

## 嘘つきコーパスの構築と分析

高島 匠平<sup>†</sup> 嶋田 和孝<sup>†</sup>

<sup>†</sup>九州工業大学情報工学部 〒820-8502 福岡県飯塚市川津 680-4

E-mail: †{s\_takabatake,shimada}@pluto.ai.kyutech.ac.jp

あらまし 家庭用ロボットの普及に伴い、ロボットと人間がコミュニケーションを行う機会が増加している。人間を傷つけないためにロボットにも嘘をつくことが求められる可能性がある。ゆえに、嘘の理解は必要不可欠である。本研究は、コンピュータによる嘘の理解の一環として、嘘の検出を行うことを目的とする。被験者ごとに変化しない環境で、表情の変化などの非言語情報から嘘の検出を行うために、嘘つきコーパスの構築を行う。また、嘘をついたときとそうでないときの違いの分析を人手と機械学習により行う。機械学習による分析では、表情認識などで用いられるLBP-TOP特徴量を用いて、嘘の検出実験を行う。

キーワード 嘘検出, 非言語情報, LBP-TOP 特徴量

## Construction and analysis of liar corpus

Shohei TAKABATAKE<sup>†</sup> and Kazutaka SHIMADA<sup>†</sup>

<sup>†</sup> Computer Science and Systems Engineering, Kyushu Institute of Technology

Kawazu 680-4, Iizuka, Fukuoka, 820-8502 Japan

E-mail: †{s\_takabatake,shimada}@pluto.ai.kyutech.ac.jp

**Abstract** Recognizing lies in interaction is one of the most important and challenging tasks in artificial intelligence. For example, it leads to development of a human-computer interaction system that detects a liar in a conversation. In another example, some systems in the future might occasionally need to lie not to hurt user's feelings, namely a white lie for trouble-free conversations. In this paper, we construct a video corpus for developing a system that detects lies in conversations. The corpus consists of 560 question-answer pairs by 9 persons; 280 true answers and 280 lies. We manually analyze the images in the corpus for recognizing lies. In addition, we apply a machine learning technique, SVMs with LBP-TOP features, into our corpus. We discuss the results of the analysis of the corpus and the machine learning approach to our corpus.

**Key words** lie detection, nonverbal information, LBP-TOP features

### 1. はじめに

嘘をつくという行為は、大学生を対象に行った調査によると一日に平均1.96回行われる日常的な行為であると報告されている[1]。家庭用ロボットの普及に伴い、ロボットと人間がコミュニケーションを行う機会が増加している。人間は相手を傷つけないために嘘をつくことがあり、ロボットにもそれが求められる可能性がある。また、ロボットに対して嘘が入力されることもある。ゆえに、ロボットは嘘を理解し、それをうまく処理しなければならない。そのために嘘の検出は必要不可欠である。また、嘘に関する研究として、人狼というコミュニケーションゲームがある。近年では人狼知能大会というものが行われ、人工知能同士が嘘を交えてのコミュニケーションを自然言語で行うようになっている。人狼において嘘を見抜くことは基本戦略である[2]といわれるほど重要である。

嘘の検出は先行研究において、発言の一部に嘘を含むかどうかを分類するタスクとして取り組まれており、主に機械学習が用いられる[3][4]。Enos[3]は英語で収録された嘘を含むCSCコーパスに対して、言語特徴量と音響特徴量を用いて分類を行っている。角森ら[4]は、CSCコーパスと、それと同様の設定で収録した日本語偽言コーパス(JDC)に対して嘘の検出を行い、英語と日本語で嘘の検出に有効な特徴量の差異を調べている。これらのコーパスは、人手により質問がされているため、被験者ごとに収録環境が異なる可能性がある。

本研究では、嘘とは「間違いである分かっていることを言うこと」とし、被験者ごとに変化しない環境で、嘘をつくことによる表情の変化などの非言語情報から嘘の検出を行うために、嘘つきコーパスの構築を行う。そのために質問に対する返答時の嘘に焦点を当て、質問の際に、モニターを用いることで、被験者ごとに環境が変わらないようにしている。また、嘘をつい



(a) 右斜め前からの撮影 (b) 左斜め前からの撮影



(c) 正面からの撮影

図 1: 撮影した画像例

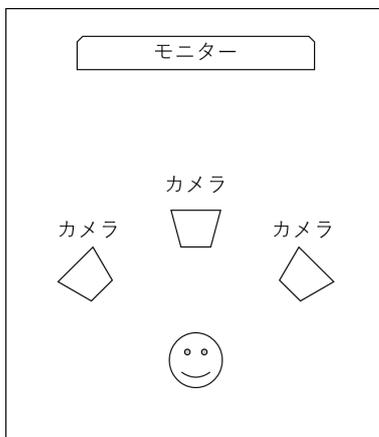


図 2: 撮影環境

たときとそうでないときの違いの分析を人手と機械学習により行う。

## 2. 嘘つきコーパスの構築

本研究では、被験者ごとに変化しない環境で、表情の変化などの非言語情報から嘘の検出を行うために、嘘つきコーパスの構築を行う。

被験者はモニターの前に座る。被験者からモニターまでの距離は 170cm である。モニターに表示される「嘘をつくか」もしくは「本当のことを言うか」の告知に従い、モニターに表示される質問に答える。この様子を被験者から見て右斜め前 (図 1a)、左斜め前 (図 1b)、正面 (図 1c) の 3 方向から撮影する。被験者とカメラとの距離は 45cm で固定している。撮影には GoPro HERO5 Session を用い、画素数は横幅 1920 ピクセル、縦幅 1080 ピクセル、フレームレートは 30 fps に設定している。収録環境を図 2 に示す。

収集したデータについて述べる。質問タイプによる嘘をついた際の違いを捉えるために、「はい」または「いいえ」で答える質問と物の名前や数値などで自由に答える質問の二種類の質問タイプを用意する。被験者は、二種類の質問タイプそれぞれ 30 問、計 60 問に対して返答をする。収集に使った質問の例を下に示す。



図 3: 嘘か本当かの告知と質問を同時に表示する場合

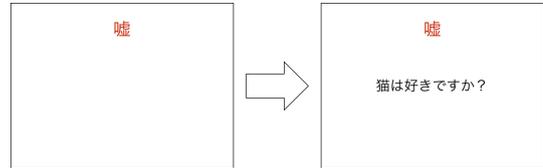


図 4: 嘘か本当かの告知を質問の先に表示する場合

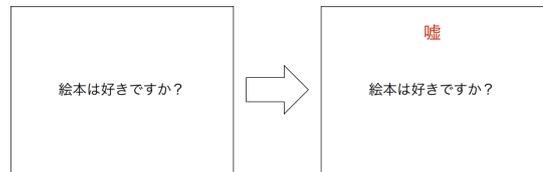


図 5: 嘘か本当かの告知を質問の後に表示する場合

- 「はい」か「いいえ」で答える質問の例
  - サッカーは好きですか?
  - バク転はできますか?
  - アメリカに行ったことはありますか?
  - 今日朝食を食べましたか?
- 自由に答える質問の例
  - 好きな動物は何ですか?
  - 海外には何回行ったことがありますか?
  - 0~9 で好きな数字はどれですか?
  - 昨日の天気は何でしたか?

それぞれの 30 のデータの中に、嘘をつくか本当のことを言うかの告知が、質問と同時に (図 3)、質問の直前 (図 4)、質問の直後 (図 5) に表示されるものが 10 ずつ含まれる。また、それぞれの質問の半数で嘘をつくように指示される。男性 7 名、女性 2 名の計 9 名を被験者として計 540 のデータを収集した。

## 3. 分析

本節では、まず 3.1 節で人手により嘘をついたときとそうでないときの違いを分析する。3.2 節では機械学習による嘘の検出を行う。

### 3.1 人手による分析

収録したデータのうち正面から撮影したデータを観察したところ、嘘をついた際に、

- 笑みの表情
- 表情が左右非対称
- 目線をそらす
- 韻律が変化する

といった特徴がみられた。図 6 に無表情のフレームを、図 7 に笑みの表情かつ目線そらしたフレームを、図 8 に表情が非対称



図 6: 無表情のフレーム



図 7: 笑みの表情かつ目線をそらしたフレーム



図 8: 表情が左右非対称になったフレーム

になったフレームを示す。図 8 では、被験者から見て左側の下唇が右側より上がっている。

嘘をついたときには、見破られるかもしれないという不安、欺瞞による罪悪感、相手を騙す喜びの 3 つの感情が生じるとされている [5]。確認された笑みの表情は、騙す喜びによるものである可能性が考えられる。また、顔面統制の技法として、修飾、調節、偽装の 3 つがあげられている [6]。修飾とは、示した表情の注釈としてもう一つの表情をあらわす事である。調整とは、本当に感じている表情の強度を実際より強めたり弱めたりすることである。偽装とは、何も感じていないのにある感情を示したり、ある感情を感じているのに表情に何も現れないようにしたり、実際に感じている感情の代わりに感じていない感情の顔貌を示したりすることである。表情が左右非対称になったときやや笑っているように見え、実際この後に図 7 のような笑みの表情が見られた。ゆえに左右非対称になったのは、笑みの表情を調整しようとしたものである可能性が考えられる。

### 3.2 機械学習による分析

人手による分析により笑みなどの表情が嘘と関係していることが分かった。そこで、笑顔などの表情分類などで使用される LBP-TOP 特徴量 [7] を機械学習の特徴量として用いる。LBP 特徴量は、画像の局所的な輝度の分布を捉えることができ、顔認識などで利用される。LBP-TOP 特徴量は、空間情報のみを扱う LBP 特徴量を時空間情報が扱えるように拡張したもので

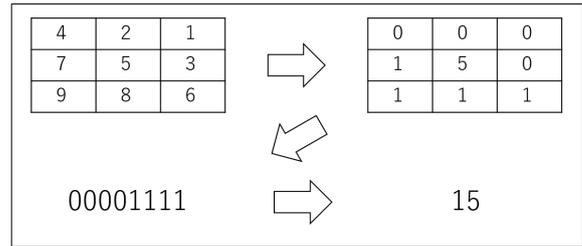


図 9: LBP 特徴量の計算の流れ

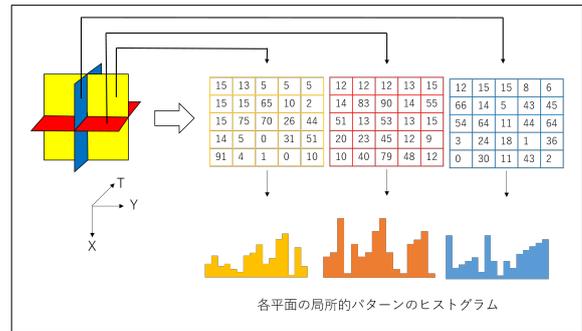


図 10: LBP-TOP 特徴量の計算の流れ

ある。

次に、LBP-TOP 特徴量の計算方法について述べる。画像シーケンスに対して、画像である XY 平面と画像の X 軸に対する時間的変化を表す XT 平面と画像の Y 軸に対する時間的変化を表す YT 平面に対して、特徴量の計算を行う。平面の各ピクセルに対して、そのピクセルの輝度の強さである画素値をその近傍ピクセルに対する閾値とみなし、近傍ピクセルの画素値が閾値よりも高い場合は 1、そうでない場合は 0 を設定する。このようにして、平面内のパターンを二進数で得る。これらを十進数に変換する。各ピクセルに対する計算の例を図 9 に示す。これらから平面ごとにヒストグラムを生成する。最後に、これらの 3 つのヒストグラムを連結させる。LBP-TOP 特徴量の計算の流れを図 10 に示す。

LBP-TOP 特徴量算出の具体的な処理について述べる。嘘をつくか本当のことを言うかの告知が行われた付近のフレームの顔領域を 5 フレーム毎に 18 フレーム分切り出す。顔領域の検出は、OpenCV<sup>(注1)</sup> の Haar 分類器を用いて機械的に行う。この際、顔領域が 18 フレーム切り出せなかったものは分析の対象外とした。これによりデータ数は 426(うち 215 が嘘)となった。切り出した顔領域をグレイスケール化する。このとき LBP-TOP 特徴量の計算の都合上縦幅を 200 ピクセル、横幅を 200 ピクセルの大きさに変更し統一する。これらから LBP-TOP 特徴量を算出する。

嘘つきコーパスに対する嘘の検出実験として、LBP-TOP 特徴量を用いた、SVM<sup>(注2)</sup> による分類を行った。評価には、leave-one-out 交差検定を用いた。表 1 にエラー率、再現率、適合率を示す。エラー率 E, 再現率 R, 適合率は P それぞれ次のように表される。

(注1) : <https://opencv.org/>

(注2) : <http://svmlight.joachims.org/>

表 1: 分析結果

特徴量	エラー率	再現率	適合率
LBP-TOP	44.84	57.21	55.41

$$E = \frac{FP + FN}{TP + FP + FN + TN}$$

$$R = \frac{TP}{TP + FN}$$

$$P = \frac{TP}{TP + FP}$$

ここで、TP は嘘を嘘に分類した数、FP は嘘でないものを嘘に分類した数、FN は嘘を嘘でないものに分類した数、TN は嘘でないものを嘘でないと分類した数である。今回提案した手法では、嘘の検出に対して十分な精度が得られなかった。LBP-TOP 特徴量のみでは、嘘をついた際の表情の変化を十分に捉えられていない可能性がある。

## 4. 考 察

### 4.1 嘘つきコーパスに対する考察

収録後に被験者に話を聞いたところ、嘘の告知がされた際に、嘘をついたつもりが後々考えると本当のことを言ってしまったというコメントがあった。正確な嘘つきコーパスの構築のために、被験者に自分の収録映像を見てもらい、言った内容と嘘か本当かが一致しているかを再度確認してもらう必要がある。

Levine [8] は嘘の定義を「欺瞞のサブセットであり、間違いであると分かっていることを言うことで他者を騙すこと」としている。本研究での定義との差は「欺瞞」という部分である。彼は欺瞞の定義を「意図的または故意にほかの人を間違えた方向に導くこと」としている。今回のコーパスは、間違いであると分かっていることを言うことに注目しており、意図的に間違えた方向に導くこと、つまり騙そうとする意思については考慮していない。騙そうとする意思を考慮したコーパスの収集を行い、今回の嘘つきコーパスと比較することで、騙そうとする意思が嘘をつくという行為にどのような影響を与えるかを確認する必要があると考えられる。

### 4.2 嘘の検出に対する考察

嘘の検出実験を行う際に、OpenCV による顔領域の検出を行ったが、顔領域が切り出せないことがあった。切り出せなかった付近のデータを確認したところ、被験者が正面のカメラの方向を向いていなかったり、手で顔の一部を覆っていたりしていた。このような場合の対応は、今後の課題である。

LBP-TOP 特徴量を用いて、機械学習による嘘の検出を行ったが、満足のいく結果ではなかった。今回使用した特徴量は人手による分析結果が十分に活かされていない。人手による分析で得られた笑みの表情、非対称な表情、目線の動きを表す特徴量を今後使用していく必要がある。また、今回は画像のみを利用したが、韻律の変化も確認されたため、視覚情報だけでなく音声情報にも注目した特徴量を使用し、有効性を確認する必要がある。また近年、画像認識で注目されている畳み込みニュー

ラルネットワークによる特徴量抽出や時系列データのパターンを認識できる再帰型ニューラルネットワークの使用も検討している。今回は、嘘か本当かの告知付近の表情の変化に注目したが、発言前や発言後の表情の変化についても分析を行う必要があると考えられる。

今回は複数の質問タイプを用いて、コーパスを構築をしたが、タイプによる嘘をついた際の違いについて分析していない。嘘の検出に有効な質問タイプの知識を得るために、それらの違いについて分析をする必要があると考えられる。

## 5. おわりに

本稿では、嘘つきコーパスの構築を行い、収録したデータから人手により確認された嘘をついた際の表情の変化に注目し、LBP-TOP 特徴量を用いて嘘の検出実験を行った。その結果、データ数 426(うち 215 が嘘) に対して、エラー率 44.84% であった。今後は人手による分析を十分に考慮した特徴量や手法を用いた嘘の検出を行いたい。

様々な種類のデータを収集したが、ついた嘘の種類の違いについては分析していない。今後はそれらの違いについて分析を行いたい。また、今回は間違いであると分かっていることを言うことに注目したが、今後は騙そうとする意思を考慮したコーパスの構築と、騙そうとする意思の有無での嘘をついた際の非言語情報の違いの発見を目指す。

## 謝 辞

本研究は科学研究 17H01840 の助成を受けたものである。

## 文 献

- [1] Bella M. DePaulo, Deborah A. Kashy, Susan E. Kirkendol, Melissa M. Wyer, and Jennifer A. Epstein. Lying in everyday life. *Journal of personality and social psychology*, Vol. 70, No. 5, p. 979, 1996.
- [2] 片上大輔, 鳥海不二夫, 大澤博隆, 稲葉通将, 篠田孝祐, 松原仁ほか. 人狼知能プロジェクト (< 特集 > エンターテイメントにおける ai). *人工知能*, Vol. 30, No. 1, pp. 65–73, 2015.
- [3] Frank Enos. *Detecting Deception in Speech*. PhD thesis, Columbia University, 2009.
- [4] 角森唯子, グラム・ニュービッグ, サクリアニ・サクティ, 戸田智基, 中村哲. 対話中における嘘検出に有効な特徴量比較と質問分析. *SIG-SLUD*, Vol. 4, No. 02, pp. 19–24, 2014.
- [5] Paul Ekman. *Telling lies: Clues to deceit in the marketplace, politics, and marriage (revised edition)*. WW Norton & Company, 2009.
- [6] Paul Ekman, Wallace V. Friesen, 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] Guoying Zhao and Matti Pietikainen. Dynamic texture recognition using local binary patterns with an application to facial expressions. *IEEE transactions on pattern analysis and machine intelligence*, Vol. 29, No. 6, pp. 915–928, 2007.
- [8] Timothy R. Levine. Truth-default theory (tdt) a theory of human deception and deception detection. *Journal of Language and Social Psychology*, Vol. 33, No. 4, pp. 378–392, 2014.