

2020 年度システム開発仕様書

VR 異性間コミュニケーション  
スキルトレーニングにおける自己紹介音声  
の好感度推定：システム仕様

指導教員  
菊池 英明 教授

早稲田大学人間科学部人間環境科学科  
学生番号 1J17D018

石田 豊  
Ishida Yutaka

2020 年 6 月 28 日

## 目次

1	はじめに	3
1.1	本研究の背景 . . . . .	3
1.2	本研究の目的 . . . . .	4
2	VR トレーニングシステム仕様	4
2.1	評価システムの作成 . . . . .	4
2.2	評価の流れ . . . . .	5
2.3	VR 環境 . . . . .	6
2.4	VR 環境の構築 . . . . .	7

## 図目次

1	推定モデル . . . . .	5
2	システム処理の流れ . . . . .	5
3	トレーニング画面 . . . . .	6

# 1 はじめに

## 1.1 本研究の背景

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 [10] 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 [6] 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」[11] 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。

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 [1] 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 [3] 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。

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 [9] 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 [7] 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。

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。

## 1.2 本研究の目的

本研究では、異性との交際を望むが、コミュニケーションに不安を抱える人々を対象に、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の獲得支援を行う。

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ことが課題である。また、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 [9] 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 [9] 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証することが課題である。また、精度向上のためには堀池らが注目した音響特徴量以外に、好感度に寄与する特徴量の検討が必要である。岡田ら [4] は、コミュニケーション能力推定で話速を使用している。話速で好感度変化は起こるのかを検証することも課題である。

上記の課題を踏まえ、本研究の目的に沿った音響特徴量の特定や追加した特徴量の好感度変化の検証を行い、重み付けを行った上で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ことが目標である。

## 2 VR トレーニングシステム仕様

本章では、VR トレーニングシステムの仕様について述べる。

### 2.1 評価システムの作成

まず、評価システムのモデル学習、作成について述べる。男性の自己紹介データを 144 音声収録し、女性評価者が評価を行った。女性評価者には 10 項目のアンケート [5] を実施し、性格傾向を取得した。女性評価者のラベルが付いた自己紹介音声を学習させる。モデルの学習について、図 1 に示す。opensmile [2] によって自己紹介音声の音響特徴量を抽出を行うのと同時に、Google speech to text によるモーラ数の抽出と発話区間検出を行い、話速を取得する。その後、特徴量を機械学習にかける。機械学習は SVM で行う。女性評価者から取得した性格傾向ごとにデータを分割した後、同様に学習し、5 つのモデルを作成する。総合モデルと性格傾向ごとのモデル、計 6 つのモデルから、自己紹介音声を評価する。

## 評価システムの作成

女性評価者の性格傾向の  
5種類の推定モデル

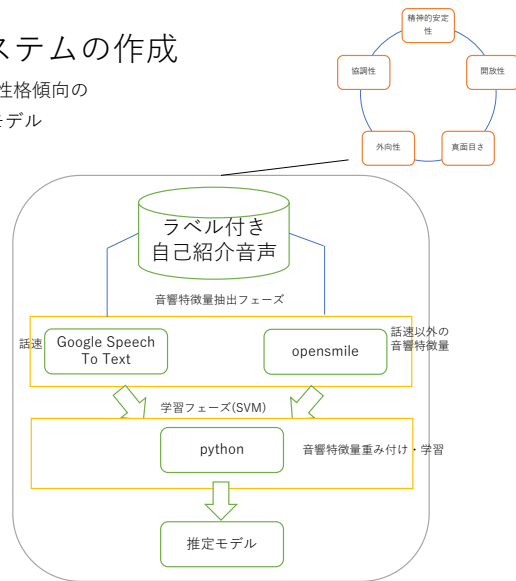


図1 推定モデル

## 2.2 評価の流れ

図2に本研究におけるシステム処理の流れを示す。システム処理の流れは、ユーザーが自己紹介音声を評価システムに入力することで、評価値を返すという流れである。女性の性格傾向評価を行い、テキストによるフィードバックとグラフを表示する。

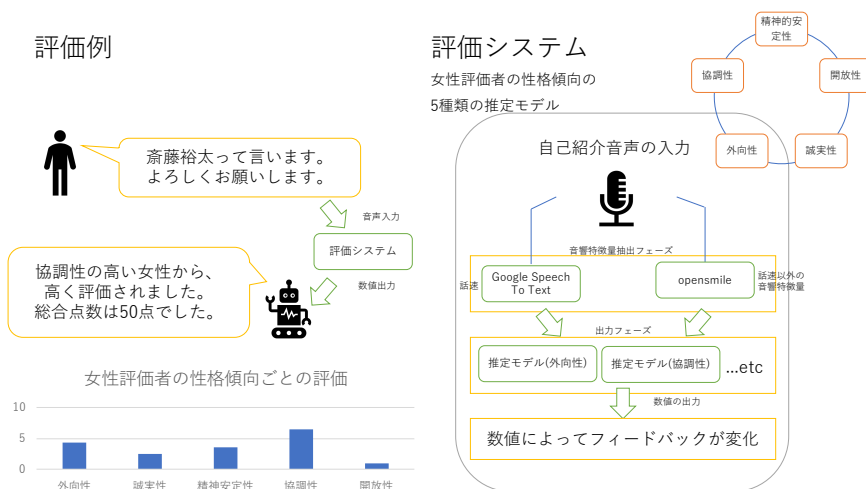


図2 システム処理の流れ

## 2.3 VR 環境

想定する VR 環境を示す。また、図 2 は実際にトレーニングを行う画面の一部である。

### 2.3.1 想定する VR 環境

- 想定場面:男女合同コンパ
- 想定場所:飲食店の個室
- 想定人数:男 3 女 3 計 6 人
- 想定内容:女性と対話を行った後、自己紹介をする。

VR 環境は実際に異性とのコミュニケーションを行う場所として、合コンを想定している。合コンは、異性とのコミュニケーションの場として周知であり、また交際を目的とする男女の出会いの場であるため、交際を見据えたコミュニケーショントレーニングの場として最もふさわしいためである。合コンの参加人数は女性 3 人、ユーザー含む男性 3 人の計 6 人を想定している。この理由として、3 対 3 で行われる合コンは、全員が話題を共有でき、席の都合など会話がしやすいという利点があり、一般的に黄金比と呼ばれている。2 対 2 の合コンは、3 対 3 に比べて必然的に会話量が多くなるため、コミュニケーションに不安を抱える人にとって、ハードルの高いトレーニングになる可能性がある。1 対 1 も同様である。本研究では、3 対 3 を採用する。また本研究では、より現実に近いトレーニングを構築するため、実写でのトレーニングを想定している。



図 3 トレーニング画面

### 2.3.2 一連の画面遷移シナリオ

- 着席場面
- 異性との雑談場面
- 自己紹介する流れ場面
- 仮男性 1 の自己紹介場面
- ユーザーの自己紹介場面

- 評価場面

より自然な自己紹介を促すため、着席場面から開始する。そして、異性から軽い雑談が始まり、自己紹介を促す会話が行われる。その後、システムであらかじめ用意してある男性の自己紹介音声を流すことで、現実感を演出し、そのあとテロップで自己紹介をユーザーに促し、ユーザーは自己紹介を行う。評価場面に移り、評価が行われるといった画面遷移を行う。

## 2.4 VR 環境の構築

VR 環境の構築には、Unity を使用する。

## 参考文献

- [1] Apicella CL, Feinberg DR, Marlowe FW. 2007. "Voice pitch predicts reproductive success in male hunter-gatherers". Biol Lett. 3:682-684.
- [2] "openSMILE:" <https://www.audeering.com/opensmile/>
- [3] 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
- [4] 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)
- [5] 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
- [6] 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.
- [7] 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", 電子情報通信学会技術研究報告 = IEICE technical report : 信学技報, 2019-03, Vol.118(495), pp.197-202
- [8] 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)
- [9] 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)
- [10] 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).
- [11] 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](https://www.myri.co.jp/research/report/pdf/myilw_report_2017_01.pdf), (閲覧日 2020 年 5 月 5 日).