

音声認識における周囲騒音の影響の除去 (中間報告)

— 音声認識の自動化機械への応用 —

小田原 幸生

機械電子部

Method of eliminating adverse effect due to Surrounding Noise in Speech Recognition (interim report)

- Application of Speech Recognition to Automatic -

Yukio ODAWARA

Mechanics & Electronics Division

1. 緒 言

高齢化の進展に伴い、福祉の充実や介護者の人手不足対策の一環として、人が命令する言葉を理解し、言葉で応答して働く、人との会話を前提とした自動化機械の開発が望まれている。そこで、平成9年度から音声認識技術の調査を行い、音声認識の障害になる周囲騒音の影響の軽減について研究を開始した。また、音声認識等の自動化への応用のためマイコン開発システムを導入し、開発環境の整備を行った。

2. 音声認識の評価試験

2.1 音声認識ボード

音声認識には、不特定話者に対応した音声認識LSIを搭載した音声認識ボードを用いた。仕様をTable 1に、写真をFig. 1示す。

音声認識の基本処理は全て上記ボードの中で行われる。このボードにマイクを装着し、パソコンのスロットに入れる。そして、BASICやC言語のプログラムを用い、音声認識する単語の辞書ファイルをJISコードのテキストファイルで登録した後、認識の開始、結果の取得などのコントロールを行う。

2.2 評価試験の結果

音声認識における騒音の影響の調査は、イチ〜ジュウ(数)、ハイ、イイエ、オン、オフ(以上)の14通りの単語について、騒音がある場合とない場合での音声認識を行った。被検者は筆者を含む6名で、各自がそれぞれの単語を1回ずつ発声し、正誤を判定した。発声のレベルは口元を騒音計のマイクから10cm離れた状態で、およそ80dBであった。結果は、騒音のない場合は認識率83%で、特に“オフ”の識別が難しく、また個人差もあった。次に騒音としてTVのアナウンサーの音声がある場合(騒音ピーク70~80dB)について試験し、認識率80%前後が得られた。騒音がある場合で認識率の低下が小さかったのは、話者が次第にマイクを口もとに近づけ、発声を

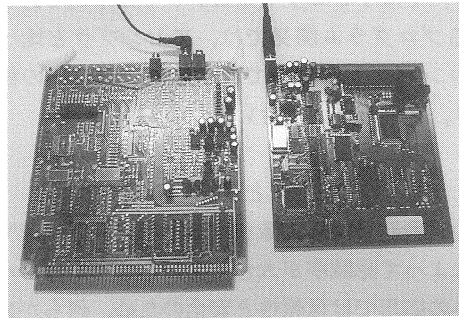


Fig. 1 音声認識ボード (右)

パソコンとのインターフェース (左)

Table 1 音声認識ボードの仕様

型 式	AUDIO-98VOICE RA		
機 能	音声認識, 音声合成		
メーカー名	有限会社スカラベ		
音声認識LSI	NEC	μ	PD77524
音声合成	〃	沖電気工業	MSM7576
使用するパソコン	NEC	PC-98シリーズ	

(特徴)

- ・事前学習が不要な不特定話者による音声認識が可能。
- ・日本語の読みのテキストファイル作成により、任意の語彙を認識できる。
- ・最大認識単語数 1000 (1単語平均音節数=5)
- ・外部ノイズ対策により騒音環境下での音声認識に対応

明瞭にしたためと思われる。TVの音声で勝手に音声認識を開始することはなかった。

3. マイコンの利用について

3.1 マイコン開発支援装置の概要

音声認識のコントローラとしてマイコンを用いるため、マイコン開発支援装置を使用する。この仕様をTable 2に、写真をFig. 2に示す。(前記の音声認識試験ではパソコンを使用した。)

3.2 プログラム・テスト

マイコンCPUの処理速度の向上と、メモリ容量の増大により、高級言語を用いたプログラム開発が一般的になっている。そこで、C言語によるプログラム開発の基本操作を確認するために、

①10組の整数データの並び替え(昇順)

②3組の実数データの浮動小数点乗除算

についてプログラムを作成し、コンパイル後の実行モジュールの大きさと実行速度を求めた。CPUは8ビット系のKL5C80A16(川崎製鉄(株))、C言語はLSI-C80を用いた。この結果をTable 3に示す。

浮動小数点演算は汎用コンピュータと比べ1000倍位の時間がかかり、やはり8ビット系CPUでは不利である。C言語によるプログラム開発では、アセンブラと比べプログラムがコンパクトに記述でき、他のCPUへの移植が容易であり、メリットが大きい。

4. まとめ

音声認識の試験結果から、識別する単語、周囲騒音や話者の慣れによって認識率が大きく変わることが分かった。自動化への応用では信頼性と安全のため、何らかのバックアップが必要である。今後、非正常性騒音の除去機能を確認し、さらに音声認識の判定で人工知能(AI)的手法の応用について検討する。

参考文献

- 1) 音声認識LSI μ PD77524データ



Fig. 2 マイコン開発支援装置

Table 2 マイコン開発支援装置の仕様

型式	PARTNER-ET model 20
接続方式	ROM型
メーカー名	京都マイクロコンピュータ(株)
ポートCPU	Z80, KC80シリーズ, Z80をコアとするASIC (ハードウェアは命令長16ビットまで対応)
開発言語	アセンブラ PROASM-II C言語 LSI-C80
パソコン	FMV TIII20
OS	ウィンドウズ95
導入年	平成9年度 (昭和60年度に導入した in-III(ソフィア・システムズ)の後継機)

Table 3 プログラム・テスト

処理内容	実行モジュール・サイズ(バイト)	実行時間(ms)
10個の整数データの並び替え(昇順)	346	2.9
実数データ x, y, zの浮動小数点乗・除算 (x * y / z)	1119	14

※CPU KL5C80A16 クロック 動作クロック 19.688(MHz)

ROMアクセス 1ウェイト挿入