



UNIVERSIDADE DO MINHO

JULHO 2016

Mestrado Integrado em Engenharia Eletrónica Industrial e Computadores

UNIDADE CURRICULAR

Processamento Digital de Sinal

Remoção de ruído de fundo e de ruído esporádico

Docente: Carlos Lima

Trabalho realizado por:

Luís Miguel Matos de Almeida-a68576

Índice

Introdução.....	0
Conceitos Teóricos	1
1-SNR – Signal to Noise Ratio	1
2-Ruido branco	1
3-Distribuição Normal	2
Funcionamento do código.....	3
Testes	6
Conclusão	10
Anexos	11
Função remove_ruido	12
Script	13

Introdução

No âmbito da disciplina de processamento digital de sinal foi proposto a implementação de um algoritmo para a detecção da fala, usando o programa Matlab utilizado nas aulas da disciplina.

Assim, ira ser realizado quatro testes:

- Segmento de fala com intervalo de silêncio e ruído de fundo;
- Segmento de fala sem intervalo de silêncio contendo um ruído esporádico depois da fala;
- Segmento de fala com intervalo de silêncio contendo ruído esporádico antes e depois da fala;
- Segmento de fala com intervalo de silêncio contendo ruido esporádico no meio da fala.

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

Conceitos Teóricos

1-SNR – Signal to Noise Ratio

Relação sinal-ruído (SNR signal to noise ratio) 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 [1].

$$\text{SNR} = \frac{P_{\text{signal}}}{P_{\text{noise}}},$$

Figura 1-definição SNR

A relação sinal-ruído é geralmente usada em sinais elétricos, mas pode também ser aplicada em sinais óticos e acústicos.

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:

$$\text{SNR} = \frac{\mu}{\sigma}$$

$$\text{SNR} = (\text{mean}(\text{gravacao})/\text{std}(\text{gravacao}));$$

Figura 2- definição alternativa

2-Ruido 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, aplica-se a uma sequência de erros aleatórios, sempre que esta tiver média e variância constantes e sem apresentar auto correlação [2].

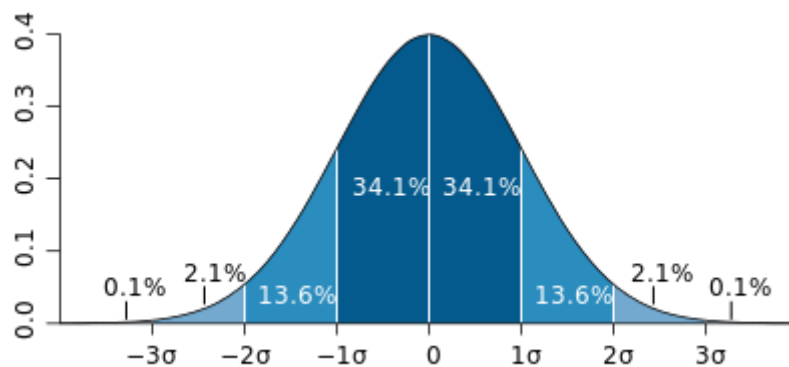
3-Distribuição Normal

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 [3].

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

Para a realização deste trabalho recorreu-se ao uso de uma função desenvolvida especificamente para este trabalho, chamada `remove_ruido()` e um script onde se introduz os parâmetros e onde são feitos os gráficos.

Esta função recebe como parâmetros as seguintes variáveis:

- `gravação` : vetor com o sinal de áudio. Gravado com a função `wavrecord()` do MATLAB;
- `tamostras_ruido`: número de amostras iniciais onde apenas se grava ruído (sem informação útil);
- `tamostras_pausas`: número de amostras entre duas palavras durante a fala de uma pessoa (geralmente 50 amostras);
- `amostras_ruido_espo`: número máximo de amostras que um ruído esporádico pode tomar (geralmente 2000 amostras).

Retorna como resultado os seguintes vetores:

- `Voz_sem_ruido`: sinal resultante sem ruído.
- `ruido_retirado`: ruído removido da gravação.

No início da função inicializam-se os vetores de retorno a 0 bem como as variáveis `cont_amostras`, `cont_amostras_espo` e `cont_pausas` também a 0.

A variável `cont_amostras_espo` conta o número de amostras de modo a conseguir verificar se o que se está a avaliar é um sinal esporádico ou não.

A variável `cont_pausa` verifica se é uma pausa entre palavras ou não.

A variável `cont_amostras` é uma auxiliar para a contagem de amostras quando se percorre o sinal recebido.

```
cont_amostras_espo=0; %variável contador de amostras de ruído
esporádico
cont_amostras=0;      %variável contador do numero de amostras
cont_pausas=0;        %variável contador de pausas entre falas
voz_sem_ruido=0;      %vetor da voz limpa
ruido_retirado=0;     %vetor do ruído
```

Após isto calcula-se o SNR com a formula apresentada no capítulo anterior.

```
SNR=(mean(gravacao)/std(gravacao));
```

Depois calcula-se um limite ao qual é chamado de *lim*, que nos vai indicar qual é a probabilidade de encontrar ruído ou fala, assumindo uma distribuição gaussiana.

```
lim=SNR*10;  
  
if abs(lim)<1  
    lim=2;  
end
```

Assumindo que nas primeiras amostras apenas se gravou ruído, calcula-se a média e o desvio padrão deste “sub-sinal” que contem apenas informação não desejada.

```
media=mean(gravacao(1:tamostras_ruído));  
desv_padrao=std(gravacao(1:tamostras_ruído));
```

De seguida, percorre-se o ficheiro de áudio do início ao fim e avalia-se o valor da gravação é maior que o limite calculado ($media + abs(lim) * desv_padrao$). Se for maior, trata-se essa amostra como uma palavra da fala, caso não seja trata-se como ruído.

Caso seja palavra fala, avalia-se sempre se o contador do ruído esporádico é maior que o valor recebido, se for, conta-se como uma palavra, caso não seja maior é removido.

O mesmo acontece para a pausa entre palavras, se o contador da pausa for maior que o valor passado por parâmetro, assume-se que é ruído.

```
for i=1:length(gravacao) %vetor da gravação é percorrido até ao fim  
    if abs(gravacao(i)) > media+abs(lim)*desv_padrao %avalia-se  
        se o valor da gravação é maior que o limite, se for maior é porque é um  
        segmento de fala e se for menor é porque é ruído  
        cont_amostras_espo=cont_amostras_espo+1;  
        cont_amostras=cont_amostras+1;  
        cont_pausas=0;  
  
        if cont_amostras_espo > amostras_ruído_espo %se o contador  
        de amostras de ruído esporádico for maior que o valor recebido, conta-  
        se como um segmento de fala, se for menor é removido  
        voz_sem_ruído=cat(1,voz_sem_ruído,gravacao(i-  
        cont_amostras+1:i));  
        cont_amostras=0;  
    end  
  
    else  
        cont_pausas=cont_pausas+1;  
        cont_amostras=cont_amostras+1;  
        if cont_pausas > tamostras_pausas %se o contador de pausas  
        for maior que o valor recebido, então houve uma pausa na fala e assume-  
        se como ruído  
        ruído_retirado=cat(1,ruído_retirado,gravacao(i-  
        cont_amostras+1:i));  
        cont_amostras_espo=0;  
        cont_amostras=0;  
        cont_pausas=0;  
    end  
end  
end  
end
```

Linha de comandos do Matlab

```
>> gravacao= wavrecord(4*8000,8000); %captura de um sinal de 4 segundos a 8khz  
>> sound(gravacao) %ouvir se a gravação ficou bem gravada  
>> plot (gravação) %ver gráfico
```

Usamos a função plot para determinar o número de amostras iniciais do ruído (tamostras_ruído), a duração do tempo de pausa permitido no segmento de voz (tamostras_pausas) e o tamanho em amostras que um ruído esporádico pode tomar (amostras_ruído_espo). A frequência que foi utilizada foi de 8kHz e o tempo de gravação foi de 4s.

De seguida, iremos para outro script, em que será chamada a função remove_ruído e feita a criação dos gráficos pretendidos.

```
tamostras_ruído=5000; %valor exemplo visto num grafico  
tamostras_pausas=50;  
amostras_ruído_espo=2000;  
  
[voz_sem_ruído,ruído_retirado]=remove_ruído(gravacao,tamostras_ruído,tamostras_pausas,amostras_ruído_espo); %saida da voz limpa e voz sem ruído  
  
a=subplot(3,1,1),plot(gravacao); %criação de 3 graficos (gravação, voz_semruído, ruído_retirado);  
title(a,'Gravação');  
b=subplot(3,1,2),plot(voz_sem_ruído);  
title(b,'Voz Limpa');  
c=subplot(3,1,3),plot(ruído_retirado);  
title(c,'Ruído');
```

Para testarmos se a voz_sem_ruído esta correta poderemos na linha de comandos fazer a função sound(voz_sem_ruído), tal como o ruído_retirado.

Testes

Em todos os testes usaram-se as seguintes palavras: “Luís Almeida”.

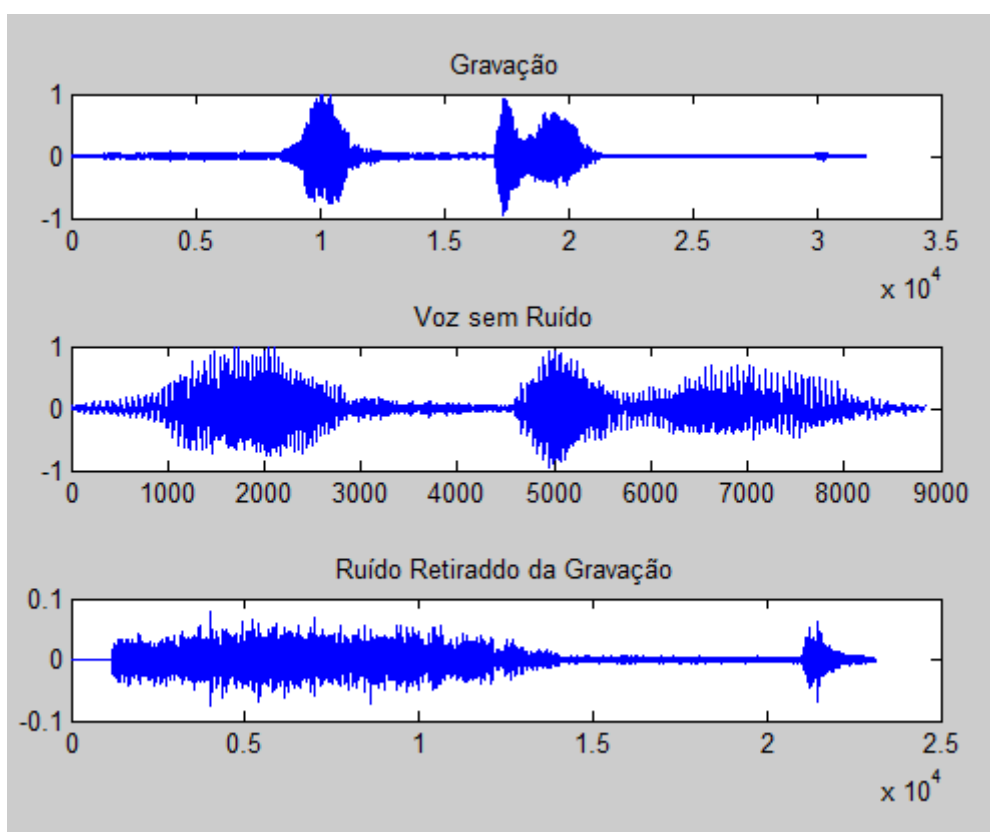
Em todas as imagens que se seguem, o primeiro gráfico, é o som gravado, o segundo é o som sem ruído (voz sem ruído) e o ultimo é apenas o ruído retirado da gravação.

Amostra 1:

Este teste mostra o resultado do segmento de fala de duas palavras, “Luís Almeida” com um intervalo de silêncio entre as duas palavras. Tem um ruído de fundo de um aspirador e um ruído esporádico de uma porta a fechar, tendo sido este ultimo um ruído não previsto.

Foi utilizado os seguintes parâmetros:

```
tamostras_ruído=8000;  
tamostras_pausas=50;  
amostras_ruído_espo=2000;
```

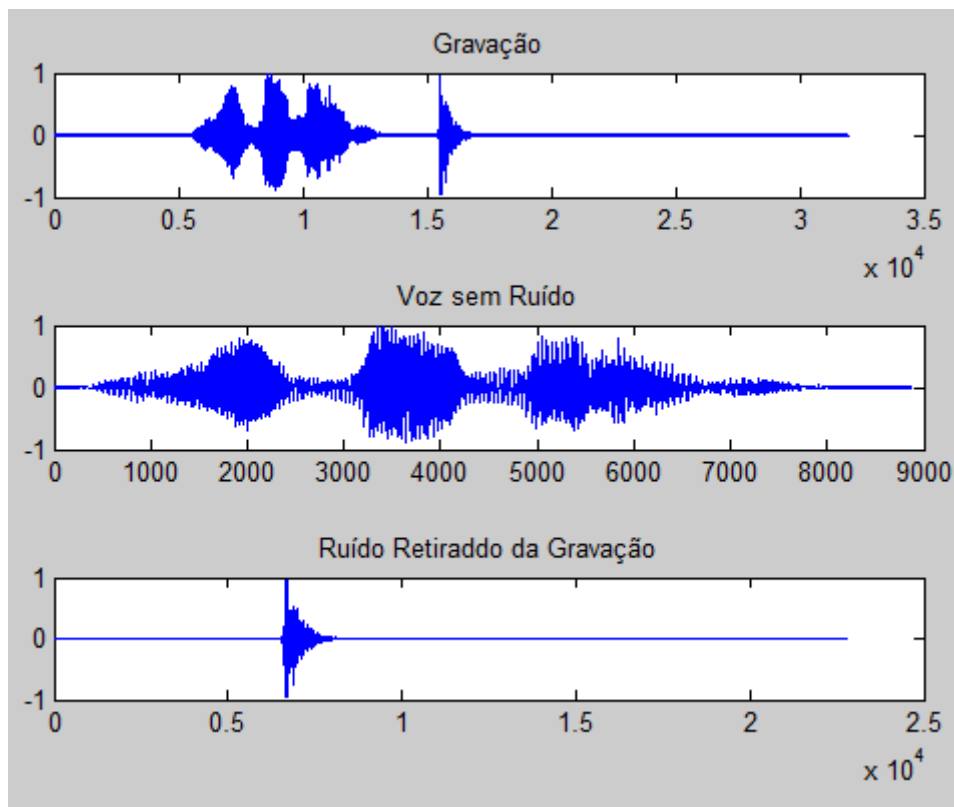


Amostra2:

Este teste mostra o resultado do segmento de fala de duas palavras, “Luís Almeida” sem um intervalo de silêncio entre as duas palavras. Tem um ruído esporádico de uma caneta a cair depois do segmento de fala das duas palavras.

Foi utilizado os seguintes parâmetros:

```
tamostras_ruído=5000;  
tamostras_pausas=50;  
amostras_ruído_espo=2000;
```

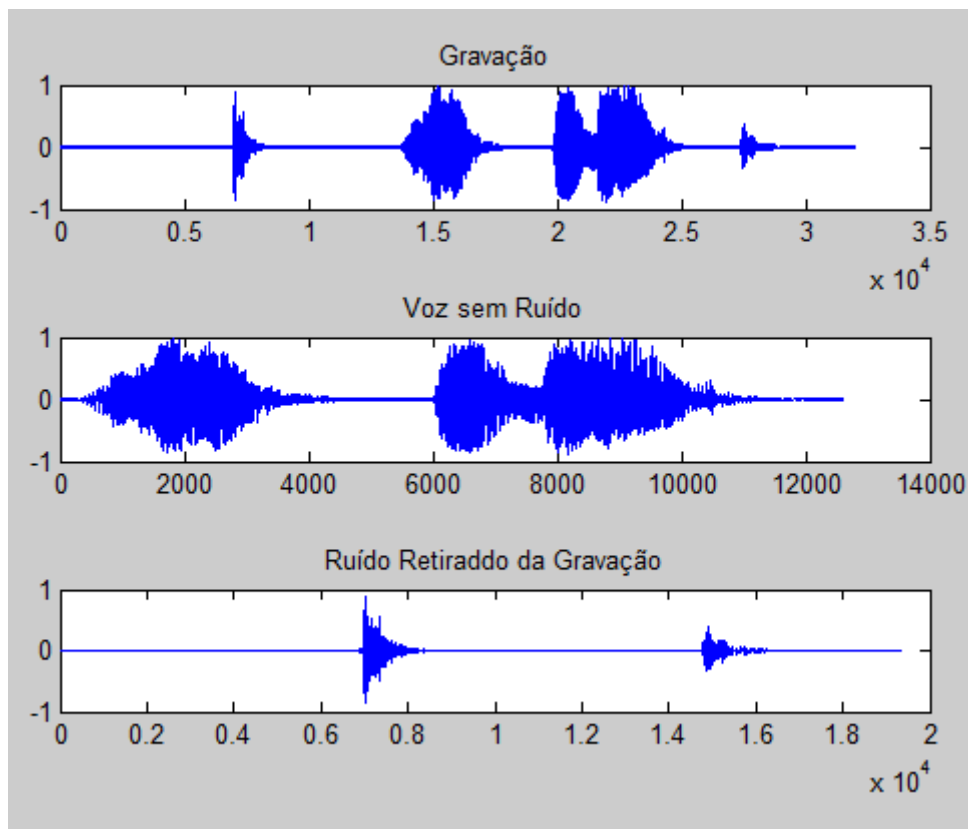


Amostra 3:

Este teste mostra o resultado do segmento de fala de duas palavras, “Luís Almeida” com um intervalo de silêncio entre as duas palavras. Tem um ruído esporádico de uma caneta a cair no início e depois do segmento de fala das duas palavras.

Foi utilizado os seguintes parâmetros:

```
tamostras_ruído=6800;  
tamostras_pausas=50;  
amostras_ruído_espo=2000;
```

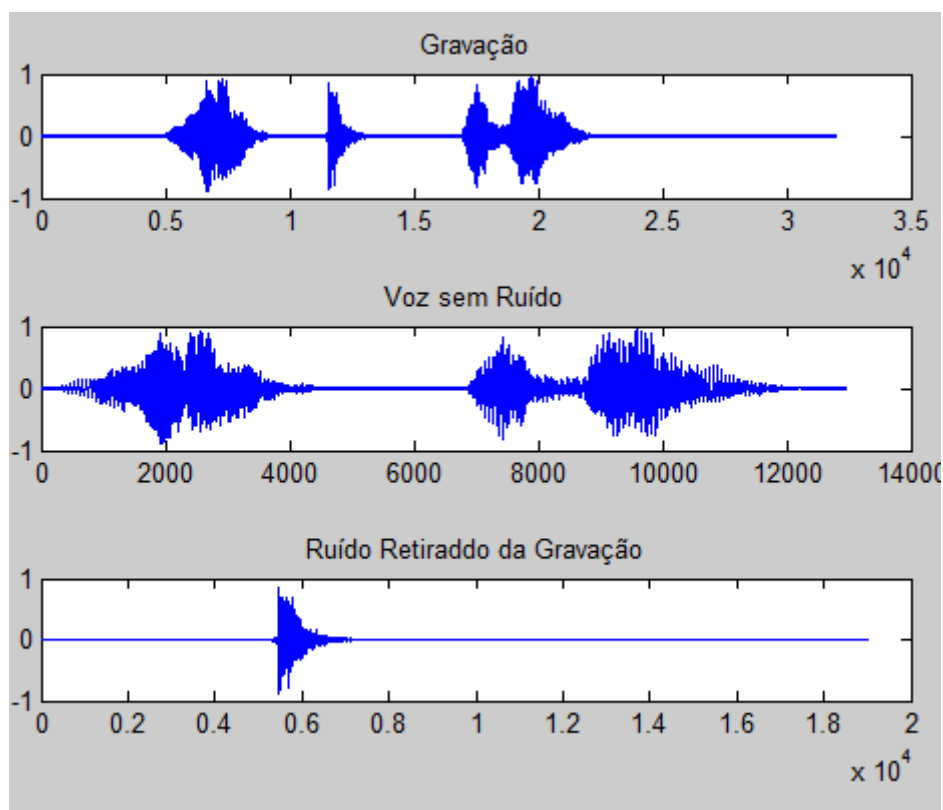


Amostra 4:

Este teste mostra o resultado do segmento de fala de duas palavras, “Luís Almeida” com um intervalo de silêncio entre as duas palavras. Tem um ruído esporádico de uma caneta a cair no meio do segmento de fala das duas palavras.

Foi utilizado os seguintes parâmetros:

```
tamostras_ruído=6800;  
tamostras_pausas=50;  
amostras_ruído_espo=2000;
```



Conclusão

Após a finalização do trabalho podemos concluir que as quatro amostras foram realizadas com sucesso, provando assim que a função funciona. Também podemos concluir que foi possível adquirir conhecimentos na ferramenta utilizada MATLAB que não foi muito abordada até agora.

Assim é possível provar a eliminação de ruído existente entre as falas, resultando uma fala limpa (sem ruído).

Neste trabalho ajudou a perceber o tratamento que leva os sons que ouvimos provenientes de rádios, televisão...

Referências

- [1] “wikipedia,” 10 Julho 2016. [Online]. Available: https://en.wikipedia.org/wiki/Signal-to-noise_ratio. [Acedido em Julho 2016].
- [2] “wikipedia,” 10 Julho 2016. [Online]. Available: https://pt.wikipedia.org/wiki/Ru%C3%ADdo_branco. [Acedido em Julho 2016].
- [3] “wikipedia,” 24 Setembro 2015. [Online]. Available: https://pt.wikipedia.org/wiki/Distribui%C3%A7%C3%A3o_normal. [Acedido em Julho 2016].

Anexos

Função remove_ruido

```
function [voz_sem_ruido,ruido_retirado]=remove_ruido(gravacao,tamostras_ruido,tamostras_pausas,amostras_ruido_espo)

    cont_amostras_espo=0;    %variável contador de amostras de ruído esporádico
    cont_amostras=0;        %variável contador do número de amostras
    cont_pausas=0;          %variável contador de pausas entre falas
    voz_sem_ruido=0;        %vetor da voz limpa
    ruido_retirado=0;       %vetor do ruído

    %Cálculo do SNR através da razão da média e do desvio
    SNR=(mean(gravacao)/std(gravacao));

    %"lim" calcula um limite, que nos vai indicar qual é a probabilidade de encontrar voz ou ruído
    lim=SNR*10;
    if abs(lim)<1
        lim=2;
    end
    %Adquirir o modelo do ruído através do cálculo da média e do desvio padrão
    media=mean(gravacao(1:tamostras_ruido));
    desv_padrao=std(gravacao(1:tamostras_ruido));

    for i=1:length(gravacao) %vetor da gravação é percorrido até ao fim
        if abs(gravacao(i)) > media+abs(lim)*desv_padrao%compara se o valor da gravação,
            cont_amostras_espo=cont_amostras_espo+1;%se for maior que o limite: maior é segmento de fala, menor é ruído
            cont_amostras=cont_amostras+1;
            cont_pausas=0;

            if cont_amostras_espo > amostras_ruido_espo%se o contador for maior que o valor recebido,
                voz_sem_ruido=cat(1,voz_sem_ruido,gravacao(i-cont_amostras+1:i));
                cont_amostras=0; %conta-se como um segmento de fala, se não é removido
            end

        else
            cont_pausas=cont_pausas+1;
            cont_amostras=cont_amostras+1;
            if cont_pausas > tamostros_pausas
                %se o contador for maior que o valor recebido, então houve uma pausa na fala e assume-se como ruído
                ruido_retirado=cat(1,ruido_retirado,gravacao(i-cont_amostras+1:i));
                cont_amostras_espo=0;
                cont_amostras=0;
                cont_pausas=0;
            end
        end
    end
end
```

Script

```
tamostras_ruído=5000;    %valor visto no grafico
tamostras_pausas=50;
amostras_ruído_espo=2000;

[voz_sem_ruído,ruído_retirado]=remove_ruído(gravacao,tamostras_ruído,tamostras_pausas,amostras_ruído_espo);
%saída da voz limpa e voz sem ruído

%criação de 3 graficos (gravação, voz_semruído, ruído_retirado);
a=subplot(3,1,1),plot(gravacao);
title(a,'Gravação');
b=subplot(3,1,2),plot(voz_sem_ruído);
title(b,'Voz Limpa');
c=subplot(3,1,3),plot(ruído_retirado);
title(c,'Ruído');
```