

誤認識データベースと用言コーパスによる助詞補正を用いた音声認識後の文章補正方式

合田 輝幸[†] 渡部 広一[†] 河岡 司[†]

[†]同志社大学工学部知識工学科

1. はじめに

近年，人間と知能ロボットとの柔軟なコミュニケーションが重要視されている．そのためには，知能ロボットが話者の言葉を多様な環境で正確に聞き取る必要がある．

本稿では，市販の音声認識ソフトから得た出力に対して誤認識データベース^[2]を用い文脈的に補正を行うことで，不特定話者に対してより精度の高い音声認識システムを実現することを目的とする．その手法として，誤認識データベースによる一致度補正，複数の音声認識ソフトの複合利用，用言コーパス^[1]による助詞補正を用いた文章補正方式を提案する．

2. 文章補正方式について

2.1 文章補正の流れ

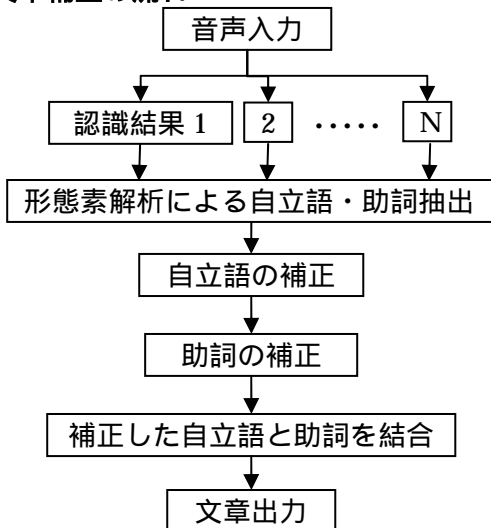


図 1：文章補正の流れ

図 1 に示すように，文章補正方式において，補正は自立語と助詞に対して行われる．

Sentence Modification Method after Speech Recognition using The Incorrect Recognition Database and Particle Modification by Declinable Word Corpus

[†]Teruyuki Gouda Hirokazu Watabe Tsukasa Kawaoka Knowledge Engineering and Computer Sciences, Doshisha University

2.2 自立語の補正

自立語の補正には誤認識データベースを用いた一致度補正方式を採用している．

2.1.1 誤認識データベース

入力語 A に対して音声認識ソフトが出力した認識語 a_i と出現割合 w_i の対の集合を求める．

$$A = \{(a_1, w_1), (a_2, w_2), \dots, (a_i, w_i)\}$$

ここで a_i を「属性」， A を「見出し語」とし，このような属性が定義された見出し語の集合を誤認識データベース（表 1）と呼ぶ．

表 1：誤認識データベースの例

見出し語	属性 1, 重み	属性 2, 重み	属性 3, 重み	...
上	うえ, 86.7	うで, 3.7	すえ, 1.5	...
酔う	よう, 45.9	ようぐ, 21.5	りょう, 6.7	...
腕	うで, 39.3	ので, 27.4	すえ, 9.4	...
胸	むね, 24.9	むねん, 24.5	うで, 10.8	...

2.1.2 一致度補正方式

音声認識ソフトから返される複数の認識語を一組とし，誤認識データベース中の任意の見出し語の属性集合と比較を行う．認識語が属性集合の中に含まれる場合，各々の属性の出現割合の総和をその見出し語の一致度とし，全ての見出し語に対して一致度を計算し，一致度が最大となる見出し語を最終的な結果として出力する．

2.2 助詞の補正

助詞の補正には多数決処理による補正を用いている．これは，複数の音声認識ソフトから得た認識結果の助詞のうち最も多く出力された助詞を補正結果とする方法である．

3. 実験環境

本研究では，音声認識ソフトに特定話者用ソフト D と不特定話者用ソフト J を用いた．入力には無線ハンドマイクを利用し，音声認識専用のマシンで入力音声を受信する方式を採用した．また受信機を複数台設置することで，ほぼ同程度の入力音声を複数の音声認識専用マシンで受信することが可能となる．本稿では 3 台の受信機とマシンを用意し，音声認識実験を行った．

4. 音声認識ソフトの複合による自立語補正

本研究では従来、一方の音声認識ソフトの認識結果のみで補正を行い精度の検証を行っていたが、ソフトウェア D とソフトウェア J では認識しやすい語に差異がある。そのため、2つのソフトから得た認識結果全てを対象として補正することにより正解率向上が期待できると考える。

そこで、音声入力に対して以下の3種類の認識結果を用いて補正を行い、精度を検証した。助詞の補正には多数決処理を用いている。

ソフトウェア D の認識結果 (6 候補)

ソフトウェア J の認識結果 (6 候補)

ソフトウェア D (3 候補) とソフトウェア J (3 候補) を合わせた認識結果 (6 候補)

また、比較対象として について補正を行う前の正解率 (認識結果 6 候補の平均値) を「補正前」として示す。被験者 15 名の認識結果を元に、上記 ~ の比較を行った。得られた結果が図 2 である。

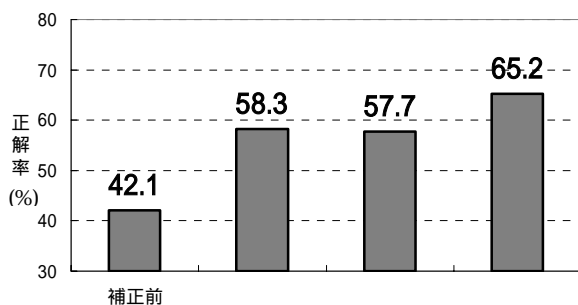


図 2：音声認識ソフトによる正解率の比較

図 2 より、 の手法の精度が、 に比べて約 7%高くなっていることがわかる。よって、一方のソフトの認識結果のみを補正するより、両方のソフトの認識結果を複合した上で補正を行う方がより効果があることが示された。

5. 用言コーパスを用いた助詞補正

2.2 節で説明した多数決処理による助詞補正方式には、認識結果に助詞が出力されない場合や全ての助詞の認識結果が異なっていた場合に補正が行えないという問題点がある。そのため認識結果の助詞に依存しない補正方法が必要となる。そこで、用言に対する主語や目的語を定義した用言コーパスを用いて補正結果の自立語から助詞を判断する手法を提案し、検証を行った。自立語の補正には 4 章で提案した複合手法を用いている。

被験者 15 名の認識結果を元に、助詞補正に用言コーパスを用いて補正を行ったところ、平均 30.6%とかなり低い正解率となった。これは本研

究で用いた用言コーパスのデータ量が少なく、適切な補正が行えなかったためだと考えられる。

そこで、上記で説明した多数決処理で補正不可能な認識結果にのみ用言コーパスを適用し、多数決処理と比較を行った。比較対象として 4 章と同じく補正前の正解率を示す。得られた結果 (被験者 15 名の平均) が図 3 である。

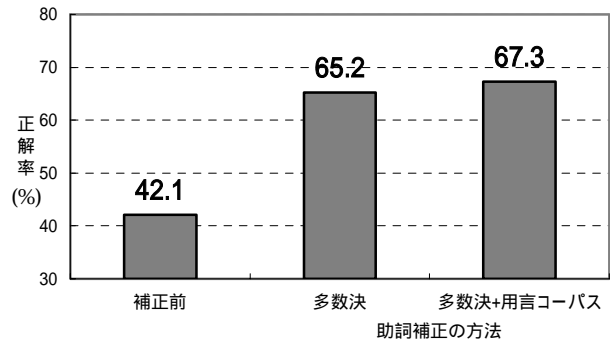


図 3：助詞補正の方法による正解率の比較

図 3 より、用言コーパスを用いることで 2.1% 正解率が向上していることがわかる。本手法において、多数決処理による補正が不可能な認識結果にのみ用言コーパスによる補正を適用することで、2つの補正方法の特性を生かし、精度を向上させることができた。

6. まとめ

ロボットと人間が自然にコミュニケーションを行う際に最も重要となるのは、正確な音声認識である。本稿では、音声認識後の文脈的補正によって高精度の音声認識を得るため、音声認識ソフトの複合利用、用言コーパスを用いた助詞補正を提案した。これらの手法を用いることで、図 3 に示すように補正前 42.1%に対して補正後 67.3%と、約 25%の補正効果が得られた。

以上より、本研究の目的である「音声認識後の文脈的補正による精度向上」を実現するシステムを構築することができた。

謝辞

本研究は文部科学省からの補助を受けた同志社大学の学術フロンティア研究プロジェクトにおける研究の一環として行った。

参考文献

- [1] NTT コミュニケーション科学研究所監修，日本語語彙体系，岩波書店，1997。
- [2] 葛谷紳，渡部広一，河岡司，“誤認識データベースを用いた単語音声認識方式”，信学技報，NLC2004-35，pp.71-76，2004。