



FCTUC FACULDADE DE CIÊNCIAS
E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

Monitorização da atividade de indivíduos

Análise e Transformação de Dados
Licenciatura em Engenharia Informática
2020/2021

Duarte Emanuel Ramos Meneses – 2019216949 - duartemeneses@student.dei.uc.pt

Inês Martins Marçal – 2019215917 - inesmarcal@student.dei.uc.pt

Patrícia Beatriz Silva Costa – 2019213995 - patriciacosta@student.dei.uc.pt

Índice

Introdução.....	3
Início do projeto	4
Pontos 1 e 2.....	4
Tratamento do sinal	5
Selecionar uma atividade dinâmica e usar diferentes tipos de janela para segmentar o sinal. Calcular a DFT desse fragmento e comparar os resultados obtidos para descobrir as diferenças entre as janelas.....	5
Cálculo da DFT do fragmento do sinal associado a cada atividade para todos os utilizadores	7
Identificar as características mais relevantes, nomeadamente espectrais, para cada atividade, analisando os resultados obtidos por atividade e por utilizador	8
Identificar o número de passos por minuto, para cada atividade dinâmica, de cada utilizador e apresentar a respetiva média e desvio padrão	10
Identificar as características, por exemplo espectrais, que permitam diferenciar as atividades estáticas, de transição e dinâmicas.....	11
Identificar as características, por exemplo espectrais, que permitam diferenciar cada atividade.....	12
Utilizar a Short-Time Fourier Transform (STFT) para obter computacionalmente as distribuições tempo-frequência para o sinal do acelerómetro Z.....	16
Conclusão	19
Bibliografia / Webgrafia	20

Introdução

A monitorização de movimentos humanos é uma atividade em expansão atualmente visto ser uma ótima forma de controlar o Homem. São já várias as áreas que se renderam aos benefícios desta ciência como o desporto: por exemplo, qualquer equipa de futebol profissional com algumas posses tem neste momento pessoas a analisar em tempo real os dados biométricos enviados por sensores colocados nos jogadores para assim conseguirem rentabilizar a performance desportiva. Também a saúde se tem apoiado nisso para controlar o estado de saúde de pacientes mais isolados. Tal como dissemos consideramos que esta é uma área em expansão e no futuro estamos certos que a monitorização de movimentos humanos estará presente em grande parte do nosso quotidiano.

Este projeto visa a análise de dados, nos domínios do tempo e da frequência, usando sinais recolhidos de acelerómetros de smartphones com o objetivo de identificar a atividade realizada. Essa atividade pode ser uma de doze: *WALKING*, *WALKING_UPSTAIRS*, *WALKING_DOWNSTAIRS* (dinâmicas), *SITTING*, *STANDING*, *LAYING* (estáticas), *STAND_TO_SIT*, *SIT_TO_STAND*, *SIT_TO_LIE*, *LIE_TO_SIT*, *STAND_TO_LIE*, *LIE_TO_STAND* (de transição).

Neste relatório vamos explicar como implementamos computacionalmente em *Matlab* as nossas soluções para analisar o que nos era pedido e vamos ainda fazer uma análise dos resultados que fomos obtendo.

Início do projeto

Pontos 1 e 2

Começamos este projeto por fazer o *download* dos sinais que tínhamos de analisar (relativos à nossa turma PL-1) disponíveis em <https://archive.ics.uci.edu/ml/machine-learningdatabases/00341/HAPT%20Data%20Set.zip>. Vimos ainda a descrição dos dados ao aceder a: <https://archive.ics.uci.edu/ml/datasets/SmartphoneBased+Recognition+of+Human+Activities+and+Postural+Transitions>. Nestes sites retiramos informações importantíssimas para o decorrer do projeto como, por exemplo, o valor da frequência de amostragem ser de 50 Hz.

Após estudar as características das experiências realizadas, tendo já os dados descarregados e as informações necessárias de cada ficheiro, partimos para a elaboração de código em *Matlab*. Começamos logicamente por importar os dados necessários (ponto 1) recorrendo ao método *importdata* já existente no *Matlab*. Posteriormente, representamos os sinais identificando a atividade a que cada fragmento correspondia utilizando o método por nós desenvolvido *exercício2* (ponto 2). Esse método, ao receber a frequência de amostragem, o conteúdo dos ficheiros *activity_labels.txt* e *labels.txt*, o sinal a representar e um *array* com [nºexperiência nºutilizador] por linha, é responsável por representar graficamente o sinal segmentado escrevendo a que atividade se refere cada porção de espetro.

A figura seguinte é o resultado da representação da experiência 1:

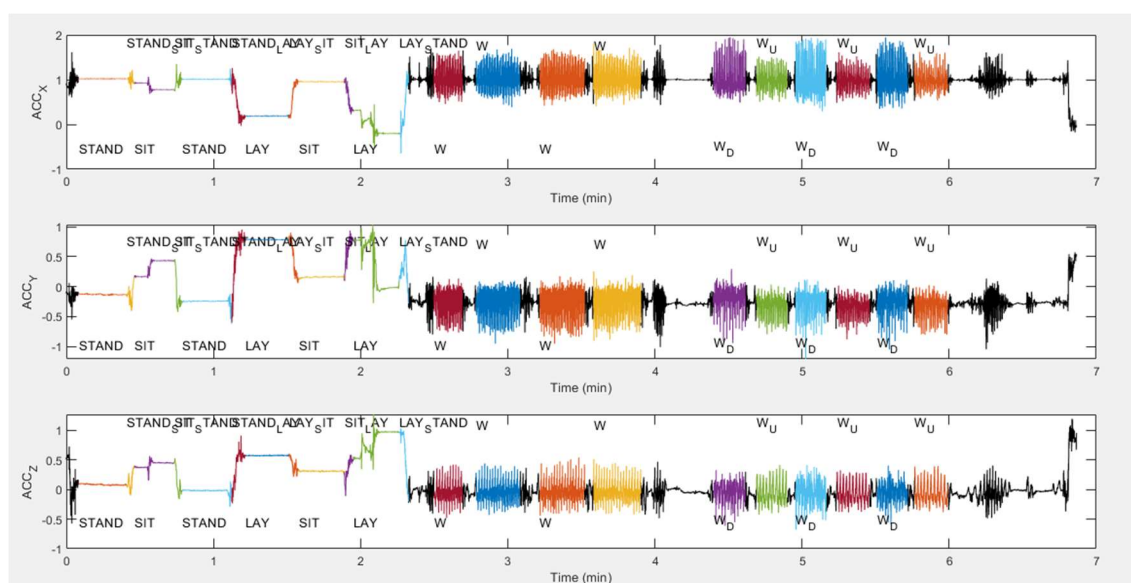


Figura 1 – Representação gráfica de sinais do acelerómetro da experiência 1 com identificação das atividades

Para esta representação tivemos de ir fazendo o *plot* à medida que cada atividade ocorria (início e fim das atividades definidos pelas colunas 4 e 5, respetivamente, do ficheiro *labels.txt*) para cada eixo (coluna 1 dos ficheiros que contêm os sinais dos acelerómetros corresponde ao eixo dos X, 2 dos Y e 3 dos Z).

Tratamento do sinal

Selecionar uma atividade dinâmica e usar diferentes tipos de janela para segmentar o sinal. Calcular a DFT desse fragmento e comparar os resultados obtidos para descobrir as diferenças entre as janelas

Para este ponto, utilizamos o método por nós definido *dft*. Este é um método que vai percorrer o sinal pretendido pelo utilizador e segmentá-lo por atividades. Se a atividade for a que se quer analisar, o programa calcula a DFT para esse segmento (atividade), aplica-lhe a janela pretendida também pelo utilizador e mostra graficamente o resultado. Decidimos escolher a atividade dinâmica “WALKING”.

Num estudo prévio sobre janelas percebemos que a aplicação destas é uma maneira muito comum de identificar temporalmente a ocorrência de certas frequências, ou seja, é uma estratégia para conseguir simultaneamente uma análise em termos de frequência e de tempo. Porém, aplicando DFT a um sinal apenas se obtêm as frequências e não as suas localizações temporais. Para contornar este problema surgiu a hipótese de aplicar DFT em janelas, isto é, em pequenas porções do sinal. Deste modo, sendo que em cada janela se identificam as frequências presentes, já dá para as reconstruir temporalmente, visto a janela “deslizar” sob o sinal.

No entanto, uma questão surge: Existem vários tipos de janela; qual utilizar? Obviamente que cada tipo de janela obtém diferentes resultados quando aplicado ao sinal. Para este projeto, começámos por testar a janela Retangular. Neste tipo de janela, percebemos rapidamente que existe um corte abrupto entre as frequências, isto é, no início e final das janelas podem aparecer valores distintos uma vez que o sinal, ao repetir-se periodicamente, faz com que apareçam as tais transições abruptas que vão resultar em frequências que não estavam no sinal original (o que leva à sua adulteração).

Exemplo desta situação é a representação seguinte, referente à primeira ocorrência da atividade dinâmica por nós escolhida (WALKING) na experiência número 1:

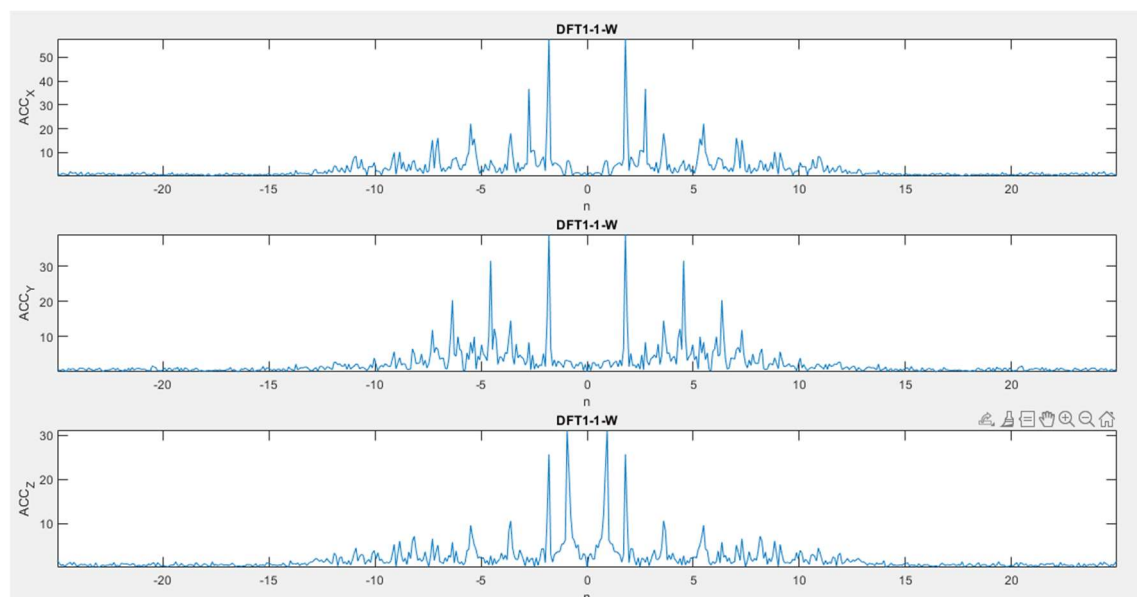


Figura 2 - Representação gráfica da DFT da 1ª ocorrência de WALKING na experiência 1 recorrendo a janela Retangular

Como se pode observar, existem cortes abruptos entre as frequências.

Para evitar isto, partimos para janelas do tipo não retangular: de *Hamming*, de *Blackman* e a Gaussiana. Este tipo de janelas faz com que haja atenuação dos extremos e consequentemente evita os cortes abruptos. A desvantagem deste tipo de janelas é que estas atenuam o sinal em que atuam. Para contornar isto, decidiu-se não percorrer as janelas de forma seguida (isto é, não começar a janela seguinte logo no fim da anterior) mas sim começar a janela seguinte um pouco antes do fim da anterior (causando sobreposição). Quando a DFT e as funções de janela são aplicadas juntamente ocorre *spectral leakage*, isto é, uma propagação localizada de frequências.

Os resultados que obtivemos para a mesma situação da representação da figura 2 foram os seguintes:

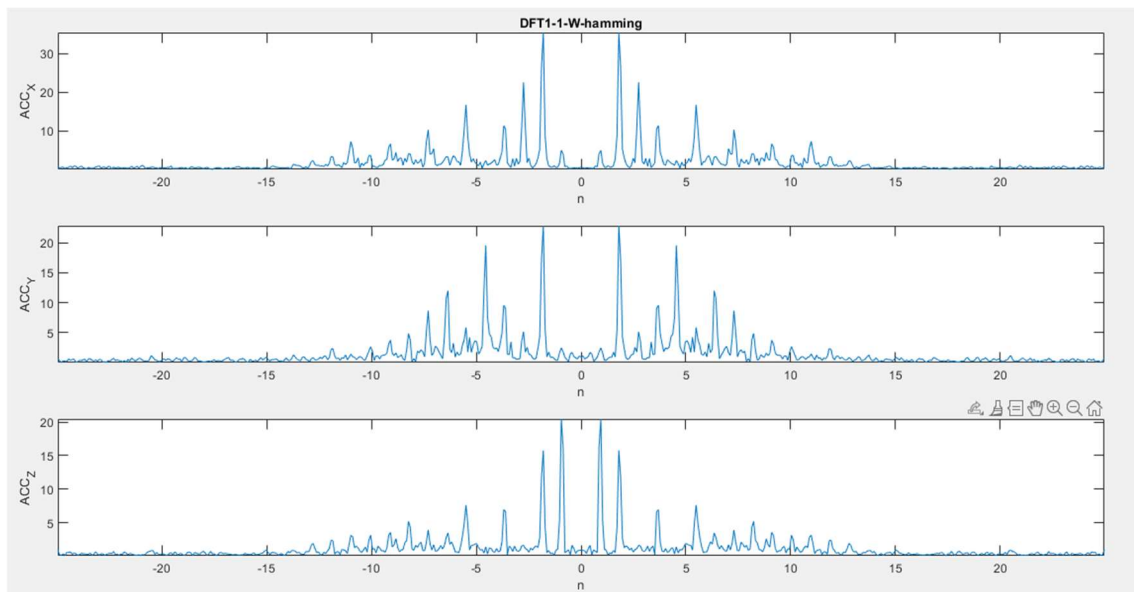


Figura 3 – Representação gráfica da DFT da 1ª ocorrência de WALKING na experiência 1 recorrendo a janela de Hamming

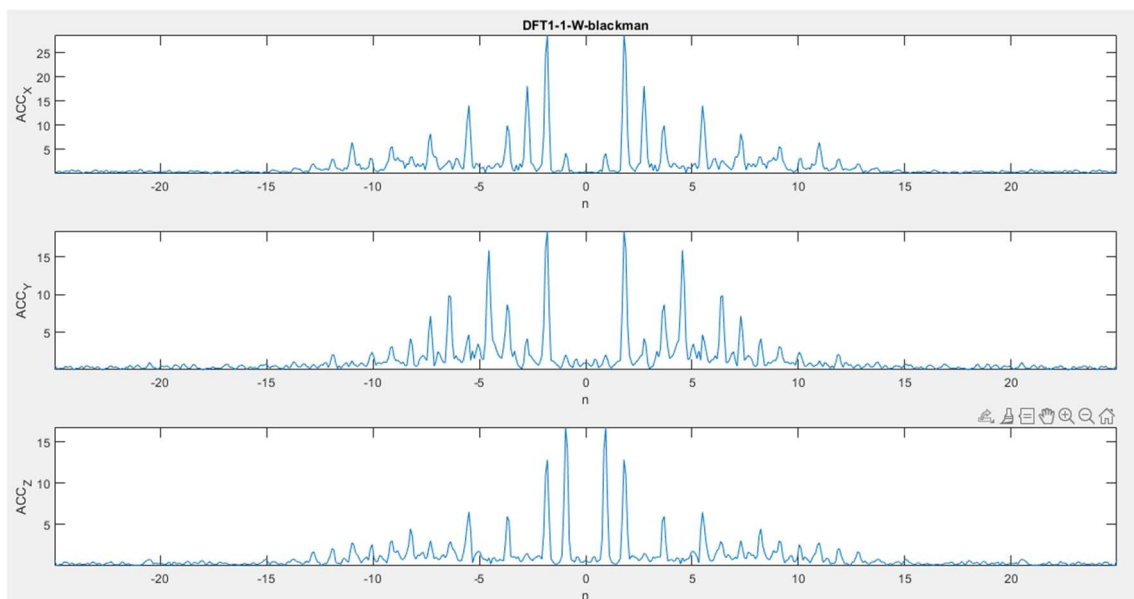


Figura 4 - Representação gráfica da DFT da 1ª ocorrência de WALKING na experiência 1 recorrendo a janela de Blackman

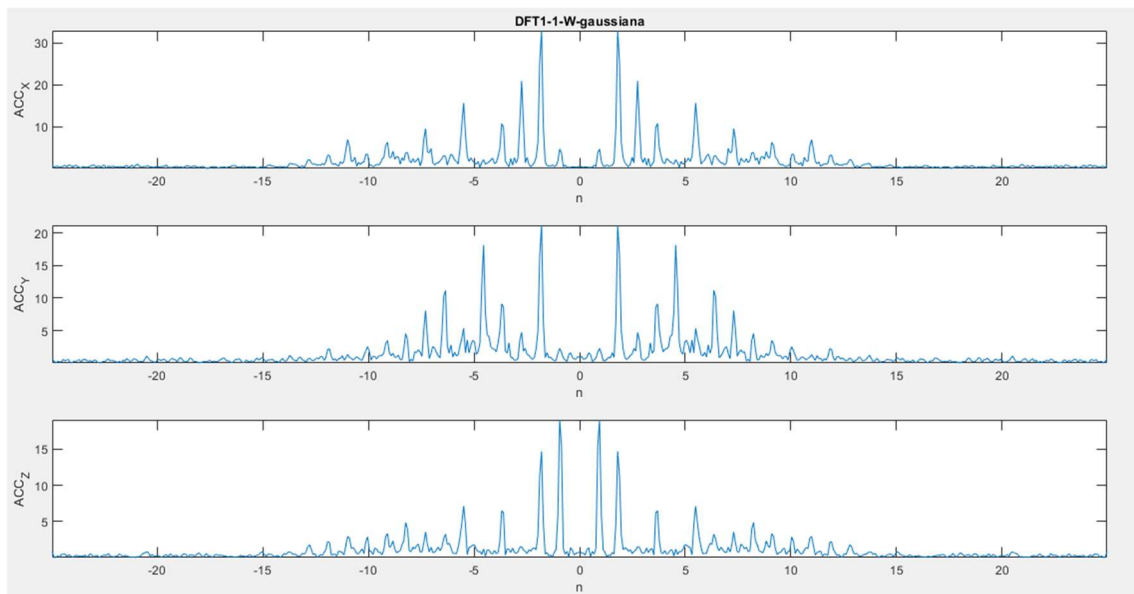


Figura 5 - Representação gráfica da DFT da 1ª ocorrência de WALKING na experiência 1 recorrendo a janela Gaussiana

Analisando as representações acima, percebemos que os três tipos de janela apresentam resultados muito parecidos entre si, sobretudo a janela de *Hamming* e a Gaussiana.

A janela de *Blackman* é a que apresenta maior atenuação das frequências, porém isto traz desvantagens. Atenuar de forma mais intensa acaba por fazer com que os *side-lobes* não anulem os a eles contíguos, o que é mau pois os *side-lobes* próximos devem ser semelhantes. Deste modo, como a Gaussiana e a de *Hamming* apresentam resultados semelhantes e sendo elas menos intensivas na atenuação que a de *Blackman*, optamos por neste projeto fazer testes utilizando a janela de *Hamming*. Em comparação com a Gaussiana são muito semelhantes, portanto qualquer uma que fosse escolhida apresentaria bons resultados.

Cálculo da DFT do fragmento do sinal associado a cada atividade para todos os utilizadores

Nesta fase do projeto, calculamos a DFT (com janela retangular) para cada atividade de todos os utilizadores. Desenvolvemos o método *dft* que recebe a frequência de amostragem, o sinal a analisar, o conteúdo do ficheiro *labels.txt*, o tipo de janela a utilizar (retangular, gaussiana, *blackman* ou *hamming*), o número de atividade a analisar, o conteúdo do ficheiro das atividades (*activity_labels*) e um *array* composto pelo número de utilizador. Com estes dados irá, para os 3 eixos dos acelerómetros, aplicar a janela escolhida pelo utilizador a cada ocorrência da atividade que se pretende analisar. Antes de utilizar o sinal para analisar, tiramos-lhe a tendência através do método já presente nas bibliotecas do *Matlab detrend*.

Os resultados seguintes são os correspondentes à atividade *WALKING_UPSTAIRS* (primeira ocorrência) do utilizador número 1 (experiência número 1):

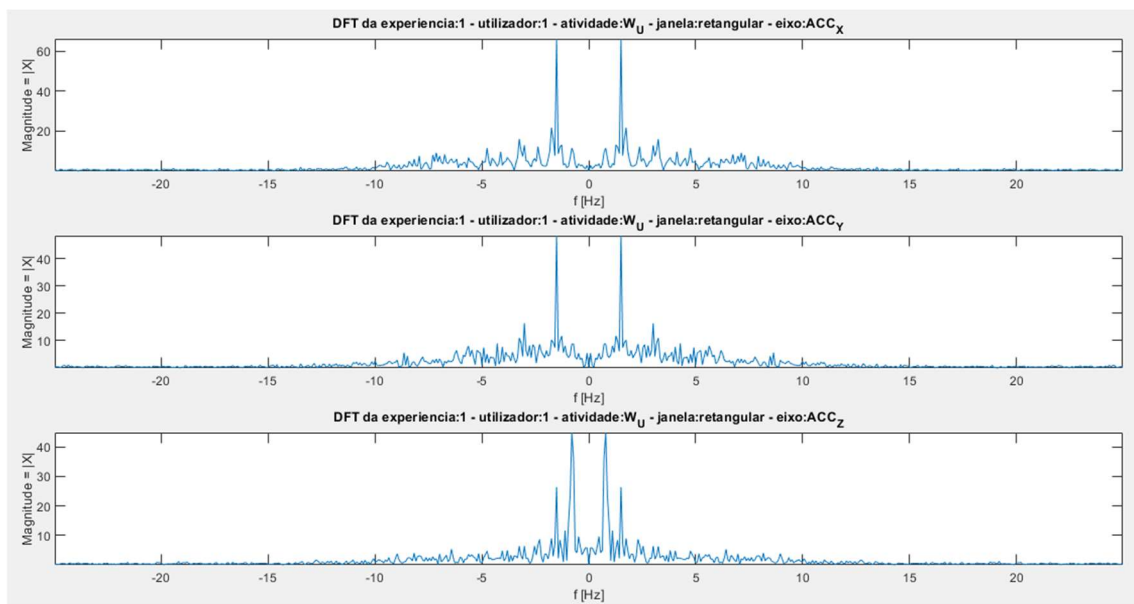


Figura 6 - Representação gráfica da DFT da 1ª ocorrência de WALKING_UPSTAIRS na experiência 1 recorrendo a janela Retangular

Identificar as características mais relevantes, nomeadamente espectrais, para cada atividade, analisando os resultados obtidos por atividade e por utilizador

É lógico que todas as atividades são diferentes umas das outras em termos de frequência, magnitude, ..., e isso leva a que se possam distinguir. Inclusivamente, podem-se identificar através de certas características espectrais. Para tal, decidimos que seria uma boa prática aplicar DFT ao sinal para calcular o espetro de frequências do sinal. No entanto, a DFT por si só só devolve magnitudes. Como esta é uma componente que depende do número de pontos da janela em questão, não é um bom valor para comparar cada atividade.

Deste modo, optamos por comparar as atividades a partir da componente amplitude e frequência dos sinais. Tendo o valor máximo da DFT, temos o valor máximo da magnitude do sinal. Definindo um limiar para que apenas as componentes relevantes do sinal sejam consideradas, conseguimos obter valores interessantes para analisar e conseguir distinguir cada atividade. Após vários testes, o limiar sendo 0.8 obteve os melhores resultados, isto é, apenas vamos considerar os valores que sejam superiores a 80% do máximo da componente magnitude. A partir daqui, para encontrar as componentes frequência bastou localizar os índices na DFT cuja magnitude está de acordo com a condição anterior, transpor esses índices para o vetor de frequências e encontrar as positivas. Já para encontrar a magnitude, utilizamos o mesmo limiar. Aqui, precisamos de calcular o c_m da Transformada de Fourier. Isso obtém-se dividindo o valor da magnitude pelo número de pontos da amostra e, caso o m seja superior a 0, multiplicamos o módulo da componente por 2.

Guardando estes valores em *arrays* próprios (um para cada eixo e atividade), conseguimos tanto fazer a média das componentes magnitude mais relevantes bem como a das frequências. Tendo isto basta representar graficamente cada valor de média (frequências num gráfico e amplitudes noutra) para poder analisar e distinguir as atividades. Este é basicamente o funcionamento do método que implementamos para este ponto (*diferenciaEstaticas*) que

recebe a frequência de amostragem, um *array* com os dados correspondentes aos sinais que queremos analisar e o conteúdo do ficheiro *labels.txt*.

Calculando a média das componentes frequência e amplitude mais relevantes para cada atividade do utilizador n.º 1, obtivemos os seguintes resultados:

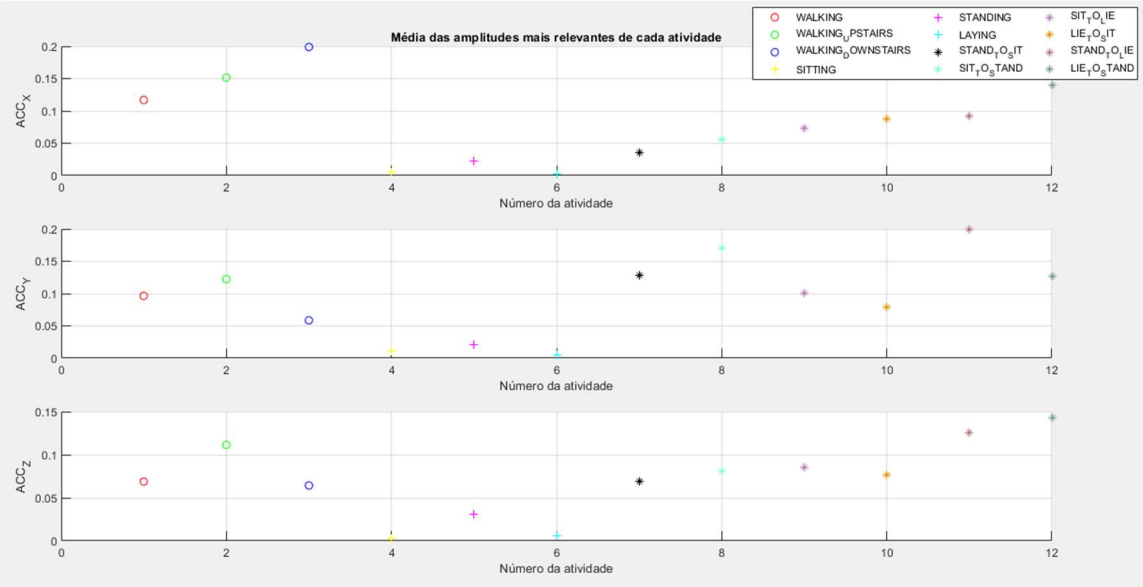


Figura 7 - Representação gráfica da média das componentes amplitude mais relevantes de cada atividade realizada pelo utilizador 1

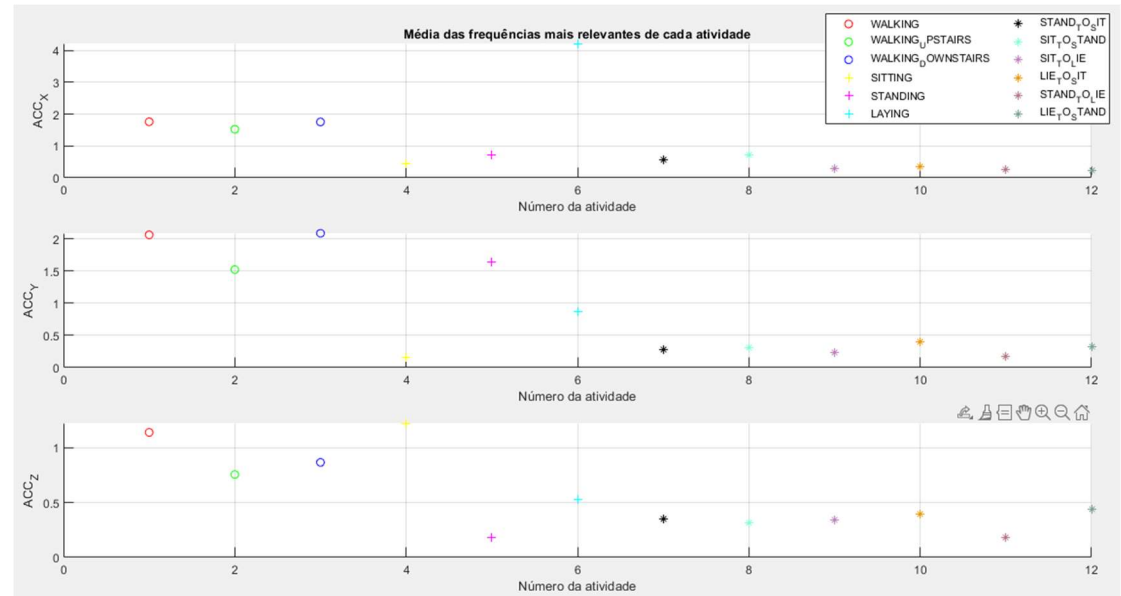


Figura 8 - Representação gráfica da média das componentes frequência mais relevantes de cada atividade realizada pelo utilizador 1

A partir dos resultados acima podemos tirar já algumas notas que virão a ser úteis mais à frente neste projeto. Numa análise superficial, conjugando a média das amplitudes mais relevantes e a média das frequências igualmente mais relevantes, conseguimos ver algumas diferenças nas atividades. Por exemplo, as atividades estáticas são as que apresentam uma média das suas componentes amplitude mais baixa, mas isso será um assunto abordado adiante.

Identificar o número de passos por minuto, para cada atividade dinâmica, de cada utilizador e apresentar a respetiva média e desvio padrão

Para este ponto, optamos por determinar o número médio de passos e respetivo desvio padrão para cada utilizador e para as experiências no geral. Para isso, desenvolvemos o método *CalcMediaDP* que ao receber a informação do ficheiro *labels.txt*, os dados a analisar e a frequência de amostragem, vai percorrer o sinal, atividade a atividade, e enviar os dados relativos a essa atividade para o método *StepsPerMinute* também por nós desenvolvido. Este método, à semelhança do método *diferenciaEstaticas*, vai através da magnitude devolvida pela DFT, calcular a frequência correspondente à maior componente magnitude. Tendo isso, para ter o número de passos por minuto, basta multiplicar a frequência obtida por 60 (1 minuto tem 60 segundos). O método *CalcMediaDP*, ao receber o valor do número de passos por minuto da atividade que esteve a considerar, vai adicioná-lo a um *array* com os valores do número de passos até então calculados da mesma atividade. No final, utiliza o método *mean* para obter a média e o *std* para obter o desvio padrão (métodos já existentes nas bibliotecas do *Matlab*).

Os resultados das experiências que tínhamos para analisar foram, de um modo geral:

ATIVIDADE	EIXO					
	X		Y		Z	
	MÉDIA	σ	MÉDIA	σ	MÉDIA	σ
WALKING	106.8244	6.0800	106.8244	6.0800	79.9841	27.0316
WALKING_UPSTAIRS	95.4954	6.8143	91.0909	14.5200	52.7444	16.9499
WALKING_DOWNSTAIRS	108.7771	10.8035	111.9341	29.7155	55.5641	16.3060

Tabela 1 – Média e desvio padrão de cada atividade dinâmica em cada eixo

Já de um modo mais específico, os resultados obtidos para cada utilizador foram:

Utilizador	ATIVIDADE	EIXO					
		X		Y		Z	
		MÉDIA	σ	MÉDIA	σ	MÉDIA	σ
1	WALKING	105.0932	3.7368	105.0932	3.7368	52.9549	2.7746
	WALKING_UPSTAIRS	91.3779	2.5942	91.3779	2.5942	44.9792	1.5943
	WALKING_DOWNSTAIRS	104.8022	6.6183	106.5097	5.2552	51.2090	3.3562
2	WALKING	107.1997	3.2547	107.1997	3.2547	107.1997	3.2547
	WALKING_UPSTAIRS	99.3117	3.7210	80.9597	25.7659	73.0540	25.9316
	WALKING_DOWNSTAIRS	111.2370	8.5585	98.3755	21.5580	68.6186	30.4907
3	WALKING	100.6734	2.8770	100.6734	2.8770	100.6734	2.8770
	WALKING_UPSTAIRS	96.8282	7.5884	96.8282	7.5884	47.1203	4.2691
	WALKING_DOWNSTAIRS	110.9219	17.2056	106.7385	17.1770	53.1108	10.1372
4	WALKING	116.0625	3.2013	116.0625	3.2013	86.1376	32.1555
	WALKING_UPSTAIRS	94.2417	9.5401	94.2417	9.5401	46.7613	5.2544
	WALKING_DOWNSTAIRS	108.8051	10.2806	133.4336	47.2653	50.8323	3.3956

Tabela 2 - Média e desvio padrão de cada atividade dinâmica em cada eixo para cada utilizador

Com isto, podemos concluir, tal como era expectável, que a maioria dos utilizadores realiza mais passos a descer as escadas (*WALKING_DOWNSTAIRS*) e, pelo contrário, realizam menos passos ao subir escadas (*WALKING_UPSTAIRS*).

À partida para este exercício consideramos 100 como um valor aceitável de número médio de passos por minuto e qualquer outro próximo de 100. Deste modo, é com alguma cautela que analisamos, por exemplo, o número médio de passos da atividade *WALKING_DOWNSTAIRS* do utilizador 1 para o eixo Y. Por muito que ao descer escadas o ritmo e a cadência de passos sejam mais elevados, consideramos que 133 passos por minuto (mais de 2 por segundo) seja demasiado e haja um afastamento do real valor do número médio de passos por minuto (daí o valor elevado do desvio padrão neste caso).

Identificar as características, por exemplo espectrais, que permitam diferenciar as atividades estáticas, de transição e dinâmicas

Pegando na função já acima mencionada *diferenciaEstaticas*, fomos calcular o valor médio das componentes amplitude e frequência para cada atividade. Graficamente, os resultados contabilizando todas as experiências foram os seguintes:

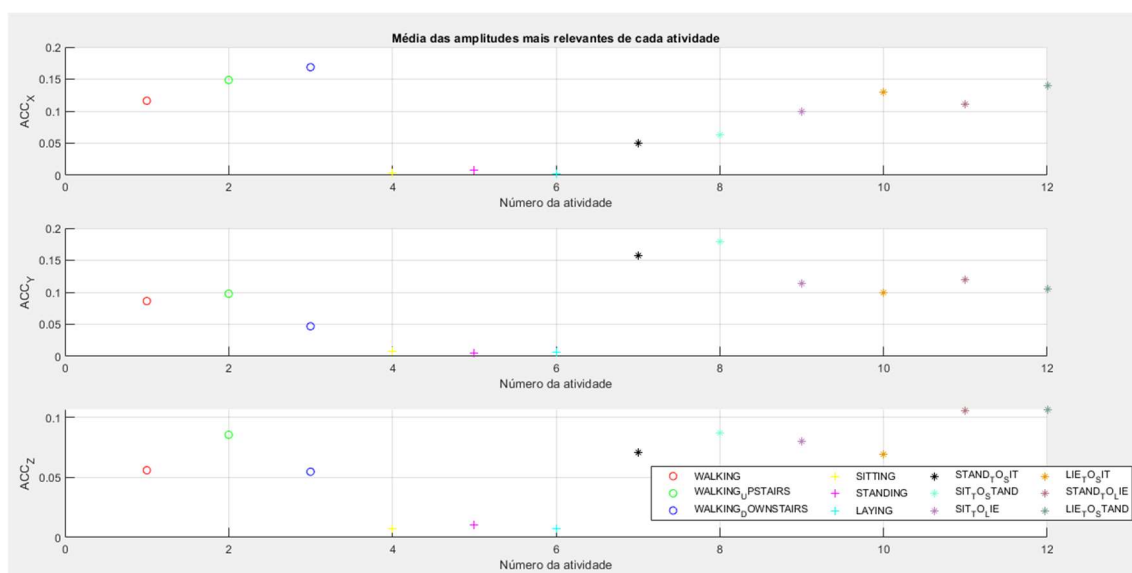


Figura 9 - Representação gráfica da média das componentes amplitude mais relevantes de cada atividade

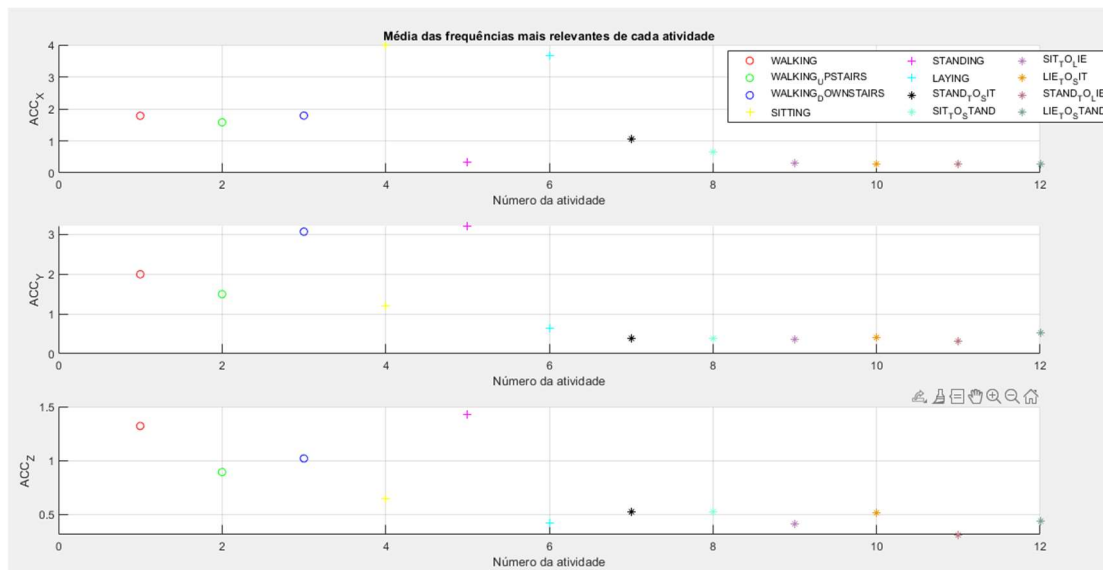


Figura 10 - Representação gráfica da média das componentes frequência mais relevantes de cada atividade

Tal como mencionamos anteriormente, as atividades estáticas são as que apresentam uma média das suas componentes amplitude mais baixas em todos os eixos. Apesar de todas as atividades apresentarem médias da componente amplitude muito próximas de 0, são as das atividades estáticas que mais se aproximam desse valor e isso é notório pelo gráfico acima (figura 9). Deste modo, conseguimos distinguir as atividades estáticas das restantes.

Falta, neste momento, arranjar alguma especificidade que consiga diferenciar as atividades de transição das dinâmicas. Se analisarmos o gráfico da figura 10 reparamos que as médias das frequências mais relevantes das atividades dinâmicas são superiores em todos os eixos comparativamente com as das atividades de transição. Com isto, conseguimos distingui-las.

Identificar as características, por exemplo espectrais, que permitam diferenciar cada atividade

Sendo os gráficos acima referentes à média das componentes amplitude e frequência mais relevantes de cada atividade, um sinal que apresente componentes semelhantes a uma atividade Y pode ser catalogado como atividade Y. Deste modo, neste ponto decidimos representar graficamente com menos granularidade para conseguir distinguir melhor cada tipo de atividade. Para isso decidimos representar, tal como acima, a média das componentes frequência e amplitude mais relevantes só que desta vez fazendo um gráfico para cada tipo de atividade, isto é, para as dinâmicas, para as estáticas e para as de transição. Os resultados foram os seguintes:

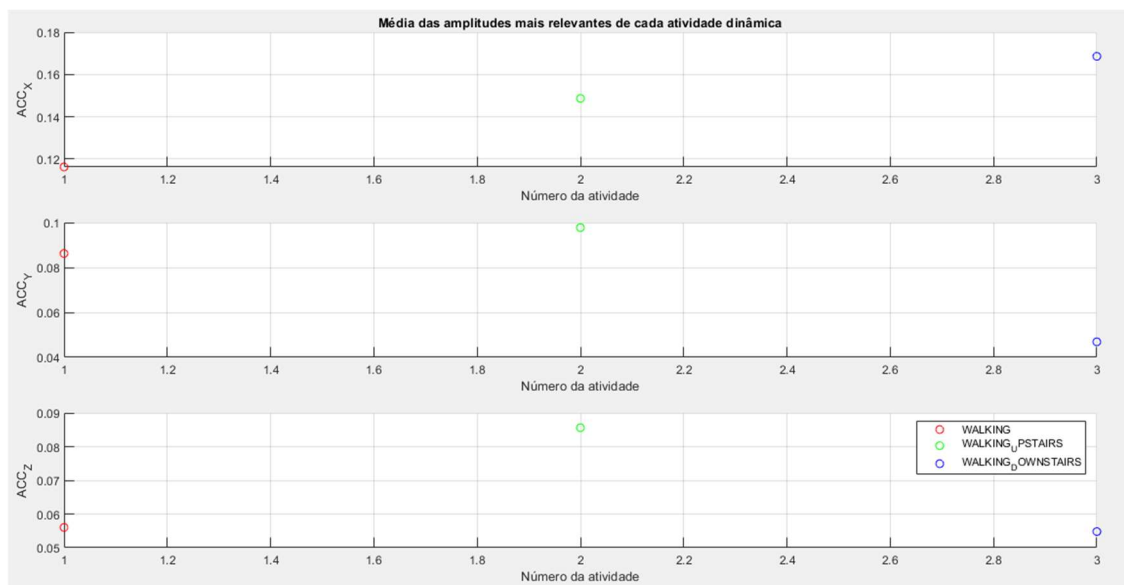


Figura 11 - Representação gráfica da média das componentes amplitude mais relevantes de cada atividade dinâmica

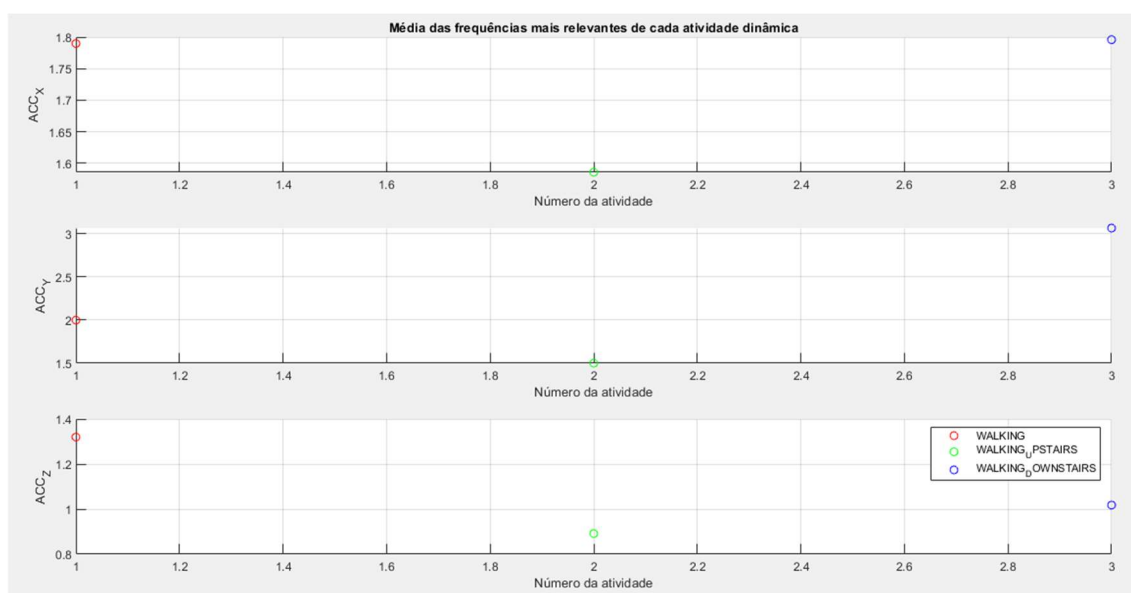


Figura 12 - Representação gráfica da média das componentes frequência mais relevantes de cada atividade dinâmica

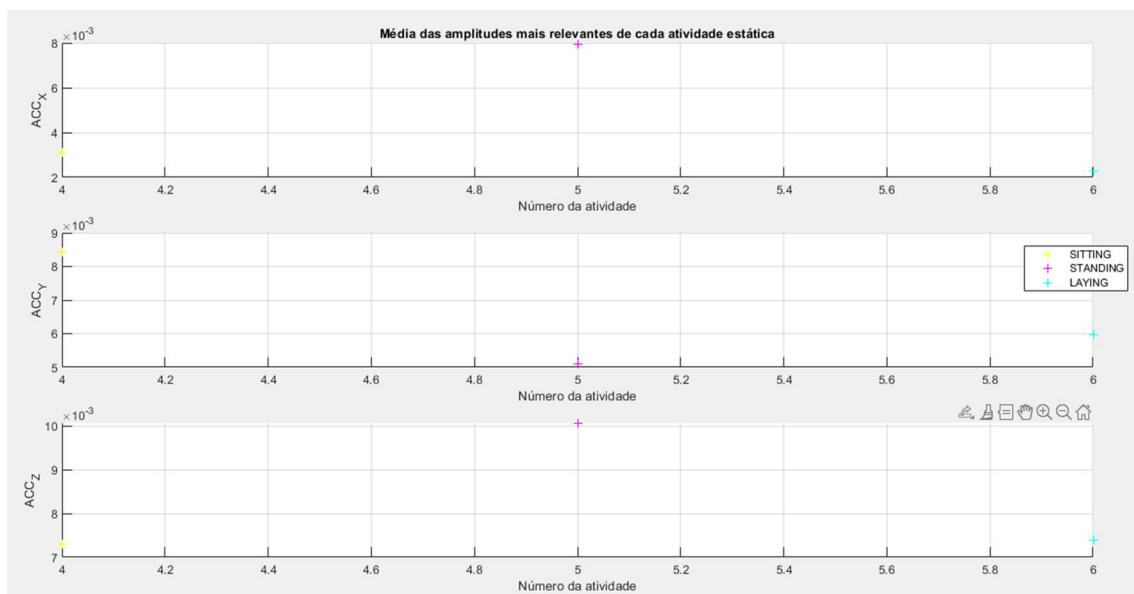


Figura 13 - Representação gráfica da média das componentes amplitude mais relevantes de cada atividade estática

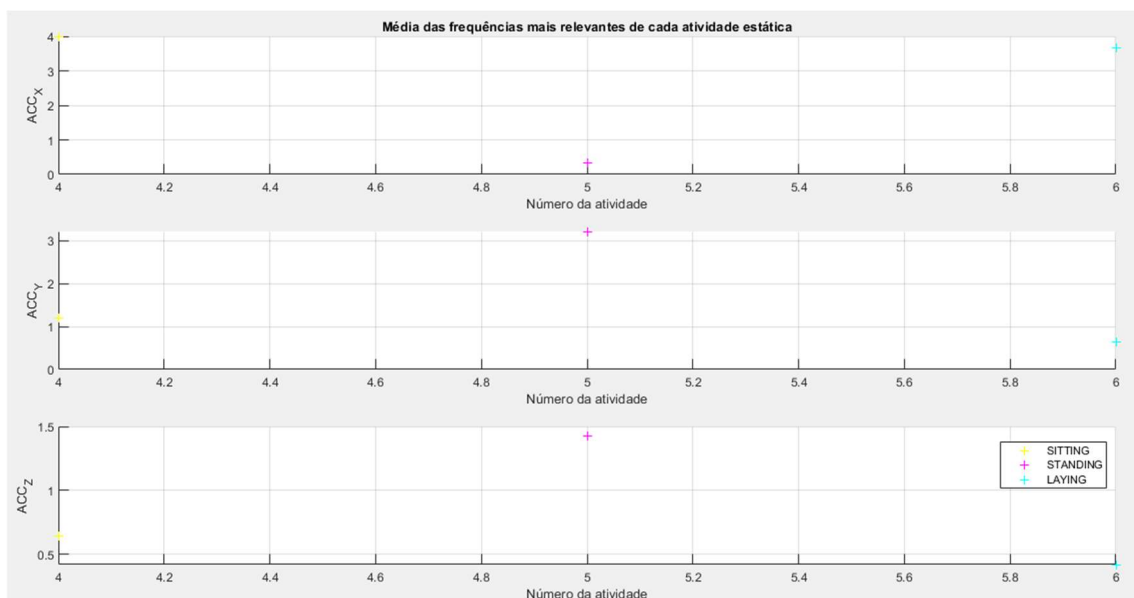


Figura 14 - Representação gráfica da média das componentes frequência mais relevantes de cada atividade estática

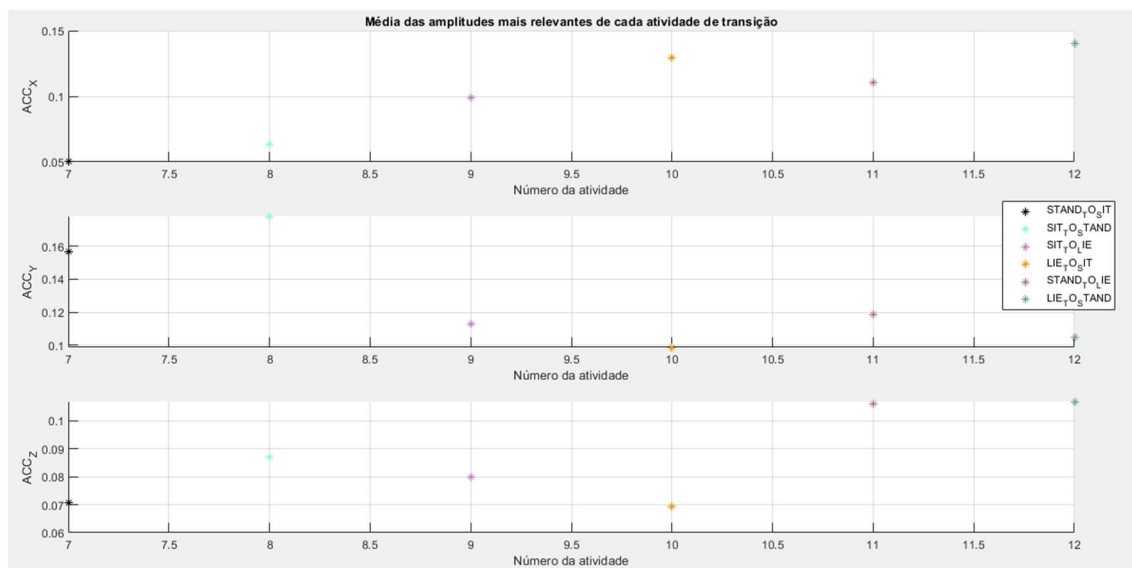


Figura 15 - Representação gráfica da média das componentes amplitude mais relevantes de cada atividade de transição

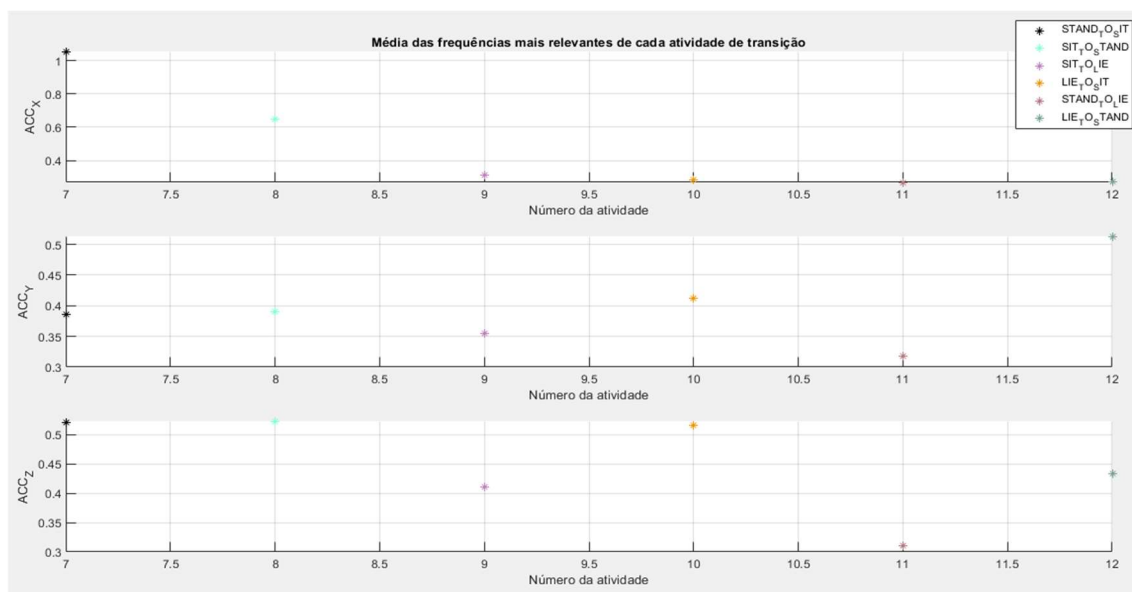


Figura 16 - Representação gráfica da média das componentes frequência mais relevantes de cada atividade de transição

Tal como podemos constatar, existem claras diferenças entre cada atividade. Se à média das frequências mais relevantes do sinal aliarmos a média das amplitudes mais relevantes do sinal, fica ainda mais fácil de identificar cada atividade. Por exemplo, nas atividades dinâmicas o *WALKING* e o *WALKING_DOWNSTAIRS* apresentam médias das componentes frequência mais relevantes muito semelhantes no eixo X. Só com este dado seria muito difícil de as distinguir. No entanto, se adicionarmos à equação o valor da média das componentes amplitude mais relevantes, rapidamente distinguimos estas duas atividades, pois no eixo X o *WALKING* apresenta valores menor de amplitude que o *WALKING_DOWNSTAIRS*. Tal como este, outros casos existem em que a utilização destas duas componentes é essencial para se conseguirem distinguir atividades.

Apesar de cada utilizador ser um utilizador e os dados variarem um pouco entre si mediante o indivíduo que realiza a experiência, decidimos fazer a média dos dados totais dos 4 utilizadores que temos de analisar, uma vez que as atividades têm determinadas características que as distinguem independentemente de quem as pratica. Deste modo, criamos a tabela seguinte com o resumo dos gráficos acima onde constam, para cada eixo e atividade, a média das componentes frequência e amplitude mais relevantes. Qualquer atividade analisada que tenha médias das mesmas componentes parecidas com as que constam na tabela, pode ser associada a determinada atividade.

ATIVIDADE	EIXO					
	X		Y		Z	
	Freq [Hz]	Amp [rad]	Freq [Hz]	Amp [rad]	Freq [Hz]	Amp [rad]
WALKING	1.7898	0.1163	1.9950	0.0862	1.3198	0.0560
WALKING_UPSTAIRS	1.5858	0.1487	1.4950	0.0978	0.8912	0.0856
WALKING_DOWNSTAIRS	1.7959	0.1686	3.0656	0.0469	1.0179	0.0548
SITTING	4.0135	0.0031	1.2085	0.0084	0.6458	0.0073
STANDING	0.3229	0.0079	3.2144	0.0051	1.4289	0.0101
LAYING	3.6738	0.0023	0.6354	0.0060	0.4195	0.0074
STAND_TO_SIT	1.0546	0.0507	0.3857	0.1568	0.5197	0.0707
SIT_TO_STAND	0.6481	0.0635	0.3908	0.1783	0.5226	0.0870
SIT_TO_LIE	0.3151	0.0988	0.3550	0.1131	0.4101	0.0798
LIE_TO_SIT	0.2832	0.1292	0.4123	0.0987	0.5148	0.0693
STAND_TO_LIE	0.2699	0.1105	0.3184	0.1189	0.3113	0.1060
LIE_TO_STAND	0.2732	0.1403	0.5132	0.1046	0.4339	0.1068

Tabela 3 - Média das componentes frequência e amplitude mais relevantes de cada atividade em cada eixo

Apesar de a visualização das diferenças entre atividades seja mais notória pelos gráficos representados acima, estes dados serão úteis para o ponto que veremos a seguir.

Utilizar a *Short-Time Fourier Transform* (STFT) para obter computacionalmente as distribuições tempo-frequência para o sinal do acelerómetro Z

Tal como já referimos acima, uma forma de localizar temporalmente as frequências num dado sinal é fazer a DFT em janelas que “deslizam” pelo espetro. Este é basicamente o conceito principal da STFT (calcular a DFT de um sinal recorrendo a janelas deslizantes).

Para fazer este ponto, implementamos o método *stft* que ao receber o sinal e a frequência de amostragem, vai calcular o tamanho de cada janela (definimos como sendo 2% do espectro total), a sobreposição das janelas (50%) e o número de pontos por janela. Tendo estes dados e o vetor de frequências calculados, o método vai passar a percorrer o sinal que o utilizador lhe enviou por janelas deslizantes. Em cada janela será aplicada a janela de Hamming (escolhida anteriormente neste projeto) e o sinal será tratado de modo a ficar sem tendência (utilizando *detrend*). Tendo isto, basta localizar onde está a frequência positiva correspondente à maior magnitude do sinal, do mesmo modo que fizemos e explicamos anteriormente quando pretendemos identificar as características mais relevantes para cada atividade. Tendo o valor dessa frequência, o método vai armazená-lo num *array* definido para o efeito. No final basta visualizar os valores pertencentes a esse *array* para ver as frequências relevantes do sinal.

Os valores que obtivemos para a experiência número 1 (eixo Z) foram os seguintes:

Columns 1 through 14													
0.2427	0.1214	0.1214	0.1214	0.1214	0.4854	0.1214	0.1214	0.1214	0.1214	0.2427	0.1214	0.1214	0.1214
Columns 15 through 28													
0.1214	0.1214	0.1214	0.1214	0.1214	0.1214	0.1214	0.1214	0.1214	0.1214	0.1214	0.1214	0.1214	0.1214
Columns 29 through 42													
0.1214	0.1214	0.1214	0.1214	0.1214	0.2427	0.3641	0.8495	0.8495	0.9709	0.8495	0.1214	0.8495	0.9709
Columns 43 through 56													
0.9709	0.8495	0.8495	0.1214	0.8495	0.9709	0.8495	0.9709	0.8495	0.7282	0.9709	0.8495	0.8495	0.8495
Columns 57 through 70													
0.8495	0.8495	0.9709	0.1214	0.1214	0.1214	0.1214	0.7282	0.7282	0.8495	0.7282	0.1214	0.7282	0.7282
Columns 71 through 84													
0.7282	0.8495	0.8495	0.9709	0.8495	0.7282	0.8495	0.7282	0.7282	0.9709	0.8495	0.8495	0.8495	0.7282
Columns 85 through 98													
0.7282	0.7282	0.7282	0.1214	0.1214	0.3641	0.8495	0.8495	0.6068	0.4854	0.6068	0.1214	0.1214	0.4854

Figura 17 - Resultado das frequências relevantes em cada janela da STFT

Analisando os resultados da figura 17, percebemos que há algumas frequências que não são muito semelhantes às que calculamos anteriormente e mostramos na tabela 3 (0,1214 Hz e 0,2427 Hz, por exemplo). Porém, existem frequências detetadas pela STFT como sendo relevantes nas janelas em questão semelhantes às médias que mostramos acima neste relatório. Exemplo disso é a frequência de 0,3641 Hz. Apesar de não termos nenhuma igual na nossa tabela, podemos supor que esta atividade se trata de STAND_TO_LIE, de SIT_TO_LIE, de LIE_TO_STAND ou de LAYING. Todas estas atividades têm frequências médias semelhantes à em questão. Já na frequência de 0,8495 Hz supomos que a atividade em causa seja WALKING_UPSTAIRS. Embora, mais uma vez, o valor da frequência não seja exatamente igual ao que temos como referência, está relativamente próximo e não existe mais nenhuma atividade com frequência média semelhante. Deste modo, podemos dizer com relativa confiança que a atividade realizada na janela cuja frequência relevante foi 0,8495 Hz é WALKING_UPSTAIRS. Pela mesma razão, supomos que a atividade da janela que devolveu 0,9709 Hz como frequência relevante seja WALKING_DOWNSTAIRS.

Já quando aparece a frequência de 0,7282 Hz existem mais dúvidas. Apesar de este valor ficar entre 2 que constam na nossa tabela de referência, fica um pouco afastado. Porém, podemos supor com alguma confiança que a atividade em questão seja ou SITTING ou WALKING_UPSTAIRS, sendo a primeira mais provável. Para finalizar, a frequência de 0,6068 Hz

muito provavelmente diz respeito à atividade de SITTING (que tem uma frequência média no eixo dos X de 0,6458 Hz).

Tal como era de esperar, identificar atividades apenas através de um dado espectral (frequência relevante) de um eixo (Z) não é nada fácil. Embora possamos ter algumas ideias do que aquela frequência pode representar em termos de atividade, é sempre muito complicado acertar a 100%.

Outra maneira de chegar ao mesmo resultado (se não fosse pedido para implementar uma função de raiz) seria aplicar o método *spectrogram* das bibliotecas do *Matlab* que devolve já o *array* das frequências mais relevantes. Apesar de esta não ser a solução que nos foi requerida, decidimos comparar os resultados obtidos pela nossa função com os devolvidos pelo método do *Matlab*.

Conclusão

A análise de dados é um ramo da ciência interessantíssimo que tem vindo a crescer ao longo dos últimos anos. Este projeto teve como intuito analisar dados nos domínios do tempo e da frequência, usando sinais recolhidos de acelerómetros de smartphones com o objetivo de identificar a atividade realizada.

Para analisar os sinais recorremos a DFT e a STFT (DFT só que recorrendo a janelas deslizantes). Um dos grandes objetivos era conseguir através da STFT distinguir as atividades através da componente frequência relevante em cada janela dessa transformada. No entanto, como utilizamos valores médios de todas as experiências (envolvendo vários utilizadores) para comparar com os das frequências relevantes da STFT, nem sempre conseguimos identificar a 100% a atividade em causa. Apesar disso, esta questão já era expectável sobretudo sendo o eixo em análise nessa questão o Z (o que limita a análise). Consideramos, assim, que realizamos tudo corretamente.

Em suma, pensamos que conseguimos analisar o pretendido recorrendo a técnicas abordadas nas aulas teóricas e prático-laboratoriais da cadeira de Análise e Transformação de Dados.

Bibliografia / Webgrafia

- Materiais disponibilizados pelo docente da cadeira de Análise e Transformação de Dados [21 de maio de 2021];
- “Window Function” in en – Wikipedia. https://en.wikipedia.org/wiki/Window_function [18 de maio de 2021].