く3つの特徴量は，その値が**小さいほど**幸福を示していることになる．また，各特徴量ごとに変動の絶対量は異なる．例えば，口元の動きのほうが，目の開閉の動きよりも大きいはずである．そのため，各特徴量に一定の重みを付ける必要がある．この重みを決定するために，まず，後の予備実験で用いる顔画像の中から，本稿の筆者の平常時の顔と笑った時の顔画像（サンプルは図2，図3に示す）を10枚ずつ無作為に選択して特徴量を測定し，その平均値の差分をそれぞれ求めた．平均値の計算結果を表1に示す．「差分比率」は口幅に対する各特徴量の「差分」の比率を示したものである．この結果より，「口幅」 h_7 や「口角下がり具合」 a_3 と比べ，「眉の高さ」 a_1 ，「目の開き具合」 a_2 の変化の絶対量は小さく，半分程度であることが分かる．

以上の考察から，笑顔スケール X を以下のように決定した．

$$X = 2h_7 - a_1 - a_2 - 2a_3 \quad (9)$$

顔の動きの度合いには個人差があることを考慮し，表1の差

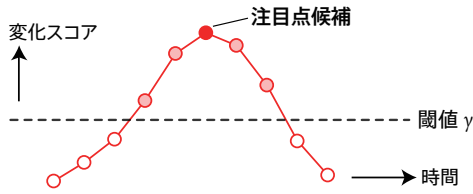


図4 注目点候補の説明

分比率をそのまま係数に用いることは避けた。また、 X の計算に用いる各特徴量は標準化を施した値を用いる。具体的には、もとの特徴量の値 $x(t)$ に対して、その平均が μ_x 、標準偏差が σ_x であるとき、 $z(t) = \frac{x(t) - \mu_x}{\sigma_x}$ で計算される値 $z(t)$ を特徴量の値として用いる。

3.3 実験の詳細

本稿の筆者の顔から作成したテスト用データ、およびケア最中の高齢者の顔画像から作成したデータについて、単一特徴量および複合特徴量である笑顔スケールの注目点抽出実験を行う。本実験では、Python 用の ChangeFinder ライブラリ¹⁾を使用する。このライブラリはパラメータとして AR モデルの次数 $order$ 、忘却率 r 、平滑化サイズ w をとる。第一段階学習ではこれらのパラメータがそのまま使われ、第二段階学習では平滑化サイズが $\text{round}(w/2.0)$ で求まる値となる。なお、今回の各実験では、パラメータを $order = 1$ 、 $r = 0.025$ 、 $w = 5$ と設定する。

ChangeFinder によって算出された変化スコア $S(t)$ のうち、図4のように、一定の閾値 γ を超えてから下回るまでの間で最もスコアが高い時点を注目点候補とする。今回の一連の実験においては、閾値 γ を変化スコア S の平均 μ_S と標準偏差 σ_S を用いて $\gamma = \mu_S + \sigma_S$ で定める。

4. 予備実験

4.1 概要

予備実験として、本稿の筆者の顔画像から作成したテスト用の特徴量データを用いた注目点の抽出実験を行った。

テスト用データは、本稿の筆者が通常の表情と笑顔の繰り返しを何回か行った映像を基に作成した。映像はノートパソコンの画面を見ながら内蔵カメラを用いて撮影した。特徴量は1秒間隔で測定し、明らかな外れ値は除去した。これにより不足したデータはその前のデータで埋め合わせている。このデータの単一特徴量および笑顔スケールに対して ChangeFinder を適用し、変化スコアの算出および注目点候補の抽出を行った。

本実験ではまず、どれだけ表情分析が効率化されるかを推定するため、特徴量データの総数 n で注目点候補数 n_p および閾値 γ を超えたデータの数 n_o を割った比率 n_p/n 、 n_o/n をそれぞれ求める。これらの値が小さいほど、確認すべきデータの数が減ることになり、分析がより効率化されるといえる。

次に、得られた注目点候補の妥当性を確認するために、目視による表情の変化点との比較を行う。まず、事前に目視により、図5のように、顔が (a) 楽しそうになった時点および (b) 楽しくなさそうになった時点を探し、イベント点として記録する。

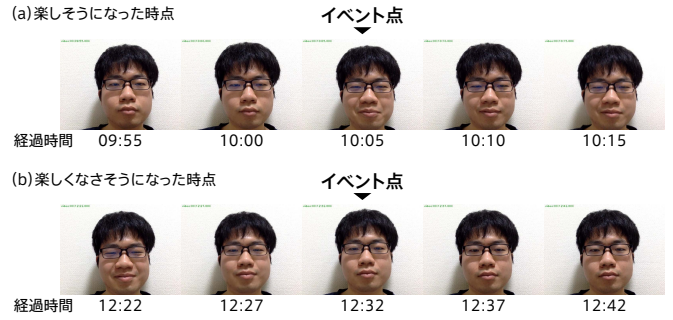


図5 イベント点の例

ただし、記録するのは明らかな顔の動きがみられる時のみとする。次に注目点候補のうち、イベント点とタイミングが一致するものを調べる。これを正解点とし、その数を正解数とする。本実験では、注目点候補がイベント点の前後10秒以内にあれば一致したとみなす。ただし、一つのイベント点に対して正解数は一回のみカウントする。そして、注目点候補の数 n_p 、正解数 n_c 、イベント点の数 n_e に対して、以下を調べる。

- **Recall** : 捉えることができたイベント点の比率 n_c/n_e
- **Precision** : 妥当な注目点候補の比率 n_c/n_p

Recall が高いほど表情変化の取りこぼしが少ないといえる。また、Precision が高いほどノイズの少ない正確な注目点抽出ができていているといえる。

なお、変化スコアの前頭部データは、SDAR アルゴリズムの学習が始まって間もない時点のため、参考にならない。そのため、先頭30秒部分のデータは閾値 γ および各尺度の計算から除外した。

4.2 結果

図6は4つの単一特徴量それぞれの、図7は笑顔スケールの変化スコアのグラフである。赤線が変化スコア、薄青線は各特徴量の生の値、黒破線は閾値 γ を示す。赤点が注目点候補を示す。そのうち、二重丸は正解点を示す。黄縦線と水色縦破線はイベント点を示し、黄色が楽しそうになったとき、水色が楽しくなさそうになったときを示す。左軸は変化スコアの値を、右軸は特徴量の値、横軸は経過時間を示す。また、表2に各尺度の計算結果を示す。

n_p/n の値はいずれの場合も0.02以下、 n_o/n の値も高くても0.1程度であった。つまり、本手法により注目点を探すべきデータを10分の1以下に絞り込むことができたといえる。

また、笑顔スケールに対して ChangeFinder を適用した際に最も正解数が多く、Recall の値も最も高くなった。グラフを見ると、2分付近、3分付近、15分30秒付近のイベント点で注目点候補が得られたのは笑顔スケールからのみであった。さらに、イベント点以外の部分における注目点候補が最も少ない、つまり Precision が最も高かったのも笑顔スケールの場合であった。これらより、複合特徴量の注目点候補は、単一特徴量の注目点候補に比べ重要な瞬間を捉えている可能性が高いといえる。

各特徴量のグラフでは、イベント点のない6分～9分部分での注目点候補が見られる。これらの点では、笑顔は見せていないが、目を見開いたり、口元を結んだりといった、細かい顔の

(注1) : <https://github.com/shunsukeaihara/changefinder>

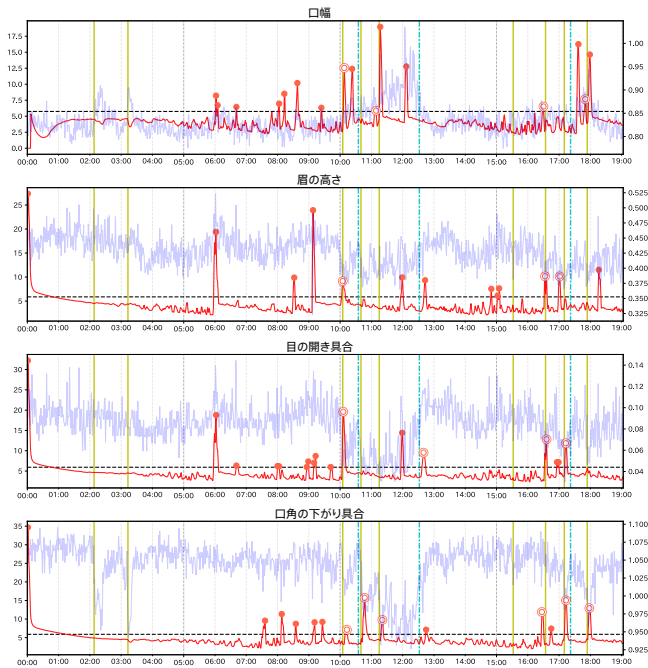


図6 各単一特徴量に対する変化スコア（テスト用データ）

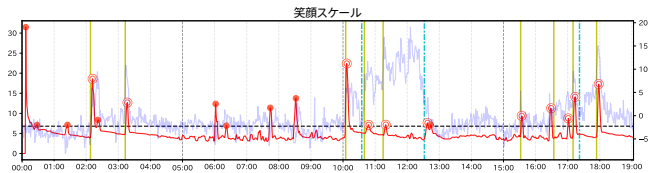


図7 笑顔スケールに対する変化スコア（テスト用データ）

表2 計算された各尺度（テスト用データ）

データの総数 $n=1114$, イベント点の数 $n_e=12$

特徴量種別	単一				複合
名称	口幅 h_7	眉の高さ a_1	目の開き 具合 a_2	口角の下がり 具合 a_3	笑顔 スケール
n_p	16	12	16	13	18
n_o	55	80	75	113	78
n_p/n	0.0144	0.0108	0.0144	0.0117	0.0162
n_o/n	0.0494	0.0718	0.0673	0.1014	0.0700
n_c	4	3	4	6	11
Recall	0.3333	0.2500	0.3333	0.5000	0.9167
Precision	0.2500	0.2500	0.2500	0.4615	0.6111

動きがあり、それらに ChangeFinder が反応していた。顔の動きを詳細にトラッキングしたデータを用いる以上、こういったノイズは多かれ少なかれ発生してしまう。ただ、笑顔スケールを用いた場合には、各特徴量ごとの細かい変動が少々抑えられるため、単一特徴量の場合と比べて重要でない時点での反応が抑えられ、結果として Precision の値が上がったと考えられる。

以上より、今回用いたテスト用データに対しては、特徴量データにおける ChangeFinder の利用と、複合特徴量の導入が有効であることが確認できた。

5. 実験

5.1 概要

本実験では、先行研究である FET の実験（2.1 を参照）で用

表3 各被験者の詳細

	年齢	性別	備考
被験者A	80	男性	表情が豊か。脳梗塞発症経験あり。
被験者B	99	女性	ご高齢であり、表情が顔に表れにくい。
被験者C	73	女性	表情が豊か。要支援。
被験者D	84	女性	表情が豊か。要支援だが会話に支障なし。
被験者E	80	女性	重度の認知症。多幸症の疑いあり。

※年齢は2018年1月30日（FET実験が行われた）時点のもの

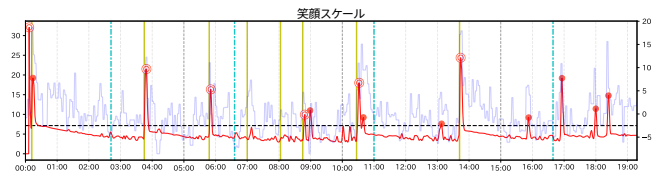


図8 笑顔スケールに対する変化スコア（被験者 A）

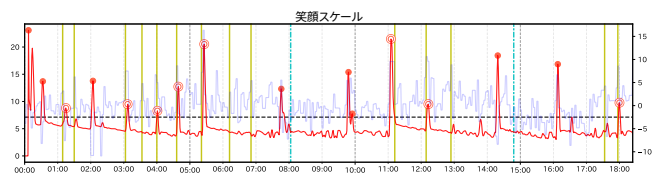


図9 笑顔スケールに対する変化スコア（被験者 C）

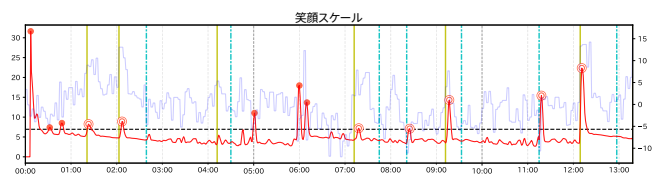


図10 笑顔スケールに対する変化スコア（被験者 D）

いられたケア最中の被験者 5 名の画像を用いて特徴量データを作成し、それらに対して注目点の検出実験を行う。そして、予備実験の場合と同様に分析が効率化したか、また妥当な注目点候補が得られたかどうかを調べる。各被験者の情報は表 3 に掲載してある。表情豊かな被験者 A, C, D を被験者群 1 と、表情の変化に乏しい被験者 B, E を被験者群 2 と分け、それぞれの群について考察を行った。なお、本研究は、神戸大学大学院システム情報学研究科の研究倫理審査委員会の許可を受けている（承認番号 R01-02）。

各被験者の顔画像は約 3 秒間隔で撮影されたものであったため、得られた特徴量データもそれに準じるものとなったが、実験条件を予備実験と合わせるため、予め各特徴量データは 1 秒間隔になるよう埋め合わせを行った。また、明らかな外れ値は除去した。予備実験と同様、先頭 30 秒部分のデータは閾値 γ , 各比率の計算から除外する。

5.2 被験者群 1 の結果

被験者群 1 についての分析結果を示す。ChangeFinder の変化スコアの算出に加えて、被験者群 1 は目視により表情の変化を確認することができるため、予備実験と同様にイベント点を調べたうえで、Recall, Precision の算出を行った。各被験者について、笑顔スケールの変化スコアのグラフを図 8～図 10 に示す。各グラフの凡例は予備実験と同一である。また、表 4～表 6 に Recall, Precision の計算結果を示す。

表 4 計算された各尺度（被験者 A）

データの総数 $n=1129$, イベント点の数 $n_e=11$

特微量種別	単一				複合
名称	口幅 h_7	眉の高さ a_1	目の開き 具合 a_2	口角の下がり 具合 a_3	笑顔 スケール
n_p	13	17	19	20	12
n_o	61	76	86	93	62
n_p/n	0.0115	0.0151	0.0168	0.0177	0.0106
n_o/n	0.0540	0.0673	0.0762	0.0824	0.0549
n_c	6	3	3	7	5
Recall	0.5455	0.2727	0.2727	0.6364	0.4545
Precision	0.4615	0.1765	0.1579	0.3500	0.4167

表 5 計算された各尺度（被験者 C）

データの総数 $n=1075$, イベント点の数 $n_e=16$

特微量種別	単一				複合
名称	口幅 h_7	眉の高さ a_1	目の開き 具合 a_2	口角の下がり 具合 a_3	笑顔 スケール
n_p	12	12	15	10	15
n_o	67	45	77	61	76
n_p/n	0.0112	0.0112	0.0140	0.0093	0.0140
n_o/n	0.0623	0.0419	0.0716	0.0567	0.0707
n_c	6	4	6	4	8
Recall	0.3750	0.2500	0.3750	0.2500	0.5000
Precision	0.5000	0.3333	0.4000	0.4000	0.5333

表 6 計算された各尺度（被験者 D）

データの総数 $n=769$, イベント点の数 $n_e=14$

特微量種別	単一				複合
名称	口幅 h_7	眉の高さ a_1	目の開き 具合 a_2	口角の下がり 具合 a_3	笑顔 スケール
n_p	10	8	7	6	12
n_o	42	38	36	61	54
n_p/n	0.0130	0.0104	0.0091	0.0078	0.0156
n_o/n	0.0546	0.0494	0.0468	0.0793	0.0702
n_c	5	2	2	3	7
Recall	0.3571	0.1429	0.1429	0.2143	0.5000
Precision	0.5000	0.2500	0.2857	0.5000	0.5833

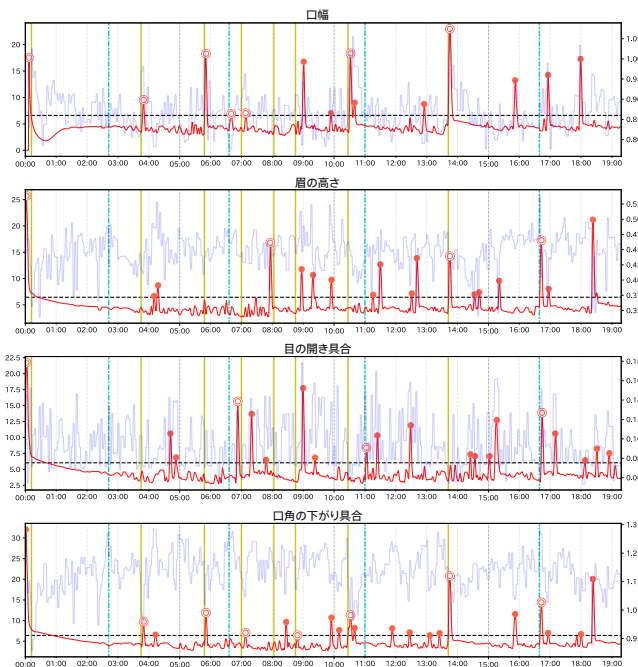


図 11 各単一特微量に対する変化スコア（被験者 A）

各被験者について、いずれの特微量を用いた場合でも n_o/n の値は高くても 0.08 程度であった。調査すべきデータを全体の 8% 以下に減らせていると考えると、分析の効率化については

達成できているといえる。本実験では元データに大きくパディングを施しているため、 n_p/n はあまり参考にならないが、値としてはすべて 0.02 以下であった。

被験者 C, D について、笑顔スケールを用いた場合に Recall および Precision の値が最も高くなった。この結果から、被験者 C, D については、予備実験と同様、複合特微量の利用が有効であったといえる。被験者 C, D の両者において、Recall の値はともに 0.5 であり、イベント点の半分が注目点候補として検出されたことになる。笑顔スケールのグラフを確認すると、イベント点の周辺でその値自体には上昇や下降が見られるものの、それらすべてが注目点候補として抽出されるわけではなかった。つまり、ChangeFinder は単純に局所的な値の変化に反応しているわけではないことが考察できる。

次に、被験者 A についてであるが、このケースでは笑顔スケールに比べて「口幅」が Recall, Precision とともに、「口角の下がり具合」が Recall において高い値となった。ここで、各単一特微量の変化スコアを図 11 に示す。笑顔スケールのグラフからは、4 分付近、10 分 30 秒付近、14 分付近など大きな特微量の変動が見られる部分に比べて、6 分 30 秒～7 分付近の小さな値の変動では変化スコアの反応がなく、イベント点も見逃していることが見て取れる。一方、「口幅」についてはこれを検出できている。単一特微量の結果は複合特微量の場合と比べて重要でない時点を注目点候補とみなす場合が多いものの、データの性質によっては複合特微量による結果を補完するために有効であると考えられる。

5.3 被験者群 2 の結果

被験者群 2 についての分析結果を示す。被験者群 2 に関しては、目視でイベント点を探すことが困難であったため、ChangeFinder の変化スコアの算出のみを行い、注目点候補として検出された点における表情を確認することで考察を行う。各被験者について、笑顔スケールの変化スコアを示すグラフを図 12, 図 13 に示す。また、 n_p/n , n_o/n の計算結果を表 7, 表 8 に示す。各グラフの凡例は予備実験や被験者群 1 の実験と同一である。

各被験者について、 n_o/n の値は常に 0.1 以下であり、被験者群 1 の場合と同様、本手法によりデータの大幅な絞り込みができたといえる。

被験者 B は表情筋の硬化により、明らかな表情の変化が見られなかった。笑顔スケールの値は変動し続けているものの、明確に上昇/下降したり、時系列の傾向が変わる時点はうまく確認できなかった。また、注目点候補付近の顔を調べると、主に口を結ぶ、口を開けるなどの口元の動きが見られた。これは、各単一特微量においても特に変動の傾向がみられない中で、複合特微量の係数が高い「口幅」や「口角下がり具合」などが強く反映され、ChangeFinder がそれらの変動に敏感に反応した結果と思われる。

次に、被験者 E について考察する。被験者 E は多幸症の症状からか、ケアの最初から終わりまで概ね楽しそうな様子を見せていた。笑顔スケールから抽出された注目点候補の付近の顔を調べると、表情が大きく動く瞬間を捉えることができた。ケア

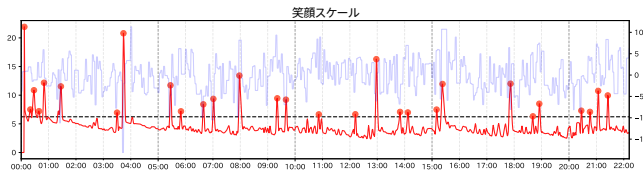


図 12 笑顔スケールに対する変化スコア（被験者 B）

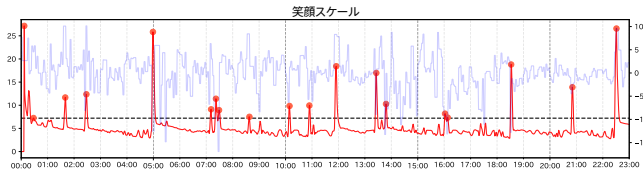


図 13 笑顔スケールに対する変化スコア（被験者 E）

表 7 計算された各尺度（被験者 B）

データの総数 $n=1303$

特微量種別	単一				複合
名称	口幅 h_7	眉の高さ a_1	目の開き 具合 a_2	口角の下がり 具合 a_3	笑顔 スケール
n_p	23	19	19	21	26
n_o	95	86	88	129	115
n_p/n	0.0177	0.0146	0.0146	0.0161	0.0200
n_o/n	0.0729	0.0660	0.0675	0.0990	0.0883

表 8 計算された各尺度（被験者 E）

データの総数 $n=1351$

特微量種別	単一				複合
名称	口幅 h_7	眉の高さ a_1	目の開き 具合 a_2	口角の下がり 具合 a_3	笑顔 スケール
n_p	14	16	20	14	17
n_o	72	70	100	71	82
n_p/n	0.0104	0.0118	0.0148	0.0104	0.0126
n_o/n	0.0533	0.0518	0.0740	0.0526	0.0607

前半～中盤（14 分以前）では、大きく笑った様子を見ることができた、また、後半部分、18 分 30 秒に注目点候補が存在するが、この地点以降、22 分 30 秒ぐらいまで笑顔含め顔の動きがあまり見られなくなった。ケアの記録によると、18 分 20 秒から 4 分間ほど本被験者は好きな音楽を聞かされていたとあり、これに真剣に聞き入っていた結果、表情が一時的に薄くなったと推測できる。

5.4 考 察

被験者群 1 については、抽出された注目点候補が、目視で確認したイベント点、つまり表情の変化タイミングとある程度合致することが確認できた。扱うデータによる精度の差や、取りこぼしも発生していたものの、膨大な時系列データの中から調査すべきデータが大きく絞られたことを考慮すると、今回用いた手法により、絶え間なく変化する特微量データの中から、注目点として適切な部分を効率的に拘えたといえ、注目点を効率よく見つけ出すヒントを得ることができるという点で、本手法は有用であると思われる。また、被験者群 1 の 3 名に共通して、変化スコアのグラフを細かく見たときに、閾値 γ を超えないものの、イベント点の周辺で、小さく変化スコアが上がっている箇所が存在する。閾値を下げればこのような箇所も検出されるようになるが、その分注目点候補の数自体も増えてしまうため、妥当な注目点を確認するコストは上がってしまう。その

ため、利用者側が自身のニーズや、分析にかけられる時間やコスト等を考慮した上で、閾値を定める必要があると思われる。

被験者群 2 については、そもそも顔の動きが小さい方に対しては、大きな表情の変化を示すような注目点候補を抽出することがうまくできなかった。しかし、注目点候補の前後において、実際に口を動かすなどの顔の動きは見られたため、感情の推定は困難にしろ、外部刺激に対する反応の大きさを探ることで、ケアへの興味を失わずにインタラクションができたかを推定すること等は可能であると思われる。一方で、多幸症とみられる方に対しては、大きく笑った瞬間や、真剣な表情になった瞬間をうまく注目点候補として捉えることができたため、本研究の手法により、一見ずっと楽しそうにしている方が、特にどういった場面で喜んでいたかを推測することができるといえる。

6. おわりに

本章では、表情センシングサービスを用いて得られる特微量データに対して ChangeFinder を用いた変化点検知手法を適用することで、注目点を効率的に探し出すことを提案した。さらに、分析をより高い精度で行うために、複合特微量の考案を行った。実際に分析を行うことにより、表情豊かな方について、妥当な注目点候補を効率的に抽出できることがわかった。また、表情の変化に乏しい方についても、顔が大きく動くタイミングを捉えることができ、ケアへの反応を調べる手がかりとなるような情報が得られた。

今後の課題として、本研究での手法の妥当性や精度の検証をより深く行うことが挙げられる。そのためには、より多くの、また、様々な背景を持つ被験者のケア映像を用いて分析・考察を行うが必要になる。また、分析結果を実際に介護に携わる方に確認してもらい、現場の目線からの意見を得たいと考えている。

謝辞 本研究の一部は JSPS 科研費 JP19H01138, JP17H00731, JP18H03242, JP18H03342, JP19K02973 の助成を受けている。

文 献

- [1] 日本経済再生本部, “未来投資戦略 2017 (案),” <http://www.kantei.go.jp/jp/singi/keizaisaisei/dai26/siryou.pdf>.
- [2] 平山孝輔, 佐伯幸郎, 中村匡秀, 安田清, “高齢者ケアの科学的な評価を目的としたユーザ定義の顔特微量の測定,” 電子情報通信学会技術研究報告, 第 119 巻, pp.17–22, March 2020. 東京, 日立製作所・中央研究所.
- [3] 佐古 嵐, 佐伯幸郎, 中村匡秀, 安田 清, “コグニティブコンピュータによる顔感情解析を用いたケア効果の定量的評価,” 電子情報通信学会技術報告書, 第 LOIS2017-89 巻, pp.105–110, March 2018. 沖縄・那覇市 IT 創造館.
- [4] S.Shanthi S.Saranya, R.Rajeshkumar, “A survey on anomaly detection for discovering emerging topics,” International Journal of Computer Science and Mobile Computing (IJCSMC), vol.3, no.10, pp.895–902, Oct. 2014.
- [5] 山西健司, データマイニングによる異常検知, 共立出版株式会社, 2016.
- [6] P. エクマン, W.V. フリーゼン, 工藤 力訳, 表情分析入門—表情に隠された意味をさぐる, 誠信書房, 1987.
- [7] 太田智美, 田村真理子, 有田真理子, 木曾奈央子, 佐伯行一, “表情分析—エクマンにより提唱されている表情の特徴との比較検討,” 滋賀医科大学看護学ジャーナル, vol.3, no.1, pp.20–24, 2005. <http://hdl.handle.net/10422/875>