

Instructions for Authors of SBC Conferences Papers and Abstracts

Evertt Arima Machado¹

¹CMCC – Universidade Federal do ABC (UFABC)
Praia Grande, SP – Brazil

arima.evertt@ufabc.edu.br

Abstract. *This paper describes the attempt to reimplementation of part of the experiments conducted by [Nickel et al, 2011] on Accelerometer-Based Biometric Gait Recognition. Although the original dataset and related algorithm are not available, we used a similar data and develop the code in order to compare results. In the process, we found appropriate to suggest some changes in the algorithm and sequence of steps taken to achieve conclusive results.*

Resumo. *Este artigo descreve a tentativa de reimplementar parte dos experimentos conduzidos por [Nickel et al, 2011] sobre Reconhecimento Biométrico baseado em Acelerômetro. Apesar dos dados e algoritmos originais não estarem disponíveis, utilizamos uma base de dados similar e desenvolvemos o código afim de comparar os resultados. No processo, acreditamos apropriado sugerir algumas mudanças no algoritmo e na sequência de passos utilizados para alcançar resultados conclusivos.*

1. Introdução

O termo Biometria tem origem na junção das palavras gregas *bios* (vida) e *metron* (medida). Seu significado implícito é a medição biológica, ou seja, é o estudo das características fisiológicas e comportamentais que são exclusivas de cada pessoa.

As características fisiológicas de uma pessoa, como a íris, a digital, a retina, a face podem mudar somente em circunstâncias extremas como acidentes ou traumas. Já as características comportamentais do indivíduo evoluem com o tempo e são suscetíveis a alterações por fatores como estresse, problemas de saúde ou sentimentos de situações, como uma situação de perigo.

As características comportamentais do indivíduo. Neste campo podemos destacar o reconhecimento do estilo de escrita ou estilometria, o reconhecimento da voz, o reconhecimento da dinâmica de digitação ou dos movimentos que se faz com o *mouse* e o reconhecimento da forma de caminhar. Em nosso estudo, nos limitaremos a este último.

O modo de caminhar é, segundo [Santos, 2017], um dos que nesta área, se destaca como sendo bastante eficaz. Contudo é salientado problemas como a baixa distinção e permanência, o que dificulta a separabilidade dos indivíduos, além da dificuldade na manutenção do sistema biométrico. Assim sendo, o sistema de reconhecimento da forma de caminhar não funciona como um sistema completo, mas como um sistema complementar a outros sistemas biométricos.

Um sistema biométrico baseado na forma de caminhar pode ser dividido em dois grandes grupos, sendo o primeiro a partir de imagens de vídeos, capturados por meio de

câmeras, cujos quadros são utilizados para descrever o caminhar de um indivíduo. O segundo modo é a partir de sensores. Neste, os dados são obtidos por meio de sensores posicionados no chão, no corpo ou em dispositivos móveis, como *smartphones*. Em nosso experimento, utilizaremos dados coletados por meio do acelerômetro de um *smartphone*.

1.1. Organização do texto

Este artigo está organizado em seis capítulos. O capítulo 2 visita trabalhos relacionados à biometria, em especial as que utilizam a biometria do modo de caminhar. O capítulo 3 explana a proposta principal deste artigo. O capítulo 4 descreve a metodologia empregada e detalhes dos dados e métodos utilizados. O capítulo 5 destaca os resultados obtidos na experimentação e por fim o capítulo 6 conclui este artigo, apontando possíveis trabalhos futuros.

2. Trabalhos Relacionados

Em geral, os sistemas biométricos comportamentais baseado na forma de caminhar são divididos em dois grandes grupos. O primeiro modo coleta os dados por meio de câmeras de vídeo que, uma vez gravado, a silhueta do indivíduo por meio das quais sua forma de caminhar é representada em um modelo matemático.

O segundo modo é a obtenção dos dados por meio de sensores, que podem estar posicionados no chão, no corpo ou em dispositivos móveis, como *smartphones*. Em nosso experimento, utilizaremos dados coletados por meio do acelerômetro de um *smartphone*.

O posicionamento de sensores no chão obtém dados significativos. Ele possui vantagens deste ponto de vista uma vez que, ao caminhar sobre uma área com sensores dedicados, permite obtenção de atributos como a força aplicada contra o solo, a massa corporal [Jenkins, 2007], além de outros dados que podem ser extraídos das imagens geradas pelos sensores que são ativados no decorrer da caminhada, como distâncias e ângulos dos pés [Santos, 2017]. A maior desvantagem é que é custoso e passivo, no sentido que possui mobilidade limitada.

Uma outra modalidade de sistema de biometria comportamental por meio de sensores é fixar sensores dedicados a partes específicas do corpo do usuário ou de sua vestimenta. Este fato o torna um sistema pouco prático, uma vez que o indivíduo dever ser preparado para, não somente a obtenção dos atributos da marcha, mas também para teste ou para qualquer utilização destes dados como forma de reconhecimento ou autenticação.

Entendemos que dispositivos móveis, em especial *smartphones* atuais que contam com sensores incorporados, atualmente massificados, são uma opção plausível para se coletar os atributos da caminhada de um indivíduo.

Neste aspecto, o acelerômetro é utilizado como principal sensor, como observa [Santos, 2017], coletando informações de forma contínua e, possibilitando uma autenticação do indivíduo também de forma contínua. O acelerômetro registra as acelerações do usuário em três direções distintas, x, y e z, durante a marcha.

De forma a trabalhar estes atributos, [Nickel *et al*, 2011] propõe a utilização de métodos estatísticos para a extração de características. Por outro lado, [Preece *et al*, 2009] afirma que experimentos demonstram que utilizar a *Fast Fourier Transform* (FFT), melhora consideravelmente os resultados obtidos.

3. Proposta do Experimento

Neste artigo, buscamos reproduzir os experimentos relatados em nosso artigo base e comparar o desempenho do classificador SVM (*Support Vector Machine*) no reconhecimento e classificação de usuários de dispositivos celulares por meio do reconhecimento comportamental, utilizando para tal fim, os dados do acelerômetro do dispositivo.

Aplicaremos técnicas distintas para a extração das características (*features*) dos dados brutos, na fase de pré-processamento dos dados, o que resultará em características diferentes para os mesmos dados.

Para a extração das características, teremos um método estatístico, no qual.

Utilizando o mesmo classificador e hiperparâmetros, pretendemos comparar os resultados obtidos de forma a elucidar qual das duas técnicas de extração de dados obtém melhor desempenho na classificação de um usuário genuíno.

Utilizaremos como classificador o SVM (*Support Vector Machine*). O ajuste inicial do experimento para hiperparâmetros foi de $C = 1.0$, kernel = “RBF” (*Radial Basis Function*) e gamma = “auto”.

Uma vez obtidos os resultados utilizaremos um algoritmo para obter um ajuste automatizado desses hiperparâmetros, fixando como entrada neste algoritmo apenas um grupo de dados, sendo primeiramente os estatísticos, no domínio do tempo e em um segundo momento, os resultantes do *Fast Fourier Transform* (FFT), no domínio da frequência. Obtemos como saída deste o kernel que mais se adequa ao grupo de dados de entrada, entre linear ou rbf, e o “C” que poderá variar entre [0.1, 0.5, 1, 5, 10]. O gamma permanecerá fixo em “auto”.

4. Metodologia Experimental

O experimento tomará como base o artigo produzido por [Nickel *et al*, 2011]. Contudo, os dados originais do experimento não estão disponíveis de forma que utilizaremos uma base de dados públicos similares. Vale salientar que substituiremos os dados do sistema de coordenadas baseada no mundo para uma baseada no usuário, conforme descrito em [Santos, 2011]. O algoritmo utilizado no experimento de [Nickel *et al*, 2011] não se encontra disponibilizado. Portanto desenvolveremos a sua implementação.

Estes dois fatores isolados já comprometerão os resultados. Contudo, buscamos confrontar as técnicas utilizadas por [Nickel *et al*, 2011], com as descritas por [Preece *et al*, 2009], de forma a estabelecer qual das duas técnicas de extração de dados obtém maior taxa de sucesso no reconhecimento de um usuário genuíno. As métricas que utilizaremos para verificar o sucesso do experimento, de forma simples, é a Taxa de Falso Acerto (FMR – *False Match Rate*) e a Taxa de Falso Não-Acerto (FNMR – *False Non Match Rate*), além da Acurácia Balanceada.

4.1. Dados

O conjunto de dados utilizados para o experimento forma coletados de 50 usuários caminhando por cinco minutos [Santos, 2017]. Os dados de atividades de não-caminhada foram descartados.

Para coletar os dados foi utilizado um *smartphone* LG Nexus 5, em duas sessões

de cinco minutos, em dias distintos. Todos os dados foram coletados utilizando uma frequência de 40Hz. Os nomes dos arquivos contêm a identificação numérica do usuário, bem como a sessão (1 ou 2). Os dados estão divididos em 3 sistemas, sendo o Sistema de Coordenadas do Dispositivo, do Mundo e do Usuário.

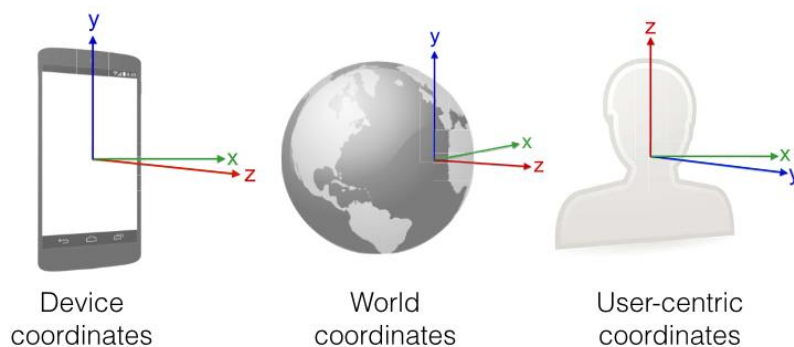


Figura 1. Diferentes sistemas de coordenadas para representar dados obtidos com o acelerômetro de dispositivos móveis. [Santos, 2017]

No primeiro, o Sistema de Coordenadas do Dispositivo (*device_coordinates*), temos os dados do acelerômetro e do vetor de rotação interpolados. O segundo sistema é o chamado Sistema de Coordenadas do Mundo (*world_coordinates*). Este refere-se aos dados do acelerômetro, rotacionado para o Sistema de Coordenadas Mundial, sem o *timestamp*, uma vez que o processo de alinhamento dos dados de aceleração e vetor de rotação sincroniza ambos os dados, resultando nas acelerações nos eixos x, y e z em reais coordenadas mundiais.

Para nosso experimento, utilizaremos o terceiro grupo de dados, o Sistema de Coordenadas baseado no Usuário. Tais valores são obtidos rotacionando os dados com base no sistema mundial para basear-se no usuário, método proposto e descrito por [Santos, 2017].

4.2. Métodos Comparados (Estatístico versus FFT)

Há, como mencionado no item 3, pretendemos comparar as métricas obtidas por meio de características estatísticas dos dados no domínio do tempo com as métricas resultantes de características extraídas pelo FFT, no domínio da frequência.

No que concerne ao domínio do tempo, os métodos estatísticos, os dados já interpolados devem ser centralizados em torno de zero. Segundo [Nickel *et al*, 2011], isto deve ser feito pois, como mencionado em 4.1, utilizamos o acelerômetro de um *smartphone* para obter os dados $\mu: \bar{s}_a(t) = \bar{s}_a(t) - \mu_a, a \in \{x, y, z\}$. Os *smartphone* não são muito bem calibrados, de forma que a gravidade deve ser retirada da aceleração vertical. Podemos fazer isto simplesmente subtraindo o valor médio de cada eixo x, y e z, de seus respectivos valores.

Após centralizado, particionamos os dados em janelas. Utilizaremos janelas de 3000ms, 5000ms e 7000ms sugeridas por [Nickel *et al*, 2011] para realizar a experiência, com uma sobreposição de 50%. Desta forma cada segmento conterá os dados de aproximadamente 3 a 7 ciclos de marcha, considerando um ciclo por segundo.

Após estes procedimentos, extrairemos os valores máximo, mínimo, média e

desvio padrão em cada segmento, das acelerações em x, y e z, bem como do vetor de magnitude, este último calculado por $s_m = \sqrt{\bar{s}_x^2 + \bar{s}_y^2 + \bar{s}_z^2}$.

Já no domínio da frequência, o método aplicando o FFT aos cinco primeiros valores das acelerações de x, y e z de cada janela, obteremos cinco números complexos como resultado. São os valores absolutos destes números complexos que utilizaremos em nosso experimento.

O FFT é aplicado somente aos cinco primeiros números pois são os mais significativos, evitando processar computacionalmente todos os dados de cada janela que resultaria em sua grande parte de ruídos.

É importante ressaltar que o tamanho das janelas. Há muita discussão sobre como dividir os dados em janelas. Uma das formas utilizadas é o ciclo da marcha. Segundo [Hoang *et al*, 2013], um ciclo é definido como “o intervalo de tempo entre duas ocorrências sucessivas de um dos eventos repetitivos” quando caminhamos.

Portando, [Hoang *et al*, 2013] entende que a sucessão de ciclos é o que define a forma de caminhar de um indivíduo. Contudo, o janelamento utilizando tal técnica é complexo e não o utilizaremos para este experimento. Para [Wang *et al*, 2007], por meio de análise dos sinais brutos de aceleração da forma de caminhar em superfície plana, a janelamento dos dados deve ser de 2 segundos.

Ao tentarmos recriar o experimento de [Nickel *et al*, 2011], usaremos para nosso experimento as mesmas janelas utilizadas, sendo elas de 3s, 5s e 7s, conforme citado no item 4.2.

Quanto ao classificador, optamos pelo SVM (*Support Vector Machine*), utilizado no trabalho de [Nickel *et al*, 2011], com **kernel** = ‘rbf’ (*Radial Basis Function*).

4.3. Conjuntos de Treino e Teste

Conforme mencionado no item 4.1, os dados dos 50 usuários foram coletados em duas sessões. Portanto, utilizaremos a primeira sessão como sessão de treinamento e a segunda sessão como teste.

A sessão 1, o conjunto de treino, utilizaremos 100% dos dados do usuário genuíno. Os dados outros 49 usuários serão carregados em memória, contudo, a utilização de 100% dos dados resultaria em um conjunto de treino desbalanceado, afetando diretamente o resultado do experimento. Para balancear o conjunto de teste, foram escolhidos aleatoriamente a quantidade exata de janelas final dos dados do usuário genuíno.

Já a sessão de teste, 100% dos dados de todos os 50 usuários foram utilizados.

5. Resultados

Trataremos separadamente os resultados para cada uma das janelas especificadas no item 4.2. Com base nos resultados de [Nickel *et al*, 2011], a primeira implementação para cada janela foi utilizando hiperparâmetros fixos para o classificador SVM, sendo o kernel = ‘rbf’ e $C = 1.0$.

A segunda implementação, os hiperparâmetros citados são definidos por meio de um *grid search*, que recebe com atributo de entrada as características de teste extraídas por meio dos métodos estatísticos, como também utilizado por [Nickel *et al*, 2011].

A terceira implementação, novamente utilizamos o *grid search*, contudo o atributo de entrada são as características de teste extraídas por meio do FFT, como indicado por [Preece *et al*, 2009].

Para todas as implementações utilizaremos $\gamma = 'auto'$.

5.1. Tamanhos de Janelas

Os resultados obtidos, de forma geral, ficaram abaixo de nossa expectativa inicial, para todos os ajustes possíveis, dado a pesquisa conduzida por [Nickel *et al*, 2011] e [Preece *et al*, 2009].

Separados pelo tamanho de janelamento, verificamos que a variação dos resultados não é muito grande, para qualquer uma das três janelas estudadas. Verificamos que, independente do tamanho da janela, ou forma de ajuste em nossos hiperparâmetros, o classificador obteve uma acurácia balanceada menor em todas as instâncias de nossas experimentações, quando utilizamos como atributos, as características extraídas por meio do FFT.

As Tabelas 1, 2 e 3 reúnem os resultados das três formas de implementação para uma janela de 3 segundos, 5 segundos e 7 segundos respectivamente. Há, em todas elas, dois valores destacados, indicando o melhor e o pior resultado obtido dentre os resultados da janela ao qual pertence. Dentre estes 6 valores, há dois valores somente em que a célula está ressaltada indicando o melhor e o pior resultado obtido dentre todos os testes realizados.

A Tabela 1, utiliza uma janela de 3 segundos, e obteve em seu escopo, os melhores resultados entre 0,857 a 0,871 para acurácia balanceada, todos referentes ao Método Estatístico de extração de características. Os piores resultados dentro do escopo de janela variam entre 0,611 a 0,587, todos no domínio da frequência. Ambos aconteceram para $C = 10$ e $kernel = "RBF"$, utilizando *grid search* sobre as características estatísticas. Este último valor, 0,587 relativo às características no domínio da frequência, refere-se também ao pior resultado geral, independente da janela ou ajuste dos hiperparâmetros.

A Tabela 2, utiliza uma janela de 5 segundos, e obteve em seu escopo, os melhores resultados entre 0,896 a 0,873 para acurácia balanceada, todos no domínio do tempo. Os piores resultados dentro do escopo de janela variam entre 0,627 a 0,63, todos no domínio da frequência. Ambos aconteceram para $C = 10$ e $kernel = "RBF"$, utilizando *grid search* sobre as características estatísticas.

Neste escopo de janela (5 segundos), [Nickel *et al*, 2011] afirmam ter conseguido os melhores resultados, utilizando o SVM ($kernel = 'RBF'$ e $C = 1.0$) e sistema de coordenadas mundiais.

Tabela 1. Métricas obtidas com uma janela de 3 segundos

	Manual	
Janela (segundos)	3	
Kernel (RBF ou Linear)	RBF	
C (0.1, 0.5, 1, 5, 10)	1	
Método	Estatístico	FFT
Total de instâncias de usuários genuínos:	198	
Total de usuário Verdadeiro/predições Verdadeiro:	189	111
Total de usuário Verdadeiro/predições Falso:	9	87
Total de instâncias de usuários impostores:	9702	
Total de usuário Falso/predições Verdadeiro:	2341	3275
Total de usuário Falso/predições Falso:	7361	6427
False Match Rate:	0.24129045557616985	0.3375592661306947
False Non Match Rate:	0.04545454545454545	0.4393939393939394
Balanced Accuracy:	0.8566274994846423	0.611523397237683
	Estatístico	
Janela (segundos)	3	
Kernel (RBF ou Linear)	RBF	
C (0.1, 0.5, 1, 5, 10)	10	
Método	Estatístico	FFT
Total de instâncias de usuários genuínos:	198	
Total de usuário Verdadeiro/predições Verdadeiro:	186	105
Total de usuário Verdadeiro/predições Falso:	12	93
Total de instâncias de usuários impostores:	9702	
Total de usuário Falso/predições Verdadeiro:	1914	3457
Total de usuário Falso/predições Falso:	7788	6245
False Match Rate:	0.19727891156462585	0.35631828488971345
False Non Match Rate:	0.06060606060606061	0.4696969696969697
Balanced Accuracy:	0.8710575139146568	0.5869923727066584
	FFT	
Janela (segundos)	3	
Kernel (RBF ou Linear)	RBF	
C (0.1, 0.5, 1, 5, 10)	5	
Método	Estatístico	FFT
Total de instâncias de usuários genuínos:	198	
Total de usuário Verdadeiro/predições Verdadeiro:	186	108
Total de usuário Verdadeiro/predições Falso:	12	90
Total de instâncias de usuários impostores:	9702	
Total de usuário Falso/predições Verdadeiro:	1979	3439
Total de usuário Falso/predições Falso:	7723	6263
False Match Rate:	0.20397856112141827	0.3544629973201402
False Non Match Rate:	0.06060606060606061	0.45454545454545453
Balanced Accuracy:	0.8677076891362605	0.5954957740672027

Tabela 2. Métricas obtidas com uma janela de 5 segundos

	Manual	
Janela (segundos)	5	
Kernel	RBF	
C	1	
Método	Estatístico	FFT
Total de instâncias de usuários genuínos:	118	
Total de usuário Verdadeiro/predições Verdadeiro:	116	69
Total de usuário Verdadeiro/predições Falso:	2	49
Total de instâncias de usuários impostores:	5782	
Total de usuário Falso/predições Verdadeiro:	1363	1913
Total de usuário Falso/predições Falso:	4419	3869
False Match Rate:	0.23573158076790038	0.3308543756485645
False Non Match Rate:	0.01694915254237288	0.4152542372881356
Balanced Accuracy:	0.8736596333448634	0.62694569353165
	Estatístico	
Janela (segundos)	5	
Kernel	RBF	
C	10	
Método	Estatístico	FFT
Total de instâncias de usuários genuínos:	118	
Total de usuário Verdadeiro/predições Verdadeiro:	115	70
Total de usuário Verdadeiro/predições Falso:	3	48
Total de instâncias de usuários impostores:	5782	
Total de usuário Falso/predições Verdadeiro:	1054	1922
Total de usuário Falso/predições Falso:	4728	3860
False Match Rate:	0.1822898650985818	0.3324109304738845
False Non Match Rate:	0.02542372881355932	0.4067796610169492
Balanced Accuracy:	0.8961432030439295	0.6304047042545832
	FFT	
Janela (segundos)	5	
Kernel	RBF	
C	10	
Método	Estatístico	FFT
Total de instâncias de usuários genuínos:	118	
Total de usuário Verdadeiro/predições Verdadeiro:	115	70
Total de usuário Verdadeiro/predições Falso:	3	48
Total de instâncias de usuários impostores:	5782	
Total de usuário Falso/predições Verdadeiro:	1054	1922
Total de usuário Falso/predições Falso:	4728	3860
False Match Rate:	0.1822898650985818	0.3324109304738845
False Non Match Rate:	0.02542372881355932	0.4067796610169492
Balanced Accuracy:	0.8961432030439295	0.6304047042545832

Tabela 3. Métricas obtidas com uma janela de 7 segundos

Manual		
Janela (segundos)	7	
Kernel	RBF	
C	1	
Método	Estatístico	FFT
Total de instâncias de usuários genuínos:	84	84
Total de usuário Verdadeiro/predições Verdadeiro:	82	53
Total de usuário Verdadeiro/predições Falso:	2	31
Total de instâncias de usuários impostores:	4116	4116
Total de usuário Falso/predições Verdadeiro:	4005	1441
Total de usuário Falso/predições Falso:	3111	2675
False Match Rate:	0.24416909620991253	0.3500971817298348
False Non Match Rate:	0.023809523809523808	0.36904761904761907
Balanced Accuracy:	0.8660106899902819	0.640427599611273
Estatístico		
Janela (segundos)	7	
Kernel	RBF	
C	5	
Método	Estatístico	FFT
Total de instâncias de usuários genuínos:	84	84
Total de usuário Verdadeiro/predições Verdadeiro:	82	56
Total de usuário Verdadeiro/predições Falso:	2	28
Total de instâncias de usuários impostores:	4116	4116
Total de usuário Falso/predições Verdadeiro:	753	1437
Total de usuário Falso/predições Falso:	3363	2679
False Match Rate:	0.18294460641399418	0.3491253644314869
False Non Match Rate:	0.023809523809523808	0.3333333333333333
Balanced Accuracy:	0.896622934888241	0.6587706511175899
FFT		
Janela (segundos)	7	
Kernel	Linear	
C	0.1	
Método	Estatístico	FFT
Total de instâncias de usuários genuínos:	84	84
Total de usuário Verdadeiro/predições Verdadeiro:	84	69
Total de usuário Verdadeiro/predições Falso:	0	15
Total de instâncias de usuários impostores:	4116	4116
Total de usuário Falso/predições Verdadeiro:	1757	2144
Total de usuário Falso/predições Falso:	2359	1972
False Match Rate:	0.4268707482993197	0.5208940719144801
False Non Match Rate:	0.0	0.17857142857142858
Balanced Accuracy:	0.7865646258503401	0.6502672497570456

As maiores variações de valores aconteceram no escopo de janela de 7 segundos. De todos os testes realizados com *grid search*, utilizando como parâmetro de entrada as características no domínio da frequência, foi a única vez que o algoritmo sugeriu a aplicação do SVM com *kernel* linear e $C < 1.0$ como melhor forma de classificação dos dados. Também foi a única vez que, utilizando esta configuração, o classificador alcançou 100% de acerto do usuário genuíno. No caso dos usuários impostores, com esta configuração, o classificador errou 42,69%.

No escopo de 7 segundos também encontramos o melhor resultado de todo o experimento, onde ocorre a maior acurácia balanceada, 0,897, utilizando o *kernel* RBF, com $C = 5$ e características estatísticas. Nesta situação, o classificador errou apenas 2,38% dos usuários genuínos e 18,29% dos usuários impostores.

6. Conclusão

Durante o experimento, conseguimos reimplementar parte dos experimentos conduzidos por [Nickel *et al*, 2011], extraíndo as características no domínio do tempo, por meio de métodos estatísticos dos dados de um acelerômetro integrado a um *smartphone*.

Como forma de verificar a acurácia do algoritmo classificador SVM, adotamos a abordagem sugerida por [Preece *et al*, 2009], aplicando aos dados do acelerômetro, o FFT, a fim de transformá-lo do domínio do tempo para o domínio da frequência.

Após realizada estas transformações de pré-processamento, iniciamos a fase de testes e classificador, janelando os dados em segmentos de 3, 5 e 7 segundos com sobreposição de 50%. O classificador escolhido foi o SVM com *kernel* “RBF”.

Como próximos trabalhos, podemos citar a revisita a este artigo, com o intuito de refazer a experiência, utilizando as coordenadas no sistema mundial. Não suficiente, um estudo mais detalhado dos hiperparâmetros utilizados na SVM se faz necessário a fim de melhor ajustar o classificador, melhorando seu desempenho.

Também como trabalho futuro, é a implantação de outros classificadores de forma a termos um melhor entendimento do desempenho de cada classificador dado determinadas características, extraídas através dos métodos contidos neste artigo. Em especial, seriam o HMM (*Hidden Markov Model*) sugerido por [Nickel *et al*, 2011] e o kNN (*k-Nearest Neighbor*), sugerido por [Preece *et al*, 2009].

Finalmente, a expansão deste experimento a outras bases de dados de acelerômetros, visando encontrar as melhores formas de melhor se extrair as características determinantes das acelerações nos eixos x, y e z obtidas por meio de *smartphones* aliado ao classificador e configuração necessária, diminuindo a complexidade do processo, e aumentando a confiabilidade desta forma de reconhecimento biométrico comportamental contínuo, tornando-o aplicável como uma camada secundária de identificação, aliada a

7. Referências

- SANTOS, Geise. Técnicas para autenticação contínua em dispositivos móveis a partir do modo de caminhar. 2017. 1 recurso online (87 p.). Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Computação, Campinas, SP. Disponível em: <<http://www.repositorio.unicamp.br/handle/REPOSIP/322672>>. Acesso em: 1 set. 2018.
- Ferreira, A., Santos, G., Rocha, A. and Goldenstrein, S. (2017) “User-Centric Coordinates for Applications Leveraging 3-Axis Accelerometer Data”, In: IEEE Sensors Journal, Vol. 17, Nº 16
- Hoang, Thang & Choi, Deokjai. (2014). Secure and Privacy Enhanced Gait Authentication on Smart Phone. TheScientificWorldJournal. 2014. 438254. 10.1155/2014/438254.
- Nyan, M.N. & Tay, Francis & Seah, K.H.W. & Sitoh, Y.Y.. (2006). Classification of gait patterns in the time–frequency domain. Journal of biomechanics. 39. 2647-56. 10.1016/j.jbiomech.2005.08.014.
- Lorena A.C., de Carvalho A.C.P.L.F. (2008) Investigation of Strategies for the Generation

- of Multiclass Support Vector Machines. In: Nguyen N.T., Katarzyniak R. (eds) New Challenges in Applied Intelligence Technologies. Studies in Computational Intelligence, vol 134. Springer, Berlin, Heidelberg
- Hoang T., Choi D., Vo V., Nguyen A., Nguyen T. (2013) A Lightweight Gait Authentication on Mobile Phone Regardless of Installation Error. In: Janczewski L.J., Wolfe H.B., Sheno S. (eds) Security and Privacy Protection in Information Processing Systems. SEC 2013. IFIP Advances in Information and Communication Technology, vol 405. Springer, Berlin, Heidelberg
- Nickel, Claudia & Brandt, Holger & Busch, Christoph. (2011). Benchmarking the performance of SVMs and HMMs for accelerometer-based biometric gait recognition. IEEE International Symposium on Signal Processing and Information Technology, ISSPIT 2011. 281-286. 10.1109/ISSPIT.2011.6151574.
- Preece, Stephen & Yannis Goulermas, John & Kenney, Laurence & Howard, David. (2009). A Comparison of Feature Extraction Methods for the Classification of Dynamic Activities From Accelerometer Data. Biomedical Engineering, IEEE Transactions on. 56. 871 - 879. 10.1109/TBME.2008.2006190.