

UNIVERSIDADE DO MINHO

MESTRADO INTEGRADO EM ENGENHARIA ELETRÓNICA INDUSTRIAL  
E COMPUTADORES

PROCESSAMENTO DIGITAL DE SINAL

---

# Detecção e Remoção de Intervalos de Silêncio na Fala

---

*Aluno*

RAFAEL MELEIRO MARQUES  
A74420

*Professor*

CARLOS MANUEL GREGÓRIO  
SANTOS LIMA

11 Junho 2019

# Conteúdo

<b>Lista de Figuras</b>	<b>iii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Motivação . . . . .	1
1.2 Objetivos . . . . .	1
1.3 Estrutura do Relatório . . . . .	1
<b>2 Conceitos Teóricos</b>	<b>2</b>
2.1 Signal-to-noise ratio . . . . .	2
2.2 Processos Estocásticos . . . . .	2
2.3 Ruído Branco . . . . .	2
2.4 Modelo do Ruído . . . . .	3
2.5 Outlier . . . . .	4
<b>3 Problema</b>	<b>5</b>
<b>4 Desenvolvimento em Matlab</b>	<b>6</b>
4.1 Aquisição do som . . . . .	6
4.2 Tamanho da janela . . . . .	6
4.3 Valor SNR . . . . .	6
4.4 Média e Variância . . . . .	7
4.5 Matriz para guardar a informação . . . . .	7
4.6 Declaração/Definição de variáveis . . . . .	7
4.7 Vetores de ruído ou fala . . . . .	8
4.8 SNR . . . . .	8
4.9 Protótipo da função . . . . .	8
<b>5 Testes e análise dos resultados</b>	<b>9</b>
5.1 Geração de ruído branco - 0 dB . . . . .	9
5.2 Geração de ruído branco - 10 dB . . . . .	10
5.3 Geração de ruído branco - 20 dB . . . . .	11
5.4 Geração de ruído branco - 30 dB . . . . .	12
5.5 Geração de ruído branco - 40 dB . . . . .	13
5.6 Geração de ruído branco - 50 dB . . . . .	14
5.7 Sinal original com adição de ruído branco . . . . .	15



## Lista de Figuras

1	Em (a) sinal limpo sem ruído, em (b) sinal com adição de ruído. . . . .	3
2	Modelo do Ruído . . . . .	3
3	Identificação no modelo do ruído, da fala e do ruído . . . . .	4
4	Identificação no modelo do ruído, da fala e do ruído . . . . .	4
5	Ruído branco de 0dB Threshold = 0.65 . . . . .	9
6	Ruído branco de 10 dB Threshold = 0.65 . . . . .	10
7	Ruído branco de 20 dB Threshold = 0.65 . . . . .	11
8	Ruído branco de 30 dB Threshold = 0.85 . . . . .	12
9	Ruído branco de 40 dB Threshold = 1.25 . . . . .	13
10	Ruído branco de 50 dB Threshold = 1.75 . . . . .	14
11	Sinal original com adição de ruído branco a 0 dB . . . . .	15
12	Sinal original com adição de ruído branco a 10 dB . . . . .	15
13	Sinal original com adição de ruído branco a 20 dB . . . . .	16
14	Sinal original com adição de ruído branco a 30 dB . . . . .	16
15	Sinal original com adição de ruído branco a 40 dB . . . . .	17
16	Sinal original com adição de ruído branco a 50 dB . . . . .	17

# **1 Introdução**

## **1.1 Motivação**

Com a realização deste trabalho pretende-se desenvolver competências no software Matlab, bem como consolidação de conhecimentos relativamente a esta matéria.

Através da prática com a ferramenta Matlab será possível visualizar alguns dos conceitos aprendidos na Unidade Curricular.

## **1.2 Objetivos**

Pretende-se implementar um algoritmo para detecção dos intervalos de silêncio na fala. Em comunicações estes intervalos são desprezados e não são enviados pois não contêm informação linguística.

Pretende-se usar o conhecido “Shewart Protocol” mas com threshold adaptativo às condições do ruído.

## **1.3 Estrutura do Relatório**

Para cumprir os objetivos no relatório será dada uma breve explicação relativamente aos conceitos teóricos.

Posteriormente será apresentado o problema proposto pelo docente e em seguida como será feita a implementação na ferramenta Matlab.

Para se conseguir verificar o correto funcionamento do código serão realizados testes para visualmente se verificar.

## 2 Conceitos Teóricos

### 2.1 Signal-to-noise ratio

SNR (signal-to-noise ratio) corresponde à relação sinal-ruído ou razão sinal-ruído, sendo usado para comparar o nível (potência) de um sinal desejado que, neste caso, corresponde à fala gravada, com o nível (potência) de ruído de fundo, presente nesse mesmo sinal.

$$SNR = \frac{Potencia\ do\ Sinal}{Potencia\ do\ Ruído}$$

Uma definição alternativa para este termo é a reciprocidade ao Coeficiente de variação, ou seja, a razão entre os parâmetros Média e Desvio Padrão, correspondentes a um determinado sinal.

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

### 2.2 Processos Estocásticos

Um processo estocástico é utilizado para analisar sinais em que não é possível saber do seu conteúdo previamente, como o problema sugerido. No entanto sabe-se o comportamento que apresentará, daí, pode-se descrever os sinais estocásticos através de um modelo probabilístico.

Podemos então considerar sinais estocásticos como sinais aleatórios, cujas amostras são independentes de todas as outras, pelo que sempre que repetimos as experiências e registamos o seu output, os valores das sucessivas amostras serão de um modo geral diferentes das amostras obtidas nas experiências anteriores.

### 2.3 Ruído Branco

O ruído branco é um tipo de ruído produzido pela combinação simultânea de sons de todas as frequências. Este é denominado de branco em analogia ao funcionamento da luz branca, dado que esta é obtida por meio da combinação simultânea de todas as frequências cromáticas. Por conter sons de todas as frequências, o ruído branco é frequentemente empregado para mascarar outros sons.

O 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 por analogia com o funcionamento da luz branca, dado que esta é obtida por meio da combinação de todas as frequências

cromáticas ao mesmo tempo. É frequentemente empregue para mascarar outros sons e é utilizado na síntese sonora, pois pode ser usado como raiz para produzir, por filtragem, formas de ondas complexas.

Em estatística, é um conceito econométrico, muito presente no estudo das séries temporais, especialmente as estocásticas discretas, onde todas as variáveis aleatórias seguem uma distribuição normal de média zero, variância constante e as covariâncias são nulas.

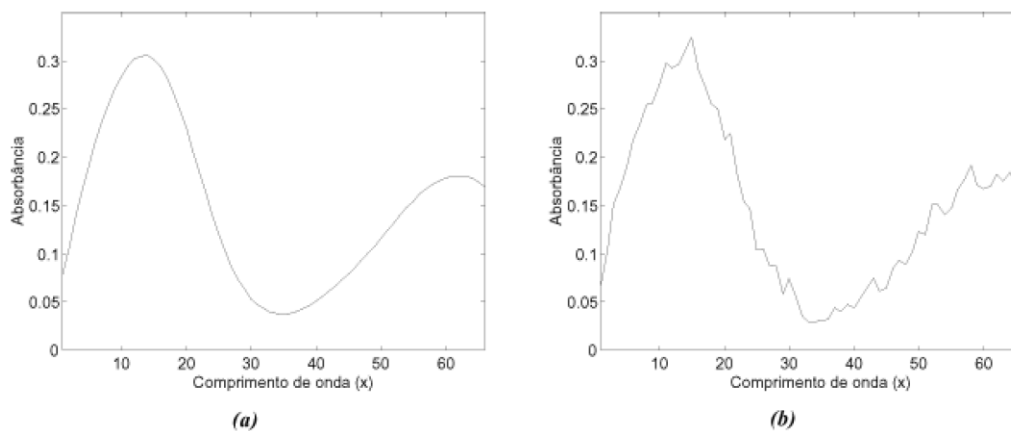


Figura 1: Em (a) sinal limpo sem ruído, em (b) sinal com adição de ruído.

## 2.4 Modelo do Ruído

O ruído é constituído por dois parâmetros fundamentais: a média e a variância, sendo estes calculados a partir das amostras retiradas do sinal. É através deles que se consegue obter o modelo do ruído:

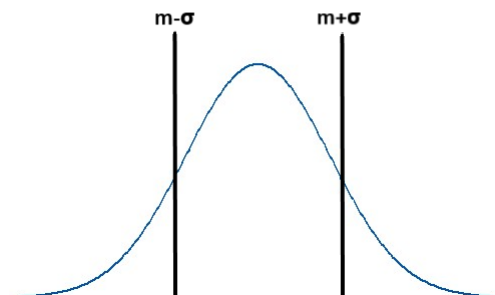


Figura 2: Modelo do Ruído

Se os valores estiverem dentro dos parâmetros  $(m-\sigma, m+\sigma)$  é sinal de ruído, se estiverem fora dos parâmetros, ou seja, no outlier, é sinal de fala.

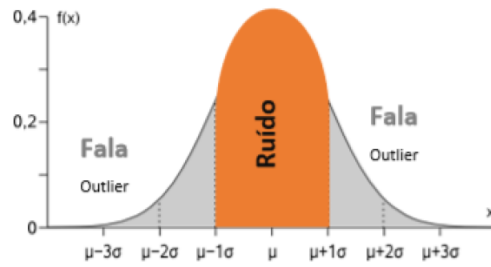


Figura 3: Identificação no modelo do ruído, da fala e do ruído

## 2.5 Outlier

Os outliers são dados que se diferenciam drasticamente de todos os outros, são pontos fora da curva. Por outras palavras, um outlier é um valor que foge da normalidade e que pode (e provavelmente irá) causar anomalias nos resultados obtidos por meio de algoritmos e sistemas de análise.

Um dos métodos de identificação de outliers mais utilizado é o desvio-padrão. Sendo que neste método, será considerado outlier um determinado valor que se encontre a uma certa quantidade de desvios-padrões da média. Essa quantidade deverá variar consoante o tamanho da amostra.

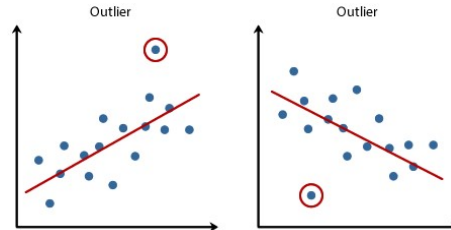


Figura 4: Identificação no modelo do ruído, da fala e do ruído



### 3 Problema

Pretende-se implementar um algoritmo para detecção dos intervalos de silêncio na fala. Em comunicações estes intervalos são desprezados e não são enviados pois não contêm informação linguística.

Pretende-se usar o conhecido “Shewart Protocol” mas com threshold adaptativo às condições do ruído. Este mecanismo requer:

1. Estudo do melhor valor de threshold para cada relação sinal-ruído. Pretende-se que faça este estudo para SNR de 0 a 50 dB com intervalos de 10 dB. Este estudo requer síntese de ruído branco, soma ao sinal e detecção dos segmentos contendo apenas ruído usando vários valores do threshold, valores entre 0.2 e 5, que devem ser colocados numa tabela. Trata-se de um estudo heurístico que deve ser justificado em relatório.
2. Cálculo automático da SNR e aplicação dos valores calculados no ponto 1.

## 4 Desenvolvimento em Matlab

De forma a cumprir o especificado pelo problema realizaram-se os seguintes passos:

### 4.1 Aquisição do som

Aquisição do som para frequência de amostragem = 4kHz durante 2 segundos.

```
Fs = 4000;
segundos = 2;

recordSound = audiorecorder(Fs,16,1);    %Criação do Objeto para Gravação
disp('Começar a falar...')
recordblocking(recordSound, segundos); %Gravação por 2 segundos no Objeto

disp('Gravacao concluida')
play(recordSound);                      %Reprodução do Objeto

input_signal = getaudiodata(recordSound); %Conversão do Objeto para um Array
```

### 4.2 Tamanho da janela

Definição do tamanho da janela, fração de 1's por janela, para considerar fala ou ruído e número de segmentos de dados por segundo.

```
tamanho_janela = 150;
fator = 0.5;
```

### 4.3 Valor SNR

Definição do valor SNR em dB para o ruído branco. Adição ao sinal capturado do ruído branco Gaussiano.

```
asnr = awgn(x,db_white_noise,'measured');
```

## 4.4 Média e Variância

Calculo da média e da variância do sinal original.

```
m=mean(x(1:length(x))); %calcular a media
v=var(x(1:length(x))); %calcular a variancia
```

## 4.5 Matriz para guardar a informação

Alocação e preenchimento (percorrendo todo o vetor asnr) de uma matriz para guardar a informação se, a cada ponto, é ruído ou fala.

```
for i=1: length(asnr)
    if(abs(asnr(i)-m)>(thresh*sqrt(v)))
        buffer_classifica(i)=1;
    else
        buffer_classifica(i)=0;
    end
end
```

## 4.6 Declaração/Definição de variáveis

1. Declaração e inicialização dos vetores de voz e ruído.
2. Definição do valor de 1's presentes na janela para ser considerado fala.
3. Definição e atribuição de valores das variáveis de controlo para gestão das janelas

```
clean=0; %buffer para a fala
noise=0; %buffer para o ruido
buffer_classifica = 1:1:length(x); %buffer para as validacoes

alfa=fr*wlen; %calcular o threshold

initw=1;
endw=initw+wlen;
prev_noise=0; %inicializar a variavel correspondente ao estado anterior
```

## 4.7 Vetores de ruído ou fala

Preenchimento dos vetores de ruído ou fala, dependendo do número de 1's presentes na janela, comparado com o valor de  $\alpha$ .

```
while(endw<length(x))
    w=buffer_classifica(initw:endw);

    if(sum(w)>alfa)
        if(prev_noise==1)
            clean=cat(1, clean, x(initw:endw));

        else
            clean=cat(1, clean, x(initw+wlen/2:endw));

        end
        prev_noise = 0;
    else
        if(prev_noise==1)
            noise=cat(1, noise, x(initw+wlen/2:endw));

        else
            noise=cat(1, noise, x(initw:endw));

        end
        prev_noise = 1;
    end

    initw=initw+wlen/2;
    endw=endw+wlen/2;

    if(endw>length(x))
        endw=length(x);
        initw=endw-initw;
    end
end
```

## 4.8 SNR

Cálculo do signal-to-noise ratio.

```
SNR = 20*log(clean_mean/noise_mean)    %calcular o SNR
```

## 4.9 Protótipo da função

```
function [noise, clean, SNR] = end_point(x, wlen,fr, thresh, db_white_noise)
```

## 5 Testes e análise dos resultados

Para os diversos testes apresentados abaixo, foi gravado um som com a frase "detecção de ruído". Para cada um dos testes é apresentada uma janela composta por três gráficos, gerada pelo matlab, que representa o sinal original, o ruído e o sinal com a informação vocal (para cada potência de sinal de ruído branco).

Para cada grupo de testes, é apresentada uma tabela com os valores de SNR consoante a variação do threshold.

### 5.1 Geração de ruído branco - 0 dB

Uma vez que para este teste o SNR é de 0dB, existe um efeito considerável do ruído no sinal original. À medida que o SNR for aumentando, o efeito do ruído sobre o sinal irá ser cada vez mais reduzido.

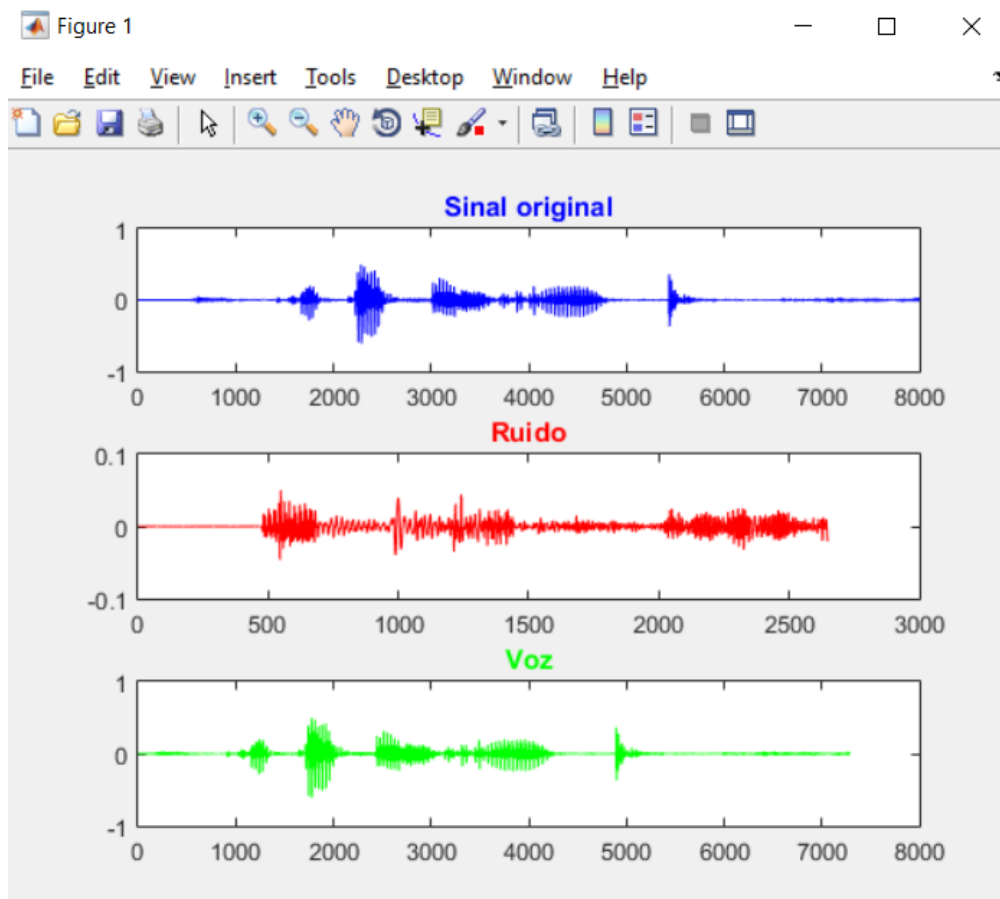


Figura 5: Ruído branco de 0dB Threshold = 0.65

Ruido Branco	Threshold	SNR
0 dB	0.65	75,1149
	0.85	48,8955
	1.25	47,2149
	1.75	55,1239

## 5.2 Geração de ruído branco - 10 dB

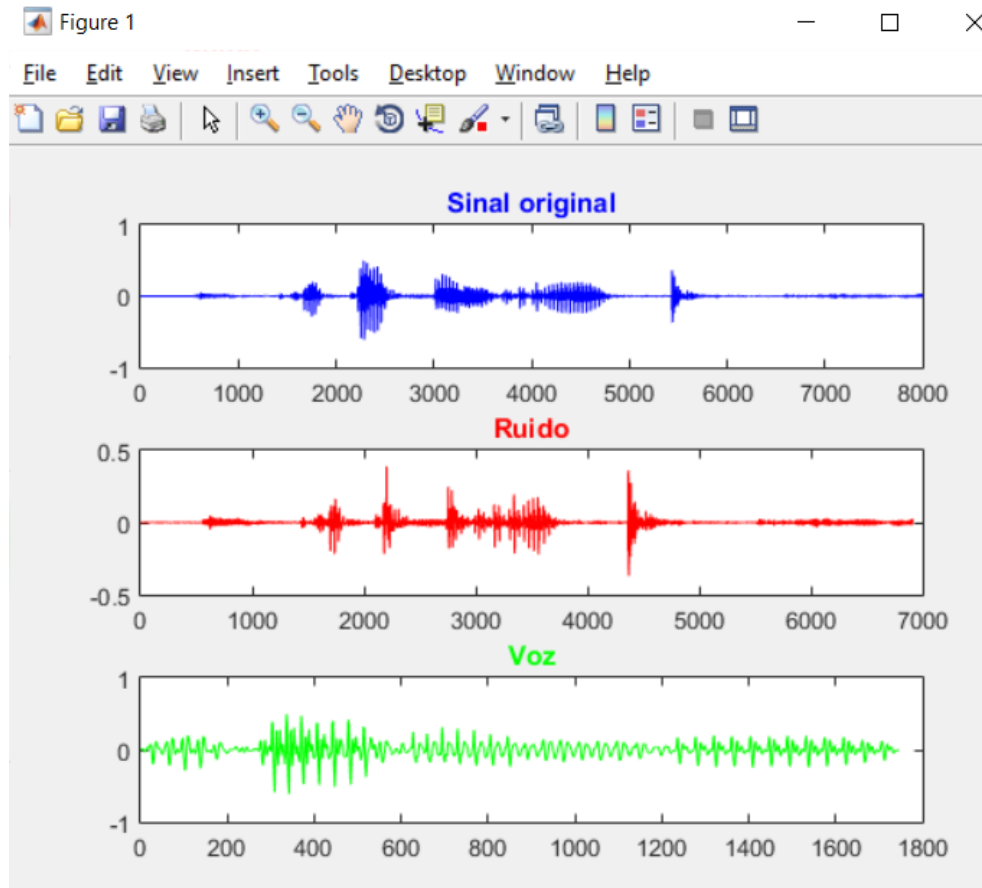


Figura 6: Ruído branco de 10 dB Threshold = 0.65

Ruido Branco	Threshold	SNR
10 dB	0.65	49,9642
	0.85	46,1777
	1.25	42,1657
	1.75	55,1239

### 5.3 Geração de ruído branco - 20 dB

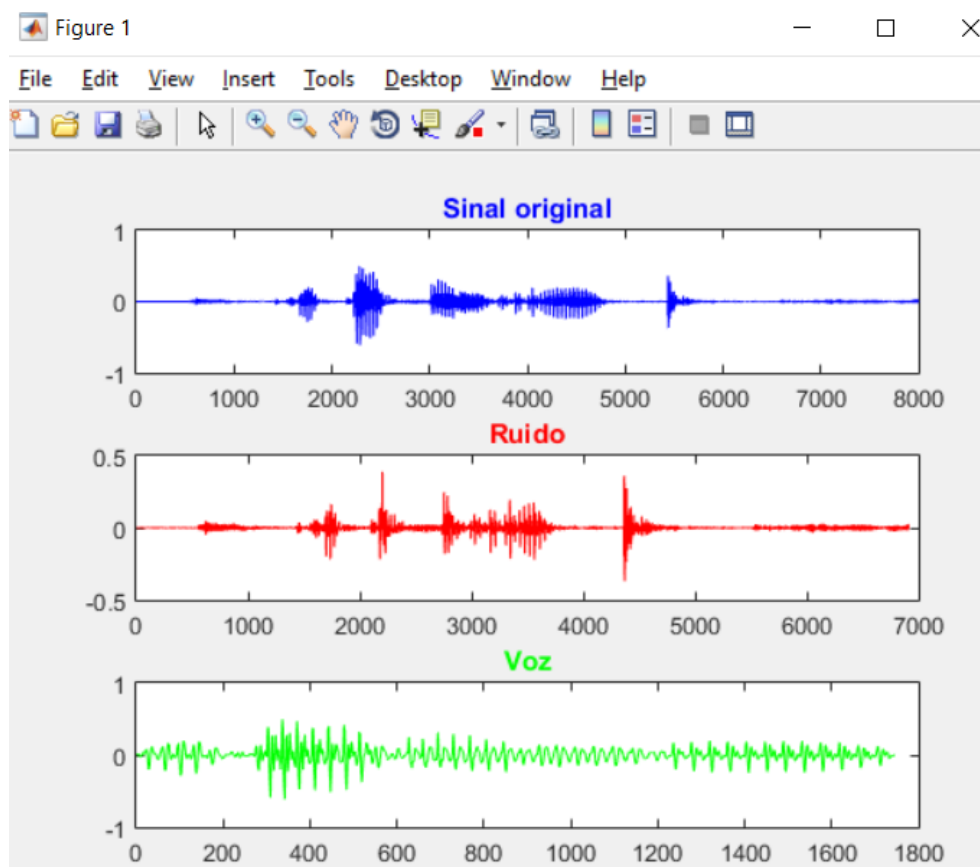


Figura 7: Ruído branco de 20 dB Threshold = 0.65

Ruido Branco	Threshold	SNR
20 dB	0.65	49,2363
	0.85	46,1777
	1.25	40,9678
	1.75	55,1239

## 5.4 Geração de ruído branco - 30 dB

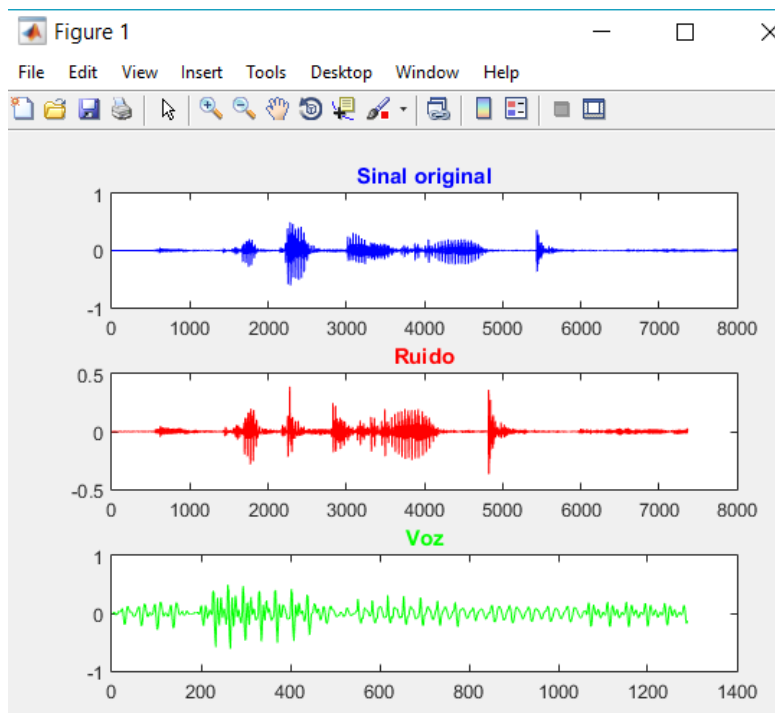


Figura 8: Ruído branco de 30 dB Threshold = 0.85

Ruido Branco	Threshold	SNR
30 dB	0.65	47,1976
	0.85	45,8278
	1.25	43,2405
	1.75	55,1239



## 5.5 Geração de ruído branco - 40 dB

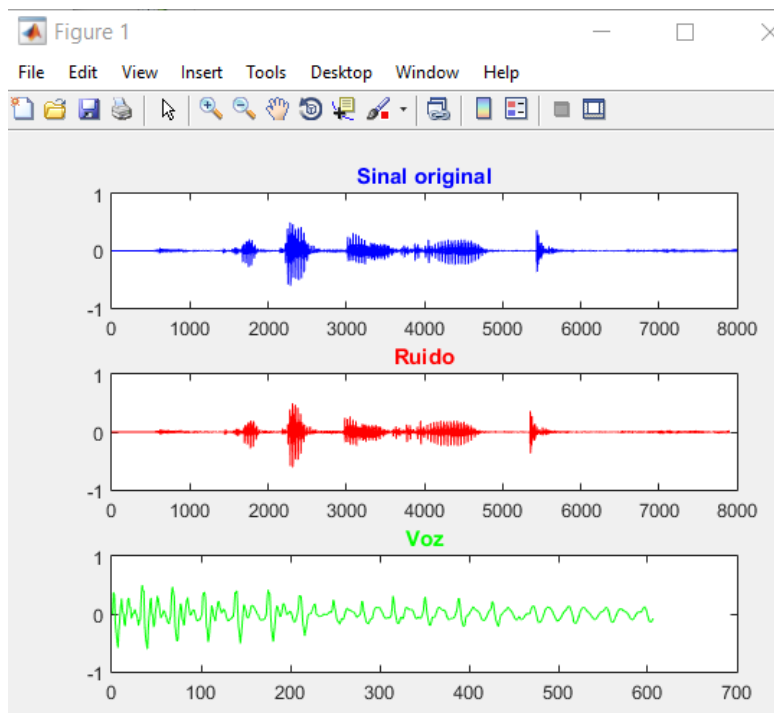


Figura 9: Ruído branco de 40 dB Threshold = 1.25

Ruido Branco	Threshold	SNR
40 dB	0.65	48,7712
	0.85	45,431
	1.25	42,1657
	1.75	55,1239

## 5.6 Geração de ruído branco - 50 dB

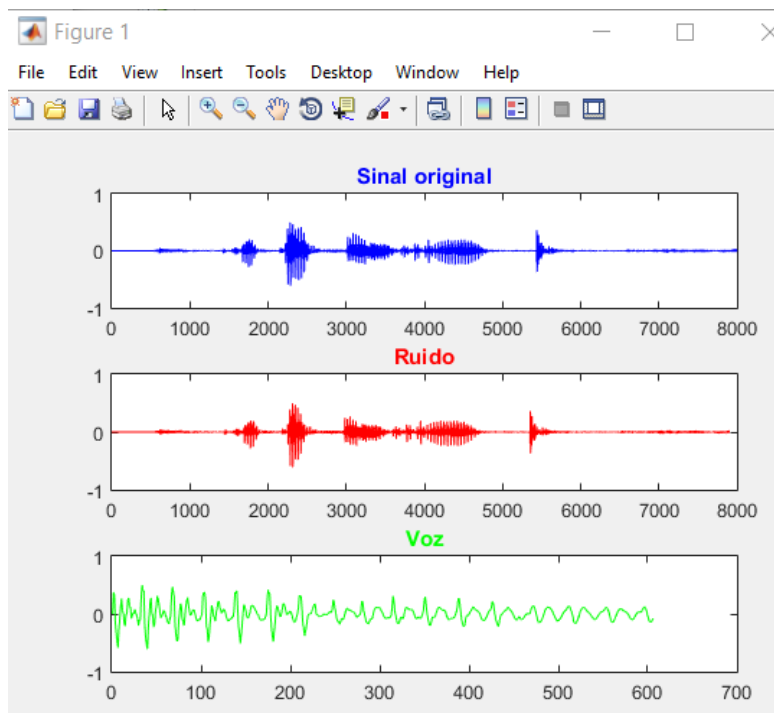


Figura 10: Ruído branco de 50 dB Threshold = 1.75

Ruido Branco	Threshold	SNR
50 dB	0.65	47,4587
	0.85	45,3873
	1.25	42,1657
	1.75	55,1239

## 5.7 Sinal original com adição de ruído branco

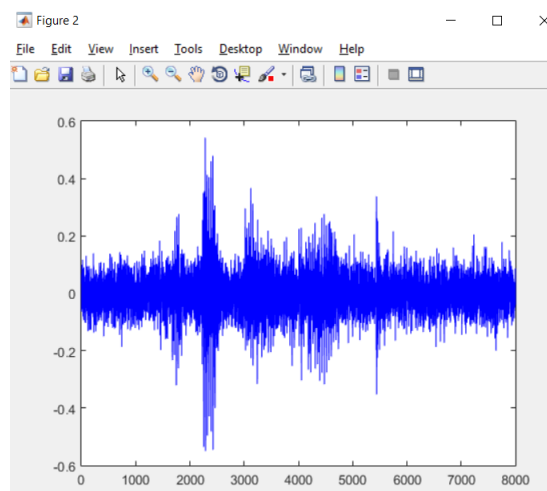


Figura 11: Sinal original com adição de ruído branco a 0 dB

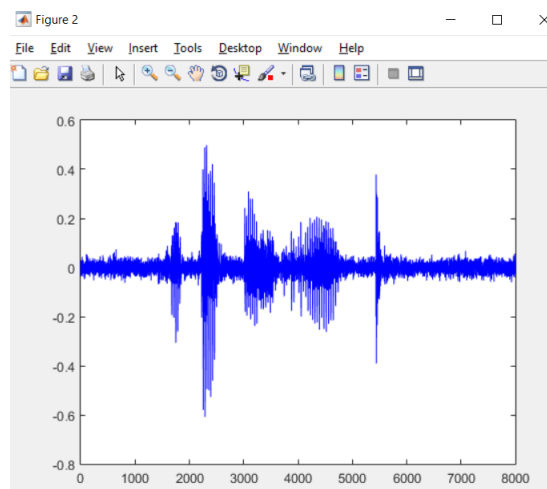


Figura 12: Sinal original com adição de ruído branco a 10 dB

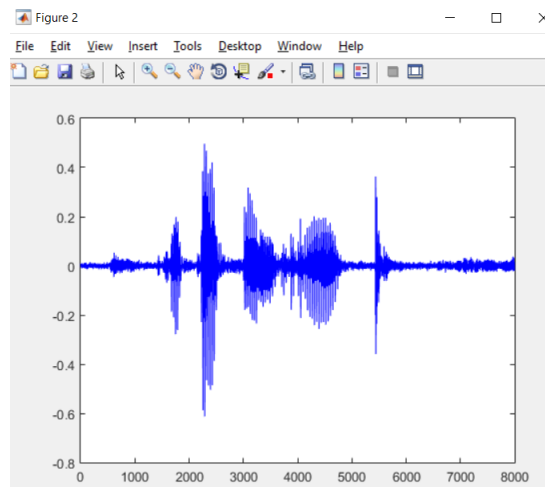


Figura 13: Sinal original com adião de ruído branco a 20 dB

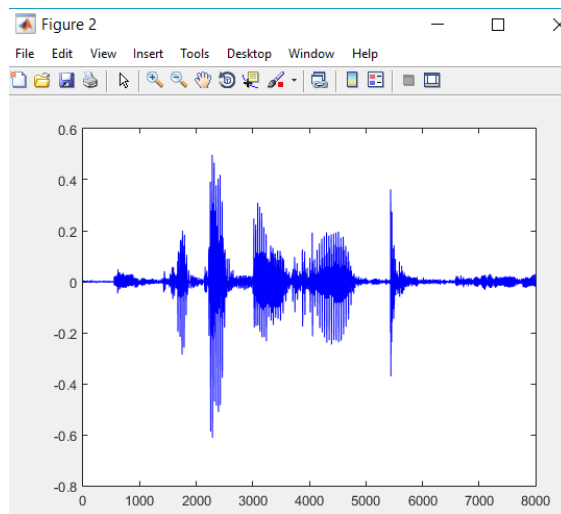


Figura 14: Sinal original com adião de ruído branco a 30 dB

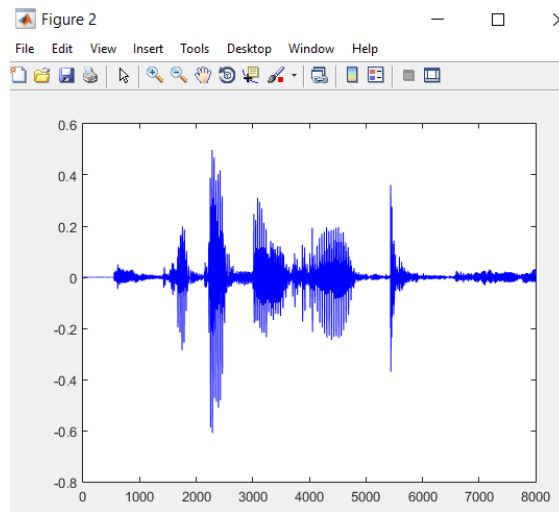


Figura 15: Sinal original com adião de ruído branco a 40 dB

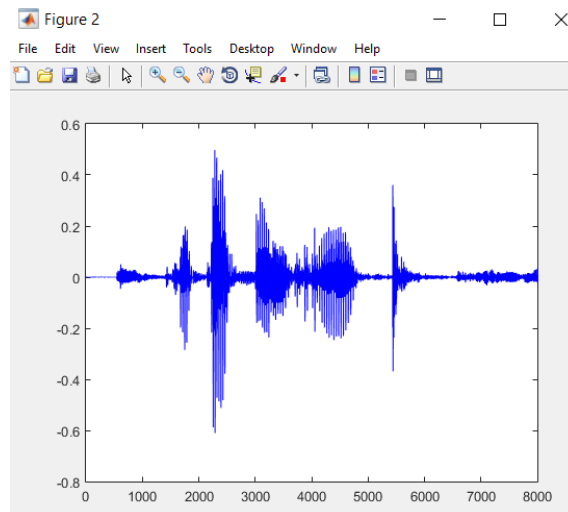


Figura 16: Sinal original com adião de ruído branco a 50 dB

## 6 Conclusão

Este trabalho ajudou a perceber o tratamento que levam os sons que ouvimos provenientes de rádios e televisões.

Após alterar os valores de threshold para o mesmo ruído branco, conclui-se que o SNR diminui. No entanto, à medida que o threshold aumenta, o sinal passa a ser considerado cada vez mais ruído. Isto deve-se ao facto dos outliers da distribuição normal serem alargados.

Através das figuras 11, 12, 13, 14, 15 e 16 também se pode concluir que à medida que a potência do sinal de ruído branco aumenta, esse mesmo ruído perde influência relativamente ao sinal original.

No final, adaptando os parâmetros da função para cada o sinal de áudio, foi exequível desagregar a parte da fala da parte do ruído do sinal.

É importante notar também que este algoritmo destina-se a remover principalmente sinais de background desnecessários, pelo que para processos mais complexos outros algoritmos teriam de ser usados.