

会話ゲームにおける微表情ベンチマークデータの構築

藤田 裕文¹ 松川 徹¹ 鈴木 英之進¹

概要：微表情とは、隠そうとしている真実の感情が露出し、短時間で無意識に生じる表情である。さらに微表情は、本人には露出しているかが分からず、相手の隠している感情を知る手がかりとして効果的である。微表情研究は、療養や交渉、尋問などの様々な場面での応用が考えられる。もっとも微表情は、短い時間で弱く発生するため、これを分析することは難しい。近年の微表情分析研究では、自然に生じた微表情を収集したベンチマークデータの重要性が広く認識されている。ただしよく用いられているベンチマークデータは、実生活の環境とはかけ離れた条件で収集されており、応用上問題がある。われわれは、会話中の嘘を見抜く会話ゲームという、従来よりも実生活の環境に近い設定下で、微表情ベンチマークデータを構築した。

Construction of micro-expression benchmark data in conversation game

HIROFUMI FUJITA¹ TETSU MATSUKAWA¹ EINOSHIN SUZUKI¹

Abstract: Micro-expressions (MEs) are brief, involuntary facial expressions which reveal genuine emotions that people do not intend to show. Moreover, MEs might be unaware to the actor, thus may provide effective clues for getting to know real emotions. Studying MEs is applicable to healing, negotiations, and interrogations. However, analyzing MEs is a hard task because they are generated rapidly and with low intensity. In recent studies on ME analysis, the importance of benchmark data with spontaneous MEs is widely recognized. Unfortunately frequently used benchmark data consist of MEs collected in environments that are far from our daily lives, causing problems in applications. In this work, we built spontaneous MEs benchmark data under a more natural setting of conversation game, which is done to detect lies.

1. はじめに

表情は、人々が感情を相手に伝えるのに重要な役割を担っている。人に感情が発生すると、その感情を表した表情が表れる。しかし、人は感情をごまかしたり、隠そうとしたりする場合があります。その場合、表情から、感情を知ることが難しくなる。その隠そうとしている感情を知る方法として、近年注目を浴びているのが微表情である。

微表情とは、人々が隠そうとする真実の感情を表しており、短時間で、無意識で起こる顔表情である [1]。1966 年、Haggard と Isaacs が、心理療法における話し合いのフィルムをスローモーション再生で調べている間に、微表情の存在を初めて発見した [2]。その後、1969 年、Ekman らが、

自殺を図ろうとしていた精神患者のカウンセリング中の動画を調べている際に、微表情を発見した [1]。その患者は、動画を通して、幸せそうに見えたが、2 フレーム (1/12s) だけ、苦悩の感情を顔に漏らしていた。それから三十年以上の間、微表情に関する論文はとても少なかった [3]。

しかしながら、最近になって、微表情が注目され始めている。その理由は、嘘発見や、過激な行動をとる人を事前に知ることの手がかりに微表情がなりうるからである [1]。心理学の分野では、Ekman が、微表情を認識する能力をあげるため、Micro Expression Training Tool (METT) を作成し [4]、警官の微表情検知の能力を向上する手助けとした。しかし、微表情は、程度が弱く、短い顔表情であるため、人間の肉眼で検出、さらには認識するのは限界がある。

一方、コンピュータビジョンの分野では、微表情を自動で認識するシステムを作ることを試みられているが、まだ、初期段階の新しい研究分野であり、十分な研究がされ

¹ 九州大学, 福岡県福岡市西区元岡 744 番地
Kyushu University, 744, Motooka, Nishi-ku, Fukuoka, 819-0395, JAPAN

ていない。微表情分析を発展させるためには、良質な微表情ベンチマークデータの作成が急務である。現在、意図的に作った微表情からなるベンチマークデータと自然に生じた微表情からなるベンチマークデータがある。しかし、公開されている微表情ベンチマークデータは数が少なく、収集された微表情が意図的ではなく自然な表情であるベンチマークデータは、SMIC[5], CASME[6], CASMEII[7] だけである。さらに、それらに収集されているのは、動画を見ている際の微表情であり、表情を露出しない指示下で撮影するなど、実生活の環境とはかけ離れている。本論文では、より自然な微表情ベンチマークデータを作成するために、このような指示をせず、ゲーム形式をとる会話中の微表情を集めた。

本論文の構成を述べる。まず2章で既存の微表情ベンチマークデータの説明を行う。次に、3章で新しい微表情ベンチマークデータをその作成法を含めて説明する。4章でまとめと次への展開を述べる。

2. 従来の微表情ベンチマークデータ

微表情ベンチマークデータは、微表情分析に必要不可欠のものである [5], [7], [8], [9], [10]。微表情分析の手法を考え、実験をすることは、ベンチマークデータなしでは行うことができず、ベンチマークデータ作成は、微表情分析において、基礎となる部分である。ベンチマークデータを作成するには、表情の撮影に加え2つの作業が必要である。1つ目は、微表情が露出する環境を作ることである。微表情が露出する環境とは、本当の感情を隠そうとし、顔の動きを抑えようとするときである。Ekmanによると、人が感情を隠そうとするとき、その本当の感情が一瞬、漏れてしまい、その漏れている表情を微表情と言う [1]。さらに、嘘をつく際に、微表情がでることも知られている [1], [11]。したがってベンチマークデータの作成を行う時でさえ、すべての人が微表情を出すとは限らない。2つ目は、微表情を撮影した動画の中から微表情を探すことである。微表情は、1/2秒以下の時間で現れ [12], [13]、程度も弱いため、スロー再生で探す必要があり、時間が非常にかかってしまう。

過去の研究では、Polikovsky らが、意図的な微表情のベンチマークデータを作成した [10]。顔の筋肉を動かす程度を小さくしつつ、7つの基本感情を表し、その後、できるだけ早く無表情に戻すように指示をして、作られたものである。微表情が非常に短い時間で起こることから、高速カメラで撮影をし、200fpsであった。このベンチマークデータには、自然ではなく、意図的に作った微表情が収集されている。

SMIC[5], CASME[6], CASMEII[7] は、自然に起こった微表情を収集したベンチマークデータである。3つのベンチマークデータの作成方法で共通していることが2つある。1つ目は、表情を抑えるように指示した後、感情を起

表 1 従来の微表情ベンチマークデータ

	Polikovsky [10]	SMIC [5]	CASME [6]	CASMEII [7]
サンプル数	42	164	195	247
人数	10	16	19	26
fps	200	100	60	200
感情クラス数	7	3	8	5
微表情の種類	意図的	自然	自然	自然

表 2 質問内容

質問内容
どんなアルバイトをしていましたか？
業務内容は何ですか？
何人ぐらい働いていましたか？
あなたの好きな科目、嫌いな科目は何ですか？

こさせる動画を見せることで、微表情を発生させている点である。2つ目は、見せた動画の種類や撮影後のアンケート、Action Units (AUs) [14] によって、その微表情がどの感情を表しているかのクラスを決めている点である。表 1 に、ベンチマークデータの基本情報をまとめた。

3. ベンチマークデータの構築

3.1 概要

微表情分析の応用として、カウンセリングや尋問などを行うときに、相手の隠している感情を推定することが挙げられる。これらは、会話をしながら行うもので、動画を見ながら行うものではない。よって、会話中に起こる微表情を収集したベンチマークデータを構築することにした。われわれは、微表情を効率よく収集する手段として、質問者が回答者の2つの回答のうちどちらが嘘であるかを当てる、会話ゲームを提案する。本論文では会話ゲームを撮影し、回答者の微表情から構成されるベンチマークデータを作成する。

3.2 データ収集設定

3.2.1 1対1の質疑応答

当初の会話ゲームは質問者1名と回答者1名の間で行われた。質問者は1つあるいは複数の回答を要する質問を行う。回答者は、前者に対しては本当の回答と嘘の回答を答え、後者に対しては本当と嘘を混在した回答を答える。質問者は嘘の回答を予想して回答者に伝え、予想の正誤によって勝敗が決まる。微表情は人が隠そうとしている感情を表しており [3]、感情を隠そうとしない場合は通常表情が出ると期待できる。会話ゲームにおいて回答者は嘘を隠そうとするため、この設定は微表情を効率的に収集できると期待した。質問内容の例を表2に示す。

もっともこの方式では、人間関係が1種類しかなくゲームが単純であるため、微表情は期待したほど出なかった。質問者の手掛かりは回答内容とその際の表情・身振りだけ

であり、回答者は自分だけに集中して嘘を隠せばよい。この方式では回答者が圧倒的に有利であり微表情を滅多に出さない。

3.2.2 1対2の会話ゲーム

そこで出題者を設けて回答者を2名に増やし、質問者がこの2名に交互に質問する形式に変更した。まず出題者が、質問者だけに2個のタイトル（例：海、学校）を伝え、回答者たちだけにそのうちの1個のタイトル（例：海）と関連する画像を1枚（例：夏のビーチの画像）を見せる。画像を見せるのは、回答者たちと質問者に共通の知識を持たせ、その後の会話を成立させるためである。その後3分間、回答者たちは画像を注意深く見て、質問を予想して小声で相談しあう。その間質問者は離れた場所にいるため、回答者たちの相談は分からない*1。その後2分間半、質問者はまず片方のタイトル（例：学校）を述べ、回答者2人対して交互に質問して（例：何階建てでしたか？）回答を得る（例：3階建てです）。回答者たちは、事前に知らされていないタイトルについては、事前の相談と関連画像なしにお互いの回答が矛盾しないように注意する必要がある、前節の設定に比較して微表情をより出しやすい。次の2分間半、質問者はもう片方のタイトルについて同じことを行う。会話ゲームにおいて回答者が極度に有利にならないように、事前に見せられた画像についての嘘は禁止した。最後に質問者は、回答者が2つのタイトルのうちどちらを見ていないかを予想し述べる。予想が当たれば質問者の勝ち、そうでなければ回答者の勝ちである。

この実験において、回答者の難しい点は、2つあり、1つ目は、回答者は1つの画像は見えていないことから、それについては、その場で、質問に対して、考えて、答えなければならなかった点である。2つ目は、質問を回答者2人に対して、交互に行い、これにより、質問者は、2人の内、1人が答えたことに関連した質問をもう1人にすることができ、回答者は、より瞬時に判断しなければならない点である。この2つの難しい点を作ることによって、回答者に認知的負担を増させ、より嘘の手がかりを強調させることが期待でき[15]、さらに、微表情が出やすくなることが期待できる。また、アイコンタクトをとり続けることを指示し、さらに、嘘の手がかりが強調されることが期待できる[15]。

タイトルが「陸上競技場」であったときの質問と応答を表3に示す。

3.3 データ収集

大学生か大学院生である男性14名と女性4名の計18名からデータを収集した。使用したカメラはGoPro HERO4 BLACKであり、設定は解像度1920×1080、フレームレートは120fpsであり、顔とカメラの距離は50-80cm程度で

*1 会話ゲームにおいて勝敗が偏っていなかったため、実際質問者は回答者たちの相談が分からなかったことが裏付けられた。

表3 質問内容

話した人	話した内容
質問者	地面はどのような感じでしたか？
回答者1	よくある、トラックの赤茶色っぽい色でした。
質問者	地面の色は茶色だけでしたか？
回答者2	近い系統の色が何種類もありました。
質問者	大会は行われていましたか？
回答者1	練習が行われていました。
質問者	どんなことをしていましたか？
回答者2	走ってはならず、準備運動をしていた。
質問者	走る競技以外の競技はしていましたか？
回答者1	準備運動段階であったが、していた。
質問者	人数はどのくらいでしたか？
回答者2	30-40人くらいいました。
質問者	収容人数はどのくらいですか？
回答者1	300-400人くらいです。
質問者	練習していた人はプロですか？
回答者1	学生でした。

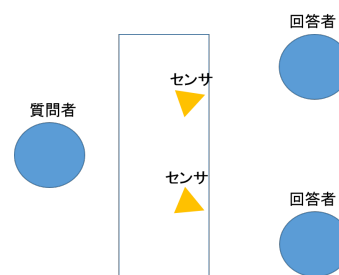


図1 データ収集の配置

あった。実験実施日は2015年12月10日、11日、12日であり、実施時間は、13:00-20:00である。合計165個の微表情を撮影した。カメラと被験者の配置を図1に示す。

3.2.2節で説明した会話ゲームを2人1組でそれぞれ3セット実施した。このゲームは、質問者の質問によって、難易度が変わる。難易度をあげるために、質問者により瞬時的な判断が必要な質問を行ってほしいと伝えた。さらに、作業の効率化を図るため、画像のタイトルは6個とし、データ収集前に質問者に見せた。質問者は、それぞれのタイトルについて基本となる質問はあらかじめ考えておき、会話ゲーム中に質問が思いつかないときにそれらを利用した。この措置により、質問の難易度が安定し、会話ゲームの質が参加者に過度に依存することなく高まった。

3.4 ラベル付け

データ収集で得られた動画から微表情の部分を取り出し、ベンチマークデータを作るために、2つの作業を行った。第一に、動画をスロー再生し、注意深く見て、微表情が出ている時点を探し、その部分を取り出した。その際の基準として、5個のAction Units[14](AU1, 2, 4, 6, 12)を用いた。具体的には、これらのAction Unitsに定められ



図 2 onset frame 図 3 apex frame 図 4 offset frame

ている顔の動作が生じている地点の動画の切り出しを行った。第二の作業は、取り出された動画を画像に変換し、その画像列から、微表情の始まりのフレーム (onset frame)、微表情の程度が最も強いフレーム (apex frame)、微表情の終わりのフレーム (offset frame) を見つけることである。取り出した微表情の画像列に対して、どこからどこまでが微表情かを定める基準は、微表情が出た顔の部分の動き始めを onset frame、動きの程度が最も大きい地点を apex frame、動きが終わった地点を offset frame とした。この際に、微表情の時間の定義が 1/2 秒以下であることから、onset frame から offset frame のフレーム数が 60 枚より大きいものも除外した。

図 2, 3, 4 に、それぞれ onset frame, apex frame, offset frame とラベル付けされた画像の例を示す。

4. おわりに

本論文では、今まで存在しなかった、会話中の微表情ベンチマークデータを構築した。従来のデータ収集法は、実験参加者に動画を見せ、その顔を撮影するというものであったが、これには実生活環境とはかけ離れているという問題があった。われわれは、新しいデータ収集の手法として会話ゲームを考案した。

次の課題として、微表情特定がある。従来のベンチマークデータの微表情は、動画を見ている際の微表情であるので、頭の位置が安定しており、微表情特定が容易だったと思われる。我々の作成したベンチマークデータの微表情は、頭の動きや口の動きなどがあり、特定が難しい。今後は、頭の動きや口の動きへ対応した、微表情特定手法を開発したい。

謝辞

本研究を進めるにあたり、データ収集にご協力いただきました鈴木研究室の皆様、データ収集参加者の皆様に心から感謝いたします。本研究の一部は JSPS 科研費 25280085, 15K12100 の助成を受けています。

参考文献

- [1] Ekman, P. and Friesen, W. V.: Nonverbal Leakage and Clues to Deception, *Psychiatry*, Vol. 32, No. 1, pp. 88–106 (1969).
- [2] Haggard, E. A. and Isaacs, K. S.: Micromomentary Facial Expressions as Indicators of Ego Mechanisms in Psychotherapy, *Methods of research in psychotherapy*,

- Springer, pp. 154–165 (1966).
- [3] Ekman, P.: Darwin, Deception, and Facial Expression, *Annals of the New York Academy of Sciences*, Vol. 1000, No. 1, pp. 205–221 (2003).
- [4] Ekman, P.: MicroExpression training tool (METT), *San Francisco: University of California* (2002).
- [5] Pfister, T., Li, X., Zhao, G. and Pietikäinen, M.: Recognising Spontaneous Facial Micro-expressions, *IEEE International Conference on Computer Vision (ICCV)*, 2011, IEEE, pp. 1449–1456 (2011).
- [6] Yan, W.-J., Wu, Q., Liu, Y.-J., Wang, S.-J. and Fu, X.: CASME Database: a Dataset of Spontaneous Micro-expressions Collected from Neutralized Faces, *IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, 2013 10th, IEEE, pp. 1–7 (2013).
- [7] Yan, W.-J., Li, X., Wang, S.-J., Zhao, G., Liu, Y.-J., Chen, Y.-H. and Fu, X.: CASME II: An Improved Spontaneous Micro-expression Database and the Baseline Evaluation, *PloS One*, Vol. 9, No. 1 (2014).
- [8] Li, X., Hong, X., Moilanen, A., Huang, X., Pfister, T., Zhao, G. and Pietikäinen, M.: Reading Hidden Emotions: Spontaneous Micro-expression Spotting and Recognition, *arXiv preprint arXiv:1511.00423* (2015).
- [9] Moilanen, A., Zhao, G. and Pietikainen, M.: Spotting Rapid Facial Movements from Videos Using Appearance-based Feature Difference Analysis, *2014 22nd International Conference on Pattern Recognition (ICPR)*, IEEE, pp. 1722–1727 (2014).
- [10] Polikovsky, S., Kameda, Y. and Ohta, Y.: Facial Micro-expressions Recognition Using High Speed Camera and 3D-gradient Descriptor, *3rd International Conference on Crime Detection and Prevention (ICDP 2009)*, IET, pp. 1–6 (2009).
- [11] Porter, S. and Ten, B. L.: Reading Between the Lies Identifying Concealed and Falsified Emotions in Universal Facial Expressions, *Psychological Science*, Vol. 19, No. 5, pp. 508–514 (2008).
- [12] Yan, W.-J., Wu, Q., Liang, J., Chen, Y.-H. and Fu, X.: How Fast are the Leaked Facial Expressions: The Duration of Micro-expressions, *Journal of Nonverbal Behavior*, Vol. 37, No. 4, pp. 217–230 (2013).
- [13] Matsumoto, D. and Hwang, H. S.: Evidence for Training the Ability to Read Microexpressions of Emotion, *Motivation and Emotion*, Vol. 35, No. 2, pp. 181–191 (2011).
- [14] Ekman, P. and Friesen, W. V.: *Facial Action Coding System: Investigator's Guide*, Consulting Psychologists Press (1978).
- [15] Vrij, A. and Granhag, P. A.: Eliciting Cues to Deception and Truth: What Matters are the Questions Asked, *Journal of Applied Research in Memory and Cognition*, Vol. 1, No. 2, pp. 110–117 (2012).