

研究計画書

「VR 異性間コミュニケーションスキルトレーニングシステムの提案」

早稲田大学人間科学部
菊池研究室所属
学籍番号 1J17D018 石田 豊

1 研究の背景

本研究では、異性間での好意の持たれるコミュニケーションに焦点を当て、音声からの好感度自動推定を実現し、それを導入した VR トレーニングシステムを作成する。

人間同士が関係を築くためには、コミュニケーションが不可欠である。関係を築くためには、相手に親しみやすいコミュニケーションを心掛けることが重要であるが、コミュニケーションに正解はなく、正しい指標も認知されていないことが課題である。2014 年のマイナビの調査 [14] では、男女計 500 人のうち、72.4% がコミュニケーションに苦手意識を持っていると回答した。苦手意識を持つと答えた人々は、緊張や会話が続かないといった悩みから、コミュニケーションに苦手意識を持つ傾向が見られ、上司や同僚・取引先などとうまくコミュニケーションできず、業務上困っているという回答が多く見られた。

異性間でのコミュニケーションにおいては、2016 年 12 月の楽天オーネットの新成人に対する調査 [9] では、新成人男女計 600 人のうち、74.7 %が「交際相手がおらず、交際相手がほしい」と回答した。また株式会社明治安田生活福祉研究所による「2017 年 男女交際・結婚に関する意識調査」[15] では、全国の 15～34 歳の男女 10、304 人を対象に、異性とのコミュニケーションが苦手か調査した。男性は全体で 50%、女性は 46.7% が苦手と回答した。以上の調査から、異性の交際相手を求める人は多く、また異性とのコミュニケーションが苦手意識を持つ人が多いことがわかる。ここで、異性間コミュニケーションスキルトレーニングの需要があると考えられる。

トレーニング内容については、合コンを想定したコミュニケーションスキルトレーニングである。これは男女の出会いの手段として認知度も高いためである。初対面の異性が向かい合い、すぐにコミュニケーションを取るといった状況が自然な設定であるといえる。合コンの場面を VR 環境で体験することで、「環境の慣れ」やコミュニケーションの問題の一因である「緊張」への対策が期待できる。

1.1 先行研究調査

音声からの好感度自動推定に関連する先行研究について述べる。

1.1.1 Attractive Voice

音声の好感度については、attractive voice の研究が古くから行われている。Apicella ら [2] の調査では、タンザニア狩猟民族において、Voice Pitch の低い男性が有意に生殖成功率が高いことも明らかになっている。また、Jovana ら [6] の実験では、113 名平均年齢 20 才の女性を対象に、男性の声を加工し、Voice Pitch を下げた男性化された声と Voice Pitch を女性らしく高くした声で、「I really like you」と「I really don't like you」の二つのセリフの計 4 種類を 8 対用意し女性被験者に無作為に聞かせた。結果、発言内容に関わらず、女性化された声以上に男性化された声に魅力を感じるという回答が有意であった。

1.1.2 好感度推定に関する国内研究

次に、好感度推定に関する国内研究の調査を行った。堀池ら [13] の実験では、現状の発話の印象を改善するための発話訓練がボイストレーナの主観に基づいて行われていることを指摘し、客観的指標を提示するため、推定モデルを作成した。音響特徴量と抑揚の操作が音声の好感度に与える影響を調査し、基本周波数とスペクトル包絡が好感度に寄与することが明らかになっている。

神山ら [10] の研究では、アノテータのラベル付与能力を考慮した電話応対音声の好感度推定モデル学習法の検討を行った。コンタクトセンタのオペレータの応対音声における従来の好感度推定モデルでは、ラベル付与誤りによる精度の低下が考慮されていなかった。長時間音声のラベル付与であり、1 音声に付与できるラベルの数は限られ、ラベル付与誤りによって精度の低い正解ラベルであるのが問題点であった。解決のアプローチとして、2 つの手法を用いた。1 つ目は、アノテータのラベル付与誤りを考慮した好感度ラベルの推定を行った学習データの構築であり、2 つ目は、本来の好感度からラベル付与誤りが発生する過程をニューラルネットワークに内包したモデル学習である。従来の好感度推定から、最大誤り削減率 12% を達成した。

1.1.3 Interspeech2012 speaker likability sub-challenge

InterSpeech 2012 speaker trait challenge[1] では、音響情報から話者特性の検出を目的とした企画であり、好感度 (likeability) 推定のサブチャレンジが実施された。好感度データベースと話者検出の Baseline システムが提供され、精度向上を目指す数多くの研究が提案されている。

Jouni ら [5] は、Baseline システムの特徴量セットの 6125 項目から、重複した内容の特徴量を排除し、必要最低限まで絞り込み、349 項目まで削減した。推定モデルは 2% の精度の向上が見られた。

Dingchao ら [3] は、Baseline システムの手法の見直しを行った。ガウス過程回帰を使用したモデルを作成し、1.1% の精度の向上がみられた。

1.1.4 コミュニケーション能力推定

コミュニケーション能力の自動評価に関する研究は近年数多く行われている。岡田ら [7] はコミュニケーション能力の推定を行った。言語情報だけでなく、発話、韻律、視線、ジェスチャー、姿勢、表情の役割が重要視された。上記の要素から特徴量を抽出し、人事採用経験者による評価で、コミュニケーション能力に関わる特徴量を明らかにした。総合的なコミュニケーション能力値に大きく寄与した特徴量は、単一では発話ターン、複数特徴量の組み合わせでは、発話ターン、発言内容、動作であることが明らかになった。

1.1.5 感情推定に関する研究

最後に、好感度推定そのものの試みではないが、手法的には類似しているため、主観的な評価を音響的特徴から推定する研究を述べる。

羽田ら [12] の実験では、感情音声コーパスである Japanese Twitter-based Emotional Speech (JTES) を使用して、ニューラルネットワークによる感情認識を行い、音響特徴量の検討を行った。無声音区間の検出が 3~6% の精度向上を可能にした。

神山ら [11] の研究では、コンタクトセンタにおける顧客の要望の緊急度推定をおこなった。従来手法では、短時間特徴と話速を組み合わせた Support Vector Machine(SVM) の手法で行っていた。緊急度におけるリズムの重要性を指摘し、従来手法では短時間特徴では短すぎ、話速のような全体平均特徴量では、リズムを捉えるのは難しいという問題があった。そこでリズム特徴を捉えるために、Envelope Modulation Spectrum

(EMS) および、音の変動の大きさを示す MFCC の時間ごとの統計量を用いて、長時間リズム特徴を抽出した。また、手法を SVM から、Recurrent Neural Network(RNN) に変更し、従来の短時間特徴量と全体特徴量話速を用いた結果と比べて 50% の誤り削減率を達成した。

2 研究の目的と目標

本研究の最終的な目的は、異性間コミュニケーションに不安を抱える人々に VR トレーニングを通じて、好意を持たれるコミュニケーションスキルを獲得してもらうことである。

これに従い、現時点での好感度推定における問題点を先行研究において調査した。Attractive Voice で提示した研究は、進化生物学的な視点であり、人間の根本的な部分では魅力的な音声には Voice Pitch が寄与していることは明らかであるが、社会的、文化的な背景が考慮されていないという問題がある。

次に、好感度推定に関する国内研究で示した研究について、堀池ら [13] の実験では、基本周波数とスペクトル包絡が好感度に寄与することが明らかであるが、ここでの好感度と本研究での好感度とは少し差異がある。堀池ら [13] が示しているのは、他者に良い印象を与える度合いであり、本研究では、交際目的で好意をもたれる度合いであることを指摘する。好感度の種類によって、特徴量やその変化率に差異があるかどうかを検討することが課題である。同様に、神山ら [10] の研究でも、好感度推定を行っている。オペレータの好感度推定と異性間コミュニケーションにおける好感度推定では、前者は好印象な音声、後者は魅力的な音声という違いがある。

また音響特徴量について、数秒単位の短い発話、単語単位、一文単位の好感度の推定や特徴量については研究事例が数多くあるが、対話では持続的に作用するのかどうか、時間的な長さを考慮した研究事例は少ない。本研究では、好感度特徴量の持続性も検証したい。

上記の課題を踏まえ、本研究の目標は、本研究の目的に沿った音響特徴量の特定と持続性の検証を行い、推定モデルの作成と評価を行う。そして、VR システムへの導入を目指す。

3 研究内容と手法

本研究では、VR 空間を用いて、合コンを想定したコミュニケーショントレーニングを行う。体験者には、VR 空間で、仮想の異性との会話を行い、好感度推定に基づくリアルタイムでのフィードバックと終了時にコミュニケーションの評価結果を提示する。コミュニケーションスキルの可視化により、コミュニケーションの改善を目指す。

また、評価者の性格傾向ごとにモデルを作成し、多角的な採点を行い、コミュニケーション指標を複数提示する。システム実現の第一段階として、短い発話である自己紹介の好感度推定を行い、第二段階では、対話における好感度推定を行う。

3.1 VR 環境の構築

VR 環境の構築には、ゲームエンジン Unity を使用する。想定環境を以下に記載する。

- 想定場面:男女合同コンパ
- 想定場所:飲食店などの宴会用の個室
- 想定人数:男女 6 人

- 想定内容:自己紹介、女性と対話を行う。

3.2 評価システムの作成

評価システムの作成には、以下の手順を行う。

- 音声収録
- 好感度評価
- 推定モデル作成

3.2.1 音声収録

本研究は男性を想定したコミュニケーショントレーニングであるため、男性の音声データを取得する必要がある。第一段階として、自己紹介音声の収録を行う予定である。内容を統一させた自己紹介スクリプトを読み上げる。

3.2.2 好感度評価

好感度評価では、女性による評価を取得する。評価時に性格傾向を取得するアンケートを行う。アンケートには、小塩ら [8] が作成した日本語版 Ten Item Personality Inventory を使用する。

3.2.3 推定モデル作成

本研究で使用する特徴量は、OpenSMILE[4] を使用し、特徴量を抽出した後、次元削減を行い、好感度を目的変数とした SVM を作成する。修士では、ニューラルネットワークを用いた学習での作成を行う。

4 具体的な計画

4.1 卒業研究

大学4年次には、異性間コミュニケーションにおける自己紹介の好感度推定システムの完成を目標とする。
5～7月：

- クラウドソーシングによるデータ収録
- クラウドソーシングによるデータ評価
- 推定モデルの作成・評価

8～10月：

- 推定モデルを搭載した VR 環境の構築
- VR システムの評価

11～12月：

- 成果をもとに、卒業論文を仕上げる。

4.2 修士

卒業研究では最初の自己紹介場面に絞るが、修士では一連の対話場面に対象を広げる。修士での課題として、異性間コミュニケーションをモデル化する必要がある。そのため、交際以前の異性間コミュニケーションの特性を明らかにすること、対話における場面ごとの好感度評価モデルを明らかにすること、評価モデルを実装した VR トレーニングシステムを設計することがあげられる。

1 年次は、先行研究調査や音声収録と評価を通して、交際以前の異性間コミュニケーションの特性を明らかにする。特徴量の特定や、好感度変化を左右する特徴量と卒業研究で使ったものとを比較する必要がある。対話場面では、自己紹介と違い、一言の好感度の重みが異なるため、検討する必要がある。

また、対話における場面ごとの好感度評価モデルを作成する。時間的な変化や会話の回数によって、好感度の重みが変わる可能性があり、複数の好感度評価モデルを作成する必要がある。対話場面にあった好感度評価モデルを作成、評価する。

2 年次は、評価モデルを実装した VR システムを作成・評価を行う。評価方法については、被験者が VR トレーニングを体験する前と後で、コミュニケーションの好感度変化が見られたかを女性による評価で判断する。評価向上が明らかになれば、本システムの有用性が証明できる。

5 期待される成果

異性とのコミュニケーションに不安を抱える男性に対して、本研究の VR システムを体験し、トレーニングしてもらうことで、コミュニケーションスキルの向上を目指す。また、一つの客観的な指標を提示することで、体験者がコミュニケーションのさらなる指標や新しい方向性を発見する手がかりとなることが、期待される研究成果である。なお完成後は、企業への提案や実用化を目指し、異性とのコミュニケーションに不安を抱える人々に利用してもらうことで、社会への貢献を果たしたい。

6 おわりに

以上のように研究を推進する計画である。SNS の普及によって、対面でのコミュニケーション場面が減少する社会で、VR のような技術を活用し、対面でのコミュニケーションをトレーニングする場を提供することで、一人からコミュニケーショントレーニングができる環境を作り上げ、コミュニケーションに苦手意識を持つ人々の助けになれば幸いである。

参考文献

- [1] Schuller, B. Steidl, S. Batliner, A. Nöth, E. Vinciarelli, A. Burkhardt, F. Van Son, R. Wening, F. Eyben, F. Bocklet, T. Mohammadi, G. Weiss, B. ” The INTERSPEECH 2012 speaker trait challenge”, 13th Annual Conference of the International Speech Communication Association 2012, INTERSPEECH 2012, 2012, Vol.1, pp.254-257
- [2] Apicella CL, Feinberg DR, Marlowe FW. 2007. ” Voice pitch predicts reproductive success in male hunter-gatherers” . Biol Lett. 3:682-684.
- [3] Dingchao Lu, Fei Sha, ” Predicting Likability of Speakers with Gaussian Processes” , 13th Annual

- Conference of the International Speech Communication Association 2012, INTERSPEECH 2012, 2012, Vol.1, pp.286-289
- [4] ” openSMILE:” <https://www.audeering.com/opensmile/>
 - [5] Jouni Pohjalainen, Serdar Kadioglu, Okko Rasanen. ” Feature Selection for Speaker Traits” . 13th Annual Conference of the International Speech Communication Association 2012, INTERSPEECH 2012, 2012, Vol.1, pp.270-273
 - [6] Jovana Vukovic, a Benedict C. Jones, Lisa DeBruine, a David R. Feinberg, Finlay G. Smith, Anthony C. Little, Lisa L. M. Welling, and Julie Main. 2010. ” Women’ s own voice pitch predicts their preferences for masculinity in men’ s voices” , Behavioral Ecology, Volume 21, Issue 4, July-August 2010, Pages 767-772
 - [7] 岡田 将吾, 松儀 良広, 中野 有紀子, 林 佑樹, 黄 宏軒, 高瀬 裕, 新田 克己, ” マルチモーダル情報に基づくグループ会話におけるコミュニケーション能力の推定” 人工知能学会論文誌 31 巻 6 号 AI30-E(2016 年)
 - [8] 小塩真司, 阿部晋吾, カトローニ ピノ, ” 日本語版 Ten Item Personality Inventory(TIPI-J) 作成の試み” , パーソナリティ研究, 2012, Vol.21(1), pp.40-52
 - [9] 株式会社オーネット広報グループ, ” 第 22 回新成人意識調査 2017 年新成人 (全国 600 人) の恋愛・結婚意識” , ことぶき科学情報, Vol.79, 2017.
 - [10] 神山歩相名, 安藤 厚志, 増村 亮, 小橋川 哲, 青野 裕司, ” アノテータのラベル付与能力を考慮した電話応対音声の好感度推定モデル学習法の検討” , 一般社団法人 電子情報通信学会 Vol.IEICE-118, No.IEICE-EA-495, IEICE-SIP-496, IEICE-SP-497, pp.IEICE-EA-197, IEICE-SIP-197, IEICE-SP-197-IEICE-EA-202, IEICE-SIP-202, IEICE-SP-202(2019.3)
 - [11] 神山歩相名, 安藤厚志, 増村亮, 小橋川哲, 青野裕司, ” 話速の変動を捉える特徴量に基づく留守録音声の緊急度推定” , 日本音響学会講演論文集 1-3-4(2019.9)
 - [12] 羽田優花, 加藤正治小坂哲夫, ” 感情音声データベース JTES を用いた音声感情認識における特徴量の検討” . 日本音響学会講演論文集 2-P-36(2019.9)
 - [13] 堀池梓哉, 森勢将雅, ” 音響特徴量と抑揚の操作が発話音声の好感度に与える影響の分析” , 日本音響学会講演論文集 2-q-27(2019.9)
 - [14] マイナビニュース, ” 72.4% の人がコミュニケーションに苦手意識-「言葉に詰まる」「緊張する」” , <https://news.mynavi.jp/article/20140506-a063/>, (閲覧日 2020 年 5 月 5 日).
 - [15] 明治安田総合研究所, ” 15～34 歳の恋愛と男女交際- 男女交際・結婚に関する意識調査より-” , https://www.myri.co.jp/research/report/pdf/myilw_report_2017_01.pdf, (閲覧日 2020 年 5 月 5 日).