

平成 28 年度 卒業研究

機能語に注目した  
音声合成朗読システムのための感情  
推定

東京理科大学 理工学部 経営工学科  
西山研究室 7413069 恒川 泰輝

指導教員 西山 裕之

# 学士論文概要

# 目 次

学士論文概要	i
第 1 章 序論	1
1.1 背景 . . . . .	1
1.1.1 音声合成 . . . . .	1
1.1.2 オーディオブック . . . . .	1
1.2 本論文の目的 . . . . .	2
1.3 本論文の構成 . . . . .	2
第 2 章 関連研究	3
2.1 音声合成 . . . . .	3
第 3 章 提案手法	4
3.1 提案手法の概要 . . . . .	4
第 4 章 実装	5
4.1 開発環境 . . . . .	5
4.2 まとめ . . . . .	5
第 5 章 評価	6
5.1 評価目的 . . . . .	6
第 6 章 結論	8
6.1 まとめ . . . . .	8
6.2 今後の展望 . . . . .	8



# 图 目 次

# 表 目 次

# 第1章 序論

序論ではまず，本論文の背景としてオーディオブックと音声合成について解説する．

## 1.1 背景

### 1.1.1 音声合成

音声合成とは，人間の音声を人工的に作り出すことである．音声合成技術は文字を読むことが困難な障害者，外国人や幼児などに画面読み上げソフトとして長く利用されてきており，言葉を発することが困難な人が代替手段として利用することも多い．さらに，21世紀に入ってからでは家電製品の音声ガイダンスや公共交通機関のアナウンス，ロボットの発話用途などとして広く使用されるようになっている．

### 1.1.2 オーディオブック

オーディオブックとは主に書籍を朗読したものを録音した音声コンテンツのことである．アメリカを中心に市場規模が拡大している．もともと車社会のアメリカなどの国では早期から市場が確立していたが，近年インターネットを介して気軽にダウンロードして楽しめる環境が整ったことなどによりアメリカとカナダの市場規模が2015年には前年比21%拡大している．また日本においても定額配信サービスが開始されており，今後さらに普及する可能性がある．しかしながら，このようなオーディオブックは書籍から音声化する際に手間やコストが電子書籍にくらべて10倍ほどかかっており2～3ヶ月ほどかかると言われている．

## 1.2 本論文の目的

音声合成技術を用いて人手で行っている朗読作業を根源的な目的である。これまでの音声合成研究の結果、単に情報を伝達する目的では十分な音質が確立されている。しかし、従来の音声合成は一文やフレーズの読み上げでは高品質な音声を実現している一方で書籍データのような長い文章では平板で淡々とした読み上げになってしまい、感情的あるいは情緒的な表現を多く含む小説などの朗読を聞くには不十分である。近年になって、喜怒哀楽といった感情の種類をパラメタとして与え表現豊かな音声合成できるソフトが販売されている。しかし、これらのパラメタは文または単語ごとに人手で設定する必要がある。短い文章など限られた場合は容易であるが、小説といった膨大な文章に対して都度人手でパラメタ調整を行うのは大変手間がかかる。そこで、本研究では文章から読み上がる感情として最適なものを推測することを目的とする。

## 1.3 本論文の構成

本章では、本論文の背景となる。。。について説明した。そのうえで、。。。のニーズと課題について説明し、それを踏まえ。。。概要に触れた後、本論文の目的を述べた。第二章では既存の関連する研究について述べる。また第三章では本論文が提案する手法の詳細を述べ、第四章でその具体的な実装について説明する。そのうえで第五章で提案手法の効果を測定するために行なった実験の結果と考察を述べ、第六章でその結論と今後の展望について述べる。



## 第2章 関連研究

本章では音声合成や朗読システム、感情推定に関する既存の研究について紹介し、それに対しての本研究の位置づけと意義を述べる。

### 2.1 音声合成

て s t

## 第3章 提案手法

本章では提案する手法の詳細について説明する。

### 3.1 提案手法の概要

本手法では予め人手で各文に対して感情ラベルを与えた教師データを作成しそのデータを機械学習することで分類機を作成する。感情ラベルは、普通、嬉しい、恥ずかしい、怒ってる、悲しいの5種類とする。未知の入力文が与えられた場合に、この中の一つを自動的に割り当てることが目的である。基本的に句点ごとに文章を分割してすべての文がそれぞれのどの感情として音声合成されるべきかを推定する。なお、分割する際は基本的に句読点で分割し、カギカッコの始まりは句読点とみなしてそこから次の文の始まりとした。また、カギカッコ内の句読点も分割の対象とした。ここで文章に依存する内容語による推定では膨大な教師データが必要になると考えられる。そこで内容語によらず文の形式である程度、感情を推定することができると仮定して、文の機能語のみを抽出して学習を行う。具体的には名詞、動詞、形容動詞、形容詞を取り除く。その後、各文は品詞分解し基本形変換を行い bags of words に変換し正解ラベルを元に学習を行う。分類機にはナイーブベイズ、ランダムフォレスト等を用いる予定である。

## 第4章 実装

本章では，提案するシステムの具体的な実装について述べる．まず使用した機材や技術情報などについて明確にし，次に提案するシステムの流れを全体像とともに説明する．

### 4.1 開発環境

### 4.2 まとめ

本章では，提案するシステムを実装するにあたり利用したハードウェア及びソフトウェア情報について述べ，具体的な手法の実装部分の詳細について述べた．

## 第5章 評価

本章では提案手法を具体的に実装したシステムを用いて，主要コンテンツの抽出性能を評価するために行なった実験の概要と結果について述べる．

### 5.1 評価目的



## 第6章 結論

本章では，本研究のまとめ，および今後の展望について述べる．

### 6.1 まとめ

本研究では

### 6.2 今後の展望

今後の展望として，

## 参考文献

- [1] 吉田 有里, 奥平 康弘, 田村 直良 (2003, August). Eliminating noisy information in web pages for data mining. In Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 296-305). ACM.
- [2] Internet Live Stats, <http://www.internetlivestats.com/>
- [3] 一般財団法人経済広報センター 情報源に関する意識・実態調査報告書, 2013
- [4] Facts Hunt, <http://www.factshunt.com/>
- [5] World Wide Web Consortium (W3C), <http://www.w3.org/>
- [6] HTML5 is a W3C Recommendation (W3C news), 2014
- [7] Yi, L., Liu, B., & Li, X. (2003, August). Eliminating noisy information in web pages for data mining. In Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 296-305). ACM.

## 付 録 A    Web ページの URL リスト