

Voice Recognition By MATLAB (November 2017)

F. D. Araujo, M. L. de Sousa, M.C.B. Araújo, M. V. de O. Guilhon, N. M. F. Mota

Abstract— This article will address the theme "voice recognition by MATLAB", which proved to be very useful. Because to trigger something using the voice can be very comfortable and practical, besides being safe. This work will explore MATLAB's own functions for the implementation of sound and recognition of that sound.

Keywords—MATLAB, Voice, Sound.

I. Introdução

A modernidade introduziu a sociedade uma gama de inovações, como sistemas de reconhecimento de movimento e controles cada vez mais sensíveis para uso em jogos. Porém ainda é comum que toda interação com o computador seja dada por meio de tecnologias como teclados, mouses e joysticks. Desde que foram criados as primeiras aplicações existe uma busca por tornar as mesmas cada vez mais interativos por meio de controles mais intuitivos, reconhecimento de movimento ou mesmo de toque. Porém grande maioria dos recursos é pouco inclusiva e depende de toque direto do usuário, por isso a procura pelo uso de reconhecimento de voz para facilitar a interação do usuário de maneira geral.

Este artigo apresenta uma demonstração de como funciona o reconhecimento fisicamente, análises de entrada e saída de dados e estudar os princípios e recursos do reconhecimento de voz por meio do software.

II. O som no MATLAB

O Matlab propicia alguns recursos para a leitura e reprodução de arquivos de sons.

`sound sound(y)` converte um vetor y em som. O vetor é escalonado automaticamente de modo que o máximo e mínimo valores em y possam corresponder à máxima e mínima amplitude permitidas pelo *hardware* de som do computador.

O som é produzido na taxa de amostragem pré-ajustada;

`sound(y, Fs)` reproduz o som numa frequência de amostragem de F_s Hz.

O comando `wavread` Permite "carregar" para o ambiente Matlab um arquivo de som; `[y,Fs,Formato]=wavread(wavefile)` carrega um arquivo em formato .wav retornando os dados amostrados na variável y , a taxa de amostragem na variável F_s e informação de formato do arquivo .wav na variável `Formato`. A informação de formato é um vetor de 6 elementos.

O `wavread` atualmente aceita somente um canal de dados de 8 bits.

O comando `wavwrite` Grava um arquivo .

`wav; wavwrite(y,Fs,wavefile);`

onde: y = Os dados amostrados a serem gravados.

F_s = Taxa na qual os dados foram amostrados.

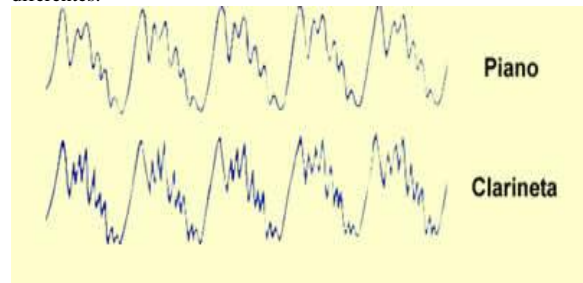
`wavefile` = Uma *string* com o nome do arquivo .wav a ser criado.

III. Conceitos na física

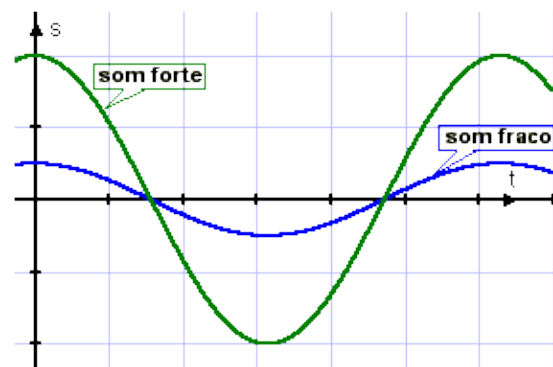
O som é uma onda mecânica, a qual possui a intensidade e frequência necessárias para ser percebida pelo ser humano. Uma onda mecânica é um tipo de onda que necessita de meios materiais, como o ar ou o solo, para a sua propagação. As frequências audíveis ao ouvido humano estão entre 16 Hz e 20000Hz (20kHz). Nessa faixa são encontrados a voz humana, instrumentos musicais, alto-falantes, etc.

Os sons podem ser caracterizados a partir da timbre, intensidade ou altura.

Timbre: É o conjunto de ondas sonoras que formam o som. Com o timbre é possível diferenciar as fontes sonoras, por exemplo a percepção de um piano e uma clarineta, que tem sons totalmente diferentes.

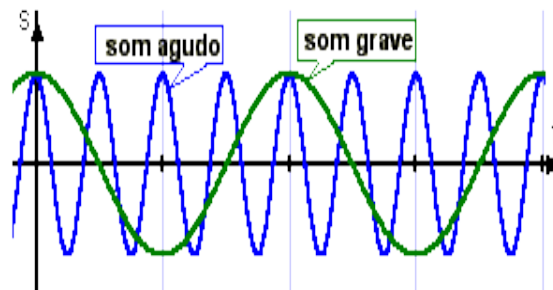


Intensidade: É a quantidade de energia transportada pelo som. De acordo com a intensidade do som, podemos dizer que ele é mais forte, ou seja, a onda possui uma maior amplitude ou mais fraca, onde a onda possui menor amplitude.



Amplitude das vibrações de uma partícula do campo ondulatório (meio).

Altura: Relacionada a frequência do som, desse modo se pode distinguir os sons mais altos como os de maior frequência e mais agudos. Os mais baixos são o de menor frequência e mais graves.



Frequências das vibrações de uma partícula do campo ondulatório (meio).

F. D. Araújo is with Universidade Estadual do Maranhão, São Luís. E-mail: felipemanager2009@hotmail.com

M. L. de Sousa is with Universidade Estadual do Maranhão, São Luís. E-mail: mateus_lima99@hotmail.com

M.C.B. Araújo is with Universidade Estadual do Maranhão, São Luís. E-mail: murilocalcastelo@hotmail.com.

M. V. de O. Guilhon is with Universidade Estadual do Maranhão, São Luís. E-mail: viniciusguilhon@gmail.com

N. M. F. Mota is with Universidade Estadual do Maranhão, São Luís. E-mail: motta.reig1@hotmail.com

III. Metodologia e Resultados

Nesta seção serão descritos os resultados da realização deste trabalho, onde, inicialmente teremos as entradas do sistema, sua base de dados, como o reconhecimento de voz funciona, e, seus respectivos resultados finais.

Para o este trabalho foi usado o MATLAB(MATrix LABoratory), um software de alta performance voltado para o cálculo numérico, onde os problemas e suas respectivas soluções são expressas matematicamente.

Utilizando comandos próprios para o trabalho com sons, como o wavread, wavwrite e sound, foi possível a aquisição dos comandos de voz para que o som seja analisado e comparado com o som gravado previamente.

III. I. Aquisições do comando de voz

Através da função wavrecord, o áudio é gravado pelo MATLAB. O bloco de Aquisição dos Comandos de Voz é responsável por pegar as falas por um microfone inserido na entrada de áudio de um PC. Este bloco basicamente é uma classe em matlab que modela um gravador de voz e tem como saída um vetor de bytes cujos níveis representam a intensidade do sinal de voz. Os parâmetros de entrada utilizados para esse gravador são:

- Tempo de gravação: 1.5s
- Taxa de amostragem: 11025
- Número de canais: 1

Essas amostragens de sinal e até mesmo o experimento em si foram feitos em uma placa de áudio Conexant SmartAudio HD e vale ressaltar que tudo foi feito sem microfone de filtragem de áudio.

IV.I. Bases de Dados

Neste programa está sendo utilizada uma base de dados onde contém trechos de falas, com sons gravados previamente, mais precisamente falando, 19 falas (.wav) de aproximadamente 1.5 segundos. O funcionamento consiste em uma entrada que está na base de dados, o processamento que será feito por métodos numéricos avançados e a saída será o reconhecimento da voz falada justamente porque anteriormente já existia esta gravação na base de dados, nada mais é do que uma linkagem.

IV.II. Reconhecimentos de um sinal de voz

Será analisada uma entrada de dados com o título 'canto', nela, será detalhado todo o funcionamento de um sistema reconhecedor de voz juntamente com os métodos usados para o funcionamento do mesmo.

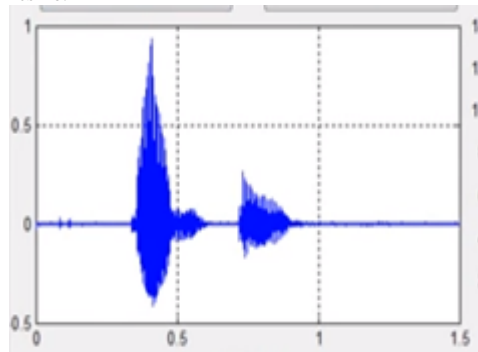


Figura 1.1: Sinal de entrada com sua taxa de amplitude em frequência(palavra canto).

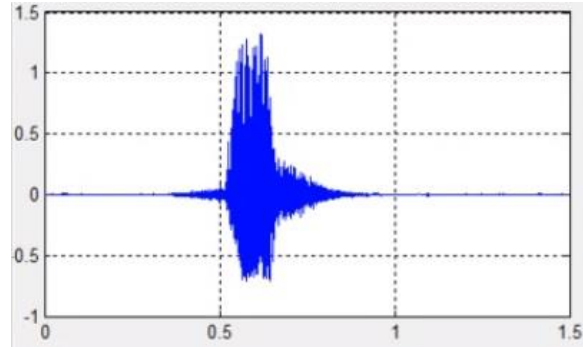


Figura 1.2: Sinal de entrada com sua taxa de amplitude em frequência(palavra sal).

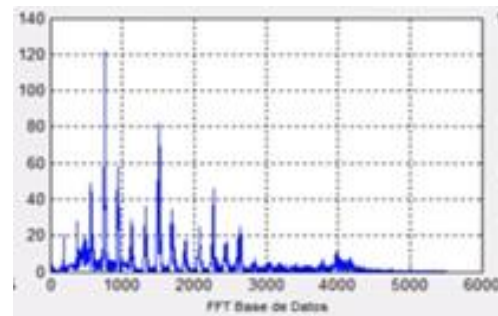


Figura 2.1: Transformada rápida de Fourier do áudio gravado no banco de dados do sistema(palavra canto).

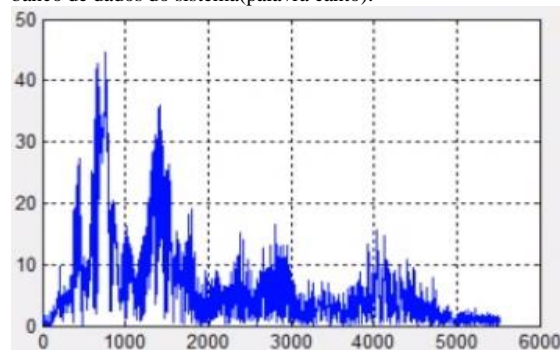


Figura 2.2: Transformada rápida de Fourier do áudio gravado no banco de dados do sistema(palavra sal).

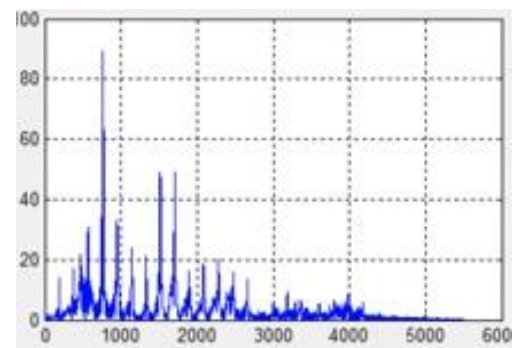


Figura 3.1: Transformada rápida de Fourier do áudio gravado no banco de dados do usuário(palavra canto).

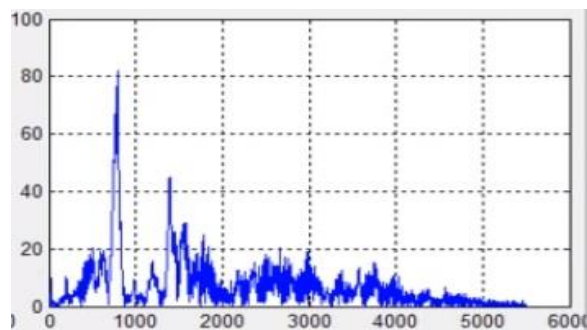


Figura 3.2: Transformada rápida de Fourier do áudio gravado no banco de dados do usuário (palavra sal).



Figura 4.1: Reconhecimento da palavra 'canto'.



Figura 4.2: Reconhecimento da palavra 'sal'.

V. Conclusão

Neste trabalho chegou-se a conclusão que o desenvolvimento de sistemas de reconhecimento de voz é viável para facilitar a interação do usuário com diversos tipos de sistemas.

Existem alguns fatores que podem influenciar no reconhecimento de voz. O timbre da voz do usuário, os equipamentos utilizados e os ruídos do ambiente. Deve-se atentar ao máximo a estes fatores para que se possa obter bons resultados.

Por fim conclui-se que o desenvolvimento de sistemas de reconhecimento de voz é benéfico para o avanço tecnológico.

Referências

HALLIDAY, D. Fundamentos de Física: Mecânica, vol 2. 9 ed. LTC, 2006.

<<http://www.mec.ita.br/~adade/Matlab/Web/som.htm>> Acesso em: 04 nov. 2017.

Siqueira J. K. RECONHECIMENTO DE VOZ PARA O PORTUGUÊS BRASILEIRO. Centro de Estudos em Telecomunicações, Rio de Janeiro, 2008.

<<http://www.ece.northwestern.edu/local-apps/matlabhelp/techdoc/ref/wavrecord.html>> Acesso em: 04 nov. 2017.

<<http://matlab.izmiran.ru/help/techdoc/ref/wavread.html>> Acesso em: 04 nov. 2017.