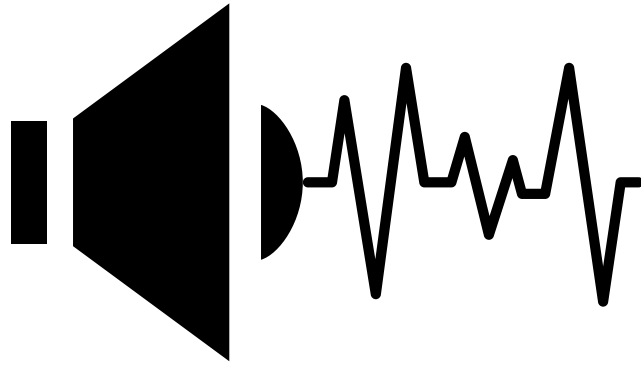


ANÁLISE E TRANSFORMAÇÃO DE DADOS



MINI-PROJETO IDENTIFICAÇÃO DE DÍGITOS ATRAVÉS DE SINAIS DE ÁUDIO

Ana Carolina Dos Santos Moraes
Nº2021222056

Fernanda Margarida Rodrigues Fernandes
Nº2021216620

Tiago Nuno Matos Cabo Lopes de Almeida
Nº2021221615

ÍNDICE:

Introdução	3
Gráficos - Exercício 4	4
Exercício 4.1/4.2/4.3	5
Exercício 5	7
Exercício 6	9
Conclusão	11

▼ INTRODUÇÃO/CONTEXTUALIZAÇÃO

Este mini-projeto tem como principal objetivo adquirir sensibilidade para as questões fundamentais da análise e transformação de dados, em particular, o cálculo da DFT e da STFT, para permitir distinguir os diferentes dígitos.

Como tal, os conceitos que aprendemos vão ser aplicados, através da elaboração de um código em "MATLAB" que permita responder às questões pedidas.

Ao longo deste trabalho são utilizados alguns conceitos teóricos, que serão a seguir apresentados, tais como: DFT (*discrete Fourier transform*) e a STFT (*short-time Fourier transform*).

É importante referir que para cada exercício (4,5,6), implementámos, digamos, uma função, na medida que facilita a observação e a compreensão do mesmo.

➡ DFT - *DISCRETE FOURIER TRANSFORM*

A Transformada Discreta de *Fourier* é uma ferramenta matemática fundamental utilizada para analisar o conteúdo de frequência de um sinal discreto no domínio do tempo. Ela transforma um sinal discreto do domínio do tempo num sinal discreto no domínio da frequência, permitindo a visualização das componentes de frequência presentes no sinal.

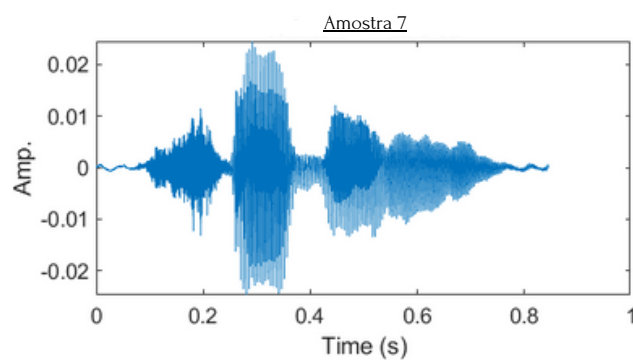
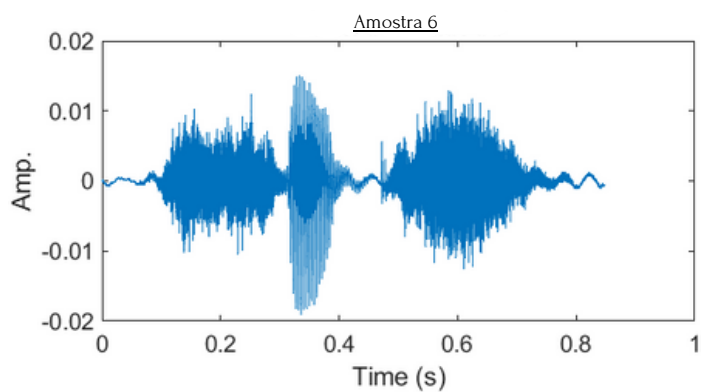
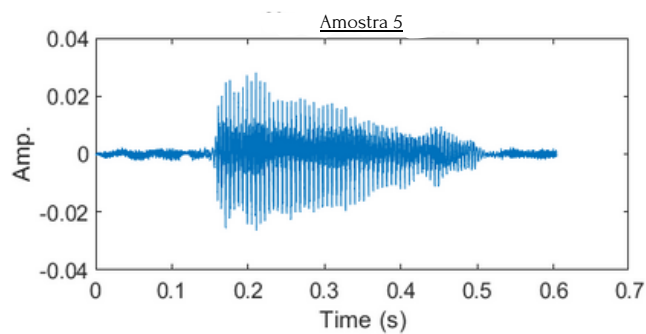
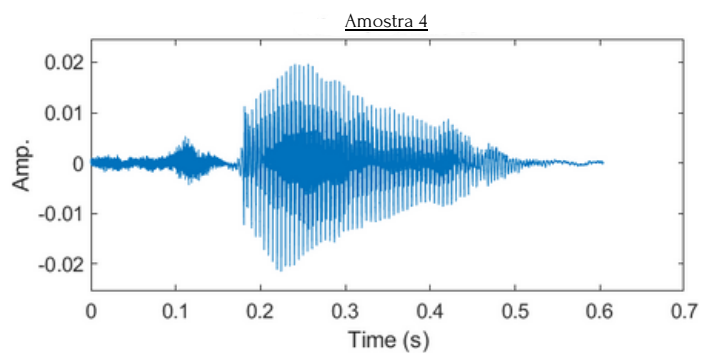
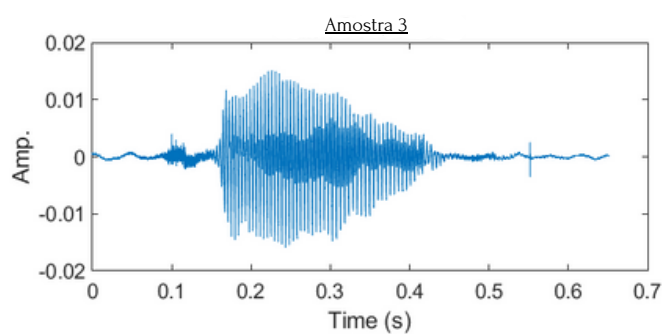
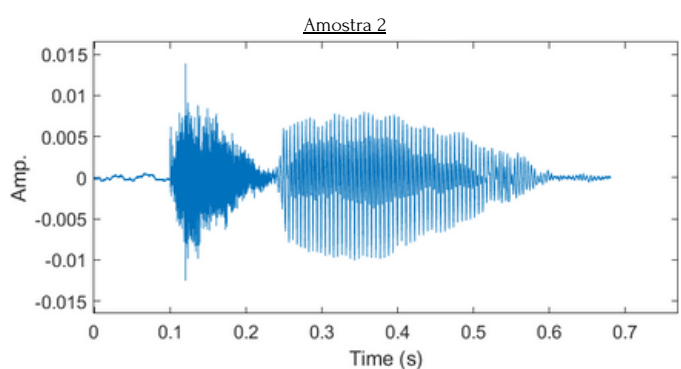
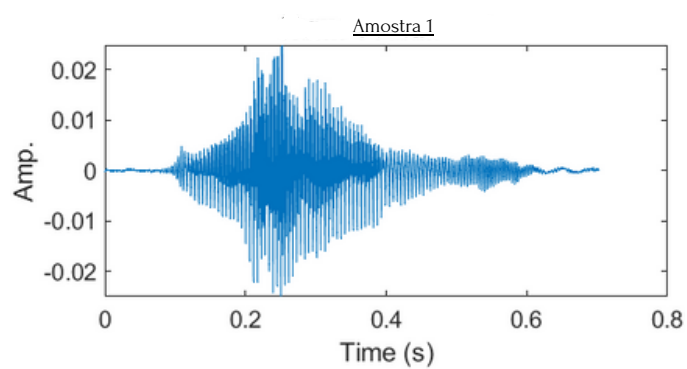
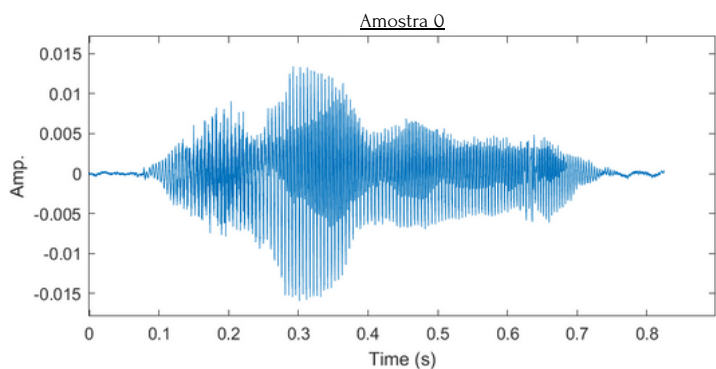
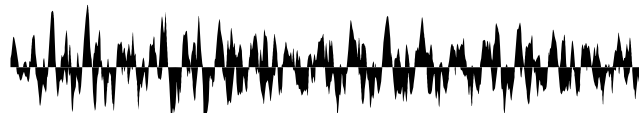
No contexto deste problema, a forma que utilizámos para o cálculo da DFT para cada dígito, foi a utilização da fórmula *FFT()* do "MATLAB". Como sabemos, a DFT produz um espectro de amplitude complexo, que contém informações sobre a amplitude e a fase de cada componente de frequência presente no sinal. Assim, para uma melhor visualização do espectro de amplitude, calculámos o valor absoluto.

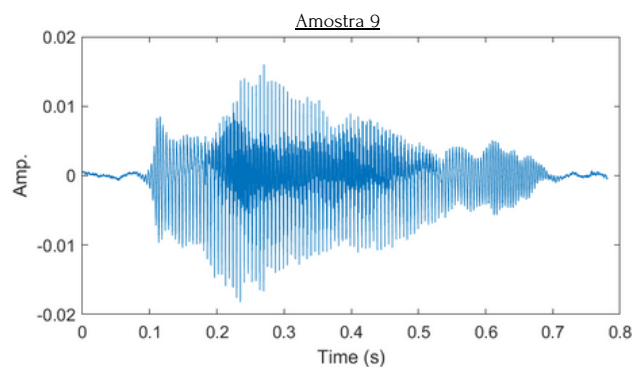
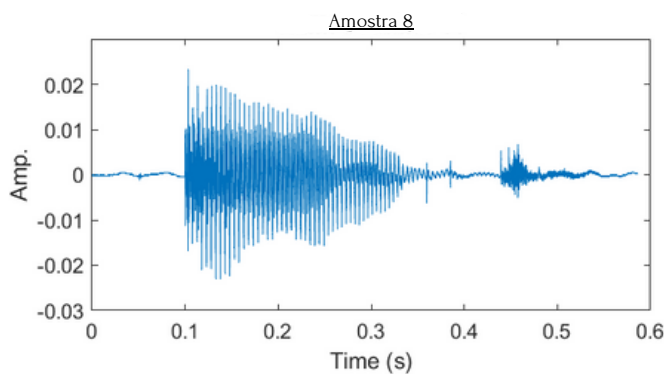
➡ STFT - *SHORT-TIME FOURIER TRANSFORM*

A Transformada de Fourier de Tempo Curto é uma técnica de análise de sinais que combina os benefícios da análise de frequência da Transformada de Fourier com a capacidade de análise de tempo do domínio do tempo. Ela permite analisar como as características de frequência de um sinal variam ao longo do tempo.

Esta Transformada divide o sinal de entrada em segmentos de curta duração e calcula a Transformada de Fourier em cada segmento. Isto é realizado devido à utilização de uma janela deslizante multiplicada pelo sinal original antes do cálculo da transformada. O resultado é uma representação em tempo-frequência do sinal, onde cada ponto representa a magnitude ou fase de uma determinada frequência num determinado momento. Ao ajustar os parâmetros da STFT, como o tamanho da janela e a sobreposição entre os segmentos, é possível controlar o equilíbrio entre a resolução temporal e a resolução espectral. Janelas maiores fornecem melhor resolução espectral, enquanto janelas menores fornecem melhor resolução temporal. No contexto deste problema, a forma que utilizámos para o cálculo da STFT para cada dígito, foi a utilização da fórmula *spectrogram()* do "MATLAB".

GRÁFICOS - EXERCÍCIO 4



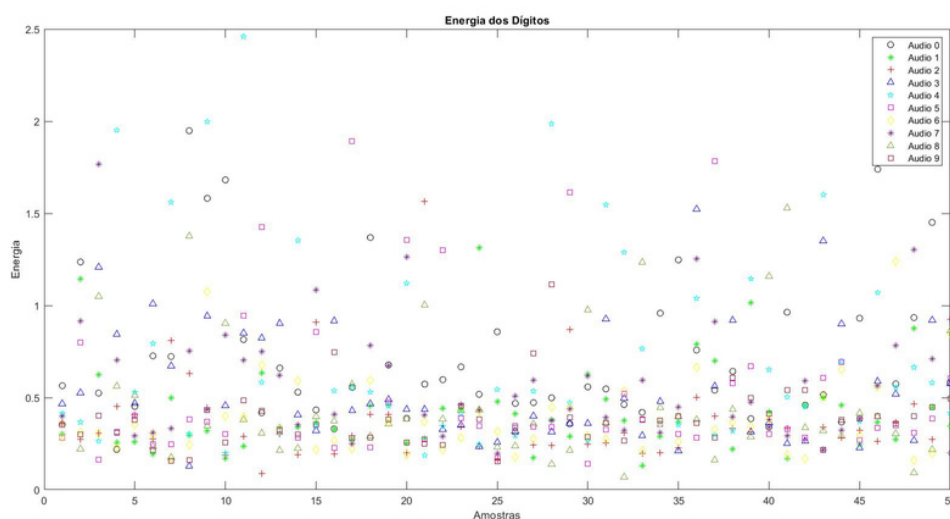


EXERCÍCIO 4.1/4.2/4.3

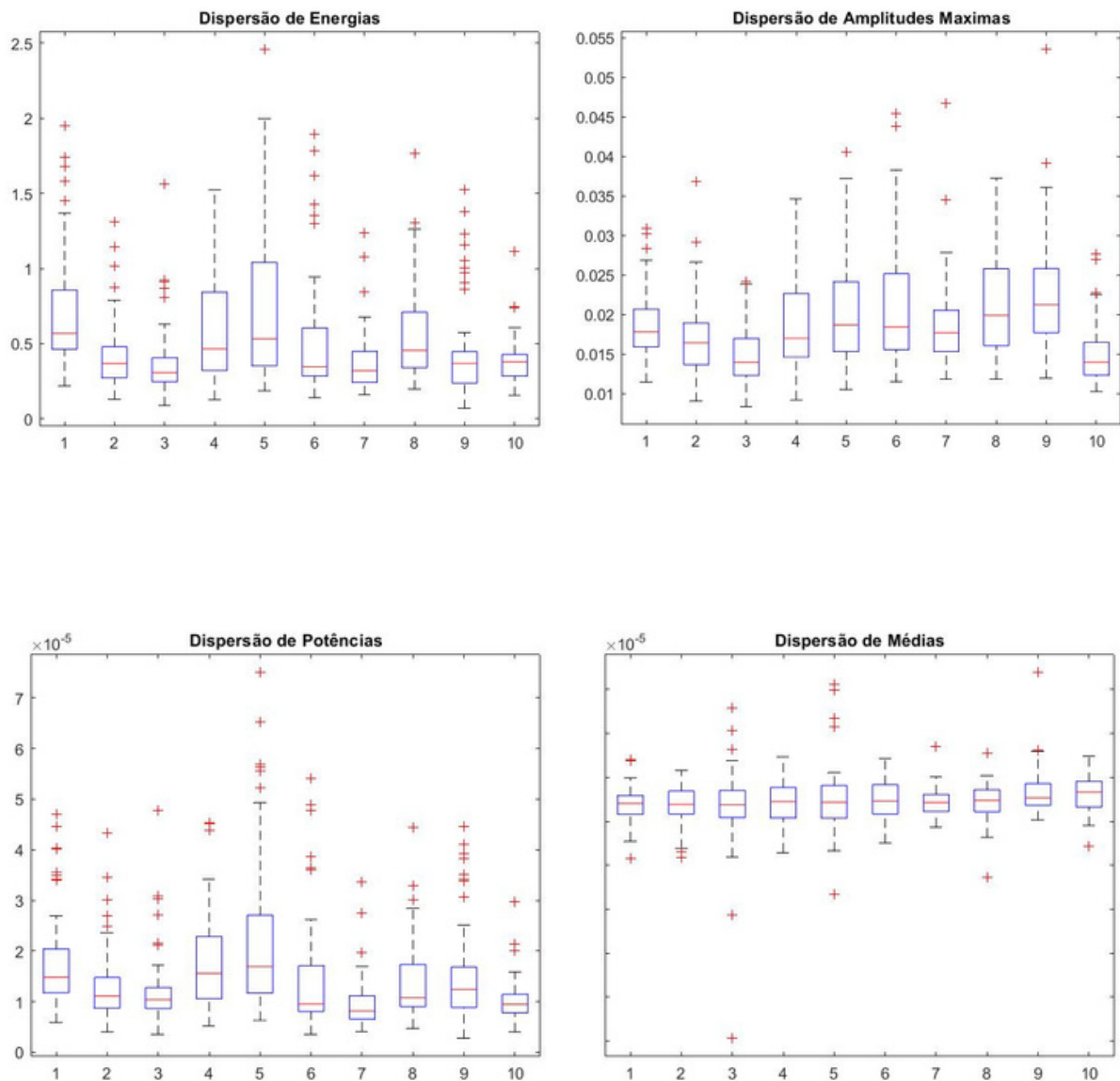


Para este exercício, o código que implementámos foi bastante simples na medida em que, percorremos cada dígito e para cada dígito, cada áudio e, como colocamos o projeto na mesma pasta dos áudios só foi necessário perceber de que forma o nome do áudio variava, para assim realizar a função `plot()`, após isso foi feita uma seleção dos gráficos que representa melhor as características de cada dígito.

Ao mesmo tempo que representávamos os gráficos, fomos calcular a energia, a amplitude máxima, a média e a potência de cada áudio, dado que, por opção, escolhemos representar estes dados por meio de uma função `boxplot()`, para ser mais perceptível as características de um determinado sinal. Também optámos por juntar todas as energias dos dígitos num gráfico, com o mesmo objetivo descrito anteriormente. Assim, encontra-se representado abaixo o resultado dos gráficos referidos acima.



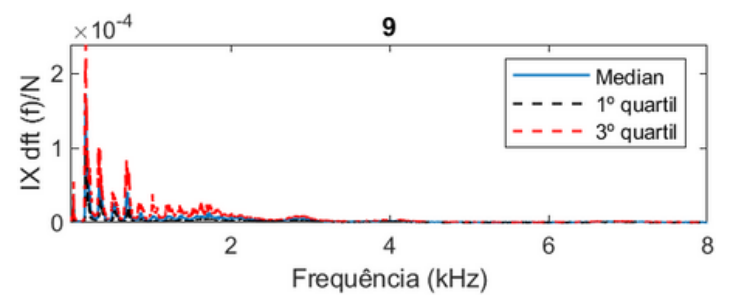
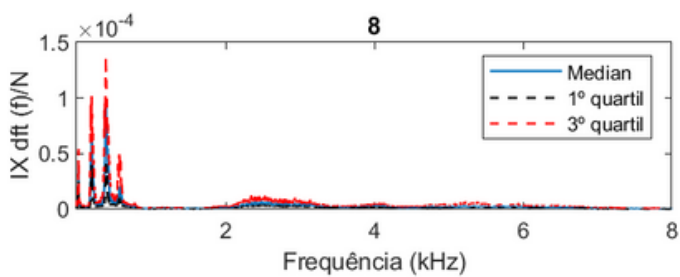
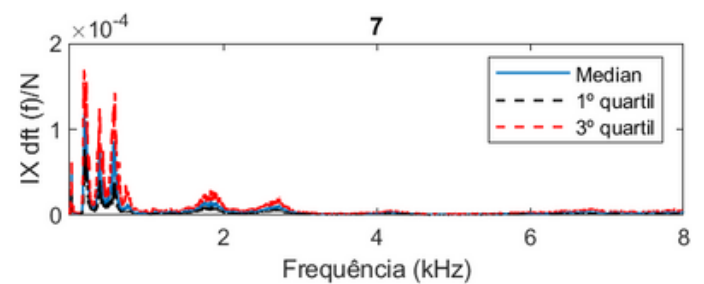
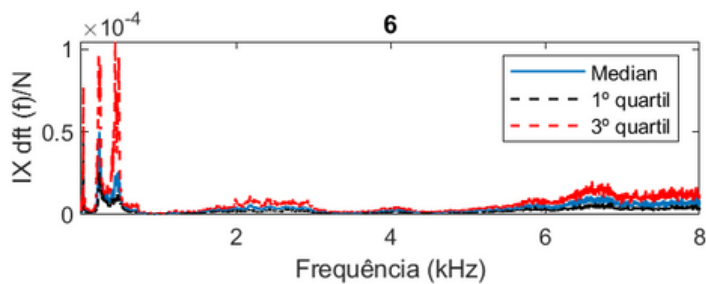
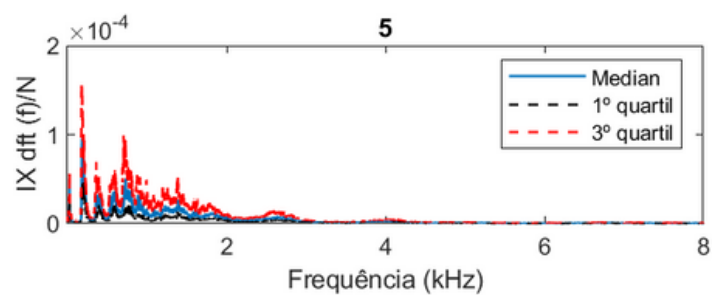
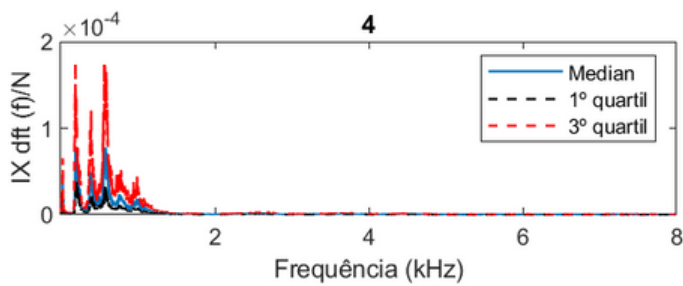
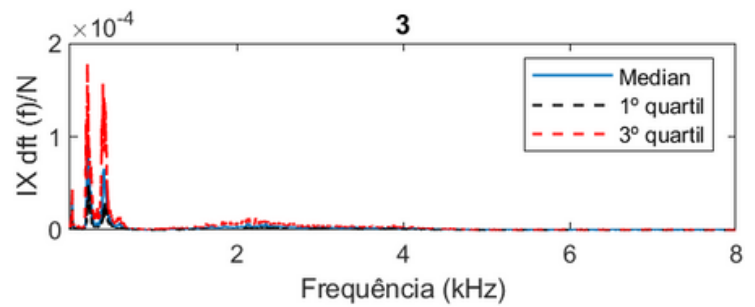
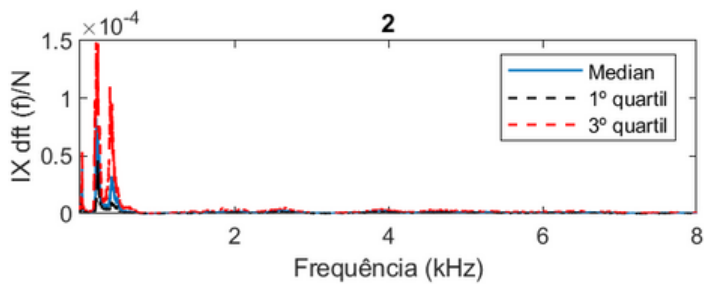
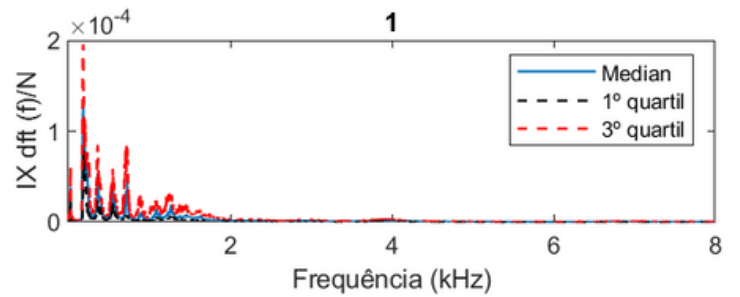
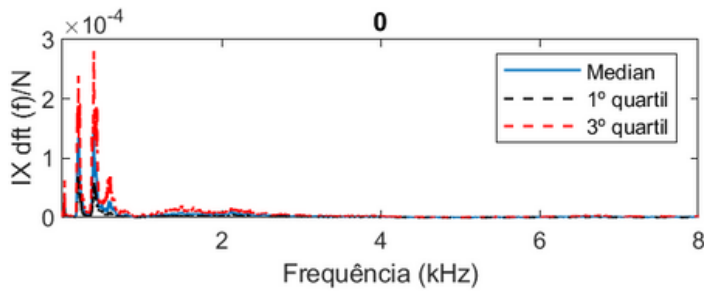
Nota: Na realização deste projeto, não implementámos as condições (*if-then-else*), visto que se torna difícil de explicar sem o recurso às imagens, pelo que esclarecemos tudo neste relatório. Nos gráficos de dispersão, o eixo do x começa no 1 e não no zero (correspondente ao áudio 0), dado que o vetor em *MATLAB* começa no índice 1. Assim o áudio correspondente é número do eixo do x menos 1.

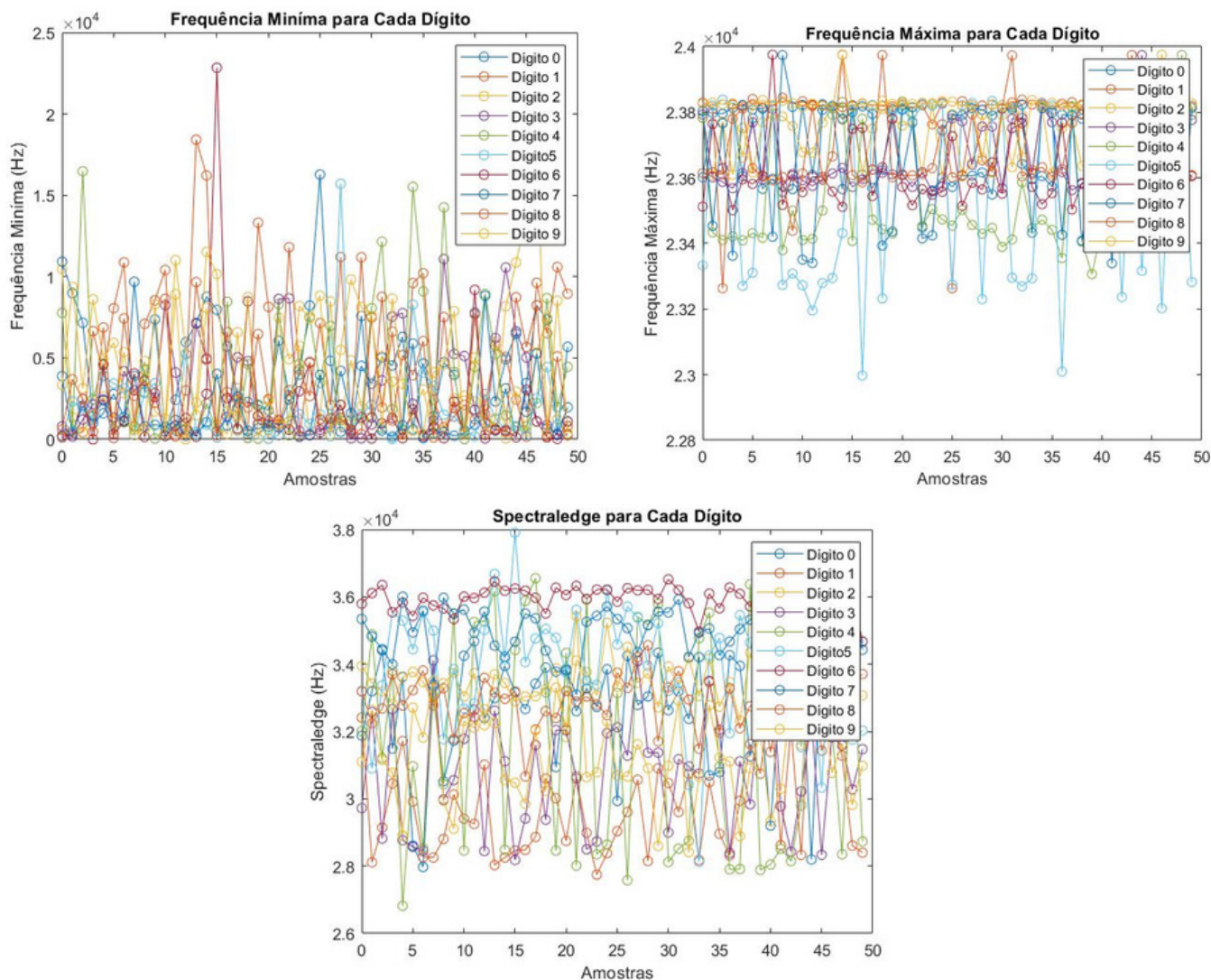


Após a observação destes gráficos, conseguimos analisar determinadas características. Esta análise pode ser feita através da distinção de pares de dígitos ou mesmo olhando para um gráfico e entender que um dado dígito tem uma determinada característica e, que mais nenhum outro dígito tem. Assim, concluímos que:

- O dígito 4, é o dígito com maiores valores de energia e de potência.
- O dígito 5, é o dígito que contém maiores valores de amplitude máxima.
- O dígito 2, é o dígito que contém menores valores de amplitude máxima.
- O dígito 6 tem uma menor dispersão de médias e uma menor dispersão de potências.
- O gráfico correspondente ao áudio 4 e o gráfico correspondente ao áudio 3, têm bastante semelhança em termos de amplitude e como ela varia consoante o tempo.
- O gráfico correspondente ao áudio 1 tem uma grande variação de amplitude, enquanto no gráfico 5 essa variação é mais constante.

▼ GRÁFICOS - EXERCÍCIO 5





Ao analisar a DFT dos sinais, podemos observar que o eixo x do gráfico representa as frequências, em kHz, e o eixo do y representa a magnitude da componente de frequência correspondente. É aqui que conseguimos retirar as frequências dominantes presentes em cada sinal e determinar a intensidade de cada uma. Dado que, como guardamos para cada dígito a magnitude de cada áudio, ao representar o gráfico da DFT de cada dígito é de realçar que é semelhante à parte do gráfico (*Mediana*) acima correspondente, assim não colocámos no relatório, mas a função de *plot* está no código. Ao calcularmos o "zero padding" estamos a garantir que todos os sinais de áudio tenham o mesmo comprimento antes de calcular a transformada de Fourier, esta aumenta a resolução da transformada de Fourier e melhora a precisão da análise espectral.

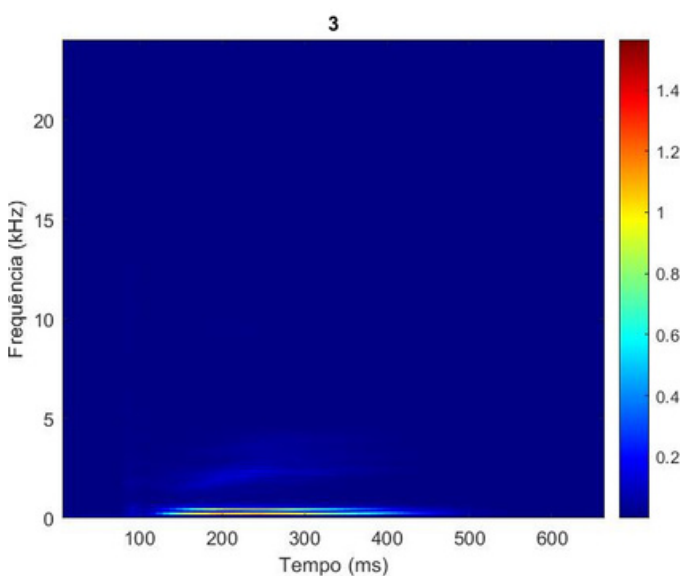
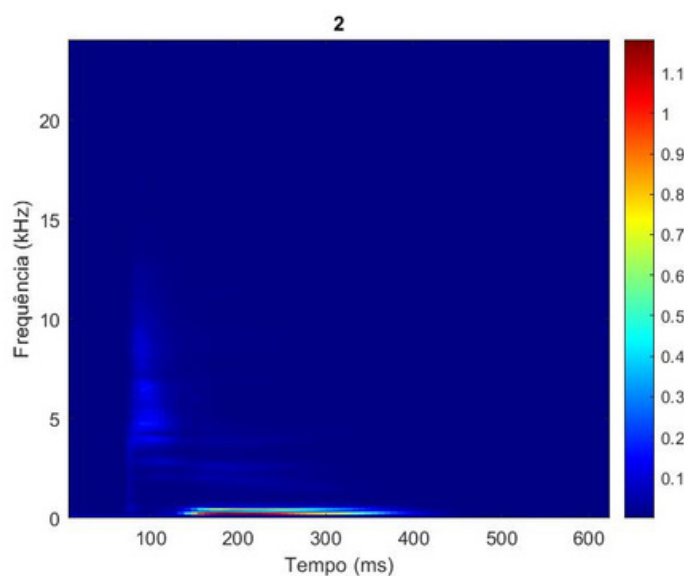
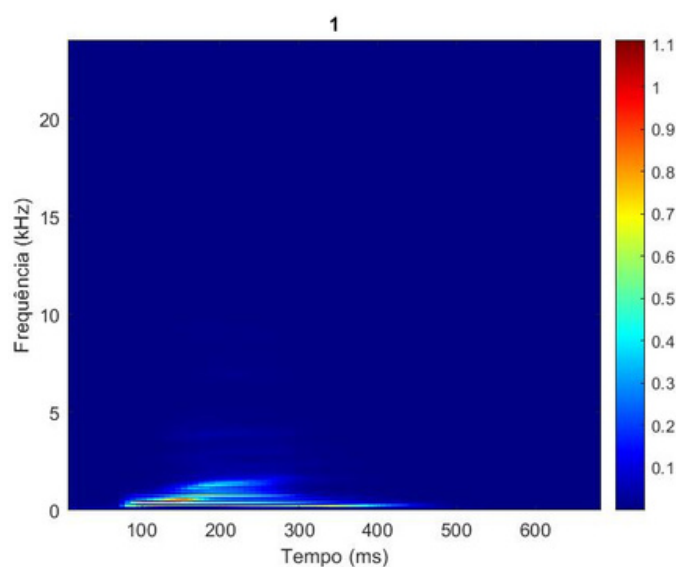
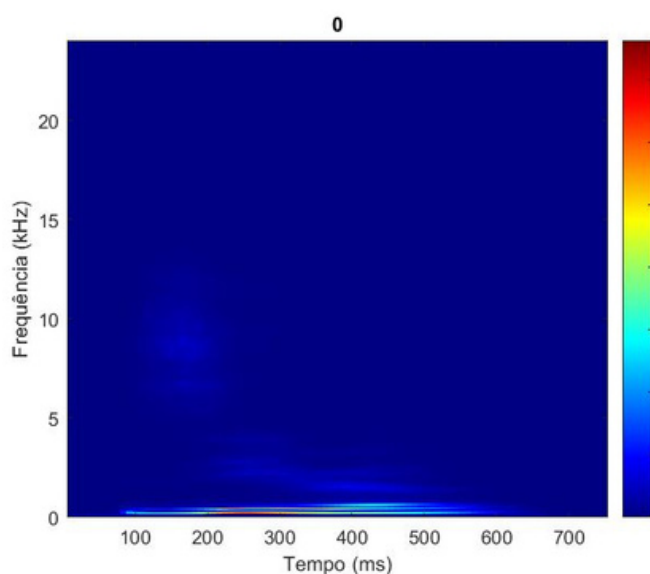
Assim, ao analisarmos os gráficos acima e, tendo em conta as características espectrais ($\max(f)$, $\min(f)$, $\text{edge}(f)$) que calculámos, podemos concluir que:

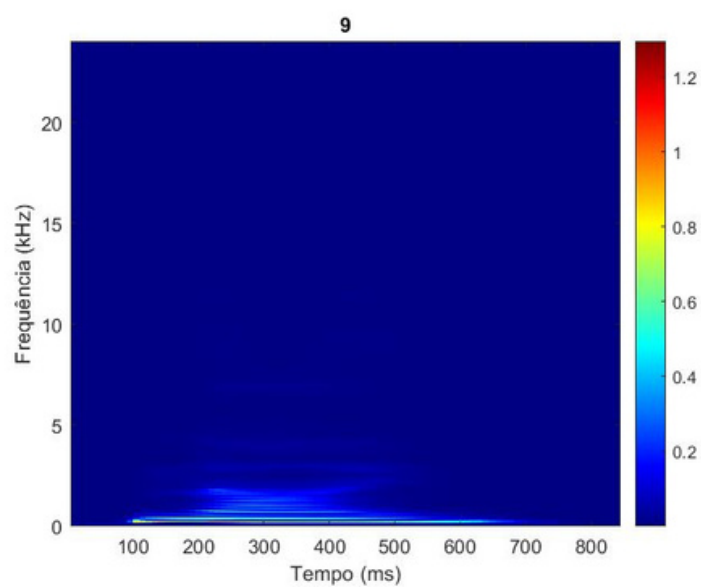
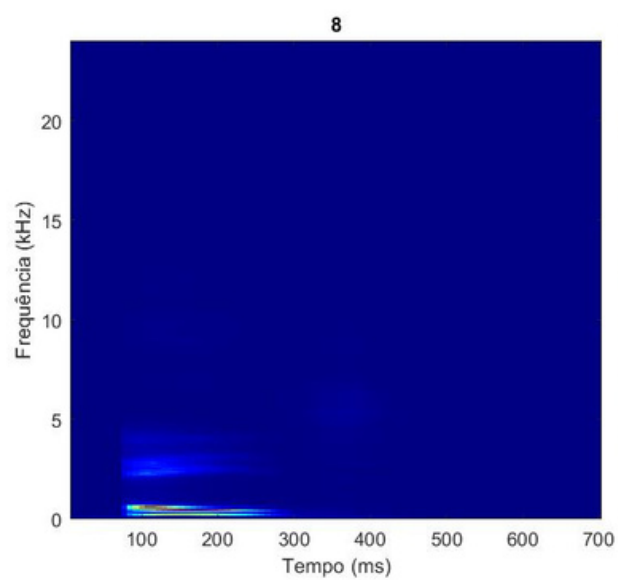
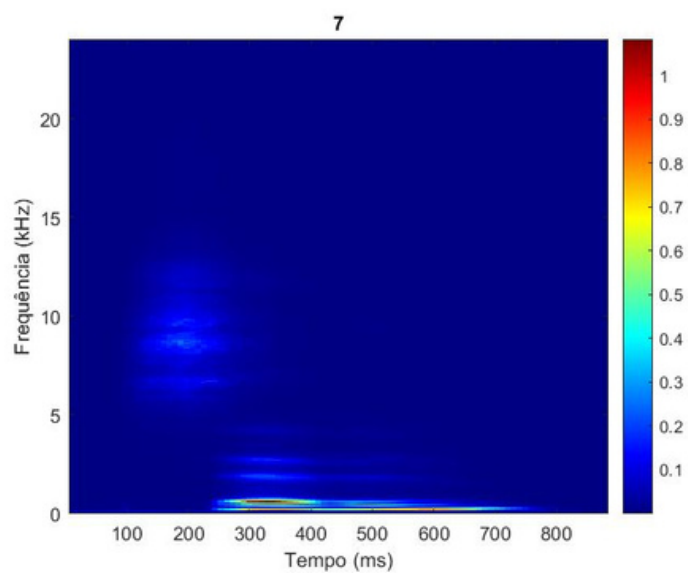
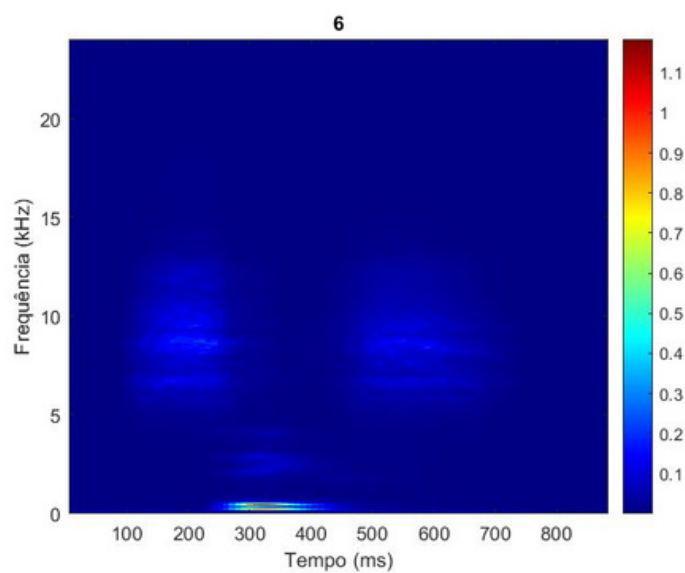
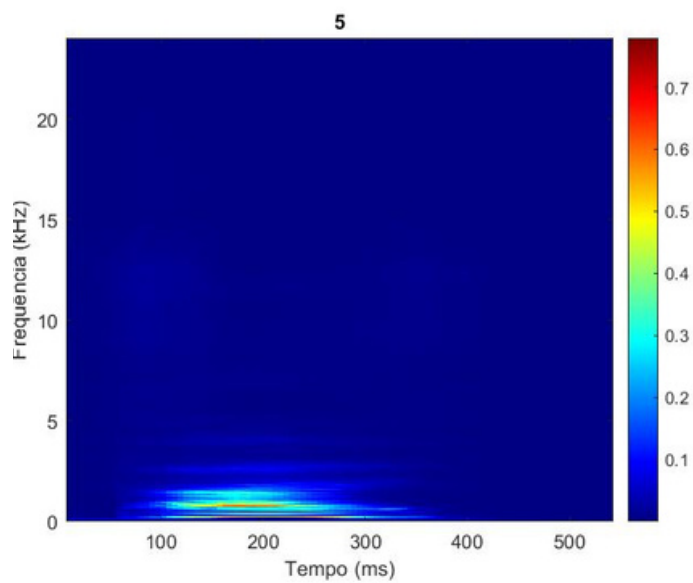
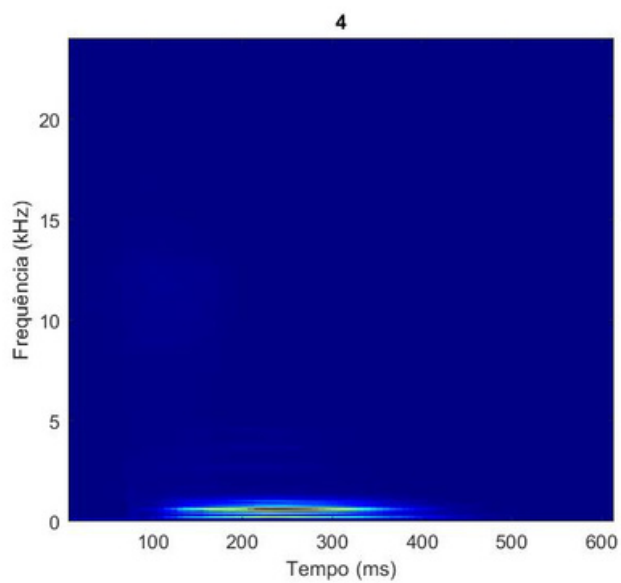
- O dígito 5 é o que apresenta menor frequência máxima e um maior pico da *spectraledge*;
- O dígito 1 é o que apresenta mais picos de frequência máxima;
- O dígito 6 tem valores semelhantes de frequência máxima entre si e os valores da *spectraledge* são bastante constantes;
- O dígito 4 é o que apresenta mais picos de frequência mínima e valores mais baixos da *spectraledge*;

Relativamente aos gráficos que apresentam os quartis e a mediana, podemos afirmar que:

- O dígito 5 oscila mais no início, enquanto o dígito 6 oscila mais no final.
- O dígito 1 e 9 apresentam semelhanças em termos de gráfico, no entanto, há sempre ligeiros desvios.

▼ GRÁFICOS - EXERCÍCIO 6





Relativamente a este ponto, existem várias abordagens, na medida em que podemos alterar o número de pontos para cálculo da FFT, e/ou pontos de sobreposição, e/ou tamanho da janela, o que pode resultar em diferentes características e resoluções espectrais nas representações de frequência dos sinais de áudio. Para isto foi útil experimentar diferentes configurações.

Nó código que executámos, após calcular todas as STFT de cada amostra, encontramos o tamanho máximo das STFT da amostra e ajustamos o tamanho, após isto é calculado a média das STFT das amostras, obtendo a STFT média do dado dígito.

Ao comparar a DFT com a STFT é de realçar algumas diferenças, como, por exemplo, a STFT analisa o sinal de áudio em pequenos segmentos de tempo usando uma janela deslizante. Ela fornece informações sobre a variação do espectro de frequência ao longo do tempo e a DFT analisa o sinal de áudio na sua totalidade, sem considerar a sua variação temporal. Ela fornece o espectro de frequência do sinal na totalidade. A STFT permite ajustar a resolução temporal por meio do tamanho da janela e do tamanho do salto (sobreposição) entre as janelas. Um tamanho de janela menor resulta numa melhor resolução temporal para capturar mudanças rápidas no sinal, enquanto um tamanho de janela maior resulta numa resolução temporal mais baixa.

Relacionando os vários espectrogramas com os gráficos resultantes do cálculo da DFT, conseguimos extrair diversas informações como:

- No dígito 6 na DFT como apresenta mais oscilações quando a frequência é maior, na STFT apresenta uma tonalidade mais clara quando a frequência aumenta. O mesmo acontece no 5, mas quando a frequência é menor.
- No geral, a maioria dos dígitos apresenta uma menor frequência no final do áudio, logo não há oscilações na DFT, o mesmo acontece na STFT, ou seja, não há alterações de tonalidade visíveis.
- O dígito 4 é relativamente constante na STFT, e na representação no DFT as oscilações são baixas, com exceção de quatro oscilações no início, o que resulta numa pequena tonalidade na STFT. Também é possível afirmar o mesmo para o dígito 3.

Em resumo, a STFT e a DFT permite-nos compreender as características temporais e espectrais de um sinal de áudio, fornecendo informações relevantes sobre a sua estrutura, padrões e mudanças ao longo do tempo. A escolha adequada dos parâmetros e a interpretação dos resultados são fundamentais para uma análise precisa e relevante do sinal de áudio. Em termos de código, podemos afirmar que para cada ponto importante, utilizámos uma função e ainda se guarda informações necessárias em 3 ficheiros ".mat".