

並列認識と概念知識を用いた単語の音声認識率向上手法

The Improvement Technique in the Rate of Speech Recognition of the Word
using Parallel Recognition and Conceptual Knowledge

葛谷 紳 目片 小百合 渡部 広一 河岡 司
Shin Kuzuya Sayuri Mekata Hirokazu Watabe Tsukasa Kawaoka

*1同志社大学工学研究科知識工学専攻

Department of Knowledge Engineering and Computer Sciences, Graduate School of Engineering, Doshisha University

As the technique of the improvement in the rate of recognition in voice word, the improvement technique in the rate of recognition by parallel processing majority judging and the improvement technique in the rate of recognition by the relevance between the words using conceptual knowledge are proposed, and an experiment shows the validity.

1. はじめに

近年、ロボットの研究は進み、ロボットが人間のパートナーとして活躍する日も遠くないだろう。現在のコンピュータにおける情報の入出力は、主にキーボード、ディスプレイを用いて行われている。しかし、コンピュータが知能ロボットとしての役割を果たすためには、音声によるロボットと人間の会話システムの実現が必要である。本研究では、市販の音声認識ソフトを利用し、人間の音声認識方法を参考にしながら、音声認識システムの精度向上を目指す。

2. 音声認識ソフトの検証

利用する音声認識ソフトは、最終的に知能ロボットに組み込むことを想定しているため、以下の条件を満たす必要がある。

- エンロール（ソフトが用意した文章を読み上げ、個人の声の特徴や喋り方の癖などの傾向を分析・登録する）が不要
- 開発キットが提供されている

本研究では、これらの条件が満たされている東芝音声認識ソフト LaLaVoice 2001[東芝 01] を利用する。

LaLaVoice 2001 には、LaLaVoice SDK (Software Development Kit) という、LaLaVoice の音声エンジンを利用して音声機能をアプリケーションに組み込むための開発ツールがある。音声合成・認識機能を実装するための ActiveX コントロール、音声合成・認識エンジンの属性設定のための各種ユーティリティが提供されている。

3. 単語認識と連続音声認識

LaLaVoice SDK の ActiveX コントロールの内、音声認識コントロールが単語認識を、連続音声認識コントロールが連続音声認識を提供する。

単語認識とは、「夕日」や「見る」のような単語を認識することで、単語登録を行った単語のみ認識できる。単語登録とは、音声認識するための単語のリストを予め辞書として、テキストファイルに記述しておくことである。このファイルをアプリ

ケーション実行時にコントロールに追加することによって、音声が入力された際、登録単語の中から第一候補を認識結果として出力する。一方、連続音声認識は、「夕日を見る」のような文章の認識を行う。この場合、コントロールが所持する辞書を用いるので、ユーザは辞書を変更できない。

単語認識と連続音声認識の単語認識精度を比較したところ、前者は約 90%、後者は約 60% の認識率であった。これは、連続音声認識は、辞書を変更出来ないため、膨大な数の単語から認識候補を選んでいることが原因である。よって、連続音声認識コントロールではユーザ側から認識精度を向上させるアプローチが困難であるので、単語認識をターゲットに研究を進める。

4. 実験環境

実験を行う上で、提案手法の比較を行うために入力する音声データは、比較対照となるシステムにおいて同一のものを使用する必要がある。そこで、音声入力は、あらかじめパソコンで録音したテストデータを使用することとした。16 ビット、48KHz、PCM でサンプリングを行い、別のパソコンから録音したテストデータを出力し、音声認識を行うパソコンに直接入力した。次章から述べる実験はすべてこの方法で行った。

5. 辞書の単語数と認識率の関係

辞書の登録単語数と認識率の関係を見るために、毎日新聞、国語辞書等から自動構築した概念ベース [広瀬 02] から、出現頻度 [Salton 75] の高い上位 1000 語、3000 語、6000 語、9000 語を抽出し単語辞書として、単語認識について調査した。

実験により、登録単語数が減るに従い、認識率が向上していることがわかった。これは単純に確率の問題であり、母集団が減ることによって考えられる。

また実験結果から以下のことがわかった。

- 似た発音、同じ音便の単語、単語の長さが同じである単語は、誤認識される可能性が高い。
例)「イヌ」と「イス」、「ヒゲ」と「いいえ」
- 認識精度は、雑音、マイクと口の角度や距離、パソコン、サウンドボード、マイクなど音声認識に関わる機器の性能に影響を受ける。
- 1つのファイルに登録できる最大登録単語数は 9997 語である。

連絡先: 葛谷 紳, 同志社大学工学研究科 知識情報処理研究室,
〒 610-0394 京都府京田辺市多々羅都谷 1-3, :0774-65-6944

6. 並列処理による多数決手法

実験より、同じプログラム・同じ入力であっても、異なる認識結果が出力される場合があることが分かっている。そこで、音声認識プログラムを複数同時に立ち上げ、それぞれが出力した認識結果を認識候補として多数決を取ることで、音声認識率を向上させるシステムを提案する。

音声認識プログラム1プロセスによる単語認識と、5プロセスによる多数決処理の単語認識を比較し、さらに、登録単語数を3000語・6000語・9000語の三種類による比較を行った。ここで扱う単語辞書は、概念ベースの出現頻度情報の上位から、それぞれの単語数分を抽出した単語からなり、各プロセスはそれぞれ同じ単語辞書を持っている。認識対象は辞書に含まれる300語である。

実験結果(図1)より、5プロセスによる多数決処理が有効であることがわかる。

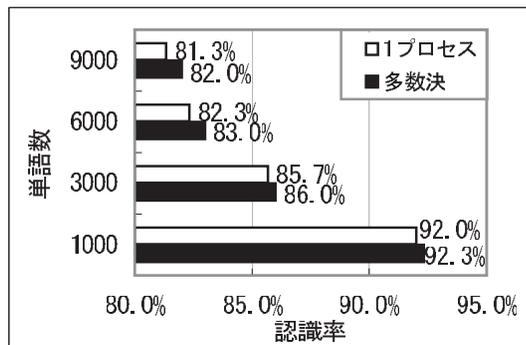


図1: 多数決を用いた認識率

7. 関連度を用いた音声認識システム

単語認識を利用して、さらに複数単語を認識する手法を提案する。これは、今後、会話などの複数単語からなる文章を認識することを考慮している。

1単語の認識率が p だとすると、単純に n 単語の認識率を見た場合、同時に n 語が正解する確率は p^n である。そのため語数が増えるにしたがって認識率が低下してしまう。

そこで、2語入力において、その2語がどれほど関連性があるかについて調べ、その2語について関連性が高ければ、それを認識結果として出力することとした(図2)。このとき、2語間の関連性を定量化する技法として関連度計算方式[渡部01]を使用することができる。

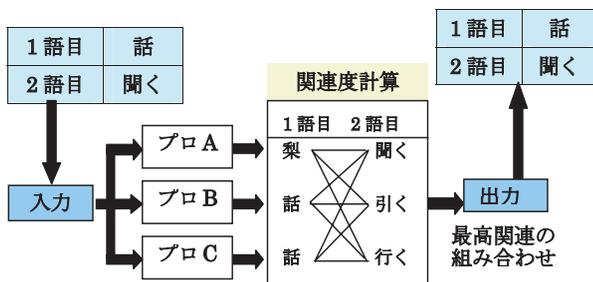


図2: 関連度計算を利用した複数単語認識

実験にて、音声認識プログラム単独と関連度を用いたシステムの単語認識率の比較を行った。関連の深い2単語のセットを100セット人手で作りセットで正解・不正解を判断する。

この実験より、関連の深い2単語を入力した場合、関連度計算を用いたシステムは、2~6%の認識率向上を得た(図3)。ただし、3つのプログラムで実行している限り、この方法が使えるのは、同音異義語を最大で2つ持つ単語までである。

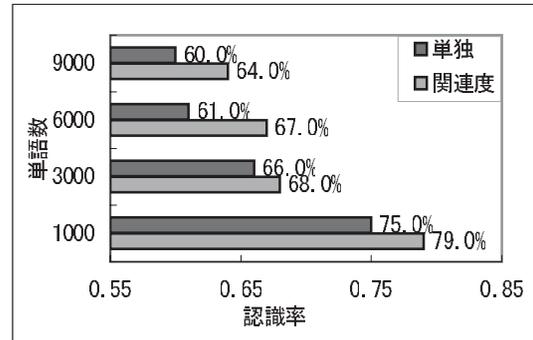


図3: 関連度計算を用いた認識率

8. おわりに

前半では並列多数決処理を用い、後半では関連度計算を用いて、単語の音声認識率向上を目指し、どちらも認識率向上を達成することが出来た。

また、関連度計算を用いることによって、意味を考慮した音声認識に一步近づいた。ただし、これは関連の深い単語を入力した場合のみ通用する。文章を構成する単語同士には様々な関係があるので、そのすべてに対応するにはさらなる工夫が必要である。

会話に適する音声認識を実現させるためには、今回の研究で用いたように、文章の意味やその場の状況も認識の判断材料の一つに加えていくことが、より知的な音声認識に近づく為には必要不可欠である。

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

- [東芝 01] 東芝: LaLaVoice2001, <http://www3.toshiba.co.jp/pc/lalavoice/index.j.htm>, 株式会社 東芝, 2001.
- [広瀬 02] 広瀬幹規, 渡部広一, 河岡司: 概念間ルールと属性としての出現頻度を考慮した概念ベースの自動精練手法, 信学技報, TL2001-49, pp.109-116, 2002.
- [Salton 75] G.Salton, A.Wong, C. S. Yang: A Vector Space Model for Automatic Indexing. Communications of the ACM, 18(11), 613-620, 1975.
- [渡部 01] 渡部広一, 河岡司: 常識的判断のための概念間の関連度評価モデル, 自然言語処理, Vol.8, No.2, pp.39-541, 2001.