

Meta 3

1. Introdução

O objetivo desta meta é analisar características espectrais extraídas de gravações de áudio de dígitos pronunciados por diferentes locutores.

A análise visa diferenciar entre dígitos usando a Transformada de Fourier de Curto Prazo (STFT) e diversas técnicas de processamento de sinal.

2. Metodologia

Nota: A linguagem usada foi Python mas será disponibilizada em zip o ficheiro com o código em MATLAB. A razão é o facto de o MATLAB não correr na máquina que realizou o trabalho.

2.1 Parametrização da STFT:

Para cada dígito, a Transformada de Fourier de Curto Prazo (STFT) foi calculada utilizando a função *spectrogram* no MATLAB, porém, adaptada em Python. Diferentes parametrizações foram exploradas, incluindo tamanho da janela, sobreposição e número de pontos para o cálculo da FFT. Os parâmetros foram variados para identificar a configuração mais adequada para a tarefa.

2.2 Extração de Características de Potência:

Características de potência foram extraídas das representações STFT usando diferentes janelas tempo-frequência. Essas características capturam a distribuição de potência em diferentes bandas de frequência ao longo do tempo.

2.3 Análise das Características Extraídas:

As características de potência extraídas foram analisadas para identificar as mais discriminativas para a diferenciação de dígitos. Diversas técnicas de visualização, como box plots, gráficos 2D e 3D, foram utilizadas para explorar o espaço de características e identificar características discriminativas-chave.

3. Resultados

3.1 Otimização dos Parâmetros da STFT:

Por meio de experiências, os seguintes parâmetros da STFT foram considerados mais eficazes:

- Tamanho da Janela: 250 amostras;
- Sobreposição: 200 amostras;
- Número de Pontos para a FFT: Determinado automaticamente com base no tamanho da janela.

3.2 Extração de Características de Potência:

Várias características de potência foram extraídas das representações STFT, incluindo:

- Espectro de Potência Médio
- Espectro de Potência Máximo
- Espectro de Potência Mínimo
- Assimetria da Distribuição de Potência
- Curtose da Distribuição de Potência
- Planicidade da Distribuição de Potência
- Fluxo Espectral
- Dispersão Espectral
- Fator de Crista Espectral
- Entropia Espectral

3.3 Análise das Características Extraídas:

A visualização das características extraídas revelou que as seguintes características forneceram a melhor discriminação entre dígitos:

1. Fator de Crista Espectral
2. Entropia Espectral
3. Assimetria da Distribuição de Potência

4. Análise de Gráficos

4.1 Box Plots:

Os box plots foram utilizados para visualizar a distribuição das características espectrais entre os diferentes dígitos. Esta representação gráfica permite uma rápida compreensão das diferenças nas características de potência entre os dígitos.

4.2 Gráficos 2D e 3D:

Além dos box plots, gráficos 2D e 3D foram empregados para explorar as relações entre diferentes características espectrais. Essas visualizações proporcionam uma compreensão mais profunda das interações entre as características e sua influência na diferenciação dos dígitos.

4.3 Identificação das Características Mais Discriminativas:

Por meio da análise visual dos gráficos, identificamos as três características que permitem uma melhor discriminação entre os dígitos. Estas foram selecionadas com base na clareza das diferenças observadas entre os dígitos e na capacidade de separação das classes.

4.4 Gráficos e o que significam:

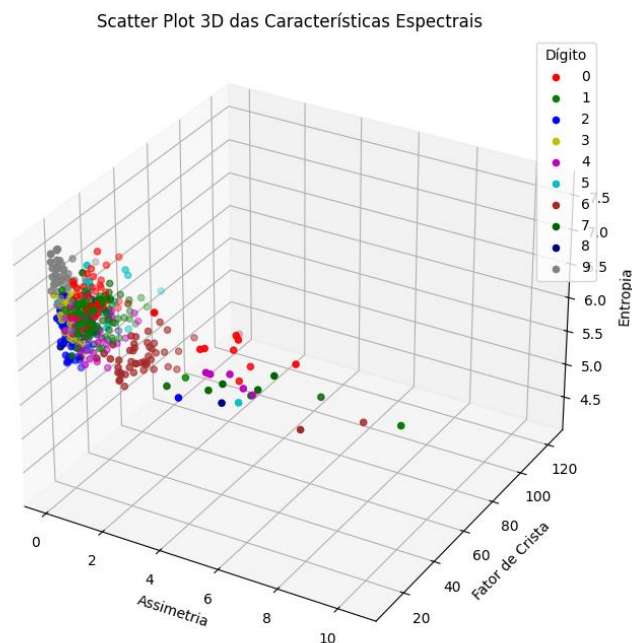


Figura 1: Representação gráfica das características de potência;

Analisando o gráfico 3D, notamos que o dígito 7 se destaca, estando isolado dos outros dígitos. Por outro lado, todos os outros dígitos se encontram juntos, o que mostra uma semelhança de valores. Este gráfico proporciona uma clara distinção entre dígitos, permitindo uma fácil diferenciação. A entropia mais alta encontra-se no dígito 9.

Quanto aos gráficos das potências, entropia, curtose, assimetria, etc, salientamos em todos o dígito 7, em que este se distingue da média em todos os casos.

5. Considerações Finais

A análise visual dos gráficos proporcionou insights valiosos sobre as características espectrais mais discriminativas para a diferenciação de dígitos. A combinação de técnicas de visualização, incluindo box plots, gráficos 2D e 3D, permitiu uma compreensão abrangente das relações entre as características e sua influência na separação das classes.

Esta análise visual complementa as análises quantitativas realizadas anteriormente e fornece uma perspectiva mais intuitiva sobre a eficácia das características espectrais na diferenciação de dígitos. As informações obtidas desta análise são fundamentais para o desenvolvimento de modelos de classificação robustos e precisos para a identificação automática de dígitos pronunciados por diferentes locutores.

Trabalho realizado por:

Carlos Miguel Matos Soares, 2020230124

