

Introdução ao Processamento de Voz e Fala

Adelino Pinheiro Silva

Universidade Federal de Minas Gerais (UFMG)
Faculade de Letras (FALE)
Introdução à Linguística Computacional

17 de dezembro de 2024



Sumário

- 1 Sumário
- 2 Aplicações
- 3 Representação digital do som
- 4 Análise em frequência
- 5 Medição no sinal de voz e fala
- 6 Encerramento
- 7 Referências

Assuntos

1 Sumário

2 Aplicações

3 Representação digital do som

4 Análise em frequência

5 Medições no sinal de voz e fala

6 Encerramento

7 Referências

Acessibilidade, Inclusão e Educação

- Interface homem máquina.
- Tradução e interpretação.
- Conversão de texto para voz e voz para texto.
- Assistentes inteligentes.



Imagen: <https://eletronjun.com.br/wp-content/uploads/2021/03/image1-1-edited.png>

Comunicação, Saúde e Segurança

- Comunicação a distância.
- Triagem e diagnóstico de distúrbios da voz.
- Avaliação de fluência.
- Biometria.
- Comparação Forense de Locutor.



Imagen: <https://encrypted-tbn0.gstatic.com/images>

Natureza, Entretenimento, Mimetização

- Reconhecimento de espécies.
- Psicoacústica.
- Autotune.
- *Deep fakes*.
- Dublagem automática.



Imagen: https://classic.exame.com/wp-content/uploads/2016/09/size_960_16_9_twitter-passaros1.jpg

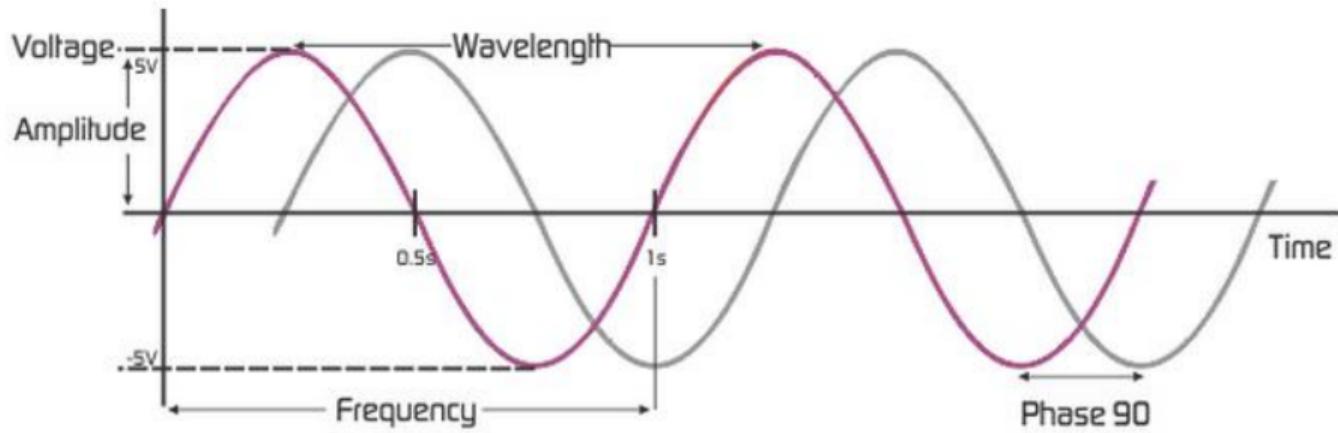
Assuntos

- 1 Sumário
- 2 Aplicações
- 3 Representação digital do som
- 4 Análise em frequência
- 5 Medição no sinal de voz e fala
- 6 Encerramento
- 7 Referências

Sinal da voz e da fala

Onda acústica: onda de pressão longitudinal por compressão e descompressão.

- Amplitude.
- Frequência.
- Fase.



Da natureza para o computador

No quadro...

- Amostragem
taxa de amostragem e limitação de banda.
- Quantização
ruído $SNR(dB) = 6B - 7,2dB$.
- Número de canais
mono, stereo, 5.1 ...
- Armazenamento
tamanho do arquivo WAV PCI.

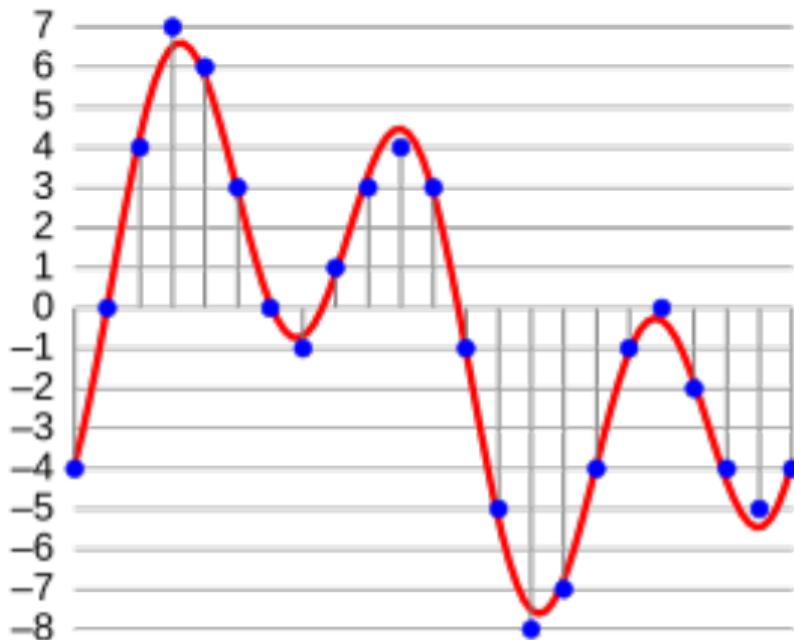


Imagem: <https://upload.wikimedia.org/wikipedia/commons/thumb/2/21/4-bit-linear-PCM.svg>

Assuntos

- 1 Sumário
- 2 Aplicações
- 3 Representação digital do som
- 4 Análise em frequência**
- 5 Medição no sinal de voz e fala
- 6 Encerramento
- 7 Referências

Definição para ondas periódicas

Decomposição de uma onda “complexa” em um conjunto de ondas simples.

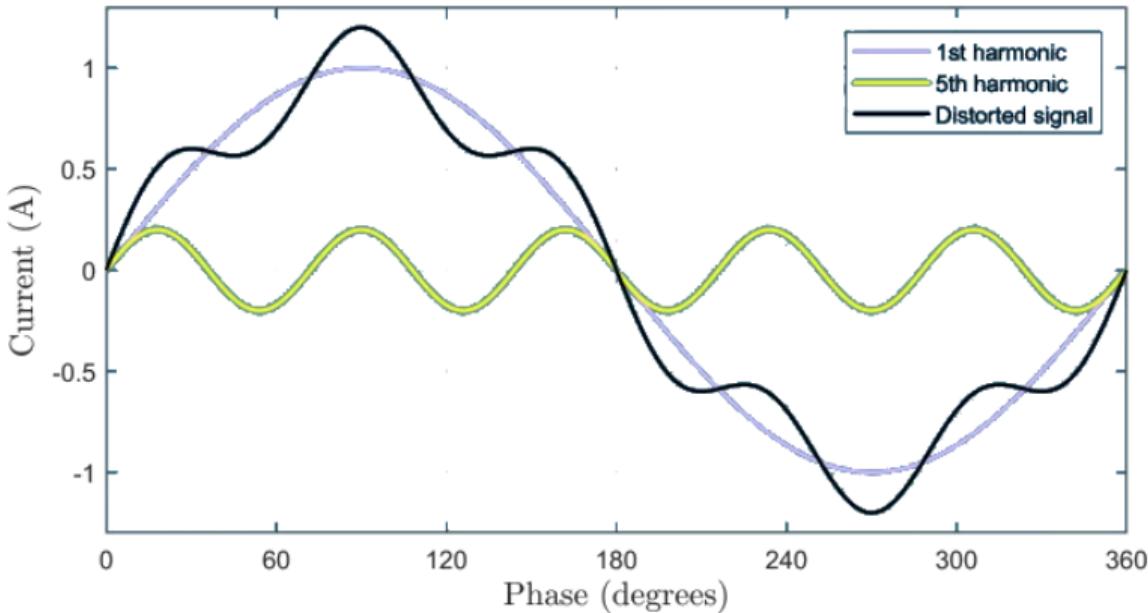


Imagem: <https://power-convertisers.net/wp-content/uploads/2022/03/>

A partir de uma frequência fundamental...

Múltiplos inteiros com mudança de amplitude e fase

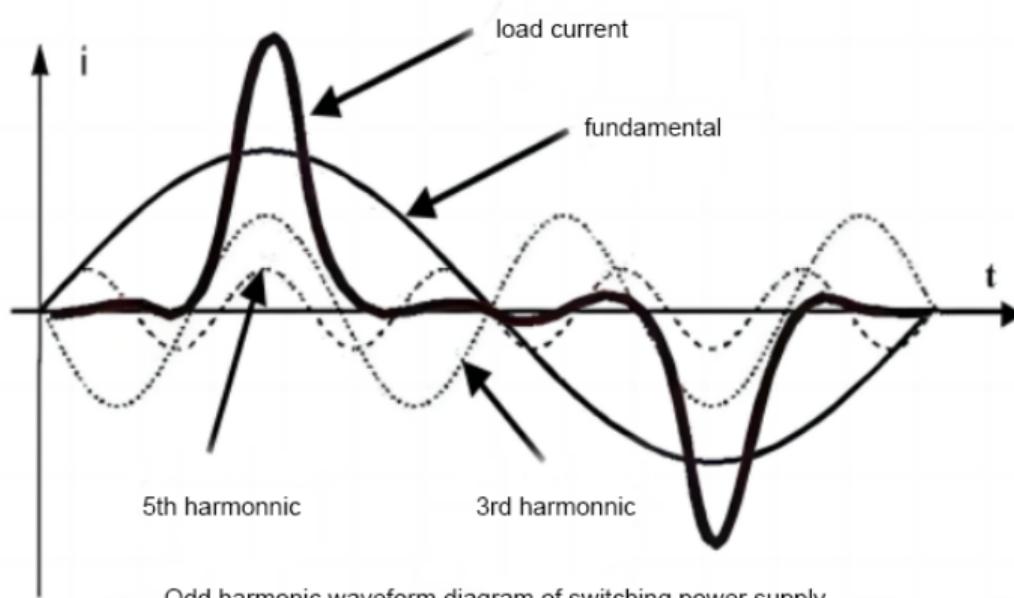


Imagem: <https://www.zddqelectric.com/js/htmledit/kindeditor/attached/20240530/>

Ferramentas matemáticas

FFT (*fast Fourier transform*)

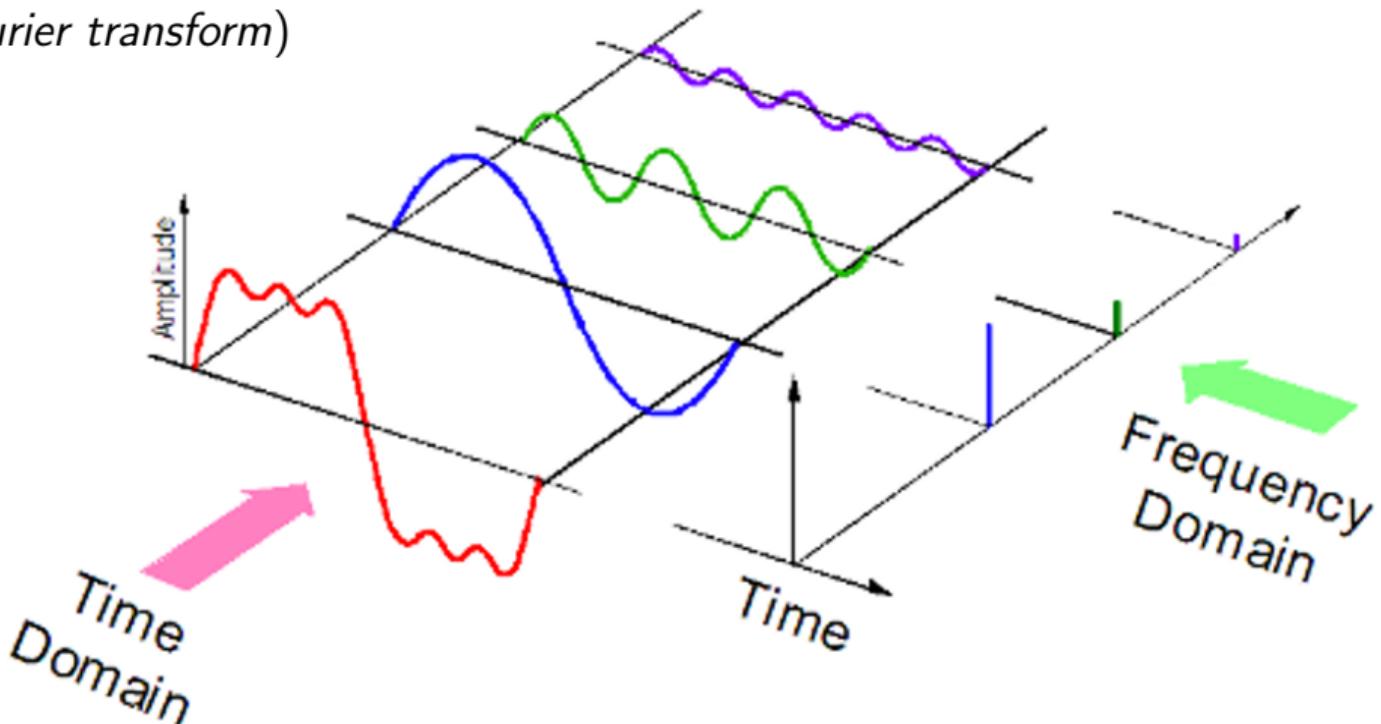


Imagem: https://miro.medium.com/v2/resize:fit:1358/1*w1mcpi8gCgQI4FrqWFJ2GA.png

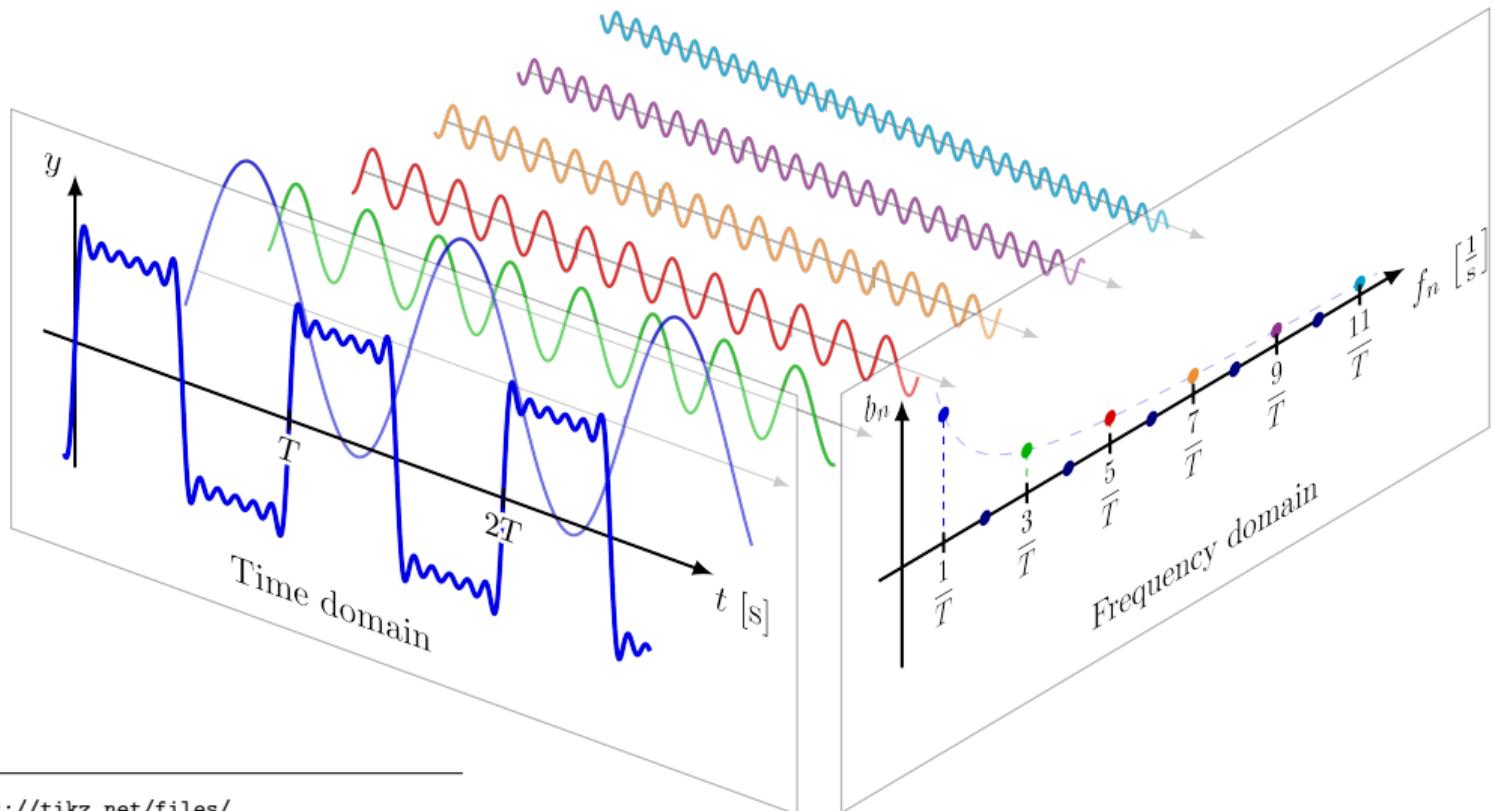
FFT (*fast Fourier transform*)

Imagen: <https://tikz.net/files/>

Espectrograma

FFT de pequenos pedacinhos do áudio

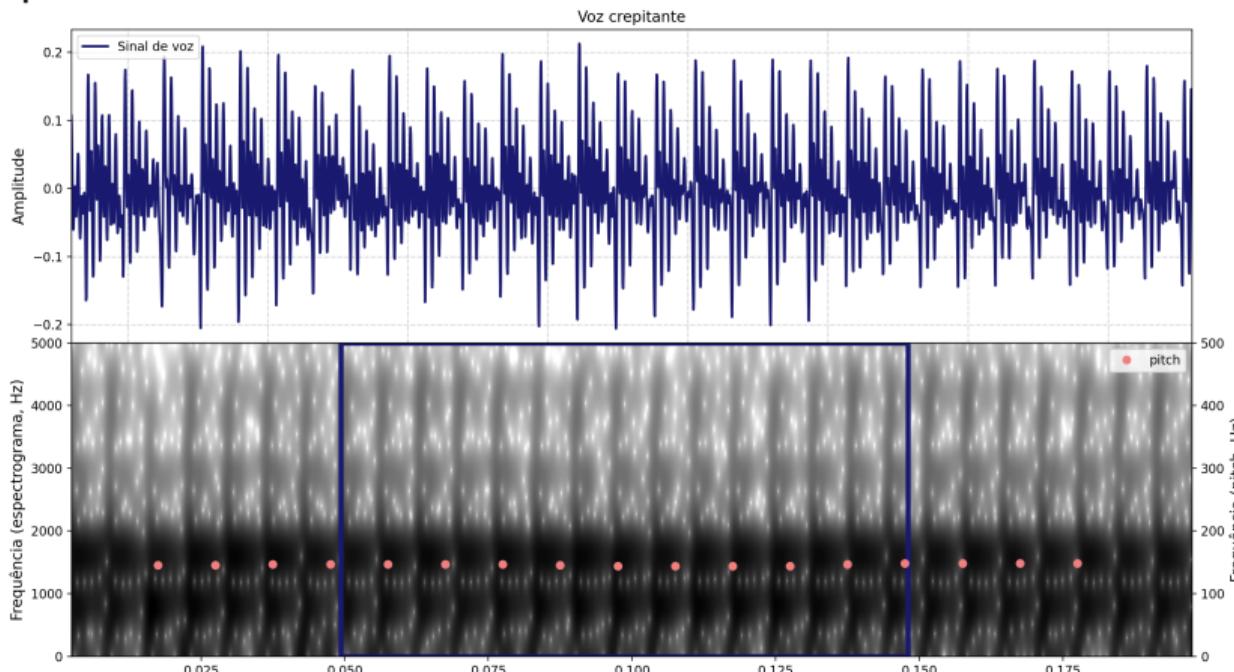


Imagem: http://www.poslin.letras.ufmg.br/diss_defesas_detalhes.php?aluno=2312

Espectrograma... No quadro

Relação de compromisso entre tempo e frequência

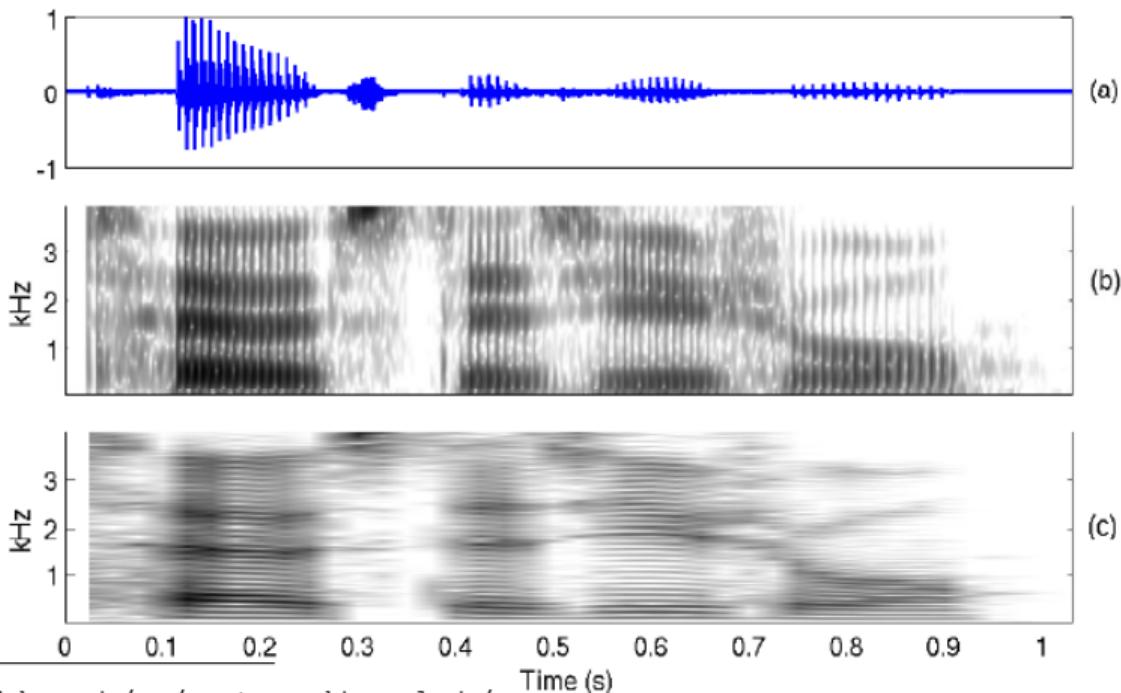


Imagem: <https://ssp-iiith.vlabs.ac.in/exp/spectrographic-analysis/>

Assuntos

- 1 Sumário
- 2 Aplicações
- 3 Representação digital do som
- 4 Análise em frequência
- 5 Medições no sinal de voz e fala
- 6 Encerramento
- 7 Referências

Começando do básico

- Duração.
- Intensidade.
- Frequência fundamental.
- Formantes.

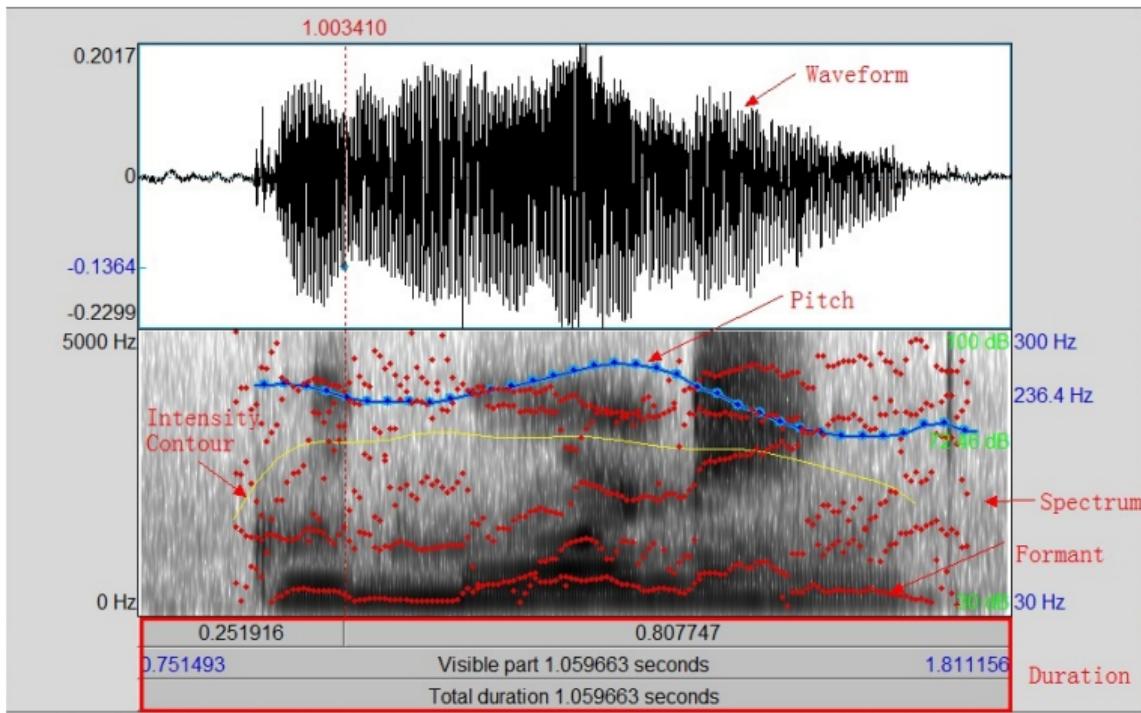
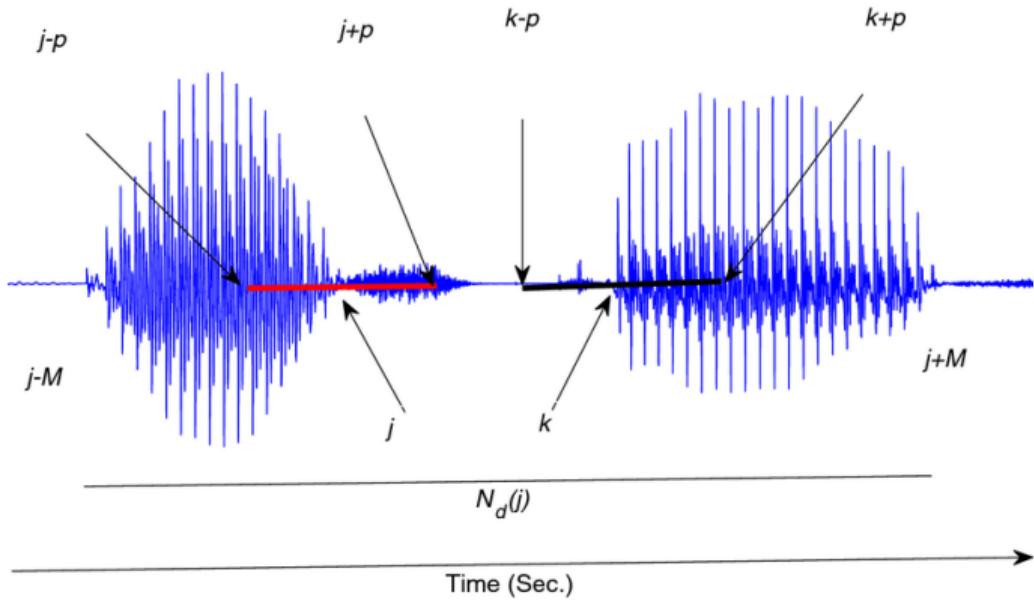


Imagem: https://corpus.edu.hk/english_pronunciation/wp-content/uploads/2020/03/

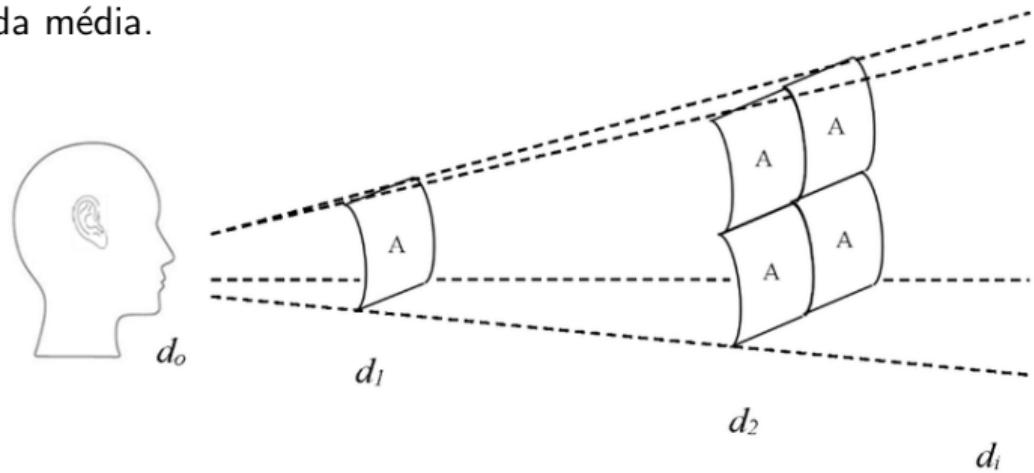
Duração

- Como é medido:
 - Intervalo entre início e fim.
- Para que serve:
 - Localização de acentos.
 - Interpretação de contexto.
 - Fluência de leitura.



Intensidade

- Como é medido:
 - Na onda: valor da raiz quadrada média.
 - SPL: razão logarítmica.
- Para que serve:
 - Localização de acentos.
 - Capacidade de fala.
 - Distúrbios respiratórios.



$$SPL(dB) = 10 \log_{10} \left(\frac{P_{med}}{p_0} \right)^2$$

$$p_0 = 20 \mu Pa$$

Frequência fundamental

- Como é medido:
 - Correlação.
 - Espetrograma.
- Para que serve:
 - Estado emocional.
 - Prosódia.
 - Distúrbios da fala.
 - Parâmetro biométrico.

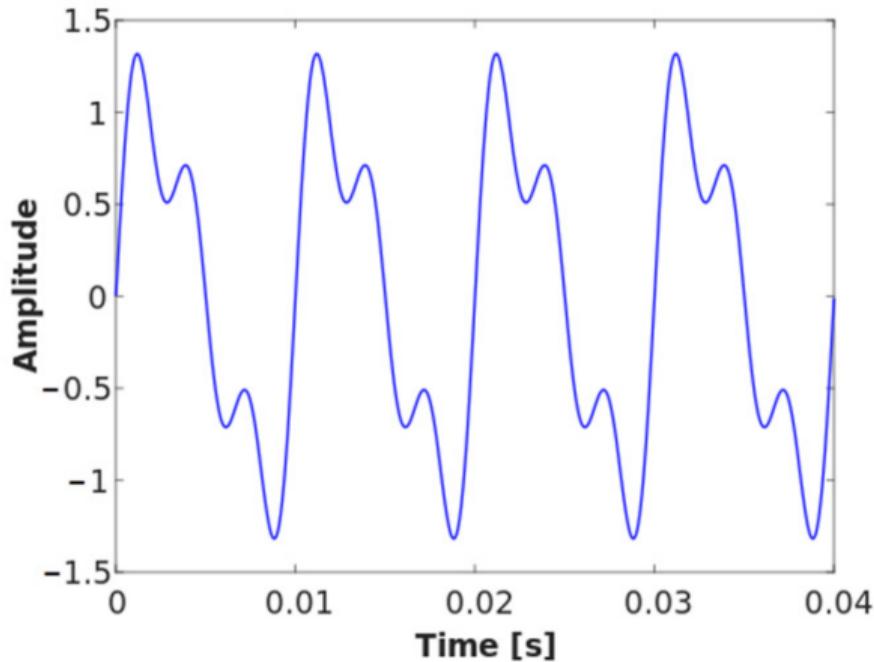


Imagen: <https://www.mdpi.com/applesci/applesci-13-08191>

Formantes

- Como é medido:
 - FFT e análise LPC.
- Para que serve:
 - Avaliação de fluência.
 - Distúrbios da fala.
 - Biometria.
 - Identificação da fala.

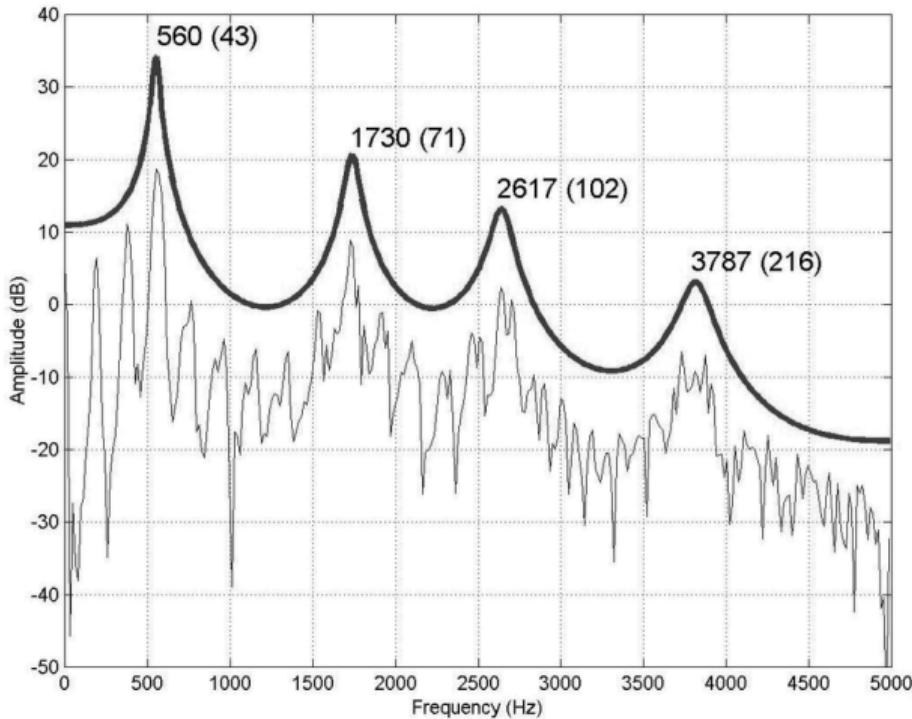


Imagem: <https://www.researchgate.net/profile/Phil-Rose/publication/274326216/>

Assuntos

1 Sumário

2 Aplicações

3 Representação digital do som

4 Análise em frequência

5 Medidas no sinal de voz e fala

6 Encerramento

7 Referências

Agradecimentos

Fim!

Contatos:

- e-mail: adelinocpp@gmail.com, adelinocpp@yahoo.com, adelino@ufmg.br ou adelino.pinheiro@policiacivil.mg.gov.br;
- Whatsapp (31) 98801-3605;
- Instituto de Criminologia - Academia de Polícia Civil de Minas Gerais. Rua Oscar Negrão de Lima, 200, Nova Gameleira, Belo Horizonte-MG. Tel.: (31) 3314-5620;
- Setor de Perícias em Áudio e Vídeo - Instituto de Criminalística. Av. Augusto de Lima, 1833, Barro Preto, Belo Horizonte-MG. Tel.: (31) 3330-1887;

Sobre este material

Esta obra está licenciada sob a licença *Creative Commons CC BY-NC-SA 4.0*

Favor fazer referência a este trabalho como:

Silva, A. P. (2024), *Introdução ao Processamento de Voz e Fala*. Online:

<https://github.com/adelinocpp/processamento-voz-fala>

```
@Misc{Silva2024,  
  title = {Introdução ao Processamento de Voz e Fala},  
  author = {Adelino Pinheiro Silva},  
  howPublished = {\url{https://github.com/adelinocppprocessamento-voz-fala}},  
  year = {2024},  
  note = {Version 1.0; Creative Commons BY-NC-SA 4.0.},  
}
```



Dúvidas



Imagen: <https://www.hevcon.com.br/duvidas-frequentes-relacionado-ao-novo-bem-e-as-alteracoes-das-mps/>.

Assuntos

1 Sumário

2 Aplicações

3 Representação digital do som

4 Análise em frequência

5 Medição no sinal de voz e fala

6 Encerramento

7 Referências

Referências I