

# RECONHECIMENTO DE VOZ PARA APLICAÇÃO EM CADEIRA DE RODAS

Arlei Barcelos, Fabrício de Carvalho Souza, Jair Moraes dos Santos, Vitor Tomaz de Aquino.

*Curso de Engenharia Elétrica com ênfase em eletrônica da Associação Educacional Dom Bosco - AEDB*

*Estrada Resende Riachuelo, 2535, Campo de Aviação, CEP27511-971.*

*Resende, Rio de Janeiro, Brasil.*

*E-mails: [arlei@csn.com.br](mailto:arlei@csn.com.br), [fabricao.souza@terra.com.br](mailto:fabricao.souza@terra.com.br), [jairsan@terra.com.br](mailto:jairsan@terra.com.br), [vitortomaz@ig.com.br](mailto:vitortomaz@ig.com.br)*

## RESUMO

De natureza documental, o presente trabalho é resultado de um estudo em revistas, livros e artigos em meio eletrônico. O objetivo é expor as origens e as técnicas de reconhecimento de voz, em especial o software IBM Via Voice, para serem implementadas em cadeiras de rodas destinadas às pessoas tetraplégicas ou com dificuldade de movimento. No programa IBM Via Voice, além do reconhecimento, também são utilizadas ferramentas como macros, que são empregadas para fazer a interface com o *Visual Basic*. Isto garante uma maior flexibilidade de programação para o controle da cadeira de rodas através da porta paralela de um computador. Alguns dispositivos micro-processados (Microcontrolador PIC) também serão abordados, elucidando as peculiaridades as características e o funcionamento de cada componente que foi testado durante a elaboração do projeto. Portanto, o conhecimento dos softwares IBM Via Voice, *Visual Basic* e de componentes como PIC, Porta Paralela, *Display*, bem como a utilização conjunta destes, apresentaram-se como aspectos fundamentais para a realização deste intento.

**Palavras-Chave:** *Reconhecimento de Voz. IBM Via Voice. Visual Basic. PIC.*

## ABSTRACT

*The present document is a result of researches on digital magazines, books and articles. The main goal is to explain the source and techniques of voice recognizer, in special the software IBM ViaVoice, used on the implementation of wheelchairs for tetraplegic people or for people with movement's difficulties. Besides the recognition, the program IBM ViaVoice works with some tools like macros to implement and make interface to the Visual Basic program. These characteristics allow the program being flexible for the wheelchair's control by the computer through a parallel port. Some components, such as a microprocessor (PIC Microcontroller), also will be shown, explaining some peculiarities, characteristics and the function of each tested component during the project's construction. Therefore, knowing the softwares IBM ViaVoice, Visual Basic operations, some components like PIC, parallel port, Display and the utilization of all this components together, presents as fundamental aspects for the achievement of this work.*

**Key-words:** *Voice Recognition, IBM Via Voice, Visual Basic, PIC.*

## 1. Introdução

Este projeto trata de estudar e elaborar um sistema de reconhecimento de voz em português para ser implementado em cadeira de rodas com comandos mecânicos, visando à autonomia e o baixo custo de instalação para pessoas que buscam liberdade de locomoção dentro de um ambiente seguro e que não comprometa a integridade física do cadeirante. Obviamente o projeto não tem o objetivo, de tornar um cadeirante (tetraplégico) independente, como sair na rua sozinho. Isto poderia causar mais problemas do que soluções. Espera-se que esta independência venha dentro do seu lar, como ter a possibilidade de locomover-se da sala para a varanda, independentemente de ajuda, e com segurança.

## 2. Desenvolvimento

Para o estudo de reconhecimento de voz, segundo [3] a escolha do método de reconhecimento depende da característica genérica do sinal de fala.

Segundo Ronald et. al [7], é possível classificar os sistemas de reconhecimento de fala pela característica de dependência de locutor. E de acordo com Roe & Wilpon [8], a dependência de locutor é, essencialmente, uma característica resultante do modo como o reconhecedor será implementado. As limitações de reconhecimento e tipos de erros que podem ser gerados são citados por Sun [9]

O programa utilizado neste projeto o programa IBM Via Voice, bem como algumas de suas funções. Para a geração de um código binário, através da saída de uma porta paralela de um computador, foi utilizado o programa Visual Basic, ficando o circuito de controle, a mecânica e a montagem da cadeira de rodas realizado por [1].

### 2.1. Definição do software de reconhecimento

A pesquisa do software a utilizar levou-se em consideração a língua falada, a robustez do reconhecimento e a interface de trabalho com outros programas devido à aplicação deste desenvolvimento ser no Brasil e de acordo com Damasceno [2] o IBM ViaVoice obteve o melhor desempenho e aplicabilidade. Segundo [5] o IBM ViaVoice permite a incorporação das suas tecnologias e funcionalidades em outras aplicações.

### 2.2. API's do IBM ViaVoice

Existem duas principais API's que o IBM ViaVoice utiliza para a aplicação deste projeto que são IBM SMAPI e IBM DMAPI.

Segundo [5] a SMAPI (Speech Manager API) é uma API que suporta apenas funções de reconhecimento de voz. A interface entre o mecanismo de reconhecimento e o software de aplicação é o SMAPI no qual é a interface principal para o mecanismo de reconhecimento do ViaVoice.

Segundo [4] a Dictation Macro API (DMAPI)<sup>1</sup> é uma função na linguagem C que permite desenvolvedores ter acesso a macros de ditado e arquivos dentro da aplicação. A DMAPI é liberada para aplicações de ditado escritas pelo ViaVoice.

Macro de ditado é um atalho para entrada de texto que consiste em um nome e uma ação.

O nome é o comando de voz que é falado para envolver a ação. Quando o nome da macro é falado, a macro de ditado é substituída pelo texto ou ação definida pela macro feita durante a programação.

---

<sup>1</sup> Para maiores informações sobre DMAPI acesse o arquivo no seguinte endereço:  
<http://www.wizzardsoftware.com/docs/dmapi%20reference%20guide.pdf>

### 2.3. Construção do dicionário

Em nossa aplicação foi necessário criar um novo dicionário para a aplicação. Existem diferentes partes na aplicação do ViaVoice que limita o que pode-se dizer. Por exemplo, caso seja necessário utilizar fala contínua utiliza-se aplicações que podem entrar com ditado contínuo, ou seja, programa como o speakpad, microsoft word entre outros. Neste projeto limitamos o reconhecimento para evitar erros de reconhecimento no qual prejudicaria o desempenho do mesmo.

### 2.4. Decodificador do IBM ViaVoice

Esta seção sumariza o processamento de áudio pelo mecanismo do IBM ViaVoice. Algumas informações cedidas por [3] segue abaixo:

- O áudio tem a taxa de amostragem de 22 KHz em tempo real.
- É armazenado em um buffer de memória interna de áudio sons acima de 30 segundos.
- O áudio armazenado em disco é usado pelo mecanismo de reconhecimento para confirmação e correção de palavras.
- O áudio é salvo em um disco a 11KB por segundos.
- O parâmetro SmSet SM\_SAVE\_AUDIO<sup>2</sup> controla se o áudio vai ser gravado em disco ou não.
- O áudio gravado no disco pode ser liberado por SmDiscardData.

Enquanto o parâmetro SM\_SAVE\_AUDIO especifica se o áudio em PCM (Pulse code modulation) é salvo em disco, ele também inclui todos os dados requeridos para fazer a correção. Estes dados incluem palavras alternativas e pontos, assim como os dados do áudio processado.

### 2.5. Macro de Navegação

Segundo [nós] esta funcionalidade tornou simples a filtragem das palavras, pois, depois de criada a macro com o nome do comando que se queira ativar, por exemplo, "direita", toda vez que for dito ao IBM a palavra "direita", este executará uma série de comandos pré-estabelecidos pelo usuário.

### 2.6. O Visual Basic

Após a escolha do IBM ViaVoice como *software* de reconhecimento de voz, precisamos criar um meio de visualizar se os comandos que estavam sendo ditos estavam sendo atendidos. Inicialmente, utilizamos o VBA do Excel devido ao fato de ter uma interface com o usuário bastante prática e por ser compatível com as macros criadas no IBM ViaVoice. Além disso, ficava mais simples de se apurar se o resultado desejado para uma determinada função estava sendo atingido ou não. Porém, a grande desvantagem do VBA consistia no fato de que este só podia trabalhar dentro dos aplicativos do Office, diferente do *Visual Basic (VB)* que trabalha em todo ambiente Windows. Este motivo, aliado ao fato de que necessitávamos criar uma interface mais prática e que pudesse trabalhar em qualquer máquina independente de ter um pacote *Office* instalado, fez com que utilizássemos o VB como a interface entre o IBM ViaVoice e a placa com a porta paralela. A tela do usuário é mostrada na Figura 1.

---

<sup>2</sup> : O comando SM relatado acima é usado em C++. Para saber mais sobre esta aplicação ir na seção 13.2.6 do link abaixo.

<http://www.wizzardsoftware.com/docs/smapi%20developers%20guide.pdf>

## 2.7. Interface Criada no VB

Com o intuito de se criar algo prático, bastante visual e que também não atrapalhasse no tempo de resposta do programa de reconhecimento, após alguns modelos criados, decidimos utilizar a janela mostrada na figura abaixo, onde o objetivo era para cada comando dado marcar a caixa correspondente ao comando recebido.

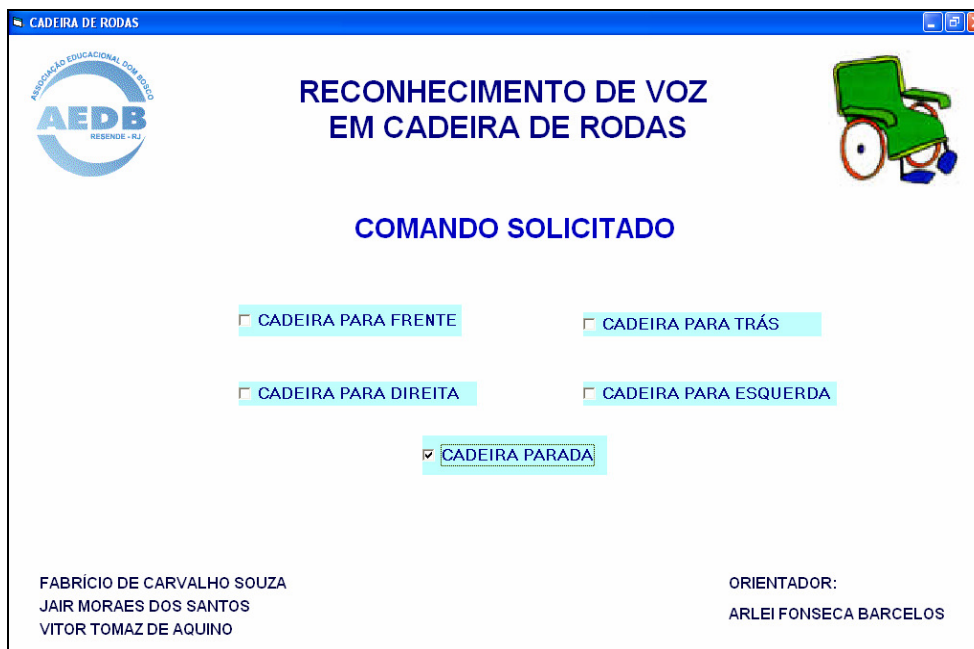


Figura 1- Tela de visualização do usuário

## 2.8. Utilizando a Porta Paralela

O Visual Basic possui algumas limitações. Ele não acessa diretamente o hardware do sistema. Todas as requisições de hardware devem ser feitas através do Windows. Por isso, o mais perto que se pode chegar da porta paralela é o Objeto Printer. Esse objeto é útil quando se quer imprimir algo, mas é inútil quando se deseja controlar um hardware conectado à porta paralela. Para conseguir controlar a porta diretamente, necessita-se do uso de algo externo ao programa. A solução para isso é o uso de alguma DLL que permita o controle da porta paralela. DLL é a sigla para *Dynamic Link Library* que quer dizer Biblioteca de Ligação Dinâmica. Existem várias à disposição na Internet e gratuitas. Em nosso projeto utilizamos a *inpout32.dll*. Com esta DLL conseguimos enviar informações através da porta paralela do computador.

## 2.9. Hardware e dispositivos aplicados

A pesquisa e o estudo detalhado de alguns dispositivos e componentes aplicados a este trabalho, por exemplo, a Porta Paralela, foi utilizada para enviar os comandos do computador à cadeira de rodas; o Micro controlador PIC [6] mostrado na Figura 2, foi utilizado para decodificar o código binário recebido da Porta Paralela e comandar o Display, e o Display LCD mostra o comando de voz dado é apresentado na Figura 3. O conjunto destes componentes se fez indispensáveis para o bom andamento do trabalho.

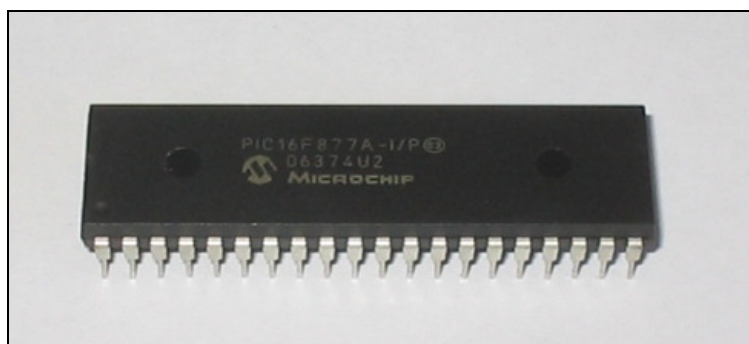


Figura 2 - Microcontrolador utilizado PIC16F877A.



Figura 3 - Display LCD 16x2.

### 2.10. Circuito de interface com a cadeira

O circuito de conexão com a cadeira foi idealizado com a utilização de um Microcontrolador PIC 16F877A da Microchip, programado para ler as entradas da Porta Paralela, um LCD 16x2 para sinalização dos comandos e um conjunto de foto acopladores tipo FOD817, para fazer o isolamento da porta paralela do restante do circuito.

## 3. Conclusão e Resultados

De acordo com os dados estatísticos e os testes realizados comprovou-se a funcionalidade da ferramenta IBM ViaVoice que pode ser utilizada, juntamente com o software Visual Basic, tanto para reconhecimento de voz quanto para acionamento da cadeira. Através desta proposta, realizou-se um estudo para diagnosticar quais ferramentas eram viáveis na utilização deste trabalho. Verificou-se que na língua portuguesa existem poucos software de reconhecimento de voz e o que obteve melhor desempenho em reconhecimento de palavras foi o IBM ViaVoice.

Outro fator importante deste programa são as funcionalidades que as API's do IBM proporcionaram na filtragem das palavras de interesse e a interface do ViaVoice com o Visual Basic. O Visual Basic foi o que obteve a melhor interface visual, facilidade de programação e acionamento do circuito de controle da cadeira.

Um estudo paralelo foi realizado para uma interface com o meio externo. A interface escolhida foi o microprocessador PIC que está sendo utilizado em larga escala, pela flexibilidade de utilização e baixo custo.

O projeto torna-se versátil devido a sua ampla área de aplicação podendo ser utilizado em: acionamento de qualquer circuito, acionamento de dispositivos em ambiente residencial, sistemas de segurança entre outros.

Desta forma, acredita-se que o objetivo deste projeto foi alcançado, uma vez que a proposta inicial foi "Reconhecimento de Voz para aplicação em Cadeira de Rodas". Além disso, este trabalho apresentou resultados satisfatórios para reconhecimento de palavras e os índices de acertos ficaram em torno de 95,2%.

No projeto foi utilizado um computador portátil para utilização com a cadeira de rodas, como proposta de aperfeiçoamento deve-se buscar a utilização de um sistema embarcado para reconhecimento de voz na cadeira de rodas.

#### 4. REFERÊNCIAS BIBLIOGRÁFICAS

[1] A. S. Pereira.; G. Vieira, J. P. "Motorização de Cadeiras de Rodas". Projeto Final de curso, AEDB, dezembro de 2007.

[2] DAMASCENO, E.F., **Implmentação de serviço de voz em ambientes virtuais**, em <<http://www.dcc.ufla.br/infocomp/artigos/v4.3/art09.pdf>>. Acessado em: 22 out. 2007

[3] FURUI S. **Digital Speech Processing, Synthesis and Recognition**. Marcel Dekker, Inc., 1989.

[4] **IBM DMAPI Developer's Guide**, em: <<http://www.wizzardsoftware.com/docs/dmapi%20reference%20guide.pdf>> , 2001. Acesso em 01 Out 2007

[5] **IBM SMAPI Developer's Guide**, em: <<http://www.wizzardsoftware.com/docs/smapi%20developers%20guide.pdf>> , 2001. Acesso em 01 Out 2007.

[6] MICROCHIP, **Microchip Corporation**, disponível em: <[www.microchip.com](http://www.microchip.com)>, 2007. Acesso em: 15 Set. 2007.

[7] RONALD A. C., et al. **Survey of the State of the art in Human Language Technology**. Disponível em <<http://cslu.cse.ogi.edu/HLTsurvey/HLTsurvey.html>> 2000. Acesso em 16 Set. 2007.

[8] ROE, D.B. e WILPON J. G., "**Voice communication between humans and machines**". National Academy of Sciences, 1994.

[9] SUN. Java™ Speech API Programmer's Guide Version 1.0 . 1998, acessado em 12 Jul. 2003

\_\_\_\_\_, **The Java Speech Markup Language Specification** <<http://java.sun.com/products/java-media/speech/forDevelopers/jsapi-guide/SpeechEngine.html>>, 1997. Acessado em 13 Jul. 2007