



UNIVERSIDADE D
COIMBRA

Relatório

Análise e Transformação de Dados 2019/2020

Mariana Loreto

2018280762

Mariana Lança

2018288500

Coimbra, maio de 2020

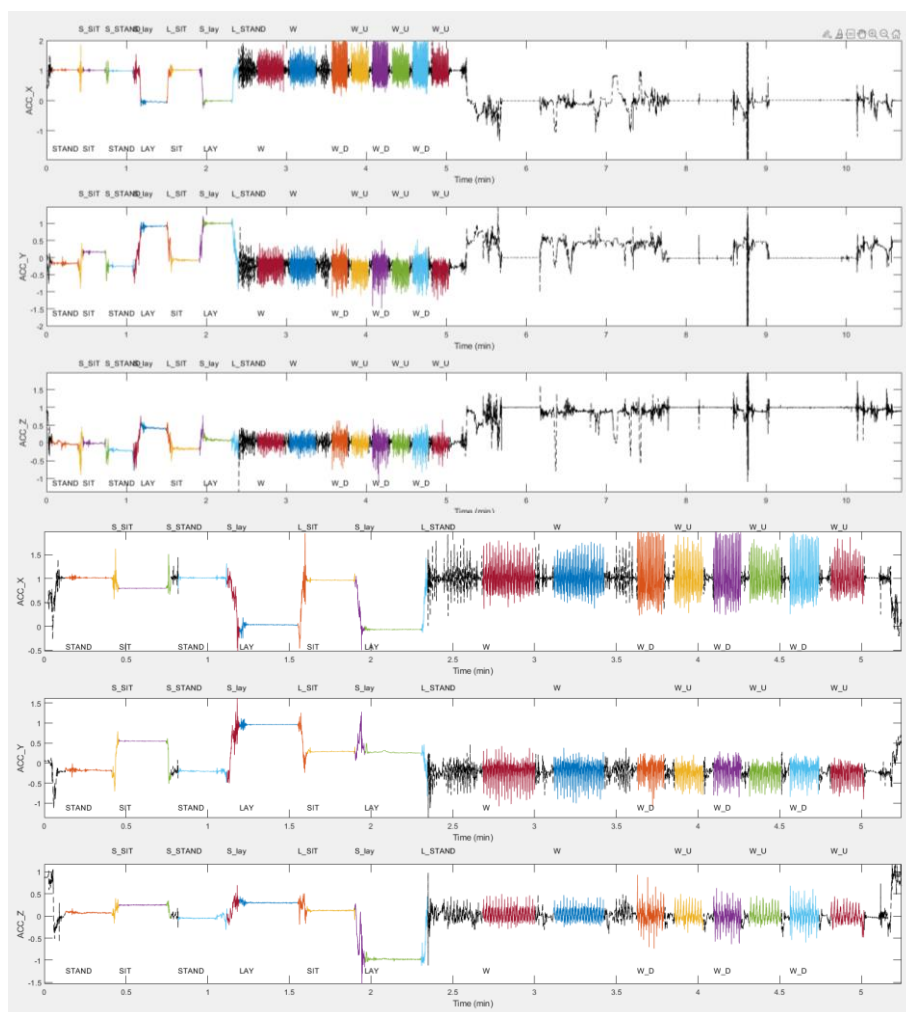
Introdução

Este projeto, realizado para a cadeira de Análise e Transformação de Dados, tem como objetivo a análise de dados provenientes da monitorização de atividade humana, bem como a identificação e classificação de cada um desses doze movimentos diferentes.

Neste trabalho, realizado na linguagem Matlab, recorreremos e aplicamos vários conceitos estudados em aula, tais como Transformadas de Fourier Discretas (DFT) e Short Time Fourier Transform (STFT) no estudo e caracterização dos vários sinais fornecidos.

Recolha dos dados a partir da Database

De modo a obter os dados pretendidos a partir dos ficheiros fornecidos (para o nosso grupo, da experiência 11 à 20) foi utilizada a função *ImportData*, própria do Matlab, e os valores foram guardados numa matriz de células, divididos em eixos e atividades, a qual denominámos de **ACC**. Esta divisão foi efetuada de modo a facilitar o estudo de cada uma das atividades individualmente. Ao mesmo tempo é realizada a representação gráfica desses mesmos sinais, identificando as várias atividades. A seguir são apresentados alguns desses resultados.



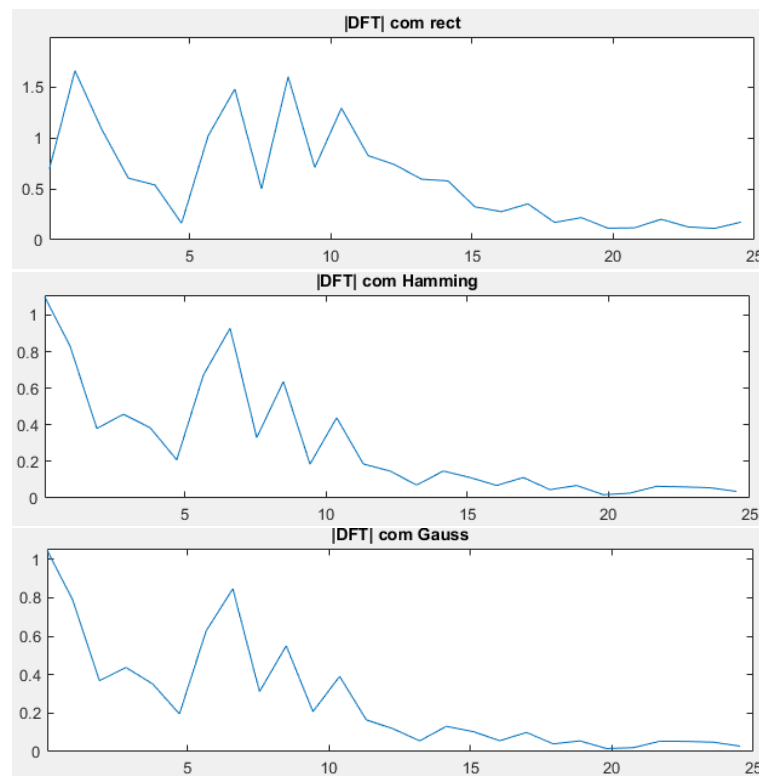
As porções do sinal a preto não estão identificadas como nenhuma das 12 atividades a estudar e não serão alvo de análise.

Tratamento do Sinal

Após a representação dos dez sinais para estudo, é-lhes aplicada uma DFT para três janelas, retangular, de *hamming* e gaussiana, passo fundamental para a escolha da janela que usada para o resto do projeto. Além disso, é também retirada a tendência, recorrendo à função *detrend*.

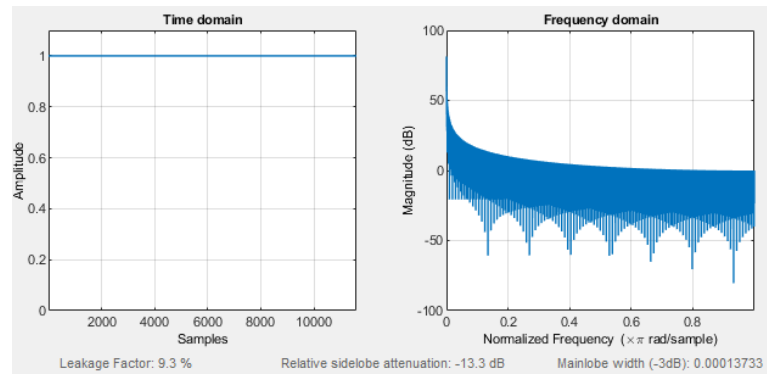
Na escolha da janela foi necessário passar por um processo de pesquisa e de testes de modo a perceber qual a melhor janela para os próximos exercícios a serem realizados neste trabalho.

Observando os gráficos seguintes, que apresentam um sinal dinâmico no eixo do Z ao qual foi aplicado as três janelas, podemos concluir que os picos diminuem à medida que a frequência diminui também. Percebemos também que o sinal da DFT com a janela de hamming e gaussiana são bastante semelhantes, sendo que há uma maior redução da magnitude dos picos (incluindo os picos médios), principalmente na gaussiana, o que é acompanhado com a diminuição do vazamento espectral (*spectral leakage*).

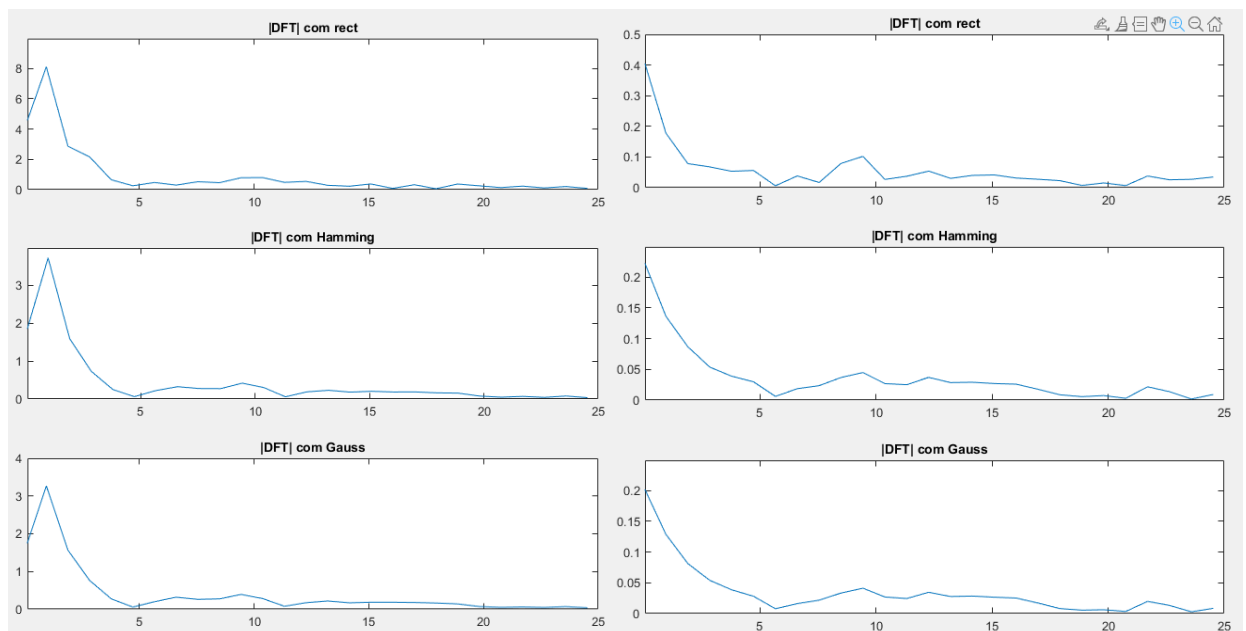
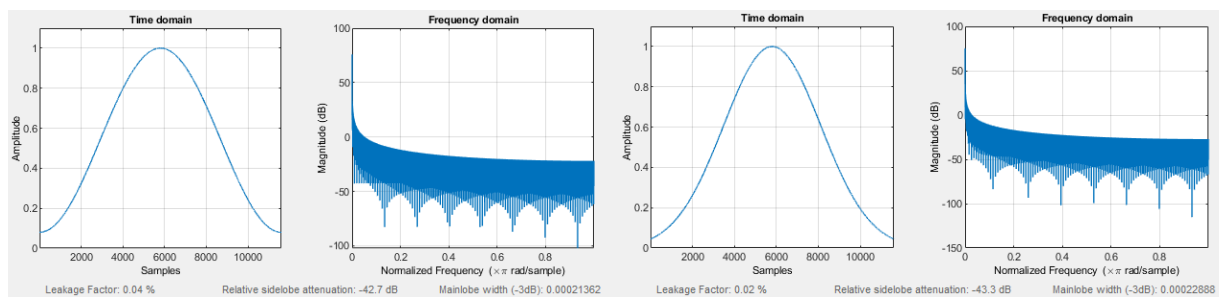


Podemos verificar, na janela retangular, um resultado bastante característico deste tipo de janela, nomeadamente, o aplanar dos topos e dos lados adjacentes ao pico, o que causa a perda de informação e pode mascarar detalhes importantes do espectro.

Como se pode ver a seguir, temos que o *leakage factor* da janela retangular é muito elevado comparando com o das restantes janelas (passando de 9.3% para 0.04%), o que faz deste tipo de janela uma má escolha e por isso não deverá ser considerada.



Já em relação às outras duas janelas, temos que a atenuação do *side lobe* é mais negativo no caso da janela gaussiana, sendo o desejado, visto que quanto mais próximo de zero este parâmetro for, menor o descontínuamento, o que é o pretendido. No entanto, procura-se também um *main lobe* estreito. Neste sentido, a janela de *hamming* fornece melhores resultados.



Com os gráficos das restantes atividades podemos tirar as mesmas conclusões anteriormente retiradas. Desta forma e fazendo o balanço com os dados adquiridos, chegou-se à conclusão de que o uso das janelas de *hamming* é uma boa opção para o cálculo dos passos, no próximo exercício.

Estudo do número de passos por minuto

O ser humano tem uma passada em torno dos 0.6Hz e os 2Hz, o que corresponde a um tempo entre os 0.5 e 1.6 segundos. Tendo em conta estes dados, foi-nos possível fazer uma estatística das atividades em relação ao número de passos por minuto.

Para este processo, foi aplicada uma DFT com janela de *Hamming* deslizante (com uma largura no intervalo 0.5-1.6s), sendo que apenas nos focámos nos picos entre os 0.6 e os 2Hz (correspondente à frequência das passadas), nomeadamente o primeiro pico. Esta janela foi aplicada com o propósito de tornar a recolha das frequências mais eficiente de modo a ultrapassar certas situações, como é o caso da variação de velocidades.

Abaixo, são apresentados os resultados das três atividades dinâmicas, nos três eixos X, Y e Z.

- Walking:

	Média	Desvio padrão
X	98.3118	25.6083
Y	101.2910	23.7087
Z	83.4161	29.0371

- Walking Upstairs:

	Média	Desvio padrão
X	101.0782	23.6522
Y	84.9057	28.8212
Z	78.8410	28.1517

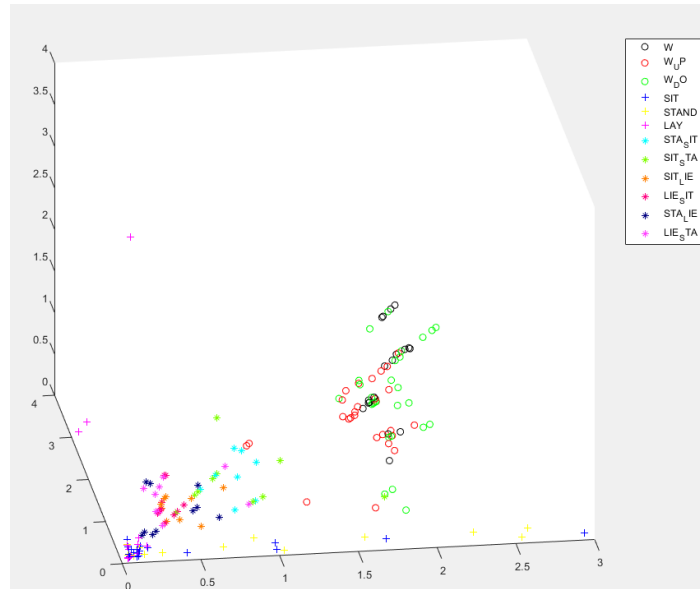
- Walking Downstairs:

	Média	Desvio padrão
X	111.2557	10.5111
Y	101.4964	23.3350
Z	85.8816	28.7857

Estudo das diferentes atividades

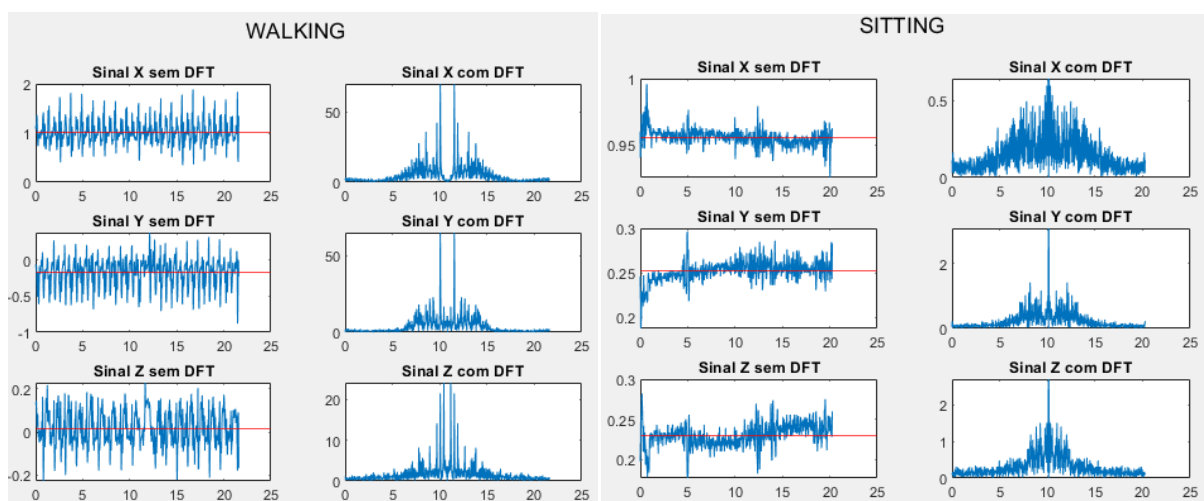
No estudo das diferentes atividades, começamos por diferenciar as dinâmicas das restantes. Para tal, gerou-se um gráfico 3D onde são apresentadas as frequências do primeiro pico dos vários sinais, para cada um dos eixos.

Como é possível visualizar pelo gráfico ao lado, há uma clara divisão entre as frequências dos sinais dinâmicos (representados por 'o'), apesar dos *outliers*, e as restantes, sendo que os pontos se encontram nas zonas de maior frequência nos vários eixos. Pelo contrário, as atividades estáticas (representadas por '+') ficam por volta dos valores mais baixos, nomeadamente no intervalo 0-0.5Hz, havendo uma distinção o menos clara nas zonas de transição.



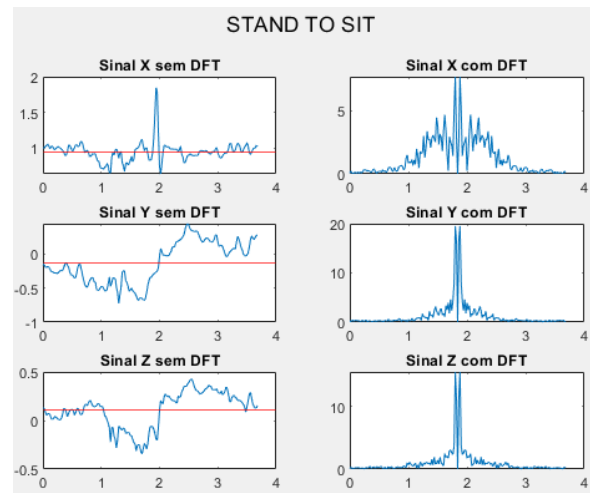
Tendo em conta que apesar de ser possível distinguir claramente as dinâmicas das estáticas não é possível tão claramente as de transição das restantes, havendo, por isso, uma baixa performance da sensibilidade e especificidade. Desse modo, é necessário recorrer a uma análise mais detalhada.

Para essa análise, começamos por comparar os valores da DFT entre as diferentes atividades:

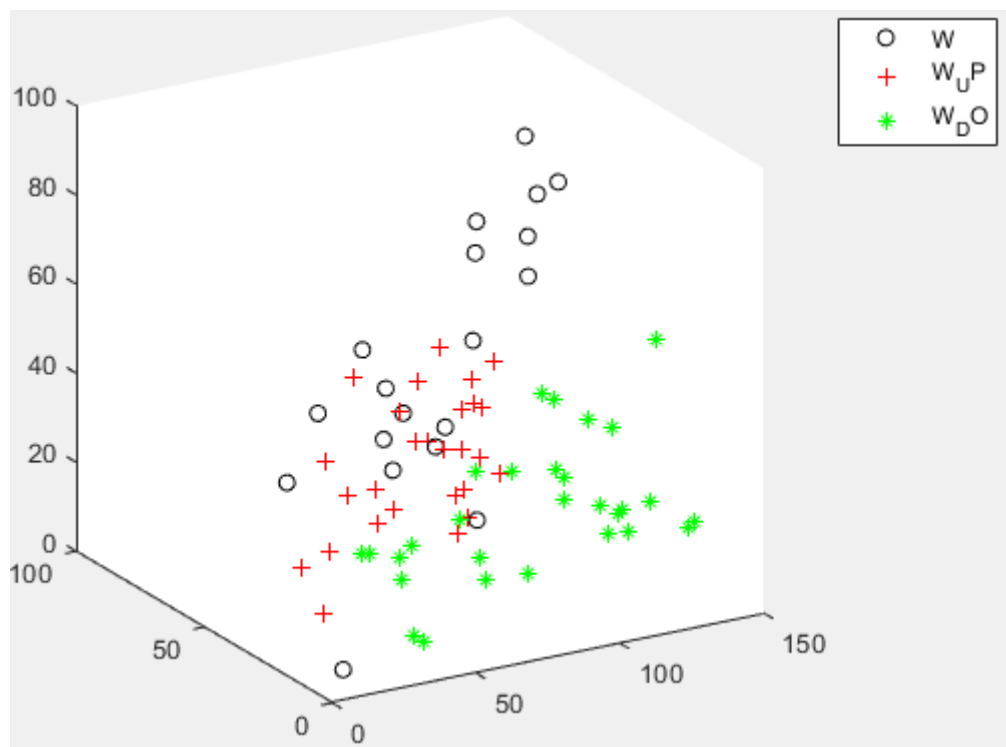


Observando os vários gráficos, podemos concluir que existe um intervalo dos coeficientes máximos para cada uma das atividades, em que as de transição se encontram no intervalo [10,20], as dinâmicas num intervalo de valores superiores, e as estáticas, inferiores.

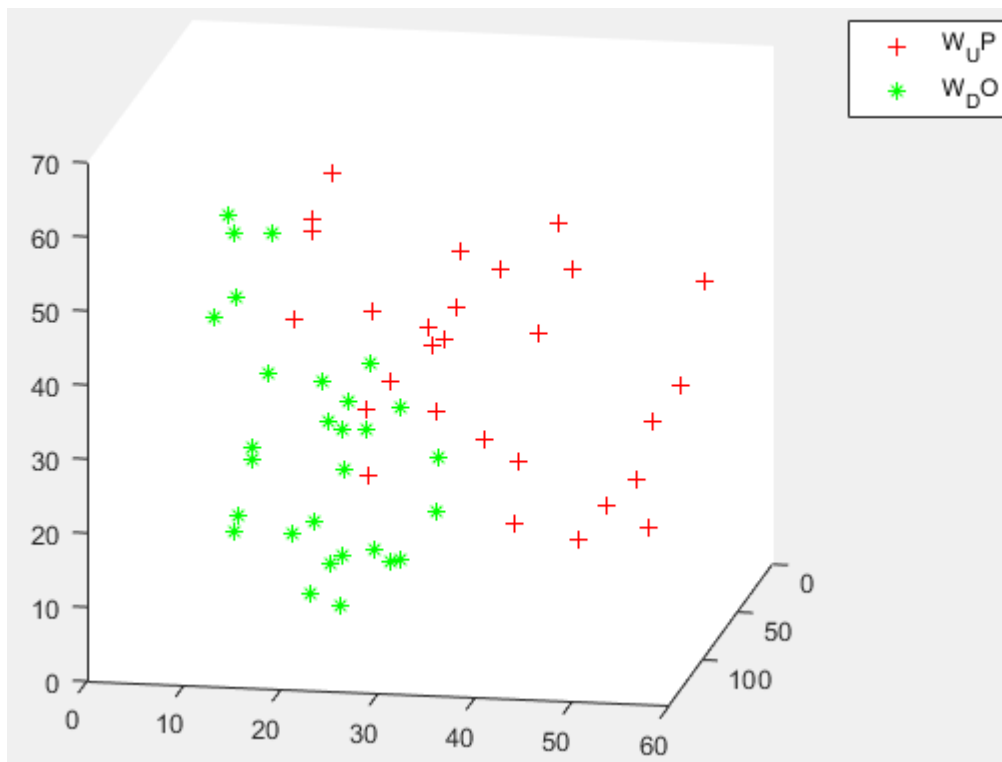
Estes resultados devem-se as transições abruptas no sinal original, que se reflete posteriormente no sinal com DFT.



De modo a diferenciar as várias atividades dinâmicas recorreu-se à apresentação da magnitude do primeiro ponto dos dados referentes a cada atividade. A seguir apresenta-se o resultado obtido:



Analisando o gráfico, é possível concluir que a componente da magnitude no eixo Z da atividade *Walking* apresenta valores superiores aos das restantes, apesar da baixa especificidade, possivelmente por uma movimentação mais evidente neste eixo.



Do mesmo modo, ao se analisar a representação dos pontos das magnitudes das atividades *walking upstairs* e *walking downstairs*, podemos notar uma congregação dos pontos desta primeira atividade em valores mais elevados de Y, enquanto que os pontos da segunda atividade se encontram concentrados na zona de valores de Y menores. Neste gráfico é bastante mais notória a divisão entre atividades, apesar de alguns *outliers*.

STFT (Short Time Fourier Transform)

Nesta última etapa, criamos a função `calcSTFT` (`name_file`, `fs`, `Tframe`, `Toverlap`) para obter as distribuições tempo-frequência para o sinal do acelerómetro no eixo Z. Em que o `name_file` corresponde ao nome do ficheiro que escolhemos e `fs` é a frequência de amostragem. O `Tframe` corresponde à largura da janela e `Toverlap` corresponde à sobreposição, ambos já foram discutidos em cima.

No cálculo da STFT usamos uma janela de Hamming deslizante. Escolhemos esta janela com base nas conclusões já discutidas anteriormente.

Analisando o gráfico abaixo, podemos observar que conseguimos separar as atividades ao longo do tempo, bem como as transições entre estas de forma clara.

No intervalo $[0, 40[$ (valores aproximados) identificamos uma atividade estática, e logo de seguida no intervalo $[40, 80[$ verificamos uma atividade de transição. Até aproximadamente ao instante 250s, estas vão-se intercalando. A partir deste instante contundo, podemos observar a transição para as atividades dinâmicas.

