

## Trabalho 2 – Modelos Probabilísticos

Thiago Koster Lago

Universidade Federal do Rio de Janeiro

[thiagokoster@poli.ufrj.br](mailto:thiagokoster@poli.ufrj.br)

<https://github.com/ThiagoKoster/Trabalho-2-Modelos>

### 1. Introdução

Síntese de voz é um processo de produção artificial de voz humana, utilizando um processo computadorizado implementado por software ou hardware. Neste trabalho, implementaremos a sintetização do fonema “ss” utilizando MATLAB.

No repositório se encontram os áudios originais e sintetizados para os fonemas “SS” E “CH”.

### 3. Função Periodogram()

Utilizando o código fornecido no artigo, criei uma função chamada Periodogram() que tem como objetivo traçar os gráficos necessários para análise e retornar os parâmetros que serão utilizados na construção do sinal sintetizado.

Essa função calcula os parâmetros  $\hat{a}[k]$  pelas equações de Yule-Walker, utiliza esses parâmetros para calcular a variância do ruído branco e traça o Periodograma do sinal de entrada e seu modelo AR PSD.

Os parâmetros de entrada da função são: O sinal a ser analisado, o número de polos para o filtro e uma *string* que será utilizada como título para o gráfico do Periodograma.

### 2. Método

Para realizar a síntese do fonema “ss”, primeiro gravei um trecho de fala pronunciando a palavra *seven*. Utilizando o software *Audacity*, separei o fonema “ss” do resto do áudio. Criando desta forma um arquivo de áudio contendo apenas o fonema desejado.

Utilizei o arquivo de áudio produzido e fiz uma amostragem de 11025 Hz.

Utilizando a função *Periodogram* criada no sinal amostrado “y\_r” foi possível obter os parâmetros do fonema e a variância do ruído branco. Com esses dados, foi criado o ruído branco  $u(n)$  e foi gerado o sinal  $x(n)$  utilizando a função de transferência

$$H_{uv} = \frac{1}{1 - \sum_{k=1}^p a[k]z^{-k}} \text{ e o ruído } u(n) \text{ como}$$

entrada do filtro.

Este filtro foi implementado utilizando a função *filter* do MATLAB.

De posse do sinal  $x(n)$ , utilizei a função *Periodogram* novamente para traçar os gráficos com intuito de realizar a comparação entre o sinal original amostrado e o sinal obtido pelo filtro.

### 3. Resultados

A seguir serão apresentados os resultados obtidos pelo método descrito a cima.

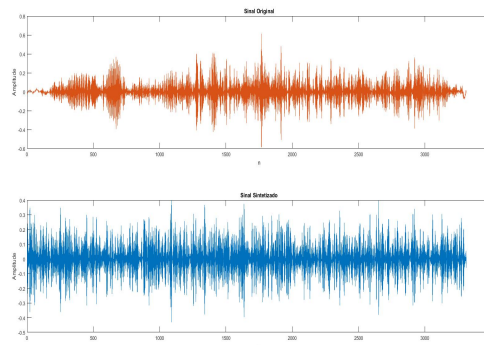
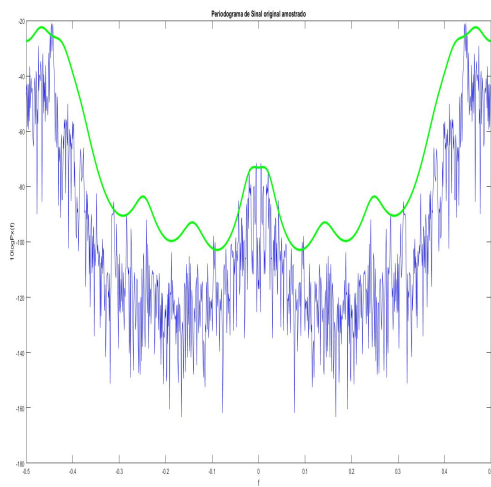


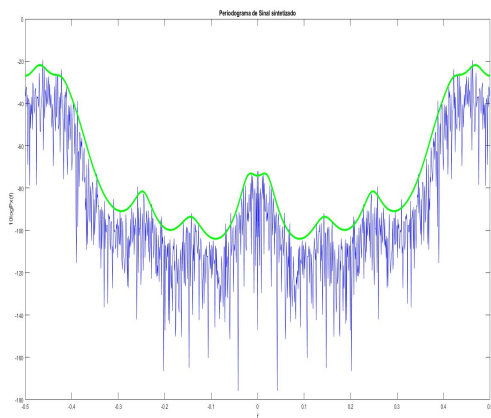
Figura 1: Sinal original amostrado (laranja) e sinal sintetizado (azul)

Como pode-se perceber, o sinal em azul possui padrões semelhantes ao original e com variância semelhante. Porém, comparar os sinais no tempo não é um método confiável pois o ouvido humano não é sensível a fase. Apesar das diferenças, os sinais são semelhantes ao ouvir os arquivos de áudios gravados.

Portanto para se obter uma comparação mais confiável, analisaremos os periodogramas e os modelos AR PSD (*Autoregressive Power Spectral Density*) para se obter melhores análises.



*Figura 2: Periodograma (azul) e modelo AR PSD (verde) do sinal original*



*Figura 3: Periodograma (azul) e modelo AR PSD (verde) do sinal sintetizado*

Agora de posse dos modelos AR PSD, podemos verificar a semelhança entre os dois sinais. Ambos modelos possuem mesmos picos com valores bem próximos.

Ouvindo os arquivos de áudio da amostra original e sintetizada não foi possível notar diferença entre os sinais. Tornando assim, o resultado do experimento satisfatório.

#### 4. Conclusão

Com este trabalho, pode-se concluir que por mais que os sinais no tempo não sejam idênticos, eles podem ser bem semelhantes devido a falta de sensibilidade à fase do ouvido humano. E que ao se analisar a densidade de potência espectral é possível perceber tal semelhança.

Ao ouvir os dois áudios, não foi possível captar nenhuma diferença tanto para o fonema “ss” apresentado neste trabalho, quanto para o fonema “ch” salvo no repositório evidenciando desta forma, o sucesso do experimento.