



Universidade do Minho

Escola de Engenharia

Mestrado Integrado em Engenharia Eletrónica
Industrial e Computadores

Processamento Digital de Sinal 2014-2015

Algoritmo que permite a separação da fala do ruído

Trabalho realizado por:

Cristian Costa a68573

Docentes: Carlos Lima

Carlos Silva

Índice

| | |
|---|----|
| Introdução..... | 3 |
| Conceitos Teóricos | 4 |
| SNR- Signal-to-Noise Ratio | 4 |
| Média | 4 |
| Variância | 5 |
| Ruído Branco | 5 |
| Distribuição Normal | 6 |
| Funcionamento do código em <i>Matlab</i> | 7 |
| Alguns Exemplos na linha de comandos do <i>Matlab</i> | 9 |
| Testes | 10 |
| Amostra 1: Sem Ruído → “Cristian Costa” | 10 |
| Amostra 2: Ruído Esporádico (bater uma palma) → “Engenharia Eletrónica” | 11 |
| Amostra 3: Ruído de Fundo (som de trovoadas) → “Engenharia Eletrónica” | 12 |
| Amostra 4: Ruído de Fundo (som de trovoadas) e Ruído esporádico (bater uma palma) → “Cristian Costa” | 13 |
| Conclusão..... | 14 |

Introdução

Este trabalho consiste em implementar um algoritmo para a detecção da fala usando o modelo da distribuição Gaussiana.

Para tal, irá ser realizado 4 testes um para cada diferente situação, tais como: um em que contém segmentos de fala com alguns intervalos de silêncio (pausa); outro onde, para além do segmento de fala irá conter um ou mais ruído esporádico; o terceiro teste irá conter ruído de fundo durante o segmento de fala e por fim um teste que contém um ruído esporádico e um ruído de fundo.

Será também, brevemente resumido, alguns conceitos teóricos lecionados nesta UC e que são fundamentais para a compreensão e realização do problema proposto.

A elaboração do trabalho foi feita através do programa *Matlab*, sendo este usado nas aulas.

Conceitos Teóricos

SNR- Signal-to-Noise Ratio

Relação sinal-ruído ou razão sinal-ruído é usado em medidas de um sinal em meio ruidoso, definido como a razão da potência de um sinal e a potência do ruído sobreposto ao sinal, isto é, compara o nível de um sinal desejado com o nível do ruído de fundo.

Quanto mais alta for a relação sinal ruído, menor é o efeito do ruído de fundo sobre a detecção ou medição do sinal.

A equação usada no código para a sua determinação:

$$\text{SNR} = (\text{mean}(\text{fich_audio})/\text{sqrt}(\text{var}(\text{fich_audio})));$$

Média

Em estatística, a média é o valor que aponta para onde mais se concentram os dados de uma distribuição. Pode ser considerada o ponto de equilíbrio das frequências, num histograma.

$$\text{Media} = \frac{\sum_{i=1}^n x_i}{n} = \frac{x_1 + \dots + x_n}{n}$$

Variância

Em estatística, a variância de uma variável aleatória é uma medida da sua dispersão estatística, indicando quão longe em geral os seus valores se encontram do valor esperado.

O seu valor corresponde ao quadrado do Desvio Padrão.

Variance = s^2

$$= \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_N - \bar{x})^2}{N}$$
$$= \frac{1}{N} \sum (x_i - \bar{x})^2$$

$$\begin{aligned}\text{Var}(X) &= E[(X - E(X))^2] \\ &= E[X^2 - 2X E(X) + [E(X)]^2] \\ &= E(X^2) - E[2X E(X)] + [E(X)]^2 \\ &= E(X^2) - 2E(X)E(X) + [E(X)]^2 \\ &= E(X^2) - 2[E(X)]^2 + [E(X)]^2 \\ &= E(X^2) - [E(X)]^2\end{aligned}$$

Ruído Branco

Ruído branco é um tipo de ruído produzido pela combinação simultânea de sons de todas as frequências. O adjetivo branco é utilizado para descrever este tipo de ruído em analogia ao funcionamento da luz branca, dado que esta é obtida por meio da combinação simultânea de todas as frequências cromáticas.

Em estatística, um ruído branco, aplica-se a uma sequência de erros aleatórios, sempre que esta tiver média e variância constantes e sem auto correlação.

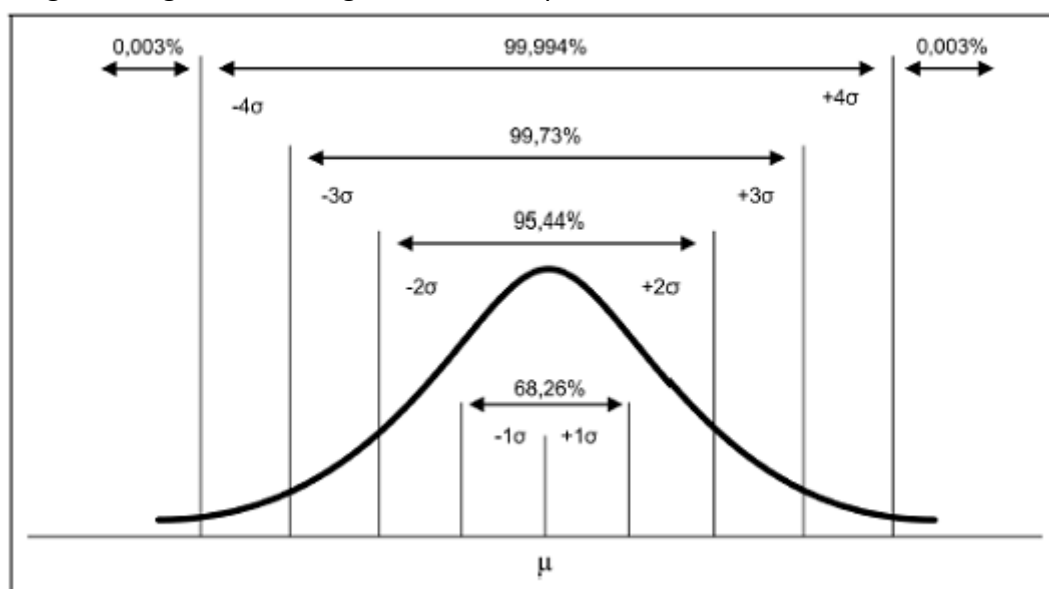
Distribuição Normal

Também conhecida como Distribuição de Gauss ou Gaussiana, é uma das mais importantes distribuições da estatística.

É inteiramente descrita por dois parâmetros: a **média** μ e o **desvio padrão** σ , ou seja, conhecendo-se estes valores consegue-se determinar qualquer probabilidade em uma distribuição Normal.

A partir desses dois parâmetros conseguimos especificar a equação da curva Normal $N(\mu, \sigma)$, sendo a área da curva normal determinada através dos mesmos.

A Figura a seguir mostra algumas áreas importantes:



A distribuição normal é simétrica em torno da média o que implica que a área à direita é igual a área à esquerda.

A média refere-se ao centro da distribuição e o desvio padrão ao achatamento da curva.

Funcionamento do código em *Matlab*

```
function [nseg_fala,nseg_ruido]=detecao_nsegfala(fich_audio,fHz,tam_ruido,tam_tempo_pausa,
tam_tempo_ruido_esporadico)
%detecao_nsegfala: função que analisa o sinal de áudio e separa o ruído
%da fala, criando dois vectores novos, (nseg_ruido) e (nseg_fala) respectivamente

%argumentos de entrada:
    %fich_audio: vector que contém um sinal audio
    %fHz: frequencia de amostragem do sinal em hertz
    %tam_ruido: tamanho do segmento do ruído
    %tam_tempo_pausa: a duração do tempo de pausa permitido no segmento de voz (geralmente 50 amostras)
    %tam_tempo_ruido_esporadico: tamanho em amostras que um ruído esporádico pode tomar

%argumentos de saída:
    %nseg_fala: vetor que guarda o segmento de fala (sem ruído)
    %nseg_ruido: vetor que guarda o segmento de ruído (ruído removido)

    temp_esporadico=0; %variavel contador de amostras de ruído esporadico
    namostras=0; %variavel contador do numero de amostras
    temp_fala_pause=0; %variavel contador da pausa entre as falas
    |(verifica se é uma pausa ou não)
    nseg_fala=0; %vetor segmento de fala
    nseg_ruido=0; %vetor segmento de ruído

%formula para o calculo do SNR
SNR=(mean(fich_audio)/sqrt(var(fich_audio)));

%Threshold calcula um limite, que nos vai indicar qual é a
%probabilidade de encontrar ruído ou fala.É calculado com base no SNR
threshold=SNR*10;

if abs(threshold)<1
    threshold=2;
end

%calcula da média e da variância
media=mean(fich_audio(1:tam_ruido));
variancia=var(fich_audio(1:tam_ruido));
```

```

%Algoritmo da função
for i=1:length(fich_audio) %o ficheiro de audio é percorrido
    if abs(fich_audio(i))>media+abs(threshold)*sqrt(variancia) %Se for
        %maior, é porque é um segmento de fala, se for menor é porque
        %é ruído
        temp_esporadico=temp_esporadico+1;
        namostras=namostras+1;
        temp_fala_pause=0;

        if temp_esporadico > tam_tempo_ruído_esporadico
            %se o contador de amostras de ruído esporadico for
            %maior que o valor recebido, conta-se como um segmento de
            %fala
            nseg_fala=cat(1,nseg_fala,fich_audio(i-namostras+1:i));
            namostras=0;
        end

    else %segmento de ruído
        temp_fala_pause=temp_fala_pause+1;
        namostras=namostras+1;
        if temp_fala_pause>tam_tempo_pausa %se o contador da pausa
            %for maior que o valor recebido, então conclui-se que
            %houve uma pausa na fala
            nseg_ruído=cat(1,nseg_ruído,fich_audio(i-namostras+1:i));
            temp_esporadico=0;
            namostras=0;
            temp_fala_pause=0;
        end
    end
end
end
end

```


Alguns Exemplos na linha de comandos do *Matlab*

```
>> variavel=wavrecord(3*8000,8000);% captura de um sinal de 3segundos a 8kz
```

```
>> sound(variavel)%reprodução do sinal audio capturado
```

```
>> plot(variavel) % plot do sinal audio capturado
```

```
%utilização da função principal
```

```
>> [f,r]=detecao_nsegfala(variavel,fHz,tam_ruido,tam_tempo_pausa,tam_tempo_ruido_esporadico)
```

```
>> subplot(3,1,1), plot(variavel);%plot do sinal original
```

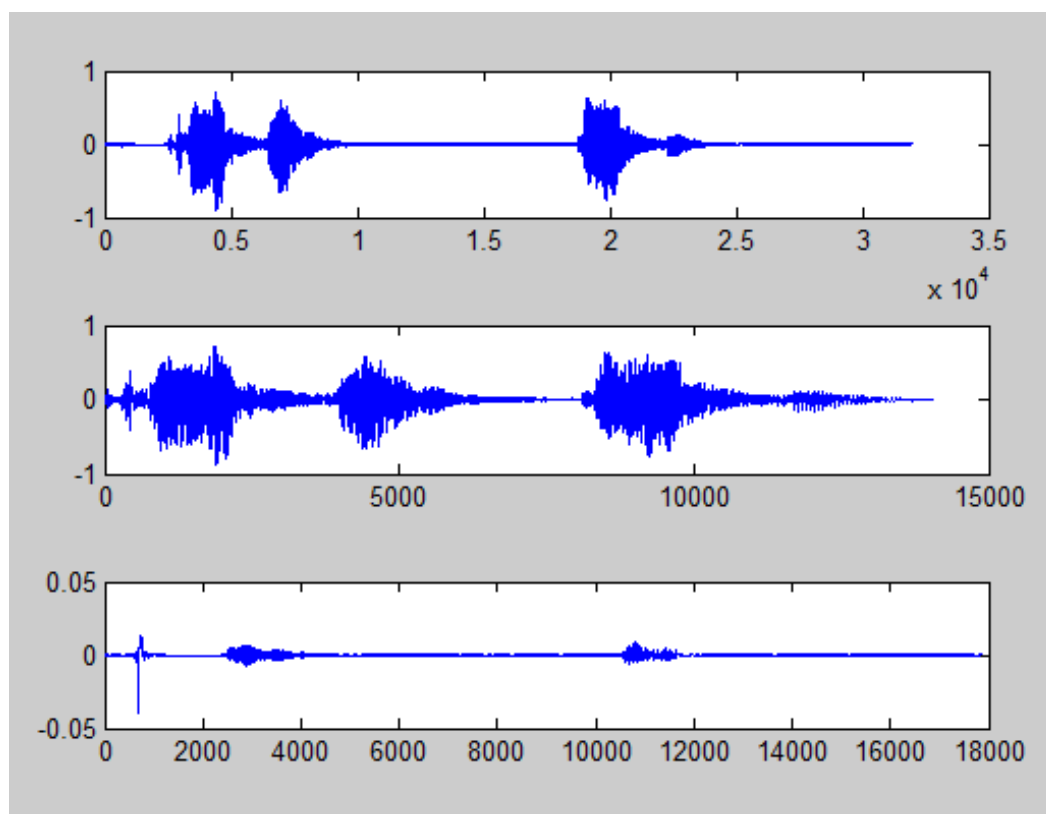
```
>> subplot(3,1,2), plot(f);%plot do sinal limpo (fala)
```

```
>> subplot(3,1,3), plot(r);%plot do sinal de ruido
```

Testes

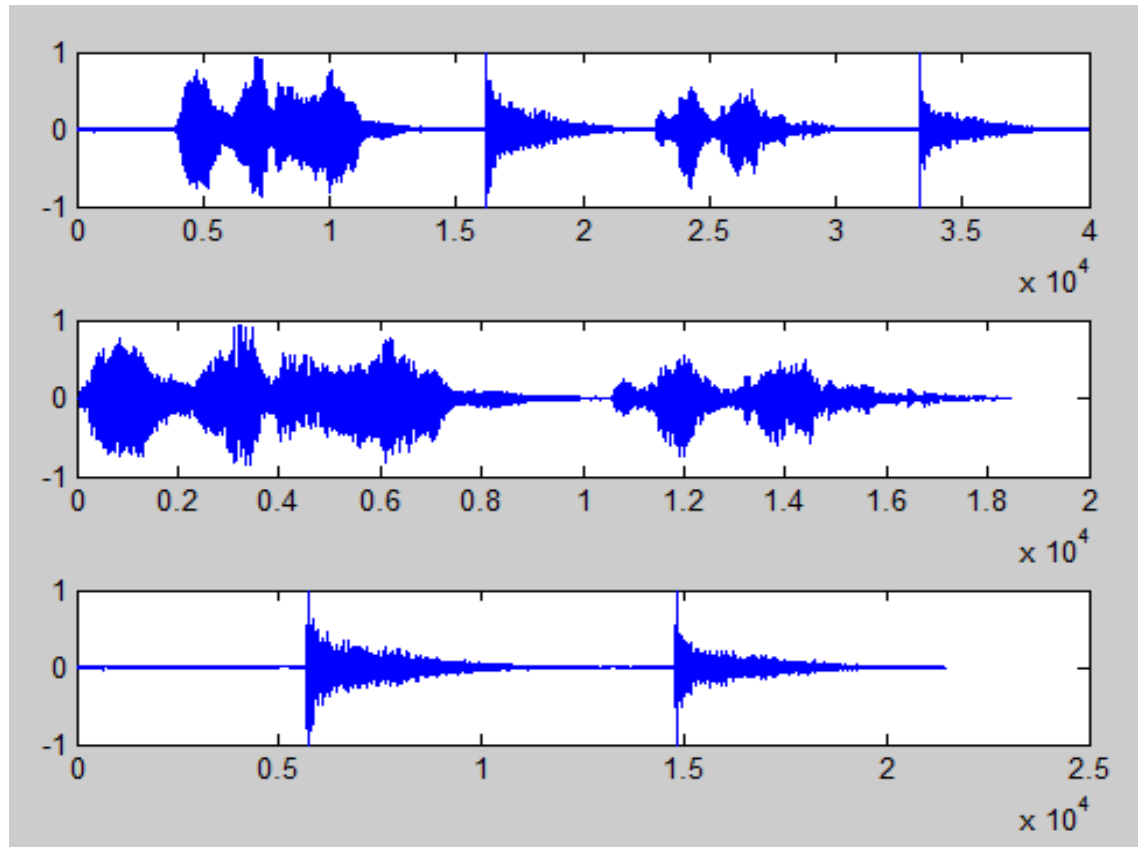
Amostra 1: Sem Ruído → “Cristian Costa”

Este teste mostra o resultado do segmento de fala de duas palavras, “Cristian Costa” com um intervalo de silêncio entre as duas palavras.



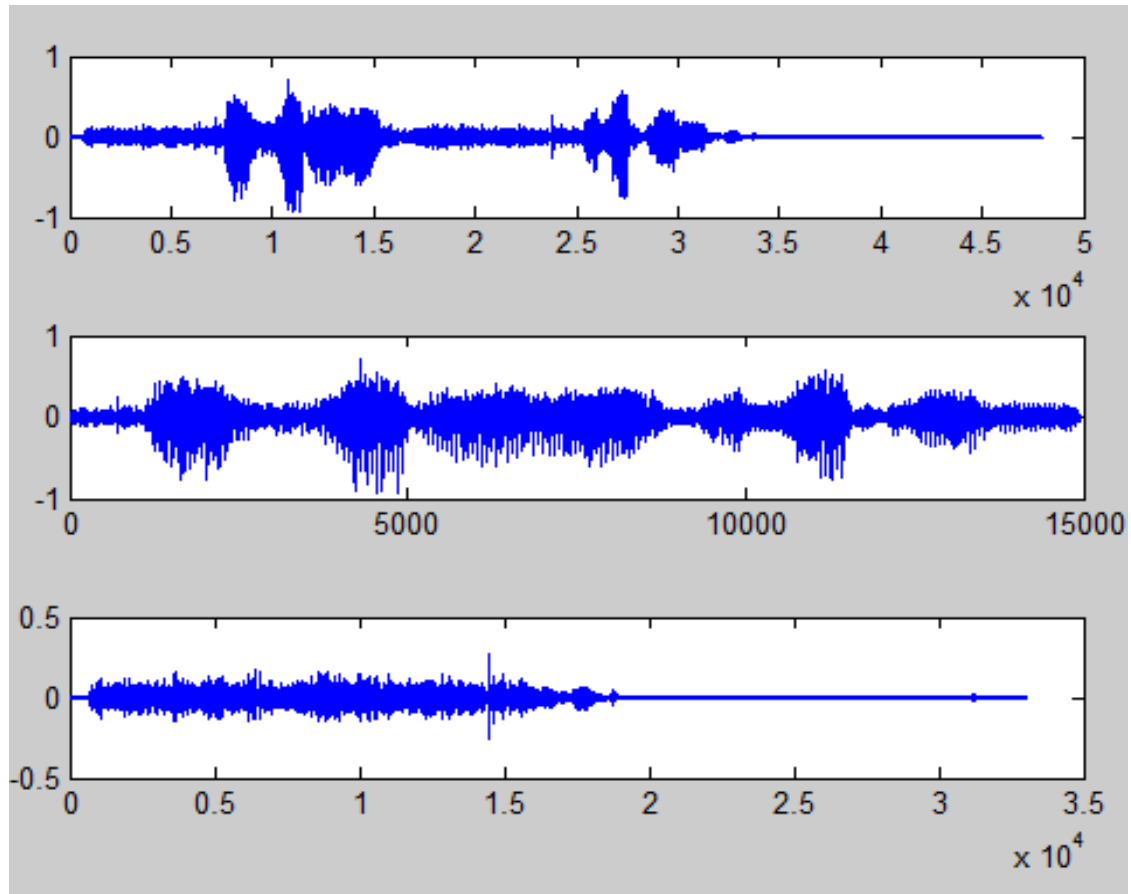
Amostra 2: Ruído Esporádico (bater uma palma) → “Engenharia Eletrónica”

Esta figura representa um segmento de fala de duas palavras, “Engenharia Eletrónica” com dois sons de bater uma palma: o primeiro ocorre no meio das duas palavras e o último ocorre no fim da última palavra.



Amostra 3: Ruído de Fundo (som de trovada) → “Engenharia Eletrônica”

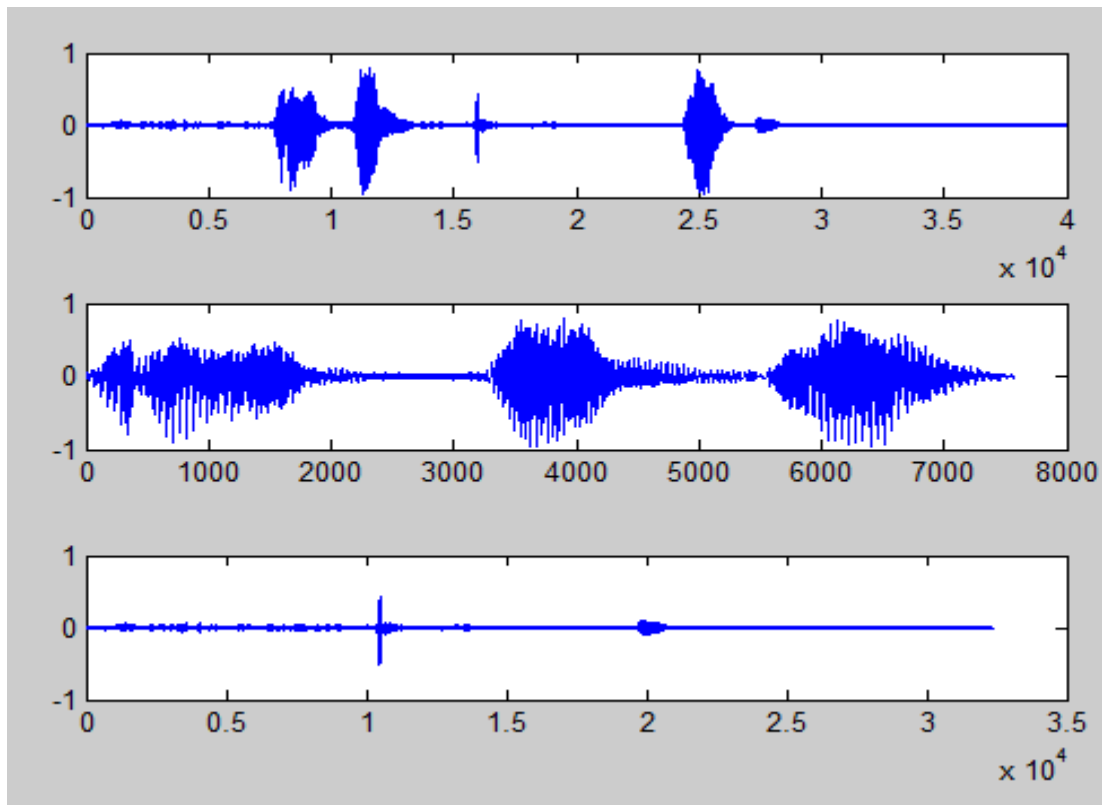
Neste teste, também foi usado um segmento de fala de duas palavras “Engenharia Eletrônica”, com um ruído de fundo, mais concretamente, o som de uma trovada.



Amostra 4: Ruído de Fundo (som de trovoadas) e Ruído esporádico (bater uma palma) → “Cristian Costa”

O último teste, consiste na utilização dos dois ruídos: som da trovoadas e o bater da palma.

O sinal áudio contém duas palavras “Cristian Costa” e o ruído esporádico ocorre no meio destas duas palavras.



Conclusão

As quatro amostras realizadas foram concebidas com sucesso, provando assim que a função funciona.

Os testes foram focados em exemplos simples e práticos, englobando as situações descritas anteriormente.

Foi assim possível provar a eliminação de ruído existente entre as falas, resultando em segmentos vozeados de forma seguida mas em que só consta a fala, sem ruído.