

音声認識環境制御装置の開発

Speech Recognizing Remote Control for Disabled Person

○ 富山県高志リハビリテーション病院 大島 淳一

キーワード: 音声認識、環境制御、自作

1. はじめに

環境制御装置は家庭電化製品の自立操作を可能にし、重度麻痺者の在宅生活の質を大きく改善させるものと考えられ、すでに製品として長く流通している。しかし比較的高額な導入費用に対する経済的補助制度がごく一部の自治体に限られ、また家電製品の買い替えでの設定変更などで、技術サポートが十分ないなどの理由で普及が進んでいない。またユーザも環境制御装置を利用する生活を体験する機会が得られず、導入の成果をイメージしにくいいため普及が進まない等の意見もあるらしい。

一方、音声認識は十年以上前にパソコンソフトとして市販され高い実用性が示され、今では Windows の標準機能となった。スマートフォンでは音声検索アプリなどへ応用が広がっているが、こちらも福祉機器関係への応用が期待される割には普及していない。

これらは一見、福祉機器関係の技術開発が近年停滞しているように見える。しかし現実には、必要な技術はすでに机の上に届いているのに、それを必要とする人に届けるための、人材、技能、知識が医療福祉教育現場で極端に欠乏していると考えるのが正しい。そこで今回、日曜大工や家庭料理のように、身近にある機材と無料で手に入る材料と、安価な機材を利用して、音声認識環境制御装置を作る方法を紹介する。

2. 準備

2.1 ハードウェア関係

モニタ、マイク、メモリ、記憶装置、CPU など音声認識に必要な一式を安価に入手するには WindowsXP パソコンが適している。また後述するソフトウェアのこれまでの蓄積を利用できる利点もある。家電製品のリモコンと同じ赤外線信号を発生する装置として、「伝の心」に採用され実績もあり、家電製品リモコンの情報が入手利用可能な、テクノツール(株)の「なんでも

IR」を使うことにした。



図1 なんでも IR(右手前)

2.2 ソフトウェア関係

WindowsXP パソコンで音声認識を行うには、日本語音声認識エンジン・開発キット、Speech SDK5.1 と日本語を使用する言語パック Speech SDK LangPack を使った。これらはいずれもマイクロソフトのサイトから無償で入手できる。また、WindowsVista以降は標準でインストールされているので特にインストールする必要はない。これらのソフトは、不特定話者のエンロールなし発声でも高い認識精度を示し、技術進歩がうかがわれる。

さらに Speech SDK5.1 に対応した、任意の言葉の認識を行うライブラリは、ラジコン自動車の音声制御用に kilala 氏がすでに開発し公開している¹⁾。(SKSpeechLib.dll) このライブラリを含む、VisualBasic.NET のサンプルソフトを使用すると、各種ソフトのインストール確認から音声認識のセットアップまで円滑に作業を進めることができる。

さらに、VisualBasic.NET(無償の Visual Studio Express で可)を使えばこのサンプルソフトに手を加えて独自の機能を追加することもできる。

```

<GRAMMAR LANGID="411">
  <RULE NAME="IRControl" TOPLEVEL="ACTIVE">
    <L>
      <P><RULEREF NAME="IR"/></P>
    </RULE>
  <RULE NAME="IR">
    <L PROPNAME="IRCMD">
      <P VAL="1">/なんでもIR/なんでもあいある/ナンデモアイアール;</P>
      <P VAL="2">/テレビ/てれび/テレビ;</P>
      <P VAL="3">/電源/でんげん/デンゲン;</P>
      <P VAL="4">/音大きく/おとおおきく/オトオオキク;</P>
      <P VAL="5">/音小さく/おとちいさく/オトチイサク;</P>
      <P VAL="6">/チャンネル/ちゃんねる/チャンネル;</P>
      <P VAL="7">/あー/ああ/アア;</P>
    </L>
  </RULE>
</GRAMMAR>

```

図2 グラマーファイルの例

```

Label10.Text = e.ItemName(1)
End If
If e.values(0) = 1 Then
  Call Nandemo()
End If
If e.values(0) = 2 Then
  SendKeys.SendWait("%2")
End If
If e.values(0) = 3 Then
  SendKeys.SendWait("%q")
End If
If e.values(0) = 4 Then
  SendKeys.SendWait("%o")
End If

```

図3 VB.NET コードの例

3. 音声認識の設定ファイルと実行コード

図2の、グラマーファイルはライブラリが音声認識の際に返す値、返す文字列、認識する文字列をひらがなとカタカナで設定し、XML形式で記述されている。これを変更して認識する言葉を自由に設定できる。

図3にコードの例をしめす。これは kilala 氏のサンプルコードに「なんでも IR アシスタント」の起動と、コマンド送信の部分を加えたもので、VB.NET で記述してある。「なんでも IR」の標準操作ソフト「なんでも IR アシスタント」は、通常マウスで操作するが、キー入力でも同様の操作ができる。そこで、SendKeys.SendWait メソッドを用いて相当するキーを「なんでも IR アシスタント」に送り「なんでも IR」を操作できる。以上の方法により、音声でリモコンが動き家電製品を操作できる。

4. 特徴とこれからの展望

ここで紹介している方法にはいくつかの特徴がある。まず使用する機器のほとんどが身近なごくありふれたものである。次にソフトの多くが無償で提供され、ライセンスなども必要としない。また多くの技術が公開から年数が経ているため相応に安定し、またネットにノウハウの蓄積も存在する。また比較的スペックの低い古いパソコンでも利用の妨げがすくなく。これらの結果と

してコストが低く、また借り物パソコン等で試用する際には機器のやりくりの苦労が少ないなど、実際の現場で役に立つ利点も多い。

SpeechSDK5.1 と SKSpeechLib を利用した音声認識には、市販の音声認識ソフトにはない特徴もある。市販品は一般ユーザを想定し、分かりやすさとシンプルさを考慮して企画設計されている。例えば多くのソフトは、enter キーを入力する音声コマンドは、「エンター」であり、このため「エンター」を発声できないと入力できない。このように市販の音声認識ソフトには、特定の言葉を発声できない不自由があると使用上不便が多くなる。

これに対して今回の方法は任意の言葉を音声コマンドに設定できるので、特定ユーザが随意的に発する識別可能な発声を自由に利用でき、持てる能力をより発揮する助けになるだろう。また今後、技術をさらに進めれば声による操作スイッチとして使えるだろう。

また、認識した音声に対して、人工音声で返事をする自動コミュニケーションへの応用も考えられる。これはごく初歩的な人工知能と位置づけられるが、現状では、一般人と普通の会話を通じて自然なやり取りするにはまだまだ研究や開発が十分ではない。しかし認知症者や小児、児童、その他特殊な事情のある人を相手にする自動コミュニケーションは近年中に実現すると予想される。

5. おわりに

12年前に ViaVoice を使った時も技術の進歩に驚かされた。一方、今回は別のことで驚いた。まず安いことと、認識の能力が比較にならないほど高いこと。さらに技術が広く公開され別用途に応用できること。互換の日本語音声認識エンジン (Julius) がすでに開発されていること。そして最後に、このことがあまり知られていないこと。スマホの音声認識は知っているのに、それを自分で作れることを知らず、問題解決につなげることが出来なかった自分に一番驚いた。

6. 参考

1) kilala: VB.NET で音声認識 (その1) 音声認識ライブラリを作る、
<http://kilala.mydns.jp/def/VBTIPS/005/>