

命令語で動作する車型ロボット — 音声識別回路の応用 —

Car Type Robot who moves in the Instruction Word
- Application of the Voice Discrimination Decision Circuit -

金野 茂 男
KINNO Shigeo

1. はじめに

簡易な電子回路での音声認識システムの構築を試みている (1)、(2)、(3)。母音に関しては、6号⁽⁴⁾に至ってようやく安定し、かつより精度の高い判定を行うシステムを構築するまでになったと思っている。6号の使用においては、それ以前のシステムと比較すると、少しやっかいな手続きが必要となってしまったが、その見返りとして母音の判断精度の大幅な向上が行えたと考えている。

今回、6号のシステムを拡張し、制御者(著者)の発する幾つかの命令語で、動作制御の出来るデモ用小型ロボットシステムを制作した。写真1にその外観を示す。ロボットは3輪車型自走ロボットである。大きさは円形の基台が直径約14cmである。前輪の2つの車輪はギア付きマイクロモーターで個別に駆動される。後尾にある尾輪は自在型である。車体に立てたアンテナ棒の先端に、コンデンサマイクが取り付けられている。このマイクに向かって「進め」、「止まれ」「後」などの命令語を発する。ロボットはこれらの命令を解読すれば、それに従った動作をするシステムとなっている。

「どの程度正しく応答するのか? 100%か?」という質問が当然出てきそうである。100%とはとても言えないが、システムの使用方法に慣れ、発生音が安定して出せるようになれば、命令語でロボットは結構正確に動くと言っておこう。また本システムを理解できれば、連語からなる多数の命令語や、指示語で本ロボットの多様な動作制御を行うことは十分に可能であることも述べておこう。

以下で、6号の応用として制作した自走ロボットで、子音を含んだ命令語をどのように解析し、判断させ、対応させている制御を行わせているの

かについて紹介しよう。

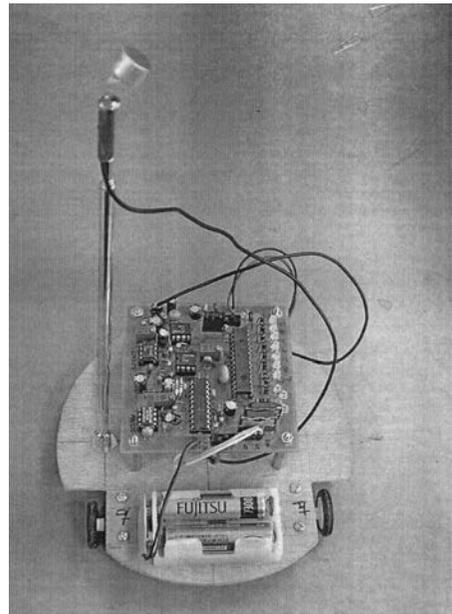


写真1 制作した車型ロボットの外観

2. 設計・制作

初期設定の面倒さはさておいて、強度相関法を用いた6号での母音の判断成功率は結構高いものがある。では、6号は子音の分析にも使えるのであろうか? 試しに適当な子音を6号に向かって発音し、その結果をLEDで視認してみた。日本語の子音は「接頭音+母音」から成り立っている。6号では、接頭音部分の分析使用は全く持っていないので、当然ながら接頭音を分析することが出来ない。が後半の母音は識別できるはずである。結果はまさにその通りであった。子音が持っている後半の母音は、6号は結構正確に識別することがわかった。とはいっても、母音だけの場合と比較すると、誤答率が高くなっている。この理由

や考察については、後述しよう。とにかく、6号では子音の接頭音の識別は出来ないが、引き続き母音の識別は良くやってくれることがわかった。

単音だけではなく、幾つかの適当な連音も入力してみた。母音部分に対する正答率は結構良いものがあった。これならば、子音の接頭音の識別機能が無くても、6号のシステムを応用して、連語からなる命令語を用いて、ロボットを制御できるシステムが構築できるものと確信した。

図1に、6号をもとにした今回のロボットの電子回路図を示す。著者は、PIC18Fシリーズのクロックパルス源とし40MHzセラロックを今まで常用してきた。だいたい前にも多数個を秋葉原で購入したのであるが、最近どういう訳か、秋葉原でこの周波数のセラロックを見かけることが無くなった。手持ち部品もなくなってきた。18Fシリーズは40MHzまで動作する。セラロックを20MHz等で置き換えてもよいが、そうするとアセンブラプログラム中の多数の箇所パラメータなどの変更が必要となる。これでは大変である。ところで、18FシリーズにはPLL回路が内蔵されていることに気がついた。外付けセラロックを10MHzとしても、このPLL回路を使用すれば、PIC内部でのクロックパルスは $10\text{kHz} \times 4 = 40\text{MHz}$ となり、この40MHzでPICは動作してくれる。PLL回路の使用方法はコンフィグレーションビットの設定で行える。この設定の仕方の詳細は18Fx2のテクニカルノートを参照すること。が、参照しなくても本システムの再現は出来よう。今回の著者の提供しているアセンブラプログラムではクロックはPLL仕様となっている。プログラムの先頭箇所にコンフィグレーションビットの設定部分がある。その箇所を一度は目にしておくべきであろう。著者の先のプログラムと比較するのも良いであろう。

従って、PICへの外付けセラロックは10MHzとなっている点、及びI/OポートCにはモータードライブ回路が付加してある。この2点だけが、6号と異なっている。他は全く同じ回路である。

車輪駆動用モーターとして、超小型精密ギアモーター「栄42D」(発売元 STL JAPAN、価格約3800円)を使用する。このモ-

ーターの仕様を表1に掲載しておく。このモーターは結構高価である。車輪駆動のモーターとしては、一般模型工作用として、ギアと一体になったモーターが安く模型店などで購入(数百円程度)できよう。このギア付きモーターは安い点は好いのであるが、ギアの部分で発生する騒音が、マイクを持っている本システムでは大問題を発生する。ギアからの雑音がマイクから混入し、誤動作すること請け合いなのである。それに対して栄42Dは高価であるが、ギアから発生する雑音は無いとはいえないが、結構小さい。本システムでの使用に当たって、その雑音はほとんど問題にならない程度である。

PICから出力されるのモータードライブ用ロジック信号で、モーターを駆動するために、モータードライバTA7268BPを使用する。モーターの駆動電源をシステム電子回路の電源と共通にすると、モーター駆動時における電源の急激な電圧降下によって、電子回路の動作が不正常になる可能性が大である。それを避けるために、モーター駆動電源は独立した別電源(単3×2本)としている。ただし、確実な電源フィルターなどが使用できれば、電源は1つで賄えるであろう。

発音する命令語として3連語からなるものとして、「止まれ」、「進め」、「後」、「左」、「右へ」を準備することにした。最後の「右へ」は「へ」が無いと2連語となるので、今の場合の3連語に該当しなくなるので1語分の「へ」を付加したものである。表2にこれらの命令語をアルファベットでの表記、そして、それから母音部分を3語母音列として表記した文字列を記載している。5つの3語母音列の文字配列を見ると、先頭の2文字だけでこれらの命令を区別できることがわかる。それが表中の右端の先頭2語文字列の欄である。即ち、上述の命令語をマイクから入力する。システムは、その命令語の先頭2語の母音列だけを判断する。その判断に基づいて、該当する制御動作指令をモーターに送ればよいことがわかって。

命令語を子音列でもって識別するのではなく、子音に付属している母音列で識別する。少しいかさまな方法のようであるが、命令語、即ち特定の連語群の識別方法の1手段であろう。

表3にPICに書き込んだアセンブラプログラムの本システムに関係した部分だけの簡単なフロー

命令語で動作する車型ロボット — 音声識別回路の応用 —

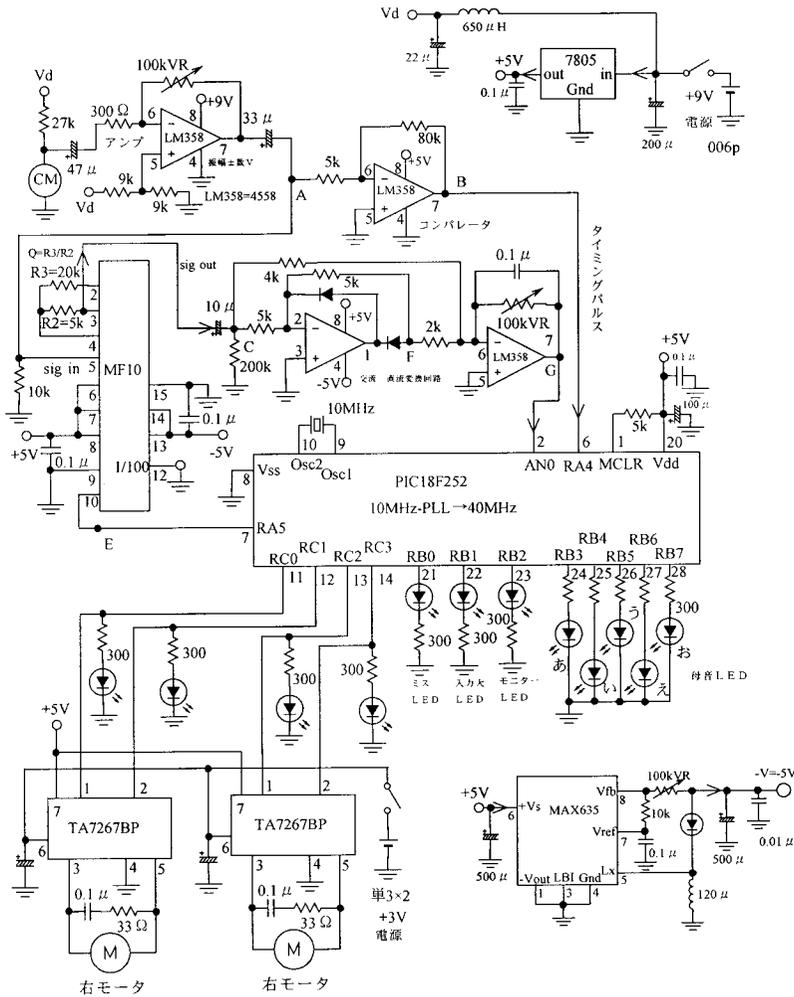


図1 車型ボットの回路図

軸の回転数	約60回転/分
最大トルク	3300gcm
軸径	3mmφ
使用電圧	1DCV~7.2DCV
重量	8.7g
大きさ	長さ34mm×幅12mm×高さ10mm
電流	70mA~400mA

表1 「栄42D」の諸元

命令語解読表	3 語母音列	先頭 2 語母音列
(1) 止まれ t o - m a - r e	→ o - a - e →	o - a
(2) 進め s u - s u - m e	→ u - u - e →	u - u
(3) 後 u - s i - r o	→ u - i - o →	u - i
(4) 左 h i - d a - r i	→ i - a - i →	i - a
(5) 右へ m i - g i - h e	→ i - i - e →	i - i

表 2 命令語の分解

チャートを示している。プログラムの大半は 6 号のそれである。6 号のアセンブラプログラムを母親として書き上げている。

一音が入力すると、コンパレータからトリガ信号が形成され、入力音量が安定するまで若干の時間遅延をさせた後、PICは周波数分析動作に移行する。6号ではこの遅延時間を約 200ms としていた。本号では、接頭音があるので、その部分の飛び越しのことも考慮して、遅延時間長は約 400ms としている。この遅延時間長の設定は可変であるので、アセンブラプログラム中で、必要かつ確実な長さに設定すればよい。

6号では、母音の入力→周波数分析→強度相関法での判断→母音LEDへの表示→少しの待機→次の母音の入力。後は繰り返しの手順となっており、1サイクルは1秒～2秒かかっている。むしろかけていると言うべきである。この時間は必要に応じてアセンブラプログラムの該当箇所を変更することで短縮できるから。6号でのこれらの手順をそのまま採用しているの、本システムでは、命令語は1語1語がゆっくりと入力されることを要求している。早口では正常動作は全く期待できないし、通常の早さでの発音速度でも速すぎると思われる。どの程度の発音速度がよいのかは、本システムを動作させ、試験して見るのが一番良いであろう。

命令語を入力すると、一語、一語毎に母音LEDが点灯する。この母音LEDの点灯状況からシステムが正しく入力音を判断しているかを視認できよう。

6号でも母音の分析はもっと短時間でできることは紹介している。アセンブラプログラム中で関係する箇所を変更するだけですませることが出来る。本プログラムでも同じ箇所を同じく変更するならば、通常の発音速度だけではなく、早口の命

令語でも正常な判断を行わせることが出来ると思われる。現時点では、特に早くする必要を感じていないので、この件は著者は試してはいないが。

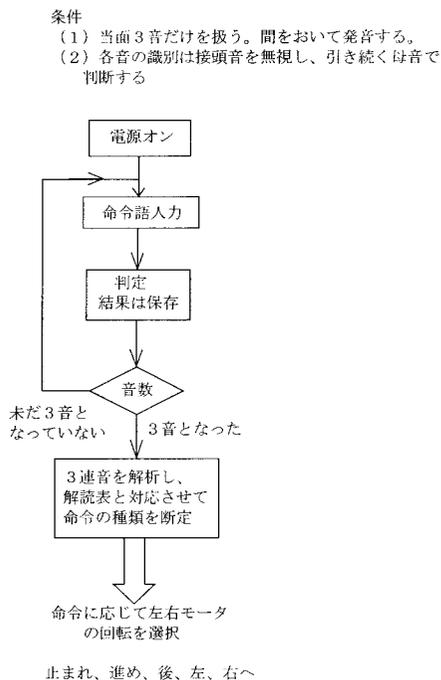


表 3 命令語判断のフローチャート

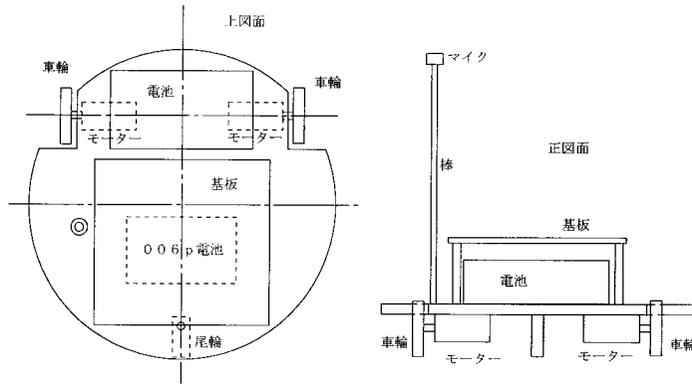


図2 台車の図面

図2に台車の図面を示している。各箇所の寸法は省略しているが、台車の直径の約14cmをめどにすればよい。

写真2、3に本ロボットをより詳細に紹介する。マイクを台車の上方高く位置させるために、ロッドアンテナを利用している。写真2の左に位置する金属棒がロッドアンテナである。都合がよいことにこのロッドアンテナは伸縮自在型である。台車の上の真正面に、モーター用の乾電池があり、その奥に、電子回路用電源の006P乾電池が位置している。

写真3からモーター及び車輪の配置状況が見て取れる。ギアモータはL字金具で基台に取り付けている。尾輪の自在車、及び2つの車輪は車などの模型ジャンクから適当なものを探し出したものである。

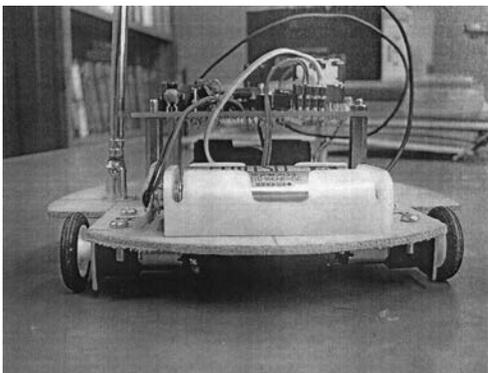


写真2 正面から

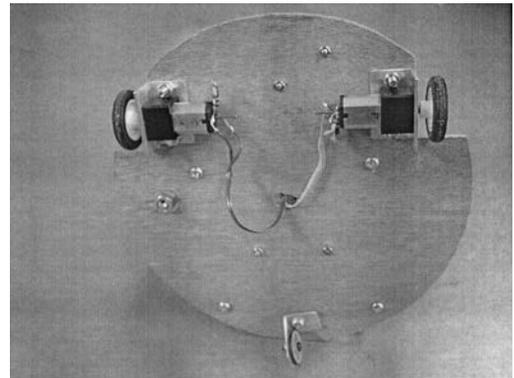


写真3 台車の底

3. システムの使用方法

電子回路を製作し、台車を試作して組み合わせ、全体を完成させる。PICには提供しているアセンブラプログラムを読者の発音母音に合うよう強度相関の比較演算部分を変更したプログラムとして、書き込む。この変更方法、手順については6号で紹介しているのでそれを参考とすればよい。従って、6号のシステムを再現していれば心強い。

- (1) まず電子回路の電源スイッチだけをオンとし、マイクから母音を入力してみよう。
- (2) 該当する母音LEDが正しく点灯するように、音声の音程などに気をつける。
- (3) ほぼ確実に母音の判断が出来ることを確認したら、一端電源をオフとする。

これから命令語での動作試験に入る。発音した命令語を解読できないで暴走する場合があるので。広いところで試験するか、あるいは台車を簡単に手で処理できるようにしておく。

(4) 「進め」を命令してみよう。成功すれば、ロボットは前進する。

(5) その状態で「止まれ」を命令してみよう。成功すれば、ロボットは止まる。

(6) 後は繰り返しである。

命令語の発音が不明瞭であったり、中途半端であった場合、その発音を取り消す必要がある。短い発音例えば、舌打ち音、手叩き音のような短時間の音がシステムに入力すると、システムは雑音として処理するようにしている。短時間音を入力させれば、システムはミスをLEDを点灯し、処理を音声再入力待ち状態に復帰させてくれる。これを利用すればよいであろう。

日本語で「橋」、「箸」、「端」は発音では「h a - s i」であるが、それらの違いは抑揚やアクセントで区別される。つまり、通常の連語では音程が途中で変化することが良くあることである。が本システムでは音程が変化しては困る。このことは理解していると思う。命令語の発音は極力努力して、一定(予約の)音程で発音するようにしなければならない。

4. おわりに

少しごまかしの方法で子音を含んだ命令語で、ロボットの制御を行うシステムの例を構築した。今後は、やはり子音を識別したい。が、本システムの延長で、子音の識別が出来そうなアイデアは未だ浮かんでこない。接頭音をどうにかして、各子音列毎に明瞭に分類したいのであるが。接頭音部分のアナログ波形をAD変換する方法がある。これについては今後取り組んでみたいと思うが、すぐやるかどうかは未定である。

以下に、思いつくことを記載する。

(1) 本システム(6号でも)による母音の判断成績は、著者の母音のうちで「あ」、「い」、「う」では極めて成績がよい。が、「え」と「お」の判断は発音に気をつけていないと、時折お互いに混乱して判断する。このことは6号に掲載して

いる母音の周波数スペクトルからも一応は納得できる。スペクトルが似ているのである。

本システムでは連語からなる命令語を母音列として取り扱い、判断を行っている。「え」と「お」の判断が混乱する可能性があることを前提とし、そのような場合に対応するようにプログラムの判断アルゴリズムに修正を加えておけば、本システムはより正確に命令語に応答できるようになる。

例を示そう。「止まれ」の先頭2語母音列は「o - a」である。先頭の「o」を「e」と判断しては、現判断基準ではロボットは何の判断も下せない。「a」は2番目にある。表2から「a」が2番目にある命令語として、「左」(i - a)もある。「i」は正確に判断できるので、1番目が「i」で、2番目が「a」ならば、入力した命令語を「左」と判断させ、1番目が「i」ではなく、2番目が「a」ならば、入力した命令語を「止まれ」と判断させることが出来る。

(2) 現システムでは3連語の命令語だけに対応させているが、4連語、5連語、それ以上の連語に対しても拡張できる。ということは、多種の命令語や指示語で、ロボットに多様な動作を行わせることが出来ということである。

(3) 6号でも触れていたが、6号の拡張である本システムでも、使用者毎に周波数スペクトル $f_1 \sim f_{10}$ の間の強度相関関係を見だし、それに基づいて使用者毎の判断アルゴリズムをアセンブラプログラムに書き込まなければならない。この手間をどうにかして軽減したい。

5号までのように、システム使用始めに、基準となる母音データを登録しておく方法がある。母音の入力により、各母音毎に $f_1 \sim f_{10}$ までのスペクトル強度データが得られる。5号まではこのデータをもとに最小2乗法を適用していた。強度相関法を適用するには得られた $f_1 \sim f_{10}$ の間の強度相関関係を見出す必要がある。この探査をプログラムの実行で行い、得られた強度相関関係データを保存しておく。その後、システムを実測モードとして動かし、入力してきた母音の強度相関関係データと保存してある基準母音データの間での比較検査から、判定を行わせればよい。この方法をやるかどうか今のところ決めてはいな

参考文献

- (1) 「音声識別 — 母音あ、い、う、え、お —」、金野茂男、小山高専電子制御工学科、2005年、11月、URLで公開済み。
- (2) 「音声識別—4号 —母音あ、い、う、え、おの識別—」、金野茂男、小山高専電子制御工学科、2006年1月、URLで公開済み。
- (3) 「音声識別—5号—母音の識別性能の更なる向上に向けて—」、金野茂男、小山高専電子制御工学科、2006年3月、URLで公開済み。
- (4) 「音声識別 —6号 —新たな識別方法の採用—」、金野茂男、小山高専電子制御工学科、2006年、3月。

著者のURL

<http://www.oyama-ct.ac.jp/D/kinnoken>

