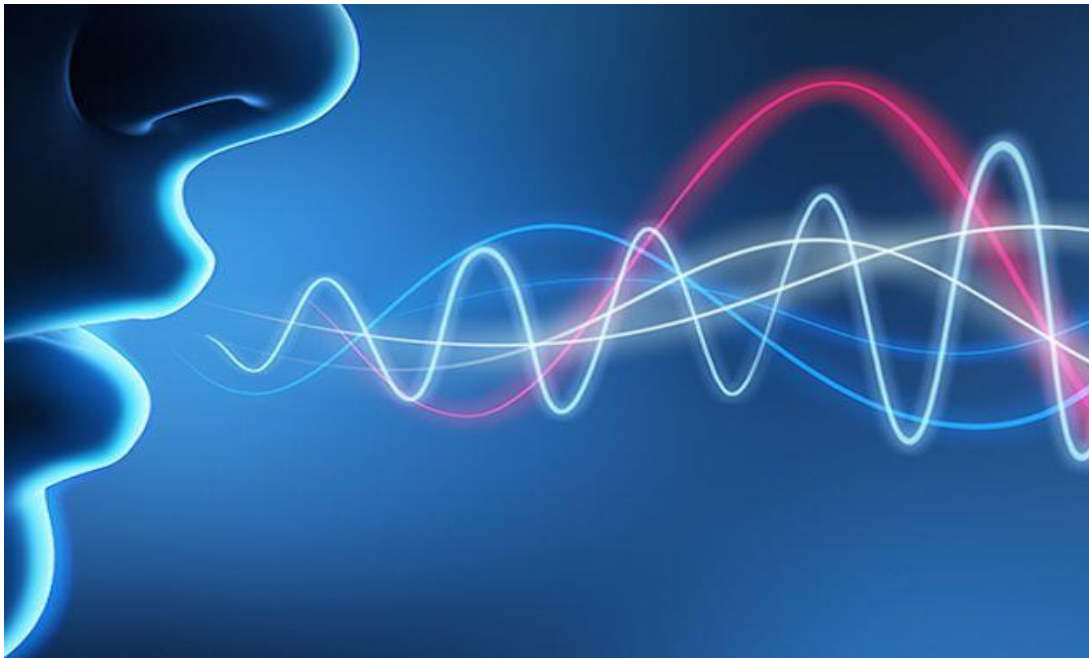




Universidade do Minho
Escola de Engenharia

Processamento Digital de Sinal

ALGORITMO PARA A DETEÇÃO DE SILÊNCIOS NA FALA



Carlos Miguel Gomes Mendes, A74338

Mestrado Integrado em Engenharia Eletrónica Industrial e Computadores

ÍNDICE

Introdução	3
Conceitos Teóricos.....	4
SNR – Signal to Noise Ratio	4
Ruído Branco	4
Processos Estocásticos	4
Modelo Gaussiano.....	5
Implementação.....	6
Testes.....	7
Teste 1 -> 10dB	7
Teste 2 -> 20dB	7
Teste 3 -> 30dB	8
Teste 4 -> 40dB	8
Teste 5 -> 50dB	9
Conclusão.....	10

INTRODUÇÃO

Este trabalho foi proposto no âmbito da disciplina de Processamento Digital de Sinal, com o objetivo de detectar os silêncios na fala através de um algoritmo em Matlab.

O estudo de cada relação sinal-ruído requer síntese de ruído branco, soma ao sinal e detecção dos segmentos contendo apenas ruído usando vários valores dos parâmetros. Através de processos estocásticos (processos aleatórios que dependem do tempo), retiramos o Modelo do Ruído (média, variância).

Ao analisar o sinal da fala, tudo o que estiver dentro dos parâmetros do Modelo que será apresentado, é ruído, o que estiver no seu outlier (fora dos limites) é fala. Desta forma é possível distinguir as componentes do sinal que são fala e as que são ruído.

CONCEITOS TEÓRICOS

SNR – Signal to Noise Ratio

A relação sinal-ruído é usado em medidas de um sinal no meio ruidoso e define a razão da potência de um sinal e a potência do ruído sobreposto ao sinal, comparando o nível de um sinal desejado com o nível do ruído de fundo.

Esta relação é normalmente usada em sinais elétricos, mas também pode ser utilizada em sinais acústicos e óticos. Quão mais elevada for a relação sinal-ruído, menor será o impacto do ruído de fundo sobre a medição ou detecção do sinal.

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

Figura 1 - Definição SNR

Ruído Branco

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 ruído em analogia ao funcionamento da luz branca, dado que esta é obtida por meio da junção em simultâneo de todas as frequências cromáticas.

Em estatística, este aplica-se a uma sequência de erros aleatórios, sempre que esta tiver média e variância constantes e não apresentar autocorrelação. O valor da média, por conveniência, será zero, podendo, no entanto, especificar-se um outro valor. Assim, o ruído branco será temporalmente homogêneo, estacionário e não dependente de instantes anteriores.

Processos Estocásticos

Um processo estocástico é utilizado para analisar sinais em que não é possível determinar o seu conteúdo previamente. Como é possível saber o comportamento que apresentará, a partir 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.

No caso do ruído é possível saber os limites a partir dos quais um segmento do sinal é efetivamente ruído ou contém informação relevante, pois este apresenta as seguintes características:

- Distribuição normal de media igual a zero
- Variância constante
- Covariância nula
- Simetria
- É não correlacionado entre diferentes estantes.

Modelo Gaussiano

Uma Distribuição Normal 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.

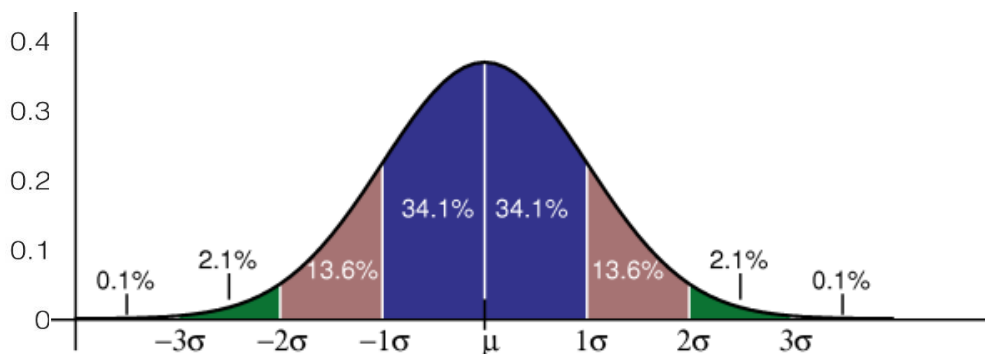


Figura 2- Representação gráfica de uma distribuição normal, em que μ e σ , representam, respetivamente, a média e o desvio padrão.

A distribuição Gaussiana 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.

IMPLEMENTAÇÃO

A implementação deste algoritmo começa-se por efetuar uma gravação de um sinal áudio. De seguida passamos a análise do sinal e então fazer o cálculo do ruído e separar este da componente fala.

O próximo passo é o cálculo do valor médio e variância, e como esses valores comparar as amostras para verificar se podem ou não ser consideradas ruído. Caso surjam erros como considerar ruído quando é fala e vice-versa. Para combater estes erros alteramos o método, em vez de uma análise amostra a amostra, fazemos a análise a segmentos. Faz-se a contabilização do número de amostras dentro dos valores da curvatura e fora, caso haja mais amostras dentro é ruído, se não é fala.

TESTES

Para a realização dos testes gravou-se um sinal áudio e realizaram-se os cinco testes seguintes com os diferentes valores do SNR.

Teste 1 -> 10dB

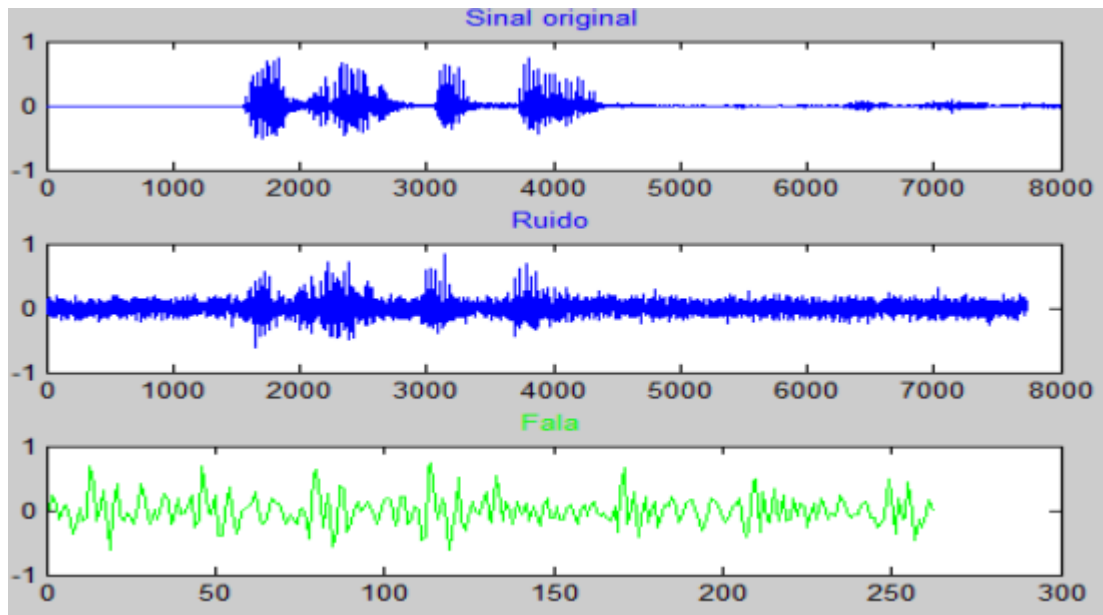


Figura 3 - Resultados obtidos com 10dB

Teste 2 -> 20dB

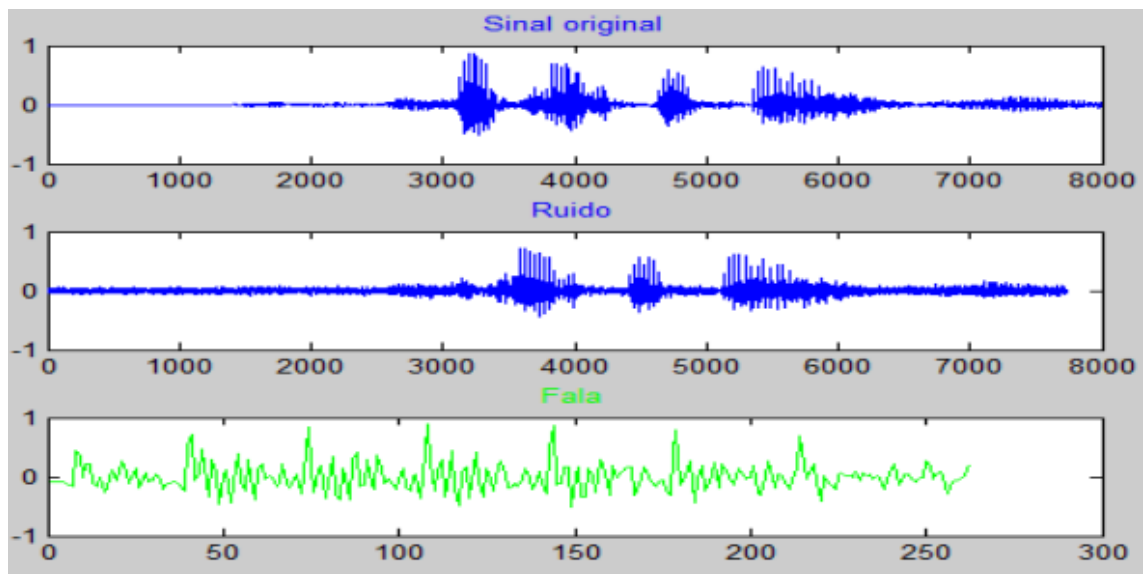


Figura 4 - Resultados obtidos com 20dB

Teste 3 -> 30dB

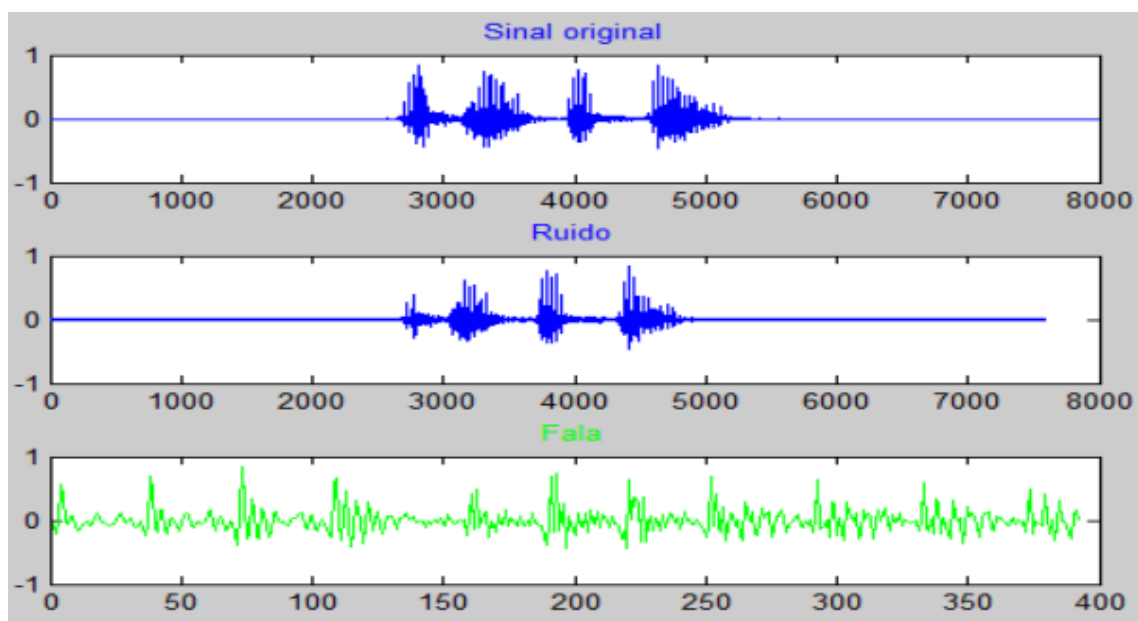


Figura 5 - Resultados obtidos com 30dB

Teste 4 -> 40dB

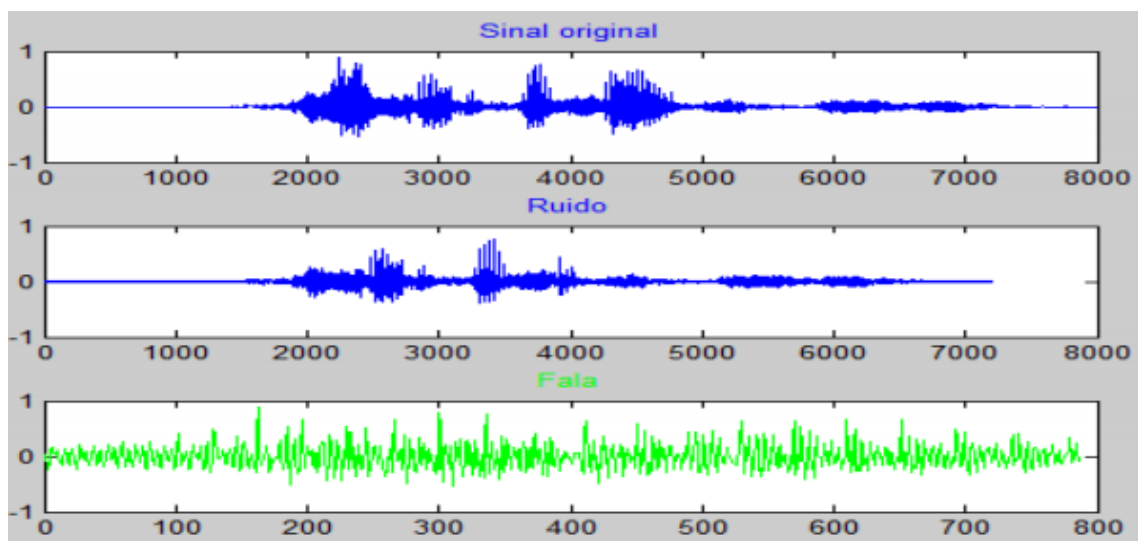


Figura 6 - Resultados obtidos com 40dB

Teste 5 -> 50dB

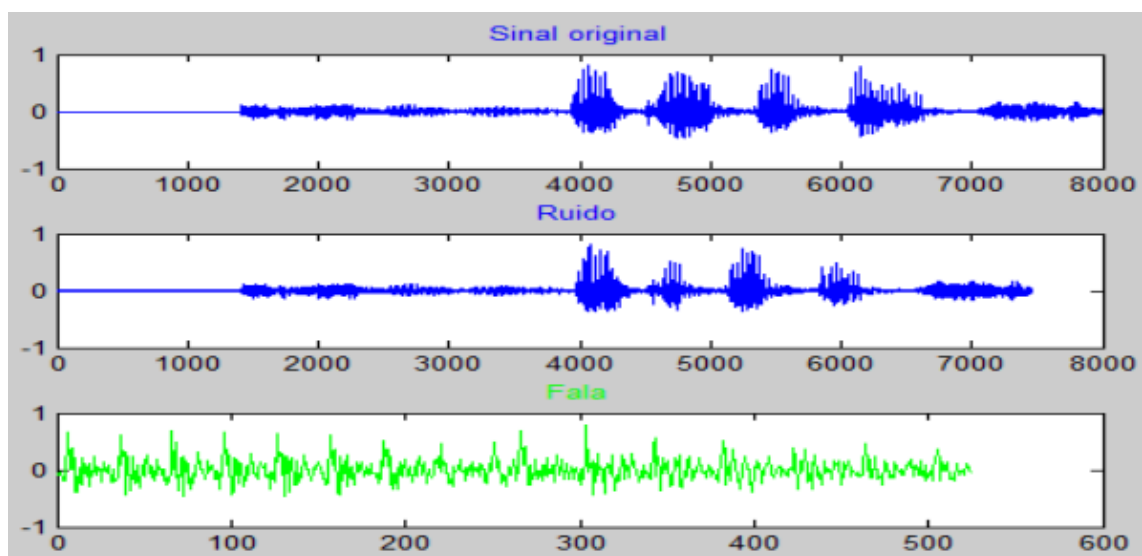


Figura 7 - Resultados obtidos com 50dB

CONCLUSÃO

Através da realização deste trabalho foi possível consolidar, de uma forma geral, todos os conhecimentos, em particular da ferramenta de desenvolvimento (*Matlab*), abordados nas aulas da unidade curricular de Processamento Digital de Sinal.

Após a realização dos testes ao algoritmo, pode-se verificar melhor o seu funcionamento, conforme a gravação, previamente.

Podendo-se, então, concluir que no caso do fator, caso este seja baixo, será guardado no “buffer_fala” todos os sons vozeados, mas também o ruído e os silêncios. Caso seja alto, poderá ocorrer a perda de alguma informação importante.

No final, adaptando os parâmetros da função para cada sinal de áudio, foi possível separar a parte da fala da parte do ruído do sinal, tendo assim um sinal mais pequeno e só com a informação desejada.