

音声電卓の言語理解技術について

和田 平司 三角田 秀実

On The Language Understanding Technology of Voice Calculator

Heiji Wada Hidemi Misumida

(所属なし)

キーワード オペアンプ技術 言語理解技術 計算機 デジタル回路

あらまし 我々は音声電卓の言語理解技術について①特徴抽出②選択について基礎的技術の研究及び実験を行なったのでここに報告する。

我々の目的は安価でシンプルな音声電卓を開発することにある。そこで基礎的な(あ)～(お)の母音のベクトルの抽出と選択について実験を行った。

1. はじめに

我々は音声電卓を安価に試作するために言語理解技術のうち①特徴抽出②選択の基礎研究を行なった。

①特徴抽出②選択について基礎的な(あ)～(お)の母音の言語についてベクトル化を図ることを試みた。これにより安価に音声電卓が可能になる。

特徴抽出、選択について我々はアクティブフィルタを用いて言語のベクトル化を図ることを試みた。

簡単な回路を附加することにより、メモリーへのダイレクトアクセスが可能になる。

2. 本論

2-1) 標本化定理について

波形の最大周波数の 2 倍を超えた周波数で標本化すれば、完全に元の波形が再構築される。

すなわち原信号に含まれる最大周波数成分を f_m とする。 $2f_m$ より高い周波数 f_s で標本化した信号は、ローパスフィルタで高周波数成分を除去することによって原信号を完全に復元できる。

そして、A/D 変換及び D/A 変換回路も必要になる。

2-2) ベクトル化について

我々はバンドパスフィルタを用いてベクトル (v_0, v_1, v_2, v_3) を構成することを試みた。

我々はベクトル (v_0, v_1, v_2, v_3) を構成する為にバンドパスフィルタを用い、出力段に Amp を設けて、デジタルパルス化を図り、音声を“0”, “1”のパルスを作り、ベクトル (v_0, v_1, v_2, v_3) を作ることにした。バンドパスフィルタの中心周波数 f_0 はチャンネル 1 $f_{01}=1000\text{Hz}$ 、チャンネル 2 $f_{02}=2000\text{Hz}$ 、チャンネル 3 $f_{03}=3000\text{Hz}$ 、チャンネル 4 $f_{04}=700\text{Hz}$ とした。

2-3) ベクトルの回路図

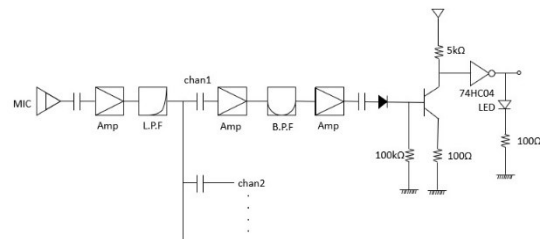


図 1 音声抽出のベクトル化のブロック・ダイヤグラム図

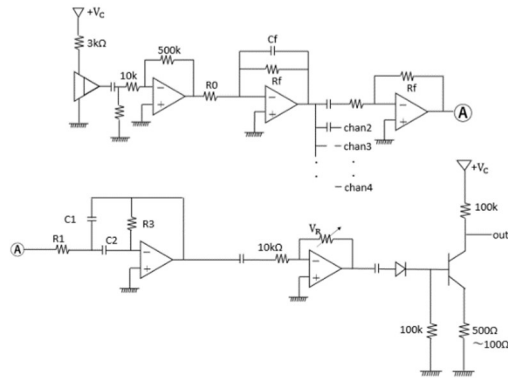


図2 音声抽出のベクトル化の回路図

2-4) ベクトル化について その3

周波数と出力電圧の関係。基礎的な(あ)～(お)の母音のベクトル化を図ることを試みた。

出力電圧と各音声のベクトルの表を表1に示す。

表1 出力電圧とベクトル (v_0, v_1, v_2, v_3) の理論点灯表

channel \ 発音	あ	い	う	え	お
1000Hz	● 1200		● 1300		
2000Hz		● 2310		● 1950	
3000Hz		● 2830 fm			
700Hz					● 860

表1-1 男性.m

channel \ 発音	あ	い	う	え	お
1000Hz	●		●		● 1060
2000Hz	● 1600		● 1500	● 2430	
3000Hz		● 2830 fm			
700Hz		●			

表1-2 女性.fm

channel (Hz) \ 発音	あ	い	う	え	お
chan1 (1.0k)	● m		● m		
chan2 (2.0k)				● m	
chan3 (3.0k)		● fm			
chan4 (0.7k)	● m	● m	● m	● m	● m

表1-3 実験値

3. 実験及び考察

3-1) 実験は図2の回路の出力に74HC04のインバータを取り付けてLEDを点灯させることにした。

発音は(あ)～(お)の母音を発音することによって、表1のベクトル化を図ることができた。

これはLEDの点灯により確認した。また各出力段の波形の例を下の図3-1、図3-2、図3-3に示す。



図3.1 chan1 (あ)の波形



図3.2 chan2 (え)の波形



図 3-3 chan4 (お) の波形

図 3-4 chan3 (い) は省略する

注) (う) の発音も省略する。

3-2) 各 channel の B.P.F のレベルを母音と channel に対比させた。

chan1	あ	≒1000Hz	} 3-1
chan2	え	≒2000Hz	
chan3	い fm	≒3000Hz	
chan4	お	≒700Hz	

3-1 の (あ) ~ (お) の発音 (フォルマント) と周波数の関係とした。

そして、3 段目の出力段のレベルを可変抵抗として Level 調整を行なった。

母音と周波数は以下の表 3 で与えられる。

Hz

	男性		女性	
	F1	F2	F1	F2
あ	790	1270	980	1640
い	280	2310	340	2830
う	320	1300	390	1550
え	520	1950	620	2430
お	510	860	650	1060

表 2-1 母音とフォルマント

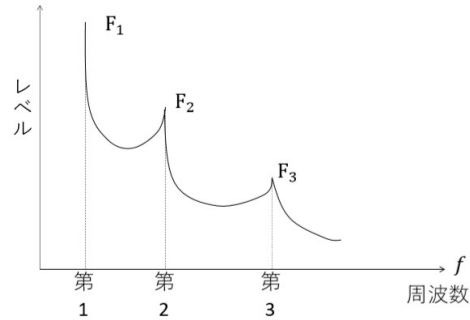


図 3 母音とフォルマント

第 1 フォルマントの周波数 (基音) は B.P.F にて減衰しているため、F₂ の第 2 フォルマントを測定できるとした。が、(お) の発音が第 1 フォルマントと一致し、(あ) ~ (お) の発音に対して、channel 4 を点灯させる。

3-3) オシレータにより、B.P.F の中心周波数 f_0 のレベルを調整して、客観的に測定する必要があるが。

我々は個人差による声の特徴が異なる誤差を踏まえて、母音のフォルマント F₂ でレベルを合わせた。

然し、今後の課題として、B.P.F の Q を鋭くすることにより、オシレータによるレベル調整を行うことを検討している。今回の論文に於いては、確証を得た所で報告を行なった。

然も、今後の試作機の製作に当たって、B.P.F の channel を 4channel から 8channel へと増すことにより、確実に音声電卓の計算が実行される。

即ち、試作機に於いて、8channel の B.P.F の後のベクトル化と、Q の鋭さによりそれは検討されることとする。

3-4) 理論値と実験値の考察

表 1 に理論値の LED の点灯表を示した。但し、表 1-1 が男性 m、表 1-2 が女性 fm である。実験値の LED の点灯表を表 1-3 に示した。

実験値は今回のフォルマントで調整したので、第 1 フォルマントは 300 ~ 800Hz であり、channel 4

の中心周波数 700Hz の LED が、(あ) ~ (お) の発音に対して全て点灯した。

理論値は (お) の LED のみが点灯するはずであるが、第 1 フォルマントのレベルが、全ての発音 (あ) ~ (お) に対して強いからである。

が此れは B.P.F の Q と出力数のレベル調整により改善される。

発音 (う) については chan4 より chan1 の方が LED が点灯すると思われるが、男性の第 2 フォルマントのばらつきにより chan4 も点灯すると考えられる。

発音 (い) と (う) は chan4 の LED は理論的には点灯しない。

フォルマントの表 2-1 はおおざっぱな目安であり、だいたい所、理論値と実験値の一致がみられた。

そこで、(あ) ~ (お) の特徴抽出はできたと考えられる。

選択として、出力数の発音 (あ) ~ (お) のデジタルパルスの理論値と一致するデータコードの場合のみデータをラッチし出力する回路を設計できるので、何回か発音をくり返すことになる。

これは、初期の音声認識の課題であった。

今回の段階では発音を理論値と一致するまでくり返す必要があり、試作機において、改善する。

3-5) 今回の実験を行った基板の写真を写真 1 に示す。

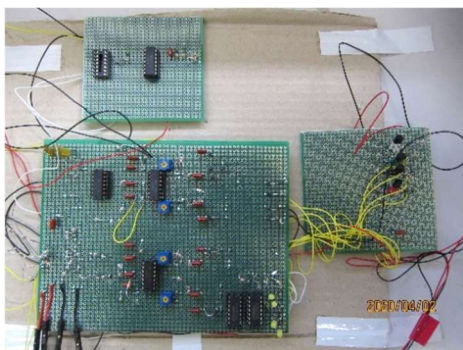


写真 1 実験の基板

4. 結論

我々は音声電卓の言語理解技術について、安価で回路構成が簡単である①特徴抽出②選択の実験を行なった。母音に対する特徴抽出と②選択は表 1-3 にしめされる通り、旨く 4channel でベクトル化が図れた。参考までに、試作機に対する表 1 の基礎的な (あ) ~ (お) の母音の言語は、今のところ 4channel でのベクトル化に於いて、多少困難なところがあるが、我々は B.P.F による channel を 8channel にすることと、最大の条件である B.P.F の Q を鋭くする必要がある。このことにより、F₁ のフォルマントを無視でき、F₂ のフォルマントを抽出可能とすることができる。然も回路を安価に設計することが出来る。

今回は表 1 による母音の特殊抽出の①抽出②選択が可能となったので報告した。

今後の課題はサンプルを増すと同時に、表 1-1、表 1-2 の①特殊抽出の選択を行なうことが必要になるが、その理論的解決について上述してある。また実際に試作機を作ることにある。

参考文献

- (1) "音声言語理解のこれまでとこれから"
電子情報通信学会誌 Vol.101 NO.9
PP.875~P.913、2015
- (2) 河原 達也 編著
"音声認識システム"
情報処理学会編