

科学研究費助成事業 研究成果報告書

平成 27 年 5 月 25 日現在

機関番号：12605

研究種目：若手研究(B)

研究期間：2013～2014

課題番号：25870570

研究課題名(和文) 微表情計測に基づく心理的不安定状態推定システムの構築

研究課題名(英文) An estimation system of psychological conditions based on a micro facial expression recognition method

研究代表者

赤木 康宏 (Akagi, Yasuhiro)

東京農工大学・工学(系)研究科(研究院)・特任准教授

研究者番号：90451989

交付決定額(研究期間全体)：(直接経費) 2,500,000円

研究成果の概要(和文)：本研究では、顔の3次元形状データに基づき、微表情と呼ばれる現象を認識するための手法を研究・開発した。本研究では、顔形状を高精度に計測したデータを用いて、人間の内部状態と連動した、微表情を発見するための手法として、機械学習による顔の運動特徴の学習および認識手法を構築した。その結果、本研究で被験者に与えた課題に関して、高精度に顔の運動を記録したデータを用いることで、人間の内部状態と連動した特徴の現れる部位を定め、約75%の精度で、これを推定することができた。一方で、顔形状の計測における平均誤差が数mm、またフレームレートを20fpsに低下させた状態では、微表情の認識が困難であるという結果を得た。

研究成果の概要(英文)：A method for recognizing a micro expression which is a small motion appearing on a face by using a high density and high frame-rate 3D reconstruction method are developed. In order to verify the set of methods to identify the site of micro expression and its accuracy, as a simple recognition problem, vocalizing unknown words was assigned to be further analyzed mechanically. As a result, by using highly accurate data of the facial movement, we were able to identify feature parts that matches with the human mental state, and able to estimate at about 75% accuracy. On the other hand, if the accuracy of the face shape measurement was decreased by few 10mm and lower the frame rate to 20fps, recognition of the micro expression became troublesome.

研究分野：コンピュータビジョン

キーワード：表情認識 微表情 トラッキング 3次元点群

1. 研究開始当初の背景

計算機による人間の顔認識はコンピュータビジョンの分野において最も活発に研究されている対象の一つである。特に表情認識技術はカメラや携帯情報通信機器等の様々な製品で利用されており、海外では表情認識技術を専門に扱うベンチャー企業等も生まれている。また、医療福祉分野においても、入院患者の見守りや感情を理解できない障害をもつ人への支援システム等の、幅広い応用が期待されている。このような一般的な表情を認識する技術は大変有用であるが、さらに人間の表情には微表情と呼ばれる無意識的な表情動作が含まれており、これが嘘などの内面的な心理状態の変化を表す指標として有用であるという研究がある。微表情を認識する技術は空港での審査、各種の窓口業務等に有用な新たな対人セキュリティ技術として注目されている。しかし、人間にとって微表情の認識には特別な訓練を必要とするほか、見逃しや誤認の可能性が高いという問題がある。また、従来の表情認識技術では微表情のような詳細な表情の差異を検出することが難しいという問題もある。そこで、計算機による表情認識技術の精度を高め、微表情の自動認識を実現する技術が求められている。

2. 研究の目的

人間の顔に表れる微小な表情変化（微表情）は、嘘や動揺などの内面的な心理状態の変化を知る手がかりとして有用であることが知られているが、一般的な人間にとっては認識が困難であるという課題がある。本研究では、表情の中から個人に固有の特徴をもつ微表情の発現箇所を自動的に検出する手法を実現し、表情から人の心理状態の変化を推定するシステムを構築する。人間の表情認識能力では発見が困難な対象者の微表情を定量的かつもろさずに検出する技術が実現されれば、表情観察に基づく新たなセキュリティ技術の実現に大きく貢献できる。

3. 研究の方法

(1) 研究方法の概要

(1-a) 識別しようとする人間の状態および実験方法

本研究では、被験者に動揺を引き起こさせ、人間の内部状態を変化させながら顔を計測する必要があるため、以下の課題を被験者に課す。被験者には3秒に1語の頻度でディスプレイに提示される単語を一定のリズムで読み上げる課題を与え、課題終了後に、出題された単語が既知であるか、または、未知であるかを回答する(図1)。

この方法を選んだ理由は、状態が2値(既知/未知)の2種類のみである単純な識別問題であり、かつ被験者が確実に自身の内部状態を実験後に回答できる方法だからである。出題する単語セットには、電総研492語^{¥cit}

を用いる。また、単語を提示する際には漢字の読み仮名が分からないという状態を避けるために、視線を大きく動かさずに視認できる位置に振り仮名を提示する。次節では、この課題回答時に被験者の顔形状を計測する方法を説明する。

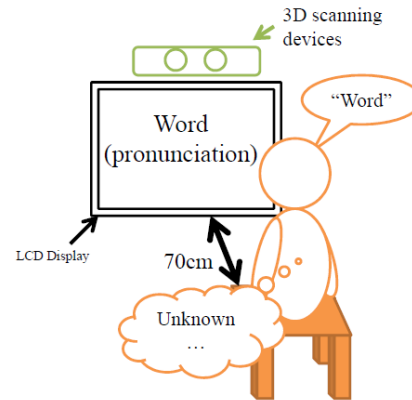
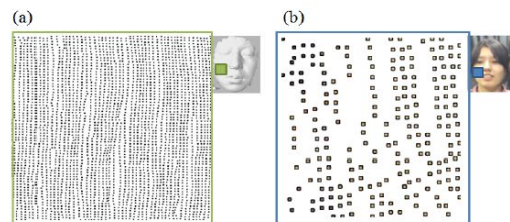


図1 単語読み上げ実験による微表情の計測方法



(a)高精度計測手法 (b)Kinect
図2 計測方法による取得点群の違い

表1 計測方法による精度・品質の違い

	計測方法	
	高精度	Kinect
解像度(mm)	0.3	2
計測点数	30,000	5,000
FPS	200	20

(1-b) 顔形状の計測方法

実験では、課題回答中の被験者の顔の3次元形状を正面から計測する。計測には、高解像度かつ高フレームレートでの撮影が可能な方法による計測(以下、高精度計測)と、Kinectによる簡易な計測を行う。前者の方法により取得した顔の変形情報は、被験者の顔に単語の未知性による変化が生じているのか、いないのかを判別する目的、および、変化が生じている領域を特定する目的で利用する。後者は、簡易な計測システムにおいて、本研究で扱う識別問題に対応可能なかを評価する目的で使用する。表1および図2、両者の手法による計測精度や計測点数の違い等を示す。計測点数は、顔を構成する点のみを抽出した後の数である。また、各単語の顔の動きデータを抽出するために、計測結果から回答者が単語を発音しているフレーム

のみを手作業で切り出す。
(1-c)特徴的な形状変化が生じている部位の特定手法

前述した Ecmann らの研究によれば、顔は部位により能動的に制御しやすい部位と、そうではない部位がある。一例をあげれば、口周辺の領域は意図的に形状を制御しやすく、眉間領域では制御が困難であり、真実の感情が表出しやすいという性質があると報告している。このように、顔の運動は部位によりその特徴が異なる可能性があるため、顔を、口元、鼻周辺、眉間周辺の3領域に分け、どの領域が本研究で扱う単語の認知に関する特徴をよく保持しているのかを調べる。具体的な方法は、各領域に、発話開始フレームの顔形状から無作為に点を選び、その点の位置での奥行き方向の変動を記録する。そして、この点の変動を特徴ベクトルとして、教師付学習におけるラベルを単語の未知性として機械学習を行い、部位毎の認識率の違いを比較する。その結果から、識別率の高い部位には、単語の未知性に関連する運動が表れているとみなす。本手法の詳細については(3)で述べる。

(2) Kinect による単語の未知性認識

(1)で説明した方法のように、高精度センサーを用いて多数の特徴点を用いれば、顔の僅かな変化を計測できるので、顔の微細な動きの認識にとって有利ではあるものの、撮影には特殊な機器を必要とし、コストがかかりすぎるといった問題もある。一方で、計測精度が低く、低フレームレートのセンサーを用いた場合には、微小な運動の変化が発見できない可能性がある。この問題に対して、(1)で述べた手法により特徴的な変形が生じている部位のみを限定できるので、計測範囲を絞り込んだ状態での、認識率の変化を調査する。この実験のために、より安価な3次元形状計測装置である Kinect を用いる。Kinect を用いる際の計測方法の工夫については、(3)で詳しく説明する。

(2-a) 高速撮影による発話特徴の検討

本章では、高精度かつ高速に撮影した顔の3次元形状データを用いて、単語発話時の、単語の認知性による形状の差異が生じているのか、また、その差異がどの部位に特徴的に表れているのか実験的に調べる。

(2-b) 高精度な顔運動情報の計測

本節で行う調査では、被験者が話す単語の違いによる影響を受けない、単語の未知性のみ由来する運動を検出する必要がある。また、微表情は被験者に固有の、“くせ”のような運動である可能性もあるので、被験者毎に個別に特徴を抽出した場合と、被験者を区別せずに特徴抽出を行った場合の2通りの調査を行う。この調査のために、複数の話者に、複数回、単語を読み上げてもらい、その時の

顔形状の計測を行う。単語には、被験者にとって既知の語7種類、未知の語7種類を用いる。また、その単語が既知/未知であるかどうかは、計測後(発生後)に被験者から申告してもらう。

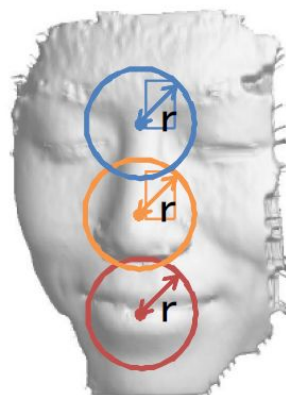


図3 認識に用いる顔の部位

(2-c) 比較調査に用いる顔の部位および運動データの学習

本研究では、顔を口元、鼻、眉間の3種の部位に分け調査を行う(図3)。この各領域から、7種の語毎にランダムに200点を抽出し、各点の100フレーム(0.5秒)間の奥行き方向変化を特徴ベクトルとして学習する。合計すると、1400地点、各100次元の特徴ベクトルが、既知語、未知語のそれぞれに与えられる。学習には、多次元かつ多数の教師データを与えた際にも効率よく学習を行うことのできる、2クラス識別器である Local Deep Kernel Learning (LDKL)を用いる。

(2-d) 微表情の個人性を考慮した場合の識別実験

まず、微表情の運動には個人性があり、他者の運動情報が外乱となる可能性を考慮し、ある被験者から収集した運動情報のみから学習を行い、学習に用いなかった、その被験者の発する単語の既知/未知性を判別する。3名の被験者に関する認識率を表2に示す。

いずれの被験者においても、鼻周辺領域の運動特徴を用いた場合が、最も認識率が高いという結果となった。特に、単語の既知性においては75%以上の認識率を示しており、この部位に単語の未知性に関する特徴が表れている可能性が示された。また、口元領域の認識率は54~65%を示しているが、今回の識別が2クラス分類であることを考慮すると、単語の未知性から生じる運動が口元領域から発見できたとは言えない。眉間領域に関しても同様に、単語の未知性を表す運動は発見できなかった。

(2-e) 微表情の個人性を考慮しない場合の識別実験

次に、被験者全員の動き特徴を区別なく学習データとして用いた場合の識別率を調査する。評価は、各個人毎に、学習に用いていな

い単語セットを用いて行う。結果を表3に示す。結果の識別率は、ここの数値の増減はあるものの、全体的には(2-d)節で行った、被験者自身のデータを学習に用いた場合に比べ低下していると言える。この結果は、微表情にはある程度の個人差があることを示している。このことは、Ekman らによる研究で述べられているように、動揺などを隠す能力には個人差があることに起因すると考えられる。

表2 同一人物の場合の識別率(%)

	被験者 A		被験者 B	
	既知	未知	既知	未知
口	62	65	54	65
鼻	79	77	77	46
眉間	44	5	72	34
	被験者 C		被験者 D	
	既知	未知	既知	既知
口	66	45	77	45
鼻	75	78	70	69
眉間	41	19	55	18
	被験者 E		被験者 F	
	既知	既知	未知	未知
口	70	13	29	62
鼻	78	57	73	74
眉間	85	51	71	21
	合計			
	既知	未知		
口	60	49		
鼻	75	67		
眉間	61	34		

表3 別人の場合の識別率(%)

	被験者 A		被験者 B	
	既知	既知	未知	未知
口	45	54	69	62
鼻	72	45	62	73
眉間	58	4	50	61
	被験者 C		被験者 D	
	既知	既知	未知	既知
口	52	75	36	54
鼻	73	79	75	76
眉間	53	15	45	21
	被験者 E		被験者 F	
	既知	既知	既知	未知
口	53	70	24	68
鼻	87	31	73	86
眉間	84	76	43	30
	合計			
	既知	未知		
口	46	64		
鼻	74	65		
眉間	55	35		

表4 低解像度・高ノイズな計測データを用いた場合の識別率(%)

被験者	口		
	既知	未知	合計
A	87	43	67
B	67	33	53
C	80	40	53
D	36	55	44
E	56	50	53
F	72	0	53
G	25	73	60

被験者	鼻		
	既知	未知	合計
A	64	78	70
B	56	60	57
C	100	50	80
D	67	67	67
E	33	67	47
F	88	29	60
G	75	54	60

被験者	眉間		
	既知	未知	合計
A	62	0	33
B	88	14	53
C	60	50	53
D	46	42	44
E	25	29	27
F	83	22	47
G	20	70	53

(3) Kinect を用いた簡易システムによる認識

(2)の結果から、鼻周辺領域を識別対象とした際には、単語の認知性の識別に有効な情報が得られるということがわかった。この知識に基づき、(1)で説明した計測方法よりもノイズが多く、フレームレートも低い計測方法である、Kinect を用いた際の認識率の変化を調査する。

(3-a) 計測データと前処理

Kinect からは、計測対象の3次元形状の他に、顔の特徴点として64点の特徴点を得ることができる(図4 黒および赤点)。その中から、鼻の輪郭形状を構成する8点を選び出し(図 赤点)、この点の奥行き方向の変動を特徴ベクトルとして学習および認識に用いる。ただし、Kinect にはノイズや計測誤差が多く含まれる可能性があるため、事前に各特徴点の動き情報を周波数解析し、10Hz以上の信号成分を除去することで、点の時系列上の変動を平滑化する処理を適用した。

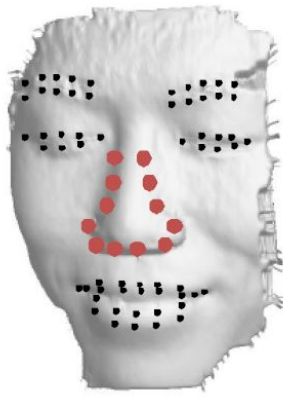


図4 認識に用いる顔の特徴点

(3-b) Kinect を用いて取得した運動特徴に基づく認識

(2)で行った実験と同様に、同一話者のデータを用いて学習と認識を行う。1人の被験者から収集したデータは、60単語分の発話データから、鼻を構成する特徴点12点の動き、40フレーム分(2.0秒)を取得した。また、学習と評価にはそれぞれ異なる単語セットを用いた。結果を表4に示す。この結果から、一部の被験者からは単語の未知性に関する情報を認識できているものの、全く特徴を捉えることのできない被験者も存在することがわかる。4.4節でも述べたように、微表情の表れにくい被験者の場合、計測精度を落とすことで特徴を捉えることが困難になることが原因と考えられる。また、口および眉間周辺領域の運動特徴ベクトルのみを用いた場合の結果も表4に示す。高精度データに基づく識別結果と同様に、鼻周辺領域に比べ、識別率は低下している。結果の中で、被験者AおよびCの口周辺データ、被験者Bの眉間周辺データを用いた結果に、80%以上の精度で問題を認識している結果がある。これは、本研究で扱う問題は2クラスであるので、入力データの多くを一方のクラスに認識してしまうような識別器が構築されると、一方のクラスのみが高い認識率が現れる。これらの結果も、他方のクラスの認識率は43%以下と低く、そのような現象が発生したものと考えられる。

4. 研究成果

本研究では、高精度な形状計測手法を用いて計測した顔の形状データに基づき、顔の一部に短時間かつ微差な動きとして表れるとされている、微表情と呼ばれる現象を認識するための手法を研究・開発した。微表情は、動揺等の人間の内部状態の変化に応じて、表情の僅かな差異として出現するという報告があり、これを認識することができれば、機械による人間理解に有用な情報として活用できる。本研究では、この微表情の表れる部位を実験的に特定し、簡易的な、顔形状情報を用いた人間の内部状態の推定試験を行い、その結果を検証した。具体的には、顔形状を

高精度かつ高密度に計測した結果を用いて、人間の内部状態と連動した、微小かつ特定の部位に生じる運動を発見するための手法として、機械学習による顔の運動特徴の学習および認識手法を構築した。その結果、本研究で被験者に与えた課題に関して、高精度に顔の運動を記録したデータを用いることで、人間の内部状態と連動した特徴の現れる部位を定め、約75%の精度で、これを推定することができた。一方で、顔形状の計測における平均誤差が2mm、またフレームレートを20fpsに低下させた状態では、微表情の認識が困難であるという結果を得た。この結果から、微表情の実用的な認識には、顔形状の計測誤差1mm以下、100fps以上の高フレームレートが必要であるという知見を得た。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 1件)

1. Yasuhiro Akagi, Ryo Furukawa, Ryusuke Sagawa, Koichi Ogawara, Hiroshi Kawasaki, "A facial motion tracking and transfer method based on a key point detection," 21th WSCG International Conference on Computer Graphics, Visualization and Computer Vision 2013, Vol.21, No.E73, pp.1-8, 2013. 査読有
<http://hdl.handle.net/11025/10656>

[学会発表](計 2件)

1. Yasuhiro Akagi, Hiroshi Kawasaki, "A method of micro facial expression recognition based on dense facial motion data," 22nd WSCG International Conference on Computer Graphics, Visualization and Computer Vision, pp. 39-44, Plzen(Czech), June, 2014. 査読有
2. 赤木康宏, 小原由華, 森永寛紀, 川崎洋, "顔の密な3次元形状の微小変化を利用した認識手法," 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, Vol.113, No.493 pp.203-208, 早稲田大学(東京), 3月, 2014. 査読無

[図書](計 0件)

[産業財産権]

出願状況(計 0件)

取得状況(計 0件)

[その他]

なし

6. 研究組織

(1) 研究代表者

赤木 康宏 (AKAGI, Yasuhiro)

東京農工大学・大学院工学府・特任准教授

研究者番号：90451989

(2) 研究分担者

なし

(3) 連携研究者

なし