UNIVERSIDADE DE LISBOA

FACULDADE DE MOTRICIDADE HUMANA

Aprendizagem Automática na Análise do Movimento e da Participação Neuromuscular em Habilidades Dinâmicas

Contributos para o estudo da relação entre Swing do Golf e Prevalência de Lombalgias

Dissertação elaborada com vista à obtenção do Grau de Doutor no ramo de Motricidade Humana na especialidade de Comportamento Motor

Orientador: Professor Doutor Pedro Luís Camecelha de Pezarat Correia

Co-orientadores: Professor Doutor Jan Maria Hendrick Cabri
Professor Doutor Pedro Miguel Picado Carvalho Serranho

Júri:

Presidente

Reitor da Universidade de Lisboa

Vogais

Professor Doutor Jan Maria Hendrick Cabri
Professor Catedrático Convidado
Faculdade de Motricidade Humana da Universidade de Lisboa

Professor Doutor João Paulo Vilas-Boas Soares Campos
Professor Catedrático
Faculdade de Desporto da Universidade do Porto

Professor Doutor Pedro Luís Camecelha de Pezarat Correia
Professor Auxiliar
Faculdade de Motricidade Humana da Universidade de Lisboa

Professor Doutor Pedro Miguel Picado Carvalho Serranho
Professor Auxiliar
Departamento de Ciências de Tecnologia da Universidade Aberta

Professora Doutora Paula Marta Pereira Bruno
Professora Auxiliar
Faculdade de Motricidade Humana da Universidade de Lisboa

Professor Doutor Wangdo Kim
Investigador Auxiliar
Faculdade de Motricidade Humana da Universidade de Lisboa

Luís Miguel Domingues Ferreira Silva

2014

UNIVERSIDADE DE LISBOA

FACULDADE DE MOTRICIDADE HUMANA

Aprendizagem Automática na Análise do Movimento e da Participação Neuromuscular em Habilidades Dinâmicas

Contributos para o estudo da relação entre Swing do Golf e Prevalência de Lombalgias

Dissertação elaborada com vista à obtenção do Grau de Doutor no ramo de Motricidade Humana na especialidade de Comportamento Motor

Orientador: Professor Doutor Pedro Luís Camecelha de Pezarat Correia

Co-orientadores: Professor Doutor Jan Maria Hendrick Cabri
Professor Doutor Pedro Miguel Picado Carvalho Serranho

Júri:

Presidente

Reitor da Universidade de Lisboa

Vogais

Professor Doutor Jan Maria Hendrick Cabri
Professor Catedrático Convidado
Faculdade de Motricidade Humana da Universidade de Lisboa

Professor Doutor João Paulo Vilas-Boas Soares Campos
Professor Catedrático
Faculdade de Desporto da Universidade do Porto

Professor Doutor Pedro Luís Camecelha de Pezarat Correia
Professor Auxiliar
Faculdade de Motricidade Humana da Universidade de Lisboa

Professor Doutor Pedro Miguel Picado Carvalho Serranho
Professor Auxiliar
Departamento de Ciências de Tecnologia da Universidade Aberta

Professora Doutora Paula Marta Pereira Bruno
Professora Auxiliar
Faculdade de Motricidade Humana da Universidade de Lisboa

Professor Doutor Wangdo Kim
Investigador Auxiliar
Faculdade de Motricidade Humana da Universidade de Lisboa

Luís Miguel Domingues Ferreira Silva

2014

Resumo

A quantificação de parâmetros temporais de eletromiografia (EMG) em tarefas dinâmicas bem como a discriminação neuromuscular de constrangimentos motores é paradigmática na literatura. Os principais constrangimentos associados à prática de golfe são o handicap e a sintomatologia de lombalgia. O objetivo desta dissertação foi conhecer até que ponto pode ser extraída informação do sinal EMG que quantifique fenómenos temporais e permita classificar os principais constrangimentos associados a uma habilidade motora dinâmica, propondo-se para esse efeito o *swing* do golfe. Para este propósito construíram-se modelos de classificação, recorrendo às máquinas de vetor suporte (SVM), para deteção do *onset* EMG e para analisar o poder discriminatório de músculos do tronco e do membro inferior nos constrangimentos referidos. Foram realizados cinco estudos, (1) a validação para a população portuguesa do questionário *Musculoskeletal Injury Questionnaire for Senior Golfers*, (2) a comparação de métodos algorítmicos automáticos de deteção (divergindo na linha de base como referência) e inspeção visual, (3) revisão de literatura sistemática qualitativa sobre a aplicação das SVM ao sinal EMG, (4) aplicação da aprendizagem automática com *features* no domínio do tempo na deteção do *onset* EMG e respetiva validação, (5) utilização de *features* da análise de quantificação de recorrência e SVM na identificação dos músculos com maior poder discriminatório e construção de modelos que classifiquem sinais EMG do *swing* relativamente ao handicap e à lombalgia. A deteção do *onset burst* no oblíquo externo (mais vincado do lado direito) e o *onset peak/burst* no reto abdominal bilateralmente apresenta elevada congruência entre métodos. Modelos SVM com quatro tipos de *features* no domínio do tempo garantem uma precisão média de 95% não diferenciando o tipo de fenómeno identificado, isto é, o *onset peak* ou *onset burst*. Caso a diferenciação seja realizada para o *onset burst*, ou para o *onset peak* quando este tende a ser coincidente com o *onset burst*, a precisão sobe para cerca de 98%. A precisão na classificação para o handicap foi de 94.4±2.7% para o *swing*, 97.1±2.3% no *backswing*, e 95.3±2.6% no *downswing*. Na lombalgia a precisão de classificação dos modelos contruídos foi de 96.9±3.8% no *swing* e 99.7±0.4% no *backswing*. Sujeitos com lombalgia apresentam diferentes estratégias de coordenação neuromuscular comparativamente a golfistas assintomáticos.

Palavras-chave: Eletromiografia, máquinas de vetor suporte, *onset*, constrangimentos motores, golfe.

Abstract

The quantification of dynamic tasks with electromyography (EMG) analogously to temporal parameters and neuromuscular constraints discrimination is paradigmatic in the literature. The main constraints associated to golf are the handicap and the low back pain. The aim of this thesis was to understand the extent to which information can be extracted from the EMG signal to quantify temporal phenomena and allows classify the main constraints associated to the golf swing, a dynamic and complex task. For this purpose were built classification models using support vector machines (SVM) for onset detection and to analyze the discriminatory power of the trunk and lower limb muscles on the mentioned constraints. Five studies were held, (1) validation for the Portuguese population of the Musculoskeletal Injury Questionnaire for Senior Golfers, (2) comparison of detection threshold algorithmics (diverging at baseline as a reference) and visual inspection, (3) systematic qualitative review about SVM application on EMG, (4) application of time domain features within machine learning for onset detection and model respective validation, (5) use of recurrence quantification analysis features within SVM to identify muscles with greater discriminatory power on handicap and low back pain. The external oblique onset burst detection (more pronounced on the right side) and bilaterally rectus abdominis onset peak/burst detection showed a high congruence between all methods. SVM models with four types of time-domain features ensure an average accuracy of 95% without differentiating the type of detection. When onset type was differentiated, onset burst and onset peak accuracy increased to 98%. The accuracy rating for the handicap was 94.4 ± 2.7% in the swing, 97.1 ± 2.3% in the backswing, and 95.3 ± 2.6% in the downswing. Low back pain accuracy of the SVM models was 96.9 ± 3.8% during the swing, and 99.7±0.4% in the backswing. Subjects with low back pain have different neuromuscular coordination strategies compared to asymptomatic golfers.

Keywords: Electromyography, support vector machines, onset, motor constraints, golf.

Agradecimentos

Ao Professor Doutor Pedro Pezarat Correia por dar significado à palavra *Professor*. Tem sido orientador para além desta dissertação tanto pela sua disponibilidade, preocupação, ensino e amizade.

Ao Professor Doutor Jan Cabri por ter cumprido a promessa de bastar aparecer no seu gabinete caso houvesse um problema, dando a solução. Obrigado por toda a aprendizagem ao mais elevado nível.

Ao Professor Doutor Pedro Serranho por ter aceite este desafio de orientação e igualmente ter contribuído para o meu percurso académico, respondendo à gratidão com "*faz parte do meu trabalho*".

À Fundação para a Ciência e Tecnologia por me ter concedido bolsa individual de doutoramento possibilitando este percurso académico.

Aos colegas do Laboratório de Comportamento Motor da Faculdade de Motricidade Humana tidos como o "grupo do golfe", Dr. João Vaz, Professora Doutora Maria António Castro, Professor Doutor Orlando Fernandes, Mestre Sérgio Marta e o Gustavo Reinaldo.

Aos docentes do Laboratório de Comportamento Motor da Faculdade de Motricidade Humana por todo apoio, preocupação e interesse no trabalho desenvolvido.

Ao Centro Interdisciplinar de Estudo da Performance Humana (CIPER) - Faculdade de Motricidade Humana da Universidade de Lisboa por terem sido a entidade acolhedora no decorrer dos trabalhos de doutoramento.

Aos golfistas que disponibilizaram o seu tempo deslocando-se ao laboratório de Comportamento Motor da Faculdade de Motricidade Humana, sem os quais este trabalho não poderia ser desenvolvido. Sujeitaram-se a recolhas que são por si morosas e em condições laboratoriais exigentes.

Á Professora Doutora Filomena Carnide pelos preciosos contributos sobre a validação de questionários.

Ao Professor Doutor Joaquim Silveira Sérgio pelas avaliações que não delegou para que tivesse mais tempo de dedicação a esta dissertação.

Ao Professor David Lindsay pelos contributos no conhecimento de lesões e outros constrangimentos associados ao golfe.

Ao Professor Doutor Norman Marwan pela disponibilidade e aconselhamento na troca de *emails* de comportamento não linear.

Aos serviços académicos e bibliotecário por todo o cuidado e disponibilidade que apresentaram, tendo sido igualmente um grandioso apoio.

Aos meus pais por terem sido pais primando por promover a "outra" educação… a dos valores.

À minha família. À minha filha Catarina. À Vilma pelo companheirismo e tudo o mais…

Deo Gratias

ÍNDICE

Índice de Tabelas

Índice de Figuras

Definições Operacionais

Feature – Termo atribuído a determinada entrada característica de um atributo que irá constituir um vetor. A *feature* é cada uma das componentes do vetor. Por sua vez, cada vetor é associado como pertencendo a cada classe.

Onset – Termo geral que refere o início de ativação do sinal EMG, ou seja, corresponde ao momento em que o músculo ativa.

Onset burst – Definição tradicional de *onset* da ativação neuromuscular, corresponde ao momento em que o músculo ativa após um período de repouso. Significa a primeira propagação dos potenciais de ação no registo EMG, independentemente de ocorrer posteriormente um pico de atividade muscular bem definido ou não.

Onset peak – Corresponde a qualquer início de ativação no registo EMG que antecede um período relevante de atividade, independentemente de anteriormente existir atividade muscular como pré-ativação. Pode corresponder a um instante *off/ on* seguido de um pico de atividade muscular, ou ao próprio *onset burst*. Neste último caso, quando a atividade que antecede o *onset* é praticamente nula existindo posteriormente um pico de atividade bem definido.

Lista de Abreviaturas

BF – *Biceps Femoris* (Bicípite Femoral).

CFS – *Correlation-based Feature Selection*.

DASDV – *Difference Absolute Standard Deviation Value* (Diferença Absoluta do Desvio – Padrão).

EMG – *Electromyography* (Eletromiografia).

FS – *Fisher Score*.

GM – *Gluteus Maximus* (Grande Glúteo).

Hc – Handicap.

HHc – *High Handicap* (Handicap Alto).

IEMG – *Integraded EMG* (Integral do EMG).

LBP – *Low Back Pain* (Lombalgia).

LHc – *Low Handicap* (Handicap Baixo).

LOG – *Logarithmic Detector* (Detetor Logarítmico).

MAV – M*ean Absolute Value* (Valor Absoluto Médio).

MUAP – *Motor Unit Action Potential* (Potencial de Ação das Unidades Motoras).

NLBP – *No Low Back Pain* (Sem Lombalgia).

ES – *Erector Spinae* (Massa Comum).

EO – *External Oblique* (Oblíquo Externo).

RA – *Rectus Abdominis* (Reto Abdominal).

RBF – *Radial Basis Function* (Função Base Radial).

RF – *Rectus Femoris* (Reto Femoral do Quadricípite).

RQA – *Recurrence Quantification Analysis* (Análise de Quantificação de Recorrência). Abreviaturas das respetivas variáveis no Capítulo II (metodologia) e no Capítulo VII.

ST – *Semitendinosus* (Semitendinoso).

SVM – *Support Vector Machines* (Máquinas de Vetor Suporte).

VAR – *Variance of EMG* (Variância do EMG).

VE – *Vastus Lateralis* (Vasto Externo do Quadricípite).

VI – *Vastus Medialis* (Vasto Interno do Quadricípite).

WL – *Waveform Length* (Comprimento do Formato da Onda).

*Outras abreviaturas serão apenas consideradas nos próprios artigos como a revisão de literatura do Capítulo V e legendas de tabelas e figuras.

INTRODUÇÃO

A eletromiografia (EMG) refere-se ao registo da atividade neuromuscular, permitindo estudar estratégias de coordenação em relação a determinado movimento. Ao longo desta dissertação serão realçados alguns paradigmas de investigação que serviram de sustentação à pesquisa desenvolvida, conduzindo posteriormente à construção de modelos de aprendizagem automática. O primeiro paradigma refere-se à deteção de parâmetros temporais no sinal EMG, nomeadamente do *onset*. Na análise deste tipo de variáveis encontramos a necessidade de aplicar algoritmos automáticos ao invés de inspeção visual, sendo esta considerada subjetiva, mas a performance desses algoritmos é aferida recorrendo a inspeção visual. A falta de reprodutibilidade entre métodos e a adaptação destes às características do sinal, como o rácio sinal-ruído e o declive do incremento da amplitude do sinal EMG, conduz a que o *onset* estudado em diferentes pesquisas possa não ser exatamente o mesmo. Esta dificuldade está bem presente no estado da arte que será apresentado nesta tese, pois a grande incidência de estudos incide sobre a problemática metodológica, existindo carência de investigação que procure, dentro do possível, quantificar a coordenação intra e intermuscular no tempo, ou seja, durante a execução motora. Devido a essa incidência metodológica e pela presença de rácios sinal-ruído mais elevados, a literatura tende a privilegiar contrações isométricas em detrimento de ações dinâmicas. Porém, pretende-se inferir para estas a metodologia aferida em contrações isométricas.

O *swing* do golfe é uma habilidade motora dinâmica complexa combinando movimento de precisão e velocidade de execução. Em habilidades motoras como o *swing* torna-se fundamental descodificar estratégias neuromusculares que expliquem a forma como o sistema neuromuscular se coordena e como essas estratégias podem ser discriminadas face a constrangimentos associados à prática desportiva. Estas lacunas na discriminação de constrangimentos em tarefas motoras dinâmicas complexas e os paradigmas encontrados nos algoritmos automáticos e inspeção visual conduziram a um pensamento: seria ideal o pesquisador interferir em determinada altura do processo para que a máquina (computador) posteriormente conhecesse o fenómeno fisiológico no sinal EMG a classificar. Este pensamento era nada mais do que uma definição da aprendizagem automática supervisionada, e daí a inclusão das máquinas de vetor suporte (SVM) nesta tese.

A contenda que motiva esta dissertação é conhecer *até que ponto pode ser extraída informação do sinal EMG que quantifique fenómenos temporais e permita classificar os principais constrangimentos associados a uma habilidade motora dinâmica.* Para o presente propósito a habilidade motora refere-se ao *swing* do golfe e os constrangimentos associados a esta modalidade são o handicap e a lombalgia. Como será possível verificar no decorrer desta dissertação, a lombalgia e o handicap são os principais constrangimentos associados à prática do golfe.

Após esta introdução, a dissertação está organizada em nove capítulos: o enquadramento teórico, o método, cinco capítulos constituídos pelos artigos realizados, uma discussão geral e as respetivas conclusões e recomendações. Em relação aos artigos, a opção foi de não colocar a publicação original no corpo da dissertação. O intuito foi o de uniformizar a configuração dos textos e permitir que os vários pontos abordados, tabelas e figuras pudessem constar nos índices facilitando assim a sua consulta.

No enquadramento teórico, Capítulo I, é realizada uma breve introdução quanto a fatores que devem estar presentes no manuseamento de EMG, nomeadamente quando se pretende informação de cada músculo com elevado teor discriminativo. Esta preocupação deve ser salientada visto procurar-se quantificar partes do sinal contendo informação na deteção do *onset* e na identificação de músculos que em determinadas fases do movimento possam estar mais associados a alterações na coordenação motora face aos mencionados constrangimentos. Prossegue-se com o estudo de algoritmos na deteção de parâmetros temporais, dos constrangimentos considerados em relação ao *swing* do golfe, entrando então no conceito de aprendizagem automática intimamente ligado à Teoria Estatística da Aprendizagem (TEA). Desta área científica realiza-se o encaminhamento para o instrumento matemático potentíssimo no reconhecimento de padrões, SVM. Termina-se numa alusão à aplicação desta técnica à EMG que será aprofundada no Capítulo VI.

O Capítulo II refere-se à metodologia, no qual se procura generalizar a mesma, face aos vários artigos, como também, explicar as técnicas utilizadas de uma forma mais exaustiva sem a limitação inerente aos artigos devido ao do número de palavras aceite para publicação. As SVM serão abordadas mais extensivamente.

O Capítulo III denomina-se por *Cross-cultural adaptation and validation of the Portuguese Survey of musculoskeletal conditions, playing characteristics and warm-up patterns of golfers*, sendo o primeiro artigo realizado. Este artigo acaba por servir dois propósitos. Para além de realizar o propósito de validação de um instrumento que permita categorizar a população golfista, apresenta os principais constrangimentos associados à prática de golfe, quer pelos dados obtidos como pelo levantamento de literatura. Os principais constrangimentos para a prática são o nível técnico ou handicap (Hc) e a incidência de lombalgia. Foram estes os constrangimentos tidos em consideração para agrupar os sinais EMG e verificar quais os músculos com poder discriminatório.

Prossegue-se no Capítulo IV com o segundo estudo, no qual é utilizado um método tradicional automático, um algoritmo de limiar que permite comparar dois métodos tendo em conta um dos principais fatores que prejudica a deteção do *onset*, a linha de base. Os dois métodos são comparados com inspeção visual sendo identificados momentos de relevância no registo EMG nos músculos do tronco durante o *swing*.

No Capítulo V surge então a aprendizagem automática propriamente dita, pelo que é realizada uma revisão de literatura exaustiva sobre a utilização de SVM no sinal EMG. Para além de permitir conhecer desenhos de investigação na literatura, fornece conhecimento sobre a metodologia, sobretudo a segmentação do sinal e a extração de *features* (características que são introduzidas no classificador).

O estudo 4 dá corpo ao Capítulo VI, explorando a aplicação das SVM na classificação do sinal EMG para determinar o *onset*. A esta altura a definição de *onset* é averiguada como o fenómeno mais relevante face à análise desenvolvida no segundo estudo com metodologia tida como tradicional. Neste capítulo apresenta-se a aprendizagem automática como método na deteção do *onset*.

O Capitulo VII é composto pelo quinto artigo, no qual através da extração de características da análise de recorrência do sinal EMG procura-se identificar os músculos com maior poder discriminatório mediante os constrangimentos à tarefa, isto é, handicap e lombalgia. São desenvolvidos modelos que permitem classificar sinais como oriundos de sujeitos de baixo e alto handicap, com e sem sintomatologia de lombalgia.

Uma discussão geral e posteriormente as conclusões e recomendações para estudos futuros findam esta dissertação, com os Capítulos VIII e IX, respetivamente.

CAPÍTULO I

1. Enquadramento Teórico

1.1. EMG e o Estudo da Função Neuromuscular

No estudo da coordenação neuromuscular procura-se determinar como o Sistema Nervoso Central (SNC) estrutura e controla os movimentos do corpo. Por sua vez, os vários segmentos corporais movem-se pela capacidade dos músculos produzirem força, o que depende dos impulsos nervosos traduzidos pelo recrutamento de unidades motoras e respetiva frequência de descarga. Os potenciais de ação na membrana da fibra muscular, o sarcolema, propagam-se até à superfície da pele pelos tecidos envolventes. Desta forma, torna-se possível o registo dessa atividade elétrica, representando o somatório dos potenciais de ação das fibras recrutadas em determinada área, consoante a zona onde foram colocados os elétrodos. Quando a sua deteção é realizada à superfície da pele, como no presente estudo, designa-se por eletromiografia de superfície.

O sinal EMG contém informação que pode ser quantificada segundo três tipos de parâmetros fundamentais: a estrutura temporal, a amplitude, e a frequência (Pezarat & Mil-Homens, 2004). Esta dissertação incidiu essencialmente no domínio do tempo, procurando quantificar e desenvolver procedimentos que retirem informação sobre a atividade neuromuscular durante habilidades motoras dinâmicas, neste caso o *swing* do golfe. A amplitude do sinal de EMG é estocástica podendo ser representada por uma distribuição Gaussiana. Esta amplitude do sinal pode variar de 0 a 10 mV (pico-a-pico) ou 0 a 1.5 mV (RMS). As frequências do sinal EMG de superfície situam-se entre 0 e 500 Hz com elevado domínio entre os 50 e 150 Hz (De Luca, 1997). Os principais mecanismos que influenciam a magnitude e a densidade do sinal são o recrutamento de potenciais de ação das unidades motoras e a frequência de disparo. Assume-se nesses mecanismos as estratégias motoras fundamentais que ajustam o processo de contração e modelam a produção de força (Konrad, 2005).

Mediante determinada problemática no comportamento motor, interessa aplicar o melhor método que garanta e preserve a informação contida no sinal EMG. Ao pretender-se categorizar problemáticas recorrentes da EMG, para simplificar, podemos enumerar duas grandes necessidades: (1) o minimizar das limitações que estão associadas à EMG, e (2) a capacidade de quantificar a informação contida no sinal. Primeiro, torna-se necessário controlar os fatores que influenciam o sinal EMG. Por exemplo, detetar ajustamentos posturais antecipatórios na ativação de músculos do tronco pode ser problemático devido ao ruído na linha de base e presença de artefactos de eletrocardiograma (Allison, 2003; Lee, Cholewicki, & Reeves, 2007). Depois coloca-se a aplicação de métodos que permitam a quantificação daquilo que realmente se pretende medir. Por vezes, os parâmetros dos próprios algoritmos são por si influenciadores de resultados e comprometem a reprodutibilidade entre estudos (Vaisman, Zariffa, & Popovic, 2010).

Grosseiramente, podemos referir que as abordagens tradicionais passam por responder numa dimensão qualitativa e quantitativa à problemática da quantificação (Konrad, 2005). A dimensão qualitativa pretende meramente perceber se o músculo está ativo e/ou a que proporção corresponde essa atividade. Focando a análise a um nível quantitativo procura-se conhecer quando é que o músculo está ativo, quanto é que o músculo está ativo e/ou se existe fadiga muscular. Esta relação quantitativa está intimamente ligada aos domínios expressos anteriormente, tempo, amplitude e frequência, respetivamente. Podem ainda ser acrescentados outros domínios como por exemplo, tempo-frequência (*wavelet*) ou abordagens que ultrapassem a análise estrutural chegando a uma incidência fenomenológica (Oskoei & Hu, 2007).

Ao conceito de quantificar a que se propõe a presente dissertação, juntam-se o discriminar e o classificar. Estas tarefas incidem no domínio do tempo e na análise do comportamento não-linear do sinal, tendo de se debater com limitações que foram encontradas nos métodos ditos tradicionais. Veja-se um simples exemplo de quantificação no estudo da intensidade. Suponhamos dois sinais aos quais pretende-se quantificar uma fase de movimento. Num dos sinais o sujeito produziu uma ativação EMG de 25% normalizado pelo seu EMG máximo até metade do período decorrido nessa janela temporal e 75% na restante metade. A média dessa janela de amostras é de 50%. Noutro sinal o sujeito produziu nas mesmas fases respetivamente 45% e 55% correspondendo também a um valor médio de 50%. A interpretação é que a atividade desse músculo nas duas situações durante essa fase foi similar, quando o comportamento do sinal é deveras diferente. Ao inferir estatisticamente para esses valores médios de intensidade está a ser desprezada informação que pode ser relevante. Quando se pretende registar diferentes estratégias neuromusculares durante a realização de uma determinada habilidade motora, estando presentes certos constrangimentos, como retirar informação consentânea do registo EMG?

Duas abordagens costumam estar presentes no estudo das relações entre o registo de EMG de superfície e as propriedades do sistema neuromuscular (Farina, Fattorini, Felici, & Filligoi, 2002): *forward* e *inverse*. A abordagem *forward* pode ser conseguida através de modelagem, permitindo prever o efeito de vários processos fisiológicos face às características do EMG de superfície. A abordagem *inverse* utiliza o registo EMG para identificar o fenómeno fisiológico que lhe está subjacente. Porém, esta última requer simplificações para reduzir o número de parâmetros e múltiplas soluções que influenciam a associação entre o sinal e a interpretação de estratégias de controlo motor. Por esse motivo, uma limitação dos modelos *inverse* deve-se ao facto das aproximações realizadas serem meramente válidas para as condições específicas que estão a ser estudadas, não permitindo a sua generalização. Sendo o sinal de EMG uma série temporal, requer métodos matemáticos que sejam capazes de detetar as suas características, sejam estas não-lineares, estocásticas, não estacionárias e determinísticas (Lei & Meng, 2012). Assim, são necessários métodos multivariados e não-lineares, ao invés de abordagens lineares que muitas vezes não são apropriadas (Marwan, Thiel, & Nowaczyk, 2002; Tolambiya, Thomas, Chiovetto, Berret, & Pozzo, 2011). Dois fenómenos aparentemente semelhantes podem conter informações sobre a sua variabilidade que permite distinguir características populacionais face a constrangimentos das mesmas.

Um ponto de partida a ser considerado refere-se ao tipo de contração (se estática ou dinâmica), pois deparamo-nos com uma situação paradigmática na abordagem no domínio do tempo, nomeadamente, em relação ao início de ativação muscular denominado por *onset*. Quando os algoritmos de deteção são testados usam-se sinais de contrações isométricas (ex.: Jöllenbeck, 2000; Solnik, Rider, & Steinweg, 2010), mas depois pretende-se inferir para habilidades motoras dinâmicas (ex.: Cole & Grimshaw, 2008a; Horton, Lindsay, & Macintosh, 2001), quando um dos fatores que influencia a prestação de algoritmos no estudo de parâmetros temporais é a linha de base (Hodges & Bui, 1996). A escolha de contrações isométricas prende-se com duas razões principais. A primeira refere-se exatamente à atividade na linha de base, pois será de esperar que o rácio sinal-ruído seja maior. O segundo aspeto é uma limitação das recolhas de eletromiografia de superfície, o movimento relativo do ventre muscular debaixo da zona onde se encontram os elétrodos e o próprio movimento da pele (Konrad, 2005). Devido a esta limitação, em estudos com contrações dinâmicas é importante que a localização dos elétrodos seja numa posição central em relação ao ventre muscular ou na zona de maior proeminência quando realizada uma contração muscular.

1.1.1. Fatores que influenciam o sinal EMG

A eletromiografia de superfície tem como grandes vantagens a facilidade de utilização, a caraterística de não ser um método intrusivo, e a capacidade de fornecer uma informação generalizada do comportamento do músculo. Porém, se por um lado apresenta estas vantagens, por outro, corresponde igualmente a um sério risco, pois *is too easy to use and consequently too easy to abuse* (De Luca, 1997, p.135), sendo necessário remover potenciais variáveis de influência para além das que se pretendem estudar. As condições experimentais de recolha são fonte de fatores que influenciam a qualidade do sinal, podendo alterar totalmente a informação contida nos dados, podendo conduzir a interpretações incorretas. As questões ambientais que levam à aquisição de ruído pelo registo EMG durante a propagação do sinal podem ser categorizadas da seguinte forma (Raez, Hussain, & Mohd-Yasin, 2006):

a) Ruído inerente ao equipamento eletrónico: este tipo de ruído não pode ser eliminado, mas pode ser reduzido pela utilização de componentes eletrónicos de elevada qualidade;

b) Ruído ambiente: deve-se à radiação eletromagnética à qual estamos expostos continuamente. O ruído ambiente pode ter amplitude superior ao sinal EMG numa ordem de um a três;

c) Artefactos mecânicos: a introdução de artefactos mecânicos leva a distorções do sinal, podendo ser produzidos na 1) zona de contacto do elétrodo e 2) nos cabos que ligam os elétrodos ao restante equipamento;

d) Ruído inerente à instabilidade de sinal: a amplitude do EMG é de natureza aleatória, sendo influenciada pela capacidade de disparo das unidades motoras.

> Na sua maioria, este tipo de ruído, tal como os artefactos mecânicos, situa-se dentro de uma frequência de 0 a 20 Hz, devendo ser removido.

O primeiro passo que o investigador deve privilegiar refere-se aos procedimentos de recolha, no sentido de evitar comprometer logo à partida a qualidade do sinal, pois as fontes de ruído são um conjunto de fatores que podem influenciar o sinal conduzindo a amostras deturpadas. Podemos classificar estes fatores em fisiológicos e não fisiológicos (Farina, Merletti, & Enoka, 2004), com as respetivas categorias (Figura 1). Os fatores fisiológicos estão relacionados com as propriedades da unidade motora e do sarcoplasma na condutividade do potencial de ação. Os fatores não fisiológicos são aqueles que têm de ser considerados logo na recolha. Estes compreendem questões morfológicas dos sujeitos e características do equipamento de deteção, assim como a relação entre ambos, como a colocação dos elétrodos. Um dos fatores que poderá exercer grande influência em estudos cujo objetivo é a discriminação por músculos é o *crosstalk*, um fator físico. Ao querer identificar a atividade de determinado músculo em relação a um constrangimento, o registo deste pode estar contaminado por atividade mio-elétrica de músculos vizinhos. Várias propostas têm sido apresentadas para quantificar o *crosstalk*, nomeadamente, a função de correlação cruzada entre dois sinais (Farmer, Halliday, Conway, Stephens, & Rosenberg, 1997; Mogk & Keir, 2003; Winter, Fuglevand, & Archer, 1994) e a quantificação pelo potencial de ação composto reflexo (reflexo H) (Mezzarane & Kohn, 2009).

A equação de correlação cruzada entre dois sinais $x(t)$ e $y(t)$ é dada por (Mogk & Keir, 2003; Winter et al., 1994):

$$R_{x,y}(\tau) = \frac{1}{T}\int_{0}^{T} x(t)\, y(t+\tau)dt$$

onde T corresponde ao comprimento dos sinais a serem correlacionados e τ ao desfasamento temporal. O índice de correlação que se situa entre os valores -1 e $+1$ resulta da normalização da equação:

$$R'_{x,y}(\tau) = \frac{R_{xy}(\tau)}{\sqrt{R_{xx}(0)R_{yy}(0)}}$$

em que $R_{xx}(0)$ e $R_{yy}(0)$ representam a auto-correlação de $x(t)$ e $y(t)$, respetivamente, quando $\tau = 0$. Com este método verifica-se que o sinal pode ser contaminado por informação vizinha dependendo do distanciamento entre elétrodos em 22-24% para 2.5cm, 4-7% para 5cm, e 1-2% quando a distância é de 7.5cm no membro inferior (Winter et al., 1994) e cerca de 40% com 3cm, 10% com 6cm, e 2.5% quando colocados a 9cm em músculos do antebraço na ação de preensão (Mogk & Keir, 2003).

O reflexo H é conseguido pela estimulação elétrica percutânea que ativa os axônios aferentes tipo Ia que inervam o músculo que poderá ser fonte de influência, induzindo uma resposta reflexa. O reflexo H é refere-se à atividade elétrica resultante do reflexo de estiramento no músculo alvo sendo seguido por um breve período de ausência de atividade EMG (Mezzarane & Kohn, 2009). A estimulação elétrica surge para contrariar as limitações de métodos que utilizam contrações voluntárias nas quais a dificuldade de separar a atividade isolada de músculos está presente, como a comparação entre dois sinais. Esta comparação tenderá a ser facilitada quando os músculos em causa são antagonistas, quer por contração voluntária ou estimulação elétrica.

A importância de quantificar ou controlar o *crosstalk* também está relacionada com o comprometimento na aplicação de métodos lineares e com a interpretação do comportamento (não linear) de determinado músculo, devido à capacidade destas interferências criarem por si alterações de carácter não linear (Choi & Kim, 2011). A redução do *crosstalk* também pode ser obtida com redução da distância entre os sensores dos elétrodos ou com outra configuração (*double differential technique)* (De Luca, Kuznetsov, Gilmore, & Roy, 2012).

A proximidade dos elétrodos da zona mio-tendinosa, das inserções, e os músculos superficiais que estão perto de regiões com grande predominância tendinosa também aumenta o risco de *crosstalk* (Merlo & Campanini, 2010). Ainda, com a agravante, de por vezes, essa interferência gerada não poder ser removida, quer através da distância entre elétrodos (inter-elétrodos e/ou inter-sensores dos elétrodos) ou recorrendo a filtragem do sinal.

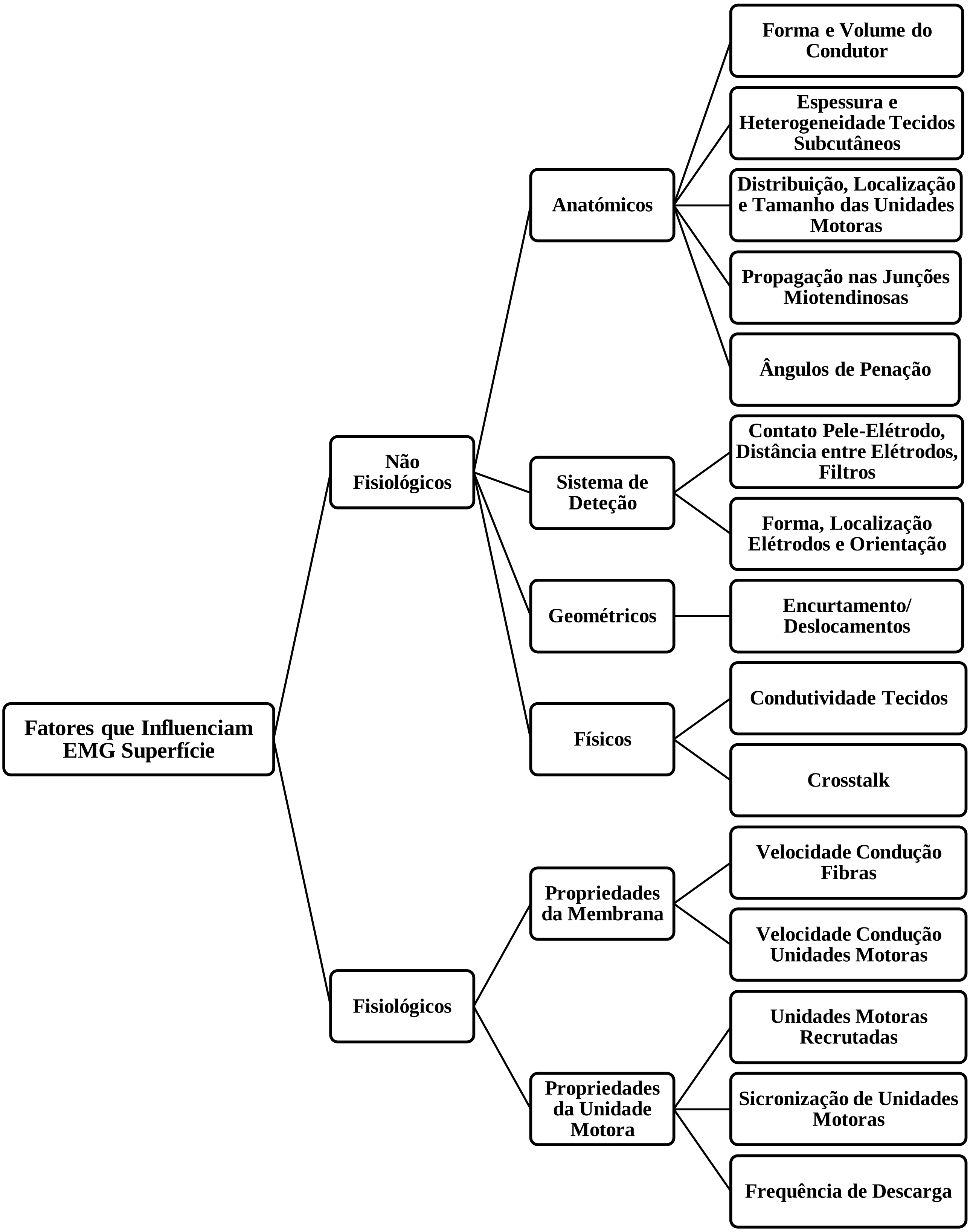


Figura 1 – Fatores que influenciam o EMG de superfície (adaptado Farina et al., 2004, p.1487).

1.1.2. Análise de parâmetros temporais em EMG

Foi a partir dos paradigmas de investigação que estão associados ao estudo dos parâmetros temporais do sinal EMG que se desenvolveu o presente trabalho de investigação. Quando se inicia qualquer estudo surge a questão sobre qual o melhor método, ou método estandardizado, em função da variável em causa. Apesar de serem identificadas como principais variáveis temporais o início de ativação (*onset*), o instante em que ocorre o pico máximo de ativação e o términus da ativação (*offset*), o primeiro tende a ser o parâmetro temporal mais estudado. Porém, apesar de existirem várias propostas para a sua determinação, as mesmas comprometem a reprodutibilidade entre estudos (Morey-Klapsing et al., 2004). Na realidade são duas as dimensões onde podem ser enquadrados os vários métodos: a inspeção visual e algoritmos de deteção (Vaisman et al., 2010).

A inspeção visual, como a própria designação indica, refere-se à determinação por parte do investigador, que ao visualizar o registo EMG marca o ponto de interesse. O problema que se coloca com este método é o tempo que o próprio requer e a subjetividade inerente ao investigador e sensibilidade para o parâmetro temporal (Jöllenbeck, 2000).

Os métodos baseados em algoritmos dividem-se em várias categorias, sendo mais usuais métodos por limiar e por detetores baseados na decisão estatística ótima. Os primeiros encontram-se na literatura com a aplicação de um limiar para a deteção do *onset* que pode ser obtido através de intervalos de confiança (Van Boxtel, Geraars, Van Den Berg-Lenssen, & Brunia, 1993), com um determinado valor de desvio-padrão adicionado a uma média do sinal (Allison, 2003; Cole & Grimshaw, 2008a; Hodges & Bui, 1996; Horton et al., 2001; Silva et al., 2013), ou por limiar fixo percentual relativo ao pico EMG máximo (Konrad, 2005). A estes algoritmos ainda tende a ser associada uma janela temporal de amostras durante a qual a intensidade do sinal se deverá manter acima do valor limiar (Hodges & Bui, 1996; Silva et al., 2013). Os detetores baseados na decisão estatística ótima referem-se essencialmente ao teste do rácio da máxima verossimilhança (*maximum likelihood ratio test*), que partem do princípio que caso o processo de geração do sinal for (parcialmente) conhecido, podem ser utilizados detetores baseados em regras de decisão estatística (Micera et al., 1998; Staude, Flachenecker, Daumer, & Wolf, 2001). Estes métodos permitem avaliar as propriedades estatísticas do sinal EMG antes e depois de uma possível mudança nos parâmetros do modelo. Recorrem a um filtro que retire a informação irrelevante o sinal y_k com $k \geq 1$ é um processo estatístico Gaussiano, aleatório e independente cujas propriedades podem ser descritas pela sua função densidade probabilidade:

$$p_\sigma(y_k) = \frac{1}{\sqrt{2\pi\sigma^2(k)}} e^{-y_k^2/2\sigma^2(k)}$$

a qual depende perfil variância $\sigma^2(k)$. A regra de decisão do modelo dependerá do prévio conhecimento sobre o perfil da variância $\sigma_0^2(k)$ e $\sigma_1^2(k, t_0)$ antes e depois da alteração ser detetada (Staude et al., 2001). Derivações deste método podem estimar a máxima verossimilhança estimando um ponto temporal desconhecido, mas tendo em conta que os restantes parâmetros são totalmente conhecidos (*Optimal estimator*), ou tendo em conta que os perfis de variância $\sigma_0^2(k)$ e $\sigma_1^2(k, t)$ são desconhecidos podendo ser estimados (*Approximated generalized likelihood-ratio detectors*). Porém, também estes autores compararam estes algoritmos com os demais por inspeção visual, e a rotina estabelecida solicita a prévia definição da janela temporal do sinal a analisar.

Devido à subjetividade inerente ao processo de inspeção visual, os estudos não devem ser realizados usando meramente este método, mas sim recorrendo a algoritmos matemáticos. Por outro lado, quando aplicado um algoritmo matemático, para que seja confiável o seu resultado, torna-se fundamental realizar inspeção visual. Acentuando mais ainda esta relação, a suposta validação de algoritmos matemáticos foi sempre realizada por recurso a inspeção visual (Solnik et al., 2010; Staude et al., 2001; Vaisman et al., 2010). Será de perguntar, afinal onde ficamos? De maior relevância coloca-se a descrição do fenómeno fisiológico a estudar, pois podem existir pré-ativações que não correspondem a nenhum padrão em particular ou que sejam redundantes para uma determinada característica a estudar. Igualmente, pode existir outra informação no tempo que estará a ser desprezada. Um estudo em que está patente a necessidade de descrever o fenómeno biológico ao longo da tarefa motora foi realizado por McGill, Chaimberg, Frost e Fenwick (2010). Neste estudo, o registo EMG é associado à fotografia de movimentos de artes marciais, sendo descritas as ações relevantes correspondentes. Não deixa porém, de estar expressa a dificuldade inerente em quantificar essa atividade muscular face a habilidades dinâmicas (McGill et al., 2010; Silva et al., 2013; Tyler & Karst, 2004).

Das características apontadas como aquelas que criam dificuldades aos métodos algorítmicos podem-se destacar o rácio sinal-ruído (Hodges & Bui, 1996; Solnik et al., 2010) e o declive com que o sinal cresce em amplitude quando em atividade (Allison, 2003). Um sinal tido como ideal para a determinação do *onset* teria uma linha de base com uma variância muito reduzida e um declive de crescimento da amplitude muito acentuado (quanto mais perpendicular à linha de base melhor). Desta forma, com diferentes limiares seria realizada a mesma deteção do *onset*. Seguindo esta lógica, encontramos propostas suplementares à deteção deste parâmetro temporal como o operador de energia Teager–Kaiser (*Teager–Kaiser energy operator*) (Li & Aruin, 2005; Solnik et al., 2010) e deteção pela técnica de alteração de momentos baseada em análise espectral singular (*change-point detection technique based on singular spectrum analysis*) (Vaisman et al., 2010).

O operador de energia Teager–Kaiser é um método não linear que permite estimar a frequência e a amplitude do sinal através de modelação AM e FM. No domínio do tempo o operador de energia Teager–Kaiser Ψ discreto pode ser definido por:

$$\Psi[x(n)] = x^2(n) - x(n+1)x(n-1)$$

onde x é o valor EMG e n o número da amostra.

A análise espectral singular é uma abordagem não paramétrica na qual é realizada uma decomposição da matriz de trajetória. Esta decomposição corresponde aos auto-vetores da matriz de desfasamento das covariâncias, sendo indicada a estrutura que melhor representa os dados subjacentes e qual a estrutura que reflete o ruído. Neste caso não é utilizado nenhum limiar, o *onset* é definido como o primeiro ponto com valor diferente de zero antes que a estatística CUSUM atinja o seu máximo. Esta técnica não requer conhecimento prévio do sinal EMG, sendo aplicada ao sinal em bruto.

Podendo ser realizadas várias análises coloca-se a questão sobre o tratamento do sinal. Usualmente, a definição tradicional de *onset* refere-se ao início da onda de atividade principal denominada por *onset burst* (Solnik et al., 2010). Por si só, outros fenómenos fisiológicos que tenham alguma relevância na execução da tarefa motora são desprezados. Outra questão a realçar refere-se aos sinais usados para análise da performance de algoritmos de deteção. Usualmente, trata-se de contrações isométricas em detrimento de contrações dinâmicas (ex.: Jöllenbeck, 2000; Li & Aruin, 2005; Solnik et al., 2010; Staude et al., 2001). Um dos motivos poderia estar associado ao deslocamento dos elétrodos através da pele que se verifica ao longo de uma ação dinâmica, alterando a área de deteção. Porém, este aspeto apresenta maior influência na amplitude do sinal comparativamente aos parâmetros temporais. Inferindo de ações isométricas para ações dinâmicas também não leva em consideração a coordenação intramuscular como a inibição seletiva.

No caso da habilidade motora analisada neste estudo, o *swing* do golfe, no melhor do nosso conhecimento, realçam-se dois estudos (Cole & Grimshaw, 2008a; Horton et al., 2001), para além daquele que é parte integrante nesta dissertação (Silva et al., 2013). Nos dois primeiros estudos, a análise temporal centra-se nos principais fatores que influenciam a execução do *swing* do golfe, o handicap (Hc) e a lombalgia.

1.2. Constrangimentos Associados à Habilidade Motora: *Swing*

Como constrangimento consideramos qualquer fator que possa influenciar ou ser influenciado quanto à prática da modalidade, nomeadamente, em relação à execução da habilidade motora em estudo. O golfe tem como objetivo colocar a bola num buraco de pequeno diâmetro com o menor número de pancadas possível. Para o golfista conseguir atingir este objetivo recorre a duas habilidades motoras: o *swing* e o *putting* (Hume, Keogh, & Reid, 2005). Na presente dissertação, a habilidade motora considerada para estudo corresponde ao *swing* do golfe, uma habilidade motora complexa que combina precisão e velocidade de execução. Interessava então conhecer os principais constrangimentos associados a esta habilidade motora, partindo de duas considerações. A primeira refere-se ao nível técnico, denominado por Hc, sendo o golfista classificado em conformidade com o Sistema de Handicap Europeu da *European Golf Association* (EGA, 2012). A segunda consideração aponta o *swing* como a tarefa motora responsável pela maioria das lesões relacionadas com o golfe (Cabri, Sousa, & Barreiros, 2009; McHardy, Pollard, & Luo, 2007; Thériault & Lachance, 1998) devendo-se ter como constrangimento a principal queixa associada, a lombalgia.

1.2.1. Handicap

Durante *o swing*, os segmentos corporais e taco combinam-se numa cadeia segmentar com o intuito de transferir energia para a bola, para que esta atinja uma velocidade e uma trajetória que maximize a distância e a precisão (Farrally et al., 2003; Keogh & Hume, 2012; Hume, Keogh, & Reid, 2005). Quando menor for a necessidade de realizar tacadas para esse propósito melhor será o nível técnico e mais baixo o Hc que poderá ser categorizado em conformidade com a tabela 1 (EGA, 2012).

Tabela 1 – Categorias de handicap.

Categorias de Handicap	Handicap Exato EGA
1	"Plus"- 4.4
2	4.5 – 11.4
3	11.5 – 18.4
4	18.5 – 26.4
5	26.5 – 36.0

Vários modelos têm sido propostos para analisar o movimento (Penner, 2003): modelo *double-pendulum*, modelo do incremento na velocidade da cabeça do taco, modelo *triple-link*. O primeiro tende a incidir sobre o *downswing* (fase de execução) assumindo duas ligações como a designação indica. A ligação superior corresponde aos ombros e braços do golfista que giram sobre um eixo central que corresponde aproximadamente a um ponto entre os ombros do golfista. A ligação inferior representa o taco, que roda em torno de um ponto localizado no centro das mãos e dos pulsos do jogador de golfe. Quando se inclui a velocidade do taco é de esperar que quanto maior for a velocidade da cabeça do taco no impacto, maior seja a distância que a bola irá percorrer. Daí, que o segundo modelo relacione variáveis retiradas do modelo *double-pendulum* com a velocidade da cabeça do taco, sendo este um parâmetro de eficácia. No modelo *triple-link*, a ligação superior representa o movimento dos ombros do jogador à medida que roda. O *link* médio representa o braço esquerdo no golfista destro. A ligação inferior é novamente constituída pelo taco de golfe. Apesar das críticas que podem resultar destes modelos (Coleman & Rankin, 2005), os mesmos permitem a extração de parâmetros, mesmo que separadamente, podendo comparar a nível da cinemática, as diferentes características na execução do *swing*. A cadeia estudada será tão eficaz quanto a sua capacidade em transferir energia para a bola no momento do impacto (Nesbit & McGinnis, 2009; Nesbit & Serrano, 2005). No entanto, o nível técnico interfere em pontos-chave que influenciam a execução do *swing* (Bradshaw, Keogh, Hume, Maulder, Nortje, & Marnewick, 2009), assim como as características morfológicas e funcionais (Keogh, Marnewick, Maulder, Nortje, Hume, & Bradshaw, 2009).

Encontramos três dimensões de medidas para análise cinesiológica do *swing*: a cinemática, cinética e a EMG. Porém, a literatura que usa EMG para estudar o *swing* incide essencialmente na população golfista profissional e/ou amadores de baixo handicap comparativamente aos de alto handicap (Marta, Silva, Castro, Pezarat-Correia, & Cabri, 2012), apesar de evidenciadas diferenças cinemáticas segundo o Hc (Gatt, Pavol, Parker, & Grabiner, 1998; Somjarod, Tanawat, & Weerawat, 2011; Zheng, Barrentine, Fleisig, Andrews, & States, 2008). Várias conclusões podem ser retiradas destes estudos. Diferentes níveis de Hc apresentam alterações significativas no joelho esquerdo quanto à flexão (Gatt et al., 1998; Somjarod et al., 2011), rotação externa (Somjarod et al., 2011) e rotação interna (Gatt et al., 1998). Golfistas profissionais destros e/ou com baixo Hc apresentam maiores amplitudes na rotação do tronco, na adução horizontal do ombro esquerdo (*lead* em golfistas destros) e na rotação externa no ombro direito (*trail* em golfistas destros) comparativamente com jogadores de Hc alto, no pico do *backswing*. No momento de impacto com a bola também são apontadas diferenças na flexão do cotovelo esquerdo (maior em Hc alto), no pulso esquerdo (ângulo entre os segmentos pulso-taco e pulso-antebraço) e na rotação do tronco (ambos maiores em Hc baixo). Da mesma forma, são apontadas diferenças face ao nível técnico para as velocidades angulares máximas obtidas no *downswing*, com maior expressão na extensão do cotovelo direito e na velocidade do taco (maior em Hc baixo, nomeadamente profissionais) (Zheng et al., 2008). Uma maior estabilidade neuromuscular nos golfistas de elevado nível técnico estará associada a uma maior capacidade na separação entre o quadril e os ombros durante o *backswing* aquando a rotação do tronco (Hellström & Tinnmark, 2008; Myers et al., 2008). Esta relação entre o segmento pélvico e o segmento dos ombros denomina-se por *X-factor* (McLean, 1992). Posteriormente, McLean (2008) desenvolveu outro conceito, o *triple X-factor*, sendo abrangidas a componente estática (*static X-factor*), a componente dinâmica quando a amplitude aumenta no início do *downswing* (*dinamic X-factor*), e esse mesmo alongamento (*X-factor strech*). Estas considerações ganham relevância visto não ser propriamente no *X-factor* onde se encontram as principais diferenças face ao nível técnico, mas sim, que golfistas de baixo Hc apresentam maior *X-factor strech* comparativamente aos golfistas de alto Hc no início do *downswing* (Cheetham, Martin, Mottram, & Laurent, 2001). Sujeitos com Hc inferior a 10 tendem a rodar os ombros excessivamente no topo do *backswing* terminando essa rotação após o quadril já ter começado a rodar na direção da bola (Burden, Grimshaw, & Wallace, 1998) para conseguir maior velocidade angular do taco e velocidade linear da cabeça do taco próximo ou durante o impacto com a bola. A ordem a que acontecem os picos de velocidade durante o *downswing* demonstra diferentes estratégias na transferência de velocidade para o taco entre profissionais e amadores (Cheetham et al., 2008). A ordem nos profissionais traduz-se em pélvis-tórax-braço, enquanto os golfistas amadores seguem a ordem pélvis-braço-tórax.

Como o próprio nome indica, no *swing* existe um balancear do lado esquerdo para o direito e deste novamente para o lado esquerdo, centrando-se a dinâmica ao nível da transferência de peso através das forças de reação ao solo. Os golfistas de baixo Hc evidenciam uma transferência de peso mais lenta comparativamente com golfistas de médio e alto Hc (Worsfold, Smith, & Dyson, 2008) no *swing*. Em relação ao *putting*, os golfistas de baixo Hc exibem menor deslocamento do centro de pressão no eixo ântero-posterior comparativamente com Hc mais elevados (Richardson, Hughes, & Mitchell, 2012).

Dos estudos de EMG que incluíram comparações entre Hc, ambos incidiram em músculos do antebraço (Farber, Smith, Kvitne, Mohr, & Shin, 2009; Glazebrook et al.,1994). Apesar de Glazebbrok et al., (1994) evidenciar similaridade entre os dois grupos, Farber et al. (2009) verificaram que sujeitos de alto e baixo handicap diferem quanto à intensidade de ativação do redondo pronador. Golfistas amadores apresentaram maior atividade muscular no braço direito (jogadores destros) durante o *foward swing* (face inicial do *downswing*). Por sua vez, os golfistas profissionais apresentaram maior atividade muscular no braço esquerdo para o mesmo músculo durante a fase de aceleração (fase final do *downswing* até ao impacto). Porém, fica presente a escassez de estudos que impliquem diferentes estratégias de ativação neuromuscular em função do Hc. Ao contrário verifica-se uma preocupação em discriminar quanto o músculo ativa nas diferentes fases do *swing* (Marta, Silva, Vaz, Bruno, & Pezarat-Correia, 2013; McHardy

& Pollard, 2005b). As regiões com maior prevalência de estudo são a cintura escapular (Jobe, Perry, & Pink, 1986, 1989; Pink, Jobe, & Perry, 1990; Kao, Pink, Jobe, & Perry, 1995) e o tronco (Bulbulian, Ball, & Seaman, 2001; Pink, Perry, & Jobe, 1993; Watkins et al., 1996), embora a grande maioria dos estudos analise apenas a intensidade do sinal EMG.

Anteriormente nesta dissertação, no ponto sobre parâmetros temporais, foi possível verificar igualmente uma incidência em músculos do tronco diferenciando populações por lombalgia (Cole & Grimshaw, 2008a; Horton et al., 2001) e não por handicap. No melhor do nosso conhecimento, os únicos dois estudos que focam o membro inferior (Bechler, Jobe, Pink, Perry, & Ruwe, 1995; Watkins et al., 1996) estudaram golfistas profissionais e/ou com handicap inferior a 5. Sabendo que o movimento realizado no *backswing* influencia a forma como os músculos do tronco são ativados, visto um maior comprimento no *backswing* estar associado a maiores níveis de ativação em músculos do tronco no início do *downswing* (Ashish, Shweta, & Singh, 2008; Bulbulian et al., 2001), poderá ser na solicitação dos músculos do tronco e membro inferior que poderão existir características que diferenciem grupos de Hc.

Conhecer os mecanismos de coordenação muscular por parte do sistema nervoso central em função do Hc, também assume relevância visto o nível técnico também estar associado à ocorrência de lesões (Batt, 1992).

1.2.2. Lombalgia

Usualmente, cinco regiões são apontadas como aquelas onde ocorre maior incidência de lesões no golfe: a região lombar, o cotovelo, o tornozelo, o ombro (nomeadamente do lado *lead*) e o pulso (Cabri, Sousa, & Barreiros, 2009; McHardy, Pollard, & Luo, 2007; Thériault & Lachance, 1998). Destas, a lombalgia tem sido referenciada como a principal queixa entre golfistas sendo por isso alvo de estudo (Cole & Grimshaw, 2008a; 2008b; Evans, Refshauge, Adams, & Aliprandi, 2005; Horton et al., 2001; Lindsay & Horton, 2002; McHardy & Pollard, 2005a; 2007; Tsai et al., 2010).

A lombalgia em golfistas japoneses de elite contribui para uma redução na participação do golfe em cerca de 55% dos sujeitos (Sugaya, Tsuchiya, Moriya, Morgan, & Banks, 1999). Destes, 51% identificaram a presença de dor do lado direito (lado *trail* em destros) e 28% do lado esquerdo (*lead* em destros), tendo os restantes não reportado lateralidade. Os fatores tidos como preditores de lombalgia em golfistas profissionais são o índice de massa corporal, estabilidade dos músculos do tronco e o comprimento dos flexores do quadril (Evans et al., 2005), défices na amplitude de movimento na rotação interna da articulação coxofemoral *lead* e na extensão da região lombar (Vad et al., 2004). Estas capacidades são diminuídas em sujeitos com historial de lombalgia. Golfistas com lombalgia realizam maior flexão do tronco durante o *address* (início do *swing*) e maior flexão lateral para o lado esquerdo durante o *backswing* (Lindsay & Horton, 2002), evidenciando menor força na extensão do tronco com velocidade angular de 60°/s e menor força na adução da coxofemoral do lado esquerdo (Tsai et al., 2010). Golfistas sem lombalgia desempenham maior amplitude de rotação do tronco e maior velocidade na flexão do tronco durante o *downswing* (Lindsay & Horton, 2002) apresentando maior atividade EMG nos músculos do tronco (Horton et al., 2001).

Será de questionar se a existência de diferentes estratégias de recrutamento neuromuscular está associada à realização do *swing* face à incidência ou prevalência de lombalgia. Desequilíbrios neuromusculares nos eretores da coluna lombar estão associados à sintomatologia de lombalgia (Renkawitz, Boluki, & Grifka, 2006). Em relação ao início do *backswing* sujeitos com lombalgia apresentam uma ativação dos eretores da coluna (EMG de superfície) antecipatória comparativamente com golfistas assintomáticos, o que pode significar uma utilização destes músculos ao invés da camada mais profunda na estabilização do tronco (Cole & Grimshaw, 2008a). Apesar destas diferenças temporais para este grupo muscular, não é no *address* onde são encontradas divergências quanto à intensidade de recrutamento. Golfistas de baixo Hc com lombalgia apresentam menor atividade da massa comum superficial no topo do *backswing* e no impacto quando comparados com jogadores sem lombalgia. Porém, quando realizada a mesma comparação com golfistas de alto Hc, apenas no topo do *backswing* são encontradas diferenças significativas e com a razão inversa (Cole & Grimshaw, 2008b). Está presente a relevância do Hc e da sintomatologia de lombalgia na população golfista, pelo menos em relação ao recrutamento deste grupo muscular. Para o oblíquo externo verificam-se na literatura resultados contraditórios, nomeadamente no lado esquerdo. Horton et al. (2001) encontraram diferenças significativas na ativação do oblíquo externo do lado esquerdo, dando-se esta ativação antes do impacto para os sujeitos assintomáticos e após para os golfistas com lombalgia. O estudo de Cole e Grimshaw (2008a) parece apontar para tendência contrária, apesar destes últimos não terem apresentado diferenças significativas entre golfistas com e sem lombalgia.

Apesar das fases do swing que costumam ser mais estudadas serem o *backswing* e o *downswing*, respetivamente, é durante o *follow-through* que são reportadas as maiores incidências de lombalgia (McHardy et al., 2007). Tal pode demonstrar uma possível fraqueza nas respostas por questionário, pois apesar da lesão ter acontecido anteriormente, apenas é percepcionada no *follow-through*. Porém, estudos que analisam a solicitação neuromuscular em golfistas com e sem lombalgia tendem a utilizar questionários associados à dor aguda como o *short-form McGill Pain Questionnaire* (ex:. Cole & Grimshaw, 2008a; 2008b; Horton et al., 2001). Sendo o conceito de lombalgia vago relativamente à sua causa, consequência e associação com a prática de golfe, a definição operacional que lhe está associada deve ser realizada de forma a diminuir a subjectividade inerente à mesma. Tendo em conta a apetência de golfistas recreativos de idade mais avançada contraírem lesões devido a declíneos na força, diminuição na flexibilidade e na coordenação, assim como a pré-existência de outras lesões (Cann, Vandervoort, & Lindsay, 2005; Lindsay, Horton, & Vandervoort, 2000) emerge a necessidade de quantificar esta população face às características do golfe. Assim surge o questionário *Musculoskeletal Injury Questionnaire for Senior Golfers* (Fox, Lindsay, & Vandervoort, 2002). Este questionário foi posteriormente utilizado para fins epidemiológicos (Palmer, Young, Fox, Lindsay, & Vandervoort, 2003). Das vantagens que podemos associar a este questionário destacam-se o histórico de lesões e hábitos de atividade física do sujeito e a mais valia de questionar o golfista sobre como este se sente relativamente à lombalgia após a realização de uma volta de golfe de 18 buracos.

Identificar diferentes estratégias de ativação muscular pode apresentar utilidade em programas de treino e de recuperação, visto que os fatores mais comuns que contribuem para lesões no golfe são a própria execução do *swing* face ao stress biomecânico que exerce, a repetibilidade do movimento, idade, equipamento, ambiente e a condição física (Sherman & Finch, 2000). Durante o *swing*, a velocidade na rotação do tronco ronda os 85º/s no *backswing* e 190º/s no *downswing* (Lindsay, Horton, & Paley, 2002), quando já houve oportunidade neste ponto de evidenciar falta de capacidade neuromuscular de rodar o tronco a elevada velocidade por sujeitos com lombalgia. A ativação dos músculos do tronco tende a apresentar períodos de incapacidade de estabilização face a falhas na ativação muscular postural antes do início do movimento da flexão do braço (Silfies, Mehta, Smith, & Karduna, 2009). A ineficácia na estabilização da região lombar foi também já associada à lombalgia (Hodges & Richardson, 1996). Sendo o *swing* uma habilidade motora complexa que envolve transferência de energia em cadeia através dos segmentos corporais até à bola (Nesbit & McGinnis, 2009; Nesbit & Serrano, 2005), a execução do *swing* corresponde a um elevado stress mecânico na região lombar, nomeadamente devido ao excessivo *X-factor* (Gluck, Bendo, & Spivak, 2008) que muitas vezes é produzido. Por este motivo são propostas alterações a nível do *swing*, sendo consideradas duas abordagens de estudo, o *swing* tradicional e o *swing* moderno (McHardy & Pollard, 2005a).

A escolha do *swing* moderno em detrimento do tradicional está associado a uma tacada mais potente e a consequente maior distância percorrida pela bola. No *swing* clássico, o *backswing* é caracterizado por uma considerável rotação da cintura escapular acompanhada pela rotação pélvica. Devido a esta última, o calcanhar esquerdo levanta do chão (em golfistas destros) (Hosea & Gatt, 1996). No *swing* moderno, os ombros ficam totalmente virados de costas em relação ao alvo quando a cintura pélvica ainda está orientada para a bola. Neste caso é produzido um elevado momento de força entre o movimento da cintura escapular e a cintura pélvica que se encontra estacionária. O movimento pélvico inicia o *downswing* e termina no *follow-through* numa posição que se denomina por *"reverse C"* (Adlington, 1996). O *swing* moderno está associado a um recrutamento dos músculos do tronco e produz maior força na região lombar comparativamente ao tradicional (Ashish et al., 2008). Os modelos designados por *linked segment* têm demonstrado que a redução de momentos articulares e forças de corte previne a ocorrência de lesões nessas estruturas (McGill, 2004). Encurtar a realização do *backswing* em cerca de 46.5º, o que permite prevenir lesão na região lombar, pode reduzir a atividade dos músculos do tronco sem prejudicar a precisão do *swing* ou reduzir a velocidade da cabeça do taco (Bulbulian et al., 2001). A relação entre flexão lateral do tronco e a velocidade axial da rotação do tronco, na forma de produto, foi uma das formas encontradas para comparar mecanismos de lesão, denominando-se por *crunch factor*. Porém, ainda não se verificaram evidências sólidas sobre a sua sensibilidade na discriminação de lombalgia (Cole & Grimshaw, 2013; Lindsay & Horton, 2002).

Coloca-se então a necessidade de desenvolver um instrumento que permita categorizar os principais constrangimentos associados a determinada habilidade, que no caso do presente estudo é o *swing* do golfe. Outro aspeto a realçar é a lacuna existente na literatura quando à discriminação da atividade muscular tendo em conta esses mesmos constrangimentos, sendo um meio de conhecer possíveis estratégias neuromusculares que são adotadas ou influenciadas pelo Hc e pela lombalgia.

1.3. Aprendizagem Automática

A aprendizagem automática ganha relevância nesta dissertação a partir da análise dos parâmetros temporais, do paradigma sobre o computador conseguir realizar a inspeção visual pelo ser humano. Coloca-se a possibilidade de realizá-la *à priori*, construíndo um modelo face ao fenómeno biológico a identificar. Entramos no domínio de conceitos como aprendizagem automática, inteligência artificial e nomeadamente, no suporte matemático constituído pela Teoria da Estatística da Aprendizagem (TEA). A aprendizagem automática assume a possibilidade da máquina poder aprender a executar determinada tarefa, que no contexto de classificação está muito associada à problemática do reconhecimento de padrões. Assim, para um conjunto de dados observáveis e classificáveis torna-se necessário considerar um problema de aprendizagem. Um problema de aprendizagem corresponde a um mapeamento de dados $\boldsymbol{x}$ segundo rótulo de dados y_i, em que $f : \mathbb{R}^n \to \{-1,1\}$. Para que determinada função $f(\boldsymbol{x})$ represente o rótulo y, define-se o problema de aprendizagem como:

Definição 1 (Problema de Aprendizagem):

O problema de aprendizagem é encontrar uma relação funcional desconhecida $f : \mathbb{R}^n \to \{-1,1\}$ *entre os objetos* $\boldsymbol{x} \in \mathcal{X}$ *e os rótulos de saída* $y \in \mathcal{Y}$ *baseado unicamente numa amostra* $(\boldsymbol{x}, y) = \big((x_i, y_i), \dots, (x_n, y_n)\big) \in (\mathcal{X}, \mathcal{Y})^n$ *de dimensão* $n \in \mathbb{N}$ *de dados independentemente e identicamente distribuídos (i.i.d.) a partir de uma distribuição* $P(\boldsymbol{x}, y)$ *desconhecida. Se o espaço de saída* $\mathcal{Y}$ *contém um número finito* $|y|$ *de elementos, a tarefa é chamada um problema de aprendizagem de classificação.*

(adaptado Herbrich, 2002, p. 18)

Partindo de um conjunto de funções, seleciona-se aquela que prevê a melhor resposta para o conjunto de dados fornecidos. Esta seleção depende da definição de n observações aleatórias, independentes e identicamente distribuídas denominadas por conjunto de treino, em que $P(\boldsymbol{x}, y) = P(\boldsymbol{x})P(y \backslash \boldsymbol{x})$. O par $(\boldsymbol{x}, y)$ é uma amostra independente desta distribuição P. Não são realizados pressupostos sobre os espaços de $\boldsymbol{x}$ e y, mas sim, em relação à forma como são gerados os pontos de treino (Luxburg & Schölkopf, 2009), sendo mantidas as seguintes proposições:

a) *A não realização de pressupostos sobre* P, sendo assumido que a distribuição probabilidade pertence a qualquer família de distribuições a estimar;

b) *Rótulos não determinísticos devido a ruídos ou a classes sobrepostas*. A distribuição P refere-se às entradas $\boldsymbol{x}$ e aos rótulos y, pelo que estes não correspondem a uma função determinística especifica de $\boldsymbol{x}_i$. Este pressuposto deve-se à possibilidade de existência de ruído derivado de uma atribuição incorreta do rótulo e/ou pela possibilidade de existir sobreposição de classes;

c) *Amostragem independente*. A amostragem deverá garantir representatividade populacional em relação ao rótulo de dados não induzindo numa determinada direção que influencie a aprendizagem. Este pressuposto está dependente dos dados a classificar, como séries temporais onde as entradas são geradas a partir de sobreposição de janelas temporais;

d) *A distribuição P é fixa*, não é assumida qualquer ordenação particular dos exemplos de treino, assim como, a distribuição de probabilidade subjacente não se altera ao longo do tempo. Porém, tal como no ponto c), quando aplicado a séries temporais este pressuposto pode ser ignorado;

e) *A distribuição P é desconhecida no momento da aprendizagem*. A aprendizagem acontece por não se conhecer P, o acesso à distribuição vai depender dos exemplos de treino.

Tendo presente os pressupostos da TEA, pretende-se construir um modelo que aprenda a partir de exemplos com a capacidade de minimizar perdas cuja robustez dependerá da capacidade de detetar as regularidades inerentes aos dados (Scholkopf, Burges, & Vapnik, 1995). Os maiores desvios que um modelo pode apresentar estão associados à falta de capacidade dos recursos observáveis em capturar as propriedades do objeto a classificar e pelo ruído existente em valores similares (Kulkarni & Harman, 2011). Sumarizando o descrito até então, o modelo de estimativa da função na TEA pode ser desenvolvido em três componentes fundamentais (Vapnik, 1999):

a) Gerador de vetores aleatórios $\boldsymbol{x} \in \mathbb{R}^n$, elaborado de forma independente a partir de uma função de probabilidade cumulativa $P(\boldsymbol{x})$, fixa e desconhecida;

b) Supervisor que devolve um vetor *output* y para cada vetor *input* $\boldsymbol{x}$ tendo em conta uma determinada função distribuição condicional $P(y \backslash \boldsymbol{x})$, também fixa e desconhecida;

c) Máquina de aprendizagem capaz de implementar um conjunto de funções $y = f(\boldsymbol{x}, \alpha), \alpha \in \Lambda$, em que f corresponde ao classificador, α aos parâmetros da função a aprender, y é a classe de classificação e $\boldsymbol{x}$ o elemento no espaço dos dados, com Λ a representar o conjunto de parâmetros possíveis.

1.3.1. Problema de aprendizagem

Torna-se possível então a construção de problemas de aprendizagem em que L será a função relacionada com a precisão do algoritmo, variando consoante o contexto, podendo este ser o reconhecimento de padrões

$$L\big(y, f(\boldsymbol{x}, \alpha)\big) = \begin{cases} -1 \; se \; y = f(\boldsymbol{x}, \alpha) \\ \;\; 1 \; se \; y \neq f(\boldsymbol{x}, \alpha) \end{cases}, \tag{1}$$

problemas de regressão

$$L\big(y, f(\boldsymbol{x}, \alpha)\big) = \big(y - f(\boldsymbol{x}, \alpha)\big)^2, \tag{2}$$

e/ou problemas da estimação de densidade

$$L\big(y, f(x, \alpha)\big) = -\log p(x, \alpha). \tag{3}$$

O conjunto de n características de um objeto de valores $x_1, x_2, \dots, x_n$ pode ser agrupado de forma a constituir um vetor característico $\boldsymbol{x}$, que será um vetor no espaço n-dimensional $\mathbb{R}^n$ denominado por espaço característico ou *feature space* (Burges, 1998; Kulkarni & Harman, 2011; Schölkopf & Smola, 2002). Assim, outra definição que deve estar presente refere-se às entradas características (*features*) e espaço das entradas características (*feature space*), nomeadamente para modelar a semelhança entre objetos através de uma função do produto interno.

Definição 2 (entradas e espaço característico)

Segundo um classificador representado pela função $f : \mathbb{R}^n \to \{-1,1\}$ *que mapeia cada componente dos objetos* $\boldsymbol{x} \in \mathcal{X}$, *cada* $\boldsymbol{x}$ *é chamado de entrada característica (feature). As várias features vão construir um espaço característico de dimensionalidade n, pelo que o espaço euclidiano característico será representado por* $\mathcal{X} \subset \mathbb{R}^n$. *No caso de se aplicar um "kernel trick" existindo a necessidade de recorrer a produtos internos num espaço de dimensão superior com mapeamento* $\Phi: \mathcal{X} \to \mathcal{H}$, *então* $\mathcal{H}$ *é denominado por espaço característico.*

(adaptado Herbrich, 2002, p. 20)

Usualmente para a representação vetorial utiliza-se **x.** Desta forma, para que não haja dificuldade na distinção terminológica entre componentes de um vetor e da representação de *features*, a partir deste ponto e nomeadamente na exposição apresentada no método, $\boldsymbol{x}_i$ corresponderá a cada componente do vetor $\boldsymbol{x}$, sendo que **x** será a representação vetorial de $\boldsymbol{x}$ no espaço característico (*feature space*). Com a inserção de dados através de Φ para o espaço $\mathcal{H}$, teremos o mapeamento $\Phi: \mathcal{X} \to \mathcal{H}$ com $x \longmapsto \mathbf{x}$.

Para atingir os objetivos da aprendizagem automática, duas grandes dimensões de problemas de aprendizagem podem ser distinguidas na literatura (Haykin, 2001; Herbrich, 2002; Luxburg & Schölkopf, 2009; Rojas, 1996): a aprendizagem não supervisionada e a aprendizagem supervisionada, conforme ilustrado na Figura 1. Na primeira não existe intervenção do investigador senão para introduzir e correr os dados, ou seja, não existem exemplos rotulados e depois aprendidos. Usualmente, o investigador desconhece quais ou quantas classes ou componentes discriminatórios vão ser produzidos após a utilização do algoritmo. A classificação por *clusters*, análise fatorial e técnicas estatísticas afins são exemplos de aprendizagem não supervisionada. Nesta dissertação, quando nos referimos a aprendizagem automática, esta reporta-se à aprendizagem supervisionada, ou seja, em determinada altura do processo, o investigador vai interferir através de exemplos de treino que estão rotulados em conformidade com a classe a que pertencem os dados. Existe um processo de treino através de exemplos de entrada-saída (rótulos) retirados de um ambiente desconhecido. Para que um algoritmo seja considerado robusto deverá "aprender" através dos exemplos fornecidos, estando porém, dependente da relação que é criada entre o ambiente fornecido e a capacidade de adaptação, sendo um ramo da inteligência artificial. Haykin (2001) refere que um sistema de inteligência artificial deverá abarcar três capacidades primordiais: (1) armazenar conhecimento, (2) aplicar esse conhecimento na resolução de problemas, e (3) adquirir novo conhecimento através da experiência.

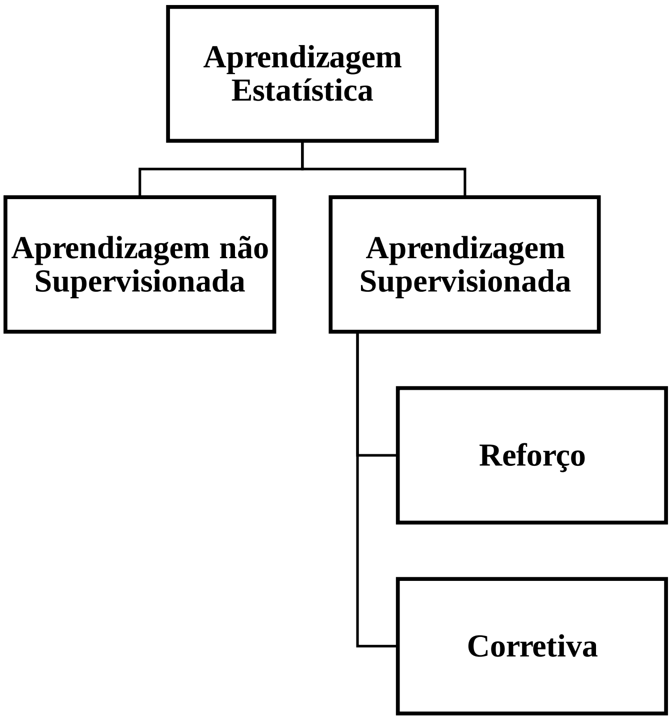


Figura 2 – Classes de algoritmos de aprendizagem (adaptado de Rojas, 1996, p.79).

Apesar do ilustrado na Figura 2, é possível encontrar na literatura a aprendizagem por reforço associada à aprendizagem não supervisionada (Haykin, 2001). A razão deve-se à aprendizagem sem reforço consistir num mapeamento de entrada e saída relacionado com a interação contínua do ambiente. O sistema terá em consideração uma sequência temporal de estímulos ou, pelo facto de existir um reforço, assume a presença de supervisão sem que as respostas tenham sido obtidas previamente. A categorização da aprendizagem por reforço dentro da aprendizagem supervisionada reporta-se após a apresentação dos exemplos de treino, podendo ser produzido o resultado pretendido ou não. Na aprendizagem com correção considera-se o erro em conjunto com o vetor de entrada, sendo determinada a magnitude de correção desse erro (Rojas, 1996).

1.3.2. Minimização do erro e dimensão Vapnik Chervonenkis

Quando nos reportamos à aprendizagem automática supervisionada, referimo-nos às máquinas de vetor suporte (SVM). Através das SVM, as abordagens dependem da separabilidade ou não separabilidade dos dados e das demais extensões, como se passará a abordar. As SVM serão desenvolvidas desde o caso de dados separáveis até à extensão usada na presente dissertação. A máquina assume que é determinística (Burges, 1998), ou seja, para uma determinada entrada **x** e respetiva escolha α, será fornecido sempre o mesmo output $f(\mathbf{x}, \alpha)$. A probabilidade $P(\mathbf{x}, y)$ descreve a relação entre os dados e os rótulos procurando a função $f(\mathbf{x}, \alpha)$ que se traduza no menor erro possível, ou seja, que forneça o menor risco expresso pelo risco esperado $R(\alpha)$ (Burges, 1998; Cristianini & Shawe-Taylor, 2000; Scholkopf et al., 1995):

$$R(\alpha) = \frac{1}{2}\int |y - f(\mathbf{x}, \alpha)| dP(\mathbf{x}, \alpha) \tag{4}$$

Como $P(\mathbf{x}, \alpha)$ é desconhecido utiliza-se o princípio da indução para inferir uma função $f(\mathbf{x},\alpha)$ para minimizar o erro, isto é, procede-se à minimização do risco empírico $R_{emp}(\alpha)$, que corresponde a um valor fixo para uma determinada escolha de α e um determinado conjunto de treino $\{\mathbf{x}_i, y_i\}$ oriundo de l exemplos.

$$R_{emp}(\alpha) = \frac{1}{2l}\sum_{i=1}^{l}|y_i - f(\mathbf{x}_i, \alpha)| \tag{5}$$

Na expressão (5) não aparece a distribuição de probabilidade, mas sim a escolha de determinado α para o par $\{\mathbf{x}_i, \alpha\}$. A parcela $\frac{1}{2}|y - f(\mathbf{x}, \alpha)|$ é denominada por função custo, que neste caso apenas assume os valores 0 e 1. Optando por um $\eta \in [0,1]$ com a probabilidade de pelo menos $1 - \eta$ é obtido o seguinte limite (Burges, 1998; Scholkopf et al., 1995; Vapnik, 1999):

$$R(\alpha) \leq R_{emp}(\alpha) + \sqrt{\frac{h\left(\log\left(\frac{2l}{h}\right) + 1\right) - \log\left(\frac{\eta}{4}\right)}{l}} \tag{6}$$

onde h é um valor inteiro não negativo designado por dimensão Vapnik-Chervonenkis (VC). A dimensão VC é uma propriedade de um conjunto de funções $\{f(\alpha)\}$ sendo por sua vez, definida pelas classes que a integram, correspondendo ao número máximo de pontos de treino que podem ser representados por $\{f(\alpha)\}$. Torna-se possibilitada a escolha de uma máquina de aprendizagem para uma determinada tarefa, o princípio fundamental da minimização do risco estrutural. Dada uma família de máquinas de aprendizagem, tendo em conta que o limite será restrito para pelo menos uma das máquinas, espera-se que pelo menos uma desempenhe a melhor performance. Desta forma, podem ser destacados o seguinte teorema e respetivo corolário (Burges, 1998):

Theorem 1: *Consider some set of m points in* $\mathbb{R}^n$. *Choose any one of the points as origin. Then the m points can be shattered by oriented hyperplanes if and only if the position vectors of the remaining points are linearly independent.*

Corollary 1: *The VC dimension of the set of oriented hyperplanes in $\mathbb{R}^n$ is n+1, since we can always choose n + 1 points, and then choose one of the points as origin, such that the position vectors of the remaining n points are linearly independent, but can never choose n + 2 such points (since no n + 1 vectors in $\mathbb{R}^n$ can be linearly independent).*

(Burges, 1998, p.4, *ipsis verbis*)

Resumindo o problema da minimização de risco, pretende-se escolher a melhor aproximação para a resposta do supervisor que meça o custo $\frac{1}{2}|y - f(\mathbf{x}, \alpha)|dP(\mathbf{x}, \alpha)$ entre a resposta y do supervisor para uma entrada $\mathbf{x}$ e a resposta fornecida pela máquina $f(\mathbf{x}, \alpha)$, que corresponderá a procurar a função com menor erro associado com a distribuição de probabilidade $P(\mathbf{x}, \alpha)$ desconhecida (Vapnik, 1999).

1.3.3. Introdução das máquinas de vetor suporte e sua evolução

No ponto anterior, foi referenciado o suporte para a construção dos problemas de aprendizagem recorrendo às SVM. As SVM foram introduzidas no início da década de 90 sendo considerado como um algoritmo de treino que *maximizes the margin between the training patterns and the decision boundary is presented* (Boser, Guyon, & Vapnik, 1992). Inicialmente, as SVM são apresentadas como uma técnica de classificação intimamente ligada ao reconhecimento de padrões, baseando-se na minimização do erro esperado de generalização através do método de validação de *leave-one-out* e através da dimensão VC, sendo desta forma avaliada a capacidade de classificação do algoritmo. Este conceito estava desenvolvido para os casos de dados separáveis, também designado por margens rígidas (Burges, 1998), sendo posteriormente apresentada uma extensão mais próxima de dados reais, o caso dos dados não separáveis ou de margens suaves (Cortes & Vapnik, 1995). Devido à problemática dos dados não-lineares, as SVM progrediram sendo categorizadas com o método *kernel* (Burges, 1998; Cristianini & Shawe-Taylor, 2000; Müller, Mika, Rätsch, Tsuda, & Schölkopf, 2001; Scholkopf et al., 1995; Scholkopf, Smola, & Muller, 1996; Schölkopf & Smola, 2002) devido à substituição do produto interno por uma função *kernel* que irá estimar esse produto interno no espaço característico de maior dimensionalidade, obtendo-se uma separação não linear no espaço de origem.

A base das SVM, o caso das margens separáveis, considera um conjunto de treino com n objetos $\mathbf{x}_i \in \mathcal{X}$ com os rótulos de dados $y_i \in \mathcal{Y}$, em que $\mathbf{x}_i \in \mathbb{R}^n$ e $y_i \in \{-1,1\}$, para $i = 1, 2, \ldots, l$ exemplos. O conjunto de treino será linearmente separável por um hiperplano que distinga as duas classes A e B (positiva e negativa, respetivamente). O conjunto de treino terá de considerar que para $(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \ldots, (\mathbf{x}_l, y_l)$, tem-se que (Boser et al., 1992):

$$\begin{cases} y_i = +1 \;\; se\ \mathbf{x}_i \in \boldsymbol{classe}\ \mathcal{X}(A) \\ y_i = -1 \;\; se\ \mathbf{x}_i \in \boldsymbol{classe}\ \mathcal{X}(B) \end{cases} \tag{7}$$

Pretende-se então determinar o melhor hiperplano que separa (neste caso mais simples de separação linear) as duas classes, designado por hiperplano ótimo. Este hiperplano será aquele cujas margens separadoras são maiores no seu espaço geométrico, motivo pelo qual se coloca um problema de maximização das margens. Os pontos $\mathbf{x}$ que se encontram sobre o hiperplano terão que satisfazer a condição (Boser et al., 1992; Burges, 1998; Cortes & Vapnik, 1995; Müller et al., 2001):

$$f(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x} + b = 0 \tag{8}$$

Podemos resumir no formato matricial que as SVM procuram a solução para o seguinte problema de otimização (Burges, 1998; Hsu, Chang, & Lin, 2010):

$$\min_{w,b,\xi} \quad \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{i=1}^{l}\xi_i \tag{9}$$

sujeita às restrições

$$y_i(\mathbf{w}^T\phi(\mathbf{x}_i) + b) \geq 1 - \xi_i \tag{10}$$

$$\xi_i \geq 0 \tag{11}$$

Este problema no caso não linear aplica um truque engenhoso denominado por "*kernel trick*", porque recorre à aplicação de uma função *kernel*. As funções mais usuais de serem utilizadas são as quatro que se seguem (Hsu, Chang, & Lin, 2010):

- Linear: $K(\mathrm{x}_i, \mathrm{x}_j) = \mathrm{x}_i^T \mathrm{x}_j$;

- Polinomial: $K(\mathrm{x}_i, \mathrm{x}_j) = (\gamma \mathrm{x}_i^T \mathrm{x}_j + r)^d, \gamma > 0$;

- RBF (RBF - *Radial Bais Function*) gaussiana: $K(\mathrm{x}_i, \mathrm{x}_j) = exp\left(-\gamma \|\mathrm{x}_i - \mathrm{x}_j\|^2\right), \gamma > 0$;

- Sigmoidal: $K(\mathrm{x}_i, \mathrm{x}_j) = \tanh(\gamma \mathrm{x}_i^T \mathrm{x}_j + r)$.

em que γ, r e d são parâmetros do *kernel*. O número de vetores suporte necessários ao usar as funções *kernel* polinomial, RBF ou sigmoide não diverge muito, sendo a RBF aquela que apresenta menor número de vetores na construção do modelo de decisão (Scholkopf et al., 1995).

No método geral desta dissertação será apresentada uma explanação mais exaustiva sobre as SVM, passando pelas SVM para dados separáveis, não separáveis (linear) e o caso utilizado, SVM para dados não lineares. Independentemente do método utilizado, outras propostas de aplicação têm sido apresentadas, como as v-SVM, o caso multiclasses e regressão.

As v-SVM foram propostas por Schölkopf, Smola, Williamson, & Bartlett (2000) denominadas por novas SVM, uma classe de algoritmos tanto para classificação como para regressão. É introduzido um parâmetro v que permite controlar o número de vetores suporte e eliminar o parâmetro C no caso da classificação e o parâmetro ε na regressão. Esta via consiste na parametrização v, no qual o parâmetro C é substituído por $v \in [0,1]$, ao desaparecer C da equação ficam os parâmetros v e ρ com o seguinte problema primal (Chen et al., 2005):

$$\min_{\mathbf{w},\xi,\rho} \quad \frac{1}{2}\|\mathbf{w}\|^2 - \nu\rho + \frac{1}{2}\sum_{i=1}^{l}\xi_i, \quad \mathbf{w} \in \mathcal{H};\ \xi \in \mathbb{R}^n, \rho \in \mathbb{R} \tag{12}$$

Com as restrições

$$\begin{cases} y_i(\mathbf{w}\cdot\mathbf{x}_i + b) \geq \rho - \xi_i \\ \xi_i \geq 0 \\ \rho \geq \ 0 \end{cases} \tag{13}$$

O parâmetro ρ está associado a $y_i(\mathbf{w}\cdot\mathbf{x}_i + b) \geq \rho - \xi_i$ porque esta sujeição divide duas classes pela margem $2\rho/\|\mathbf{w}\|$ quando $\xi = 0$. Para explicar a significância de ν, Chen et al. (2005) introduzem o conceito de margem de erro (*margin error*), que corresponde aos pontos de treino que contêm erro ou estão dentro da margem, com $\xi_i > 0$. A fração da margem de erros é dada pela equação:

$$R_{emp}^{\rho}[g] := \frac{1}{n}|\{i: y_i g(\boldsymbol{x}_i) < \rho\}| \tag{14}$$

onde

$$g(\boldsymbol{x}) = sgn\sum_{i=1}^{l} y_i\,\alpha_i k(\boldsymbol{x},\boldsymbol{x}_i) + b \tag{15}$$

A outra extensão refere-se ao conceito de multiclasses, que pode ser realizado tanto através de separadores lineares como por *kernel* (Hsu & Lin, 2002). Tanto podem ser usados classificadores binários combinados entre si, ou então, métodos que considerem todos os dados para otimização, nomeadamente, todas as classes. Destes métodos podem ser referidas técnicas como "*one-against-all*", "*one-against-one*", gráficos acíclicos direcionados de SVM (*Directed acyclic graph SVM*) e "*all-together*" (Müller et al., 2001). Aliadas a estas técnicas, outras extensões podem ser utilizadas para diminuir a complexidade na classificação. Por exemplo, Li, Yang, Jiang, Liu e Cai (2012) recolheram sinal de 32 canais de *features*, logo com dimensionalidade $\mathbb{R}^{32}$, na construção de um algoritmo SVM para classificar 17 gestos dos dedos. Aplicaram as SVM multiclasses "*one-against-one*" obtendo um total de 136 classificadores binários com $k(k-1)/2$, onde k corresponde ao número de classes, aplicando posteriormente *majority voting*. Partindo de um classificador binário para dados linearmente separáveis, as SVM evoluíram para uma ampla diversidade de aplicações permitindo tarefas supervisionadas no reconhecimento de dados não lineares e com maior dimensão de classes a classificar.

Após a decisão de recorrer à aprendizagem automática através das SVM, urgia a necessidade de conhecer o estado da arte sobre a utilização destas no sinal EMG, nomeadamente, em relação ao reconhecimento de padrões. Para além de conhecer as problemáticas de estudo envolvidas, era preciso perceber quais as metodologias que estavam associadas, quer em relação aos parâmetros respeitantes ao algoritmo SVM como nas restantes dimensões de estudo, nomeadamente, no que respeita a parâmetros temporais e ao aproveitamento das características de não linearidade do sinal EMG. Será apresentada na metodologia uma explicação sobre os três pontos base das SVM, margens separáveis, margens suaves e SVM não linear. Neste enquadramento teórico não será desenvolvida a análise sobre a literatura que apresenta aplicações das SVM ao sinal EMG porque tal constitui um artigo integrante neste trabalho, a revisão de literatura sistemática qualitativa apresentada no Capítulo V.

Referências

Adlington, G. S. (1996). Proper swing technique and biomechanics of golf. *Clinical Sports Medicine*, *15*, 9–26.

Allison, G. T. (2003). Trunk muscle onset detection technique for EMG signals with ECG artefact. *Journal of Electromyography and Kinesiology*, *13*(3), 209–216. doi:10.1016/S1050-6411(03)00019-1.

Ashish, A., Shweta, S., & Singh, S. J. (2008). Comparison of lumbar and abdominal muscle activation during two types of golf swing: An EMG analysis. *RICYDE. International Journal of Sport Science*, *IV*(12), 59–71.

Batt, M. E. (1992). A survey of golf injuries in amateur golfers. *British Journal of Sports Medicine*, *26*(1), 63–65.

Bechler, J.R., Jobe, F. W., Pink, M., Perry, J., & Ruwe, P. A. (1995). Electromyographic analysis of the hip and knee during the golf swing. *Clinical Journal of Sport Medicine*, *5*(3), 162–166.

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992) A Training algorithm for optimal margin classfiers. In *Proceedings of the Fifth Annual Workshop on Computational Learning Theory* (pp. 144–152).

Bradshaw, E. J., Keogh, J. W., Hume, P. A., Maulder, P. S., Nortje, J., & Marnewick, M. (2009). The effect of biological movement variability on the performance of the golf swing in high- and low-handicapped players. *Research Quarterly for Exercise & Sport*, *80*(2), 185–196.

Bulbulian, R., Ball, K. a, & Seaman, D. R. (2001). The short golf backswing: effects on performance and spinal health implications. *Journal of Manipulative and Physiological Therapeutics*, *24*(9), 569–575. doi:10.1067/mmt.2001.118982.

Burden, A. M., Grimshaw, P. N., & Wallace, E. S. (1998). Hip and shoulder rotations during the golf swing of sub-10 handicap players. *Journal of Sports Sciences*, *16*(2), 165–176. doi:10.1080/026404198366876.

Burges, C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, *2*, 121–167. Retrieved from http://link.springer.com/article/10.1023/A:1009715923555.

Cabri, J., Sousa, J. P., & Barreiros, J. (2009). Golf-related injuries: A systematic review. *European Journal of Sport Science*, *9*(6), 353–366. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/17461390903009141.

Cann, A. P., Vandervoort, A. A., & Lindsay, D. M. (2005). Optimizing the benefits versus risks of golf participation by older people. *Journal of Geriatric Physical Therapy*, *28*(3), 85-92.

Cheetham, P. J., Martin, P. E., Mottram, R. E., & Laurent, B. F. St. (2001). The importance of stretching the "X-factor" in the downswing of golf : The "X-factor stretch". In P. R. Thomas (Ed.), *Optimising performance in golf* (pp. 192–199). Brisbane, Australia: Australian Academic Press.

Cheetham, P. J., Rose, G. A., Hinrichs, R. N., Neal, R. J., Mottram, R. E., Hurrion, P. D., & Vint, P. F. U. S. O. C. (2008). Comparison of kinematic sequence parameters between amateur and professional Golfers. Chapter 4. In D. Crews & R. Lutz (Eds.), *Science and Golf V: Proceedings of the World Scientific Congress of Golf* (pp. 30–36). Phoenix, Arizona: The LPGA Foundation.

Chen, P.-H., Lin, C.-J., & Scholkopf, B. (2005). A tutorial on nu-support vector machines. *Applied Stochastic Models in Business and Industry*, *21*(2), 111–136. doi:10.1002/asmb.537.

Choi, C., & Kim, J. (2011). Synergy matrices to estimate fluid wrist movements by surface electromyography. *Medical Engineering & Physics*, *33*(8), 916–923. doi:10.1016/j.medengphy.2011.02.006.

Cole M, & Grimshaw P (2008a) Trunk muscle onset and cessation in golfers with and without low back pain. J*ournal of Biomechanics*, *41*:2829–2833. doi:10.1016/j.jbiomech.2008.07.004.

Cole, M. H., & Grimshaw, P. N. (2008b). Electromyography of the trunk and abdominal muscles in golfers with and without low back pain. *Journal of Science and Medicine in Sport*, *11*(2), 174–181. doi:10.1016/j.jsams.2007.02.006.

Cole, M. H, & Grimshaw, P. N. (2013). The crunch factor's role in golf-related low back pain. *The Spine Journal*, Oct 12.

Coleman, S. G. S., & Rankin, A. J. (2005). A three-dimensional examination of the planar nature of the golf swing. *Journal of Sports Sciences*, *23*(3), 227–234 . doi:10.1080/02640410410001730179.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. doi:10.1007/BF00994018 .

Cristianini, N., & Shawe-Taylor. (2000). *An introduction to support vector machines and other kernel-based learning methods*. Cambridge: Cambridge University Press.

De Luca, C. J. (1997). The use of surface electromyography in biomechanics. *Journal of Applied Biomechanics*, *13*, 135–63.

De Luca, C. J., Kuznetsov, M., Gilmore, L. D., & Roy, S. H. (2012). Inter-electrode spacing of surface EMG sensors: reduction of crosstalk contamination during voluntary contractions. *Journal of biomechanics*, *45*(3), 555–561. doi:10.1016/j.jbiomech.2011.11.010.

EGA. (2012). *Sistema de handicap EGA 2012-2016*. Epalinges/Lausanne: Edição Federação Portuguesa de Golfe.

Evans, K., Refshauge, K. M., Adams, R., & Aliprandi, L. (2005). Predictors of low back pain in young elite golfers: A preliminary study. *Physical Therapy in Sport*, *6*(3), 122–130. doi:10.1016/j.ptsp.2005.05.003.

Farber, A. J., Smith, J. S., Kvitne, R. S., Mohr, K. J., & Shin, S. S. (2009). Electromyographic analysis of forearm muscles in professional and amateur golfers. *The American Journal of Sports Medicine*, *37*(2), 396–401. doi:10.1177/0363546508325154.

Farina, D., Fattorini, L., Felici, F., & Filligoi, G. (2002). Nonlinear surface EMG analysis to detect changes of motor unit conduction velocity and synchronization. *Journal of Applied Physiology*, *93*(5), 1753–1763. doi:10.1152/japplphysiol.00314.2002.

Farina, D., Merletti, R., & Enoka, R. M. (2004). The extraction of neural strategies from the surface EMG. *Journal of Applied Physiology* (Bethesda, Md. : 1985), *96*(4), 1486–1495. doi:10.1152/japplphysiol.01070.2003.

Farmer, S. F., Halliday, D. M., Conway, B. A, Stephens, J. a, & Rosenberg, J. R. (1997). A review of recent applications of cross-correlation methodologies to human motor unit recording. *Journal of Neuroscience Methods*, *74*(2), 175–187.

Farrally, M. R., Cochran, A. J., Crews, D. J., Hurdzan, M. J., Price, R. J., Snow, J.T., & Thomas, P. R. (2003). Golf science research at the beginning of the twenty-first century. *Journal of Sports Science*, *21*(9), 753-65.

Fox, E., Lindsay, D., & Vandervoort, A. (2002). Musculoskeletal injury questionnaire for senior golfers. In *Science and Golf IV. Proceedings of the world scientific congress of golf* (pp. 88–99). St. Andrews, Scotland: Routledge Inc.

Gatt, C. J., Pavol, M. J., Parker, R. D., & Grabiner, M. D. (1998). Three-dimensional knee joint kinetics during a golf swing. Influences of skill level and footwear. *American Journal of Sports Medicine*, *26*(2), 285–294.

Glazebrook, M., Curwin, S., Islam, M., Kozey, J., & Stanish, W. (1994). Medial epicondylitis. An electromyographic analysis and an investigation of intervention strategies. *American Journal of Sports Medicine*, *22*(5), 674–679.

Gluck, G. S., Bendo, J. A, & Spivak, J. M. (2008). The lumbar spine and low back pain in golf: a literature review of swing biomechanics and injury prevention. *The Spine Journal*, *8*(5), 778–788. doi:10.1016/j.spinee.2007.07.388.

Haykin, S. (2001). *Redes neurais. Princípios e prática*. (2ª ed.). Santana: ARTMED.

Hellström, J., & Tinnmark, F. (2008). The association between stability and swing kinematics of skilled high school golfers. In D. Crews & R. Lutz (Eds.), *Science and Golf V: Proceedings of the World Scientific Congress of Golf* (pp. 37–43). Phoenix, Arizona: The LPGA Foundation.

Herbrich, R. (2002). *Learning kernel classifiers. Theory and algorithms*. Massachusetts: Massachusetts Institute of Technology.

Hodges, P. W., Richardson, C. A. (1996). Inefficient muscular stabilization of the lumbar spine associated with low back pain: a motor control evaluation of transversus abdominis. *Spine*, *21*, 2640–2650.

Hodges, P., & Bui, B. (1996). A comparison of computer-based methods for the determination of onset of muscle contraction using electromyography. *Electroencephalography and Clinical Neurophysiology*, *101*, 511–519.

Horton, J. F., Lindsay, D. M., & Macintosh, B. R. (2001). Abdominal muscle activation of elite male golfers with chronic low back pain. *Medicine and Science in Sports and Exercise*, *33*(10), 1647–1654.

Hosea, T. M. & Gatt, C. J. Jr. (1996). Back pain in golf. *Clinical Sports Medicine*, *15*, 37–53.

Hsu, C., Chang, C., & Lin, C. (2010). A practical guide to support vector classification. Retrieved from https://www.cs.sfu.ca/people/Faculty/teaching/726/spring11/svmguide.pdf

Hsu, C.-W. & Lin, C.-J. (2002). A Comparison of methods for multiclass support vector machines. *IEEE Transctions on Neural networks*, *13*(2), 415–425.

Hume, P. A, Keogh, J., & Reid, D. (2005). The role of biomechanics in maximising distance and accuracy of golf shots. *Sports medicine*, *35*(5), 429–449.

Jobe, F. W., Perry, J., & Pink, M. (1989). Electromyographic shoulder activity in men and women professional golfers. *American Journal of Sports Medicine*, *17*(6), 782-787.

Jobe, F., Moynes, D., & Antonelli, D. (1986). Rotator cuff functions during a golf swing. *American Journal of Sports Medicine*, *14*(5), 388–92.

Jöllenbeck, T. (2000). Methodological limitations of EMG-based bio-mechanical motion analysis. *ISBS – Conference Proceedings Archive, 18 International Symposium on Biomechanics in Sports*.

Kao, J., Pink, M., Jobe, F., & Perry, J. EMG analysis of the scapular muscles during a golf swing. *American Journal of Sports Medicine*, *23*(1), 19-23.

Keogh, J. W. & Hume, P. A. (2012). Evidence for biomechanics and motor learning research improving golf performance. *Sports Biomechanics*, *11*(2), 288-309.

Keogh, J. W., Marnewick, M. C., Maulder, P. S., Nortje, J. P., Hume, P. A., & Bradshaw, E. J. (2009). Are anthropometric, flexibility, muscular strength, and endurance variables related to clubhead velocity in low- and high-handicap golfers? *Journal of Strength & Conditioning Research*, *23*(6), 1841-50.

Konrad, P. (2005). *The ABC of EMG: a practical introduction to kinesiological electromyography*. USA: Noraxon Inc.

Kulkarni, S. R., & Harman, G. (2011). Statistical learning theory: a tutorial. *Wiley Interdisciplinary Reviews: Computational Statistics*, *3*(6), 543–556. doi:10.1002/wics.179.

Lee, A. S., Cholewicki, J., & Reeves, N. P. (2007). The effect of background muscle activity on computerized detection of sEMG onset and offset. *Journal of Biomechanics*, *40*(15), 3521–3526. doi:10.1016/j.jbiomech.2007.05.012.

Lei, M., & Meng, G. (2012). Nonlinear Analysis of Surface EMG Signals. In M. Lei & G. Meng (Eds.), *Computational intelligence in electromyography analysis - A perspective on current applications and future challenges*. (pp. 119–174). INTECH.

Li, N., Yang, D., Jiang, L., Liu, H., & Cai, H. (2012). Combined use of FSR sensor array and SVM classifier for finger motion recognition based on pressure distribution map. *Journal of Bionic Engineering*, *9*(1), 39–47.

Li, X., & Aruin, A. (2005). Muscle activity onset time detection using teager-kaiser energy operator. *Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference Shanghai, China* (Vol. 7, pp. 7549–7552). Shanghai, China. doi:10.1109/IEMBS.2005.1616259.

Lindsay, D. M., & Horton, J. F., & Paley, R. D. (2002). Trunk motion of male professional golfers using two different golf clubs. *Journal of Applied Biomechanics*, *18*, 366–373.

Lindsay, D. M., Horton, J. F., & Vandervoort, A. A. (2000). Review of injury characteristics, aging factors and prevention programmes for the older golfer. *Sports Medicine*, *30*(2): 89-103.

Lindsay, D., & Horton, J. (2002). Comparison of spine motion in elite golfers with and without low back pain. *Journal of Sports Sciences*, *20*(8), 599–605. doi:10.1080/026404102320183158.

Luxburg, U. Von, & Schölkopf, B. (2009). Statistical Learning Theory: Models, concepts, and results. In D. M. Gabbay, S. Hartmann, & J. Woods (Eds.), *Handbook of the history of logic*. (Vol. 10, pp. 651–706). Elsevier BV.

Marta, S., Silva, L., Castro, M. A., Pezarat-Correia, P., & Cabri, J. (2012). Electromyography variables during the golf swing: a literature review. *Journal of Electromyography and Kinesiology*, *22*(6): 803–813. doi:10.1016/j.jelekin.2012.04.002.

Marta, S., Silva, L., Vaz, J., Bruno, P., & Pezarat-correia, P. (2013). Electromyographic Analysis of Trunk Muscles during the Golf Swing Performed with Two Different Clubs. *International Journal of Sports Science & Coaching*, *8*(4): 779–787.

Marwan, N., & Kurths, J. (2004). Cross Recurrence Plots and their Applications. In Benton CV, (Ed.). *Mathematical physics research at the cutting edge* (pp. 101–139). Nova Science Publisher, Inc.

Marwan, N., & Thiel, M., & Nowaczyk, N. R. (2002). Cross recurrence plot based synchronization of time series. *Nonlinear Processes in Geophysics*, *2*(9), 325–331.

McGill, S. M. (2004). Linking latest knowledge of injury mechanisms and spine function to the prevention of low back disorders. *Journal of Electromyography and Kinesiology*, *14*(1), 43–47. doi:10.1016/j.jelekin.2003.09.012.

McGill, S., Chaimberg, J., Frost, D., & Fenwick, C. (2010). Evidence of a double peak in muscle activation to enhance strike speed and force: an example with elite mixed martial arts fighters. *Journal of Strength and Conditioning Research*, *24*(2), 348–357. doi:10.1519/JSC.0b013e3181cc23d5.Evidence.

McHardy, A., & Pollard, H. (2005a). Lower back pain in golfers: a review of the literature. *Journal of Chiropractic Medicine*, *4*(3), 135–143. doi:10.1016/S0899-3467(07)60122-0.

McHardy, A., & Pollard, H. (2005b). Muscle activity during the golf swing. *British Journal of Sports Medicine*, *39*(11), 799–804. doi:10.1136/bjsm.2005.020271.

McHardy, A. J., Pollard, H. P., & Luo, K. (2007). Golf-related lower back injuries: an epidemiological survey. *Journal of Chiropractic Medicine*, *6*(1), 20–26. doi:10.1016/j.jcme.2007.02.010.

McLean, J. (1992). Widen the gap. *Golf Magazine*. *De*c.

McLean, J. (2008). Triple X-factor. *Golf Digest*, *Jan*.

Merlo, A., & Campanini, I. (2010). Technical Aspects of Surface Electromyography for Clinicians, *The Open Rehabilitation Journal*, *3*, 98–109.

Mezzarane, R. A., & Kohn, A. F. (2009). A method to estimate EMG crosstalk between two muscles based on the silent period following an H-reflex. *Medical Engineering & Physics*, *31*(2009), 1331–1336. doi:10.1016/j.medengphy.2009.09.005.

Micera, S., Sabatini, A. M., & Dario, P. (1998). An algorithm for detecting the onset of muscle contraction by EMG signal processing. *Medical Engineering & Physics*, *20*(3), 211–215.

Mogk, J. P. M., & Keir, P. J. (2003). Crosstalk in surface electromyography of the proximal forearm during gripping tasks. *Journal of Electromyography and Kinesiology*, *13*, 63–71. doi:10.1016/S1050-6411(02)00071-8.

Morey-Klapsing, G., Arampatzis, A., & Brüggemann, G. P. (2004). Choosing EMG parameters: comparison of different onset determination algorithms and EMG integrals in a joint stability study. *Clinical Biomechanics* (Bristol, Avon), *19*(2), 196–201.

Müller, K.-R., Mika, S., Rätsch, G., Tsuda, K., & Schölkopf, B. (2001). An introduction to kernel-based learning algorithms. *IEEE TRANSACTIONS on Neural Networks*, *12*(2), 181–201.

Myers, J., Lephart, S., Tsai, Y.-S., Sell, T., Smoliga, J., & Jolly, J. (2008). The role of upper torso and pelvis rotation in driving performance during the golf swing. *Journal of Sports Sciences*, *26*(2), 181–188. doi:10.1080/02640410701373543.

Nesbit, S. M., & McGinnis, R. (2009). Kinematic analyses of the golf swing hub path and its role in golfer/club kinetic transfers. *Journal of Sports Science & Medicine*, *8*(2), 235–246.

Nesbit, S. M., & Serrano, M. (2005). Work and power analysis of the golf swing. *Journal of Sports Science & Medicine*, *4*, 520–533.

Oskoei, M. A., & Hu, H. (2007). Myoelectric control systems—A survey. *Biomedical Signal Processing and Contro*l, *2*(4), 275–294. doi:10.1016/j.bspc.2007.07.009.

Palmer, J. L., Young, S. D., Fox, E., Lindsay, D. M., & Vandervoort, A. A. (2003) Senior Recreational Golfers: a survey of musculoskeletal conditions, playing characteristics, and warm-up patterns. *Physiotherapy Canada*, *55*(2), 79–85.

Penner, A. R. (2003). The physics of golf. *Reports on Progress in Physics*, *66*(2), 131–171.

Pezarat-Correia, P. & Mil-Homens, P. (2004). *A Electromiografia no estudo do movimento humano*. Cruz Quebrada: Edições FMH.

Pink, M., Jobe, F., & Perry, J. (1990). Electromyographic analysis of the shoulder during the golf swing. *American Journal of Sports Medicine*, *18*(2), 137–140.

Pink, M., Perry, J., & Jobe, F. (1993). EMG analysis of the trunk in golfers. *American Journal of Sports Medicine*, *21*(3), 385–388.

Raez, M. B. I., Hussain, M. S., & Mohd-Yasin, F. (2006). Techniques of EMG signal analysis: detection, processing, classification and applications. *Biological Procedures Online*, *8*(1), 11–35.

Renkawitz, T., Boluki, D., & Grifka, J. (2006). The association of low back pain, neuromuscular imbalance, and trunk extension strength in athletes. *The spine journal*, *6*(6), 673–83. doi:10.1016/j.spinee.2006.03.012.

Richardson, A. K., Hughes, G., & Mitchell, A. C. S. (2012). Center of pressure excursion during the golf putting stroke in low, mid and high handicap golfers. *International Journal of Golf Science*, *1*, 127–139.

Rojas, R. (1996). *Neural networks: A systematic introduction*. Berlin: Springer-Verlag.

Schölkopf, B., & Smola, A. (2002). *Learning with kernels*. Massachusetts: Massachusetts Institute of Technology. Retrieved from http://dip.sun.ac.za/~hanno/tw796/lesings/mlss06au_scholkopf_lk.pdf

Scholkopf, B., Burges, C., & Vapnik, V. (1995). Extracting Support Data for a Given Task. In U. M. Fayyad & R. Uthurusamy (Eds.), *First International Conference on Knowledge Discovery & Data Mining*. (pp. 252–257). Menlo Park: AAAI Press.

Schölkopf, B., Smola, A. J., Williamson, R. C., & Bartlett, P. L. (2000). New support vector algorithms. *Neural Computation, 12,* 1207–1245.

Scholkopf, B., Smola, A., & Muller, K. R. (1996). *Nonlinear component analysis as a kernel eigenvalue problem* (pp. 1–18). Tübingen.

Sherman, C. A., & Finch, C. F. (2000). Preventing injuries to competitive and recreational adult golfers: what is the evidence? *Journal of Science and Medicine in Sport / Sports Medicine Australia*, *3*(1), 65–78.

Silfies, S. P., Mehta, R., Smith, S. S., & Karduna, A. R. (2009). Differences in feedforward trunk muscle activity in subgroups of patients with mechanical low back pain. *YAPMR*, *90*(7), 1159–1169. doi:10.1016/j.apmr.2008.10.033.

Silva, L., Marta, S., Vaz, J., Fernandes, O., Castro, M. A., & Correia-Pezarat, P. (2013). Trunk muscle activation during golf swing: Baseline and threshold. *Journal of Electromyography and Kinesiology*, *23*(5), 1174–1182. doi:10.1016/j.jelekin.2013.05.007.

Solnik, S., Rider, P., & Steinweg, K. (2010). Teager–Kaiser energy operator signal conditioning improves EMG onset detection. *European journal of Applied Physiology*, *2010*(110), 489–498. doi:10.1007/s00421-010-1521-8.

Somjarod, M., Tanawat, V., & Weerawat, L. (2011). The Analysis of knee joint movement during golf swing in professional and amateur golfers. *International Science Index. World Academy of Science, Engineering and Technology*, *53*, 495–498.

Staude, G., Flachenecker, C., Daumer, M., & Wolf, W. (2001). Onset detection in surface electromyographic signals: A systematic comparison of methods. *Journal on Applied Signal Processing*, *2001*(2), 67–81.

Sugaya, H., Tsuchiya, A., Moriya, H., Morgan, D. A., & Banks, S. A. (1999). Low back injury in elite and professional golfers: An epidemiologic and radiographic study. In A. J. Cochran & M.R. Fairally (Eds.), *Science and Golfl III: Proceedings of the World Scientific Congress of Golf* (pp. 83-91). Champaign, IL: Human Kinetics.

Thériault, G., & Lachance, P. (1998). Golf injuries. An overview. *Sports Medicine*, *26*(1), 43–57. doi:10.2165/00007256-199826010-00004.

Tolambiya, A., Thomas, E., Chiovetto, E., Berret, B., & Pozzo, T. (2011). An ensemble analysis of electromyographic activity during whole body pointing with the use of support vector machines. *PloS one*, *6*(7), e20732. doi:10.1371/journal.pone.0020732.

Tsai, Y.-S., Sell, T. C., Smoliga, J. M., Myers, J. B., Learman, K. E., & Lephart, S. M. (2010). A comparison of physical characteristics and swing mechanics between golfers with and without a history of low back pain. *Journal of Orthopaedic and Sports Physical Therapy*, *40*(7), 430–438. doi:10.2519/jospt.2010.3152.

Tyler, A., & Karst, G. (2004). Timing of muscle activity during reaching while standing: systematic changes with target distance. *Gait & Posture*, *20*, 126–133. doi:10.1016/j.gaitpost.2003.07.001.

Vad, V. B., Bhat, A. L., Basrai, D., Gebeh, A., Aspergren, D. D., & Andrews, J. R. (2004). Low back pain in professional golfers: the role of associated hip and low back range-of motion deficits. *American Journal of Sports Medicine*, *32*(2), 494–497. doi:10.1177/0363546503261729.

Vaisman, L., Zariffa, J., & Popovic, M. R. (2010). Application of singular spectrum-based change-point analysis to EMG-onset detection. *Journal of Electromyography and Kinesiology*, *20*(4), 750–760. doi:10.1016/j.jelekin.2010.02.010.

Van Boxtel, G. J. M., Geraars, L. H. D., Van Den Berg-Lenssen, M. M. C., & Brunia, C. H. M. (1993). Detection of EMG onset in ERP research. *Psychophysiology*, *30*(4), 405–412. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1469-8986.1993.tb02062.x/abstract.

Vapnik, V. N. (1999). An Overview of Statistical Learning Theory. *IEEE Transactions on Neural Networks*, *10*(5), 988–999.

Watkins, R. G., Uppal, G. S., Perry, J., Pink, M., & Dinsay, J. M. (1996). Dynamic electromyographic analysis of trunk musculature in professional golfers. *The American Journal of Sports Medicine*, *24*(4), 535–538.

Winter, D., Fuglevand, A., & Archer, S. (1994). Crosstalk in surface electromyography: Theoretical and practical estimates. *Journal of Electromyography and Kinesiology*, *4*(1), 15–26. doi:10.1016/1050-6411(94)90023-X.Crosstalk.

Worsfold, P., Smith, N., & Dyson, R. (2008). Weight Transfer Time Moderation by Golf Handicap. Chapter 9. In D. Crews & L. Rafer (Eds.), *Science and Golf V: Proceedings of the World Scientific Congress of Golf* (pp. 67–73). Phoenix, Arizona: The LPGA Foundation.

Xie, H.-B., Wang, Z.-Z., Huang, H., & Qing, C. (2003). Support vector machine in computer aided clinical electromyography. In *Machine Learning and Cybernetics*, *2003 International Conference* (Volume 2), 2-5 Nov, (pp. 1106–1108). doi:10.1109/ICMLC.2003.1259649.

Zecca, M., & Micera, S. (2002). Control of multifunctional prosthetic hands by processing the electromyographic signal. *Critical Reviews*, *30*(4-6), 459–485.

Zheng, N., Barrentine, S. W., Fleisig, G. S., Andrews, J. R., & States, U. (2008). Kinematic analysis of swing in pro and amateur golfers. *International Journal of Sports Medicine*, *29*(6), 487–93. doi:10.1055/s-2007-989229.

CAPÍTULO II

2. Metodologia

2.1. Apresentação do Problema

Do enquadramento teórico apresentado são realçados alguns paradigmas de investigação que servem de base de sustentação à pesquisa desenvolvida, nomeadamente, em relação à construção de modelos de aprendizagem automática. O primeiro está relacionado com a determinação do *onset*. Se por um lado, são utilizados algoritmos de deteção automática por se considerar a inspeção visual subjetiva (Jöllenbeck, 2000), por outro, sempre que os mesmos são aplicados é realizada correção por inspeção visual, sendo estes testados recorrendo a inspeção visual. Outro aspeto que encontramos no estudo de parâmetros temporais refere-se às várias propostas de algoritmos de deteção. O *onset* determinado está totalmente dependente dos parâmetros incluídos nesse mesmo algoritmo (ex.: nos algoritmos por limiar, o desvio padrão) e com a relação destes com a variabilidade do sinal, como o declive de incremento da amplitude (Allison, 2003) e o rácio sinal-ruído (Hodges & Bui, 1996; Solnik et al., 2010), nomeadamente pela influência da linha de base (Hodges & Bui, 1996; Lee et al., 2007; Staude et al., 2001). Para além de comprometer a reprodutibilidade (Vaisman et al., 2010) também conduz a limitações sobre o fenómeno fisiológico a estudar por tendencialmente a discussão basear-se em questões metodológicas em detrimento de análises fenomenológicas. Os algoritmos devem ser ajustados a habilidades motoras dinâmicas, visto o sinal conter informação que poderá estar associada a várias ações dentro da mesma habilidade motora. Um exemplo deste aspeto é o fenómeno do pico duplo (McGill et al., 2010), que depende de ajustamento entre a transição de uma fase de preparação em relação a outra de execução e ao momento de impacto. Neste último estudo, os autores fizeram corresponder ao sinal fotografias do movimento de forma a descodificar o comportamento motor em movimento de artes marciais.

No *swing* do golfe, encontramos a preocupação de discriminar e caracterizar a golfistas com lombalgia no domínio dos parâmetros temporais (Cole & Grimshaw, 2008; Horton, Lindsay, & Macintosh, 2001) baseando-se o momento temporal (*onse*t ou *offset*) nas principais fases que compõem esta habilidade motora. Porém, é na dinâmica e essencialmente na cinemática que são destacadas características que diferem populações com e sem lombalgia (Gluck et al., 2008; Lindsay & Horton, 2002; Tsai et al., 2010; Vad et al., 2004), quando esta queixa física apresenta maior incidência na população golfistas (Evans, Refshauge, Adams, & Aliprandi, 2005; Gluck et al., 2008; McHardy, Pollard, & Luo, 2007; McHardy & Pollard, 2005). Dentro dos parâmetros utilizados para categorizar a lombalgia destacam-se o *X-factor* (Cheetham et al., 2001; Hume et al., 2005) e o *Crunch-factor* (Cole & Grimshaw, 2013). Para além da lombalgia, também o Hc é referenciado como constrangimento que influencia a execução do *swing* (Worsfold et al., 2008; Zheng et al., 2008). O estado da arte sobre a quantificação do sinal EMG no domínio no tempo e respetiva utilização do sinal para classificar diferentes constrangimentos em habilidades motoras evidencia carências, nomeadamente no *swing*.

Mediante o mencionado anteriormente surge o pensamento de que seria ideal que o computador conhecesse as características das séries temporais que reportam o EMG, chegando-se à aprendizagem automática com os seguintes pontos:

(1) Exigência de algoritmos automáticos ao invés de inspeção visual por esta ser subjetiva, quando a performance dos algoritmos automáticos é testada recorrendo a inspeção visual;

(2) Carência na discussão sobre os fenómenos fisiológicos privilegiando estudos metodológicos, não existindo um algoritmo *standard*, estando presente a falta de reprodutibilidade entre estudos;

(3) Estudos metodológicos sobre a deteção do *onset* são realizados em contrações isométricas, sendo posteriormente realizadas inferências em ações dinâmicas;

(4) Lacunas na discriminação de constrangimentos em tarefas motoras dinâmicas complexas e respetiva análise sobre as diferentes estratégias neuromusculares que contemple o comportamento não linear do sinal EMG.

(5) Necessidade de instrumento que permita a categorização das abordagens anteriores, nomeadamente em relação à forma como poderá ser quantificada a lombalgia, associando-a à prática desportiva.

Face aos pontos expostos surge o problema: ***pode ser extraída informação do sinal EMG que quantifique fenómenos temporais e permita classificar os principais constrangimentos associados a uma habilidade motora dinâmica?*** Para o presente propósito a habilidade motora dinâmica refere-se ao *swing* no golfe e os constrangimentos associados a esta modalidade são o handicap e a lombalgia. O método matemático de classificação considerado para a resolução desta problemática corresponde às máquinas de vetor suporte.

2.2. Objetivos

Após a exposição do problema torna-se possível direcionar um conjunto de questões que daí derivam constituindo os seguintes objetivos:

a) Validar para a população portuguesa um instrumento que permita categorizar a população golfista quanto aos hábitos de atividade física e incidência de lesões, com ênfase na relação entre a lombalgia e a prática de golfe;

b) Verificar por métodos baseados em algoritmos de deteção tradicionais, a influência de diferentes referências de linha de base (principal fonte de ruído na deteção do *onset*) quantificando a percentagem da amplitude máxima a que essa deteção ocorre;

c) Identificar na execução do *swing* fenómenos temporais que possam ser considerados como relevantes para o estudo de padrões EMG que sirvam de suporte para decisões de estudo futuras;

d) Conhecer a metodologia inerente à aplicação de aprendizagem automática, especificamente das SVM, ao sinal EMG, como extração de segmentação, extração de *features* e parâmetros inclusos no algoritmo;

e) Caracterizar qual/quais o(s) melhore(s) conjunto(s) de *feature(s)* para a classificação SVM na deteção do *onset* tido anteriormente como o mais relevante;

f) Conhecer os intervalos que caracterizam os melhores parâmetros para o modelo SVM em termos de separabilidade e definição da função *kernel*;

g) Perceber se o aumento da dimensão da amostra, ao considerar todos os níveis de handicap, altera a qualidade do classificador na deteção do *onset*;

h) Conhecer dos músculos estudados, quais aqueles que melhor discriminam a existência de constrangimentos (handicap e lombalgia) e em que medida têm poder discriminativo (poderá ser fator quantitativo de diferentes estratégias neuromusculares);

i) Analisar a utilidade de diferentes relações de *features* de recorrência no estudo discriminatório da ativação neuromuscular mediante constrangimentos da tarefa (handicap e lombalgia);

j) Perceber a relação entre a precisão dos modelos SVM e o número de vetores suporte, ou seja, a maior ou menor facilidade com que o classificador conseguiu realizar o seu propósito.

2.3. Participantes

A amostra deste estudo pode ser dividida em duas categorias de participação, quanto à validação do questionário inicial, ou pela participação em recolhas laboratoriais. Nem todos os sujeitos que responderam ao questionário estiveram presentes em laboratório. A dimensão amostral ao longo dos estudos realizados em laboratório variou mediante a caracterização ou especificidade da amostra. A dimensão da amostra no estudo 1 foi de 61 sujeitos, dos quais foram considerados para os restantes estudos exploratórios 29 sujeitos.

No estudo 2 participaram 8 golfistas de Hc médio (idade = 52.0 ± 7.4 anos; Hc = 15.7 ± 3.2). Neste caso os golfistas foram instruídos para realizarem 5 repetições do *swing* com o taco *pith* (<100m), e outras 5 com o ferro 4 (>150m) em sequência alternada, totalizando 80 sinais do *swing* para estudo.

No quarto estudo foram recrutados 12 golfistas destros tendo sido divididos em dois grupos mediante o Hc: 6 com Hc baixo (Hc = 1.4 ± 2.5 < 5) e seis com Hc alto (Hc = 24.6 ± 4.2 > 18). Neste estudo foram consideradas quatro repetições do *swing* para cada jogador, executadas com o ferro 7 (>100 m). As amostras posteriormente sujeitas a análise correspondem a instantes, tendo sido recolhidos 1.001 instantes por repetição.

Desta forma, após a extração de *features*, foram totalizados 24024 instantes para análise (6 indivíduos x 1001 instantes por julgamento x 4 ensaios) em cada Hc e 48048 quando analisados modelos que incluam os dois grupos de Hc.

Quanto ao estudo 5, houve a participação de 21 golfistas, sendo posteriormente divididos por grupos consoante o Hc e a sintomatologia de lombalgia após a realização de uma volta de 18 buracos. Um dos sujeitos foi comum a ambas as análises (Hc e lombalgia). Para a análise sobre a lombalgia obteve-se a participação de 12 sujeitos (idade = 52.5 ± 12.1 anos; Hc = 15.1 ± 12.0), dos quais cinco formaram o grupo com lombalgia e sete o grupo sem lombalgia. A homogeneidade entre os grupos foi verificada para a idade $(U = 9.0; p = 0.189)$, Hc $(U = 15.5; p = 0.785)$, tempo de prática $(U = 9.0; p = 0.246)$, massa corporal $(U = 9.0; p = 0.699)$, e altura $(U = 9.0; p = 0.909)$. A discriminação de lombalgia foi realizada a partir da validação para a população portuguesa (estudo 1) do questionário *Musculoskeletal Injury Questionnaire for Senior Golfers* (Fox, Lindsay, & Vandervoort, 2002). Os participantes selecionados para o grupo sem lombalgia foram aqueles que responderam 0% na questão sobre a presença de desconforto lombar após 18 buracos e sem qualquer diagnóstico de lesão músculo-esquelética associada a lombalgia. Os participantes com mais de 65% de lombalgia integraram o outro grupo. Para os grupos Hc foram considerados 10 participantes, 5 de baixo Hc (Hc = 0.7 ± 2.2 < 5) e outros 5 de alto Hc (Hc = 24.3 ± 4.6 ≥ 18).

Para a construção dos grupos segundo Hc foram tidas em consideração as recomendações da *European Golf Association* (EGA, 2012).

Antes de qualquer procedimento experimental, os indivíduos foram autorizados a realizar algumas repetições, permitindo a realização de aquecimento e uma melhor adaptação para a tarefa a executar. As repetições do *swing* foram realizadas em cima de um tapete de relva artificial de golfe com características de alta absorção. Nenhum dos sujeitos apresentou alguma contraindicação no momento em que realizaram os *swings*. Todos os procedimentos foram explicados e foi assinado o respetivo termo de consentimento (Apêndice 1). Este estudo foi aprovado pelo Conselho de Ética da Faculdade de Motricidade Humana (Universidade de Lisboa) (Apêndice 4).

2.4. Procedimentos

2.4.1. Procedimentos para recolha e processamento dos sinais EMG

Os sinais EMG foram recolhidos com superfícies de deteção Al/AgCl AMBU® BlueSensor N (AMBU, Ballerup, Dinamarca) com formato em disco com 10 mm de diâmetro ligados a elétrodos ativos (Plux, Lisboa, Portugal) e a sistema de telemetria bioPLUX® research 2010 (Plux, Lisboa, Portugal) Os sinais EMG foram amplificados com uma banda passante entre 10 e 500 Hz, capacidade de rejeição do sinal comum de 110 dB e impedância de *input* superior a 100 MW. No processo de digitalização foi usada uma frequência de amostragem de 1000 Hz. Após armazenamento, os sinais EMG foram submetidos a filtragem digital com uma banda passante entre 10 e 490 Hz e retificados. Quando realizada suavização (estudo 2 e estudo 4) foi aplicada uma frequência de corte de 12 Hz com filtro *Butterworth* de 4ª ordem. Os sinais EMG foram normalizados pela máxima amplitude do conjunto de repetições para cada sujeito. Todos os sinais EMG em bruto foram sujeitos, previamente ao início do processo de processamento, a inspeção visual de forma a aferir a sua qualidade para integrar o estudo. O tratamento e processamento dos dados foram executados em MATLAB ® software V.R2010a (Mathworks Inc., Natick Massachusetts, USA).

As superfícies de deteção foram colocadas na zona de maior proeminência do ventre muscular em contração após a adequada preparação da pele (depilação, abrasão e limpeza com álcool). Os músculos analisados foram medidos bilateralmente com exceção do reto femoral (RF), vasto interno (VM), vasto externo (VL) e oblíquo externo (EO), os quais foram medidos apenas do lado esquerdo no estudo 4. No estudo 5, o RF e o EO foram medidos bilateralmente, não tendo sido considerados o grande glúteo direito (GG) e a massa comum (MC) do lado direito. O reto abdominal apenas foi considerado no 2º estudo, tendo sido analisado bilateralmente. Os locais de colocação dos elétrodos são os que a seguir se descrevem (Hermens et al., 1999; Horton et al., 2001):

- Reto abdominal (RA – *rectus abdominis*): lateral a 3 cm do umbigo;
- Bicípite femoral (BF – *bíceps femoris*): a 50% da linha que une a tuberosidade isquial e o epicôndilo lateral da tíbia;

- Semitendinoso (ST – *semitendinosus*): a 50% da linha que une a tuberosidade isquial e o epicôndilo medial da tíbia;
- Grande glúteo (GM – *gluteus maximus*): a 50% da linha que liga o sacro ao grande trocânter no fémur;
- Massa comum (ES – *erector spinae*): 3 cm lateralmente do processo espinhoso da L3;
- Reto femoral (RF – *rectus femoris*): a 50% da linha que une a espinha ilíaca antero-superior ao bordo superior da patela;
- Vasto interno (VM – *vastus medialis*): a 80% da linha que une espinha ilíaca antero-superior ao espaço articular frontal do bordo anterior do ligamento medial;
- Vasto externo (VL – *vastus lateralis*): a 2/3 da linha que liga a espinha ilíaca antero-superior ao bordo lateral da patela;
- Oblíquo externo (EO – *external oblique*): 15 cm lateralmente ao umbigo.

O elétrodo terra foi colocado sobre o manúbrio.

2.4.2. Procedimentos para recolha e processamento dos dados cinemáticos

Três câmaras de alta velocidade de marca Basler A602fc (Basler Vision Technologies, Ahrensburg, Germany) de 100 Hz foram colocadas em localização anterior, posterior e superior oblíqua. Uma quarta câmara Casio Ex-FH20 (Casio, Tokyo, Japan) de 1000 Hz foi colocada frontalmente à bola com o intuito de determinar o momento de impacto do taco na bola. Duas marcas refletoras foram colocadas nos tacos (superior e inferior) de acordo com o proposto por Horton et al. (2001) para posterior identificação das três principais fases do *swing* (Bechler et al., 1995; Pink et al., 1993; Watkins et al., 1996): (1) *Backswing* – do início do *swing* (*address*) até ao topo, sendo a fase de preparação; (2) *Downswing* – do topo do *backswing* até ao momento do impacto, sendo a fase de execução; (3) *Follow-Through* – do impacto até ao final do *swing*, pelo que se caracteriza pela fase final que resulta após o impacto. Os vídeos das repetições realizadas foram captados e posteriormente processados com o *software* SIMI 3D Motion system (SIMI Reality Motion System GmbH, Unterschleissheim, Germany).

2.4.3. Segmentação e extração de "*features*"

2.4.3.1. Segmentação

Como segmentação podemos definir o procedimento executado ao longo de uma série temporal que extrai um conjunto de amostras, sendo contabilizado tanto pelo número de amostras que dependerá da taxa de amostragem como por uma janela temporal, cuja unidade costuma ser em milissegundos. Neste ponto incluímos procedimentos referentes aos estudos II, IV e V, apesar de ter sido no estudo IV onde foi realizado o procedimento de segmentação propriamente dito.

No estudo II, tendo em consideração as fases de preparação (*backswing*) e execução (*downswing*) e a linha de base, foram utilizados dois métodos para a deteção de *onset*. No método A, a média de referência para o cálculo do limiar foi determinada pela atividade da linha de base no sinal EMG entre duas contrações voluntárias máximas. O método B envolveu a atividade EMG da própria repetição de recolha do *swing*, sendo recolhida antes dos 500 ms que precedem o início do *backswing* uma janela temporal de 1000 ms. Posteriormente foram acrescentados 3 desvios-padrão acima da média da linha de base. Adicionalmente foram colocadas duas restrições para ambos os métodos: 1. a deteção do *onset* começa 150 ms antes do *backswing*; 2. a deteção de um segundo *onset* começa 150 ms antes do *downswing*. Uma janela de 50 amostras (50 ms) move-se até que a derivada positiva e o limiar estabelecido seja atingido. A referência para o cálculo do *onset* foi o impacto.

Através da revisão de literatura do estudo 3 (Capítulo V), foi possível conhecer a metodologia associada ao controlo mio-elétrico, nomeadamente no que se refere à segmentação do sinal e extração de *features*. Das conclusões retiradas verificou-se que as *features* devem ser extraídas a partir de vários segmentos numa janela temporal para preservar a estrutura do sinal e não de amostras individuais, pois representaria uma perda significativa de informação (Hudgins, Parker, & Scott, 1993). Por sua vez, Oskoei & Hu (2007) referem que o reconhecimento de padrões, após os procedimentos de amplificação, filtragem e digitalização do sinal, seguem quatro passos fundamentais: segmentação de dados, extração de entradas (*features*), classificação e a função do controlador (Figura 3).

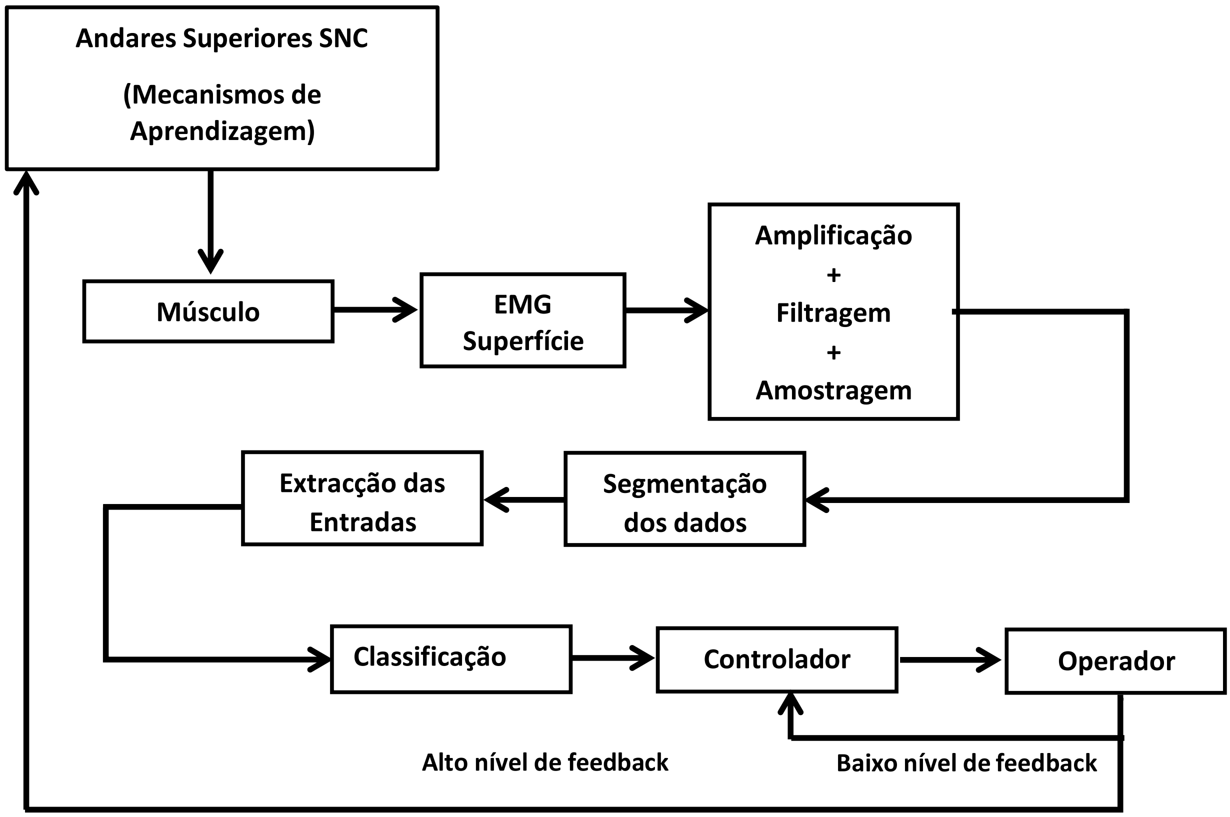


Figura 3 – Fases de tratamento de EMG no reconhecimento de padrões (adaptado de Oskoei & Hu, 2007. p. 277)

Esta esquematização serve de base para os estudos 4 (Capítulo VI) e 5 (Capítulo VII), nomeadamente para o estudo 4. Neste estudo foi considerada a janela temporal de 200 amostras (200 ms, pois a taxa de amostragem foi de 1000 Hz). Esta decisão foi suportada pela contribuição de Oskoei e Hu (2008) que consideraram 200 ms como a fronteira entre utilização de segmentação contínua ou *overlapped*. Apesar da segmentação *overlapped* estar associada a valores superiores a 200 ms, foi considerado este valor por ser uma janela fronteira entre os dois tipos de segmentação, como também, devido ao objetivo deste estudo reportar-se à deteção do *onset*. O *lag* considerado entre cada janela foi de 5 ms. Ou seja, de 5 ms em 5 ms procedeu-se à extração de um vetor de 200 amostras.

No último estudo, o fluxograma pode ser visualizado no Capítulo VII, o fracionamento correspondeu ao *swing* na totalidade e a divisão de cada fase, em conformidade com o retratado no ponto 2.4.2. neste capítulo.

2.4.3.2. Extração de "features" no domínio do tempo

Após a realização da segmentação do sinal procedeu-se à extração de *features* tendo sido estudadas seis tipos de *features* no domínio do tempo (Phinyomark, Limsakul, & Phukpattaranont, 2009; Phinyomark, Phukpattaranont, & Limsakul, 2012): valor absoluto médio (MAV – *mean absolute value*), comprimento do formato da onda (WL – *waveform length*), diferença absoluta do desvio-padrão (DASDV - *difference absolute standard deviation value*), variância do EMG (VAR – *variance of EMG*), integral do EMG (IEMG - *integrated EMG*) e o detetor logarítmico (LOG - *Log detector*).

⇒ Valor absoluto médio (MAV)

Considerando um conjunto de amostras (instantes) $\{x_1, x_2, \dots x_N\}$ dentro de uma janela temporal de dimensão N, corresponde ao valor absoluto médio do sinal dessa janela. Neste caso, a janela temporal definida corresponde a determinada parcela de sinal em torno do instante considerado.

$$MAV = \frac{1}{N}\sum_{i=1}^{N}|x_i| \tag{16}$$

⇒ Comprimento do formato da onda (WL)

Corresponde ao comprimento cumulativo da forma de onda durante determinado segmento de tempo. O WL está relacionado com a amplitude da forma de onda, tendo em conta a frequência e o tempo e pode ser visto como uma espécie de diferença finita de aproximação ao valor absoluto da primeira derivada (sem normalização pelo intervalo de tempo), sendo dada por:

$$WL = \sum_{i=1}^{N-1} |x_{i+1} - x_i| \tag{17}$$

⇒ Diferença absoluta do desvio-padrão (DASDV)

A diferença absoluta do desvio-padrão corresponde à diferença média dada pela norma euclidiana das diferenças entre valores consecutivos, normalizando-as pelo comprimento do intervalo de tempo.

$$DASDV = \sqrt{\frac{1}{N} \sum_{i=1}^{N-1} |x_{i+1} - x_i|^2} \tag{18}$$

⇒ Variância do EMG (VAR)

Foi uma das primeiras formas de quantificar o sinal EMG, tal como através do MAV. Usualmente é definida como o valor médio do quadrado do desvio-padrão do sinal. Tendo em conta que a média do sinal EMG tende para zero, pode ser calculada recorrendo à soma dos quadrados dos valores do sinal a dividir pela dimensão do intervalo menos 1.

$$VAR = \frac{1}{N-1} \sum_{i=1}^{N} x_i^2 \tag{19}$$

⇒ Integral EMG (IEMG)

O integral EMG é definido como a área que se encontra sob a curva do sinal de EMG retificado (valores absolutos), ou seja, consiste no integral do valor absoluto do sinal de EMG bruto. Refere-se à determinação da área delimitada pela curva do sinal retificado (Pezarat-Correia & Mil-Homens, 2004).

$$IEMG = \sum_{i=1}^{N} |x_i| \tag{20}$$

⇒ Detetor logarítmico (LOG)

É um detetor de não-linearidade que proporciona uma estimativa da força de contração do músculo com base no logaritmo.

$$LOG = e^{\frac{1}{N}\sum_{i=1}^{N}\log(|x_i|)} \tag{21}$$

Após a extração de *features* foram construídas as respetivas matrizes para introdução nas SVM.

2.4.3.3. Análise de quantificação de recorrência (RQA-recurrence quantification analysis)

Esta abordagem corresponde ao quinto estudo (Capítulo VII). Os gráficos de recorrência (RP – *recurrence plots*) correspondem a uma técnica matemática de análise não linear que permite visualizar a recorrência contida dentro de sistemas dinâmicos, proposto por Eckmann, Kamphorst e Ruelle (1987), baseado no teorema de recorrência de Poincaré. É assumido que para um conjunto de sistemas dinâmicos, por mais arbitrários que sejam os fenómenos ocorridos, as trajetórias retornam um número infinito de vezes com uma certa proximidade de quase todos os pontos iniciais. A este fenómeno chama-se recorrência de estados (Thiel, Romano, & Kurths, 2004). Considerando um conjunto de vetores $\{\vec{x_i}\}_{i=1}^{N}$ pertencentes a um determinado sistema, é descrito como o conjunto de séries que representam a trajetória no denominado espaço fase (*phase space*). O espaço fase é uma representação gráfica na qual o tempo está oculto, com um atractor que corresponde ao conjunto para onde tendem as trajetórias (Marwan, Carmenromano, Thiel, & Kurths, 2007). Desta forma, os gráficos de recorrência vão assumir uma matriz composta por valores “0” e “1”.

$$R_{i,j} = \begin{cases} 1: \vec{x}_i \approx \vec{x}_j \\ 0: \vec{x}_i \not\approx \vec{x}_j \end{cases} \quad i, j = 1 \ldots, N \tag{22}$$

Se $R_{i,j} = 1$, o estado é recorrente, sendo marcado graficamente com um ponto preto. Com $R_{i,j} = 0$, o estado não é recorrente, pelo que, a marcação no gráfico de recorrência corresponde a um ponto branco. Um gráfico de recorrência de uma série temporal com N pontos é uma matriz $N \times N$ constituída por pontos pretos e brancos. Assim, as RP para séries temporais com $t = i\Delta t$ é definida por:

$$R_{i,j} = \Theta\left(\varepsilon_i - \left\|\vec{x}_i - \vec{x}_j\right\|\right), \tag{23}$$

onde ε_i é a distância do instante i ou o raio de vizinhança, $\|\cdot\|$ é a norma de vizinhança e $\Theta(x)$ corresponde à função Heaviside. Neste estudo foi utilizada a norma Euclidiana e o raio de vizinhança $\varepsilon_i = 10\%$ da média do espaço fase (Marwan et al., 2007).

Um passo importante na análise não linear é a reconstrução do espaço fase, requerendo dois parâmetros fundamentais, a incorporação de dimensionalidade m e o desfasamento τ, cujo estrangeirismo será adotado denominando-se por *delay*. O m refere-se ao dimensionar com que a representação dinâmica deverá ser projetada, o que corresponde a um número mínimo de variáveis de modo a formar um espaço fase adequado a partir de uma determinada série de temporal. Um τ adequado entre pontos de tempo sequenciais com d-dimensão é um valor inteiro selecionado para minimizar a interação entre os pontos das séries de temporais (Webber & Zbilut, 2005). Usando estes valores de *embedding*, um vetor no espaço fase é reconstruído por (Marwan et al., 2007):

$$\vec{x}_i = \sum_{k=0}^{m-1} \xi_{i+k\tau} + \vec{e}_k \tag{24}$$

Para este estudo foi considerado a mediana do valor de τ encontrado por *mutual information*, tendo sido posteriormente ajustado o valor de m usando a técnica *false nearest neighbors*. Foi consistente $\tau = m = 2$.

As medidas RQA foram executadas recorrendo ao pacote MATLAB® ToolBox 5.17 (R28.20). Assim, dado um determinado número N de pontos da trajetória do espaço fase, N_l corresponde ao número de linhas diagonais na *plot* de recorrência, N_v é o número de linha verticais, *P(l) e P(v)* são o histograma do comprimento das linhas diagonais e verticais, respetivamente. Foram consideradas as seguintes *features* (Marwan et al. 2007; Zbilut & Webber, 2006):

Taxa de Recorrência (RR) é uma medida com base na densidade de recorrência, depende do número médio de vizinhos de cada ponto da trajetória:

$$RR = \frac{1}{N^2} \sum_{i,j=1}^{N} R_{i,j} \tag{25}$$

Determinismo (DET) é uma medida baseada nas linhas diagonais, depende do histograma $P(l)$, que por sua vez corresponde à relação entre os pontos recorrência que formam linhas diagonais de todos os pontos de recorrência:

$$DET = \sum_{l=l_{min}}^{N} lP(l) \Bigg/ \sum_{l=1}^{N} lP(l) \tag{26}$$

Divergência (DIV) é o inverso da L_{max}, outra medida que tem por base as linhas diagonais:

$$DIV = \frac{1}{L_{max}} \tag{27}$$

onde L_{max} é o comprimento da linha diagonal de maior cumprimento $L_{max} = \max(\{l_i; i = 1, \ldots N_l\})$, e $N_l = \sum_{l \geq l_{min}} P(l)$ é o total de linhas diagonais.

Entropia (ENT) é uma medida baseada nas linhas diagonais e representa a entropia de Shannon da probabilidade $p(l) = P(l)/N_l$ para encontrar a linha de comprimento l na *plot* de recorrência. Expressa a complexidade da *plot* de recorrência em relação às linhas diagonais.

$$ENT = -\sum_{l=l_{min}}^{N} p(l) \ln p(l) \tag{28}$$

Laminaridade (LAM) é uma métrica baseada nas linhas verticais. Corresponde à relação entre os pontos de recorrência das estruturas verticais e total dos pontos de recorrência computados. Usa-se para os v que excedam o comprimento mínimo v_{min} devido a uma diminuição na influência do movimento tangencial (Marwan et al. 2007).

$$LAM = \sum_{v=v_{min}}^{N} vP(v) \Big/ \sum_{v=1}^{N} vP(v) \tag{29}$$

Trapping Time (TT) corresponde ao comprimento médio das linhas verticais:

$$TT = \sum_{v=v_{min}}^{N} vP(v) \Big/ \sum_{v=v_{min}}^{N} P(v) \tag{30}$$

Também foram considerados os rácios entre o *determinismo* e a *taxa de recorrência* (DET/RR), conhecido por RATIO (esta denominação não será usada para não confundir com o outro rácio considerado) e entre a *laminaridade* e o *determinismo* (LAM/DET) (para verificar a utilidade entre informação sobre recorrência, linhas diagonais e verticais).

2.4.3.4. *Seleção e ponderação de "features"*

A seleção e ponderação de *features* foram aplicadas aos estudos 4 e 5, consistindo num procedimento estatístico que pondera as várias medidas face à sua relevância e redundância para o fenómeno a classificar. Consideramos também vários conjuntos de *features*, de diferentes dimensões. A formação de conjuntos de *features* em grupo de dois (F2), quatro (F4) e seis (F6) foi baseada na ponderação pelo *Fisher Score* (Duda, Hart, & Stork, 2001) e no algoritmo *Correlation-based Feature Selection* (Hall, 1999), para ordenar as *features* quanto à sua importância para a classificação.

O algoritmo *Fisher Score* (FS) atribui uma ordenação a um conjunto de *features* consoante a sua adequação para a classificação, em função do índice de Fisher. Dado um vetor de rótulos $y = \{y_1, y_2, \dots, y_c\}$ contendo c classes, o FS para cada *feature* i é definido por:

$$FS(f_i) = \frac{\sum_{j=1}^{c} n_j \left(\mu_{i,j} - \mu_i\right)^2}{\sum_{j=1}^{c} n_j \sigma_{i,j}^2} \quad (31)$$

onde μ_i é a média da feature f_i, n_j o número de amostras na $j$ésima classe, $\mu_{i,j}$ e $\sigma_{i,j}$ correspondem à média e à variância de f_i na classe j, respetivamente (Zhao et al., 2010).

O algoritmo *Correlation-based Feature Selection* (CFS), ao contrário do FS, considera o pressuposto que a seleção de *features* para aprendizagem automática deve-se processar tendo em conta a correlação entre as mesmas (Hall, 1999). Se a correlação entre cada um dos componentes de um teste em relação a determinada variável a classificar for conhecida, a inter-relação entre cada par de componentes para classificação pode ser estimada recorrendo a

$$M_s = \frac{k\overline{r_{cf}}}{\sqrt{k + k(k-1)r_{\overline{ff}}}} \tag{32}$$

onde M_s é o mérito heurístico do conjunto de *features* S, dependendo da correlação entre a soma da componentes e a variável a classificar e k é o número de componentes. $\overline{r_{cf}}$ é a média das correlações entre as componentes em relação à variável a classificar e $r_{\overline{ff}}$ corresponde à média da inter-correlação entre as componentes. O numerador fornece informação sobre a discriminação por classe para determinando conjunto de *features*, o denominador caracteriza a redundância (Hall, 1999).

As *features* foram ponderadas segundo os dois algoritmos de seleção de *features* para cada músculo da seguinte forma:

1) Ordem de ponderação segundo o FS;
2) Pontuação de cada *feature* segundo resiliência ao surgir como discriminatória no classificador CFS.

Para o estudo do Capítulo VII (quinto estudo) foi apenas utilizado o algoritmo CFS, o qual serviu de base na construção dos conjuntos de *features* RQA que representam cada músculo estudado.

2.5. Máquinas de Vetor Suporte

As máquinas de vetores suporte (SVM – *Support Vector Machines*) são algoritmos avançados de otimização matemática, baseados na aprendizagem supervisionada, introduzidos no início da década de 90 (Boser et al., 1992), estendendo-se posteriormente para dados não separáveis (Cortes & Vapnik, 1995). O caso mais simples que serviu como base para a construção destes algoritmos foi a aplicação a dados separáveis também denominados por margens rígidas. Começa-se a explicação deste ponto até à aplicação não linear que foi utilizada no presente estudo, fazendo então referência aos parâmetros utilizados.

2.5.1. SVM com dados separáveis

No capítulo I, referiu-se o ponto de partida para perceber as SVM, quando estas se reportam a dados separáveis como o caso binário, ou seja, com duas classes. Considerando um conjunto de treino com n objetos $\mathbf{x}_i \in \mathcal{X}$ com os rótulos de dados $y_i \in \mathcal{Y}$, em que $\mathbf{x}_i \in \mathbb{R}^n$ e $y_i \in \{-1,1\}$, para $i = 1,2,\dots,l$ exemplos, o conjunto de treino é linearmente separável por um hiperplano que distinga segundo a equação (7) como pertencentes a uma das duas classes A e B através do conjunto de treino $(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_l, y_l)$ (Boser et al., 1992; Cristianini & Shawe-Taylor, 2000). Então, um hiperplano irá separar os dados segundo as duas classes pretendidas. Colocando-se a problemática sobre qual o hiperplano ótimo que garante essa separabilidade, procura-se aquele que garante a maior maximização das suas margens no espaço geométrico, pelo que, os pontos que se encontram sobre esse hiperpl ano satisfazem a equação (8) (Boser et al., 1992; Burges, 1998; Cortes & Vapnik, 1995; Cristianini & Shawe-Taylor, 2000; Scholkopf et al., 1995). Onde **w** corresponde ao vetor normal ao hiperplano, $|b|/\|\mathbf{w}\|$ representa a distância perpendicular do hiperplano à origem, $\|\mathbf{w}\|$ é a norma euclidiana de **w**, e $\mathbf{w} \cdot \mathbf{x}$ é o produto escalar entre os vetores **w** e **x**. O algoritmo do vetor de suporte escolhe o hiperplano de separação com maior margem (Burges, 1998; Cristianini & Shawe-Taylor, 2000) que satisfaz:

$$\begin{array}{ll} \mathbf{w} \cdot \mathbf{x}_i + b \geq +1 \quad para \quad y_i = +1 \\ \mathbf{w} \cdot \mathbf{x}_i + b \leq -1 \quad para \quad y_i = -1 \end{array} \tag{33}$$

Sendo preciso conhecer o vetor **w** que maximiza a margem, consideram-se os hiperplanos que definem as classes expressos por

$$H_{-1} = \{\mathbf{x} \in \mathbb{R}^n : w \cdot \mathbf{x} + b = -1\}, \qquad H_1 = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{w} \cdot \mathbf{x} + b = 1\}, \tag{34}$$

em que a distância entre os hiperplanos é dada por $\frac{2}{||\mathbf{w}||}$. Ao minimizar $||\mathbf{w}||^2$ encontram-se o par de hiperplanos com a máxima margem sujeita às restrições (Burges, 1998; Cristianini & Shawe-Taylor, 2000):

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1 \geq 0 \quad \forall_i \tag{35}$$

Os vetores que se encontram sobre os hiperplanos H_{-1} e H_1 formam as margens maximizadas e são denominados por vetores suporte. Supondo que obtemos para os hiperplanos H_{-1} e H_1 os vetores suporte $\mathbf{x}_{-1}$ e $\mathbf{x}_1$, respetivamente. A projeção da distância entre os hiperplanos H_{-1} e H_1 pode ser representada pela seguinte equação (Lorena & Carvalho, 2007):

$$(\mathrm{x}_{-1} - \mathrm{x}_1)\left(\frac{\mathbf{w}}{||\mathbf{w}||} \cdot \frac{\mathbf{x}_{-1} - \mathbf{x}_1}{||\mathbf{x}_{-1} - \mathbf{x}_1||}\right) \tag{36}$$

Como pretendemos a diferença entre $\mathbf{x}_{-1} - \mathbf{x}_1$, face à condição que $H_{-1} : \mathbf{w} \cdot \mathbf{x}_{-1} + b = 1$ e $H_1 : \mathbf{w} \cdot \mathbf{x}_1 + b = -1$, a diferença será $\mathbf{w} \cdot (\mathbf{x}_{-1} - \mathbf{x}_1) = 2$ e obtém-se:

$$\frac{2(\mathbf{x}_{-1} - \mathbf{x}_1)}{||\mathbf{w}||\,||\mathbf{x}_{-1} - \mathbf{x}_1||} \equiv \frac{2}{||\mathbf{w}||} \tag{37}$$

O hiperplano ótimo será aquele que satisfaz as desigualdades $y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1 \geq 0$, minimizando a norma $||\mathbf{w}||$:

$$\min_{w,b} \frac{1}{2} ||\mathbf{w}||^2 \tag{38}$$

São então introduzidos multiplicadores de Lagrange positivos α_i convertendo as condições anteriores num problema de otimização (Burges, 1998; Cristianini & Shawe-Taylor, 2000; Müller et al., 2001; Scholkopf et al., 1995):

$$\alpha_i(y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1) > 0 \tag{39}$$

Como a restrição se repete para os diferentes pontos no espaço $\mathcal{X}$, concretiza-se a sua soma:

$$\sum_i^l \alpha_i(y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1) > 0 \tag{40}$$

Agregando este termo à função objetivo, origina a função Lagrangiana na sua formulação primordial (Burges, 1998; Cristianini & Shawe-Taylor, 2000; Müller et al., 2001):

$$\mathcal{L}(\mathbf{w}, b, \alpha) = \frac{1}{2}\|w\|^2 - \sum_i^l \alpha_i(y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1) \tag{41}$$

Torna-se necessário minimizar $\mathcal{L}(\mathbf{w}, b, \alpha)$ em relação a $\mathbf{w}$ e b que significa as derivadas parciais de $\mathcal{L}(\mathbf{w}, b, \alpha)$ serem igualadas a zero com as restrições $\alpha_i \geq 0$. Se a restrição $y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1 \geq 0$ é violada, então $y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1 < 0$, neste caso $\mathcal{L}(\mathbf{w}, b, \alpha)$ pode aumentar pelo aumento do parâmetro α_i. Ao mesmo tempo, $\mathbf{w}$ e b terá de ser escolhido de forma a diminuir $\mathcal{L}(\mathbf{w}, b, \alpha)$ e evita-se que esta função se torne num número arbitrariamente elevado e negativo (Schölkopf & Smola, 2002), obtendo-se um "ponto sela" segundo:

$$\frac{\partial}{\partial b}\mathcal{L}(\mathbf{w}, b, \alpha) = 0 \qquad e \qquad \frac{\partial}{\partial \mathbf{w}}\mathcal{L}(\mathbf{w}, b, \alpha) = 0 \tag{42}$$

Passando à formulação dual, começamos por impor (Boser et al., 1992; Burges, 1998; Chen, Lin, & Schölkopf, 2005; Lorena & Carvalho, 2007; Schölkopf & Smola, 2002) que:

$$\sum_{i=1}^{l} \alpha_i y_i = 0 \tag{43}$$

$$\mathbf{w} = \sum_{i=1}^{l} \alpha_i y_i \mathbf{x}_i \tag{44}$$

obtendo o seguinte problema de otimização com formulação dual, ao substituir estas equações em $L(\mathbf{w}, b, \alpha)$:

$$\max_{\alpha} \quad \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} \alpha_i \alpha_j y_i y_j (\mathbf{x}_i \cdot \mathbf{x}_j) \tag{45}$$

com as restrições:

$$\begin{cases} \alpha_i \geq 0, \qquad \forall i = 1, \dots, n \\ \sum_{i=1}^{l} \alpha_i y_i = 0 \end{cases} \tag{46}$$

Trocando os sinais:

$$\min_{\alpha} \quad \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} \alpha_i \alpha_j y_i y_j (\mathbf{x}_i \cdot \mathbf{x}_j) - \sum_{i=1}^{l} \alpha_i \tag{47}$$

O “ponto sela” que deriva de $\mathcal{L}$ em respeito às variáveis primais é obtido pelas condições de Karush-Kuhn-Tucker (KKT) (Kuhn & Tucker, 1951), determinando $\mathbf{w}$, em que o vetor suporte será uma expansão do subconjunto de treino, os dados que possuem $\alpha_i > 0$ determinam as margens H_{-1} e H_1 do hiperplano por se encontrarem sobre os mesmos. Os restantes dados da equação $\alpha_i(y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1) = 0$ consideram $\alpha_i = 0$ não sendo utilizados para o cálculo de **w** (Burges, 1998; Chen, Lin, & Schölkopf, 2005; Hofmann, Schölkopf, & Smola, 2008; Schölkopf & Smola, 2002). O valor de b é calculado através dos vetores suporte expressos por $\alpha_i(y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1) = 0$ (Lorena & Carvalho, 2007; Schölkopf & Smola, 2002):

$$b = \frac{1}{nSV} \sum_{\mathbf{x}_i \in \boldsymbol{SV}} \frac{1}{y_j} - \mathbf{w} \cdot \mathbf{x}_j \qquad (48)$$

em que nSV corresponde ao número de vectores suporte, podendo a equação ser expandida pela substituição de **w** obtendo-se:

$$b = \frac{1}{nSV} \sum_{\mathbf{x}_i \in \boldsymbol{SV}} \left(\frac{1}{y_j} - \sum_{\mathbf{x}_i \in \boldsymbol{SV}} \alpha_i y_i \mathbf{x}_i \cdot \mathbf{x}_j \right) \qquad (49)$$

O hiperplano que separa os dados pela maior margem será aquele com maior capacidade de generalização. Porém, como a existência de dados linearmente separáveis não é comum em tratamentos reais, o ruído existente nos dados dificulta, ou até impossibilita a utilização deste método no formato explanado, considerando então a extensão para dados não separáveis, na qual pode ser considerada também separação linear (onde é aceite uma margem de erro), ou então a sua aplicação não linear.

2.5.2. SVM com dados não separáveis: margens suaves

Para estender os procedimentos a margens suaves torna-se necessário introduzir variáveis de folga às restrições, também denominado por custo. Este custo compreende a introdução de variáveis não negativas $\xi_i \geq 0, i = 1, \dots, l$ exemplos (Burges, 1998; Cortes

& Vapnik, 1995). Assume-se que os dados não podem ser separados totalmente por um hiperplano, pelo que as restrições vão passar a ter o seguinte formato:

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1 \geq -\xi_i \tag{50}$$

com $\xi_i \geq 0$. O problema considerado em (41) passa a assumir o seguinte formato, com o adicionar da função custo à função objetivo (Burges, 1998; Cortes & Vapnik, 1995; Cristianini & Shawe-Taylor, 2000):

$$\min_{w,b} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^{l} \xi_i^{\sigma} \tag{51}$$

adicionando as restrições $0 \leq \alpha_i \leq C$. O valor de $\sigma > 0$ na equação anterior será igual a um sendo o menor valor possível para que o problema de programação quadrática expresso em (45) tenha solução única (Cortes & Vapnik, 1995), permitindo a vantagem em relação aos multiplicadores de Lagrange, pelo que o problema dual é expresso da seguinte forma:

$$\max_{\alpha} \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} \alpha_i \alpha_j y_i y_j (\mathbf{x}_i \cdot \mathbf{x}_j) \tag{52}$$

sujeito às restrições:

$$\begin{cases} 0 \leq \alpha_i \leq C, \qquad \forall i = 1, \dots, n \\ \sum_{i=1}^{n} \alpha_i y_i = 0 \end{cases} \tag{53}$$

com a solução:

$$\mathbf{w} = \sum_{\mathbf{x}_i \in \boldsymbol{SV}} \alpha_i y_i \, \mathbf{x}_i \tag{54}$$

em que SV corresponde ao número de vetores suporte. Recorrendo novamente às condições Karush-Kuhn-Tucker obtém-se o seguinte problema segundo a função primal de Lagrange (Burges, 1998; Cristianini & Shawe-Taylor, 2000; Hofmann et al., 2008):

$$\mathcal{L}(\mathbf{w}, b, \alpha) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l} \xi_i - \sum_{i}^{l} \alpha_i (y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1 + \xi_i) \; - \sum_{i=1}^{l} \eta_i \tag{55}$$

com $\alpha_i, \eta_i \geq 0 \; \forall_i \in n$, visto que α_i, η_i são multiplicadores de Lagrange introduzidos para reforçar a positividade de ξ_i. Para realizar a função dual de $\mathcal{L}(\mathbf{w}, b, \alpha)$ procede-se à identificação das condições de primeira ordem em $(\mathbf{w}, b)$ através das seguintes restrições:

$$\frac{\partial}{\partial \mathbf{w}} \mathcal{L} = \mathbf{w} - \sum_{i=1}^{l} \alpha_i y_i \, \boldsymbol{x_i} = 0 \tag{56}$$

$$\frac{\partial}{\partial b} \mathcal{L} = -\sum_{i=1}^{l} \alpha_i y_i = 0 \tag{57}$$

$$\frac{\partial}{\partial \xi_i} \mathcal{L} = C - \alpha_i + \eta_i = 0 \tag{58}$$

$$(y_i(\mathbf{w} \cdot \mathbf{x}_{\boldsymbol{i}} + b) - 1 + \xi_i) = 0 \tag{59}$$

$$\alpha_i(y_i(\mathbf{w} \cdot \mathbf{x}_{\boldsymbol{i}} + b) - 1 + \xi_i) = 0 \tag{60}$$

$$\eta_{i\xi_i} = 0 \tag{61}$$

$\alpha_i \in [0, C], \forall_i \in n$

Porém, este caso e o anterior consideram que é possível separar as variáveis por um hiperplano, o que na maioria dos casos reais não é provável devido à localização no espaço das *features* em relação aos dados a classificar.

2.5.3. SVM com dados não separáveis: não linear e "*kernel trick*"

Na aplicação não linear é realizado um mapeamento Φ dos dados para um espaço euclidiano $\mathcal{H}$ de dimensão superior (Burges, 1998; Cristianini & Shawe-Taylor, 2000; Müller et al., 2001):

$$\Phi: \mathbb{R}^{\mathrm{n}} \longrightarrow \mathcal{H}$$

$$\mathbf{x} \longmapsto \Phi(\mathbf{x})$$

As capacidades de separação deste procedimento são suportadas pelo teorema de Cover (Cover, 1965), permitindo o aumento da dimensionalidade sem que este signifique maior complexidade (Müller et al., 2001). Passa-se a processar o exemplo de treino $(\Phi(\mathbf{x}_1), y_1), (\Phi(\mathbf{x}_2), y_2), \ldots, (\Phi(\mathbf{x}_l), y_l) \in \mathcal{H} \times \mathcal{Y}$. Como existe dificuldade em utilizar de forma explícita Φ recorre-se a uma forma engenhosa de lidar com este problema, procedendo-se à substituição do produto interno $\mathbf{x}_i \cdot \mathbf{x}_j$ por uma função $K(\mathbf{x}_i, \mathbf{x}_j)$ na formulação dual em que K corresponde à função kernel que irá satisfazer pressupostos evidenciados pelo teorema de Mercer. Destes pressupostos destacam-se a positividade definida e a simetria: se k é uma função *kernel* contínua de um operador inteiro positivo, pode-se construir um mapeamento para um espaço $\mathcal{H}$ onde k age como um produto interno (Cristianini & Shawe-Taylor, 2000; Herbrich, 2002; Hofmann et al., 2008). Este procedimento designa-se por *"kernel trick"*, permitindo obter o produto interno no espaço $\mathcal{H}$ sem precisar de definir explicitamente o mapeamento Φ, através de

$$K(\mathbf{x}_i, \mathbf{x}_j) = \Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j) \tag{62}$$

O exemplo mais evidenciado na literatura (ex.: Burges, 1998; Hofmann et al., 2008; Lorena & Carvalho, 2003; 2007; Müller et al., 2001) é a transferência de dados de $\mathbb{R}^2$ para $\mathbb{R}^3$, em que:

$$\Phi: \mathbb{R}^2 \longrightarrow \mathbb{R}^3$$
$$\Phi(\mathbf{x}) = (x_1, x_2) \longmapsto (z_1, z_2, z_3) \coloneqq \left(x_1^2, \sqrt{2}x_1x_2, x_2^2\right) \tag{63}$$
$$\mathbf{w} \cdot \Phi(\mathbf{x}) + b = w_1x_1^2 + w_2\sqrt{2}x_1x_2 + w_3x_2^2 + b = 0$$

podendo agora ser aplicado um hiperplano linear no espaço de três dimensões. Sobre as mesmas restrições já mencionadas para o caso de margens suaves lineares, o problema de otimização passará a ter a seguinte forma:

$$\max_{\alpha} \sum_{i=1}^{l} \alpha_i - \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l} \alpha_i\alpha_j y_i y_j \left(\Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j)\right) \tag{64}$$

com o classificador,

$$g(x) = sgn(f(x)) = sgn\left(\sum_{\mathbf{x}_i \in \boldsymbol{SV}} \alpha_i y_i \Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j) + b\right) \tag{65}$$

Procedendo ao “*kernel trick*” que permite recorrer ao produto interno entre dois espaços característicos **x** e **y**:

$$\left(\boldsymbol{\Phi}(\mathbf{x}) \cdot \boldsymbol{\Phi}(\mathbf{y})\right) = \left(x_1^2, \sqrt{2}x_1x_2, x_2^2\right)\left(y_1^2, \sqrt{2}y_1y_2, y_2^2\right)^T$$
$$= \left((x_1, x_2)(y_1, y_2)^T\right)^2$$
$$= (\mathrm{x} \cdot \mathrm{y})^2 \coloneqq \boldsymbol{k}(\mathbf{x},\mathbf{y})$$

Pode-se generalizar à função (Müller et al., 2001; Scholkopf et al., 1996; Schölkopf & Smola, 2002):

$$\boldsymbol{k}(\mathbf{x},\mathbf{y}) = (\mathbf{x} \cdot \mathbf{y})^{\boldsymbol{d}} \tag{66}$$

Considerando a notação anterior, tem-se $k(x_i, x_j)$ ao invés de k(x,y) na forma dual do problema, pelo que:

$$\max_{\alpha} \quad \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} \alpha_i \alpha_j y_i y_j \, \boldsymbol{k}(\mathbf{x}_i, \mathbf{x}_j) \tag{67}$$

Sujeito às restrições já mencionadas para o caso não separável e com a solução (Begg, Palaniswami, Member, & Owen, 2005):

$$g(x) = sgn\big(f(x)\big) = sgn\left(\sum_{\mathbf{x}_i \in \boldsymbol{SV}} \alpha_i y_i \, \boldsymbol{k}(\mathbf{x}_i, \mathbf{x}_j) + b \right) \tag{68}$$

No presente estudo optou-se pela função RBF (*Radial Basis Function*) gaussiana (Hsu, Chang, & Lin, 2010):

$$K\big(\mathbf{x}_i, \mathbf{x}_j\big) = exp\left(-\gamma \left\| \mathbf{x}_i - \mathbf{x}_j \right\|^2\right), \gamma > 0 \tag{69}$$

em que γ é o parâmetro kernel. O número de vetores suporte necessários ao usar as funções *kernel* polinomial, RBF ou sigmoide não diverge muito, sendo a RBF aquela que apresenta menor número de vetores na construção no modelo de decisão (Scholkopf et al., 1995). No quarto estudo (Capítulo IV) e no quinto (Capítulo V) foram realizadas pesquisas de rede (para determinar os melhores parâmetros *kernel* e custo) utilizando para validação o método matemático *5-fold cross-validation*.

No quarto estudo, os intervalos de pesquisa compreenderam duas pesquisas, uma ampla e outra refinada, com os seguintes intervalos:

- Pesquisa ampla: $C = 2^{-2:1.25:8}$; $\gamma = 2^{-7:1.25:3}$.

- Pesquisa refinada: dependendo dos valores obtidos ao melhor valor determinado na pesquisa ampla foi colocado um espaçamento h = 0.25 tanto à esquerda como à direita desse valor.

No quinto estudo, os intervalos foram relativos a uma pesquisa ampla com as seguintes características:

- Pesquisa ampla: $\boldsymbol{C} = 2^{-5:2:17}$; $\boldsymbol{\gamma} = 2^{-9:2:5}$.

Nos estudos onde foram aplicadas SVM a razão entre repetições de treino e teste foi similar, correspondendo a 80% e 20% dos dados respetivamente.

2.6. Estatística Complementar

No primeiro estudo, para além das tarefas de tradução e retroversão da língua original para português e vice-versa e reunião de peritos, realizou-se a aplicação do questionário com o intervalo máximo de duas semanas procedendo-se ao teste re-teste. A validade foi aferida pelo coeficiente estatístico Kappa de Cohen para itens com escalas nominais e com o coeficiente correlação intraclasse (2,1) (ICC) *two way random single measures* (2,1) para itens com escalas quantitativas [20-22]. O r de Pearson foi utilizado em conivência com o ICC. Esta técnica também foi aplicada no estudo 2 para determinar a fiabilidade entre métodos, mas ICC do tipo *two-way mixed absolute agreement*. O r de Pearson voltou a ser utilizado no estudo 4 correlacionando a precisão com a percentagem de vetores suporte.

Os pressupostos de normalidade foram confirmados recorrendo ao teste Shapiro-Wilk. A homogeneidade das variâncias foi testada pelo teste de Levene.

A ANOVA de medidas repetidas *three-way* foi utilizada para comparar métodos, músculos e os tacos no estudo 2.

No estudo 4 recorreu-se à ANOVA mista tendo como fatores de medidas repetidas a precisão e a percentagem de vetores suporte obtidas com os três conjuntos de *features* e como fatores independentes os grupos de Hc (alto, baixo e ambos).

A MANOVA não paramétrica foi aplicada nos quarto e quinto estudos. No quarto estudo serviu para verificar diferenças entre os grupos Hc em relação aos parâmetros C e γ associados aos modelos SVM. No quinto estudo foi aplicada para confirmar as diferenças na taxa de recorrência nos vários músculos por grupos de Hc e sintomatologia de lombalgia. Para este teste foi realizada uma transformação por *ranks* e a estatística de teste χ^2 foi calculada recorrendo ao Traço Pillai, sendo o *p_value* corrigido. O *p_value* correspondente à estatística de teste F das ANOVAs foi corrigido pela estatística de teste H_F com distribuição χ^2.

Utilizou-se o teste de Friedman e respetivas comparações múltiplas de Dunn para verificar a existência de diferenças na localização central nos parâmetros custo C para dados não separáveis SVM e *kernel* γ usados nos diferentes grupos de *features* (F2, F4 e F6) do estudo 4. Esta técnica voltou a ser utilizada no estudo 5 para comparar a taxa de recorrência na discriminação por handicap e lombalgia.

Nos estudos que assim o justificaram, as comparações múltiplas foram realizadas por recurso ao teste de Bonferroni para as medidas repetidas. A esfericidade foi verificada com o teste de Mauchly e quando esta não se verificou, os graus de liberdade foram corrigidos pelo Épsilon de Greenhouse-Geisser. Como a amostragem tem natureza não probabilística realizou-se sempre que necessário a estimativa do efeito da dimensão (η_p^2), assim como, a potência do teste (π).

O nível de significância considerado foi de 5% sendo corrigido pelo número de ANOVAS quando necessário. Com exceção do estudo 5, o processamento estatístico tido como complementar foi realizado no programa IBM-SPSS 19.0 (IBM Corporation, New York, USA).

2.7. Plano de Trabalho

Após evidenciar a metodologia apresenta-se a organização da presente dissertação quanto aos artigos que resultam da mesma (Tabela 2). Os estudos 1 ao 5 constituem respetivamente, os capítulos III a VII, em conformidade com o já explanado na introdução. Na Tabela 2 pode ser verificado o processo realizado ao longo desta dissertação. Inicia-se por um estudo cuja importância está na categorização desta população, passando-se posteriormente para a análise do fenómeno temporal através de métodos clássicos aperfeiçoados pela inclusão de restrições. Após esta estruturação de base, os três estudos que se seguem incidem na aprendizagem automática, nomeadamente, na utilização das SVM.

Tabela 2 – Resumo da compilação de artigos.

Capítulo/ Estudo	Principal Objetivo	Variáveis Independentes	Variáveis Dependentes
Capítulo III/ Estudo 1 **Cross-cultural Adaptation and validation of the Portuguese Survey of Musculoskeletal Conditions, Playing Characteristics and Warm-up Patterns of Golfers**	Validação para a população portuguesa.	Hábitos de Atividade Física Características do Golfista Histórico de lesões	Capacidade Categórica e/ou Grau de cada Item
Capítulo IV/ Estudo 2 **Trunk Muscle Activation during Golf Swing: Baseline and threshold**	Deteção do *onset* no *swing* comparando linhas de base.	Linhas de Base	Onset (ms) % Deteção do Pico Máximo
Capítulo V/ Estudo 3 **EMG-based Classification with Support Vector Machine: A Review**	Revisão de literatura sobre a aplicação de SVM à EMG.	SVM EMG	Categorias de Estudo Metodologia
Capítulo VI/ Estudo 4 **Machine Learning Classification of EMG Temporal Parameters: Support Vector Machines on Onset**	Verificar a performance das SVM na deteção do *onset* durante o *swing*.	Sinal de Repouso e de Atividade Handicap	*Features* no domínio do tempo Precisão e Nº Vetores Suporte dos Modelos/Músculos
Capítulo VII/ Estudo 5 **Recurrence Quantification Analysis and Support Vector Machine for Handicap and Low Back Pain EMG classification in dynamic skills: swing**	Analisar o poder discriminatório da função neuromuscular na discriminação por handicap e lombalgia.	*Swing* e fases do *Swing* Handicap Lombalgia	*Features* RQA Precisão e Nº Vetores Suporte dos Modelos/Músculos

Referências

Allison, G. T. (2003). Trunk muscle onset detection technique for EMG signals with ECG artefact. *Journal of Electromyography and Kinesiology*, *13*(3), 209–216. doi:10.1016/S1050-6411(03)00019-1.

Bechler, J.R., Jobe, F. W., Pink, M., Perry, J., & Ruwe, P. A. (1995). Electromyographic analysis of the hip and knee during the golf swing. *Clinical Journal of Sport Medicine*, *5*(3), 162–166.

Begg, R. K., Palaniswami, M., Member, S., & Owen, B. (2005). Support vector machines for automated gait classification. *IEEE Transactions on Biomedical Engineering*, *52*(5), 828–838.

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992) A Training algorithm for optimal margin classfiers. *Proceedings of the Fifth Annual Workshop on Computational Learning Theory* (pp. 144–152).

Burges, C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, *2*, 121–167. Retrieved from http://link.springer.com/article/10.1023/A:1009715923555

Cheetham, P. J., Martin, P. E., Mottram, R. E., & Laurent, B. F. St. (2001). The importance of stretching the "X-Factor" in the downswing of golf : The "X-Factor Stretch". In P. R. Thomas (Ed.), *Optimising performance in golf* (pp. 192–199). Brisbane, Australia: Australian Academic Press.

Chen, P.-H., Lin, C.-J., & Scholkopf, B. (2005). A tutorial on nu-support vector machines. *Applied Stochastic Models in Business and Industry*, *21*(2), 111–136. doi:10.1002/asmb.537.

Cole, M., & Grimshaw, P. (2008). Trunk muscle onset and cessation in golfers with and without low back pain. *Journal of Biomechanics*, *41*, 2829–2833. doi:10.1016/j.jbiomech.2008.07.004.

Cole, M. H, & Grimshaw, P. N. (2013). The crunch factor's role in golf-related low back pain. *The Spine Journal*, *Oct 12*.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. doi:10.1007/BF00994018.

Cover, T. M. (1965). Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE Transactions on Electronic Computers*, 326–334.

Cristianini, N., & Shawe-Taylor. (2000). *An introduction to support vector machines and other kernel-based learning methods*. Cambridge: Cambridge University Press.

Duda, R., Hart, P., & Stork, D. (2001). *Pattern classification* (2nd Ed). New York: John Wiley & Sons.

Eckmann, J.-P. , Kamphorst, O. S., & Ruelle, D. (1987). Recurrence plots of dynamical systems. *Europhysics. Letters*, *4*(9), 973–977.

EGA. (2012). *Sistema de handicap EGA 2012-2016*. Epalinges/Lausanne: Edição Federação Portuguesa de Golfe.

Evans, K., Refshauge, K. M., Adams, R., & Aliprandi, L. (2005). Predictors of low back pain in young elite golfers: A preliminary study. *Physical Therapy in Sport*, *6*(3), 122–130. doi:10.1016/j.ptsp.2005.05.003.

Fox, E., Lindsay, D., & Vandervoort, A. (2002). Musculoskeletal injury questionnaire for senior golfers. *Science and Golf IV. Proceedings of the World Scientific Congress of Golf* (pp. 88–99). St. Andrews, Scotland: Routledge Inc.

Gluck, G. S., Bendo, J. A., & Spivak, J. M. (2008). The lumbar spine and low back pain in golf: a literature review of swing biomechanics and injury prevention. *The Spine Journal*, *8*(5), 778–88. doi:10.1016/j.spinee.2007.07.388.

Hall, M. A. (1999). *Correlation-based feature selection for machine learning*. (Doctoral dissertation). University of Waikato.

Herbrich, R. (2002). *Learning kernel classifiers. Theory and algorithms*. Massachusetts: Massachusetts Institute of Technology.

Hermens, H. J., Merletti, & Freriks. B. (1996). European Activities on Surface Electromyography. European Activities on Surface Electromyography, *Proceedings of the First General SENIAM Workshop*. SENIAM project, published by Roessingh Research and Development.

Hodges, P., & Bui, B. (1996). A comparison of computer-based methods for the determination of onset of muscle contraction using electromyography. *Electroencephalography and Clinical Neurophysiology*, *101*, 511–519.

Hofmann, T., Schölkopf, B., & Smola, A. J. (2008). Kernel methods in machine learning. *The Annals of Statistics*, *36*(3), 1171–1220. doi:10.1214/009053607000000677.

Horton, J. F., Lindsay, D. M., & Macintosh, B. R. (2001). Abdominal muscle activation of elite male golfers with chronic low back pain. *Medicine and Science in Sports and Exercise*, *33*(10), 1647–54.

Hsu, C., Chang, C., & Lin, C. (2010). A practical guide to support vector classification. Retrieved from https://www.cs.sfu.ca/people/Faculty/teaching/726/spring11/svmguide.pdf

Hudgins, B., Parker, P., & Scott, R. N. (1993). A new strategy for multifunction myoelectric control. *IEEE Transactions on Biomedical Engineering*, *40*(1), 82–94.

Hume, P. A, Keogh, J., & Reid, D. (2005). The role of biomechanics in maximising distance and accuracy of golf shots. *Sports medicine*, *35*(5), 429–49.

Jöllenbeck, T. (2000). Methodological limitations of EMG-based bio-mechanical motion analysis. *ISBS – Conference Proceedings Archive, 18 International Symposium on Biomechanics in Sports.*

Kuhn, H. W., & Tucker, A. W. (1951). Nonlinear programming. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability* (pp. 481–492).

Lee, A. S., Cholewicki, J., & Reeves, N. P. (2007). The effect of background muscle activity on computerized detection of sEMG onset and offset. *Journal of Biomechanics*, *40*(15), 3521–3526. doi:10.1016/j.jbiomech.2007.05.012.

Lindsay, D., & Horton, J. (2002). Comparison of spine motion in elite golfers with and without low back pain. *Journal of Sports Sciences*, *20*(8), 599–605. doi:10.1080/026404102320183158.

Lorena, A. C., & Carvalho, A. C. P. L. F. (2003). *Introdução às Máquinas de Vetores Suporte*. (Relatórios Técnicos do ICMC).

Lorena, A. C., & Carvalho, A. C. P. L. F. De. (2007). Uma introdução às support vector machines. *RITA*, *XIV*(2), 43–67.

Marwan, N., Carmenromano, M., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics Reports*, *438*(5-6), 237–329. doi:10.1016/j.physrep.2006.11.001.

McGill, S., Chaimberg, J., Frost, D., & Fenwick, C. (2010). Evidence of a double peak in muscle activation to enhance strike speed and force: an example with elite mixed martial arts fighters. *Journal of Strength and Conditioning Research*, *24*(2), 348–357. doi:10.1519/JSC.0b013e3181cc23d5.Evidence.

McHardy, A., & Pollard, H. (2005). Lower back pain in golfers: A review of the literature. *Journal of Chiropractic Medicine*, *4*(3), 135–143. doi:10.1016/S0899-3467(07)60122-0.

McHardy, A. J., Pollard, H. P., & Luo, K. (2007). Golf-related lower back injuries: An epidemiological survey. *Journal of Chiropractic Medicine*, *6*(1), 20–6. doi:10.1016/j.jcme.2007.02.010.

Müller, K.-R., Mika, S., Rätsch, G., Tsuda, K., & Schölkopf, B. (2001). An introduction to kernel-based learning algorithms. *IEEE TRANSACTIONS on Neural Networks*, *12*(2), 181–201.

Oskoei, A. M., & Hu, H. (2007). Myoelectric control systems—A survey. *Biomedical Signal Processing and Control*, *2*(4), 275–294. doi:10.1016/j.bspc.2007.07.009.

Oskoei, M. A., & Hu, H. (2008). Support vector machine-based classification scheme for myoelectric control applied to upper limb. *IEEE Transaction on Biomedical Engineering*, *55*(8), 1956–1965.

Pezarat-Correia, P. & Mil-Homens, P. (2004). *A Electromiografia no estudo do movimento humano*. Cruz Quebrada: Edições FMH.

Phinyomark, A., Limsakul, C., & Phukpattaranont, P. (2009). A novel feature extraction for robust EMG pattern recognition. *Journal of Computing*, *1*(1), 2151–9617.

Phinyomark, A., Phukpattaranont, P., & Limsakul, C. (2012). Feature reduction and selection for EMG signal classification. *Expert Systems with Applications*, *39*(8), 7420–7431. doi:10.1016/j.eswa.2012.01.102.

Pink, M., Perry, J., & Jobe, F. (1993). EMG analysis of the trunk in golfers. *American Journal of Sports Medicine*, *21*(3):385–388.

Schölkopf, B., & Smola, A. (2002). *Learning with kernels*. Massachustts: Massachusetts Institute of Technology. Retrieved from http://dip.sun.ac.za/~hanno/tw796/lesings/mlss06au_scholkopf_lk.pdf

Scholkopf, B., Burges, C., & Vapnik, V. (1995). Extracting Support Data for a Given Task. In U. M. Fayyad & R. Uthurusamy (Eds.), *First International Conference on Knowledge Discovery & Data Mining*. (pp. 252–257). Menlo Park: AAAI Press.

Scholkopf, B., Smola, A., & Muller, K. R. (1996). *Nonlinear component analysis as a kernel eigenvalue problem* (pp. 1–18). Tübingen.

Solnik, S., Rider, P., & Steinweg, K. (2010). Teager–Kaiser energy operator signal conditioning improves EMG onset detection. *European journal of Applied Physiology*, *2010*(110), 489–498. doi:10.1007/s00421-010-1521-8.

Staude, G., Flachenecker, C., Daumer, M., & Wolf, W. (2001). Onset detection in surface electromyographic signals: A systematic comparison of methods. *Journal on Applied Signal Processing*, *2001*(2), 67–81.

Thiel, M., Romano, M. C., & Kurths, J. (2004). How much information is contained in a recurrence plot? *Physics Letters A*, *330*(5), 343–349. doi:10.1016/j.physleta.2004.07.050.

Tsai, Y.-S., Sell, T. C., Smoliga, J. M., Myers, J. B., Learman, K. E., & Lephart, S. M. (2010). A comparison of physical characteristics and swing mechanics between golfers with and without a history of low back pain. *Journal of Orthopaedic and Sports Physical Therapy*, 40(7), 430–438. doi:10.2519/jospt.2010.3152.

Vad, V. B., Bhat, A. L., Basrai, D., Gebeh, A., Aspergren, D. D., & Andrews, J. R. (2004). Low back pain in professional golfers: the role of associated hip and low back range-of motion deficits. *American Journal of Sports Medicine*, *32*(2), 494–497. doi:10.1177/0363546503261729.

Vaisman, L., Zariffa, J., & Popovic, M. R. (2010). Application of singular spectrum-based change-point analysis to EMG-onset detection. *Journal of Electromyography and Kinesiology*, *20*(4), 750–760. doi:10.1016/j.jelekin.2010.02.010.

Watkins, R. G., Uppal, G. S., Perry, J., Pink, M., & Dinsay, J. M. (1996). Dynamic electromyographic analysis of trunk musculature in professional golfers. *The American Journal of Sports Medicine*, *24*(4), 535–8.

Webber, C. L. J., & Zbilut, J. P. (2005). Recurrence quantification analysis of nonlinear dynamical systems. In M. A. Riley & G. C. Van Orden (Eds.), *Tutorials in contemporary nonlinear methods for the behavioral sciences* (pp. 26–94). Retrieved 2012, from http://www.nsf.gov/sbe/bcs/pac/nmbs/nmbs.jsp.

Worsfold, P., Smith, N., & Dyson, R. (2008). Weight transfer time moderation by Golf handicap. Chapter 9. In D. Crews & L. Rafer (Eds.), *Science and Golf V: Proceedings of the World Scientific Congress of Golf* (pp. 67–73). Phoenix, Arizona: The LPGA Foundation.

Zbilut, J. P., & Webber Jr, C. L. (2006). Recurrence quantification analysis. *Wiley Encyclopedia of Biomedical Engineering*. John Wiley & Sons, Inc.

Zhao, Z., Morstatter, F., Sharma, S., Alelyani, S., Anand, A., & Liu, H. (2010). Advancing feature selection research-ASU feature selection repository. Arizona State University. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Advancing+Feature+Selection+Research+-+ASU+Feature+Selection+Repository#0

Zheng, N., Barrentine, S. W., Fleisig, G. S., Andrews, J. R., & States, U. (2008). Kinematic analysis of swing in pro and amateur golfers. *International Journal of Sports Medicine*, 29(6), 487–493. doi:10.1055/s-2007-989229.

CAPÍTULO III

3. Cross-cultural Adaptation and Validation of the Portuguese Survey of Musculoskeletal Conditions, Playing Characteristics and Warm-up Patterns of Golfers

Artigo submetido: ***Journal of Back and Musculoskeletal Rehabilitation***

Silva, L., Castro, M. A., Marta, S., Almeida, J., Vaz, J., Cabri, J., & Pezarat-Correia, P. Cross-cultural adaptation and validation of the Portuguese Survey of musculoskeletal conditions, playing characteristics and warm-up patterns of golfers. *Journal of Back and Musculoskeletal Rehabilitation.*

Cross-cultural adaptation and validation of the Portuguese Survey of musculoskeletal conditions, playing characteristics and warm-up patterns of golfers

Silva, Luís; Faculdade de Motricidade Humana, Universidade de Lisboa

Castro, Maria António; Coimbra College of Health Technology, Polytechnic Institute of Coimbra

Marta, Sérgio; Faculdade de Motricidade Humana, Universidade de Lisboa

Almeida, José; Faculdade de Motricidade Humana, Universidade de Lisboa

Vaz, João; Faculdade de Motricidade Humana, Universidade de Lisboa

Cabri, Jan, The Norwegian School for Sport Sciences, Dept. Physical Performance

Pezarat-Correia, Pedro; Faculdade de Motricidade Humana, Universidade de Lisboa

Abstract

BACKGROUND: The Survey of musculoskeletal conditions, playing characteristics and warm-up patterns of golfers (MSK Golfers) was developed in Canada due to the lack of information concerning musculoskeletal conditions affecting senior golfers. The questionnaire is a short golfer-reported with 37 continuous and categorical variables.

OBJECTIVE: The purpose of this study was to translate and cross-culturally adapt the MSK Golfers into Portuguese and to test it construct validity and reproducibility.

METHODS: The MSK Golfers was translated and then it was tested for psychometric properties. 61 golfers with at least 1 year of practice in golf aged between 14 and 70 years were recruited. Validity of the MSK Golfers was assessed by evaluating data quality (missing, floor and ceiling effects), reproducibility analyses included intraclass correlation coefficients (ICC) (2,1) and Cohen's Kappa coefficient.

RESULTS: The ICC values for continuous items ranged from 0.634 to 0.998 with the exception of one item of Golf Activity items. Kappa statistics for categorical items ranged between 0.714 and 1.00.

CONCLUSIONS: The Portuguese version of the Survey of musculoskeletal conditions, playing characteristics and warm-up patterns of golfers showed a high reliability for a golf population with an age range of 14 and 70 years.

Keywords: Questionnaire, Sport Medicine, Musculoskeletal Injuries, Golf

3.1. Introduction

Golf is practiced by people of several ages, different skill levels, wide amount of practice, and many objectives as competition, leisure, tourism, and marketing events. Although golf is considered a sport with a moderate aerobic component, during the swing, the body segments and club combine an energy chain to transfer to the ball the necessary velocity and trajectory that maximize distance and accuracy [1,2]. Usually, the variables used to measure swing outcome are ball displacement, shot accuracy, club head velocity, and club face angle, being better the performance when golfers focused on those swing outcomes rather than a given body movement [3]. However, to achieve these objectives, the golfer applies extreme muscle strength [4] at maximum ranges of motion [5], and with high velocity joint movement [6,7]. The repeated execution of the swing during a golf course can explain the high incidence of golf-related injuries. Lindsay et al. [8] reported that more than 70% of golf players experienced injuries resulting in playing at an unsatisfactory skill level during a short period of time.

Age-related golf practice, rates of participation, declines in strength, flexibility and coordination, as well the efficacy of conditioning programs for improve performance and prevent injuries are major areas of research in the senior golfer [8,9]. Playing golf provides a stimulus of moderate intensity exercise for seniors, but musculoskeletal injuries can result from unsafe participation and/or due to preexisting musculoskeletal disorders. Cann et al. [9] considered four specifics concerns for the senior golfer: injury rehabilitation coordinated by therapists, warm up routines, club-fitting/coaching on proper technique, and pre-season conditioning programs. However, there is a lack of studies separating the effects of swing modification from physical rehabilitation, as well the effectiveness of changes in swing leading for a decrease in complains as low back pain [10]. The main complaint reported is low back pain, followed by elbow and wrist [6,11,12] but the golf-related practice risks and health benefits have not been fully established and the literature available about golf injuries is merely descriptive and controversial [12]. There is a need for tools to help understand how exercise habits and injuries incidence may be related, about the golf injuries mechanisms. For instance, the role of specific physical activity, skill, and overuse (as repeated swing trials in drive range) in the prevention and rehabilitation of acute and chronic disorders [13] or changes on muscle activation, as erector spinae, in low back pain golfers comparing with asymptomatic [14]. Literature expresses the need for an instrument that categorize the golf population, namely, the senior player [8,15]. Studies about temporal parameters in golfers with and without low back pain used the McGill Pain Questionnaire [14,16]. There was lack of instruments to the golfers' characterization, mostly relating low back pain with the golf practice.

The Survey of Musculoskeletal Conditions, Playing Characteristics and Warm-up patterns of golfers was developed by Fox et al. [15] due to the lack of information concerning musculoskeletal conditions affecting senior golfers. The questionnaire is a short golfer-reported with 37 continuous (eg. golf handicap, minutes of warm-up) and categorical variables (eg. presence or absence of any injuries in the past 3 years during golf activities which lead to stop or modify the golf practice). It takes about 10 minutes to answer and gives information about musculoskeletal conditions affecting players, their playing characteristics and warm-up patterns. Palmer et al. [17] applied this survey to one hundred members of golf courses reporting five groups of musculoskeletal (MSK) conditions: 36.8% with rheumatologic conditions, 18.4% with tendinitis, 18.4% of muscles strains, 10.5% golfers stated vertebral/disk pathologies, and 15.8% answered others individual conditions. Half of the respondents indicated to have suffered from MSK conditions, also a large portion of golfers showed spending a little time with warm-up and stretching before performing the swing.

To be able to characterize Portuguese golf players and their usual behavior on golf, this Survey of Musculoskeletal Conditions, Playing Characteristics and Warm-up patterns of Golfers (MSK Golfers) was translated and cross-culturally adapted into the Portuguese population and tested the Portuguese version for psychometric properties in terms of validity and reproducibility in golfers.

3.2. Methods

3.2.1. Study design

The study was carried out in a two-step procedure; firstly, the MSK Golfers was translated and cross-culturally adapted; and secondly, the Portuguese version of MSK

Golfers was tested for psychometric properties in a cross-sectional design with a 2-week follow-up for test-retest.

3.2.2. Translation and cross-cultural adaptation

Translation and cross-cultural adaptation of the MSK Golfers was performed according to international guidelines [18,19] under license of the author. The English MSK Golfers was translated into Portuguese independently by two Portuguese native translators (one doctor with experience in sports and one professional translator). The obtained translations were discussed in a first consensus panel to achieve the first preliminary version before it was translated back to English.

Two translators and native speaker of English, who were blinded to the original MSK Golfers, performed the back-translation. A second panel consisting of a physiotherapist, a golf coach and the researchers in our group reviewed all translations. The committee discussed discrepancies until consensus on a pre-final version was achieved.

The goal of the pre-final Portuguese MSK Golfers was that it should be as concise and easy to understand as possible. Pre-final Portuguese version of MSK Golfers was tested on six golfers for reviewing. None of the golfers had difficulties in understanding the meaning of items or responses. A suggestion to adapt question number according to Portuguese golf season was made. A third consensus panel completes the final version of MSK Golfers. The Portuguese version is presented in Additional file 1: Appendix.

3.2.3. Participants

A total of 61 golfers were included in the study. Eligible participants were golfers with at least 1 year of practice on golf, aged 14 and over, and were able to speak, read and write in Portuguese. Exclusion criteria were practice shorter than 1 year. Baseline characteristics of the test-retest subgroup are presented in Table 1. Comparing with the original, the Portuguese version considers a higher age range. 6.6% of the participants had 18 years old or lower, 45.9% were between]18,40] years, 31.1% between]40,60], and the other 16.4% up to 60 years of age. The differences found between genders are associated to the Portuguese reality in practicing golf. All participants received written and oral information about the study. Signed informed consent was obtained from all participants. The study was approved by the Ethical Committee of Faculty of Human Kinetics - Technical University of Lisbon (FMH – UTL) in 2011.

Tabela 3 [Table 1 - Subject Characteristics].

	N	Age (years)	Height (m)	Mass (kg)	Experience (years)
Male	56	42.4±14.76	1.75±0.1	78.6±10.6	11.1±6.1
Female	5	34.6±19.9	1.66±0.1	65.6±7.6	8.2±4.7
Total	61	41.8±15.2	1.74±0.1	77.6±11.0	10.9±6.0

3.2.4. Procedures and measures

Participants consent to participate by filled in the MSK Golfers test and retest preferably with a two-week interval.

The MSK Golfers comprises 37 items focusing on golfers habits. The original MSK Golfers was evaluated in senior golfers (above 50 years) and was found to be a reliable and valid measure of golfers' playing conditions [15]; 6 questions of personal categorization, 7 questions about history of illness and injury, 3 questions about the golf game, 2 questions about golf swing, 6 questions about golf course, 4 questions about playing amount, 6 questions about warm-up and exercises and 3 questions about golf injuries.

3.2.5. Statistical analysis

The validation procedures were similar to the original in order to facilitate reproducibility, and comparison. Agreement between answers of the two occasions test retest was assessed with a two-week interval. This period was chosen to avoid biases in the answers due to factors such as fatigue, learning, or memory effects, and close enough periods to avoid genuine changes in the variables that are pretended to measure. Descriptive statistics for quantitative variables as "*handicap*" or "*On average how many yards do you hit your driver*", and proportions of transformed ordinal variables such as "*Typically, how often are you aware of low back pain after golfing 18 holes*?" were presented by mean ± standard deviation. Nominal qualitative variables as "*Do you routinely perform any of your golf stretches away from the course/practice range*?" (Portuguese version) were presented by the respective frequencies.

Validity of the MSK Golfers was assessed by evaluating data quality (missing, floor and ceiling effects). Test-retest reliability was analyzed with the Cohen's Kappa statistic for items with nominal scales and with Intraclass Correlation Coefficients two-way random single measures (2,1) (ICC) for items with quantitative and transformed ordinal scales [20-22]. All data were processed in SPSS 19.0 [23]. Pearson's r was also calculated for items with categorical scales to provide a reference for assessing ICC.

3.3. Results

MSK Golfers showed acceptable psychometric properties in terms of comprehensibility, consistency (agreement), construct validity, and reproducibility when applied to golfers with an age range between 14 - 70 years. The first difference between this cultural validation and the original is the widening of the amplitude of range. The second important aspect is the inclusion of a question to complement the survey part about golf-related injuries, namely the question 35 "*Have you suffered any injuries in the past 3 years while playing or practicing golf, which caused you to stop or modify your game?*". Whether the respondent answers "*yes*", in the Portuguese version he should refer "*How many time have you stopped*". The unit measure of this question is "*days*".

The reliability results of the Past Medical History items showed a high agreement between the two occasions (Table 2) with the item about knowing any *contraindication to activity* showing the lower Kappa value (0.659), not letting, however, being an acceptable reliability. *Heart condition*, *chest pain with activity* and, *chest pain at rest* Kappa's values were not shown because there's a perfect agreement.

Tabela 4 [Table 2 -Reliability of Past Medical History Items].

	Frequency (occ1)		Frequency (occ2)		
	Yes (%)	**No (%)**	**Yes (%)**	**No (%)**	***Kappa***
Heart Condition	-	100	-	100	-
Chest Pain with Activity	-	100	-	100	-
Chest Pain at Rest	-	100	-	100	-
Loss of Balance or Loss of Consciousness	5.1	94.9	10	90.0	0.733
Orthopaedic Problem Aggravated by Activity	16.7	83.3	16.4	83.6	0.880
Ingesting Blood Pressure or Heart Medication	13.1	86.9	11.5	88.5	0.924
Known Contraindication to Activity	3.3	96.7	1.6	98.4	0.659

The items that measure the level of subject (Table 3) showed a high agreement between the two occasions with the worse value being the ICC for the "*On average, how many yards do you hit your driver*" showing an ICC of 0.837. The variable "*Are your golf clubs customized to fit your golf swing*" (fitting) showed a difference between the two occasions of 3.3%, but with the high ICC of 0.997.

Tabela 5 [Table 3 - Reliability of Subject Skill Items].

	Mean±SD (occ1)		**Mean±SD (occ2)**		**ICC**	**r**
Years Played	10.93±6.08		10.89±6.13		0.998	0.998
Driver Range	197.46±28.89		196.39±31.09		0.837	0.838
7-Iron Range	119.34±21.46		120.45±21.42		0.970	0.971
Handicap	17.54±9.25		17.62±9.24		0.997	0.997
	Frequency (%)		**Frequency (%)**		***Kappa***	
	Yes	**No**	**Yes**	**No**		
Fitting	36.1	63.9	32.8	67.2	0.855	
Handed	-	100	-	100	-	
Swing Side	98.4	1.6	98.4	1.6	1.0	

About the Transportation on Golf Course Items (Table 4), only the "*On average, how often do you push your clubs around the course on a cart*?" item had an ICC lower than 0.9 (ICC=0.705). All the others items had ICC values upper than 0.9.

Tabela 6 [Table 4 - Reliability of Transportation on Golf Course Items by proportion of rounds].

	Mean±SD (occ1)	Mean±SD (occ2)	ICC	r
Using Power Cart	20.08±28.93	20.25±28.92	0.992	0.992
Carring Clubs	42.13±40.32	44.18±40.77	0.943	0.944
Pulling Clubs	29.84±32.54	31.48±32.79	0.952	0.952
Pushing Clubs	8.55±18.98	10.74±22.41	0.705	0.714
Using Electric Cart	12.46±29.41	11.15±27.51	0.946	0.948

All Reliability of Frequency of Golf Activity Items (Table 5) showed an ICC upper than 0.7, with the exception of the Class Sessions Off Season (ICC=0.456), being the lowest value found in the entire questionnaire. Reliability of Frequency of Golf Activity items achieved a maximum ICC value of 0.920. For these items it was noted that the lowers handicap tend not to respond because they have a daily practice.

Tabela 7 [Table 5 - Reliability of Frequency of Golf Activity Items].

	Mean±SD (occ1)	Mean±SD (occ2)	ICC	r
Rounds Played				
Early Season	5.60±4.76	5.95±5.49	0.914	0.924
Mid Season	7.33±5.86	8.20±7.26	0.881	0.907
Late Season	5.64±4.53	6.55±6.18	0.773	0.819
Off Season	6.29±6.67	6.31±6.79	0.863	0.862
Practice Sessions				
Early Season	4.71±5.36	4.67±4.7	0.888	0.894
Mid Season	5.42±8.42	5.16±5.42	0.796	0.872
Late Season	4.98±5.52	4.70±4.84	0.917	0.925
Off Season	5.14±6.83	4.95±4.81	0.761	0.805
Putting Freq.				
Early Season	3.55±4.91	3.40±3.67	0.901	0.939
Mid Season	4.09±8.12	3.84±4.58	0.717	0.834
Late Season	3.67±5.02	3.48±4.09	0.920	0.939
Off Season	4.60±7.60	3.76±4.33	0.775	0.906
Class Sessions				
Early Season	1.51±2.89	1.24±1.38	0.634	0.816
Mid Season	1.52±3.62	1.48±2.59	0.903	0.952
Late Season	1.55±2.85	1.28±1.57	0.708	0.837
Off Season	1.76±4.50	1.20±1.34	0.456	0.836

Interestingly, despite the time used for warm-up (Table 6), ranging 9 minutes, this item showed an ICC of 0.793, leading that some golfers had no specific warm-up routine habits. The time of the warm-up used for stretching was 3.95 and 3.64 minutes for the first and second occasion, respectively (ICC=0.832). The higher concordance found for the reliability of warm-up and exercise habits items was in the question "*Do you routinely participate in a cardiovascular conditioning program apart from golfing*?" showing a Kappa of 0.964, but the others nominal variables present a Kappa higher than 0.7.

Tabela 8 [Table 6 - Reliability of Warm-up and Exercise Habits].

	Mean±SD (occ1)		Mean±SD (occ2)		ICC	r
Warm-up	9.15±5.82		9.02±4.66		0.793	0.810
Stretching	3.95±3.61		3.64±3.98		0.832	0.849
	Frequency (%)		**Frequency (%)**		***Kappa***	
	Yes	**No**	**Yes**	**No**		
Stretching During Round	24.6	75.4	27.9	72.1	0.746	
Stretching Away from Course	39.3	60.7	39.3	60.7	0.794	
Strengthening Program	48.3	51.7	52.5	47.5	0.797	
Endurance Program	61.7	38.3	61.0	39.0	0.964	

Table 7 present the reliability for the injury items. The new added question "*How many time have you stopped*" achieved an ICC of 0.672, and the golfers with injuries reported having stopped 25 and 27 days for occasion 1 and 2, respectively. The question that relates low back pain to sports practice also presented a high ICC of 0.905. Both the general injury incidence and musculoskeletal injuries incidence were upper than 20%, showing a high kappa.

Tabela 9 [Table 7 - Reliability of Injury Items].

	Mean±SD (occ1)		Mean±SD (occ2)		ICC	r
Stop Playing because Injury	24.55±9.34		27.27±13.30		0.672	0.714
Low Back Pain after 18 holes	17.54±20.81		18.64±24.91		0.905	0.919
	Frequency (%)		**Frequency (%)**		***Kappa***	
	Yes	**No**	**Yes**	**No**		
Injury Incidence	21.7	78.3	25.4	74.6	0.906	
Musculoskeletal Injury	23.3	76.7	23.3	76.7	0.814	

3.4. Discussion

The MSK Golfers for the Portuguese population showed a high reliability. In this study was taken into account the recommendations stated by Portney & Watkins [21] for ICC values. A good reliability is considered for ICC equal or greater than 0.75, however, the authors stated that for most clinical measurements, reliability should exceed ICC of 0.90 to ensure reasonable validity. Yet, the present instrument it is essential of characterization. In this study the variables measured using ICC, 44.5% were higher than ICC = 0.9 and 72.4% upper than ICC = 0.7. Also, in the original version, the named "*skill*" items were highly reliable with ICC's upper than ICC = 0.90, and "*transportation on the course*" variables ranging between ICC = 0.692 and ICC = 0.921. The Portuguese validation showed for the "skill" ICC's higher than 0.9 and close to "1.0" (table 3), only the drive range had an ICC = 0.837. For the *transportation on golf course items*, the Portuguese version also showed an ICC's greater than 0.9 with the exception of "Proportion of Rounds Push" item (ICC = 0.705).

The reliability of Past Medical History items were similar to the original version with the Kappa showing a high agreement, but the first three items (*heart condition*, *chest pain with activity*, and *chest pain at rest*) in the present study showed a frequency of 100% with the answer "no" in both occasions. The Kappa values for these variables are not presented because it is determined by cross-tabulating ratings for two coders, and the agreement expected by chance is determined by the marginal frequencies of each rate [20]. In the original version, the respondents reporting "*orthopedic problem aggravated by activity*" were more than double (39.5%) comparing with the present study (16.7%). This group of items was based on the Physical Activity Readiness (PAR-Q) form from Public Health Agency of Canada and the Canadian Society for Exercise Physiology [24] validated for people with age ranging from 15 to 69 years. In this study one respondent had 14 years, whereby is a limitation. Fox et al. [15] stated similar age-related limitation with the Par-Q (British Columbia Ministry of Health, 1978 cited [15]), because their version was validated for individuals aged 18 to 69 and they had participants aged 70 years and older. They reported that the use of Par-Q only as an accepted measure of general health, not as a screening device and not critical in these study types.

Comparing Portuguese and Canadian versions relative to "*warm-up and exercise habits*", Portuguese spend more time warming-up prior to playing, with 9.15±5.82

minutes in first occasion and 9.02±4.66 minutes in the second occasion, while the Canadian golfers mean values on occasion one and two were 4.47±0.59 and 4.07±0.54 minutes, respectively. The ICC's both version were similar, ICC = 0.793 in the Portuguese version and ICC =0.750 for the Canadians. The amount of this period which was devoted to stretching showed for occasions one and two, 3.95±3.61 and 3.64±3.98 for the Portuguese's (ICC = 0.832) and 2.90.49 and 2.7±0.50 minutes for Canadian's, respectively (ICC = 0.791). Palmer et al. [17] considered seniors golf warm-up activities surprisingly limited. The time spent on both warm-up (16.6%) and stretching (36%) was less than 1 minute, and over than 75% of the total participants spent less than five minutes. Golfers with MSK conditions spend more time warming-up than those who not have experienced those problems.

The imbalance in the relationship between the sample sexes, besides being related to the actual Portuguese golf population. In the literature it can be found sample differences with more percentage of males' golfers [25,26], but shoulder, knee and elbow injuries associated to overuse were found to be similar [11,27]. The findings of the most likely area of injury in female golfers were elbow and lower back, related with overuse, poor conditioning, and incorrect swing mechanism [28].

In this study, the injury incidence was low comparing with the data report by Lindsay et al [8] and Fox et al [15]. However, both validations showed a relative absence of self-identified musculoskeletal injuries linked to golf participation. These differences could be associated with the inclusion of young athletes in the present study. Fox et al. [15] had a participants with an age of 50 years and older, and one-half of their respondents stated a preexisting musculoskeletal condition affecting golf practice. Cabri et al. [12] reported that the persistence of golf practice for long periods could be the main reason for the prevalence of injuries in ages between 50 and 65 years. This population has more probability of injury due to relevant changes as decrease strength, flexibility, coordination and balance. Also, many of the senior golfers have already experienced preexisting MSK conditions [17]. Recreational golfers were not exposed to the repetitive stresses experienced by professional players who spend significant time practicing at the driving range, the putting green, and playing the course [15,28]. Recreational golfer usually had more options varying the type of swing, way of transport and total duration of play [17].

Controversially, injury rates in amateur golfers may vary between mean values of 1.19 to 1.31 per year, while the professionals had higher prevalence with ratios near 2 injuries per year due to the time spend playing [27], rather than technical aspects or warm-up habits [28,29].

The health benefits of golf are related with the amount of physical activity during the course, as the way of transportation. The respective items showed a higher reliability than the original. In the Canadian version, the proportion using power cart was approximately 71%, unlike the 17% reported by Palmer et al. [17] and 20% of the present study, according to which the respondents carry their clubs approximately 43% of the rounds. Similar, Palmer et al. [17] reported that two-thirds of the respondents referred a preference for walking the golf course at least 50% of the time. The careless driving and inadequate use of power carts were associated to traumatic injuries [30,31].

The question about the incidence of low back pain after 18 holes showed a mean value of about 18% of the time. This result is lower than expected, since low back pain is the major complaint among golfers [32-35] being the main cause for a reduction in playing and skill level decrease in elite golfers [32]. Nevertheless, it is necessary take into account the specificity of the question when it ask “after 18 holes”, in other words, after a golf course. Besides the question specify, Palmer et al. [17] also found higher incidence of low back pain with 42% of the respondents with complaints.

3.5. Conclusion

The Survey of Musculoskeletal Conditions, Playing Characteristics and Warm-up Patterns of Golfers for the Portuguese population showed a high reliability. The question about the golf practice time lost due to an injury included showed a good reliably as well. In the Portuguese version the age range is larger than the original version, but the agreement remains. This survey demonstrates an excellent applicability and reliability being valid and useful for the characterization of the Portuguese golf population.

Acknowledgments

The project “Neuromuscular activity in the golf swing with implications for the practice and in the prevention of overuse injuries” was supported by the Portuguese Foundation for Science and Technology (FCT) fund (PTDC/DES/105176/2008).

The authors would like to thank to David Lindsay and Filomena Carnide for their comments on the manuscript.

References

[1] Farrally MR, Cochran AJ, Crews DJ, Hurdzan MJ, Price RJ, Snow JT, et al. Golf science research at the beginning of the twenty-first century. J Sports Sci. 2003; 21(9): 753-765.

[2] Hume PA, Keogh J, Reid D. The role of biomechanics in maximising distance and accuracy of golf shots. Sports Med. 2005; 35(5):429-49.

[3] Keogh JW, Hume PA. Evidence for biomechanics and motor learning research improving golf performance. Sports Biomech. 2012; 11(2):288-309.

[4] Hosea TM. and Gatt CJ. Back pain in golf. Clin Sports Med. 1996; 15: 37-53.

[5] Hosea TM, Gatt CJ, Galli KM, Langrana NA, Zawadasky JP. Biomechanical analysis of the golfer's back. In: Cochrane AJ. Science and Golf. London: Chapman and Hall; 1990. pp. 43-48.

[6] McCarroll M, Retting AC, Shelbourne K. Injuries in the amateur golfer. Phys Sportsmed. 1990; 18: 122-126.

[7] Zheng N, Barrentine SW, Fleisig GS, Andrews JR. Kinematic Analysis of Swing in Pro and Amateur Golfers. Int J Sports Med. 2008; 29(6): 487-493.

[8] Lindsay DM, Horton JF, Vandervoort AA. A Review of Injury Characteristics, Aging Factors and Prevention Programmes for the Older Golfer. Sports Med. 2000; 30(2): 89-103

[9] Cann AP, Vandervoort AA, Lindsay DM. Optimizing the benefits versus risks of golf participation by older people. J Geriatr Phys Ther. 2005; 28(3): 85-92.

[10] Gluck GS, Bendo JA, Spivak JM. The lumbar spine and low back pain in golf: a literature review of swing biomechanics and injury prevention. Spine J, 2008; 8(5), 778-788.

[11] Gosheger G, Liem D, Ludwig K, Greshake O, Winkelmann W. Injuries and overuse syndromes in golf. Am J Sports Med. 2003; 31(3): 438-443

[12] Cabri J, Sousa J, Kots M, Barreiros J. Golf-related injuries: a systematic review. Eur J Sport Sci. 2009; 9(6):353–66.

[13] Suckel A. [Sports medicine analysis of golf "swing" and lesions occurring during golf practice]. portverletz Sportschaden. 2002; 16(1):31-5

[14] Cole MH, Grimshaw PN. Electromyography of the trunk and abdominal muscles in golfers with and without low back pain. J Sci Med Sport. 2008; 11(2): 174-181.

[15] Fox E, Lindsay DE, Vandervoort AA. Musculoskeletal Injury Questionnaire for Senior Golfers. Science and Golf IV. Proceedings of the World Scientific Congress of Golf. St. Andrews, Scotland, July 22-26, 2002. pp. 88-99.

[16] Horton JF, Lindsay DM, Macintosh BR. Abdominal muscle activation of elite male golfers with chronic low back pain. Med Sci Sports Exerc. 2001; 33(10): 1647-1654.

[17] Palmer JL, Young SD, Fox E, Lindsay DM., Vandervoort, A. A. Senior Recreational Golfers: a survey of musculoskeletal conditions, playing characteristics, and warm-up patterns. Physiother Can. 2003; 55(2): 79-85.

[18] Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures. Spine. 2000; 25:3186-3191.

[19] Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol. 1993; 46:1417-1432.

[20] Hallgren, KA. Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. Tutor Quant Methods Psychol. 2012; 8(1): 23–34.

[21] Portney L, Watkins M. Foundations of Clinical Research: Applications to Practice. 3rd ed. Upper Saddle River, New Jersey: Pearson Prentice Hall; 2009.

[22] Shrout PE, Fleiss JL. Intraclass Correlations: uses in assessing rater reliability. Psychol Bull. 1979; 86(2):420-428.

[23] McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. Psychol Methods. 1996; 1(1):30–46.

[24] ACSM. Acsm's Guidelines for Exercise Testing and Prescription. 8th ed. Baltimore: Lippinkott Williams & Wilkins; 2010.

[25] Bechler J, Jobe F, Pink M, Perry J, Ruwe P. Electromyographic analysis of the hip and knee during the golf swing. Clin J Sport Med. 1995; 5(3):162–5.

[26] Bulbulian R, Ball K, Seaman D. The short golf backswing: effects on performance and spinal health implications. J Manipulative Physiol Ther. 2001; 24(9):569–75.

[27] Theriault G, Lachance P. Golf injuries. An overview. Sports Med. 1998; 26(1): 43-57.

[28] McCarroll JR. The frequency of golf injuries. Clin Sports Med, 1996; 15(1): 1-7.

[29] Metz JP. Managing Golf Injuries. Technique and equipment changes to aid treatment. Phys Sportsmed. 1999; 27(77), 41-58.

[30] Rahimi SY, Singh H, Yeh DJ, Shaver EG, Flannery AM, Lee MR. Golf associated head injury in the pediatric population: a common sports injury. J Neurosurg. 2005; 102(2 Suppl): 163-166.

[31] Watson DS, Mehan TJ, Smith GA, McKenzie LB. Golf cart-related injuries in the U.S. Am J Prev Med. 2008; 35(1): 55-59.

[32] Finch C, Sherman C, James T, Farrally M, Cochran A. The epidemiology of golf injuries in Victoria, Australia: evidence from sports medicine clinics and emergency department presentations. In: M. Farrally & A. Cochran, editors. Science and Golf III: Proceedings of the World Scientific Congress of Golf. Champaign, Ill.: Human Kinetics; 1999. pp. 73-82.

[33] McHardy A, Pollard H, Luo K. Golf injuries: a review of the literature. Sports Med. 2006; 36(2):171–87.

[34] Sugaya H, Tsuchiya A, Moriya H, Morgan DA, Banks SA. Low back injury in elite and professional golfers: An epidemiologic and radiographic study. In: M. Farrally & A. Cochran, editors. Science and Golf III: Proceedings of the World Scientific Congress of Golf. Champaign, Ill.: Human Kinetics; 1999. pp. 83-91.

[35] Vad VB, Bhat AL, Basrai D, Gebeh A, Aspergren DD, Andrews JR. Low back pain in professional golfers: the role of associated hip and low back range-of motion deficits. Am J Sports Med. 2004; 32(2): 494-497.

Appendix 1 Portuguese version

Universidade Técnica de Lisboa
Faculdade de Motricidade Humana

Questionário sobre condições músculo-esqueléticas em jogadores de golfe adultos
(Fox, Lindsay & Vandervoort,2002)

Em primeiro lugar, gostaríamos de lhe fazer algumas perguntas sobre si:

1. Qual é a sua data de nascimento? ___________ _____________ ______________
(dia) (mês) (ano)

2. Qual é a sua altura? ____________________________cm

3. Qual é o seu peso? ____________________________kg

4. Sexo masculino ou feminino? (deve escolher a opção adequada assinalando-a com um círculo no número correspondente à sua resposta)
   1. Masculino
   2. Feminino

5. Há quantos anos joga golfe? ___________ anos/ meses **(riscar o que não interessa)**

6. É destro ou esquerdino? (Círculo)
   1. Destro
   2. Esquerdino

Gostaríamos agora de o questionar acerca do seu histórico de doenças e lesões:

7. O seu médico alguma vez lhe disse que tinha um problema cardíaco e que só deveria fazer actividade física recomendada por um médico? (Círculo)
   1. Sim
   2. Não

8. Sente dor não peito quando faz exercício físico? (círculo)
   1. Sim
   2. Não

9. No último mês, alguma vez teve dor no peito quando não estava a fazer actividade física? (Círculo)
   1. Sim
   2. Não

10. Alguma vez sentiu tonturas chegando a perder o equilíbrio ou alguma vez desmaiou?
    1. Sim
    2. Não

11. Tem algum problema ósseo ou articular que possa ser agravado pela prática de actividade física? (Círculo)
    1. Sim
    2. Não

12. Toma medicamentos (por exemplo, diuréticos) para a tensão arterial ou para problemas cardíacos? (Círculo)
    1. Sim
    2. Não

13. Conhece algum outro motivo que o possa impedir de fazer exercício físico? (círculo)
    1. Sim
    2. Não

Em caso afirmativo, especifique, por favor: ___________________________________

De seguida, gostaríamos de fazer algumas perguntas sobre o seu jogo/prática de golfe:

14. Em média, quantos metros bate com o seu driver? ________________________

15. Em média, quantos metros bate com o seu ferro 7? ________________________

16. Qual é o seu handicap? _______________________

(Refira o seu handicap aproximado, caso não tenha um oficial)

Gostaríamos agora de o questionar sobre o seu swing:

17. O seu swing é destro ou esquerdino? (Círculo)
    1. Esquerdino
    2. Destro
18. Os seus tacos de golfe estão adaptados (fitting) ao seu swing? (Círculo)
    1. Sim
    2. Não

Gostaríamos de perguntar como se desloca no campo de golfe:

19. Em média, com que frequência usa o *buggie* durante o percurso de jogo? (Círculo)

0% do tempo ...... 15% .... 30% .... 50% .... 65% .... 80% .... **100% do tempo**

20. Em média, com que frequência transporta os seus tacos durante o percurso de jogo? (Círculo)

0% do tempo ...... 15% .... 30% .... 50% .... 65% .... 80% .... **100% do tempo**

21. De que lado do corpo transporta os seus tacos? (Círculo)
    1. Esquerda
    2. Direito
    3. Alternado
    4. Duas alças
22. Em média, com que frequência puxa os seus tacos num trolley durante o percurso de jogo? (Círculo)

0% do tempo ...... 15% .... 30% .... 50% .... 65% .... 80% .... **100% do tempo**

23. Em média, com que frequência empurra os seus tacos num trolley durante o percurso de jogo? (Círculo)

0% do tempo ...... 15% .... 30% .... 50% .... 65% .... 80% .... **100% do tempo**

24. Em média, com que frequência usa um trolley eléctrico para transportar os seus tacos? (Círculo)

0% do tempo ...... 15% .... 30% .... 50% .... 65% .... 80% .... **100% do tempo**

Gostaríamos de saber qual a sua regularidade na prática de golfe.

25. Em média, quantas voltas joga num só mês durante as seguintes épocas:

Época inicial (Abril-Maio) _______________________

Época média (Junho-Agosto) _______________________

Época final (Setembro-Outubro) _______________________

Fora de época (Novembro-Março) _______________________

26. Em média, quantas vezes num só mês vai para o *driving range* treinar durante as seguintes épocas:

Época inicial (Abril-Maio) _______________________

Época média (Junho-Agosto) _______________________

Época final (Setembro-Outubro) _______________________

Fora de época (Novembro-Março) _______________________

27. Em média, quantas vezes num só mês treina o putting?

Época inicial (Abril-Maio) _______________________

Época média (Junho-Agosto) _______________________

Época final (Setembro-Outubro) _______________________

Fora de época (Novembro-Março) _______________________

28. Em média, quantas vezes num só mês tem aulas com um profissional de golfe?
Época inicial (Abril-Maio) ________________________
Época média (Junho-Agosto) ________________________
Época final (Setembro-Outubro) ________________________
Fora de época (Novembro-Março) ________________________

Gostaríamos de o questionar sobre o seu aquecimento.

29. Em média, quanto tempo dedica ao seu aquecimento antes de jogar ou praticar? __________________ minutos
30. Quanto tempo do aquecimento dedica a alongamentos? ____________________minutos
31. Uma vez iniciada uma volta, costuma realizar alguns alongamentos específicos para o golfe? (círculo)
    1. Sim
    2. Não

Agora, apenas algumas perguntas sobre outros exercícios que possa fazer.

32. Por rotina, executa alguns dos seus alongamentos específicos para o golfe quando não está a praticar golfe? (círculo)
    1. Sim
    2. Não
33. Costuma realizar exercícios de força? (círculo)
    1. Sim
    2. Não
34. Costuma realizar treino cardiovascular, para além do golfe? (círculo)
    1. Sim
    2. Não

Finalmente, gostaríamos de lhe perguntar sobre as lesões sofridas durante a prática de golfe.

35. Sofreu ALGUMA lesão nos últimos 3 anos, durante o jogo ou prática de golfe, que o tenha levado a parar ou modificar o seu jogo? (círculo)
    1. Sim _______________ (refira o tempo de paragem) ____________ (dias)
    2. Não

em caso afirmativo, diga-nos qual/quais a(s) parte(s) e lado do seu corpo ficou lesionada

____

e a designação clínica ou diagnóstico de cada lesão, se souber (ex., “cotovelo de tenista”, desconforto lombar)

____

36. Por norma, com que frequência sente dor lombar depois de fazer 18 buracos? (círculo)
0% do tempo …15% .... 30% .... 50% .... 65% .... 80%............ 100% do tempo
37. Sofreu ALGUM problema muscular ou articular nos últimos 3 anos que tenha afectado a sua prática de golfe? (círculo)
    1. Sim
    2. Não

em caso afirmativo, diga-nos qual/quais a(s) parte(s) e lado do seu corpo ficou afectado(s)

____

e a designação clínica ou diagnóstico de cada problema, se souber.

____

Muito obrigado por ter respondido a este questionário

CAPÍTULO IV

4. Trunk Muscle Activation during Golf Swing: Baseline and Threshold

Publicado no ***Journal of Electromyography and Kinesiology***

Silva, L., Marta, S., Vaz, J., Fernandes, O., Castro, M. A., Correia-Pezarat, P., António, M., et al. (2013). Trunk muscle activation during golf swing: Baseline and threshold. *Journal of Electromyography and kinesiology*, *23*(5), 1174–1182.

Abstract

Journal of Electromyography and Kinesiology 23 (2013) 1174–1182

Contents lists available at SciVerse ScienceDirect

Journal of Electromyography and Kinesiology

journal homepage: www.elsevier.com/locate/jelekin

ELSEVIER

Trunk muscle activation during golf swing: Baseline and threshold

Luís Silva [a,*], Sérgio Marta [a], João Vaz [a], Orlando Fernandes [b], Maria António Castro [c], Pedro Pezarat-Correia [a]

[a] Faculty of Human Kinetics, Technical University of Lisbon, Portugal
[b] University of Évora, Portugal
[c] Coimbra College of Health Technology, Polytechnic Institute of Coimbra, Portugal

ARTICLE INFO

Article history:
Received 5 October 2012
Received in revised form 24 April 2013
Accepted 28 May 2013

Keywords:
EMG
Threshold
Trunk
Onset
Golf swing

ABSTRACT

There is a lack of studies regarding EMG temporal analysis during dynamic and complex motor tasks, such as golf swing. The aim of this study is to analyze the EMG onset during the golf swing, by comparing two different threshold methods. Method A threshold was determined using the baseline activity recorded between two maximum voluntary contraction (MVC). Method B threshold was calculated using the mean EMG activity for 1000 ms before the 500 ms prior to the start of the Backswing. Two different clubs were also studied. Three way repeated measures ANOVA was used to compare methods, muscles and clubs. Two way mixed Intraclass Correlation Coefficient (ICC) with absolute agreement was used to determine the methods reliability.

Club type usage showed no influence in onset detection. Rectus abdominis (RA) showed the higher agreement between methods. Erector spinae (ES), on the other hand, showed a very low agreement, that might be related to postural activity before the swing. External oblique (EO) is the first being activated, at 1295 ms prior impact. There is a similar activation time between right and left muscles sides, although the right EO showed better agreement between methods than left side. Therefore, the algorithms usage is task and muscle dependent.

4.1. Introduction

Several approaches have been proposed for EMG onset detection; however there is no standardized method and its application is mainly done in motor skills with isometric contraction (Farina and Merletti, 2000), as they present better reproducibility (Lee et al., 2011). A complex motor skill such as a golf swing combines both power and precision. The purpose of the golfer is to place a ball inside a small hole with the least hits possible (Hume et al., 2005). Although it is not considered an intensive and exhausting sport, skeletal-muscle stress and demand are associated with high injury incidence (Cabri et al., 2009).

The study of onset muscle activity can provide information regarding the temporal organization and coordination of a set of muscles at use during a task (De Luca, 1997). In explosive and precise motor tasks, as throwing, the trunk muscles sequence plays an important role in the organization of the proximo-distal sequence in order to transfer energy (Hirashima et al., 2002). This mechanism leads to an increase of speed in distal segments. The movement of different body segments will depend on the motor programming of the central nervous system, which translates into a specific sequence, intensity and muscle time activation. In subjects with low back pain, the reaction time (activation to movement initiation) of abdominal muscles tends to increase as upper limb task complexity increases, due to postural organization (Hodges, 2001).

Most studies on trunk muscle EMG activity during a golf swing have focused on intensity parameters (Pink et al., 1993; Watkins et al., 1996). Only two studies have analyzed the EMG activation onset (Horton et al., 2001; Cole and Grimshaw, 2008). Both have used a threshold detection algorithm, and compared trunk muscles between symptomatic and asymptomatic golfers' lower back pain. Horton et al. (2001) used seven standard deviations (SDs) above baseline, with a 200 ms window (i.e. time interval considered for a group of samples). Although they did not find differences for the amplitude of abdominal activity between the two groups, asymptomatic subjects activated the left external oblique (EO) significantly earlier than the symptomatic, in respect to the start of the backswing. Cole and Grimshaw (2008) have set the onset at 1 SD above baseline, with a 50 ms moving window. Their results did not present significant differences between the two groups for EO, but the erector spinae (ES) was activated significantly sooner for golfers with low back pain. The difficulty in comparing studies is related to the different algorithms criteria, which compromises the reproducibility of the results (Morey-Klapsing et al., 2004; Jöllenbeck, 2000). This is particularly evident for threshold algorithms (Staude et al., 2001).

Onset detection can be divided into two categories: visual inspection (VI) and detection algorithms (Vaisman et al., 2010; Hug, 2011). Visual inspection requires a time-consuming work and the precision of the results depends on the researcher's expertise; therefore being a subjective process (Jöllenbeck, 2000) and its use being rather paradigmatic. However, the lack of a goldstandard measurement used to validate the algorithms, leads to visual inspection being used to assess the precision of threshold algorithms. Algorithms Detection can be classified into threshold algorithms (Van Boxtel et al., 1993; Hodges and Bui, 1996; Jöllenbeck, 2000; Allison, 2003) and as statistically optimized algorithms (Micera et al., 1998; Staude et al., 2001), as maximum likelihood.

The usual definition of onset refers to the initial activity register of the motor units' action potentials (Solnik et al., 2010). The different phases that make up complex motor skills would require different approaches for the meaning of EMG signal. McGill et al. (2010) characterizing a double-peak intensity phenomenon in motor skills such as kicking in martial arts. This phenomenon could be associated with the muscular actions during the different phases of those tasks. For golf swinging several phases can be discriminated, such as the preparation (backswing), execution (downswing) and result (follow-through). Some authors have opted to include descriptive and qualitative movement analysis, due to activity characteristics and particular muscle actions (Hirashima et al., 2002; McGill et al., 2010).

The precision with which a certain algorithm detects the onset is influenced by the background activity level, signal-to-noise ratio activity (Hodges and Bui, 1996; Staude et al., 2001), and onset rate of signal amplitude (Allison, 2003). Hug (2011) states that threshold algorithms vary in 1, 2, and 3 SD or between 15% and 25% of the activity's maximum peak. Other threshold algorithm approaches have considered onset to be the moment in which signal voltage/ intensity surpasses the confidence interval upper limit in a fixed number of samples (Van Boxtel et al., 1993). Hodges and Bui (1996) which have compared onset detection algorithms with different options of low pass filters 10, 50, 500 Hz combined with different sampling windows 10, 25, 50 ms and standard deviations 1, 2, 3 SD for different background activity levels. The most adequate combinations for cutoff frequency, sample window and SD were are 50 Hz/25 ms/3SD and 50 Hz/50 ms/1SD. This clearly demonstrated that excessive smoothing leads to loss of information, and that insufficient smoothing is associated with an onset detection delay.

Parameters knowledge on what constitutes the detection of algorithms is crucial on EMG temporal analysis. However, this analysis should not be restricted to isometric contractions. Temporal activity should take into account the dynamic motor skills phases, identifying key moments of motor coordination.

Golfers often wonder whether the swing is always the same when using different clubs. Swing phase time seems to be similar, but the club speed could be different (Egret et al, 2003), although there is a lack of knowledge on the activation timing in using different clubs.

The aim of this study is to analyze the temporal activity during the golf swing given the preparation phase (backswing) and execution phase (downswing) by comparing the use of two different baselines, activity threshold methods and visual inspection. Moreover, we intend to investigate whether or not the usage of different clubs leads to changes in the onset detection.

4.2. Method

4.2.1. Participants and task

Eight male right-handed amateur golfers (52.0 ± 7.4 years old; handicap of 15.7 ± 3.2) were instructed to perform five precision swings with pitching (<100 m) and five long range swings with iron 4 (>150 m) in an alternate sequence (n = 80). Before any experimental procedure, subjects were allowed to perform some repetitions, in order to enable a better adaptation to the task and to warm up. The swings were carried out on top of an artificial grass golf carpet with high absorption features. Subjects did not have any limitations for playing golf. All the procedures were explained and a consent form was signed. This study was approved by the Ethics Committee of the Faculty of Human Kinetics (Technical University of Lisbon).

4.2.2. Video recording and analysis

Three high speed Basler A602fc cameras (Basler Vision Technologies, Ahrensburg, Germany) at 100 Hz were placed in position as to determine swing phases. A fourth Casio Ex-FH20 camera (Casio, Tokyo, Japan) at 1000 Hz was placed in front of the ball, in order to determine the instant of impact. Two reflective tapes (Horton et al., 2001) were placed on the club to divide the swing in three phases (Bechler et al., 1995; Pink et al., 1993; Watkins et al., 1996). (1) Backswing – from the beginning until the top of the swing; (2) Downswing – from the top until impact; and (3) Follow-Through – from impact until the end of the swing. SIMI 3D Motion system (SIMI Reality Motion System GmbH, Unterschleissheim, Germany) was used for EMG-synchronized 3D kinematic analysis.

4.2.3. EMG procedures

EMG data was collected with active surface electrodes (Al/AgCl, disk shape 10 mm of diameter) and bioPLUX_ research 2010 telemetric equipment (Plux, Lisbon, Portugal). EMG data was collected with a 1000 Hz sampling frequency, amplified with a bandpass between 10 and 500 Hz, common-mode rejection ratio (CMRR) of 110 dB and input impedance greater than 100 MX. After stored, data was digitally filtered (10–490 Hz) and, full-wave rectified. Smoothing with a low pass filter (12 Hz, Butterworth 4th order digital filter) was applied and submitted to visual inspection comparison. EMG data processing was performed with MATLAB_ V.R2010a software (Mathworks Inc., Natick Massachusetts, USA). Skin was properly prepared by means of hair removal, abrasion and alcohol cleaning. The electrodes were placed with a 20 mm center-to-center distance and applied in parallel to the muscle fibers: rectus abdominis (RA), 3 cmlaterally from the umbilicus; external oblique (EO), 15 cm laterally from the umbilicus; erector spinae (ES), 3 cm laterally from the L3 spinous process (Horton et al., 2001). Muscle contraction was performed in order to visualize the muscle belly. The ground electrode was placed on the manubrium.

Three to four second-long maximum voluntary contractions (MVCs) were collected to determine baseline activity between two maximum voluntary contractions (MVCs):

RA – in supine position, the participant performed trunk flexion at 30_, keeping the knees at 90_ and the hip at 70_, with a researcher applying resistance on the shoulders, while another researcher bilaterally stabilized the lower limbs; EO – in lateral position, with hands on the chest and flexed legs (stabilized), the participant produced a lateral trunk flexion against the resistance presented by the researcher; ES – in prone position, with lower limbs stretched and pelvis fixated, the participant performed trunk extension against the bilateral shoulder resistance presented by the researcher (Konrad, 2005; McGill, 1991). The participants were verbally encouraged during these tasks to improve their performance.

4.2.4. Onset detection

Two methods for onset detection were used, the difference being on the threshold determination. With method A, the mean for threshold was determined with the baseline activity registered between two MVC (Fig. 1 A). Determination with method B involved the mean EMG activity for 1000 ms before the 500 ms prior to the start of the Backswing (Fig. 1 B). The threshold value was derived using 3 SD above the baseline mean amplitude. Two additional criteria 1. and 2. were applied after threshold calculation for both methods (A and B): 1. the search for onset starts 150 ms before Backswing; 2. The search for a second onset starts 150 ms before Downswing. Then a 50-sample window moved until a positive derivative was obtained, in addition to having met the established threshold. These restrictions were used to ensure that the method selects an onset related to a peak activity instead of slightly signal oscillations. The reference for onset calculation was the impact. EMG onset was also detected through visual inspection performed by two experienced researchers.

4.2.5. Statistical analysis

Data were statistically processed with IBM SPSS Statistics 19.0 (IBM Corporation, New York, USA). Descriptive statistics were presented with average ± standard deviation and coefficient of variation. Normality test of the data was performed with the Shapiro–Wilk test. Three-way repeated measures ANOVA was used to compare methods, muscles and clubs. Two-way mixed Intraclass Correlation Coefficient (ICC) with absolute agreement was used to determine the methods reliability. Repeated measures ANOVA Multiple comparisons were done with Bonferroni test, and sphericity was assessed with Mauchly's test. When there was no sphericity, the degrees of freedom were corrected with Greenhouse–Geisser test. The significance level was set at 5%.

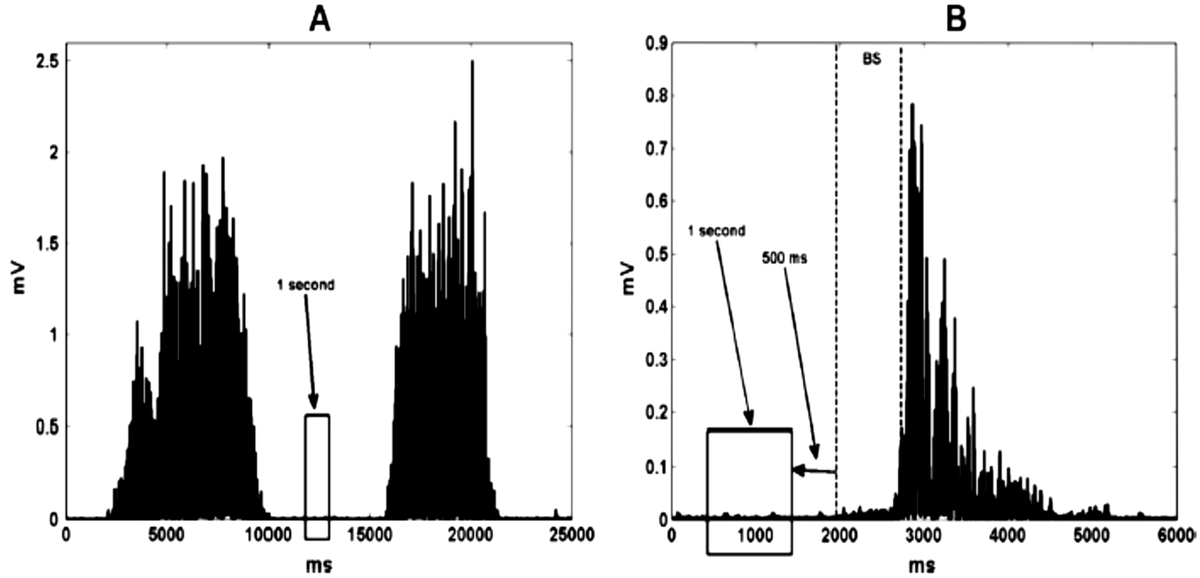


*Figura 4 [**Fig. 1**. (A) Baseline MVC method A; (B) Baseline trial method B; BS: Backswing].*

4.3. Results

Table 1 shows the descriptive for onset for method A, method B and VI for the restriction 150 ms before the backswing (A1, B1 and VI1) and for restriction 150 ms before the downswing (A2, B2, VI2). Data are present for onset results (ms) and percentage of onset detection related to the EMG maximum activity (%Pmax) across muscles side and club. In Fig. 2 it can be seen that the restriction 1 corresponds to the onset search starting before backswing (called onset burst), and the restriction 2 the search starts before downswing (called onset peak) due to a phase change in motor task. Three-way ANOVA showed no significant differences between clubs on the detection of the onset detection with the restriction before the backswing (right side muscles (R) - $F_{(1,7)}$=2.552, p=0.154, η_p^2=0.267, π=0.282; left side muscles (L) - $F_{(1,11)}$=1.910, p=0.194, η_p^2=0.148, π=0.244). For the onset detection with the restriction before the downswing no significant differences were found (α=0.05/8) either for the club and both muscles sides (R - $F_{(1,35)}$=5.035, p=0.031, η_p^2=0.126, π=0.588; L - $F_{(1,30)}$=4.843, p=0.036, η_p^2=0.139,

π=0.568). There were no significant interactions of the club in onset detection with restriction 1 and 2 in relation to muscles (R - $F_{(1.104,7.727)}$=3.247, p=0.109, η_p^2=0.317, π=0.523; L - $F_{(2,22)}$= 1.724, p=0.202, η_p^2=0.136, π=0.322) (R - $F_{(1.408,49.266)}$=4.980, p=0.009, η_p^2=0.125, π=0.686; L - $F_{(2,60)}$= 1.803, p=0.174, η_p^2=0.057, π=0.363) neither to the methods (R - $F_{(2,14)}$=1.976, p=0.175, η_p^2=0.220, π=0.340; L - $F_{(2,22)}$= 0.441, p=0.649, η_p^2=0.039, π=0.113) (R - $F_{(1.490,52.152)}$=3.280, p=0.059, η_p^2=0.086, π=0.519; L - $F_{(2,60)}$= 0.012, p=0.988, η_p^2=0.057, π=0.363), respectively.

Significant differences were found in onset detection for the three muscles with both restrictions, 1. (R - $F_{(2,14)}$=5.985, p≤0.001, η_p^2=0.461, π=0.798; L: $F_{(2,22)}$= 47.482, p≤0.001, η_p^2=0.812, π=1.0) and 2. (R - $F_{(2,70)}$=151.941, p≤0.001, η_p^2=0.813, π=1.0; L: $F_{(1.568,60)}$= 82.583, p≤0.001, η_p^2=0.734, π=1.0). The interaction between muscles and methods were also significant for both, restriction 1 (R - $F_{(4,28)}$=13.819, p≤0.001, η_p^2=0.664, π=1.0; L: $F_{(1.842,20.258)}$=9.395, p≤0.001, η_p^2=0.461, π=0.999), and restriction 2 (R - $F_{(2.537,88.799)}$=68.524, p≤0.001, η_p^2=0.662, π=1.0; L: $F_{(2.648,79.431)}$=71.085, p≤0.001, η_p^2=0.703, π=1.0), showing that onset detection is influenced by the method and depend of the studied muscle. Multiple comparisons showed a similarity between method B and VI for the general data of right side for onset detection with restriction 1 (p=1.0).

Figure 3 shows the errors bars for onset detections in the RA, EO and ES with method A, method B and visual inspection. The results for ICC, coefficient of variation and Inter-Item Correlation Matrix are presented in table 2.

4.4. Discussion

This study analyzed the temporal activity during the golf swing, given the phases of preparation and execution, with the traditional definition for onset, but also knowing important timing moments against different muscle actions that are associated with these swing phases. As the algorithms' parameters influence the results, we compare two different threshold baseline-related algorithms and visual inspection. Additionally, it was also verified if the club usage influences the performance method in onset detection and timing parameters of the trunk muscles. This problem arose from a sequence of influences derived from many factors. The first factor refers to the difficulty in comparing different studies, due to a low reproducibility method (Allison, 2003; Morey-Klapsing et al., 2004; Staude et al., 2001), when there is some diversity among algorithms (Hug, 2011).

Tabela 10 [Table 1- Descriptive mean and standard deviation onset type and method].

Method/ Onset Type			A1		B1		VI1
			Onset (ms)	EMG peak %	Onset (ms)	EMG peak %	Onset (ms)
Muscle	Side	Club	Mean [Std. Deviation]				
RA	Right	4-iron	-1320.24 [276.44]	3.96 [3.59]	-1092.56 [189.91]	4.58 [3.41]	-1039.00 [87.12]
		Pitch	-1332.44 [230.03]	4.40 [4.07]	-1129.59 [190.71]	4.94 [4.24]	-1031.62 [103.68]
	Left	4-iron	-1062.78 [274.94]	10.09 [7.57]	-1003.19 [256.78]	10.69 [8.32]	-914.81 [138.83]
		Pitch	-1020.80 [271.39]	11.49 [9.60]	-883.73 [196.81]	12.39 [9.62]	-894.83 [108.35]
EO	Right	4-iron	-1359.05 [270.83]	5.98 [3.74]	-1402.38 [259.07]	5.09 [3.35]	-1347.92 [240.48]
		Pitch	-1289.76 [310.11]	6.37 [4.71]	-1335.63 [273.05]	5.56 [3.87]	-1302.32 [256.60]
	Left	4-iron	-1193.76 [158.92]	11.25 [5.22]	-1290.74 [140.32]	6.73 [3.60]	-1275.50 [134.05]
		Pitch	-1191.50 [184.48]	12.04 [5.28]	-1263.89 [147.23]	7.05 [3.68]	-1250.67 [111.47]
ES	Right	4-iron	-1328.29 [210.93]	22.58 [9.89]	-1010.59 [178.66]	31.48 [16.31]	-1273.00 [157.97]
		Pitch	-1325.19 [148.08]	23.04 [8.26]	-1056.06 [138.11]	35.91 [12.36]	-1244.25 [138.35]
	Left	4-iron	-1345.35 [158.75]	20.32 [9.84]	-1090.17 [157.96]	31.92 [16.13]	-1247.04 [125.00]
		Pitch	-1357.10 [126.60]	23.10 [12.47]	-1060.05 [187.60]	31.85 [15.54]	-1260.00 [111.78]
Method/ onset type			**A2**		**B2**		**VI2**
RA	Right	4-iron	-442.24 [27.00]	8.11 [6.04]	-442.24 [27.00]	8.11 [6.04]	-442.53 [44.31]
		Pitch	-434.62 [35.46]	7.74 [6.53]	-433.97 [35.84]	7.77 [6.58]	-441.71[39.80]
	Left	4-iron	-426.30 [44.87]	18.84 [12.10]	-425.46 [44.22]	18.94 [12.22]	-433.51 [56.89]
		Pitch	-407.53 [55.23]	17.97 [11.47]	-406.87 [54.96]	17.97 [11.47]	-419.93 [68.82]
EO	Right	4-iron	-351.44 [55.82]	25.27 [16.06]	-352.28 [56.54]	25.23 [16.09]	-339.46 [71.45]
		Pitch	-373.55 [54.46]	27.19 [21.24]	-377.29 [56.25]	27.05 [21.31]	-367.74 [54.98]
	Left	4-iron	-364.58 [71.20]	40.74 [23.25]	-364.58 [71.20]	40.74 [23.25]	-341.26 [90.09]
		Pitch	-364.39 [64.72]	41.23 [26.39]	-365.33 [63.71]	41.16 [26.45]	-342.47 [73.26]
ES	Right	4-iron	-374.18 [72.26]	2.59 [2.89]	-237.71 [28.14]	19.78 [14.19]	-298.65 [42.29]
		Pitch	-372.25 [78.01]	5.45 [4.44]	-240.94 [51.02]	24.42 [12.79]	-303.13 [55.19]
	Left	4-iron	-328.00 [67.84]	4.82 [3.49]	-208.39 [31.39]	20.90 [14.90]	-302.04 [65.97]
		Pitch	-341.05 [70.01]	4.28 [2.34]	-205.10 [40.13]	19.95 [12.11]	-296.80 [58.80]

RA – rectus abdominis; EO – external oblique; ES – erector spinae; R – right; L – left.

Another factor that was considered refers to the tendency of preferably testing algorithms in maximum voluntary contractions or in monoarticulate actions (e.g. Jöllenbeck, 2000; Wong and Ng, 2005; Soylu and Arpinar-Avsar, 2010), instead of complex dynamic actions. Due to the phases that make up complex actions, the EMG is associated with qualitative changes from one phase to the following phases, which suggests different muscles assuming different roles throughout motor skills (Cordo et al. 2003; McGill et al., 2010). In a golf swing, which is a complex dynamic action, so far there have been few studies on temporal parameters (Horton et al., 2001; Cole and Grimshaw, 2008).

4.4.1. Onset detection methods and maximum peak percentage

When considering the task phases, backswing and downswing, they can be addressed as two temporal onsets: onset burst and onset peak (Fig. 2). The initial activity of the motor unit action potentials (Solnik et al., 2010) will be the onset burst, whereas the on/off muscle activity occurring before the peak in the execution phase will be considered as the onset peak in this study. The relevance of this phenomenon is to study motor skills and the performance accuracy in detection algorithm.

Discriminating the baseline during the relaxation between two consecutive MVCs and the swing trial background activity, a good agreement was found for the onset peak among the three methods for RA on both sides the left (R – ICC = 0.958; L – 0.957) and the right EO (ICC = 0.906), however in the last one the agreement with VI was lower than for the RA. For the RA, consistency given by the variation coefficient was also similar between methods (Table 2). By observing the detection of onset burst for the RA, results (R – ICC = 0.563; L – 0.0.479) which are different and conflicting than the onset peak. First, the onset burst does not occur in all subjects as can be observed in Fig. 1 for the right RA. Second, when the activity of the onset burst is considered, it is detected after the backswing. The phenomenon that was repeated in all subjects is the onset peak, but could be simultaneous with the onset burst in several cases because there was no activity besides the baseline, and the detection made for the restriction 1 was not validated by the VI. The pattern for the right RA starting its activation 438 ms before impact (8% Pmax) regardless of the threshold method used for determination (method A or B), and 442 ms with VI. The results are somewhat similar in the left RA, the method A and B showed 418 (8% Pmax) ms, and VI 427 ms.

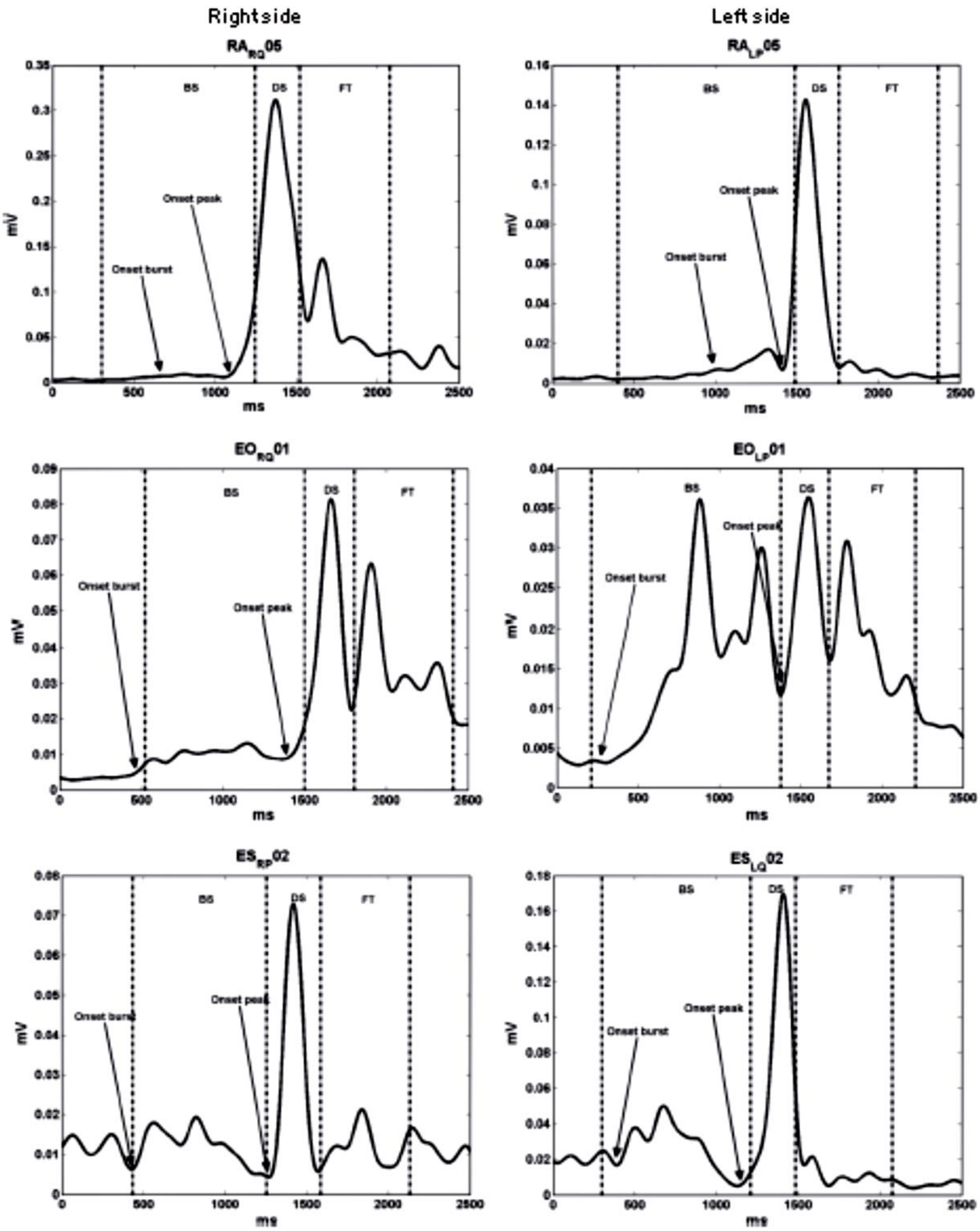


*Figura 5 [**Fig. 2**. Examples of onset burst and onset peak].*

RARQ05 – right rectus abdominis 4-iron trial 5; RALP05 – left rectus abdominis pitch trial 5; EORQ01 – right external oblique 4 - iron trial 1; EOLP01 – left external oblique pitch trial 1; ESRP02 – right erector spinae pitch trial 2; ESLQ02 – left erector spinae 4-iron trial 2.

EO onset peak, besides the good correspondence among methods (R – ICC = 0.906; L – 0.0.837), the inter correlation with VI varied from 0.517 to 0.681 and was detected at 26%Pmax for the right EO and 41%Pmax for the left EO, both at ± 364 ms before the downswing, much for method A and B. Onset burst detected in the EO had a high agreement (R – ICC = 0.906; L – 0.0.837) therefore being similar to the VI, as was the case for the onset peak found in the RA. Another similarity refers to the percentage of peak that occurs at very low values between 5%Pmax in the right EO with method B and 12%Pmax in the left EO with method A. The EO is the first muscle which is being activate, independently of the side, about 1260 ms before impact with method A and 1324 ms with method B.

For the ES it becomes difficult to distinguish where the burst onset is, because this muscle is already activated due to the forward leaning of the trunk during the address and the moments preceding the swing. This is coincident with the high background activity, then method B presents a high threshold. The onset burst was detected at 22%Pmax with method A and 33% with method B and the agreement with VI was low (R – ICC = 0.365; L – ICC = 0.214), but method A showed a better correlation. A higher signal-to-noise ratio leads to a better resemblance between the methods, and the reverse situation otherwise (Staude et al., 2001). Both automatic methods detected the onset peak in ES, which is a timing pattern in all the subjects characterized by intensely turning off/turn on the muscular activation. Nevertheless, the value agreement amongst the three methods was lower (R – ICC = 0.374; L – ICC = 0.277), method A being nearest to VI. A high EMG activity baseline is common in postural tasks, leading to an onset detection delay in respect to visual inspection (Hodges and Bui, 1996). The use of different methods for ES demonstrated that method B delays onset detection, in comparison to method A. Method A detected onset peak at a mean of 351 ms before impact (4.4%Pmax) and method B at 220 ms (21%Pmax), while the VI showed 347 ms. The concordance between threshold methods depended on the onset signal amplitude rate. A small rate of signal amplitude implies major temporal changes when different thresholds are used (Allison, 2003).

It is clear in this study that the cut percentage is different from muscle to muscle, regardless of the methods. This leads to adopting differentiation settings for the onset, which depends on the features of muscle signal under study. Hodges and Bui (1996) had already observed the same, which is related to the signal-to-noise ratio, and makes it difficult to precisely pinpoint onset detection (Solnik et al., 2008, 2010). The use of MVC rest activity no longer interferes with this ratio.

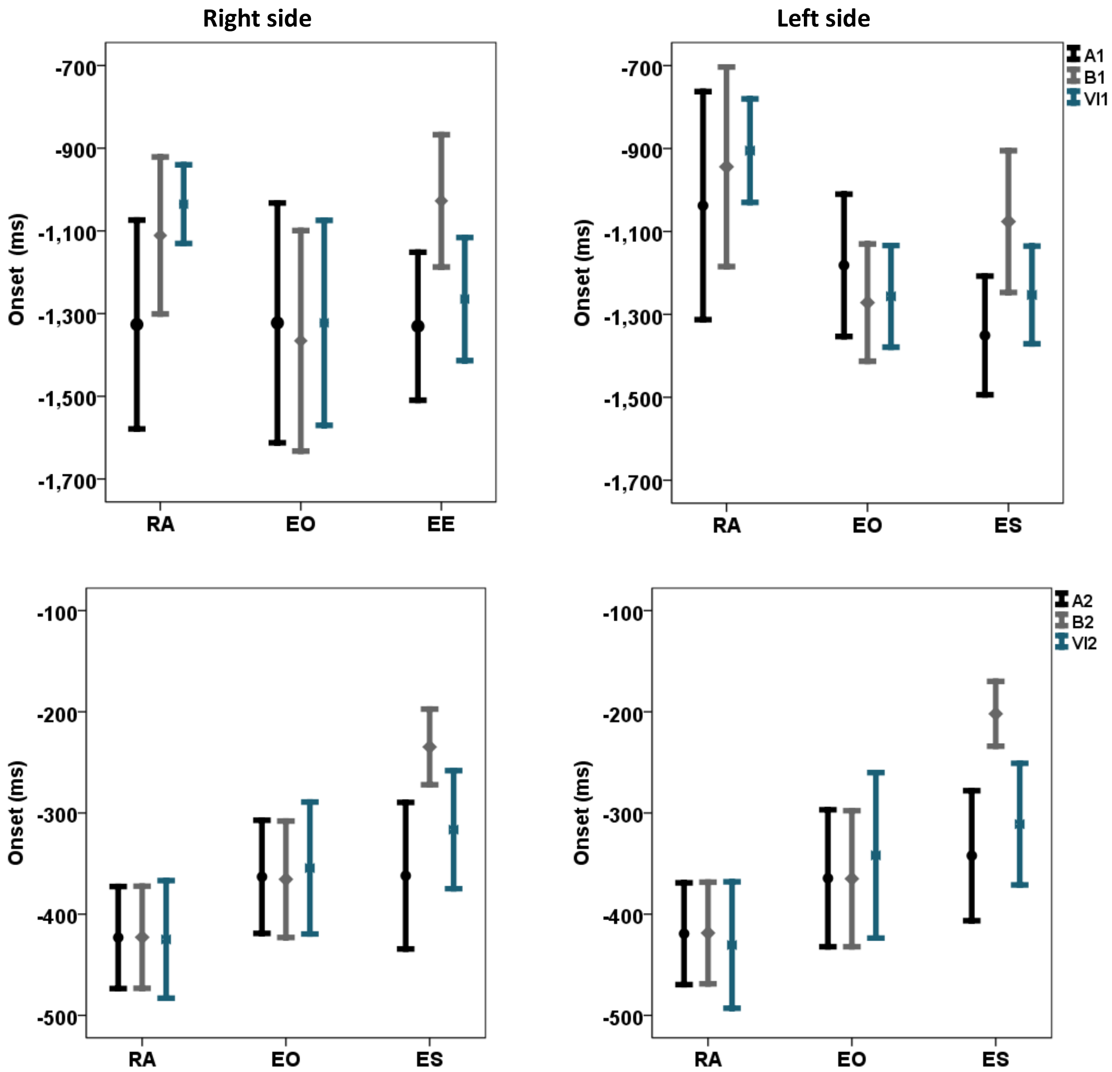


*Figura 6 [**Fig. 3**. Error bars methods A (A1, A2), method B (B1, B2) and visual inspection (IV1, IV2)].*
RA – rectus abdominis; EO – external oblique; ES – erector spinae.

Onset detection translated into maximum peak percentage is below the values presented by Hug (2011) for all the muscles that were studied, with the exception of the RA and ES, with method A. These results confirm that percentage threshold methods should use relatively low maximum peak amplitude percentage values, as stated by Jöllenbeck (2000). Nonetheless, this author studied monoarticular limb motor skills. In the present study, the results diverge between muscles due to different behaviors regarding mechanic demand.

4.4.2. EMG peak, mechanic demand and club

EMG represents the muscles' electric activity, but that activity can be associated with other functions which are not directly related to the movement itself (Vaisman et al., 2010), such as postural and/or adjustment. The study of different peaks can be useful as a good guidance in the motor behavior, but it is not advisable to compare absolute latency time with other methods (Wong and Ng, 2005).

For RA, we found the existence of a double peak in some participants, although there is consistency regarding maximum peak or the phenomenon being studied is very noticeable. In the the RA, maximum peak occurred at 170 ± 26 ms before impact on the right side and at 130 ± 81 ms on the left side. The ES was the muscle that showed the maximum activity nearest to impact, 126 ± 36 ms on the right side and 85 ± 31 ms before impact on the left side. For EO, maximum peak detection presents a higher variability. Firstly, due to the existence of three activity peaks. Another cause is the intra and inter subject regarding the moment of maximum activity, which could happen sometimes in backswing and others in downswing. According to Vaisman et al. (2010) it seems there is a tendency for a higher variability on trunk muscles when comparing them to limb muscles.

Multiple peaks can be associated with auxiliary actions to the agonist activity and with the action that is divergent in the different phases. General muscle activity pattern becomes complex for time patterns, with a qualitative change from one phase to the next, which suggests different function performance (Cordo et al., 2003). For EO, we found an almost simultaneous activation for onset burst in both sides, as if it were a co-contraction process. Theoretically, we would expect an earlier activation for these muscles, i.e. for EO on the left side, due to the fact that the trunk rotation movement was initiated by the right side. Hirashima et al. (2002) analyzed trunk, shoulder and upper limb muscle action for a ball throw. They confirmed that EO contralateral to the arm which performs the throw is activated first, regarding EO ipsilateral.

Tabela 11 [Table 2 - Statistical correlation and agreement between methods].

Muscle	Methods	Side	Onset (ms) |CV|	Onset (ms) ICC [IC95%]	Inter-Item Correlation Matrix		
			0.19			**A1**	**B1**
RA	A1 B1 VI1	R	0.17	0.563 [0.128, 0.768]	**B1**	0.630	
			0.09		**VI1**	0.597	0.471
			0.26				
		L	0.25	0.479 [0.239, 0.658]		0.247	
			0.14			0.293	0.385
			0.22				
EO	A1 B1 VI1	R	0.19	0.963 [0.945, 0.976]		0.909	
			0.19			0.908	0.913
			0.14				
		L	0.11	0.897 [0.759, 0.947]		0.826	
			0.10			0.837	0.884
			0.14				
ES	A1 B1 VI1	R	0.15	0.365 [-0.037, 0.645]		0.151	
			0.12			0.766	0.003
			0.11				
		L	0.16	0.214 [-0.089, 0.480]		0.025	
			0.09			0.399	0.133
			0.07			**A2**	**B2**
RA	A2 B2 VI2	R	0.07	0.958 [0.939, 0.972]	**B2**	0.999	
			0.09		**VI2**	0.842	0.841
			0.12				
		L	0.12	0.935 [0.903, 0.957]		0.995	
			0.15			0.795	0.794
			0.19				
EO	A2 B2 VI2	R	0.16	0.906 [0.863, 0.937]		0.981	
			0.18			0.681	0.683
			0.17				
		L	0.18	0.837 [0.757, 0.893]		0.999	
			0.24			0.517	0.518
			0.20				
ES	A2 B2 VI2	R	0.17	0.374 [-0.029, 0.629]		0.271	
			0.16			0.424	0.547
			0.20				
		L	0.17	0.277 [-0.060, 0.537]		0.121	
			0.21			0.484	0.367

RA – rectus abdominis; EO – external oblique; ES – erector spinae; R – right; L – left.

What happens in the transition from one phase to the other can be an important factor to be studied. Whereas the preparation phase will be subject to adjustment during the task, the movements in ballistic contractions are pre-programmed. When the motor skill that is being studied encompasses preparation and execution phases, the contraction relaxing contraction relationship could affect EMG behavior, due to the stretch-shortening cycle properties (McGill et al., 2010). For a golf swing, the transition from backswing to downswing is related to energy storage and transfer, in order to favor segment acceleration, followed by deceleration before impact (Cheetham et al., 2008). The double peak phenomenon is associated to different actions in those phases (McGill et al., 2010). The first peak can be associated with the action of stabilizing the trunk before movement starts, whereas the second is related to the dynamic action with its maximum peak occurring next to impact. The results indicate that a threshold algorithm can actually take the phenomenon which is being studied into consideration. Although the threshold is still a critical parameter, it is possible to find literature that suggests an adjustment to the algorithmic criteria of the data under study (Leader et al., 1998).

In the present study, the club does not interfere with different onset detections, although the effect size and power in ANOVA tends to be moderate. Egret et al. (2003) neither found significant differences for the duration of the golf swing performed with three different clubs (driver, 5-iron and pitch), nor for the partial time of golf swing phases. However, there were significant differences for club speed. This leads us to recommend that future studies approach temporal parameters together with kinematic analysis of motor chains.

4.5. Conclusion

The results of this study clearly show the baseline parameter influence in onset detection, during dynamic and complex motor activities. It is also verified that detection performance is dependent on the muscle which is being studied. RA presents a better correlation and concordance between methods, which means, baseline choice did not interfere with onset detection. For EO, onset detection it was similar with both methods, although detection occurred at different levels of maximum activity amplitude. Due to the postural task, ES had a high background activity, delaying onset detection for method B due a high threshold. Although ES was activated before the beginning of the movement, an instant activation which follows a pattern throughout trials was found. This pattern corresponds to the proximity of maximum activity to impact. In all the muscles, the right side showed a better agreement than the left side. Trunk muscles tend to increase their activity near impact. Different thresholds can correspond to similar onset detection, in case the specific muscle activity is translated into a high onset rate of amplitude.

The use of different clubs does not influence temporal parameters of trunk muscles during a golf swing.

Conflict of Interest

None declared.

Acknowledgments

The project ‘‘Neuromuscular activity in the golf swing with implications for the practice and in the prevention of overuse injuries’’ was supported by the Portuguese Foundation for Science and Technology fund (PTDC/DES/105176/2008).

References

Allison GT. Trunk muscle onset detection technique for EMG signals with ECG artefact. J Electromyogr Kinesiol 2003; 13(3):209–16.

Bechler JR, Jobe FW, Pink M, Perry J, Ruwe PA. Electromyographic analysis of the hip and knee during the golf swing. Clin J Sport Med 1995 Jul;5(3):162–6.

Cabri J, Sousa JP, Kots M, Barreiros J. Injuries in golf: a systematic review. EJSS 2009;9(6):353–66.

Cheetham PJ, Rose GA, Hinrichs RN, Neal RJ, Mottram RE, Hurrion PD, et al. Comparison of kinematic sequence parameters between amateur and professional golfers. In: Lutz R, Crews D, editors. Science and Golf V. Proceedings of the World scientific congress of Golf. Arizona: Ironwood Lithographers; 2008. p. 30–6.

Cole MH, Grimshaw PN. Trunk muscle onset and cessation in golfers with and without low back pain. J Biomech 2008; 41(13):2829–33.

Cordo PJ, Gurfinkel VS, Smith TC, Hodges PW, Verschueren SM, Brumagne S. The situp: complex kinematics and muscle activity in voluntary axial movement. J Electromyogr Kinesiol 2003; 13(3):239–52.

De Luca CJ. The use of surface electromyography in biomechanics. J Appl Biomech 1997; 13:135–63.

Egret CI, Vincent O, Weber J, Dujardin FH, Chollet D. Analysis of 3D kinematics concerning three different clubs in golf swing. Int J Sports Med. 2003; 24(6):465–70.

Farina D, Merletti R. Comparison of algorithms for estimation of EMG variables during voluntary isometric contractions. J Electromyogr Kinesiol 2000; 10(5):337–49.

Hirashima M, Kadota H, Sakurai S, Kudo K, Ohtsuki TJ. Sequential muscle activity and its functional role in the upper extremity and trunk during overarm throwing. Sports Sci 2002; 20(4):301–10.

Hodges PW, Bui BH. A comparison of computer-based methods for the determination of onset of muscle contraction using electromyography. Electroencephalogr Clin Neurophysiol 1996; 101(6):511–9.

Hodges PW. Changes in motor planning of feedforward postural responses of the trunk muscles in low back pain. Exp Brain Res 2001; 141(2):261–6.

Horton JF, Lindsay DM, Macintosh BR. Abdominal muscle activation of elite male golfers with chronic low back pain. Med Sci Sports Exerc 2001; 33(10):1647–54.

Hug F. Can muscle coordination be precisely studied by surface electromyography? J Electromyogr Kinesiol 2011; 21(1):1–12.

Hume PA, Keogh J, Reid D. The role of biomechanics in maximising distance and accuracy of golf shots. Sports Med 2005; 35(5):429–49.

Jöllenbeck J. Methodological limitations of EMG-based bio-mechanical motion analysis. ISBS – conference proceedings archive, 18 International symposium on biomechanics in sports; 2000.

Konrad P. The ABC of EMG: a practical introduction to kinesiological electromyography. USA: Noraxon Inc.; 2005.

Leader 3rd JK, Boston JR, Moore CA. A data dependent computer algorithm for the detection of muscle activity onset and offset from EMG recordings. Electroencephalogr Clin Neurophysiol 1998; 109(2):119–23.

Lee J, Jung MY, Kim SH. Reliability of spike and turn variables of surface EMG during isometric voluntary contractions of the biceps brachii muscle. J Electromyogr Kinesiol 2011; 21(1):119–27.

McGill SM. Electromyographic activity of the abdominal and low back musculature during the generation of isometric and dynamic axial trunk torque: implications for lumbar mechanics. J Orthop Res. 1991; 9(1):91–103.

McGill SM, Chaimberg JD, Frost DM, Fenwick CM. Evidence of a double peak in muscle activation to enhance strike speed and force: an example with elite mixed martial arts fighters. J Strength Cond Res 2010; 24(2):348–57.

Micera S, Sabatini AM, Dario P. An algorithm for detecting the onset of muscle contraction by EMG signal processing. Med Eng Phys 1998; 20(3):211–5.

Morey-Klapsing G, Arampatzis A, Brüggemann GP. Choosing EMG parameters: comparison of different onset determination algorithms and EMG integrals in a joint stability study. Clin Biomech (Bristol, Avon) 2004; 19(2):196–201.

Pink M, Perry J, Jobe F. EMG analysis of the trunk in golfers. Am J Sports Med 1993; 21(3):385–8.

Solnik S, DeVita P, Rider P, Long B, Hortobágyi T. Teager–Kaiser Operator improves the accuracy of EMG onset detection independent of signal-to-noise ratio. Acta Bioeng Biomech 2008; 10(2):65–8.

Solnik S, Rider P, Steinweg K, DeVita P, Hortobágyi T. Teager–Kaiser energy operator signal conditioning improves EMG onset detection. Eur J Appl Physiol 2010; 110(3):489–98.

Soylu AR, Arpinar-Avsar P. Detection of surface electromyography recording time interval without muscle fatigue effect for biceps brachii muscle during maximum voluntary contraction. J Electromyogr Kinesiol 2010; 20(4): 773–6.

Staude G, Flachenecker C, Wolf W. Onset detection in surface electromyographic signals: a systematic comparison of methods. EURASIP J Appl Signal Process 2001; 2001(2):67–81.

Vaisman L, Zariffa J, Popovic MR. Application of singular spectrum-based change point changepoint analysis to EMG-onset detection. J Electromyogr Kinesiol 2010; 20(4):750–60.

Van Boxtel GJ, Geraats LH, Van den Berg-Lenssen MM, Brunia CH. Detection of EMG onset in ERP research. Psychophysiology 1993; 30(4):405–12.

Watkins R, Uppal G, Perry J, Pink M, Dinsay J. Dynamic electromyographic analysis of trunk musculature in professional golfers. Am J Sports Med 1996; 24(4):535–8.

Wong YM, Ng GY. The double peak-to-peak analysis for determining EMG onset of muscle contraction. Electromyogr Clin Neurophysiol 2005; 45(5):267–71.

CAPÍTULO V

5. EMG-based Classification with Support Vector Machine: A Review

Submetido no ***Journal of Electromyography and Kinesiology***

Silva, L., Castro, M. A., Serranho, P., Cabri, J., Pezarat-Correia, P. EMG-based Classification with Support Vector Machine: A Review. *Journal of Electromyography and kinesiology*

EMG-based Classification with Support Vector Machine: A Review

Silva, Luís; Faculty of Human Kinetics, Technical University of Lisbon

Castro, Maria António; Coimbra College of Health Technology, Polytechnic Institute of Coimbra

Serranho, Pedro; Universidade Aberta

Cabri, Jan; Norwegian School of Sport Sciences

Pezarat-Correia, Pedro; Faculty of Human Kinetics, Technical University of Lisbon

Abstract

The support vector machines (SVM) are a mathematical technique of supervised learning with an increasing use in pattern recognition tasks. The electromyography (EMG) is a very useful tool in the quantification and description of neuromuscular activity. A literature review using the keywords "electromyography" and "support vector machines" was held. From 41 articles, three main categories, myoelectric control, neuromuscular disorders and kinesiological analysis were found. Upper limb, namely the forearm muscles, are preferentially studied muscle body region. This tendency to collect data forearm is dominant in studies of the myoelectric control. C-SVM with radial basis function (RBF) kernel is the most chosen, and when used multiclass models it appears the use of one-against-one and one-against-all. In myoelectric control, time-domain features are mainly used, varying the used channels between 2 to 10 or from a surface block. Generally, were found the use of time-domain and time-scale features than frequency-domain features. Segmentation windows vary between 32 and 256 ms. There is a lack of studies that considering both transient and steady-state states of EMG information. SVM showed high accuracy in classification tasks, and when compared with others technique had greater performance.

Keywords: Electromyography, SVM, Pattern Recognition, Myoelectric Method.

5.1. Introduction

Surface electromyography (EMG) is an experimental technique that records the sum of motor unit action potentials (MUAP) generated by the active muscle fibers that are detected on the skin and their physiological variations (Merletti & Farina, 2008). The motor unit action potentials (MUAPs) are the electrical difference between the ground electrode and the electric activity captured on the skin with a second electrode placed above the muscle belly. In spite of the non-intrusive advantages of EMG, the signal is affected by several factors that influence its quality. De Luca (1997) categorized these factors into the following classes: causative, intermediate and deterministic. The causative factors are those directly related with the basic properties of the electrodes (extrinsic factors) and the physiologic, anatomical, and biomechanical characteristics of the region where they are placed (intrinsic factors). The intermediate factors represent the physiological and physical phenomena influenced by causative factors, such as filtering and crosstalk. Deterministic factors are related with signal information and record power, such as the motor unit firing rate, amplitude, and duration of the MUAPs. All these factors compromise the fidelity of EMG records.

There are two general approaches to study the relations between EMG and the properties of the neuromuscular system: forward and inverse (Farina et al., 2004). The forward approach allows predicting the effect of several physiological processes on EMG features. The inverse approach uses the EMG to identify the physiological parameters. Whatever the approach considered, the influence of these factors is the major drawback of the EMG pattern recognition, since they can lead to a poor recognition and a decrease in discriminatory accuracy (Phinyomark, Nuidod, Phukpattaranont, & Limsakul, 2012).

The algorithms applied on EMG have been essential in myoelectric control and clinical diagnostic by means of discrimination of a given pathology between subjects. Several statistical methods most commonly found in the literature are: evidence accumulation procedure with fuzzy mapping function to transform distances for the application of the Dempster–Shafer theory of evidence (Park & Lee, 1998), principal component analysis (Beck, Stock, & Defreitas, 2011), artificial neuronal networks (ANN) (Cao, Lin, & Huang, 2010), linear discriminant analysis (LDA) (Subasi, 2012), multilayer perceptron (Oskoei & Hu, 2007), and support vector machines (SVM). The latter has been reported as the one with better performance (Khushaba, Kodagoda, Takruri, & Dissanayake, 2012; Oskoei & Hu, 2008).

SVM started in the field of statistics learning theory with Vapnik and Chervonenkis in the late 60s and had a major development from the early 90s (Boser et al., 1992; Vapnik, 1999), namely, with the introduction of the soft margin classifier (Cortes & Vapnik, 1995). SVM emerged from statistical learning theory based on the principle of structure risk minimization, and it is a useful method for data classification and regression. Given data from a so-called training set, the aim of SVM is to produce a classification model based on training data which predicts the target values of a different test set data with unknown classification (Hsu et al., 2010). The researcher is the one that tells the machine how the classification works, and this process is called supervised learning (Evgeniou, Pontil, & Poggio, 2000).

The main purpose of statistics learning theory is to develop mechanisms (algorithms) according to which certain tasks can be learned within the inductive inference and generalization capability (Luxburg & Schölkopf, 2009). Looking at the binary classification case, where there are only two classes, or in other words, two possible labels in the label space $\mathcal{Y}$ for each element of the input space $\mathcal{X}$. This is the general notion of classification problems using an input-label training independent identically distributed (*i.i.d.*) data pairs according to an unknown probability distribution $P(\mathbf{x}, y)$ to estimate the function $f: \mathbb{R}^N \rightarrow \{-1,1\}^n$, where $(\mathbf{x}_i, y_i) \in \mathbb{R}^N \times \mathcal{Y}$, in order to classify with the higher possible of accuracy the given training examples of $(\mathbf{x}, y)$. If a cost-function $c(f(\mathbf{x}), y)$ is assumed, the best function f is the one that minimizes the *functional risk* or *expect error* of that cost function, given by the following (Burges, 1998; Müller et al., 2001; Schölkopf & Smola, 2002; Vapnik, 1999):

$$R(f) = \int c(f(\mathbf{x}), y) dP(\mathbf{x}, y) \tag{1}$$

As it is not possible to directly minimize the expected risk, since the probability distribution $P(\mathbf{x}, y)$ is unknown, the choice between classifiers functions corresponds to minimizing the training error. This can be defined by the average error rate measured in the training set with fixed and finite number of observations, called *empirical risk*.

$$R_{emp}(f) = \frac{1}{n} \sum_{i=1}^{l} \frac{1}{2} |f(x_i) - y_i| \tag{2}$$

To control the complexity of a function class, the *Vapnik–Chervonenkis* (VC) theory and the *structural risk minimization principle* are considered. The VC dimension gives consistency to understand the capacity of a given set of functions. Assuming that the training set can be separable by a hyperplane, the chosen separation functions follow the form:

$$f(x) = \mathbf{w} \cdot \mathbf{x} + b \tag{3}$$

where $\mathbf{w} \cdot \mathbf{x}$ is the dot product between the two vectors **w** and **x,** the first one corresponding the normal vector to the hyperplane and $b/\|\mathbf{w}\|$ is the perpendicular distance from the hyperplane to the origin. The above equation divides the input space into two regions (Cristianini & Shawe-Taylor, 2000):

$$g(\mathbf{x}) \;= \begin{cases} +1 \; if \; \mathbf{w} \cdot \mathbf{x} + b \; > 0 \\ -1 \; if \; \mathbf{w} \cdot \mathbf{x} + b < 0 \end{cases} \tag{4}$$

Therefore, the separating hyperplane classifier is defined by

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1 \geq 0 \quad \forall_i = 1, \dots, n \tag{5}$$

The optimization problem to determine the hyperplane with the maximum margin is given by the smaller norm of the weight vector **w**

$$\min_{\mathbf{w},b} \frac{1}{2} \|w\|^2 \tag{6}$$

subject to the set of constraints (5). Introducing the Lagrange multipliers $\alpha_i \geq 0$, the Lagrangian is given by

$$L(\mathbf{w}, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_i^n \alpha_i (y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1) \tag{7}$$

The Lagrangian has to be minimized in relation to the variables $\mathbf{w}, b$ with the constrains $\alpha_i \geq 0$ (Burges, 1998; Schölkopf & Smola, 2002), resulting in a “saddle point”:

$$\frac{\partial}{\partial b} L(\mathbf{w}, b, \alpha) = 0 \qquad e \qquad \frac{\partial}{\partial \mathbf{w}} L(\mathbf{w}, b, \alpha) = 0 \tag{8}$$

Under the Kuhn–Tucker conditions, the Lagrange multipliers are non-zero only when the constraints are saturated (Lucas, Gaufriau, Pascual, Doncarli, & Farina, 2008). The explanation above refers to the traditional case of separable data. When applied to non-separable data, the solution is not feasible, and it is necessary to relax the constraints. With the introduction of non-negative cost parameters ξ_i (Burges, 1998; Hofmann et al., 2008), the problem becomes:

$$\mathcal{L}(\mathbf{w}, b, \alpha) = \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{n}\xi_i - \sum_{i}^{n}\alpha_i(y_i(\mathbf{w}\cdot\mathbf{x}_i + b) - 1 + \xi_i) - \sum_{i=1}^{n}\eta_i\xi_i \tag{9}$$

where $\alpha_i, \eta_i \geq 0\ \forall_i \in n$, and α_i, η_i are Lagrange multipliers introduced to enforce the positivity of ξ_i. The dual function is then subject to

$$\frac{\partial}{\partial \mathbf{w}}\mathcal{L} = \mathbf{w} - \sum_{i=1}^{n}\alpha_i y_i\,\mathbf{x}_i = 0 \tag{10}$$

$$\frac{\partial}{\partial b}\mathcal{L} = -\sum_{i=1}^{n}\alpha_i y_i = 0 \tag{11}$$

$$\frac{\partial}{\partial \xi_i}\mathcal{L} = C - \alpha_i + \eta_i = 0 \tag{12}$$

$$(y_i(\mathbf{w}\cdot\mathbf{x}_i + b) - 1 + \xi_i) = 0 \tag{13}$$

$$\alpha_i(y_i(\mathbf{w}\cdot\mathbf{x}_i + b) - 1 + \xi_i) = 0 \tag{14}$$

$$\eta_{i\xi_i} = 0 \tag{15}$$

$$\alpha_i \in [0, C], \forall_i \in n$$

The linear case represented before, where a hyperplane is considered for data space separation, only tolerates a limited amount of noise and number of outliers, so these cases require a non-linear approach. That is why, SVM can be divided in the linear separable case, linear non-separable case and nonlinear SVM. The nonlinear SVM is necessary when a separation hyperplane is not a good choice for separation surface. In this approach the data mapping is transferred to an infinite dimensional Euclidean space "$\mathcal{H}$" where the separation hiperplane is performed $\Phi : \mathbb{R}^N \rightarrow \mathcal{H}$ and $\mathbf{x} \rightarrow \Phi(\mathbf{x})$, replacing the scalar products. However, this theoretical constructions does not demand Φ to be known due to the "*kernel trick*" through the Kernel functions so $\left(\Phi(\mathbf{x}) \cdot \Phi(\mathbf{y})\right) =: k(\mathrm{x}{\cdot}y)$ (Müller et al., 2001; Scholkopf et al., 1996; Schölkopf & Smola, 2002).

Other approaches have been developed. The case of "*new SVM*", called v-SVM, proposed by Schölkopf, Smola, Williamson, and Bartlett (2000), is a class of algorithms both for classification and regression. The parameter v can control the number of support vectors and eliminate the parameter C in the case of classification and ε parameter for the error in regression.

Apart from noise affecting the signal quality, robustness in pattern recognition can be influenced by features extraction. In the context of SVM, a feature is a numerical dimension of the data space, and they should be chosen in order to provide data as separable as possible in the constructed data space. Features should be extracted from various segments of the signal to preserve its structure and not from individual samples, which leads to a significant loss of information (Hudgins, Parker, & Scott, 1993). Following the procedures of amplification, filtering and digitization of the signal, Oskoei and Hu (2007) report that pattern recognition in EMG is carried out through four steps: data segmentation, feature extraction, classification and the controller action.

The first concern with signal segmentation is its length. A small segment length is associated with deviations and variations in the estimation of features due to loss of signal information. A long length requires a high computational cost, which may impair the performance of real-time rating (Oskoei & Hu, 2008). Two fundamental methods can be considered for segmentation: *disjoint* and *overlapped windows* (Englehart & Hudgins, 2003; Oskoei & Hu, 2007; Oskoei & Hu, 2008). The first one should be applied to segments of 200 ms or less, whereas the *overlapped window* length exceeds 200 ms (Oskoei & Hu, 2008). Due to the stochastic and non-stationary nature of EMG signal, the decision on the segment length must take into account the existence of transition and stationary states. The segmentation techniques described are usually related with steady states, since they are associated with better accuracy.

There are three categories of features extraction methods: *time domain*, *frequency domain* and *time-scale domain*. Descriptions can be seen in several EMG and pattern recognition literature (Chen, Zhu, & Zhang, 2010; Englehart et al., 2003; Oskoei & Hu, 2007; Oskoei & Hu, 2008; Zecca, Micera, Carrozza, & Dario, 2002). The method for feature extraction is critical since factors that influenced EMG quality and the type of feature extraction can influence the classification accuracy. Tkach et al. (2010) studied the stability of time-domain EMG features during changes in signals according to three common disturbances of EMG signal source: *changing electrode location*, *variability of muscle contraction effort*, and *muscle fatigue*. They found that electrode location shift and varying effort level significantly decrease the classification accuracy. The autoregressive coefficient and cepstrum coefficient were the most stable algorithms for individual classification. However, sets of different feature types are most robust than individual ones. The variance, v-Order, logarithmic detector, and EMG histogram were the best for shifting electrodes. For the variability of muscle contraction effort, the feature set wave length, slope sign change, logarithmic detector and autoregressive coefficient showed the best performance.

The aim of this study is to analyze and compare several published works where SVM was applied for EMG pattern recognition, taking into account SVM techniques, SVM parameters, study type, muscles and tasks analyzed as well as the achieved accuracy.

5.2. Method

For this literature review a search, select and analysis protocol was performed following the procedures shown in figure 1.

5.2.1. Selection criteria

The first eligibility criterion required for study inclusion in this review was the classification of EMG patterns with SVM, regardless of the addressed problem. Thus, the keywords for the performed search were "*electromyography*" and *"support vector machine*". All types of studies in which the classification method is SVM and features are derived from the EMG signal were accepted. The expected outcomes were the SVM method (i.e. C-SVM, nu-SVC, one-class SVM, epsilon-SVM) and additional procedures (multiclass one-against-one, one-against-all, directed acyclic graph SVM, all-together, k-fold cross-validation), Kernel formulation (i.e. linear, polynomial, radial basis function, sigmoid), classification accuracy cost parameter C in non-separable case, and Kernel values.

The final inclusion criteria were (1) containing EMG data features used for SVM classification or regression; (2) papers in English; (3) EMG studies performed in human subjects with and without pathology. We then used the following exclusion criteria: (1) despite using EMG, the SVM method is only used on other variables; (2) description of procedures regarding EMG and SVM use was not clear; (3) articles not based on a specifically aimed methodology.

5.2.2. Database searching

A systematic search using the combined keywords ‘‘*electromyography”* and *“support vector machine*’’, was performed on the electronic databases PubMed, SportDiscus, ISI Web of Knowledge and Science Direct (Figure 1). Then, the procedure was as follows: (1) checking the indexes, references and similar documents; (2) eliminating the duplicates; (3) rejecting non relevant articles for the review aim and other review articles; (4) cross-checking with GoogleScholar and IEEE procedures of referenced articles relevant for the purpose of the present study.

5.3. Results

The search results gave rise to a total of 244 articles; after eliminating the indexes, references, and duplicates 73 articles were eliminated and therefore 171 were obtained for manual screening. A careful reading was performed to determine if the articles were in accordance with inclusion and exclusion criteria. A total of 42 articles were withheld for results. Figure 2 shows publications considered in this review over the years.

The results and discussion were organized in three main categories: myoelectric control, neuromuscular disorders, and kinesiological analysis.

Tables 1 to 4 show the goals, methods, and main results of the selected studies.

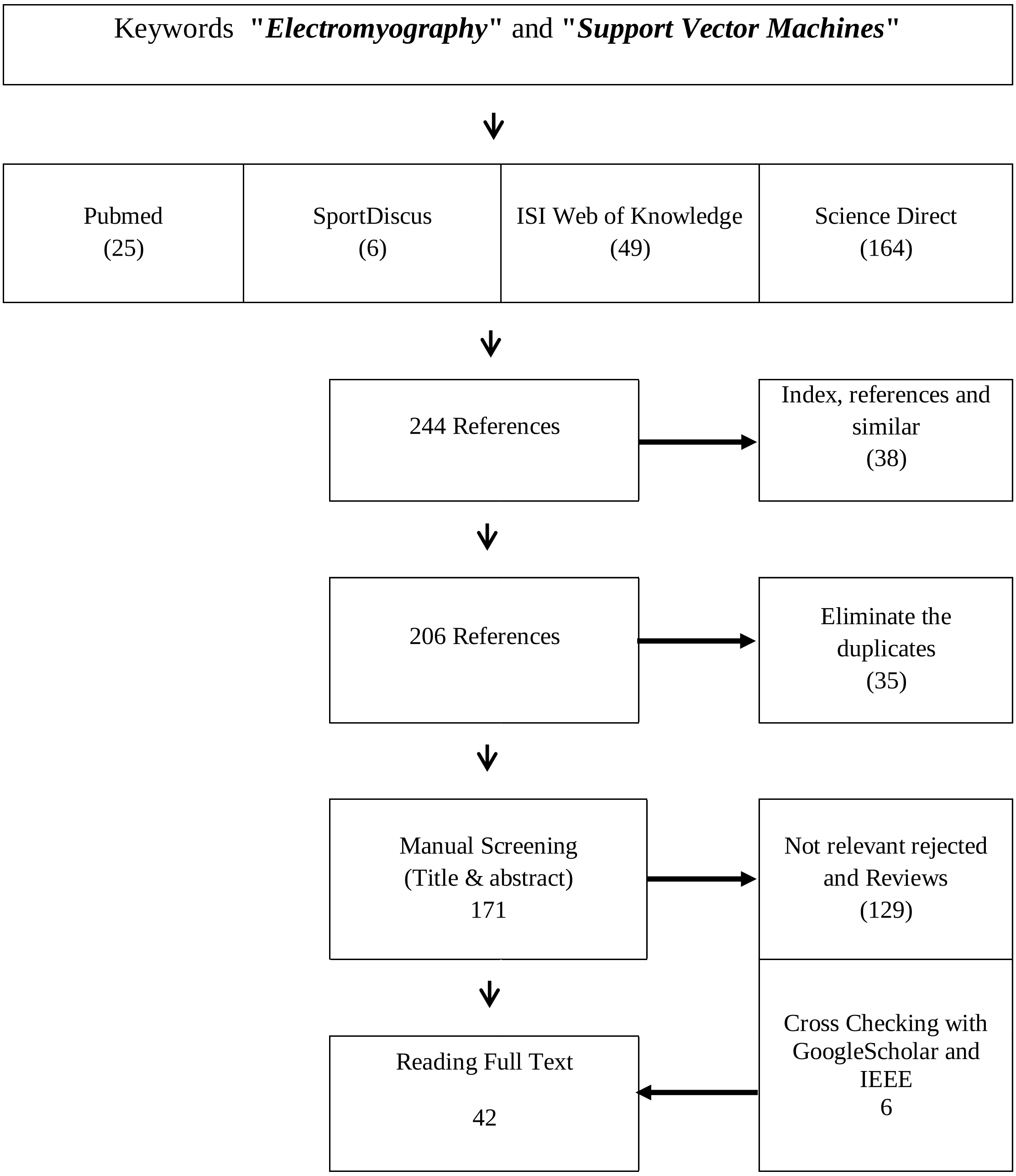


*Figura 7[**Fig. 1**. Flow chart].*

5.4. Discussion

5.4.1. Studies design and objective

The aim of this review study was to analyze how SVMs were applied on EMG pattern recognition on literature, including the considered parameters and achieved accuracy. The major applications of supervised learning in the study of EMG patterns were myoelectric control, neuromuscular disorders discrimination, and kinesiological analysis. The objective with pattern recognition associated to gestures classification is similar, but the reason behind pattern recognition was taken into account.

Early studies were about neuromuscular disorders discrimination (Güler & Koçer, 2005; Xie, Wang, Huang, & Qing, 2003), comparing the performance of SVM with back-propagation ANN, and about the classification of multi-channel surface EMG signals generated by the variation of the number of channels (Crawford, Miller, Shenoy, & Rao, 2005). To the best of our knowledge, in 2006 the number of articles was maintained, and there is a lack of EMG-based recognition with SVM in 2007. After that, there is an increase of research using this method.

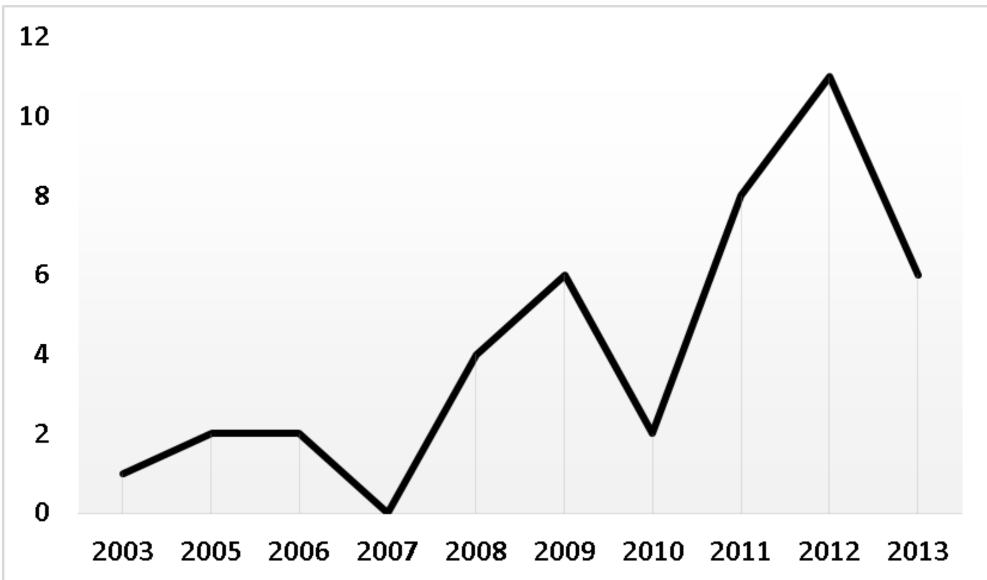


*Figura 8 [**Fig. 2**. Evolution of pattern recognition publications].*

5.4.1.1. Myoelectric control

Studies of myoelectric control aim to recognize neuromuscular patterns through EMG leading to the development of active prostheses that perform tasks with some degree of difficulty with the highest possible accuracy. As shown in this review, studies tend to focus on forearm and hand movements (Table 1). Forearm electrode placement aims at the recognition of hand gestures, and can be classified into three major categories: gross hand, wrist and arm movement recognition; individual finger activation and movement detection; and multiple finger gesture classification (Chen & Wang, 2013). The only study found that focused on the lower limb (Huang, Zhang, Hargrove, Dou, Rogers, & Englehart, 2011) shows better accuracy results of neuromuscular–mechanical fusion than methods that use single data source for both stance phase and swing phase of gait locomotion.

The major tasks regarding the upper limb are related to grasp recognition. Castellini and Smagt (2009) stated three essential concerns when operating hand movement down to the level of single fingers: the patient's ability to enforce the correct grasping type, the control of the amount of force involved in the grasping task, and the patient feedback. EMG used for feed-forward control, as well as accuracy, requires real time adaptability and speed, hence the importance of machine learning. That information depends on the features discrimination power. Surface EMG features for recognition depend on both membrane fiber properties and motor unit control process. Their extraction requires advanced methods for signal analysis, in order to enable decoding of signal information (Merletti & Farina, 2008). The success of classification system is the result of a combination of several factors (Crawford, Miller, Shenoy, & Rao, 2005): (1) careful selection of actions to classify due to the representative intuitiveness for the control task, (2) selection of muscle activity recording sites that are relevant to the actions, (3) simplicity in feature representation and real-time performance, and (4) state-of-the-art classification method-based SVM. The combination of these factors is critical for the classification of any data set. Increasing information about the type of data could be a fundamental task, but in the case of myoelectric control it must rely on the computational cost. The patient must be able to control the prosthesis accurately in real-time.

In addition to the control of prostheses using the recognition of EMG signals, the refinement of mechanical movements by robotic structures emerges as another goal, namely, the control of a robotic hand (Yang et al., 2009), and the control of a wrist exoskeleton in real-time (Khokhar, Xiao, & Menon, 2010). In both cases, as expected, the training accuracy showed higher values when compared with test sets.

Another factor that can be highlighted in studies of myoelectric control is the existence of a clear preference for surface EMG rather than deep EMG. This can be explained by two main reasons. The first is due to surface EMG being a non-invasive method. The second is related to the extracted MUAPs information. Since the surface electrodes capture the electrical activity produced by the muscle beneath the skin, the recording EMG is the sum of active muscle fibers. Instead, the deep EMG collects the activity from fiber sets surrounding the detection needle/wire, which reduces the amount of information. Apart from the important relevance of deep EMG in obtaining the activity of deeper muscles, when comparing surface with intramuscular myoelectric signals, surface electrodes are suitable to represent the amplitude of deep muscles of the torso within a 15-20% RMS difference (McGill, Juker, & Kropf, 1996). Thus, surface EMG detection shows a larger association with recognition of motor skills, which is evident in the literature produced about myoelectric control.

5.4.1.2. Neuromuscular diagnostic

The papers found in this category essentially aim at the diagnostic of neuromuscular disorders using SVM approach on EMG, dividing participants in three core groups: healthy, neuropathy, and myopathy. Only two articles do not fit this division, focusing on juvenile myoclonic epilepsy (Goker et al., 2012) and motor epileptic seizures recognition (Conradsen et al., 2012).

Neuromuscular disorders are associated with faults in the motor units elements, which makes it important to identify causes of disease as myogenic or neurogenic. The waveforms of each MUAP characterize the features that should be sufficiently discriminative to allow determination of pathology type, differing between myogenic and neurogenic disorders. The myogenic record showed a short-lasting, low-amplitude and multiphasic potential due to atrophy, which corresponds to less membrane potentials. The MUAPs of neurogenic can be characterized as long-lasting, high-amplitude and also multiphasic, since the pathology leads to the loss of axons and motoneurons

(Dobrowolski, Wierzbowski, & Tomczykiewicz, 2012). In the earliest study considered in this review for this category, both needle and surface electrodes were used, and the diagnostic was supported with biopsy if necessary (Güler & Koçer, 2005). These authors found that neuropathic patients had EMG signals with high power spectrum density, and myopathic patients showed lower power spectrum density. The normal group power spectrum density was lying between those values. The previous analysis of the groups' characteristics sought to find the best information of EMG content to feed the classifier. Katsis, Goletsis, Likas, Fotiadis, and Sarmas (2006) organized their study in three steps: preprocessing of EMG signal to find areas of low activity (candidate MUAPs); MUAPs clustering, detection and decomposition of superimposed MUAPs; and finally, classification. From 50 EMG needle signals, 2969 candidate MUAPs were obtained with a sampling threshold ranging between 30 to 100 μV. Unlike studies on myoelectric control where surface EMG assumes a significant role, in neuromuscular disorders both surface and deep EMG are used. The preprocessing can influence the classification results due to changes in the temporal structure, amplitude and frequency of the signal. To discriminate population characteristics, the quantification of EMG signals is not focused on just one characteristic, which explains why studies of neuromuscular disorders do not meet neither time domain nor frequency features. The majority of studies mention the quantification that fed the SVM classifier was through time-frequency features, i.e. wavelet transform (Conradsen et al., 2012; Dobrowolski et al., 2012; Subasi, 2012, 2013).

The typical source of classification errors associated with neuromuscular disorders discrimination is due to the borderline cases, such as the initial disease stages or muscle disorder specificity involving just a part of a muscle (Dobrowolski et al., 2012). To avoid subjectivity related to these errors, Dobrowolski et al. (2012) used MUAPs wavelet decomposition to determine a set of new features which had better discrimination information. After SVM analysis, there is a single decision parameter called the Wavelet Index, diagnosing the subject to one of the three groups (healthy, myogenic or neurogenic). The best advantage of associating wavelet analysis is the capacity to concentrate signal energy in few coefficients. Subasi (2012) compared different types of machine learning methods to classify neuromuscular disorders. Results showed that the fuzzy-SVM associated with wavelet transforms had better recognition rate, insensitivity to overtraining, and higher reliability than LDA, ANN and radial basis network. The major difference between SVM and fuzzy-SVM is related to the cost C, since this parameter is due to the product between the cost C of fuzzy-SVM and the fuzzy membership $0<s_i\leq1$ related with each training sample x_i.

Others approaches that have been considered are MUAP clustering (Güler & Koçer, 2005), multi-scale amplitude modulation–frequency (Katsis et al., 2006), scanning EMG (Goker et al., 2012), and recurrence quantification analysis (Sultornsanee, Zeid, & Kamarthi, 2011). The study showing best accuracy was from Sultornsanee, Zeid, and Kamarthi (2011) which found the highest classification for the myopathy group with a performance of 100%, and a mean accuracy for all groups of 98.28% in the tibialis anterior muscle.

In addition to accuracy in SVM studies of neuromuscular disorders diagnostic, specificity and sensitivity metrics are widely used. The specificity is the ratio between the number of healthy subjects correctly classified and the number of total healthy subjects. The sensitivity corresponds to the ratio between the number of subjects correctly classified with a given disorder and the total number of subjects with that disorder (Goker et al., 2012; Subasi, 2012, 2013). In automatic multi-modal seizure detection, sensitivity corresponds to the fraction of seizures that are correctly classified; other variables, such as latency and false detection rate, can be used (Conradsen et al., 2012). Latency is the elapsed time between the start and detection of a seizure, and the false detection rate quantifies the simulated seizures that were falsely detected per hour.

5.4.1.3. Kinesiologic analysis

The last category included in this review is kinesiologic analysis: the EMG-based analysis of movement. This category addresses the recognition of changes in EMG signal that may be associated with characteristic aspects of some specific movement. If the authors stated that use SVM pattern recognition is for prosthesis control (even showing any prosthesis), then that study was considered myoelectric control. On the other hand, if the authors generalized the aim of pattern recognition, the paper was included in kinesiologic analysis. Although most studies in this category focus on forearm muscles and wrist movements (Alkan & Günay, 2012; Futamata, Nagata, & Magatani, 2012; Tavakolan, Xiao, & Menon, 2011), it is possible to come across a broader muscle analysis than the one in myoelectric control. Two studies analyzed lower limb (Stirling, von

Tscharner, Kugler, & Nigg, 2011; Tolambiya et al., 2011), one included trunk muscles (Tolambiya et al., 2011), and other study focused on the participation of biceps on elbow flexion (Natarajan, Wininger, Kim, & Craelius, 2012).

Stirling, Von Tscharner, Kugler and Nigg (2011) analyzed whether the EMG signal undergoes changes during running that represent muscle properties and movement strategies, and also, if these changes are related to effort level. They found systematic changes in EMG features during a prolonged run, which are related to the runner effort level. In motor control studies there are a large number of variables and relations between them. SVM is a method that could find differences underlying several task variations (Tolambiya et al., 2011). These authors compare reaching movements with and without constraints with the aim of analyze the EMG classification in those movements, allowing identification of some of the features that are adapted when constraints are placed on a complex motor task. Considered constraints were postural (knee extended and reduced base of support) and focal (imposed straight finger trajectory and imposed semicircular finger trajectory). They found better accuracy classification from leg rather than trunk muscles in an ambitious setup of a whole body pointing task. These differences between trunk and lower limb are also noticed in the study of temporal parameters due to the higher variability of trunk muscles (Vaisman et al., 2010).

Using a binary SVM, Tolambiya et al. (2011) found that the postural knee extension and the focal finger trajectory showed higher classification accuracy than others tasks, with 97% and 91%, respectively. Tavakolan et al. (2011) sought to know whether age can interfere with performance of a radial basis function SVM classifier, combining eight hand movements. Movement classification of younger subjects showed higher accuracy when compared to older subjects, despite kernel function parameters being similar for both groups in pronation and supination classification. The parameter cost was 10 for 83% of subjects and the mean kernel gamma was 0.8. Alkan and Günay (2011) also studied pronation and supination of the forearm but comparing SVM with discrimination analysis method. Despite the high accuracy from both methods, SVM reached 99% of performance. However, Yan, Wang, and Xie (2008) found that pronation motion is more difficult to discriminate than supination using fuzzy least squares-SVM.

Tabela 12 [Table 1 - SVM on EMG myoelectric control forearm muscles].

Authors	Objective	Method				Main findings
		Participants	*Channels*	*Task*	*Classification, segmentation and features*	
Crawford, Miller, Shenoy, & Rao, 2005	To analyze myoelectric control in robotic prosthetic hands and limbs with multiple degrees of freedom.	*3 healthy subjects x 5 sessions*	7	*(1) grasp-release, (2) left-right, (3) left-right-up-down, (4) left-right-up-down-grasp release, (5) all 8 classes*	*linear C-SVM* *Multiclass classification* *Grid search C≤1000* *Amplitude of steady-state 128 sample windows at 2048 HZ.*	Left-right and grasp-release degrees of freedom more stable Simple tasks less time than complex ACC 92-98%
Bitzer & Smagt, 2006	To classify movements of open and close hand from the forearm EMG.	-	10	*6 movements: thumb flexion and extension, the index finger, and the middle, ring and little fingers together (approximation)*	*SVM multiclass, OAA.* *Linear(C=0.1) and RBF kernel (C=20)* *Posterior probabilities* *1 data point = 5 samples*	Relax position: $\bar{p} = 0.94$ ($0.89 \leq p \leq 0.97$) Train relax/ test prone: $\bar{p} = 0.74$ Train prone/ test relax: $p > 0.94$ Train prone/ test prone: $p = 0.87$ (different sets $p > 0.9$). Session independence (relax/prone train): relax: $p = 0.95 \pm 0.03$, $\varepsilon t = 0.55 \pm 0.21$ pronation: $p = 0.94 \pm 0.04$, $\varepsilon t = 0.41 \pm 0.20$
Castellini, Smagt, Sandini, & Hirzinger, 2008	To compare three different machine learning approach and related it to the type of grasp and related force.	*1 male x 15 sessions*	10	*Opposing the thumb and index, middle, ring or the thumb and all other fingers*	*NN, SVM and LWPR* *Gaussian RBF*	Grasp types ACC: 90% NRMSE: 10%
Lucas, Gaufriau, Pascual, Doncarli, & Farina, 2008	(1) To prove that wavelet is suitable for myoelectric control (2) To study impact of mother wavelet in discrimination	*6 subjects*	8	*Wrist flexion, wrist extension, hand supination, pronation, opening, and closing*	*Multiclass SVM, OAA.* *RBF, 5-fold cross-validation* *DWT features*	Discrimination error rate average 5%. Not possible to select a mother wavelet Averaged misclassification rate: 4.7% (optimization), 11.1% (Daubechies), 16.4 % (worst wavelet), 18.4% (spectral)
Oskoei & Hu, 2008	SVM classification of upper limb motions using myoelectric signals, and segmentation and features accuracy.	*11 healthy*	4	*2 x Isotonic flexion, extension, abduction, adduction, and keeping the hand straight*	*Multiclass C-SVM, OAO* *SL≤ 200 ms disjoint segmentation* *SL> 200 ms overlapped segmentation.* *Majority vote* *MAV, RMS, WL, VAR, ZC, SSC, WAmp, MAV1, MAV2, PS, AR2, AR6. FMN, FMD*	TD more stable than TF AR features improve with segment length increase WL best feature TD sets MAV+WL+ZC+ SSC and MAV + WL better ACC SVM better than LDA and MLP
Castellini & Smagt, 2009	To compare machine learning approaches on forearm EMG.	-	10 surface	*Opposing the thumb and: index, middle, ring or the thumb and all other fingers*	*NN, SVM (RBF), LWPR* *Disjoint segmentation*	Electrode displacement determinant role to accuracy Cross session accuracy is highly correlated to the *average minimum inter-sample distance*

Tabela 12 [Table 1 - SVM on EMG myoelectric control forearm muscles (continuation)].

Authors	Objective	Method				Main findings
		Participants	***Channels***	*Task*	***Classification, segmentation and features***	
Castellini, Gruppioni, Davalli, & Sandini, 2009	To study how phantom limb posture could be discriminated using EMG and force signals.	*3 hand amputees*	5	*Grasping postures*	*PCA (pre-processing)* *Sample-by-sample EMG* *Multi-clas SVM and $\epsilon - SVM$* *Grid-search: $C = 10^{1\ldots3}$; $\gamma = \frac{10^{0\ldots2}}{5}$*	Accuracy classification: 95.74% to 79.72% RMSE regression: 6.54% to 17.76% Lower SV%: classification: 24.59% to 15.83% Lower SV%: regression: 12.98% to 8.28%
Fontana & Chiu, 2009	To design an intelligent classifier that incorporates in the training process both variations on the electrode placement and the time-frequency representation.	*6 healthy*	6	*Wrist abduction, adduction, extension, flexion and rest*	*Multiclass directed acyclic graph RBF* *DWT each 256 points*	Minimum of 3 sessions to ACC > 89%. ACC Extension and rest > 90% ACC Abduction and adduction > 85%
Chen, Zhu, & Zhang, 2010	To propose discriminant bispectrum features to classify hand and wrist motions.	*5 males* *2 females*	6	*8 classes of hand and wrist motion*	*LLC, GMM, MLP, and C-SVM (RBF)* *Window of 256 ms/ increment of 128 ms.* *AR, PSD, RMS, Hudgins'TD*	DBS = 98.18% DBS+AR = 98.46% AR + RMS = 97.03% Order: DBS + AR > DBS > AR + RMS > TD > PSD > AR > RMS
Yang et al., 2009	To presents anthropomorphic robot hand controlled by the EMG signals.	-	6 Otto Bock	*27 different hand gesture modes: basic and 3 extended postures*	*Multiclass SVM-OAO (RBF)*	C = 32; γ=0.125 (grid search) Training ACC = 99.2% ACC CV = 73.5% ACC each group = 82.5%
Khokhar, Xiao, & Menon, 2010	To use EMG pattern recognition to control a wrist exoskeleton in real-time situation.	*8 subjects*	4 Surface	*12 protocols:* *MVC;* *10'' at 40-50%, 100% MVC torque* *19 classes [10-50%]MVC torque*	*Multiclass SVM-OAO (RBF)* *Window of 250 ms/ increment of 125 ms* *RMS, AR, WL.*	19 classes C = [45-90]; γ = [0.7-1]; ACC = 91.5[88.11-94.26]% T ACC = 88.2[84-93.57]% 13 classes C = [50-90]; γ = [0.7-1]; ACC = 97.97[95.83-99.72]% T ACC = 96.52[94.05-99.47]%
Boschmann & Platzner, 2012	To analyze the accuracy of LDA, SVM and kNN in electrode shift using.	*1 healthy male*	96 Block	*11 hand and wrist movements*	*LDA, SVM, k-NN* *Overlapped window of 100 ms with 50 ms increment* *MAV, WL, ZC, SSC*	SVM or kNN need 32 sensors to compensate 1cm shift SVM and kNN accuracy drop below 60% 2 cm shift LDA decrease more with electrode shifting

Tabela 12 [Table 1- SVM on EMG myoelectric control forearm muscles (continuation)].

Authors	Objective	Method				Main findings
		Participants	*Channels*	*Task*	*Classification, segmentation and features*	
Fligge, Urbanek, & Smagt, 2012	To improve accuracy and ability to distinguish a large number of objects separating shape, size and weight-related attributes of the object using forearm EMG.	*2 female* *4 male*	50 surface block.	*Two experiments 3 methods:* *(1) Grasp one of eight objects from 16 different.* *(2) Grasp three different weighted 1l bottles and two 0.5 bottles.*	*Multiclass SVM, OAO* *20-fold CV (experiment 1)* *10-fold CV (experiment 2)* *3 features methods* *Normalization in time (n=15), iEMG, and division for EMGmáx*	Objects shape may be distinguished in the first half of the task, but distinguishability of object size improves over the whole reaching movement EMG data gathered before the hand contact with the object can be used to obtain information on object properties Mean accuracy of methods: 87.7% (A), 87.8% (B), and 87.1% (B).
Khushaba, Kodagoda, Takruri, & Dissanayake, 2012	To investigate the discriminating accuracy between individual and combined fingers movements using surface EMG signals to control prosthetic hand.	*6 males* *2 females*	2 Surface	*6 x ten classes of individual and combined fingers movements including flexion, pinching and hand close.*	*Bayesian fusion: SVM + kNN* *C-SVM classification,* *RBF ($C = 8$; $\gamma = 12/n$ (features).* *Disjoint and overlapped segmentation with majority vote* *50 – 100 – 150 ms segments* *SSC, ZC, WL, HTD, SS, AR. Set features LDA projection*	LDA projection and SVM+kNN needs only two EMG channels Large number of voting decisions does not always improve the classification SVM+kNN post processing showed the use of an analysis window length of 100 ms with 7-to-9 voting decisions had better accuracy
Chen & Wang, 2013	To classify multiple finger movements. To propose MKL - SVM.	*3 females* *3 males*	4 Surface	*10 finger chinese number 0.5'' gestures [0-9]*	*k-NN, LDA, QDA, SVM, and MKL-SVM* *32 samples window (64 ms)* *Hudgins'TD, ACCC, SPM,*	ACCC + k-NN: $\bar{x} = 64.7\%; max = 92\%$ TD + LDA=$\bar{x} = 95.9\%; max = 99\%$ SPM + SVM:$\bar{x} = 95.03\%; max = 98.03\%$ 3F + MKL-SVM:$\bar{x} = 97.93\%; max = 99.67\%$ TD and 3F better performance
(Khushaba, Kodagoda, Liu, & Dissanayake, 2013)	To present a control scheme that discriminate EMG from forearm to avoid driver distraction.	*8 participants*	8 Surface	*14 classes of fingers gestures*	*LDA, RegTree, NB C-SVM RBF ($C = 8$; $\gamma = 12/n$ features).* *ZC, SSC, SS, RMS, MAV, IAV, AR, HTD* *OFNDA, FNDA and comparisons*	*OFNDA, FNDA with an average of less than 7% error rates* Error rates decrease with increasing channels
Kumar, Poosapadi Arjunan, & Singh, 2013	To propose single channel EMG classification with twin SVM of finger flexion.	*11 able-bodied* *1 trans-radial amputated*	-	*Four finger flexions*	*Linear TSVM* *RMS local wavelet maxima*	Able-bodied accuracy: 93% Able-bodied sensitivity: 94% Amputated accuracy: 81% Amputated sensitivity: 84%

Legend: ACC – accuracy; ACCC - autocorrelation and cross-correlation coefficients; AR – autoregressive; DBS – discriminant bispectrum; FMD – median of frequencies; FMN – mean of frequencies; FNDA - fuzzy neighborhood discriminant analysis; HTD - Hjorth time domain; IAV – integral absolute value; k-NN - k-nearest neighbor; LDA – linear discriminant analysis; LWPR – locally weighted projection regression; MAV – mean absolute value; MKL-SVM – multiple kernel learning SVM; NN – neural network; OAA – one-against-all; OAO – one-against-one; OFNA - orthogonal fuzzy neighborhood discriminant analysis; PS - power spectrum; QDA – quadratic discriminant analysis; RBF – radial basis function; RMS – root mean square; SPM - spectral power magnitudes; SS - sample skewness; SSC - slope sign changes; SVM – support vector machines; TD – time-domain; TSVM – twin support vector machines; VAR – variance; WAmp - William amplitude; WL – waveform length; ZC - zero crossings.

Tabela 13 [Table 2 - SVM on EMG myoelectric control hip muscles].

Authors	Objective	Method				Main findings
		Participants	*Channels*	*Task*	*Classification, segmentation and features*	
Huang, H., Zhang, F., Hargrove, L. J., Dou, Z., Rogers, D. R., Englehart, 2011	To develop an interface based on neuromuscular–mechanical fusion to recognize ambulation tasks and predict task transitions.	*5 unilateral transfemoral amputations*	7–9 surface	*Hip muscles* *Six gait locomotion mode and five transition mode*	*LDA, multiclass C-SVM, OAO* *RBF* ($\gamma = 1/n$) *Majority vote method* *Sliding window of 150 ms, and 12 ms increment.* *MAV, SSC, WL, ZC.*	Neuromuscular–mechanical fusion better than methods with a single data source. Low classification accuracies in the swing phase compared to the stance phase (kinematics may help). Stance phase ACC: 99% Swing phase ACC: 95%

Legend: ACC – accuracy; LDA – linear discriminant analysis; MAV – mean absolute value; OAO – one-against-one; RBF – radial basis function; SSC - slope sign changes; SVM – support vector machines; WL – waveform length; ZC - zero crossings.

Tabela 14 [Table 3 - SVM vs EMG neuromuscular diagnostic].

Authors	Objective	Method				Main findings
		Participants	*EMG*	*Muscles and task*	*Classification segmentation and features*	
Güler & Koçer, 2005	To analyze the performance of SVM and back-propagation ANN in diagnostic of neuromuscular disorders.	*19 healthy* *20 neuropathy* *20 myopathy*		*4'' biceps brachii and hypothenar group* *Maximum forearm flexion* *Maximum finger abduction*	*ANN and SVM* *AR Burg model*	SVM ACC: 94.4% SVM ROC area: 0.954 ANN ACC: 91.6 % ANN ROC area: 0.927
Katsis, Goletsis, Likas, Fotiadis, & Sarmas, 2006	To propose an EMG decomposition method and MUAP classification of neuromuscular disorders.	*10 healthy* *20 neuropathy* *20 myopathy*	*Surface*	*Biceps brachii* *30% MVC* *Elbow flexion*	*Multiclass SVM OAO* *RBF* $(C = 2^{-2.25};\ \gamma = 2^{6.25})$. *MUAP clustering* *Decomposition of superimposed MAUPs*	ACC: 86.14% Success rate 95.24%
Kaur, Arora, & Jain, 2009	To classify neuromuscular disorders using different segmentation techniques.	*3 healthy* *5 myopathy* *4 motor neuron disease*	*Deep*	*Biceps brachii muscle* *5'' 30% MVC*	*Multiclass SVM OAA* *MUAP clustering:* *- peaks of MUAP* *- beginning/ ending extraction point* *-DWT* *AR features*	Segmentation rate Peaks- 95.9% Extortion points – 75.4% DWT – 66.6% MUAP clustering – 93.13% AR – 100%
Sultornsanee, Zeid, & Kamarthi, 2011	To explore an approach for the diagnosis of neuromuscular disorders using RCA and SVM.	*Set1:29 healthy 88 neuropathy 75 myopathy* *Set2:15 healthy 50 neuropathy 35 myopathy*		*Tibialis anterior*	*Multiclass SVM (RBF)* *RQA features:* *-Recurrence rate* *-Determinism* *-Laminarity*	SVM ACC: 98.28% Healthy: 97.72% Neuropathy: 97.10% Myopathy:100%
Christodoulou et al., 2012	To discriminate neuromuscular disorders using multiscale AM-FM features.	*20 normal* *11 myopathy* *9 neuropathy*	*Surface*	*Biceps brachii* *10%, 30%, 50%, 70% and 100% MVC,*	*kNN, SOM and SVM (RBF)* Multiscale AM-FM	SVM - ACC: 75%,: SEN: 65% SPE: 85% Method: all forces[10 -100]% kNN 58[53-63]% SOM 60[55-63] SVM 78[68-75]
Conradsen et al., 2012	To propose an automatic epilepsy seizure detection system.	*10 healthy* *1 epileptic*	*Surface*	*7 channels bilateral* *Stimulated seizures:* *Myoclonic* *Versive-asymmetric* *Tonic-clonic*	*C-SVM (linear)* *C=0.8* *DWT, WPT* *(ACM+ANG+EMG)*	DWT (range all [range EMG]) SEN:100% [49-100] FDR: 0-1.7 [1.1-50] LAT: 0.75''[0.75-11] WPT (range all [range EMG]) SEN:100% [39-97] FDR: 0-18 [2.9-33] LAT: 1'' [1-1.9]

Table 14 [Table 3 - SVM vs EMG neuromuscular diagnostic (continuation)].

Authors	Objective	Method				Main findings
		Participants	*EMG*	*Muscles and task*	*Classification segmentation and features*	
Dobrowolski, Wierzbowski, & Tomczykiewicz, 2012	To propose a quantitative EMG method for classification of neuromuscular disorders.	*30 myogenic* *70 neurogenic* *200 healthy*	*Deep*	*Deltoid muscle* *(comparing other muscles as 1st dorsal interosseous, vastus lateralis and tibialis anterior)*	*Two linear C-SVM* *C_0=1000* *Symlet 4 wavelet Scalogram* *(Compare Symlet4,6 vs Daubechies4,5 vs Coiflet2)*	Myogenic SEN: 98.4% Neurogenic SEN: 98.7% SPE: 100% Total error 0.6%.
Goker et al., 2012	To classify myoclonic epilepsy through scanning EMG with machine learning algorithms.	*9 JME* *10 Control*	*Deep*	-	*NN, DT, NB* *C-SVM (RBF)* *C=1; γ=1/n (attributes)* *LCS, max, min, var, std, mean, median, mode, peak-peak voltage*	SVM SEN: [0.841-0.932] SVM SPE: [0.639-0.852] SVM ACC: [0.796-0.828] DT ACCmax: 0.992 NN ACCmax:0.954 NB ACCmax:0.923
Subasi, 2012	To compare different SVM and others methods for the diagnostics of neuromuscular disorders.	*7 myogenic* *13 neurogenic* *7 healthy*	*Deep*	*Biceps brachii* *Isometric elbow flexion 30% MVC*	*LDA,ANN,RBFN, DT* *SVM and FSVM (RBF, OAA)* *C=300; γ=0.2 (SVM)* *C=250; γ=0-2 (FSVM)* *DWT - Daubechies 4* *Subband decomposed EMG*	SVM (I-E cv) SPE: 90.25 - 86.75% Myo SEN: 97.75 - 81.5% Neu SEN: 99 - 98.5% $\overline{ACC}$: :96.67/ - 92.25% F-SVM(I-E cv) SPE: 95.25 - 90.25% Myo SEN: 98.25 - 94.5% Neu SEN: 99.5 - 95.75% $\overline{ACC}$: 97.67 - 93.5% *Others lower ACC with min LDA 91.67 – 79.1%*
Subasi, 2013	To propose PSO-SVM to increase discriminating accuracy in neuromuscular disorders.	*7 myogenic* *13 neurogenic* *7 healthy*	*Deep*	*Biceps brachii* *Isometric elbow flexion 30% MVC*	*k-NN, RBFN,* *SVM and PSO-SVM(RBF)* *DWT* *Subband decomposed EMG*	SVM SPE: 93.5% Myo SEN: 97.8% Neu SEN: 99% $\overline{ACC}$:96.75% PSO-SVM SPE: 95.25% Myo SEN: 98% Neu SEN: 99% $\overline{ACC}$: 97.41% *k-NN $\overline{ACC}$: 95.17%;, RBFN $\overline{ACC}$: 94.08%*

Legend: ACC – accuracy; ACM – accelerometer; AM-FM: amplitude modulation–frequency modulation; ANG – gyroscopes angular velovity; ANN – artificial neural network; AR – autoregressive; DT – decision tree; DWT – discrete wavelet transform; FSVM – fuzzy SVM; k-NN - k-nearest neighbor; LCS - length of cross section; LDA – linear discriminant analysis; MAV – mean absolute value; NB - Naïve Bayes; NN – neural network; OAA – one-against-all; OAO – one-against-one; PSO-SVM – particle swarm optimization-SVM; RBF – radial basis function; RBFN - radial basis network; RQA – recurrence quantification analysis; SEN - sensitivity; SPE – specificity; SOM - self-organizing map; SVM – support vector machines.

Tabela 15 [Table 4 - SVM vs EMG pattern recognition and kinesiology analysis].

Authors	Objective	Method			Main findings
		Participants	*Muscles and task*	*SVM, segmentation and features*	
Yan, Wang, & Xie, 2008	To propose MI-based selection and fuzzy LS-SVM to motion discrimination.	*30 subjects*	*Surface forearm* *Fist clench, fist stretch, wrist pronation, and wrist supination*	*Fuzzy LS-SVM, NN* *EMG activity by skidding window* *WPT-Coiflet 4* *MI-based feature selection* *Compare CSI with PCA, SFS, BE*	Fuzzy LS-SVM ACC: [0.9425 – 0.9742] NN ACC: [0.8971 – 0.9292] Pronation more difficult to discriminate CSI more time-cosuming
Shi, Zheng, Chen, & Xie, 2009	To propose a bone-muscle level system to describe wrist angle and thickness of the extensor carpi radialis.	*6 males* *1 female*	*Extensor carpi radialis* *Wrist extension–flexion* *(15, 22.5, and 30 cycles/min)*	*ANN, LS-SVM* *one-element LR*	(mean) – Proposed / BP-ANN / LS-SVM R^2 – 0.96 / 0.96 / 0.96 SRME (º) – 7.26 / 7.26 / 7.23 RRMS E (%) – 16.42 / 16.42 / 15.87
Stirling, Von Tscharner, Kugler, & Nigg, 2011	To verify if EMG signals change during running and if these changes are related to the effort level.	*15 female recreational runners*	*TA, GM, VL, ST* *Running 95% of ventilatory threshold*	*C – SVM* *$C = 10^m$ (m=1:0.5:4)* *k-CV (k=2:2:20)* *Time–frequency analysis using wavelets*	Recognition rates(mean[max], C, k-CV): TA – 89.2[93.3]%, 100, 4 GM – 88.3%[91.4], 316,4 VL – 84.6[89.9]%, 10^4,12 ST – 94.0[94.4]%, 10^4, 14
Tavakolan, Xiao, & Menon, 2011	To know the feasibility of hand postures classification in the elderly.	*12 seniors (70 Y)* *7 young (27 Y)*	*4 channels of forearm* *Combination of several hand movements dividing in 8 protocols*	*Multiclass SVM, OAO.* *RBF, 8-fold cross-validation* *C = (0,100)* *γ =(0,3).* *Segments of 250 ms (256 samples) and increase of 125 ms (128 samples)* *WL ,RMS, AR*	Seniors mean accuracy - 90.62%. Young mean accuracy – 97.6% Mean of C = 14.2 Median of C = 10 Mean of γ = 0.8 Parameters C and γ similar both populations 4 channels sEMG suitable
Tolambiya, Thomas, Chiovetto, Berret, & Pozzo, 2011	To know how support vector machines can be utilized in the analysis of electromyographic data underlying a whole body pointing task.	*10 healthy males*	*24 muscles of lower limb and trunk* *6 trials x Reaching movements pointing task towards a target without and with constrains (postural and focal)*	*Binary SVM* *vector of 200 elements each muscle* *Wigner kernel* *5-fold cross validation*	Without constrain from: knee extended – 96.7% Reduced base of support – 79.4% Straight finger trajectory – 90.6% Semicircular finger trajectory - 87.6%

Tabela 15 – [Table 4 - SVM vs EMG pattern recognition and kinesiology (continuation)].

Authors	Objective	Method			Main findings
		Participants	*Muscles and task*	*SVM, segmentation and features*	
Alkan & Günay, 2012	Classifications of EMG signal with five discriminant functions and a SVM.	-	*Elbow flexion, elbow extension, forearm pronation and forearm supination*	*SVM and DA* *10-fold cross validation* *MAV* *32 ms window length*	SVM ACC 99% DA ACC 96-98%
Futamata, Nagata, & Magatani, 2012	To evaluate the discriminant performance of multiclass SVM on hand motion using EMG.	*5 subjects*	*6 channels* *Forearm* *18 hand movements*	*CDA, Linear and non-linear SVM (RBF PSO-SVM OAO)* *Overlapped segmentation with 60 ms lag* *IEMG of 300ms*	3 channels accuracy ≤ 90% 5, 6 channels better accuracy [88.56, 99.56]
Natarajan, Wininger, Kim, & Craelius, 2012	To analyze elbow kinematics influence in arm flexion using EMG features.	*14 right-handed (healthy)*	*biceps* *Elbow flexion*	*PCA, SVM* *8 EMG features of 3 types: (independent. EMG, temporal, special)* *-peak EMG amplitude* *-peak EMG timing* *-peak EMG location*	Sensitivity = 81.22% Specificity = 86.14% Peak EMG features combination showed high performance
Ziai & Menon, 2011	To compare the performance of regression for estimation of isometric joint torque using EMG.	*8 males* *3 females* *(healthy)*	*8 channels from forearm* *Wrist isometric flexion/extension reaching different MVC levels (25% and 50%).*	PBM, OLS, RLS, SVM, ANN, LWPR, and $\varepsilon - SVR$ *Gaussian Kernel, 8-fold cross-validation* *Accuracy measured by NRMSE and* R_a^2.	LWPR lower R_a^2 and higher *NRMSE*. All models were sensitive to time and electrode displacement. 5 and 8 channels similar accuracy 2 channels lower R_a^2 and higher *NRMSE* Resampling reduces the training time OLS trained faster

Legend: ACC – accuracy; AR – autoregressive; CSI - cluster separation index; DA – discriminant analysis; ANN – artificial neural network; DWT – discrete wavelet transform; FSVM – fuzzy SVM; LS – least squares; LWPR – locally weighted projection regression; MAV – mean absolute value; MI – mutual information; OAO – one-against-one; OLS – ordinary LS; PBM – physiological based model; PSO-SVM – particle swarm optimization-SVM; RBF – radial basis function; RLS – regularized LS; RMS – root mean square; SVM – support vector machines; WL – waveform length; WPT - wavelet packet transform.

5.4.2. Segmentation and feature extraction

5.4.2.1. Segmentation

A segment is the length of time for myoelectric data acquisition considered for feature extraction. Besides signal processing, the first concern of the researcher is the source of the signal on which to perform the segmentation and feature extraction. In segmentation there are two aspects that require investigator decisions according to the proposed problem: EMG source (steady-state and/or transient state) and window length. Crawford et al. (2005) collected the steady-state EMG of a static gesture, then extracted features, and used linear SVM classification output to control the movement of a robot arm. They stated two ways of EMG application in the field of prosthetics control: (1) using the amplitude of the steady state EMG signal in a constant muscle effort, where a single channel is used as a binary switch or control a degree of freedom in amplitude; (2) using discrete actions where the temporal structure of the transient EMG activity is submitted to classification. Segmentation on steady state is performed during a period of continuous muscle contraction. The transient state is a period that contains information about the beginning of muscle contraction (the onset). The relation between task and the temporal parts of the signal could increase the accuracy of muscular action in object manipulation. For example, EMG data gathered before the hand contacts with the object can be used to obtain information on object properties (Fligge, Urbanek, & Smagt, 2013). The advantage of using steady state is a higher accuracy than transient state, and a more stable accuracy for different segment lengths **(**Huang, Zhang, Hargrove, Dou, Rogers, & Englehart, 2011)**.** In both cases, EMG from multiple channels is needed to provide the largest information and discriminatory ability (Englehart & Hudgins, 2003). So, the relation between the quality of this information and feature extraction assumes an important role, since it influences the interpretation of results. However, if some signal characteristics can be related with the amount of strength and precision, rather than with its classification, regression analysis is necessary to build a map from EMG to the force task (Castellini & Smagt, 2009), as well as other kinematic and kinetic variables (Huang et al., 2011).

The decision about steady or transient EMG states depends on the problem under study. In myoelectric control there are some studies where neighboring segments to transient periods are eliminated with the intent of increasing accuracy. However a question arises: Could it be a good solution to discriminate gross movements, but at the same time losing information about the way the central nervous system coordinates muscles in finer tasks? For example, in gait analysis, the complexity of labeling the sample windows from transitions between locomotion modes is due to the lack of a well-defined moment of transition that would serve as separation between such modes (Huang et al., 2012).

As stated before, there are two main methods for segmentation: disjoint and overlapped. Oskoei and Hu (2007) state that a length of an adjacent segment plus the processing time generating the control must be no more than 300 ms, due to real-time constraints. After the decision on the length of the segment, it is necessary to apply the technique that will perform data segmentation. Oskoei and Hu (2008) evaluated SVM classification in myoelectric signals of upper limb using disjoint segmentation for segment length equal or less than 200 ms (50, 100, 150 and 200 ms), and overlapped segmentation for higher windows (300 and 500 ms). These authors found that for single features, the accuracy of classification decreases with the shortening of segment length, since it yields more bias and variance. Khushaba et al. (2012) used continuous segmentation performing post-processing majority vote. They concluded that LDA projection and k-nearest neighbor SVM needs only two EMG channels performing post-processing majority vote, but a large number of voting decisions does not always improve the classification. Majority voting is a post-processing of continuous segmentation that takes into account the acceptable delayT_d. For a given decision point, the vote decision includes the previous and the last m samples, whose value is the class with the better number of occurrences in $2m + 1$ decision points. Then, the number of decisions used is determined by the processing time (τ) and T_d (Englehart & Hudgins, 2003). These authors stated that with no post-processing, T_d and τ tend to be the same, it is just the time from the onset of myoelectric purpose until the control system decision. With post-processing, T_d is the period with all decisions. The response time goal is given by

$$\tau \cdot m \leq T_d.$$

Accuracy improves with a great number of decisions in post-processing despite the increase in processing time. The processing delay increases with the window length, being most pronounced and linear from 100 ms (Khushaba et al., 2012).

Englehart and Hudgins (2003) studied the size of window length acceptable for classification without losing relevant information. Several values were found in the literature from an inferior limit of 32 ms (Alkan & Günay, 2012; Chen & Wang, 2013), to a superior limit of 250-256 ms (Chen et al., 2010; Khokhar et al., 2010; Tavakolan et al., 2011). Intermediate values between those were also found (Boschmann & Platzner, 2012; Khushaba et al., 2012). Despite the length in milliseconds being well established, is not enough to reference only window size without information about the sample rate. The number of samples that each window contains will depend on the sample rate, and the feature quality depends of that number of samples. Another segmentation approaches were about the identification of MUAP peaks, and determination of the beginning and the end of extraction points (Kaur et al., 2009).

5.4.2.2. Feature extraction and channels

The classifier is fed with features rather than with the raw signal with the expectation of improving classification efficiency. The choice and the process of features extraction is the most critical step in pattern recognition analysis, since they should contain information EMG behavior belonging to a certain class. The task of removing irrelevant and redundant features is also crucial to achieve an optimal combination decreasing the dimensionality of the feature space leading to lower complexity and better discrimination (Phinyomark, Limsakul, & Phukpattaranont, 2009; Phinyomark, Phukpattaranont, & Limsakul, 2012).

Researchers choose between time-domain, frequency domain or time-scale features, having in mind the structural analysis, or a combination of these domains considering a phenomenological analysis (see Oskoei & Hu, 2007). A good explanation about time-domain and frequency domain features can be found in the paper of Phinyomark, Phukpattaranont, and Limsakul (2012). These authors showed a high use of time-domain features with several comparisons between these ones and other feature types. The main categories of articles that use time-domain are from myoelectric control and kinesiologic analysis rather than neuromuscular disorders that tend to focus on time-scale features.

Chen et al. (2010) proposed the use of discriminant bispectrum features reaching a precision of 98% in hand and wrist motion classification. The feature sets showing higher accuracy were the discriminant bispectrum and autoregressive coefficient. These results are similar to those found by Oskoei and Hu (2008) since autoregressive coefficient features improve with higher segment length, and Chen et al. (2010) used a 256 ms window. Güler and Koçer (2005) used autoregressive Burg model achieving the model order by minimizing an Akaike Information Criterion given by $Nln(\rho_p) + 2p$ where p is the model order, N the number of data points and ρ_p the error variance. The classification obtained showed a suitable option for neuromuscular disorders discrimination. However, Oskoei and Hu (2008) reported that time-domain single features are more stable than time-frequency single features, and time-domain set of features showed better accuracy than single features. The classification order found by Chen et al. (2010) is presented in Table 1.

Signal non-linear characteristics contain information that can bring several combinations of specific movements translation. Extracted features should contain that information. Chen, Zhu, and Zhang (2010) suggest that the non-Gaussian and non-linear features extracted from EMG signals provide useful information in relation to hand and wrist motion classification. These authors considered that bispectrum analysis is an important higher order statistical tool for detecting deviation from a Gaussian process, and investigate non-linear interactions between two harmonic signal waves. However, one must bear in mind several assumptions, since the EMG is not merely Gaussian and high-order statistics methods cannot be discarded when the marginal probability distribution of the EMG signals is symmetric (Orosco, Lopez, & di Sciascio, 2013). Yet, Chen and Wang (2013) stated that high-order statistics methods usually perform well for stationary or weak-sense stationary EMG with continuous muscle contraction. In other cases, it is therefore preferable to use time-domain features, autocorrelation and cross-correlation coefficients and/or spectral power magnitudes features. Since the analyzed tasks involved in the classification are due to motor behavior, it seems reasonable to assume that they are continuous and derivable up to any arbitrary order, and it is expected that small variations in point samples close to each other are detectable by machine learning. Castellini and Smagt (2009) removed from the training set samples there were too close to each other (batch standardization) according to a suitable notion such as the Mahalanobis distance as the inter-sample distance. They found that this procedure reduces the training set size without degrading performance.

In time-frequency analysis the wavelet decomposition is a common method used for feature extraction. Along with SVM classification, it significantly simplifies the diagnostic process and increases accuracy due to the use of a single classification parameter. Lucas et al. (2008) applied a two-class SVM approach to analyze if wavelet signal representation is suitable for myoelectric control, and if the mother wavelet selection is common between the signals of different motion recognition (Lucas et al., 2008). Nevertheless, it is difficult to have a common mother wavelet between different channels, and vector reducing to a single decision parameter after SVM analysis seems be a “goofy” approach. The Wavelet Index showed higher accuracy when compared to methods parameters as amplitude, area and duration (Dobrowolski et al., 2012). In the study of Subasi (2012) three steps of analysis were mentioned. Decomposition of EMG signal into different frequency bands using discrete wavelet transform was done, then features were extracted from the decomposed signals. This was done to increase accuracy; afterwards, the unknown EMG signal was classified as normal, myogenic or neurogenic. Despite the discrete wavelet transform features being the most used wavelet transformation methods, it is also possible to find wavelet packed transformation (Yan et al., 2008) and both discrete wavelet transform and packed transformation (Conradsen et al., 2012) approaches. Discrete wavelet transform and packed transformation provide a good frequency resolution at low frequencies and good time resolution at high frequencies (Conradsen et al., 2012).

There are fewer approaches to the phenomenological analysis and the combination of kinesiologic variables. Nevertheless, we can highlight the recurrence quantification analysis (RQA) to discriminate neuromuscular disorders (Sultornsanee et al., 2011), and a partial phenomenological analysis that involves peak EMG based-features in the analysis of kinematics (Conradsen et al., 2012; Huang et al., 2012; Natarajan et al., 2012; Shi et al., 2009), and EMG scanning (Goker et al., 2012).

An EMG recurrence plot is a representation of the nonlinear EMG signal; some authors have focused on the following features of recurrence plot (Sultornsanee et al., 2011): *recurrence rate*, *determinism*, and *laminarity*. RQA is a mathematical technique to analyze nonlinear data and quantify the number and duration of recurrences of a dynamic system represented by its state space trajectory (Marwan, Thiel, & Nowaczyk, 2002; Marwan, Carmenromano, Thiel, & Kurths, 2007; Marwan & Kurths, 2004; Zbilut & Webber, 2006).

Conradsen et al. (2012) combined accelerometer data and gyroscope angular velocity as complementary data to EMG information to distinguish simulated epilepsy seizures from normal activities using wavelet transformation. Natarajan et al. (2012) extracted features considering three main categories: independent EMG parameters, temporal relationships, and spatial relationships. The first ones are basic time-domain features, such as mean, standard deviation and integral of EMG signal over time. The so-called temporal relationship refers to the normalized time correspondence of peak EMG and the cumulative duration above 10% normalized EMG maximum threshold. The spatial relationships include the cumulative angular distance crossing until normalized EMG was above the mentioned threshold, and the angle corresponding to EMG peak. Shi, Guo, Hu, and Zheng (2012) used a bone-muscle lever system that describes the relationship between wrist angle and extensor carpi radialis thickness. These approaches allow for better feature information and greater scope of classification tasks.

Goker et al. (2012) developed EMG scanning for discrimination of juvenile myoclonic epilepsy from healthy subjects. This method introduces the length of cross section and maximum amplitude parameters. Thus it is expected to obtain both spatial and temporal information. Nevertheless, SVM showed lower accuracy than others such as feed-forward neural networks, decision tress, and Naïve Bayes (achieving 100% of detection sensitivity).

Another approach found for feature extraction is multi-scale amplitude modulation–frequency modulation (Christodoulou et al., 2012). For each signal, the amplitude modulation–frequency modulation histogram features were extracted, getting values as instantaneous amplitude, instantaneous phase, and instantaneous frequency. The authors highlight a larger number of outliers in the case of neuromuscular disorder and point out that the most significant differences were found in the higher frequencies.

One can compare the type of features (structural or phenomenological analysis), but also the number of channels that are considered, since all will contain an amount of information to classify the phenomenon. The appropriate number of channels, as the number of features must consider the highest accuracy with the lower time cost, namely in real-time cases. The use of less than three channels is associated with lower accuracy in both classification (Futamata et al., 2012) and regression tasks (Ziai & Menon, 2011), and classification performance with four channels shows a suitable choice to discriminate as much as 15 different gestures (Tavakolan et al., 2011). However, Khushaba et al. (2012) reported that LDA projection and k-nearest neighbor SVM needs only two EMG channels performing post-processing majority vote, but a large number of voting decisions does not always improve the classification. On the other hand, calculating mutual information reduces the feature set without compromising classification accuracy. For fuzzy least squares-SVM, it seems that 8-10 channels is a good choice (Yan et al., 2008).

5.4.3. SVM considerations

The decision about the selection of the SVM model and parameters depends on class number and data type. Other methodologies can be added to the algorithm in order to maximize performance. Several approaches have been proposed to study pattern recognition in EMG data, such as Bayesian fusion (Khushaba et al., 2012); fuzzy-SVM (Subasi, 2012); fuzzy least squares SVM (Yan et al., 2008); particle swarm optimization SVM (Futamata et al., 2012; Subasi, 2013); twin SVM (Kumar et al., 2013); multiple kernel learning SVM (Chen & Wang, 2012).

Besides these approaches, and even after the process of segmentation and features extraction, usually a simple issue remains: linear SVM or non-linear SVM. The nonlinear shows a better performance when supplied with less information comparing to the linear, but with more than four channels linear could be a better choice, especially due to lower computational time (Futamata et al., 2012). Dobrowolski et al. (2012) used a penalty cost C value started on 1000 (high value) and all Lagrange parameters satisfied the $\alpha_i < C$ condition, showing that training data sets are linearly separable in neuromuscular disorders classification with wavelet transformation features, therefore nonlinear classifiers are not necessary.

The most widely used SVM model was the non-linear with radial basis function function, in which case it becomes important to choose the kernel parameter γ (or σ depending the literature) in addition to the penalty cost C. Usually, a grid-search is performed to find the best parameters. In Subasi (2012), the original SVM algorithm was used to find the optimal parameters, verifying as better values $\gamma = 0.2$ and $C = 300$ with a minimum of 42 support vectors. For fuzzy-SVM the values of $\gamma = 0.2$ and $C = 250$ were obtained, slightly differing in penalty cost.

One option found in some studies for the calculation of γ is related to the inverse of the number of attributes (Goker et al., 2012; Huang et al., 2011; Khushaba et al., 2012). It would be desirable to avoid the computational cost associated with the grid search and still have the best choice of parameters to prevent misclassification. The estimation of class-conditional probabilities for a given data point can be useful in reducing the number of false classifications due to noisiness in the data. These factors are important for transitional periods when some data between different steady states could not be detectable, and due to the possibility of adapting the classifier to gestures that are easier to classify (Crawford et al., 2005). We also note that not all studies specified grid search procedures for the finding the best parameters.

The studies that showed higher accuracy were about neuromuscular discrimination, achieving 100% using multiclass SVM with autoregressive coefficient features (Kaur et al., 2009), and RQA (Sultornsanee et al., 2011).

5.4.4. Bias factors

Factors as inter-subject variability, posture, muscle fatigue or effort level, and electrode displacement have been primary limitations associated with classification degradation (Castellini & Smagt, 2009; Tkach et al., 2010), namely in myoelectric control and kinesiologic analysis, due to the influence on the stability of features. The primary bias factors in time-domain feature stability of biceps and triceps are electrode shifting and the variability of muscle contraction effort rather than fatigue (Tkach et al., 2010). These factors also compromise the reproducibility of results across studies. Boschmann and Platzner (2012) compared different classification algorithms in electrode shift and found that a displacement of 2 cm will drop accuracy below 60% for SVM and k-nearest neighbor. For a shift of 1 cm, an increased channel number can be a good solution. SVM and k-nearest neighbor do not decrease the accuracy when compensated with 32 channels. Still, LDA always decreases with electrode displacement, being the most sensitive method. SVM showed the higher performance. In the finger motion recognition, Bitzer and Smagt (2006) stated that inaccuracies due to electrode displacement can occur with slight shifts between two different moments such as the end of a finger action and the beginning of a new one. They define that a 300ms window from the end of a finger action is a tolerable error.

The physiologic relation between muscle activity and movement assumes a relevant challenge to the performance of SVM (or other classifiers). The region in which the electrodes are placed is also critical due to EMG cross-talk and EMG patterns such as co-activation. As shown here, most studies focus on the forearm, especially in myoelectric control. The cross-talk and co-activation are present in most upper limb movements (Scheme & Englehart, 2011). As stated before, the number of channels and the availability of suitable musculature also provides a relevant contribution for the power of each classification feature. Crawford et al. (2005) sought to know if all the used channels of EMG data contribute to the classification accuracy and how many channels or electrodes are needed for control a given number of degrees of freedom. To this end, they reduced the number of channels one by one as the slightest error in means of cross validation. They found that for better accuracy, more degrees of freedom require more channels, and the most stable degrees of freedom were left-right and grasp-release. However, there is an intra-subjects variability in the order in which the channels should be dropped. The sites where electrodes are placed and the subject morphology influence the higher or lower amplitude of the EMG record, which is a reason for normalization procedures.

The variation in measures within the same sample can also occur due to intra-session variability and due to a change in the type of motion (Chen et al., 2010). In myoelectric control, the session independence assumes an important role since the patient needs a great performance over time, not only when the train model is constructed. However, using six channels from the forearm, at least three sessions assure an accuracy higher than 89% (Fontana & Chiu, 2009). The main concern is associated with different postures than those used in the training model due to the electrodes displacement (Bitzer & Smagt, 2006; Smagt et al., 2009). That is why it is important to study trained and untrained data into SVM models when the body segment is in different postures.

5.5. Conclusion and Practical Applications

Three major fields of study were identified in literature where SVM learning on EMG signals was used: myoelectric control, neuromuscular disorders discrimination, and kinesiologic analysis. Several approaches were proposed achieving high accuracy in classification. However, there is a clear lack of studies focused on the analysis of the relation between nervous system coordination and the mechanical output in sport gestures, as well as the classification of population practitioners (i.e. skill level, low back pain) using supervised learning. The main scientific areas of application are the classification of neuromuscular disorders and myoelectric control. The SVM showed high performance in this task of classification discriminating gestures on EMG- based features. When compared with no supervised learning techniques, SVM showed better accuracy.

Conflict of Interest

None declared

Acknowledgments

The project "Neuromuscular activity in the golf swing with implications for the practice and in the prevention of overuse injuries" was supported by the Portuguese Foundation for Science and Technology (FCT) fund (PTDC/DES/105176/2008).

References

Alkan A, Günay M. Identification of EMG signals using discriminant analysis and SVM classifier. Expert Syst Appl. 2012;39(1):44–47

Beck TW, Stock MS, Defreitas JM. Paired Pattern Classification of Electromyographic Intensity Patterns During Concentric and Eccentric Muscle Actions. Clinical Kinesiology 2011;65(4):76–82.

Bitzer S, Smagt P van der. Learning EMG control of a robotic hand: Towards Active Prostheses. In: Proceedings of the 2006 IEEE International Conference on Robotics and Automation. Orlando, Florida; 2006:2819–2823.

Boschmann A, Platzner M. Reducing classification accuracy degradation of pattern recognition based myoelectric control caused by electrode shift using a high density electrode array. In: Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society.Vol 2012.; 2012:4324–7.

Boser BE, Guyon IM, Vapnik VN. A Training Algorithm for Optimal Margin Classfiers. In: Proceedings of the Fifth Annual Workshop on Computational Learning Theory.; 1992:144–152.

Burges C. A tutorial on support vector machines for pattern recognition. Data Min Knowl Discov. 1998;2:121–167.

Cao J, Lin Z, Huang G. Composite function wavelet neural networks with extreme learning machine. Neurocomputing. 2010;73(7-9):1405–1416.

Castellini C, Gruppioni, E, Davalli A, Sandini G. Fine detection of grasp force and posture by amputees via surface electromyography. Journal of Physiology, Paris 2009; 103(3-5): 255–62.

Castellini C, Smagt P van der. Surface EMG in advanced hand prosthetics. Biol Cybern. 2009;100(1):35–47.

Castellini, C., Smagt, P. van der, Sandini, G., & Hirzinger, G. Surface EMG for force control of mechanical hands. In: IEEE International Conference on Robotics and Automation Pasadena, CA, USA, May 19-23, Pasadena 2008; 725–730.

Chen X, Wang ZJ. Pattern recognition of number gestures based on a wireless surface EMG system. Biomed Signal Process Control. 2013;8(2):184–192.

Chen X, Zhu X, Zhang D. A discriminant bispectrum feature for surface electromyogram signal classification. Med Eng Phys. 2010;32(2):126–35.

Christodoulou CI, Kaplanis P a., Murray V, Pattichis MS, Pattichis CS, Kyriakides T. Multi-scale AM–FM analysis for the classification of surface electromyographic signals. Biomed Signal Process Control. 2012;7(3):265–269.

Conradsen I, Beniczky S, Wolf P, Kjaer TW, Sams T, Sorensen HBD. Automatic multi-modal intelligent seizure acquisition (MISA) system for detection of motor seizures from electromyographic data and motion data. Comput Methods Programs Biomed. 2012;107(2):97–110.

Cortes C, Vapnik V. Support-vector networks. Mach Learn. 1995;20(3):273–297.

Crawford B, Miller K, Shenoy P, Rao R. Real-Time Classification of Electromyographic Signals for Robotic Control. In: Proceeding of AAAI.Vol Proceeding.; 2005:523–528.

Cristianini N, Shawe-Taylor. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge: Cambridge University Press; 2000.

De Luca CJ. The Use of Surface Electromyography in Biomechanics. J Appl Biomech. 1997;13:135–163.

Dobrowolski AP, Wierzbowski M, Tomczykiewicz K. Multiresolution MUAPs decomposition and SVM-based analysis in the classification of neuromuscular disorders. Comput Methods Programs Biomed. 2012;107(3):393–403.

Englehart K, Hudgins B. A robust, real-time control scheme for multifunction myoelectric control. Biomed Eng IEEE 2003;50(7):848–854.

Evgeniou T, Pontil M, Poggio T. Statistical Learning Theory : A Primer. Int J Comput Vis. 2000;38(1):9–13.

Farina D, Merletti R, Enoka RM. The extraction of neural strategies from the surface EMG. J Appl Physiol. 2004;96(4):1486–95.

Fligge N, Urbanek H, Smagt P van der. Relation between object properties and EMG during reaching to grasp. J Electromyogr Kinesiol. 2013; 23(2): 402–410.

Fontana JM, Chiu a. WL. Control of Prosthetic Device Using Support Vector Machine Signal Classification Technique. Am J Biomed Sci. 2009;1(4):336–343..

Futamata M, Nagata K, Magatani K. The evaluation of the discriminant ability of multiclass SVM in a study of hand motion recognition by using SEMG. In: Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Vol 2012.; 2012:5246–9.

Goker I, Osman O, Ozekes S, Baslo MB, Ertas M, Ulgen Y. Classification of juvenile myoclonic epilepsy data acquired through scanning electromyography with machine learning algorithms. J Med Syst. 2012;36(5):2705–2711.

Güler NF, Koçer S. Use of Support Vector Machines and Neural Network in Diagnosis of Neuromuscular Disorders. J Med Syst. 2005;29(3):271–284.

Hofmann T, Schölkopf B, Smola AJ. Kernel methods in machine learning. Ann Stat. 2008;36(3):1171–1220.

Hsu C, Chang C, Lin C. A practical guide to support vector classification. 2003-2010; Retrieved from https://www.cs.sfu.ca/people/Faculty/teaching/726/spring11/svmguide.pdf

Huang H, Liu H, Li N, et al. An anthropomorphic controlled hand prosthesis system. J Zhejiang Univ Sci C. 2012;13(10):769–780.

Huang, H., Zhang, F., Hargrove, L. J., Dou, Z., Rogers, D. R., Englehart KB. Continuous Locomotion-Mode Identification for Prosthetic Legs Based on Neuromuscular–Mechanical Fusion. IEEE Trans Biomed Eng. 2011;58(10):2867–2875.

Hudgins B, Parker P, Scott RN. A new strategy for multifunction myoelectric control. IEEE Trans Biomed Eng. 1993;40(1):82–94.

Katsis CD, Goletsis Y, Likas a, Fotiadis DI, Sarmas I. A novel method for automated EMG decomposition and MUAP classification. Artif Intell Med. 2006;37(1):55–64.

Kaur G, Arora A, Jain V. Multi-class support vector machine classifier in EMG diagnosis. WSEAS Trans Signal Process. 2009;5(12):379–389.

Khokhar ZO, Xiao ZG, Menon C. Surface EMG pattern recognition for real-time control of a wrist exoskeleton. Biomed Eng Online. 2010;9:41:1–17.

Khushaba RN, Kodagoda S, Takruri M, Dissanayake G. Expert Systems with Applications Toward improved control of prosthetic fingers using surface electromyogram (EMG) signals. Expert Syst Appl. 2012;39(12):10731–10738.

Kumar DK, Poosapadi Arjunan S, Singh VP. Towards identification of finger flexions using single channel surface electromyography--able bodied and amputee subjects. J Neuroeng Rehabil. 2013;10(50):1–7.

Lucas M-F, Gaufriau A, Pascual S, Doncarli C, Farina D. Multi-channel surface EMG classification using support vector machines and signal-based wavelet optimization. Biomed Signal Process Control. 2008;3(2):169–174.

Luxburg U Von, Schölkopf B. Statistical Learning Theory: Models, concepts, and results. In: Gabbay DM, Hartmann S, Woods J, eds. Handbook of the History of Logic.Vol 10. Elsevier BV; 2009:651–706.

Marwan N, Carmenromano M, Thiel M, Kurths J. Recurrence plots for the analysis of complex systems. Phys Rep. 2007;438(5-6):237–329.

Marwan N, Kurths J. Cross Recurrence Plots and their Applications. In: Benton CV, ed. Mathematical Physics Research at the Cutting Edge. Nova Science Publisher, Inc.; 2004:101–139.

Marwan N, Thiel M, Nowaczyk NR. Nonlinear Processes in Geophysics Cross recurrence plot based synchronization of time series. Nonlinear Process Geophys 2002; 2(9):325–331..

McGill S, Juker D, Kropf P. Appropriately placed surface EMG electrodes reflect deep muscle activity (psoas, quadratus lumborum, abdominal wall) in the lumbar spine. J Biomech Biomech. 1996;29(11):1503–1507.

Merletti R, Farina D. Surface EMG processing: Introduction to the special issue. Biomed Signal Process Control. 2008;3(2):115–117.

Müller K-R, Mika S, Rätsch G, Tsuda K, Schölkopf B. An Introduction to Kernel-Based Learning Algorithms. IEEE Trans Neural Networks. 2001;12(2):181–201.

Natarajan GS, Wininger M, Kim NH, Craelius W. Relating biceps EMG to elbow kinematics during self-paced arm flexions. Med Eng Phys. 2012;34(5):617–624.

Orosco EC, Lopez NM, di Sciascio F. Bispectrum-based features classification for myoelectric control. Biomed Signal Process Control. 2013;8(2):153–168.

Oskoei AM, Hu H. Myoelectric control systems—A survey. Biomed Signal Process Control. 2007;2(4):275–294.

Oskoei MA, Hu H. Support Vector Machine-Based Classification Scheme for Myoelectric Control Applied to Upper Limb. IEEE Trans Biomed Eng. 2008;55(8):1956–1965.

Park SH, Lee SP. EMG pattern recognition based on artificial intelligence techniques. IEEE Trans Rehabil Eng. 1998;6(4):400–5.

Phinyomark A, Limsakul C, Phukpattaranont P. A novel feature extraction for robust EMG pattern recognition. J Comput. 2009;1(1):2151–9617.

Phinyomark A, Phukpattaranont P, Limsakul C. Feature reduction and selection for EMG signal classification. Expert Syst Appl. 2012;39(8):7420–7431.

Phinyomark a., Nuidod a., Phukpattaranont P, Limsakul C. Feature Extraction and Reduction of Wavelet Transform Coefficients for EMG Pattern Classification. Electron Electr Eng. 2012;122(6).

Scheme E, Englehart K. Electromyogram pattern recognition for control of powered upper-limb prostheses: State of the art and challenges for clinical use. J Rehabil Res Dev. 2011;48(6):643–659.

Scholkopf B, Smola A, Muller KR. Nonlinear Component Analysis as a Kernel Eigenvalue Problem. Tübingen; 1996:1–18.

Schölkopf B, Smola A. Learning with kernels. Massachustts: Massachusetts Institute of Technology; 2002.

Schölkopf B, Smola AJ, Williamson RC, Bartlett PL. New Support Vector Algorithms. Neural Comput. 2000;12:1207–1245.

Shi J, Guo J-Y, Hu S-X, Zheng Y-P. Recognition of finger flexion motion from ultrasound image: a feasibility study. Ultrasound Med Biol. 2012;38(10):1695–704.

Shi J, Zheng Y, Chen X, Xie H. Modeling the relationship between wrist angle and muscle thickness during wrist flexion-extension based on the bone-muscle lever system: a comparison study. Med Eng Phys. 2009;31(10):1255–60.

Stirling LM, von Tscharner V, Kugler PF, Nigg BM. Classification of muscle activity based on effort level during constant pace running. J Electromyogr Kinesiol. 2011;21(4):566–571.

Subasi A. Classification of EMG signals using PSO optimized SVM for diagnosis of neuromuscular disorders. Comput Biol Med. 2013;43(5):576–86.

Subasi A. Medical decision support system for diagnosis of neuromuscular disorders using DWT and fuzzy support vector machines. Comput Biol Med. 2012;42(8):806–15.

Sultornsanee S, Zeid I, Kamarthi S. Classification of Electromyogram Using Recurrence Quantification Analysis. Procedia Comput Sci. 2011;6:375–380.

Tavakolan M, Xiao ZG, Menon C. A preliminary investigation assessing the viability of classifying hand postures in seniors. Biomed Eng Online. 2011;10(1):79.

Tkach D, Huang H, Kuiken T. Study of stability of time-domain features for electromyographic pattern recognition. J Neuroeng Rehabil. 2010;7:21:1–13.

Tolambiya A, Thomas E, Chiovetto E, Berret B, Pozzo T. An ensemble analysis of electromyographic activity during whole body pointing with the use of support vector machines. PLoS One. 2011;6(7):e20732.

Vaisman L, Zariffa J, Popovic MR. Application of singular spectrum-based change-point analysis to EMG-onset detection. J Electromyogr Kinesiol. 2010;20(4):750–760.

Smagt P van der, Grebenstein M, Urbanek H, et al. Robotics of human movements. J Physiol Paris. 2009;103(3-5):119–32.

Vapnik VN. An Overview of Statistical Learning Theory. IEEE Trans Neural Networks. 1999;10(5):988–999.

Xie H-B, Wang Z-Z, Huang H, Qing C. Support vector machine in computer aided clinical electromyography. In: Machine Learning and Cybernetics, 2003 International Conference on (Volume 2), 2-5 Nov.; 2003:1106–1108.

Yan Z, Wang Z, Xie H. The application of mutual information-based feature selection and fuzzy LS-SVM-based classifier in motion classification. Comput Methods Programs Biomed. 2008;90:275–284.

Yang D, Zhao J, Gu Y, et al. An Anthropomorphic Robot Hand Developed Based on Underactuated Mechanism and Controlled by EMG Signals. J Bionic Eng. 2009;6(3):255–263.

Zbilut JP, Webber Jr CL. Recurrence Quantification Analysis. Wiley Encycl Biomed Eng. 2006;(2):1–9.

Zecca M, Micera S. Control of multifunctional prosthetic hands by processing the electromyographic signal. Crit Rev 2002;30(4-6):459–485.

Ziai A, Menon C. Comparison of regression models for estimation of isometric wrist joint torques using surface electromyography. J Neuroeng Rehabil. 2011;8(1):56.

CAPÍTULO VI

6. Machine Learning Classification of EMG Temporal Parameters: Support Vector Machines on Onset.

Submetido no ***Journal of Medical Systems***

Silva, L., Vaz, J. V., Serranho, P., Marta, S., Cabri, J., Pezarat-Correia, P. Machine Learning Classification of EMG Temporal Parameters: Support Vector Machines on Onset. *Journal of Medical Systems.*

Machine Learning Classification of EMG Temporal Parameters: Support Vector Machines on Onset.

Silva, Luís; Faculdade de Motricidade Humana, Universidade de Lisboa
Vaz, João Rocha; Faculdade de Motricidade Humana, Universidade de Lisboa
Serranho, Pedro; Departamento de Ciências e Tecnologia, Universidade Aberta
Marta, Sérgio; Faculdade de Motricidade Humana, Universidade de Lisboa
Cabri, Jan; Norwegian School of Sport Sciences
Pezarat-Correia, Pedro; Faculty of Human Kinetics, Technical University of Lisbon

Abstract

Background: Several algorithms have been proposed in literature for onset detection but results reproducibility are compromised by a lack of physiologic discussion. This article introduces machine learning as a temporal detector of onset in EMG signal. **Objective**: To analyze the performance of time-domain features by building Support Vector Machines (SVM) models to classify onset on EMG. **Method**: Surface EMG was collected from 12 muscles (trunk and lower limb) of 12 golfers, 6 with low handicap (Hc = 1.4±2.5<5) and 6 with high handicap (Hc = 24.6±4.2>18) during the golf swing. The signal was segmented by overlapped segmentation, and the time-domain features were extracted from each window to characterize each time instant. The classification was considered with three sets of 2, 4 and 6 features (F2, F4 and F6) to build the model. After conducting grid-search, the best parameters for the radial basis function kernel SVM classifier were selected for each muscle and each set of features, establishing therefore the best classification model. **Results**: The high, low, and total Hc had an accuracy of 90.3±4%, 90.8±4.9% and 89.4±3.7% for F2, 94.9±2.5%, 95.0±3.3%, 93.5%±3.2% for F4, and 95.2±2.4% 95.1±3.2% and 93.6±3.3% to F6, respectively. **Conclusions**: The radial basis function parameters, accuracy, and the number of support vectors tends to be similar between F4 and F6, differing from F2. Four features ensure an accuracy greater than 90%. This study shows that time-domain features of EMG can serve as basis in building models for onset detection, though they could be improved by introducing other features type.

Keywords: Electromyography, SVM, time-domain features, automatic learning, Onset

6.1. Introduction

Electromyography (EMG) is a useful technique to characterize neuromuscular activation. The most common approaches analyze the EMG structure in order to quantify timing, amplitude, and frequency parameters. The main EMG timing parameters are onset, time of maximum peak, and offset. Onset is the most studied. The relevance of identifying time parameters, such as onset, provide information regarding the temporal organization and coordination of muscle related to a given motor behavior [1].

Several detection methods have been proposed for onset determination, but there is a lack of a standardized methodology [2], and do not allow the analysis of other timing characteristics that might have physiological significance. By definition, the onset is the first recorded activity of motor unit action potentials [3]. Despite the relevance of this concept, some authors have shown concern with descriptive information about the task-related physiological phenomena within EMG signal timing [4-6]. Silva et al. [5] introduced the concept of onset peak to discriminate the most relevant muscle activation pattern within and between subjects' trials, beyond the onset burst, during golf swing. McGill et al. [6] described the phenomenon of the double peak and contract-relax-contract pulsing cycle in martial arts to investigate how speed strength is accomplished. Tyler and Karst [4] analyzed time muscle activity of non-focal muscles for reaching tasks performed while standing, and with target distances variations. Nonetheless, the main studies about onset only focus on how to determinate the start of muscle activation and on the precision of the different methods under the influence of the stochastic nature of the EMG signal.

The approaches for onset detection can be categorized into two major dimensions: visual inspection (VI) and computerized detection algorithms [5]. The first were used to establish the fidelity of the onset detection computerized algorithms, but is a time-consuming job and its results depends on the researcher's observation, so subjective [7]. The most commonly used detection techniques are threshold algorithms depending on standard deviation [7-9] or confidence interval [10], and maximum likelihood or approximated generalized likelihood-ratio detectors [11]. Both threshold and likelihood methods are influenced by the background muscle activity [5], [8], [12]. The average latency artifacts near the period of muscle inactivity and activity ratio were 8 ms for approximated generalized likelihood-ratio detectors and 46 ms for threshold standard deviation methods [12]. As the fidelity of information taken from the signal is affected by the signal to noise ratio and distortion of the signal, several improved methods were also proposed, such as Teager–Kaiser energy operator [3], and wavelet transformation [13]. These aspects lead to the following issues: (1) the performance of computerized algorithms are accessed by VI; (2) usually information is collected only for the initiation of activation; (3) threshold algorithms use a sliding window to avoid bias on detection due to similarities; (4) results have a great dependency on the algorithm threshold parameters; (5) approximated generalized likelihood-ratio detectors are dependent of the likelihood ratio test and search for the maximum in test function. We propose a setting where the computerized algorithm computes characteristics of interest from a time window in the neighborhood of a central instant and then classify each instant by class of activity: $y = -1$(rest) or $y = 1$(activity). Then, a transition from -1 to 1 would represent the onset time. This idea carries on over to theory of machine learning, namely supervising learning. While in unsupervised learning there are no a priori labeled examples, the supervised learning process starts from a labeled training set [14-17] .

Support Vector Machines (SVM) are optimization algorithms based on supervised learning presented in 1992 as a *training algorithm that maximizes the margin between the training patterns and the decision boundary* [18]. After several extensions where considered as soft margins [19], the non-linear "kernel trick" through the Kernel functions so that $\left(\Phi(\mathbf{x}) \cdot \Phi(\mathbf{y})\right) =: k(\mathrm{x}{\cdot}y)$ [20], and extensions of a class of algorithms both for classification and for regression as the υ-SVM [21], and multiple kernel learning [22]. SVM has revealed to be a powerful tool in the classification of EMG patterns in myoelectric control [22–28], neuromuscular disorders [29–32], and on kinematic analysis [33–35]. The onset detection is not considered as a myoelectric control pattern recognition problem, such as proportional control, threshold control, and finite state machines [36], being a non-pattern recognition-based myoelectric control category. However, there are results in the literature that lead to the possibility of considering onset detection as a pattern recognition problem [37].

To the best of our knowledge, the only study about onset time and SVM-based classification [38] did not focus on EMG but in the onset time of the movement. Authors used kinematic variables like force and torque, isometric voluntary contraction, and the features were directly extracted from raw data. Applying pattern recognition methods to detect a point in time gives rise to two concerns. First, if the classification problem performance decreases when transient periods are included [24], [39]. In myoelectric control usually the nearest points are removed, but in onset detection this may create problems due to the loss of information. Second, we need to consider the structure of features that will characterize each instant. As stated by Hudgins, Parker, and Scott (1993, p. 86) [40], the success of pattern classification system depends almost entirely on the choice of features used to represent the continuous time waveforms.

This implies that the ideal is the computer to learn what the researcher wants to study and then adjust the algorithm to that classify phenomenon. So, the aim of this article is to propose the Support Vector Machine (SVM), a machine learning approach to solve the problem of onset detection in EMG signal, starting with time-domain features.

6.2. Method

6.2.1. Participants and task

Twelve right handed golfers were divided in two groups by handicap (Hc): six with low handicap $Hc < 5$ (1.4 ± 2.5) and six with high handicap $Hc > 18$ (24.6 ± 4.2), based on the Handicap System of European Golf Association [41]. The golfers were instructed to perform four long range swings with seven iron (>150 m) carried out on top of an artificial grass golf carpet with high absorption features. As per trial were considered 1001 instants after the features extraction, thus totaling 24024 instants to analysis (6 subjects x 1001 instants per trial x 4 trials) each Hc and 48048 both low and high Hc. Subjects were allowed to perform some repetitions to a better adaptation to the task and to warm up. Participants did not have any limitations for playing golf. All the procedures were explained and a consent form was signed. This study was approved by the Ethics Committee of the Faculdade de Motricidade Humana (Universidade de Lisboa).

6.2.2. EMG procedures

After preparing the skin properly by means of hair removal, abrasion and alcohol cleaning, the electrodes were placed with a 22 mm center-to-center distance and applied parallel to the muscle fibers in muscles from both body sides [42]: rectus femoris (RF) - At 50% on the line from the anterior spina iliaca superior to the superior part of the patella;

vastus medialis (VM) - At 80% on the line between the anterior spina iliaca superior and the joint space in front of the anterior border of the medial ligament; vastus lateralis (VL) - At 2/3 on the line from the anterior spina iliaca superior to the lateral side of the patella; biceps femoris (BF) - At 50% on the line between the ischial tuberosity and the lateral epicondyle of the tibia; semitendinosus (ST) - At 50% on the line between the ischial tuberosity and the medial epicondyle of the tibia; gluteus maximus (GM) - At 50% on the line between the sacral vertebrae and the greater trochanter just in the middle of the buttocks well above the visible bulge of the greater trochanter; external oblique (EO) - 15 cm laterally from the umbilicus; erector spinae (ES), 3 cm laterally from the L3 spinous process. Previously to electrodes placement, muscle contraction was performed in order to visualize the muscle belly. The ground electrode was placed on the manubrium.

EMG data was collected with bioPLUX® research 2010 telemetric equipment (Plux, Lisbon, Portugal) using active surface electrodes (Al/AgCl, disk shape 10 mm of diameter) and surfaces of detection AMBU® BlueSensor N (shape 30 x 22 AMBU, Ballerup, Denmark). Signal was amplified with a bandpass between 10 and 500 Hz, common-mode rejection ratio (CMRR) of 110 dB and input impedance greater than 100 MΩ. Data were collected with a sampling frequency of 1000Hz. Data was digitally filtered (10–490 Hz), full-wave rectified, and smoothed with a low pass filter (12 Hz, Butterworth 4th order digital filter). All data was submitted to visual inspection comparison.

6.2.3. Video data recording, processing and kinematic analysis

Three high speed Basler A602fc cameras (Basler Vision Technologies, Ahrensburg, Germany) at 100 Hz were placed in anterior, posterior and superior oblique positions to determine swing phases. A fourth Casio Ex-FH20 camera (Casio, Tokyo, Japan) at 1000 Hz was placed in front of the ball to determine the instant of impact. Two reflective tapes were placed on the club to divide the swing in three phases [43], [44]: Backswing – from the beginning until the top of the swing; Downswing – from the top until impact; and Follow-Through – from impact until the end of the swing. SIMI 3D Motion system (SIMI Reality Motion System GmbH, Unterschleissheim, Germany) was used for EMG-synchronized 3D kinematic analysis.

6.2.4. Segmentation and features extraction

The segmentation is a process which continuously collects windows of signal containing a number of samples. In this study we choose to consider windows of 200 ms with an increment of 195 ms (lag of 5 ms) by overlapped window at a sample rate of 1000 Hz [45]. This choice meant to avoid losing signal information over time. A smaller segment length could lead to a loss of information, and a longer length requires a high computational load. It is expected that each window sample corresponds to the information of its central point and neighbor's variability.

After segmentation of the signal, six types of time-domain features were extracted [46], [47]: mean absolute value (MAV), waveform length (WL), standard deviation difference absolute value (DASDV), variance of EMG (VAR), integrated EMG (IEMG), and logarithmic detector (LOG).

⇒ Mean absolute value (MAV): assuming a set of samples $\{x_1, x_2, \dots x_N\}$ within the given window, where N corresponds to the last observation, the MAV is given by:

$$MAV = \frac{1}{N}\sum_{i=1}^{N}|x_i| \tag{1}$$

⇒ Waveform length (WL): corresponds to the cumulative length of the waveform over a given window. The WL is related to the amplitude of the waveform regarding to frequency and time.

$$WL = \sum_{i=1}^{N-1}|x_{i+1} - x_i| \tag{2}$$

⇒ Standard deviation difference absolute value (DASDV): is the value of the standard deviation of the wavelength.

$$DASDV = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N-1}|x_{i+1} - x_i|^2} \quad (3)$$

⇒ Variance of EMG (VAR): usually is defined as the average value of the square of the standard deviation of the signal. Given that the average EMG signal tends to zero, can be calculated using the sum of the squares of the values of the signal divided by the size of the interval minus 1.

$$VAR = \frac{1}{N-1}\sum_{i=1}^{N} x_i^2 \quad (4)$$

⇒ Integrated EMG (IEMG): is defined as the area that lies under the curve of the rectified EMG signal (absolute values), i.e. is the integral of the absolute value of the raw EMG signal. Through the absolute value, the noise will cause that the integral has a constant increase within a time window, as it consists in determining the area bounded by the curve of the rectified signal.

$$EMG_I = \sum_{i=1}^{N}|x_i| \quad (5)$$

⇒ Logarithmic detector (LOG): is a non-linear detector that gives an estimate of muscle contraction force based on logarithm.

$$Log = e^{\frac{1}{N}\sum_{i=1}^{N}\log(|x_i|)} \quad (6)$$

The clusters formation in two (F2), four (F4) and six (F6) groups of features was based on the weighting by Fisher Score [48] and Correlation-based Feature Selection algorithm [49].

The Fisher Score (FS) scores each feature with respect to its importance to the classifier, having in account the given labels. Therefore, given a label vector $y = \{y_1, y_2, \dots, y_c\}$ with *c* classes, the Fischer score for each feature *i* is defined by

$$FS(f_i) = \frac{\sum_{j=1}^{c} n_j (\mu_{i,j} - \mu_i)^2}{\sum_{j=1}^{c} n_j \sigma_{i,j}^2} \tag{7}$$

where μ_i is the average of the feature f_i, n_j is the number of samples in the jth class, $\mu_{i,j}$ and $\sigma_{i,j}$ correspond to the mean and the variance of f_i in class c, respectively [50].

The Correlation-based Feature Selection algorithm (CFS) is another possibility for feature selection. The CFS considers that the selection of features for automatic learning should be taking into account the correlation between the features. If the correlation between each of the components of a test for a certain rating variable is known, the inter-relationship between each pair of components for classification can be estimated using:

$$M_s = \frac{k\overline{r_{cf}}}{\sqrt{k + k(k-1)r_{\overline{ff}}}} \tag{8}$$

where M_s is the heuristic merit of a set of features S, depending on the correlation between the total of the components and the rating variable, and k is the number of components. $\overline{r_{cf}}$ is the average of the correlations between the components related to the rating variable, and $r_{\overline{ff}}$ corresponds to the inter-correlation mean between components. The numerator provides information about how much a certain set of features have discriminate power, the denominator characterize the feature redundancy [49].

6.2.5. Support vector machines

SVM were introduced by Vapnik and Chervonenkis in the late 60's early 70's having a boost in the early 90's [18], [51], namely, with the introduction of the soft margin classifier [19] for non-separable data. The binary classification was the first SVM application, considering therefore only two classes. The aim is to find the best hyperplane which separates the data by class in the d-dimensional feature space. For that purpose a training set is considered, where the label from the label set $\mathcal{Y}$ is known for each element of the input space $\mathcal{X}$ within the feature space. This is the broad idea of classification problems using a training independent identically distributed (*i.i.d.*) data (input and label) pairs according to an unknown probability distribution $P(\mathbf{x}, y)$. Given l training examples $(\mathbf{x}_i, y_i)$ with $i = 1, \ldots, n$ where the inputs $\mathbf{x}_i \in \mathbb{R}^{\boldsymbol{d}}$ **(**d *– dimensional***)** and the label $y_i \in \{-1,1\}$, the hyperplanes which separate the two classes in $\mathbb{R}^{\boldsymbol{d}}$ are defined by a vector $\mathbf{w}$ (orthogonal to the hyperplane) and a bias b as:

$$\mathbf{w} \cdot \mathbf{x} + b = 0 \tag{9}$$

where $\mathbf{w}$ corresponds to the norm vector to the hyperplane, $|b|/\|\mathbf{w}\|$ is the perpendicular distance from the hyperplane to the origin, $\|\mathbf{w}\|$ is the euclidean norm of $\mathbf{w}$, and $\mathbf{w} \cdot \mathbf{x}$ is the dot product between vectors $\mathbf{w}$ and $\mathbf{x}$. In the case of linearly separable data, the algorithm of support vector chooses the separating hyperplane with the largest margin, satisfying:

$$\begin{array}{lll} \mathbf{w} \cdot \mathbf{x}_i + b \geq +1 & when & y_i = +1 \\ \mathbf{w} \cdot \mathbf{x}_i + b \leq -1 & when & y_i = -1 \end{array} \tag{10}$$

The optimal hyperplane is the one that derives from those inequalities where the norm $\|\mathbf{w}\|$ is minimal, ie, an optimization problem with the conditions:

$$\min_{w,b} \frac{1}{2} \|\mathbf{w}\|^2 \tag{11}$$

subject to the constrains

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) - 1 \geq 0 \tag{12}$$

which in turn, can be converted to a quadratic optimization problem by introducing positive Lagrange multipliers that lead to the Dual Formulation:

$$\mathcal{L}(\mathbf{w}, b, \alpha) = \min_{\alpha} \ \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j (\mathbf{x}_i \cdot \mathbf{x}_j) - \sum_{i=1}^{n} \alpha_i \tag{13}$$

subject to

$$\begin{cases} \alpha_i \geq 0, \qquad \forall i = 1, \dots, n \\ \sum_{i=1}^{n} \alpha_i y_i = 0 \\ \mathbf{w} = \sum_{i=1}^{n} \alpha_i y_i \mathbf{x}_i \end{cases} \tag{14}$$

To extend this algorithm to non-separable data it becomes necessary to allocate some slack restrictions. This is achieved by the introduction of a cost parameter C and non-negative variables $\xi_i \geq 0, i = 1, \dots, n$ examples. The cost function is now added to the objective function.

$$\min_{w,b} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^{n} \xi_i^{\sigma} \tag{15}$$

with the constrains $0 \leq \alpha_i \leq C$. Therefore, the dual formulation becomes

$$\mathcal{L}(\mathbf{w}, b, \alpha) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{n} \xi_i - \sum_{i} \alpha_i (y_i (\mathbf{w} \cdot \mathbf{x}_i + b) - 1 + \xi_i) - \sum_{i=1}^{n} \eta_i \xi_i \tag{16}$$

where $\alpha_i, \eta_i \geq 0 \ \forall_i \in n$ because α_i, $\eta_i$are Lagrange multipliers introduced to enforce positivity of ξ_i. The non-linear learning problem takes place at $\mathcal{H}$ space with the example sample $(\Phi(\mathbf{x}_1), y_1), \Phi(\mathbf{x}_2), y_2) \dots, \Phi(\mathbf{x}_n), y_n) \in \mathcal{H} \times \mathcal{Y}$, where Φ represents the map between the original feature space $\mathcal{X}$ and the new (higher dimensional) feature space $\mathcal{H}$. In this way, one can show that under some assumptions on K, the the inner product $\mathbf{x}_i \cdot \mathbf{x}_j$ on $\mathcal{H}$ is equal to a kernel function $K(\mathbf{x}_i, \mathbf{x}_j)$. Therefore, with a proper choice of K, one can have a nonlinear separator in $\mathcal{X}$ that relies on the same theory for the linear separator in a space $\mathcal{H}$ that does not need to be constructed. The problem is then formulated as:

$$\max_{\alpha} \ \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j \left(\Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j) \right) \tag{17}$$

subject to

$$\begin{cases} \alpha_i \geq 0, \qquad \forall i = 1, \dots, n \\ \sum_{i=1}^{n} \alpha_i y_i = 0 \\ \mathbf{w} = \sum_{i=1}^{n} \alpha_i y_i \mathbf{x}_i \end{cases} \tag{18}$$

and the classifier is defined by the signal of

$$f(x) = \sum \alpha_i y_i K(x, x_i) \tag{19}$$

A complete SVM tutorials can be found in the literature [20], [52–54] for an extensive explanation.

6.2.6. SVM validation

A 5-fold cross-validation was used to determine the generalization error of each model, and two grid-search steps to find the best cost functional constant (C) and best gamma of the Kernel function (γ):

$$K(\mathbf{x}_i, \mathbf{x}_j) = exp\left(-\gamma \left\|\mathbf{x}_i - \mathbf{x}_j\right\|^2\right), \gamma > 0. \tag{20}$$

The former grid-search is a coarse grid to find the region of parameters with higher classification accuracy. The second is a finer one on the neighborhood of best parameters of the previous step. Using an exponential growing sequences of C and γ parameters in the cross-validation, various pairs C and γ are considered to look for the best accuracy [55]. Three experiments were concluded to outwit the skill bias into the models and to avoid time-consuming due to the computational cost:

1. Extensive research: $C = 2^{-2:1.25:8}; \gamma = 2^{-7:1.25:3}$.
2. Refined search: depending on the values obtained at the best value determined on extensive research was used with spacing $h = 0.25$ both to the left and right side of those best C and γ values.

Figure 1 shows the flow chart of the research design after EMG collected.

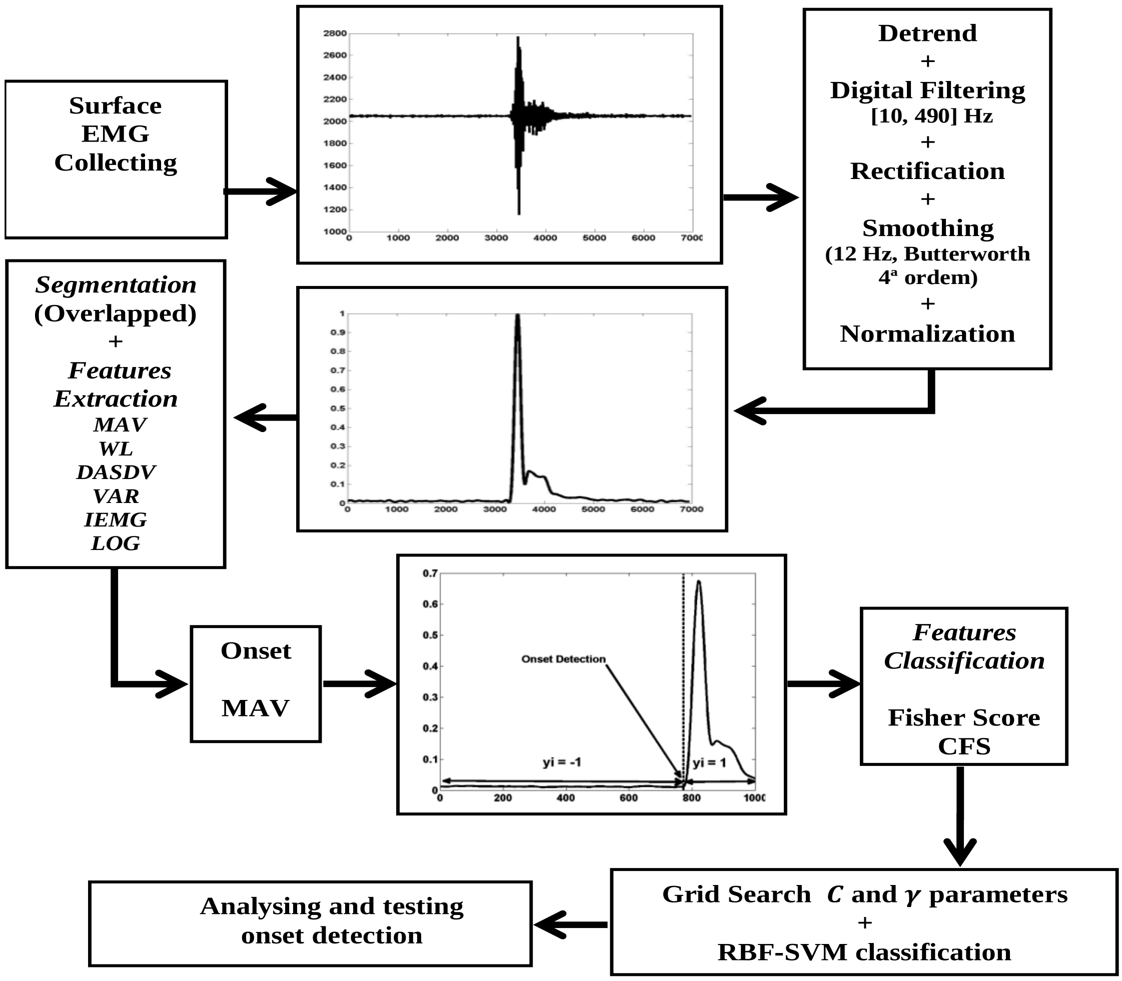


*Figura 9 [**Fig. 1**. – Flow chart EMG research design for EMG processing].*

Mean absolute value (MAV), waveform length (WL), standard deviation difference absolute value DASDV), variance of EMG (VAR), integrated EMG (IEMG), and logarithmic detector (LOG), correlation-based feature selection algorithm (CFS).

6.2.7. Complementary statistics

We also carried out statistical inference, besides the descriptive presentation, features extraction, and SVM procedures. The normality assumptions were confirmed using the Shapiro – Wilk test. Nonparametric MANOVA was used to verify differences between the Hc groups into parameters C and γ that are the basis to construct the SVM models. For this test a rank transformation was performed. The χ^2 test statistics for nonparametric MANOVA was calculated using the Pillai's Trace, and the *p-value* was corrected. Friedman ANOVA was applied to non-separable cost parameters C and Kernel γ to verify differences between features set (F2, F4 and F6). Multiple comparisons were performed.

Mixed ANOVA was used to analyze repeated measures feature sets, and differences between Hc groups (high, low and both) about SVM models accuracy and relative number of support vectors. The homogeneity of variances was tested by Levene's test. Multiple comparisons were performed by Bonferroni test for repeated measures, and the Turkey test for independent samples. The sphericity was checked with Mauchly's test and when not confirmed, the degrees of freedom were adjusted by the Greenhouse-Geisser Epsilon.

The correlation between the classification accuracy and the relative number of support vectors in each model was performed using the Pearson correlation coefficient.

Complementary Statistics were performed on the software IBM - SPSS 19.0 (IBM Corporation, New York, USA). The level of significance was 5 %, being corrected by the number of ANOVA when necessary.

6.3. Results

6.3.1. Features and SVM parameters

The features considered relevant for classification by CFS were the MAV, WL, DASDV and VAR, while the remaining features were redundant. This result is consistent for all muscles. The FS reinforced these results, with few changes in the order of WL and DASDV (as the second or third most relevant) in some muscles. Thus, three sets of features that served as characteristic vectors of SVM have been chosen:

F2: [MAV, WL], most relevant group of two features;

F4: [MAV, WL, DASDV, VAR], most relevant group of four features, being VAR the least one;

F6: [MAV, WL, DASDV, VAR, IEMG, LOG], a group with all studied features.

Figure 2 displays the boxplot for C and γ parameter. Nonparametric MANOVA showed no significant differences for the three groups of Hc both C and γ parameters each features set ($\chi^2(12) = 16.415; N = 36; p = 0.17$). Friedman ANOVA showed significant differences between feature sets for parameter ($\chi^2(2) = 42.318; p < 0.001, N = 36$). Multiple comparisons show that these differences are significant from the F2 set to both F4 and F6 ($p < 0.001$). Models with the sets F4 and F6 do not differ ($p = 1.0$). Similarly to the cost parameter C, significant differences were also found for the kernel parameter γ ($\chi^2(2) = 19.528; p < 0.001, N = 36$), and again with F2 models responsible for these differences in relation to both F4 ($p = 0.002$) and F6 ($p = 0.001$). Also, there is similarity between F4 and F6 sets ($p = 1.0$).

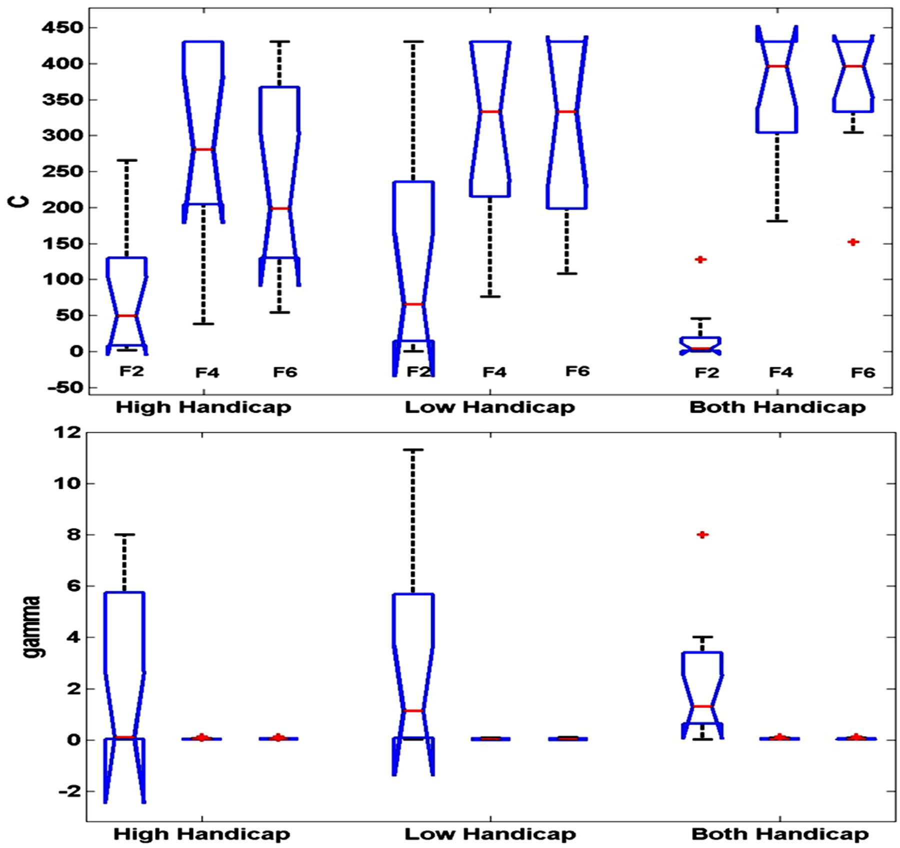


*Figura 10 [**Fig. 2**. Boxplot for cost parameter C and γ by handicap group].*

6.3.2. SVM accuracy

Considering all handicap groups and the onset as the most relevant activation (burst or peak), the algorithm shows an accuracy of 90.3±4%, 90.8±4.9% and 89.4±3.7% using F2 for the groups of high, low and all handicaps, respectively. When F4 and F6 are considered, the mean accuracy rises by about 5%. For F4, the mean classification accuracy was 94.9±2.5% for high handicap and 95.0±3.3% for low handicap. For a higher sample size considering the union of the two groups, SVM models had an accuracy of 93.5%±3.2%. When increasing the number of features by considering F6, results were similar with 95.2±2.4%, 95.1±3.2% and 93.6±3.3% for high, low and all handicap groups. Figure 3 shows an example of grid search results.

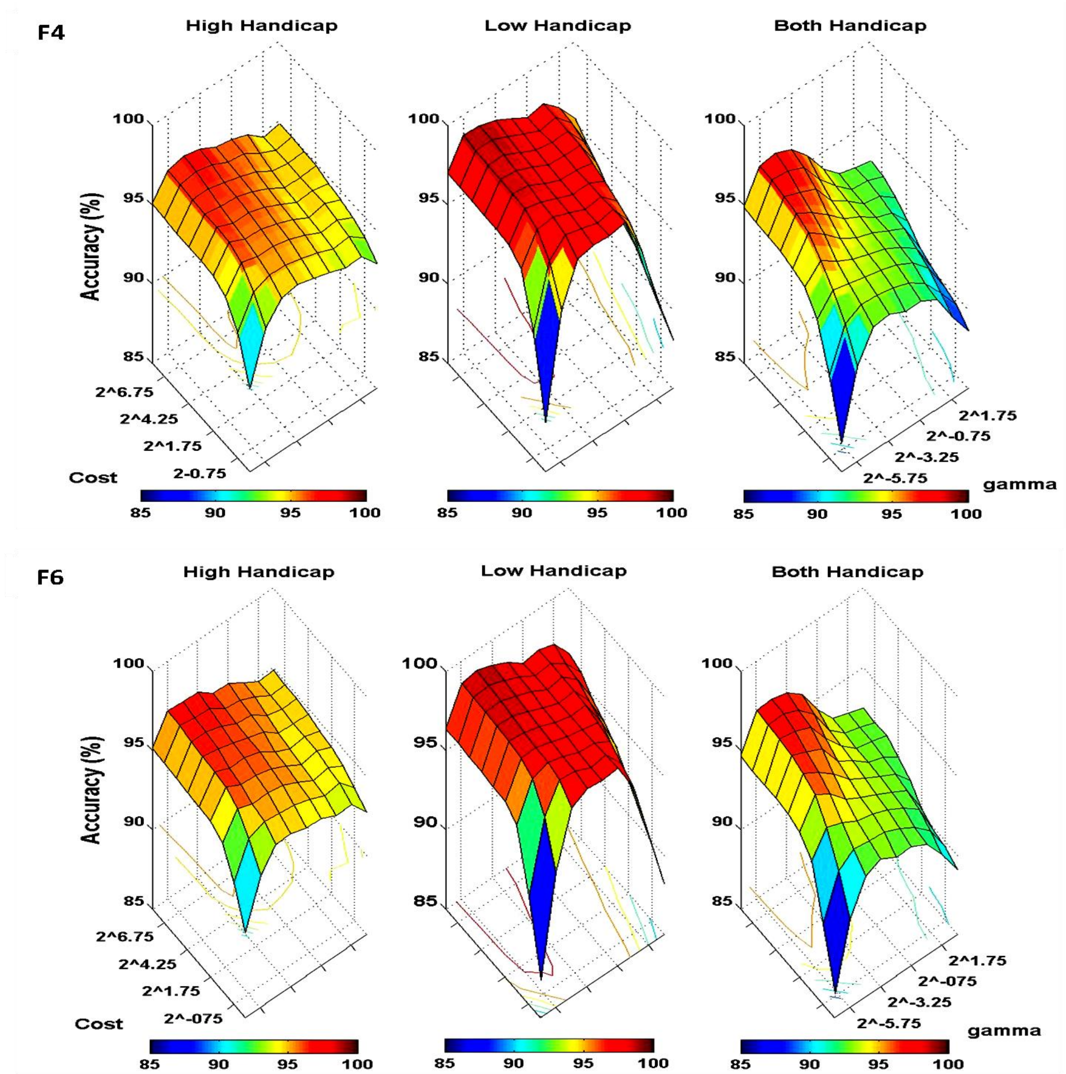


*Figura 11 [**Fig. 3**. C e γ Grid-search for left rectus femoris, F4 and F6 by handicap group].*

Mixed ANOVA showed significant differences in accuracy between the sets of features used for classification ($F_{(1.030,33.984)} = 213.328; p < 0.001; \eta^2 = 0.866;\ \pi = 1.0$) in all pairs. However, no significant differences were found between F4 and F6 ($p = 0.11$), but when compared to F2 significant differences ($p < 0.001$) are found. The interaction with the defined groups of handicap had no significant effect ($F_{(2.060,33.984)} = 0.435; p = 0.657;\ \eta^2 = 0.026;\ \pi = 0.116$), and also, there were no significant differences between groups ($F_{(2,33)} = 0.700; p = 0.504;\ \eta^2 = 0.041;\ \pi = 0.158$).

6.3.3. Relative number of support vectors

Figures 4 and 5 show boxplot and scatter, respectively, with the percentages rate of classification and number of support vectors. In tendency, the higher the accuracy the smaller number of support vectors. The percentage of support vectors required for the different classification models shows similar results comparing with those obtained previously for accuracy.

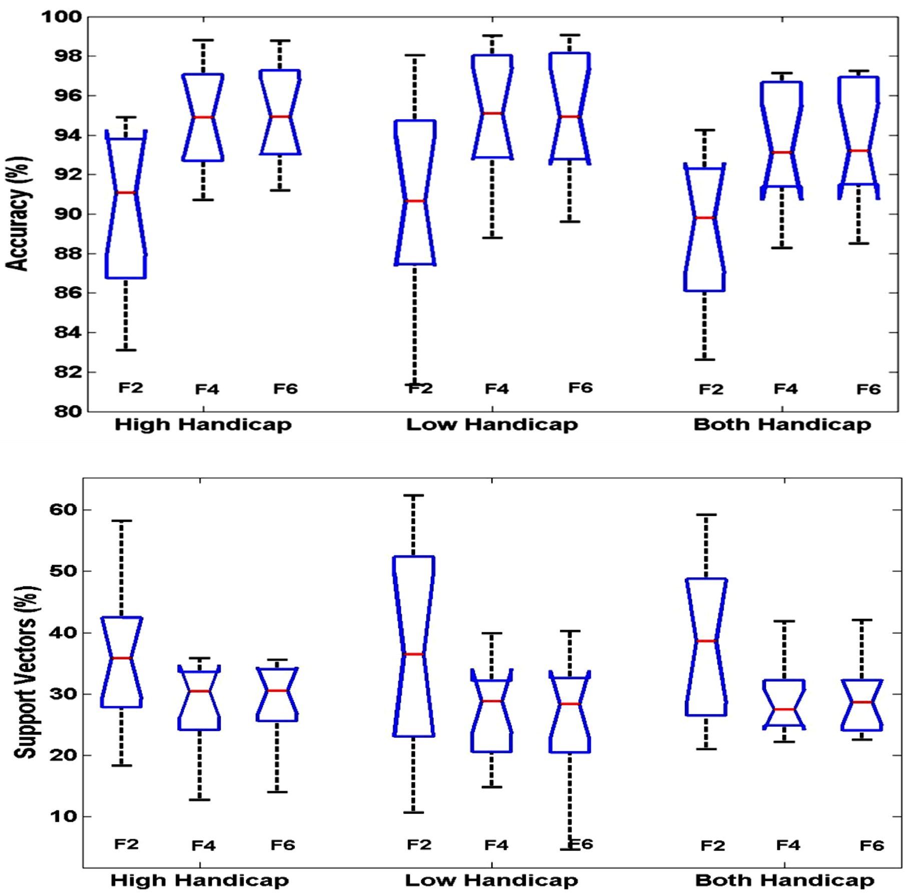


*Figura 12 [**Fig. 4.** Boxplot for accuracy and support vectors by handicap group].*

Tabela 16 [Table 1 – Parameters cost and gamma (median[min; max])].

	Handicap	F2	F4	F6	$\chi^2(gl)$	p	Pairwise Comparisons (p ajusted)		
								F2	**F4**
C	High	$2^{5.63}[2^{0.75}, 2^{8.05}]$	$2^{8.13}\ [2^{5.25}, 2^{8.75}]$	$2^{7.63}\ [2^{5.75}, 2^{8.75}]$	9.911 (2)	0.007**	**F4**	0.013*	-
							F6	0.057	1.0
	Low	$2^{6.02}\ [2^{-3}, 2^{8.75}]$	$2^{8}\ [2^{6.25}, 2^{8.75}$	$2^{8}\ [2^{6.75}, 2^{8.75}]$	14.683 (2)	0.001**		0.005**	-
								0.009**	1.0
	Both	$2^{1.77}\ [2^{-2.5}, 2^{7}]$	$2^{8.63}\ [2^{7.5}, 2^{8.75}]$	$2^{8.63}\ [2^{7.25}, 2^{8.75}]$	18.957 (2)	<0.001**		0.002**	-
								0.000**	1.0
	Total	$2^{4.52}\ [2^{-3}, 2^{8.75}]$	$2^{8.34}\ [2^{5.25}, 2^{8.75}]$	$2^{8.25}\ [2^{5.75}, 2^{8.75}]$	42.318 (2)	<0.001**		0.000**	-
								0.000**	1.0
								F2	**F4**
γ	High	$2^{-3.25}\ [2^{-5.25}, 2^{3}]$	$2^{-4.5}\ [2^{-5}, 2^{-3}]$	$2^{-4.5}\ [2^{-5}, 2^{-3}]$	2.390 (2)	0.303	**F4**	-	-
							F6	-	-
	Low	$2^{0.19}\ [2^{-6}, 2^{3.5}]$	$2^{-4.5}\ [2^{-5.5}, 2^{-3.5}]$	$2^{-4.5}\ [2^{-5.5}, 2^{-3.25}]$	7.860 (2)	0.020*		0.057	-
								0.074	1.0
	Both	$2^{0.38}\ [2^{-5.5}, 2^{3}]$	$2^{-4.5}[2^{-5.5}, 2^{-3}]$	$2^{-4.62}[2^{-5.25}, 2^{-3}]$	11.128 (2)	0.004 **		0.032*	-
								0.024*	1.0
	Total	$2^{-0.08}\ [2^{-6}, 2^{3.5}]$	$2^{-4.5}[2^{-5.5}, 2^{-3}]$	$2^{-4.5}[2^{-4.5}, 2^{-3}]$	19.528 (2)	<0.001**		0.002**	-
								0.001**	1.0

* $p < 5\%$; ; ** $p < 1\%$.

The sets of features differ significantly ($F_{(1.081,35.664)} = 28.995; p < 0.001;\ \eta^2 = 0.468;\ \pi = 1.0$) with no interaction by the different groups of handicap ($F_{(2.161,35.664)} = 0.370; p = 0.709;\ \eta^2 = 0.022;\ \pi = 0.107$). When F2 was considered, 35.8 ± 11%, 37.3 ± 17.6% and 38.7 ± 13.4% of input were required as support vectors for high, low, and all handicap, respectively.

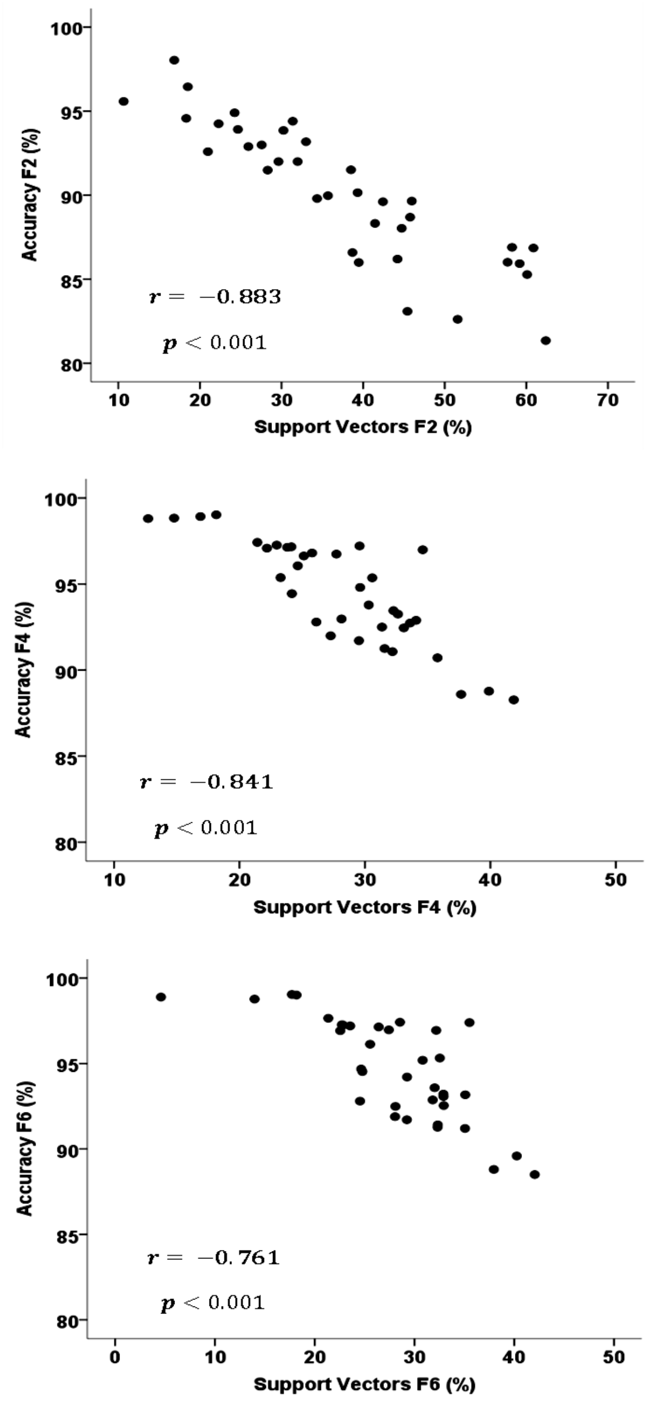


*Figura 13 [**Fig. 5**. – Scatter plot of accuracy vs support vectors by feature use].*

When the F4 and F6 sets were considered, the percentage of support vectors decreases by about 10% in the three groups: high handicap (F4 - 26.3±7.6%; F6 - 28.9±6.5%), low handicap (F4 - 26.8±7.6%; F6 - 26.5±9.6%), and all elements (F4 - 29.2±5.9%; F6 - 29.4±6.2%). Through multiple comparisons it can be seen that the use of F6 did not means less effort by the algorithm in relation to F4 (p = 1.0). However, F2 differs from both F4 and F6 ($p < 0.001$). As for accuracy, there were no differences between handicap groups($F_{(2,33)} = 0.196; p = 0.823;\ \eta^2 = 0.012;\ \pi = 0.078$).

6.4. Discussion

The aim of this study was to verify how the recognition of EMG temporal patterns using time domain features in RBF-SVM is reliable for onset detection. It is intended to give theoretical support for the construction of future models to detect both onset burst and onset peak, allowing the study of motor commands in different tasks. In this particular case, the golf swing has been studied due to its complexity and dynamic characteristics, which requires both accuracy and explosion force. Increasing the discussion about neuromuscular temporal parameters could be one goal of this study, but the main overcome challenge is to overtake the lack of reproducibility between studies [7], [5]. Within the various proposals that have been discussing the automatic detection quality leaving out the analysis of the physiological phenomenon itself. This aspect is still present in the study of McGuill et al. [6] about the phenomenon of the double peak, with a presentation of the synchronized signal to the respective motion picture being taken. When mentioned the concept of onset, this tends to refer only to the first start of muscle activation through the EMG recording, being a concept that may offer little information when it is applied to dynamic and complex motor tasks. This aspect is due to the great algorithm dependency and to consider each motor skill as a whole when it could be composed by different muscle actions. That reason leads to the concept of onset peak as any onset that can occurs before a relevant EMG peak [5].

The only study which used SVM for onset detection was by Soda et al. [56], but not on EMG. It was applied on kinematic variables, raw data and isometric voluntary contraction. In the literature we can find a tendency to study isometric muscle contractions. The reason is related to the movement of the skin changing the detection site in dynamic actions [1] and the ease of detection when an algorithm is tested in isometrics contractions. However, the price to pay is the lack of inference for dynamic actions, fundamental within the sport movement and functional training with the purpose of clinical gains. Given the need to resort to visual inspection to validate the accuracy of the different automatic algorithms, the mathematical models construction of machine learning must have the information regarding the phenomenon to study. To classify the phenomena represented by the classes, the selected features must contain representative information of each class.

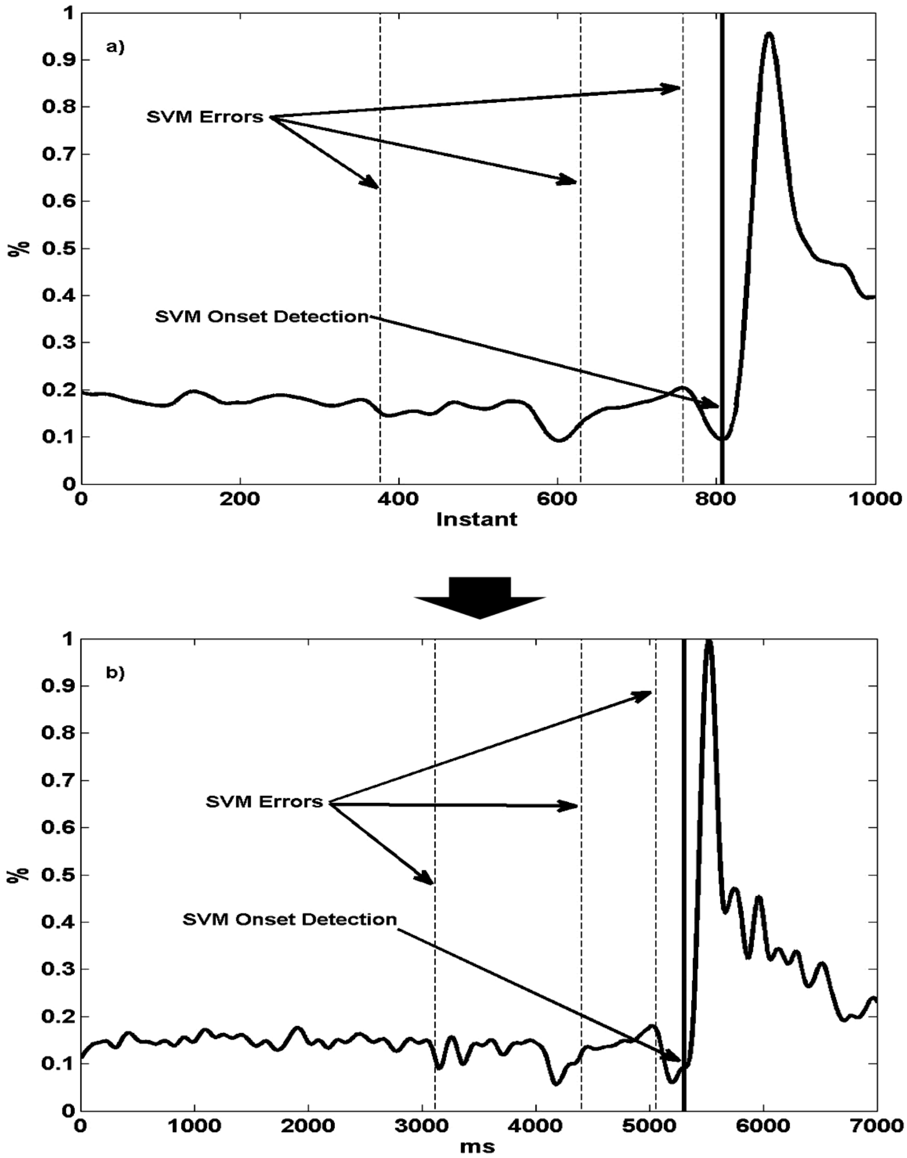


*Figura 14 [**Fig. 6**. Onset detection for left vastus medialis with SVM-F4L model on low handicap subject. a) Detection on features matrix b) Transposition to absolute signal in milliseconds].*

6.4.1. Sets of time domain features

Three sets of features (F2, F4 and F6) were used, with respectively 2, 4 and 6 features considering each model. Its grouping was based in its relevance for the classification, relying on weighting by the Fisher Score [48] and the algorithm Correlation-based Feature Selection [49]. The analysis of dynamics motor skills demands a set of features that provides greater stability within the time window selected. F2 was constituted by two most relevant features, but the F4 and F6 sets showed higher accuracy. Given the similar accuracy obtained by F4 and F6, it can be noted that F6 may contain redundant features with the features studied. However, other proposals could be considered. Tkach et al. [57] analyzed in the biceps and triceps brachial muscles combinations of different time-domain features, and found that the four features WL, slope sign change, LOG and autoregressive coefficients were the more stable set when the level of effort was changed. The same study also found that a set of four features, VAR, LOG, v-Order and EMG histogram showed stability when the classification was subject to changes in the electrodes location. Apart from the WL, VAR and LOG, just MAV are common between the present study and the study of Tkach et al. [57], showing in the last a high accuracy when no disturbance was applied. Other features studied by these authors that show a high rating power were the autoregression coefficients and cepstrum coefficients. However, the inclusion of such features increase the dimensionality of the feature vector, which in turn, increases the complexity of the classification and consequently the computational cost. This is the reason why these features were not included in the present study. Comparing with the F2 set, as the one who had the two best rated features, these results are in agreement as those found by Oskoei & Hu [45], with MAV and WL set showing a high discriminatory power and more stability for different windows length. The set of four features constituted with MAV, WL, zero crossing and slope sign change was the best set. DASDV was not studied by these authors. However, when subjected to increases in the number of trials, DASDV feature exhibits good stability and robustness in classification tasks [58].

The present study aimed to illustrate the performance of time-domain features in the classification of the EMG signals for the onset detection. The features used showed classification accuracy with a mean value of 90.2 %, 94.5 % and 94.6 % for F2, F4 and

F6 sets, respectively. Since time-domain features have better performance in the classification of EMG signal compared with features in the frequency domain [45], [58], they can serve as a basis for automatic classification models, even if other features domains are added. The performance of different types of features may, however, vary in relation to the classifier used in particular features in the time-scale domain [58]. The disadvantage of time-domain features is associated with the fact that they are calculated using the amplitude, which therefore included signal interferences [47].

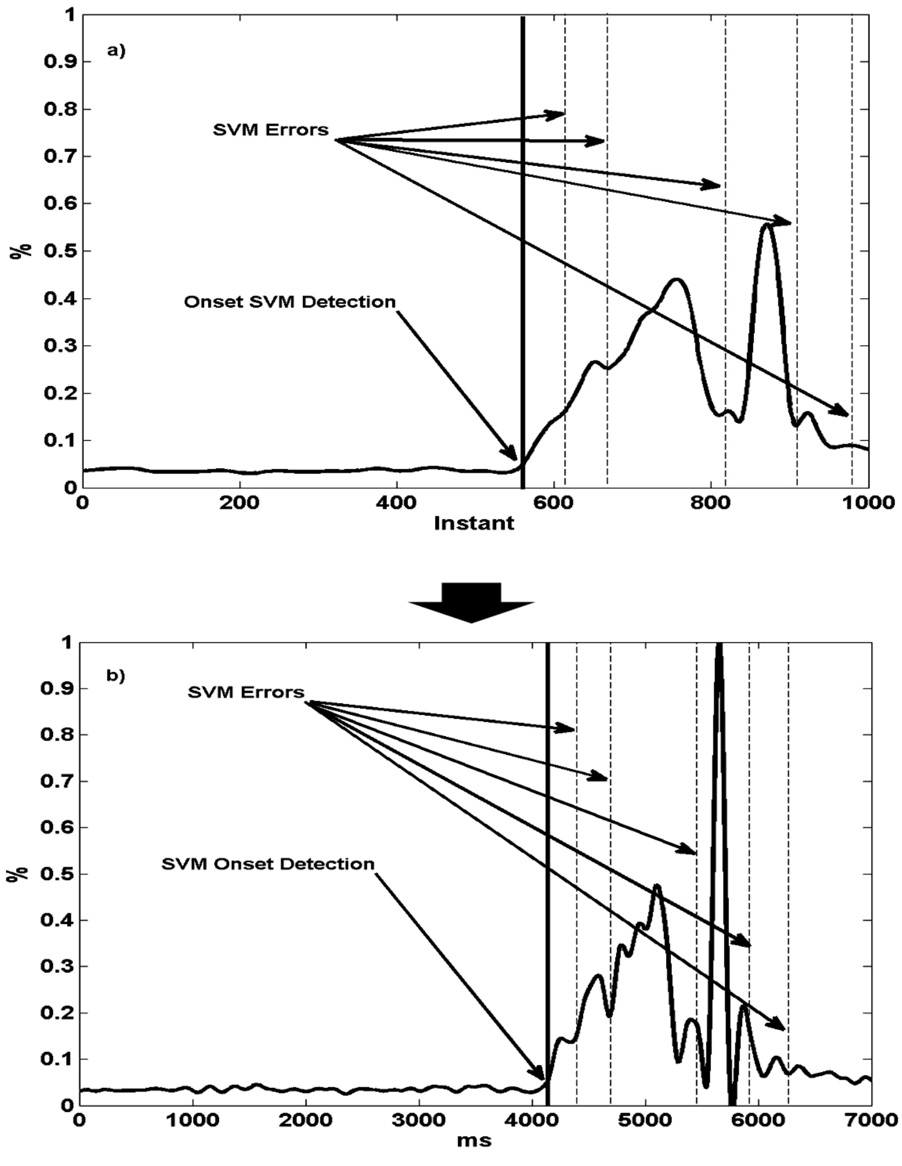


*Figura 15 [**Fig. 7.** – Onset detection for left external oblique with SVM-F6H model on high handicap subject. a) Detection on features matrix b) Transposition to absolute signal in milliseconds].*

A caution regarding the choice of features should be taken. The studies discussed above focused on myoelectric control. Therefore, the goal was to classify motion with features constructed by quantifications of time windows. The goal in this study was to classify time instants within the same EMG signal (as rest or activity).

6.4.2. SVM parameters

To the best of our knowledge, the SVM were never used as a tool for onset detection, since the EMG onset is regarded as a phenomenon that is not within the pattern recognition [36]. Nonetheless, the present study has as starting point the assumption that time instants can be represented by features containing the information regarding their classification in two classes: one before the main activity (before onset), and the other as the main activity (after onset). The nonlinear classifier with RBF kernel was used in this study, since it is the most applied amongst studies using SVM classification, such as in myoelectric control [25–27], [59–61], or in neuromuscular disorders classification [29], [31], [62–64], extending its application to the classification of hand postures taking into account aging [37].

The RBF function is the kernel recommended for nonlinear mapping when the accuracy of others kernel function is unknown [55], [65]. The advantage of using a Kernel function with format $k(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle$ is because allows building algorithms in inner products spaces. It is expected that the kernel class can be written in the positive definitive format, for the equation $k(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle$ be be satisfied and maintained the primal-dual relationship of the SVM algorithm [66], [67]. The Wendland proposition ensures that the Gaussian kernel is positive defined, as the RBF Kernel [66]. Moreover, equations as the linear and sigmoid kernel, within a certain range, tend to behave as a RBF [67]. The polynomial kernel uses more hyperplanes than RBF when the number of hyperplanes influences the complexity in model selection, and the RBF develop less numerical difficulties [65].

The choice of (C, γ) parameters depends on the classification accuracy matrix obtained in the grid-search. With the exception of F2 set which showed greater instability among models, there is a trend for the best RBF parameters varying C between $[2^{5.25}, 2^{8.75}]$ and γ in $[2^{-6.25}, 2^{-1.25}]$, taking into account a spacing of 0.25 in the exponent. There is a tendency for $\gamma < 1$ and C showing higher values, prowling $C = 2^{8.25}$. These parameters are difficult to compare in the literature, since they depend on data and are related to classification. For example, commonly are found in EMG values as $C = 2^5, \gamma = 2^{-3}$ [59], $C = 2^{[5.5,6.5]}, \gamma = 2^{[-0.5,0]}$ in myoelectric control [68], $C = 2^{\approx 8.25}, \gamma = 2^{\approx -2.32}$ in neuromuscular disorders discrimination [63], and in hand movements' classification comparing elderly with young showed $C = 2^{\approx 3.828}, \gamma = 2^{\approx -0.32}$[37] and $C = 2^{6.25}$; $\gamma = 2^{-2.25}$ [62].

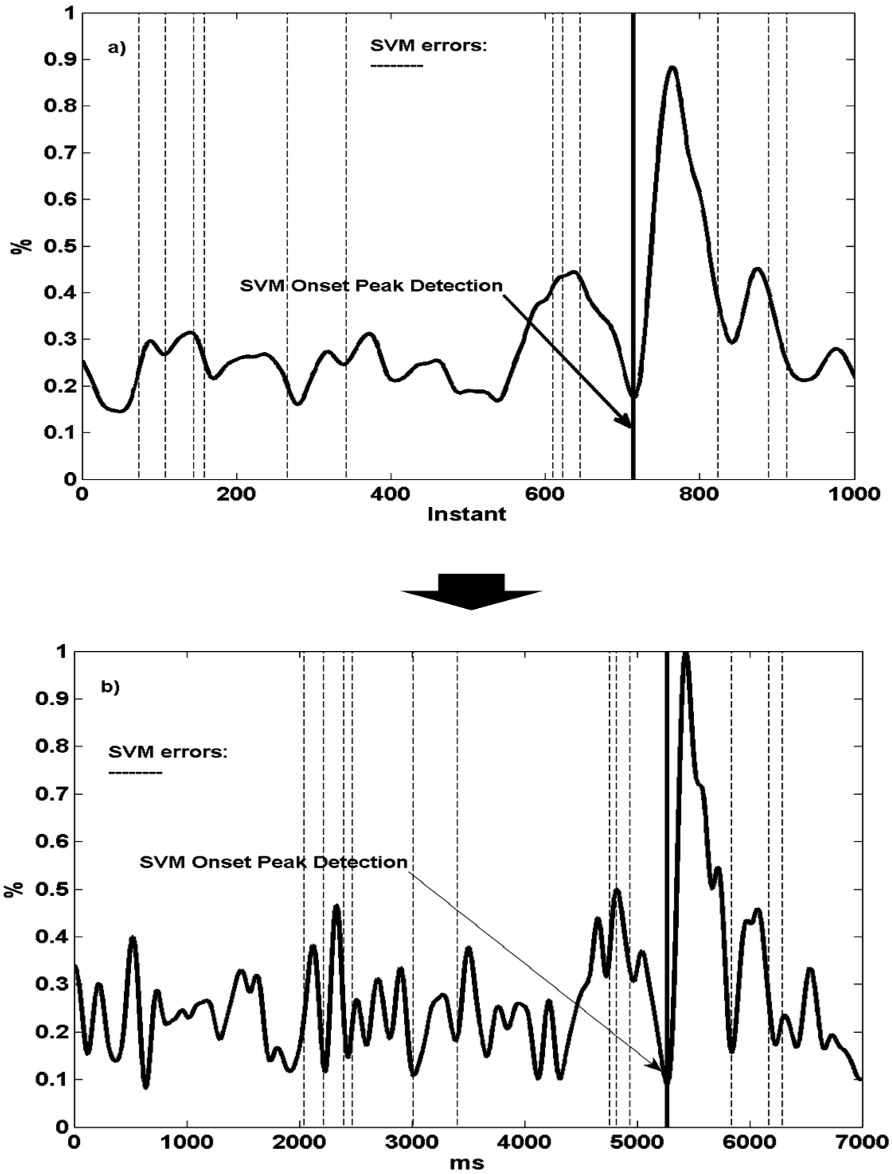


*Figura 16 [**Fig. 8.** – Onset detection for right erector spinae with SVM-F6T model on high handicap subject. a) Detection on features matrix b) Transposition to absolute signal in milliseconds].*

Other approaches have been presented for γ parameter calculation, such as the inverse or the ratio of the number of attributes $\gamma = \frac{1\ or\ 12}{n}$ [23], [24], [32], [69].

6.4.3. Onset detection and classification accuracy

From the three articles that explore the onset detection during golf swing, two focus on the problem of low back pain [70], [71] the other compares two methods varying the baseline that serves as reference [5], but all methods considered threshold algorithms. The concept of onset peak was introduced by Silva et al. [5] considering that EMG signal has moments with relevant activity that should not be overlooked. It is possible to find three physiological phenomena of onset, as there were considered different onset definitions based on the activity with greater relevance. This point led to labeling the two classes (before and after activity) whose separation (onset) is defined according to three different criteria: burst onset coincides with the peak onset, onset burst without considering any activity peak in particular, and onset peak with preceding activity.

Previous publications on onset detection address mainly how to improve the quality in its detection, such as the Teager Kaiser operator [3], a wavelet decomposition [13], and singular spectrum-based change-point analysis [2]. The application of different methods is closely linked to the signal-to-noise ratio as it influences the accuracy of the results. Thus, the activity of the baseline is identified as a major disturbance factor on onset detection by automatic algorithms [5], [8], [9]. So, the performance of classification models is associated with the base line white Gaussian noise. In turn, the accuracy on classification is dependent on the ability of the features to represent each class. In turn, the discrimination ability of features will increase as the signal-to-noise ratio increases [47]. This is well seen in cases in which the temporal phenomenon to detect was the onset peak, when there is previous electrical activity recorded. This phenomenon occurred for all subjects in the ES muscle and this is considered to be a pattern in accordance with previous research [5]. Muscles whose detected phenomena were coincident with the onset burst and the baseline showed fewer errors, always included the true onset. As an average value of 95% with time-domain features was reached, these errors are easily controlled with the inclusion of phenomenological features, or basic resources of mathematical analysis.

6.5. Conclusions

Four time-domain features may be used to classify the instants before and after the onset with a mean accuracy of 95%. Time-domain features as the mean absolute value, waveform length, difference absolute standard deviation value, and the variance of EMG can serve as a basis to build machine learning algorithms for onset detection. The onset peak with high pre activation needs the inclusion of more information to reduce classification errors.

Regarding handicap groups, SVM parameters C and γ do not differ significantly. However, as sample increases, the differences between features set accentuates, namely when comparing four or six features sets with two time-domain features set.

The accuracy shown by the SVM and the relative number of support vectors in the models do not differ by group of handicap. Again, the differences are related to the set of features used only. A two features set requires a greater number of support vectors and presents worse accuracy than the use of the four or six features sets. Great accuracy was correlated with less support vectors, meaning that the easier the classification, the better the performance of the algorithm.

The definition of onset in this study was the beginning of the more relevant activity since it represented a pattern in many trials. This procedure leads to the identification of three onset phenomena, which resulted in a varying difficulty in automatic onset definition mentioned before. Any onset burst, coincident with a greater peak of activity or not, has a high power of classification and detection. When an onset peak with previous activity is aimed, the accuracy falls, showing the increasing difficulty of detection (although the correct onset has always been identified among false trues). Thus, two recommendations related to support vector machines have to be observed. The first refers to the type of the phenomenon studied, if the onset burst or onset peak. When it is desired the detection of the onset burst onset, the use of support vector machines can be a binary classifier of sample points as belonging to the rest class or to the activity class. When the goal is the detection of the onset burst and a particular onset peak, it may be useful the multiclass support vector machines considering the instants before and after each time parameter to detect. The second recommendation is in response to this problem of detecting the onset peak. This study provides references to a set of time-domain features. It is recommended to conduct studies covering more time-domain features with phenomenological features, such as the distance of each point to a particular reference or a wavelet transform. So, keep a base with time-domain features and increase the features number with other type of information.

6.6. Practical Considerations

This approach may be reproduced with others sets of data, and for this, we considered the following key steps:

(1) Describe the operational definition of the temporal phenomenon to be detected (determined peak onset, onset burst ...);

(2) Define the size of the segmented windows considered around each time instant to compute the features and segmentation type;

(3) Justify the method of features extraction (own signal, the time-domain, phenomenological…);

(4) Select the relevant features and eliminate the redundant;

(5) Conduct a grid-search and determine the penalty and kernel parameters;

(6) Check for the transition moments between classes as the onset.

The study of models having time-domain features as base in addition of phenomenological features is recommended. Also, it is recommended to apply this directly to the EMG signal processing with different types of treatment, despite the disadvantages in the representation of the signal and the higher computational cost. This approach tends to be of great utility to allow combining clinical diagnosis to the study of neuromuscular coordination.

Conflict of Interest

None declared

Acknowledgments

The project “Neuromuscular activity in the golf swing with implications for the practice and in the prevention of overuse injuries” was supported by the Portuguese Foundation for Science and Technology (FCT) fund (PTDC/DES/105176/2008).

References

1. De Luca CJ (1997) The use of surface electromyography in biomechanics. J Appl Biomech13:135–63.

2. Vaisman L, Zariffa J, Popovic MR (2010) Application of singular spectrum-based change-point analysis to EMG-onset detection. J Electromyogr Kinesiol 20(4):750–60. doi:10.1016/j.jelekin.2010.02.010.

3. Solnik S, Rider P, Steinweg K (2010) Teager–Kaiser energy operator signal conditioning improves EMG onset detection. Eur J Appl Physiol 110:489–98. doi:10.1007/s00421-010-1521-8.

4. Tyler A, Karst G (2004) Timing of muscle activity during reaching while standing: systematic changes with target distance. Gait Posture 20: 126-133. doi:10.1016/j.gaitpost.2003.07.001.

5. Silva L, Marta S, Vaz J, Fernandes O, Castro MA, Correia-Pezarat, P. (2013) Trunk muscle activation during golf swing: Baseline and threshold. J Electromyogr Kinesio 23(5), 1174–1182. doi:10.1016/j.jelekin.2013.05.007.

6. McGill S, Chaimberg J, Frost D, Fenwick C (2010) Evidence of a double peak in muscle activation to enhance strike speed and force: an example with elite mixed martial arts fighters. J Strength Cond Res 24(2):348–57. doi:10.1519/JSC.0b013e3181cc23d5.

7. Jöllenbeck T (2000) Methodological limitations of EMG-based bio-mechanical motion analysis. ISBS – conference proceedings archive, 18 International symposium on biomechanics in sports.

8. Hodges P, Bui B /1996) A comparison of computer-based methods for the determination of onset of muscle contraction using electromyography. Electroencephalogr Clin Neurophysiol 101:511–9.

9. Allison GT (2003) Trunk muscle onset detection technique for EMG signals with ECG artefact. J Electromyogr Kinesio13(3):209–16.

10. Van Boxtel GJM, Geraars LHD, Van Den Berg-Lenssen MMC, Brunia CHM (1993) Detection of EMG onset in ERP research. Psychophysiology 30(4):405–12. doi:10.1111/j.1469-8986.1993.tb02062.x.

11. Staude G, Flachenecker C, Daumer M, Wolf W (2001) Onset detection in surface electromyographic signals: A SystematicComparison of Methods. EURASIP J. on Applied Signal Processing (2):67–81.

12. Lee AS, Cholewicki J, Reeves NP (2007) The effect of background muscle activity on computerized detection of sEMG onset and offset. J Biomech 40(15):3521–6. doi:10.1016/j.jbiomech.2007.05.012.

13. Vannozzi G, Conforto S, Alessio TD (2010 Automatic detection of surface EMG activation timing using a wavelet transform based method. J Electromyogr Kinesio 20(4):767–72. doi:10.1016/j.jelekin.2010.02.007.

14. Haykin S. (2001) Redes Neurais. Princípios e Prática. 2a ed. ARTMED, Santana.

15. Luxburg U Von, Schölkopf B. (2009) Statistical learning theory: Models, concepts, and results. In: Gabbay DM, Hartmann S, Woods J, editors. Handbook of the History of Logic. Elsevier BV pp. 651–706.

16. Herbrich R. (2002) Learning Kernel Classifiers. Theory and Algorithms. Massachusetts Institute of Technology, Massachusetts.

17. Rojas R. (1996) Neural Networks: A Systematic Introduction. Springer-Verlag, Berlin.

18. Boser BE, Guyon IM, Vapnik VN (1992) A Training algorithm for optimal margin classfiers. Proceedings of the Fifth Annual Workshop on Computational Learning Theory. pp. 144–52.

19. Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273–97. doi:10.1007/BF00994018.

20. Müller K-R, Mika S, Rätsch G, Tsuda K, Schölkopf B (2001) An introduction to kernel-based learning algorithms. IEEE Trans Neural Netw 12(2):181–201.

21. Schölkopf B, Smola AJ, Williamson RC, Bartlett PL (2000) New support vector algorithms. Neural Comput12:1207–45.

22. Chen X, Wang ZJ (2013) Pattern recognition of number gestures based on a wireless surface EMG system. Biomed Signal Process Control 8(2):184–92. doi:10.1016/j.bspc.2012.08.005.

23. Khushaba RN, Kodagoda S, Takruri M, Dissanayake G (2012) Expert systems with applications toward improved control of prosthetic fingers using surface electromyogram (EMG) signals. Expert Syst Appl 39(12):10731–8. doi:10.1016/j.eswa.2012.02.192.

24. Huang, H., Zhang, F., Hargrove, L. J., Dou, Z., Rogers, D. R., Englehart KB (2011) Continuous locomotion-mode identification for prosthetic legs based on neuromuscular–mechanical fusion. IEEE Trans Biomed Eng 58(10):2867–75. doi:10.1109/TBME.2011.2161671.

25. Lucas M-F, Gaufriau A, Pascual S, Doncarli C, Farina D. Multi-channel surface EMG classification using support vector machines and signal-based wavelet optimization. Biomed Signal Process Control 3(2):169–74. doi:10.1016/j.bspc.2007.09.002.

26. Bitzer S, Smagt P van der (2006) Learning EMG control of a robotic hand: Towards active prostheses. Proceedings of the 2006 IEEE International Conference on Robotics and Automation. Orlando, Florida. pp. 2819–23.

27. Castellini C, Smagt P van der, Sandini G, Hirzinger G (2008) Surface EMG for force control of mechanical hands. IEEE International Conference on Robotics and Automation Pasadena CA USA, May 19-23, Pasadena, pp. 725–30.

28. Fligge N, Urbanek H, Smagt P van der (2013) Relation between object properties and EMG during reaching to grasp. J Electromyogr Kinesio 23(2):402–10. doi:10.1016/j.jelekin.2012.10.010.

29. Sultornsanee S, Zeid I, Kamarthi S (2011) Classification of electromyogram using recurrence quantification analysis. Procedia Comput Sci 6:375–80. doi:10.1016/j.procs.2011.08.069.

30. Dobrowolski AP, Wierzbowski M, Tomczykiewicz K (2012) Multiresolution MUAPs decomposition and SVM-based analysis in the classification of neuromuscular disorders. Comput Methods Programs Biomed 107(3):393–403. doi:10.1016/j.cmpb.2010.12.006.

31. Christodoulou CI, Kaplanis P a., Murray V, Pattichis MS, Pattichis CS, Kyriakides T (2012) Multi-scale AM–FM analysis for the classification of surface electromyographic signals. Biomed Signal Process Control 7(3):265–9. doi:10.1016/j.bspc.2012.01.001.

32. Goker I, Osman O, Ozekes S, Baslo MB, Ertas M, Ulgen Y (2012) Classification of juvenile myoclonic epilepsy data acquired through scanning electromyography with machine learning algorithms. J Med Syst 36(5):2705–11. doi:10.1007/s10916-011-9746-6.

33. Stirling LM, Von Tscharner V, Kugler PF, Nigg BM (2012) Classification of muscle activity based on effort level during constant pace running. J Electromyogr Kinesio 21(4):566–71. doi:10.1016/j.jelekin.2011.02.005.

34. Alkan A, Günay M (2012) Identification of EMG signals using discriminant analysis and SVM classifier. Expert Syst Appl 39(1):44–7. doi:10.1016/j.eswa.2011.06.043.

35. Tolambiya A, Thomas E, Chiovetto E, Berret B, Pozzo T (2011) An ensemble analysis of electromyographic activity during whole body pointing with the use of support vector machines. PloS one 6(7):e20732. doi:10.1371/journal.pone.0020732.

36. Oskoei AM, Hu H (2007) Myoelectric control systems—A survey. Biomed Signal Process Control 2(4):275–94. doi:10.1016/j.bspc.2007.07.009.

37. Tavakolan M, Xiao ZG, Menon C (2011) A preliminary investigation assessing the viability of classifying hand postures in seniors. Biomed Eng Online 10(1):79. doi:10.1186/1475-925X-10-79.

38. Soda P, Mazzoleni S, Cavallo G, Guglielmelli E, Iannello G (2010) Human movement onset detection from isometric force and torque measurements: a supervised pattern recognition approach. Artif Intell Med 50(1):55–61. doi:10.1016/j.artmed.2010.04.008.

39. Scheme E, Englehart K (2011) Electromyogram pattern recognition for control of powered upper-limb prostheses: State of the art and challenges for clinical use. J Rehabil Res Dev 48(6):643–59. doi:10.1682/JRRD.2010.09.0177.

40. Hudgins B, Parker P, Scott RN (1993) A new strategy for multifunction myoelectric control. IEEE Trans Biomed Eng 40(1):82–94.

41. EGA (2012) *Sistema de Handicap EGA 2012-2016*. Epalinges/Lausanne: Edição Federação Portuguesa de Golfe.

42. Hermens HJ, Merletti, Freriks B (1996) European activities on surface electromyography. European activities on surface electromyography, Proceedings of the first general SENIAM workshop. SENIAM project, published by Roessingh Research and Development.

43. Watkins RG, Uppal GS, Perry J, Pink M, Dinsay JM (1996) Dynamic electromyographic analysis of trunk musculature in professional golfers. Am J Sports Med 24(4):535–8.

44. Jobe FW, Perry J, Pink M (1989) Electromyographic shoulder activity in men and women professional golfers. Am J Sports Med 17(6):782–7.

45. Oskoei MA, Hu H (2008) Support vector machine-based classification scheme for myoelectric control applied to upper limb. IEEE Trans Biomed Eng 55(8):1956–65.

46. Phinyomark A, Phukpattaranont P, Limsakul C (2012) Feature reduction and selection for EMG signal classification. Expert Syst Appl 39(8):7420–31. doi:10.1016/j.eswa.2012.01.102.

47. Phinyomark A, Limsakul C, Phukpattaranont P (2009) A novel feature extraction for robust EMG pattern recognition. Journal of Computing 1(1):2151–9617.

48. Duda R, Hart P, Stork D (2001) Pattern Classification. 2nd ed. John Wiley & Sons, New York.

49. Hall MA (1999) Correlation-based Feature Selection for Machine Learning. Dissertation, University of Waikato.

50. Zhao Z, Morstatter F, Sharma S, Alelyani S, Anand A, Liu H (2010) Advancing feature selection research-ASU feature selection repository. Arizona State University. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Advancing+Feature+Selection+Research+-+ASU+Feature+Selection+Repository#0 Accessed 20 June 2013.

51. Vapnik VN (1999) An overview of statistical learning theory. IEEE Trans Neural Netw 10(5):988–99.

52. Burges C (1998) A tutorial on support vector machines for pattern recognition. Data Min Knowl Discov 2: 121–167. http://link.springer.com/article/10.1023/A:1009715923555 Accessed 2 May 2012.

53. Schölkopf B, Smola A (2002) Learning with kernels [Internet]. Massachustts: Massachusetts Institute of Technology, Massachusetts. http://dip.sun.ac.za/~hanno/tw796/lesings/mlss06au_scholkopf_lk.pdf Accessed 2 May 2012.

54. Cristianini N, Shawe-Taylor (2000) An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge University Press, Cambridge.

55. Hsu C, Chang C, Lin C (2010) A practical guide to support vector classification. https://www.cs.sfu.ca/people/Faculty/teaching/726/spring11/svmguide.pdf Accessed 2 May 2012.

56. Soda P, Mazzoleni S, Cavallo G, Guglielmelli E, Iannello G (2010) Human movement onset detection from isometric force and torque measurements: A supervised pattern recognition approach. Artif Intell Med 50 (2010) 55–61. doi:10.1016/j.artmed.2010.04.008.

57. Tkach D, Huang H, Kuiken T (2010) Study of stability of time-domain features for electromyographic pattern recognition. J Neuroeng Rehabil 7:21. Doi:10.1186/1743-0003-7-21.

58. Phinyomark A, Quaine F, Charbonnier S, Serviere C, Tarpin-Bernard F, Laurillau Y (2013) EMG feature evaluation for improving myoelectric pattern recognition robustness. Expert Syst Appl 40(12):4832–40. doi:10.1016/j.eswa.2013.02.023.

59. Castellini C, Smagt P van der (2009) Surface EMG in advanced hand prosthetics. Biol Cybern 100(1):35–47. doi:10.1007/s00422-008-0278-1.

60. Chen X, Zhu X, Zhang D (2010) A discriminant bispectrum feature for surface electromyogram signal classification. Med Eng & phys 32(2):126–35. doi:10.1016/j.medengphy.2009.10.016.

61. Yang D, Zhao J, Gu Y, Wang X, Li N, Jiang L, et al (2009) An anthropomorphic robot hand developed based on underactuated mechanism and controlled by EMG signals. J Bionic Eng 6(3):255–63. doi:10.1016/S1672-6529(08)60119-5.

62. Katsis CD, Goletsis Y, Likas a, Fotiadis DI, Sarmas I (2006) A novel method for automated EMG decomposition and MUAP classification. Artif Intell Med 37(1):55–64 doi:10.1016/j.artmed.2005.09.002.

63. Subasi A (2012) Medical decision support system for diagnosis of neuromuscular disorders using DWT and fuzzy support vector machines. Comput Biol Med 42(8):806–15.

64. Subasi A (2013) Classification of EMG signals using PSO optimized SVM for diagnosis of neuromuscular disorders. Comput Biol Med 43(5):576–86.

65. Chang C, Lin C. LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology 2:1–39. http://dl.acm.org/citation.cfm?id=1961199 Accessed 2 May 2012.

66. Hofmann T, Schölkopf B, Smola AJ (2008) Kernel methods in machine learning. Ann Stat 36(3):1171–220. doi:10.1214/009053607000000677.

67. Lin H, Lin C (2003) A study on sigmoid kernels for SVM and the training of non-PSD kernels by SMO-type methods. http://www.csie.ntu.edu.tw/~cjlin/papers/tanh.pdf Accessed 1 Jan 2012.

68. Khokhar ZO, Xiao ZG, Menon C (2010) Surface EMG pattern recognition for real-time control of a wrist exoskeleton. Biomed Eng Online 9:41. doi:10.1186/1475-925X-9-41.

69. Khushaba RN, Kodagoda S, Liu D, Dissanayake G (2013) Muscle computer interfaces for driver distraction reduction. Comput Methods Programs Biomed 110(2):137–49. doi:10.1016/j.cmpb.2012.11.002.

70. Cole M, Grimshaw P (2008) Trunk muscle onset and cessation in golfers with and without low back pain. J Biomech 41:2829–33. doi:10.1016/j.jbiomech.2008.07.004.

71. Horton JF, Lindsay DM, Macintosh BR (2001) Abdominal muscle activation of elite male golfers with chronic low back pain. Med Sci Sports Exerc 33(10):1647–54.

CAPÍTULO VII

7. Recurrence Quantification Analysis and Support Vector Machines for Golf Handicap and Low Back Pain EMG classification in dynamic skills.

Submetido no ***Journal of Electromyography and Kinesiology***

Silva, L.M., Vaz, J.R., Serranho, P., Marta, S., Cabri, J., Pezarat-Correia, P. Recurrence Quantification Analysis and Support Vector Machines for Golf Handicap and Low Back Pain EMG classification. *Journal of Electromyography and kinesiology*.

Recurrence Quantification Analysis and Support Vector Machines for Golf Handicap and Low Back Pain EMG classification.

Silva, Luís; Faculdade de Motricidade Humana, Universidade de Lisboa

Vaz, João Rocha; Faculdade de Motricidade Humana, Universidade de Lisboa

Castro, Maria António; Coimbra College of Health Technology, Polytechnic Institute of Coimbra

Serranho, Pedro; Departamento de Ciências e Tecnologia, Universidade Aberta

Cabri, Jan; Norwegian School of Sport Sciences

Pezarat-Correia, Pedro; Faculty of Human Kinetics, Technical University of Lisbon

Abstract

The quantification of non-linear characteristics of electromyography (EMG) must contain information allowing to discriminate neuromuscular strategies during dynamic skills. In golf, both handicap (Hc) and low back pain (LBP) are main factors associated with the occurrence of injuries. The aim of this study was to analyze the accuracy of support vector machines SVM on EMG-based classification to discriminate Hc (low and high handicap) and LBP prevalence (with and without LPB) in the main phases of golf swing. For this purpose recurrence quantification analysis (RQA) features of the trunk and the lower limb muscles were used to feed a SVM classifier. Recurrence rate (RR) and the ratio between determinism (DET) and RR showed a high discriminant weight. The Hc classifications accuracy for the swing, backswing, and downswing were 94.4±2.7%, 97.1±2.3%, and 95.3±2.6%, respectively. For LBP, the accuracy was 96.9±3.8% in the swing, and 99.7%±0.4% in BS. External oblique (EO), biceps femoris (BF), semitendinosus (ST) and rectus femoris (RF) showed high accuracy depending on the laterality within the phase. RQA features and SVM showed a high capacity in discriminating muscles within swing phases by Hc and by LBP. Low back pain golfers showed less neuromuscular coordination strategies than asymptomatic.

Keywords: SVM, RQA, Golf, neuromuscular discrimination, patterns recognition

7.1. Introduction

Electromyography (EMG) is widely used for neuromuscular pattern characterization. Researchers who use EMG in the time domain usually choose variables related to intensity, duration, and muscle activation sequence. The main problem of surface EMG, when applied to motor behavior studies, is its quantification according to the description of the physiologic phenomenon. For this purpose, as a time series the EMG signal requires methods that are able to detect the nonlinear characteristics, for instance, the stochastic, non-stationary and deterministic behavior (Lei and Meng, 2012). Thus, multivariate and non-linear methods are needed, rather than linear approaches that are often not appropriate (Marwan et al., 2002; Tolambiya et al. 2011).

Recurrence plots were introduced by Eckmann et al.(1987), who described a graphical tool for measuring the time consistency of a dynamic system. Recurrence Quantification Analysis (RQA), was developed later by Webber and Zbilut (1994) and is regarded as a useful method of nonlinear data analysis. Dynamic systems are explained by the time evolution of the phase space trajectory. The vector $\vec{x}(t)$ in a d-dimensional space, formed by the d variables $x_1(t), x_2(t), \dots, x_d(t)$ about the state of a system in time, is called phase space. This vector is specified by its velocity vector $\dot{\vec{x}}(t)$ as it is moving in time and in a certain direction (Marwan et al., 2007).

The application of RQA on EMG has proven to be a sensitive tool to study muscle fatigue, showing high sensitivity to changes in muscle status like motor unit synchronization when compared with the traditional spectral analysis. It is highly correlated with spectral variables in the biceps brachii (Farina et al., 2002; Filligoi and Felici 1999), the extensor carpi radialis (Del Santo et al., 2006), and the back muscles (Ikegawa et al., 2000). These afore mentioned studies suggested that an increase in the percentage of determinist is related with an increase in fatigue. Also the percentage of recurrence is influenced by the degree of synchronization and conduction velocity (Farina et al., 2002).

Support Vector Machines (SVM) are complex and advanced algorithms in the field of supervised learning (Boser et al., 1992; Vapnik, 1999), used in classification and regression problems. SVM showed a high accuracy on EMG-based features when compared with others classifiers of myoelectric control (Castellini and Smagt, 2009), neuromuscular disorders (Güler and Koçer, 2005), and kinesiologic analysis (Shi et al., 2009). In addition to the classification, SVM have also showed high ability in the regression of kinetic parameters using EMG in grasping postures (Castellini et al., 2009) and in isometric wrist flexion and extension (Ziai and Menon, 2011). Sultornsanee et al. (2011) applied both RQA features and SVM to distinguish neuromuscular disorders in three groups, i.e. healthy, myopathy, and neuropathy, reaching an accuracy between 93.33% and 100%.

The golf swing is a dynamic complex task requiring both power and accuracy (Hume et al., 2005). The handicap (Hc) and low back pain (LBP) are associated with changes in the golf swing (Cabri et al, 2009; Lindsay et al, 2000). Lindsay et al. (2000) stated that more than 70% of golf players experienced injuries resulting in playing at an unsatisfactory skill level during a short period of time. LBP has been implicated as the major complaint of golfers as well as the body region associated with a larger incidence of injuries (Cabri et al., 2009; Lindsay et al., 2002; McHardy et al., 2007). This led to a growing interest in quantifying the factors that can be influenced by LBP (Gluck et al., 2008; Vad et al., 2004), as the e.g. EMG onset (Cole & Grimshaw 2008a; Horton et al., 2001).

The aim of this study was to identify the accuracy of SVM on EMG-based classification to discriminate Hc (low and high handicap) and LBP (with and without LPB) in the main phases of golf swing: preparation (backswing), execution (downswing) and reestablishment (follow-through). Additionally, we intended to determine 1) which muscles have discriminatory power and 2) which of the RQA features are the most relevant.

7.2. Method

7.2.1. Subjects

Twenty-one golfers performed eight trials with two clubs (pitch, 7-iron). The subjects were divided by Hc and by LBP perception after conducting an 18-hole golf course (Table 1). Twelve subjects (age = 52.5±12.13 years; Hc = 15.12±12) were assigned to two groups: with (five golfers) and without (seven golfers) LBP. Age ($U = 9.0$; $p = .189$; $g = .851$), handicap ($U = 15.5$; $p = .785$; $g = .316$), and golf practice time ($U = 9.0$; $p = .246$; $g = .726$) were homogeneous between LBP groups. Ten subjects were assigned in two other groups: five with low handicap (LHc) $Hc < 5$ (0.7 ± 2.2) and five with high handicap (HHc) ($Hc \geq 18$ (24.3 ± 4.6). One subject was assigned to both LBP and HC analysis. Hc was based on the European Golf Association recommendations (EGA, 2012), and was taken of the assumption that SVM parameters are similar at different ages (Tavakolan et al., 2011). For the LBP discrimination the Musculoskeletal Injury Questionnaire for Senior Golfers (Fox et al., 2002), Portuguese version (Silva, 2014) was used. This questionnaire allows quantifying the percentage of how often golfers are aware of LBP after golfing 18 holes. Participants selected in the no LBP (NLBP) group answered "0%" and not reported any musculoskeletal injury diagnosis with LBP. Participants with more than 65% of LBP were enlisted in the LBP group. The LBP subjects were asymptomatic (pain free) during the tests.

Tabela 17 [Table 1 – Participants characteristics].

	Groups Mean[Standard Deviation]			
	Low Back Pain		Handicap	
	Without	With	Low	High
Age (years)	48.1[11.7]	58.6 [10.9]	30.4 [7.0]	44 [10.9]
Mass (kg)	82.5 [8.9]	80.7 [6.6]	70.6 [4.9]	80.6 [19.4]
Height (m)	1.75 [0.05]	1.75 [0.05]	1.72 [0.06]	1.72 [0.08]
Handicap (Hc)	14 [4.4]	16 [7.5]	0.7 [2.2]	24.3 [4.6]
Golf Practice (years)	7 [2.8]	12.5 [10.4]	19.2 [4.4]	5.9 [5.3]

7.2.2. Task

The participants were instructed to perform the swing trials on an artificial turf golf mat with high shock absorption characteristics. Subjects were advised to take into consideration their average distances with the two clubs looking to a target. The golfers assessed their own shots being considered those that were the best performed. The two clubs had graphite shafts of standard length. The rest between trials was 1 minute.

7.2.3. EMG procedures

EMG data was collected with bioPLUX® research 2010 telemetric equipment (Plux, Lisbon, Portugal) using active surface electrodes (Al/AgCl, disk shape 10 mm of diameter) and surfaces of detection AMBU® BlueSensor N (shape 30 x 22 AMBU, Ballerup, Denmark). The EMG signals were collected with sampling frequency of 1000Hz, filtered with a bandpass filter between 10 and 500 Hz, common-mode rejection ratio (CMRR) of 110 dB and input impedance was greater than 100 MΩ. After storage, the data were digitally filtered (10–490 Hz), full wave rectified, and amplitude normalized to the subject's peak over the trials.

The skin was prepared by hair removal, abrasion and alcohol cleaning and muscle contraction was performed before fixation in order to better visualize the muscle belly. The electrodes were placed with a 20 mm center-to-center distance and applied in parallel to the muscle fibers, bilaterally on rectus femoris (RF); biceps femoris (BF); semitendinosus (ST); external oblique (EO); and unilaterally on the left gluteus maximus (GM); erector spinae (ES), as described by Hermens et al. (1996). The ground electrode was placed on the manubrium. Muscle selection was based on its side relevance in golf swing (Gatt et al.1998; Gulgin et al., 2009).

7.2.4. Video data recording, processing and kinematic analysis

Three high-speed Basler A602fc cameras (Basler Vision Technologies, Ahrensburg, Germany) at 100 Hz were placed in anterior, posterior and superior oblique positions to determine the swing phases. A fourth Casio Ex-FH20 camera (Casio, Tokyo, Japan) at 1000 Hz was placed in front of the ball, in order to determine the instant of impact.

Two reflective tapes were placed on the club to divide the swing in three phases (Hume et al., 2005): (1) the Backswing – from the beginning until the top of the swing; (2) the Downswing – from the top until impact; and (3) the Follow-Through – from impact until the end of the swing. SIMI 3D Motion system (SIMI Reality Motion System GmbH, Unterschleissheim, Germany) was used for EMG-synchronized 3D kinematic analysis.

7.2.5. Recurrence quantification analysis

Recurrence plots is a mathematical technique that allows visualizing the recurrence of dynamical systems. A set of vector $\{\vec{x_i}\}_{i=1}^{N}$ of a system, is described by a series that represents a trajectory in the phase space. Recurrence plots depends on the matrix:

$$R_{i,j} = \begin{cases} 1: \vec{x}_i \approx \vec{x}_j \\ 0: \vec{x}_i \not\approx \vec{x}_j \end{cases} \quad i, j = 1 \ldots, N \qquad (1)$$

So, RP for time-discrete variables, where $t = i\Delta t$, is defined as

$$R_{i,j} = \Theta\left(\varepsilon_i - \left\|\vec{x}_i - \vec{x}_j\right\|\right) \qquad (2)$$

where ε_i is the cut-off distance, $\|\cdot\|$ is a given norm, and $\Theta(x)$ the Heaviside function. For this study the Euclidian norm was used as neighborhood method with a threshold $\varepsilon_i = 10\%$ of the mean phase space (Marwan et al., 2007) The phase space reconstruction requires two main parameters, embedding dimension (m) and time delay (τ). The τ value was found by Mutual information ($\tau = 2$), and then m value adjusted by false nearest neighbors, than a vector in phase space is reconstructed by (Marwan et al., 2007);

$$\vec{x}_i = \sum_{k=0}^{m-1} \xi_{i+k\tau} + \vec{e}_k \qquad (3)$$

RQA features were extracted with Toolbox 5.17 (R28.20; PIK Potsdam). Given a *N* number of points on the phase space trajectory, N_l is the number of diagonal lines in the recurrence plot, N_v the number of vertical lines in the recurrence plot, and *P(l)* and

P(v) are the histogram of the line lengths of diagonal and vertical lines, respectively. So, the features extracted are (Marwan et al., 2007):

Recurrence rate (RR), a measure based on recurrence density that depends on the average number of neighbors of each point on the trajectory:

$$RR = \frac{1}{N^2} \sum_{i,j=1}^{N} R_{i,j} \qquad (4)$$

Determinism (DET) is based on diagonal lines, depends on the histogram $P(\varepsilon, l)$ and corresponds to the ratio of recurrence points that form diagonal lines in relation to all recurrence points:

$$DET = \sum_{l=l_{min}}^{N} lP(l) / \sum_{l=1}^{N} lP(l) \qquad (5)$$

Divergence (DIV) is the inverse of L_{max}, another measure based on diagonal lines:

$$DIV = \frac{1}{L_{max}} \qquad (6)$$

where L_{max} is the length of the longest diagonal line $L_{max} = \max(\{l_i; i = 1, \dots N_l\})$, and $N_l = \sum_{l \geq l_{min}} P(l)$ is the total number of diagonal lines.

Entropy (ENT) is based on diagonal lines and it means the Shannon entropy of the probability $p(l) = P(l)/N_l$ to find a diagonal line of exactly length l in the recurrence plot.

$$ENT = -\sum_{l=l_{min}}^{N} p(l) \ln p(l) \qquad (7)$$

Laminarity (LAM) is based on vertical lines, corresponds to the ratio between the vertical structures recurrence points and the all set of recurrence points computed.

$$LAM = \sum_{v=v_{min}}^{N} vP(v) / \sum_{v=1}^{N} vP(v) \qquad (8)$$

Trapping Time (TT) is the average length of the vertical lines:

$$TT = \sum_{v=v_{min}}^{N} vP(v) / \sum_{v=v_{min}}^{N} P(v) \qquad (9)$$

The ratio between *determinism* and *recurrence rate* (DET/RR), and between *laminarity* and *determinism* (LAM/DET) was also considered to verify the utility of the relationship between recurrence, vertical and diagonal lines.

7.2.6. Support vector machines

Support vector machines are a useful machine learning tool developed in the 90's (Boser et al., 1992; Cortes and Vapnik, 1995) for solving classification problems. Given a training set of inputs x_i and the output label y_i, where $x_i \in \mathbb{R}^n$ and $y \in \{-1,1\}^l$, the support vector machines search the solution of the following optimization problem (see Burges, 1998; Hsu et al., 2010):

$$\min_{w,b,\xi} \quad \frac{1}{2} w^T w + C \sum_{i=1}^{l} \xi_i \qquad (10)$$

subject to

$$y_i(w^T \phi(x_i) + b) \geq 1 - \xi_i \qquad (11)$$

$$\xi_i \geq 0$$

The training vectors x_i were mapped into a higher dimensional space with the kernel radial basis function (RBF) expressed as follows:

$$K(x_i, y_i) = \exp\left(-\gamma\|x_i - x_j\|^2\right), \gamma > 0 \qquad (12)$$

The grid search was performed with $C = 2^{-5:2:17}$ and $\gamma = 2^{-9:2:5}$ based on previous recommendations (Hsu et al., 2010). In both analysis, Hc and LBP, the trails were separated in two parts, 80% for training and the others 20% for test. High discriminatory power was considered for muscles (alone) with an accuracy equal or higher than 85%.

Figure 1 summarizes the data procedures in this study.

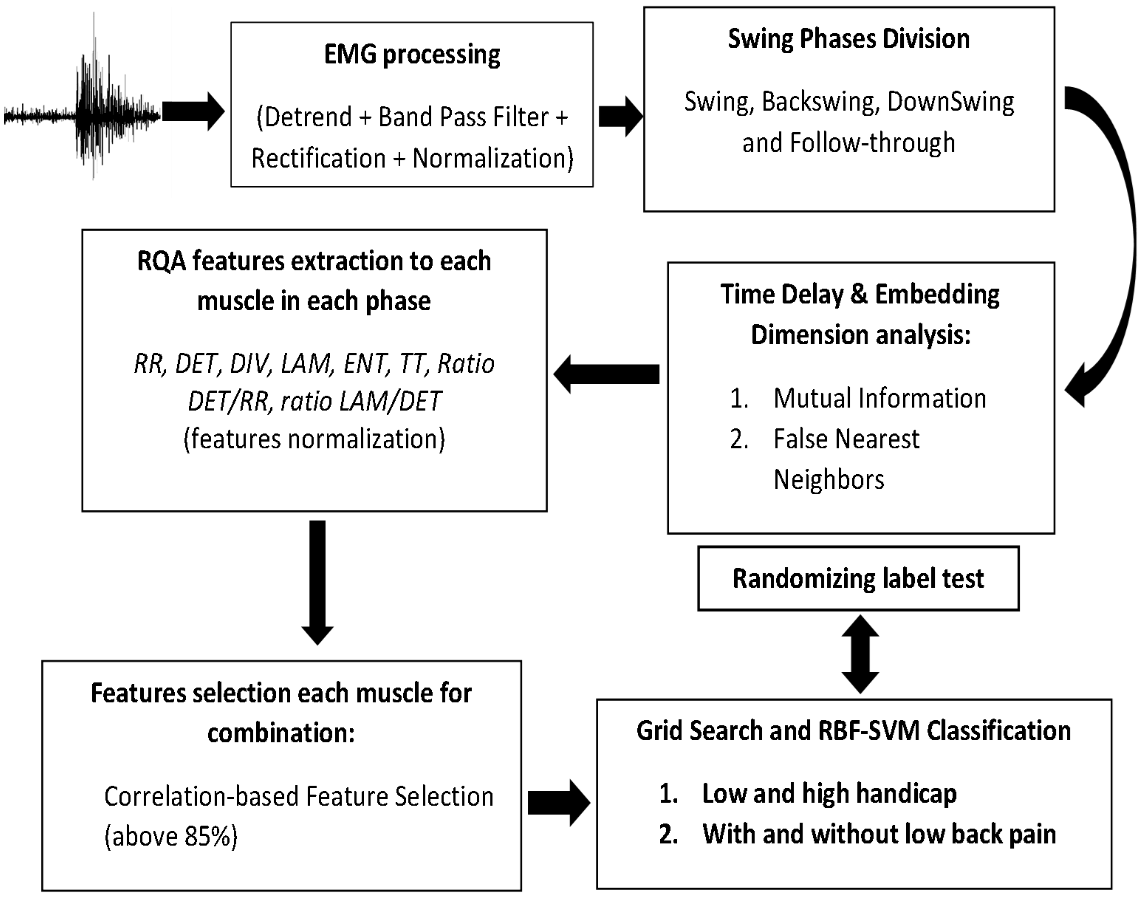


*Figura 17 [**Fig. 1.** Flow chart research procedures].*
Hc – Handicap; LBP – Low back pain; RR – Recurrence rate; DET – Determinism; DIV – Divergence; ENT – Entropy; LAM – Laminarity; TT – Trapping time.

7.2.7. Complementary statistical analysis

Four normalization steps were performed: EMG amplitude; Z-score for RQA analysis; time scale by the minimum phase duration; features that feed the SVM classifier.

The Mann-Whitney U test was performed to control the homogeneity between LPB and NLBP, also the effect size was analyzed by Hedges' g. Nonparametric MANOVA was performed to analyze the values of EMG recurrence rate of muscles between groups. The test statistics χ^2 for nonparametric MANOVA was calculated using the Pillai's Trace, and the *p-value* was corrected for analysis of muscles RR between groups. F statistics test *p-value* ANOVA was corrected by approximate χ^2distribution H_F test. Friedman test multiple comparisons were performed for RR during swing and backswing, each group alone ($\alpha = 5\%$).

Features were selected using the Correlation-based Feature Selection algorithm (Hall, 1999; Zhao et al., 2010).

Data processing was performed with MATLAB® R2013a (Mathworks Inc., Natick Massachusetts, USA).

7.3. Results

The main features selected for both Hc and LBP (Figure 2) were DET/RR, RR, and DIV. The DET/RR ratio were selected 72.5% and 60% times, the RR 60% and 55%, and DIV 37.5% and 45%, for Hc and LBP, respectively. However, DIV showed no relevance on backswing when it was used to discriminate trials of LBP and NLBP.

The backswing was the major discriminant phase. Follow-through showed no discriminatory capacity in any of the situations studied, and downswing did not show enough capability to classify LBP (considering accuracy ≥ 85%). Figure 3 shows the grid search for two channels, four and five channels. Figure 4 shows features sets with the channels with higher accuracy by phase, for Hc and LBP. The classification accuracy for Hc was 94.4±2.7 [91.6-98.1%] during the swing, 97.1±2.3 [94.9-99.1%] in the backswing, 95.3±2.6 [91.6-97.2%] during the downswing. During backswing, the right

EO showed the greatest accuracy (94.9%), followed by right BF (94.4%), left ST (88.9%), right ST (86.9%), and right RF (86.9%). For LBP, the results were 96.9±3.8 [90.9-100%] for the swing, and 99.7%±0.4 [99.3-100%] in the backswing. The left BF showed the higher accuracy (99.3%) followed by the left ST (88.7%), right BF (86.6%), and left RF (85.9%).

Nonparametric MANOVA showed significant differences in RR between Hc for swing $(\chi^2(2) = 66.528; N = 97; p < .001)$ and backswing $(\chi^2(2) = 54.912; N = 97; p < .001)$. Similar results were found for the LBP, both in swing $(\chi^2(2) = 135.8; N = 176; p < .001)$ as in backswing $(\chi^2(2) = 124.425; N = 176; p < .001)$ phases.

The muscles with greater RR sensitivity between Hc groups during the swing were BF, GM, EO, ST ($p < .001$) and RF ($p = .027$) of the leftside and right ST ($p = .024$). In the backswing, between Hc groups, the BF, EO of the right side, and ES, RF ($p < .001\ for\ all$) and ST ($p = .007$) of the left side showed significant differences. The remain muscles each phase showed similarity ($p \in [.083, .098]$) between Hc groups.

RR differences were found between LBP symptomatic and asymptomatic for almost all muscles ($p < .001; GM: p = .023$) for both phases with the exception of left side GM ($p = .135$), EO and ST ($p = .132$), and right RF ($p = .083$) in the swing, and right side EO ($p = .585$), RF ($p = .892$) and ST ($p = .862$) in the backswing.

The RR showed significant differences within muscles in the swing $(LHc: \chi^2_F(9) = 171.27, p < .001;\ HHc: \chi^2_F(9) = 113.37, p < .001)$ and backswing $(LHc: \chi^2_F(9) = 200.95, p < .001;\ HHc: \chi^2_F(9) = 94.2, p < .001)$ for Hc groups. Also, LBP symptomatic and asymptomatic showed significant differences in the swing $(NLBP: \chi^2_F(9) = 306.48, p < .001; LBP: \chi^2_F(9) = 350.32, p < .001)$ and backswing $(NLBP: \chi^2_F(9) = 496.22, p < .001; LBP: \chi^2_F(9) = 335.49, p < .001)$. Figures 5 and 6 show the multiple comparisons.

Tabela 18 [Table 2 – Individual left muscle RQA and SVM parameters by handicap].

Serie (n=132)		Selected Features	SVM Results		
			$C-\gamma$	%SV	ACC
RF	Swing	RR, DIV, TT	2^7-2	23.4	**91.6**
	BS	RR, DIV, DET/RR	$2-2^{-3}$	72.9	**86.9**
	DS	RR, DET, DIV, ENT, LAM, TT, DET/RR, LAM/DET	2^3-2	42.1	**87.9**
	FT	RR, DIV	$2^{13}-2^{-1}$	63.6	65.4
EO	Swing	RR, ENT, DET/RR	$2-2^{-1}$	61.7	**85.1**
	BS	RR, LAM, DET/RR	$2^{13}-2^3$	44.9	80.4
	DS	RR	$2^{-1}-2^5$	86.9	64.5
	FT	DET, ENT, LAM, TT	2^7-2^5	86.0	59.8
BF	Swing	RR, ENT, DET/RR	2^9-2^{-1}	49.5	77.6
	BS	DIV, DET/RR	2^7-2^{-1}	34.6	84.1
	DS	RR, DET/RR	2^3-2^{-1}	51.4	**85.0**
	FT	DET	$2-2^{-7}$	88.8	65.42
ST	Swing	RR, DET, LAM	2^3-2	43.9	81.1
	BS	RR, DET/RR	2^9-2	34.6	**88.8**
	DS	RR, ENT, TT, DET/RR	$2-2$	38.3	**91.6**
	FT	RR, ENT, TT, DET/RR	$2^{-1}-2^3$	88.7	80.4
GM	Swing	RR, ENT, TT	$2^{11}-2$	43.0	84.1
	BS	DET, DIV, LAM, TT, DET/RR, LAM/DET	2^3-2	57.9	80.4
	DS	DET/RR	$2^{13}-2^{-1}$	54.2	78.5
	FT	RR, ENT, DET/RR	2^3-2^{-1}	61.7	83.2
ES	Swing	DET, DIV, LAM, TT, LAM/DET	2^3-2	48.6	83.2
	BS	RR, LAM, DET/RR	2^9-2^5	59.8	77.6
	DS	DIV, DET/RR	2^9-2	55.1	75.7
	FT	DET/RR	2^5-2^3	87.9	60.0

Legend: ACC – Accuracy; %SV – Percentage of support vectors; RF – Rectus femoris; EO – External oblique; BF – Biceps femoris; ST – Semitendinous; GM – Gluteus maximus; ES – Erector spinae; BS – Backswing; DS – Downswing; FT – Follow-through. Bold – Equal or higher than 85%.

Tabela 19 [Table 3 – Individual right muscle RQA and SVM parameters by handicap].

Serie (n=132)		Selected Features	SVM Results		
			$C-\gamma$	%SV	ACC
RF	Swing	DIV, LAM, DET/RR	$2^{13}-2^{3}$	31.0	78.2
	BS	DET/RR	$2^{-1}-2^{-5}$	53.5	66.7
	DS	RR, DET/RR	$2^{7}-2^{-1}$	27.5	79.5
	FT	DIV, LAM, DET/RR, LAM/DET	$2^{5}-2$	54.9	73.1
EO	Swing	RR, DIV, DET/RR	$2^{5}-2$	33.1	82.1
	BS	ENT, LAM/DET	$2^{11}-2$	14.1	**94.9**
	DS	DET, LAM, DET/RR, LAM/DET	$2^{7}-2$	20.4	**85.9**
	FT	ENT, DET/RR	$2^{11}-2$	39.7	80.7
BF	Swing	RR, DIV, TT	$2^{3}-2^{-1}$	57.0	**86.0**
	BS	DIV, TT, DET/RR	$2^{-1}-2^{-1}$	52.3	**94.4**
	DS	TT, DET/RR, LAM/DET	$2^{-3}-2$	87.9	82.2
	FT	RR, DET/RR, LAM/DET	$2^{-3}-2$	94.4	70.1
ST	Swing	RR, DIV, LAM, DET/RR, LAM/DET	$2^{3}-2^{3}$	57.9	80.4
	BS	RR, LAM, TT, LAM/DET	$2^{13}-2^{3}$	36.5	**86.9**
	DS	RR, DIV, TT, DET/RR, LAM/DET	$2^{13}-2^{3}$	39.3	82.2
	FT	DET, ENT, TT, DET/RR, LAM/DET	$2^{-1}-2^{-1}$	86.9	74.8

Legend: ACC – Accuracy; %SV – Percentage of support vectors; RF – Rectus femoris; EO – External oblique; BF – Biceps femoris; ST – Semitendinous; GM – Gluteus maximus; ES – Erector spinae; BS – Backswing; DS – Downswing; FT – Follow-through. Bold – Higher than 85%.

Tabela 20 [Table 4 – Individual left muscle RQA and SVM parameters by low back pain].

	Phase (n=176)	RQA Selected Features	SVM Results		
			$C-\gamma$	%SV	ACC
RF	Swing	RR, DIV, TT, DET/RR	$2^{13}-2$	33.1	**87.3**
	BS	LAM, DET/RR	$2-2^{-1}$	51.4	**85.9**
	DS	RR, DET, DIV, ENT, LAM/DET	$2^{7}-2^{3}$	52.1	77.5
	FT	LAM, DET/RR	$2-2^{3}$	73.9	62.0
EO	Swing	RR, ENT, TT	$2^{5}-2^{-1}$	63.4	68.3
	BS	RR, DET/RR	$2^{8}-2$	45.8	82.4
	DS	RR, DET, DIV	$2^{5}-2$	76.7	70.4
	FT	RR, DET, ENT, LAM, TT	$2^{-5}-2^{-5}$	82.4	63.4
BF	Swing	RR, DET, DIV, LAM, DET/RR, LAM/DET	$2^{7}-2$	38.7	83.1
	BS	RR, DET/RR	$2^{-3}-2^{-1}$	42.3	**99.3**
	DS	DIV, ENT, LAM, TT, LAM/DET	$2^{5}-2^{3}$	49.3	84.5*
	FT	RR, DET, ENT, LAM, TT, DET/RR, LAM/DET	$2^{15}-2^{3}$	49.3	74.6
ST	Swing	RR, DIV, ENT, TT, DET/RR	$2^{9}-2$	26.8	**90.9**
	BS	RR, DET, LAM, TT	$2-2^{-3}$	83.8	**88.7**
	DS	ENT, LAM/DET	$2^{13}-2^{-1}$	62.7	66.2
	FT	DET, ENT, TT, DET/RR	$2^{-1}-2$	89.4	62.0
GM	Swing	ENT, DET/RR, LAM/DET	$2^{7}-2$	34.5	**88.0**
	BS	ENT, DET/RR	$2^{3}-2^{-1}$	50.7	78.9
	DS	DIV, ENT, LAM, DET/RR, LAM/DET	$2^{11}-2$	42.3	77.5
	FT	DET, DIV, ENT, LAM, TT, DET/RR, LAM/DET	$2^{5}-2^{3}$	43.7	79.6
ES	Swing	RR, DET, ENT, LAM/DET	$2^{3}-2^{-1}$	43.7	**85.2**
	BS	RR, DET, LAM, DET/RR, LAM/DET	$2^{5}-2$	52.8	76.1
	DS	DIV, LAM	$2^{-1}-2^{-1}$	69.0	75.4
	FT	RR, DIV, DET/RR	$2^{15}-2$	52.1	73.9

Legend: ACC – Accuracy; %SV – Percentage of support vectors; RF – Rectus femoris; EO – External oblique; BF – Biceps femoris; ST – Semitendinous; GM – Gluteus maximus; ES – Erector spinae; BS – Backswing; DS – Downswing; FT – Follow-through. Bold – Higher than 85%.

Tabela 21 [Table 5 – Individual right muscle RQA and SVM parameters by low back pain].

Serie (n=176)		Selected Features	SVM Results		
			$C-\gamma$	%SV	ACC
RF	Swing	RR, DIV, LAM, LAM/DET	$2^{8}-2^{-1}$	67.6	77.5
	BS	ENT, LAM, DET/RR	$2-2$	66.9	72.5
	DS	DET, DIV, ENT, LAM/DET	$2^{3}-2$	74.7	72.5
	FT	DET, DIV, DET/RR, LAM/DET	$2-2$	78.9	74.0
EO	Swing	RR, TT, DET/RR	$2^{15}-2^{-1}$	38.7	73.2
	BS	DET/RR, LAM/DET	$2^{15}-2$	42.3	79.6
	DS	RR, DIV	$2^{-3}-2^{-3}$	100	63.4
	FT	RR	$2^{5}-2^{3}$	90.1	62.7
BF	Swing	DET, DIV, LAM, TT, DET/RR	$2^{9}-2$	42.3	81.0
	BS	RR, DET, LAM, TT, DET/RR	$2^{3}-2$	46.5	**86.0**
	DS	RR, DIV	$2^{-5}-2^{-9}$	100	57.0
	FT	DIV, DET/RR	$2^{3}-2^{5}$	94.4	72.5
ST	Swing	RR, DET, DIV	$2^{13}-2$	29.6	**86.6**
	BS	TT, DET/RR, LAM/DET	$2^{9}-2$	45.1	78.2
	DS	RR, ENT, LAM, TT	$2^{9}-2$	56.6	72.5
	FT	DET/RR	$2^{-1}-2$	55.6	64.1

Legend: ACC – Accuracy; %SV – Percentage of support vectors; RF – Rectus femoris; EO – External oblique; BF – Biceps femoris; ST – Semitendinous; GM – Gluteus maximus; ES – Erector spinae; BS – Backswing; DS – Downswing; FT – Follow-through. Bold – Higher than 85%.

7.4. Discussion

This study intended to identify the quantitative power accuracy from RQA features to classify Hc and LBP, using SVM as classifier, in critical phases of golf swing. Analyzing the discriminatory power between muscles due to several constrains gives information about neural coordination.

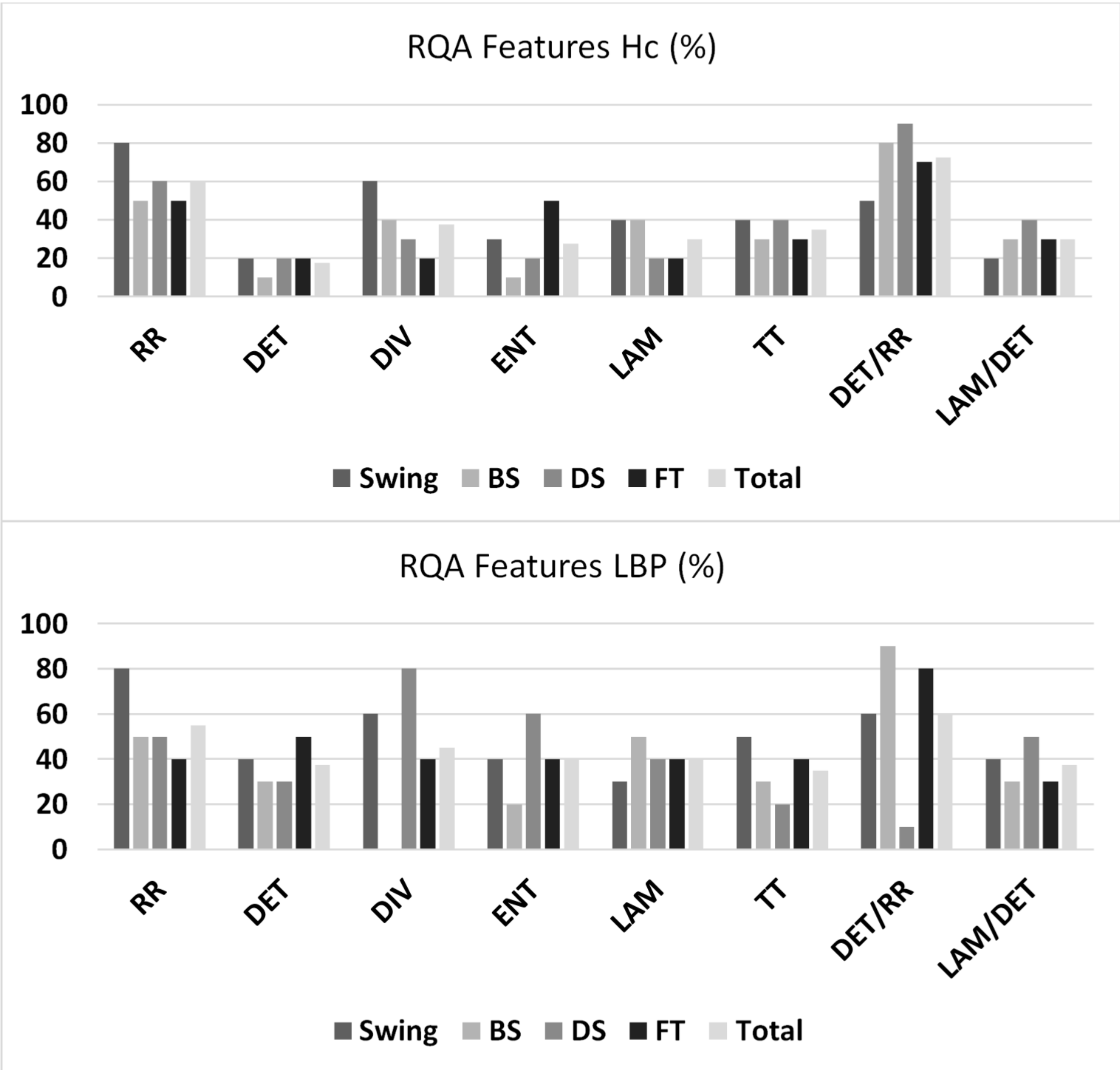


*Figura 18 [**Fig. 2**. Correlation-based feature selection for Hc and LBP].*

Hc – Handicap; LBP – Low back pain; RR – Recurrence rate; DET – Determinism; DIV – Divergence; ENT – Entropy; LAM – Laminarity; TT – Trapping time; BS – Backswing; DS – Downswing; FT – Follow-through.

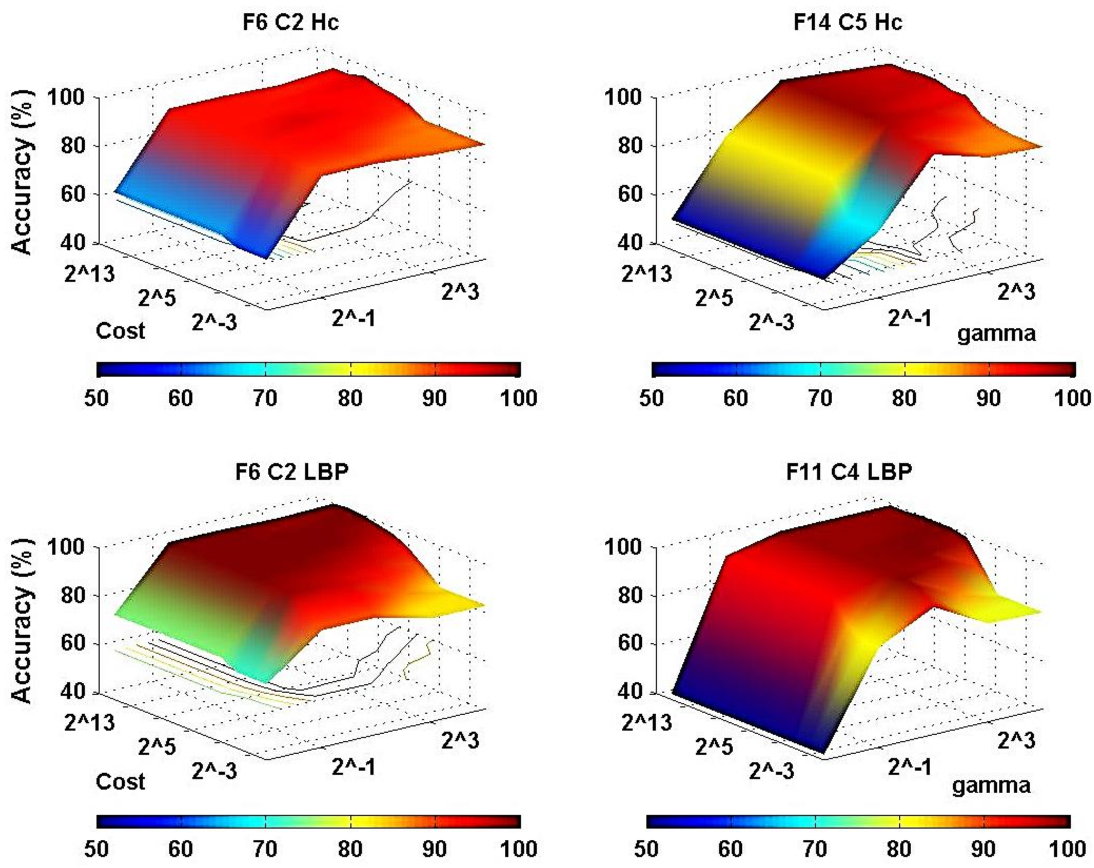


*Figura 19 [**Fig. 3.** Grid Search for different number of features and channels by Hc and LBP].*

Hc – Handicap; LBP – Low back pain; F number– Number of features; C number – Number of channels.

7.4.1. Features selection

Features are vectors that contain information, which enables a (better or worse) classification. Therefore, it is a critical aspect. The recurrence plot visualization of the rectified EMG signal for each phase is similar to a mixture between drift and disrupted non-stationary state. Two issues related to dimensionality must be present: (1) the features with better representation for each muscle (features dimensionality) (2) the muscles that best discriminate each phase (channels).

The increase of accuracy when the number of channels was equal or greater than three are similarly to literature (Tavakolan et al., 2011), besides different purposes of classification. The use of four channels seems suitable to classification problems based on EMG. In the present study, the LBP group achieved 100% with two and four channels during the backswing. Sultornsanee et al. (2011) also found accuracies of 100% with RQA features but only using RR, DET and LAM to classify neuromuscular disorders. It is believed that the classification of neuromuscular disorders may be easier due the myogenic and neurogenic differences. The myogenic have less membrane potentials due to atrophy (short-lasting and high amplitude record), and neurogenic due to the loss of axons (long-lasting and high-amplitude record) (Dobrowolski et al., 2012).

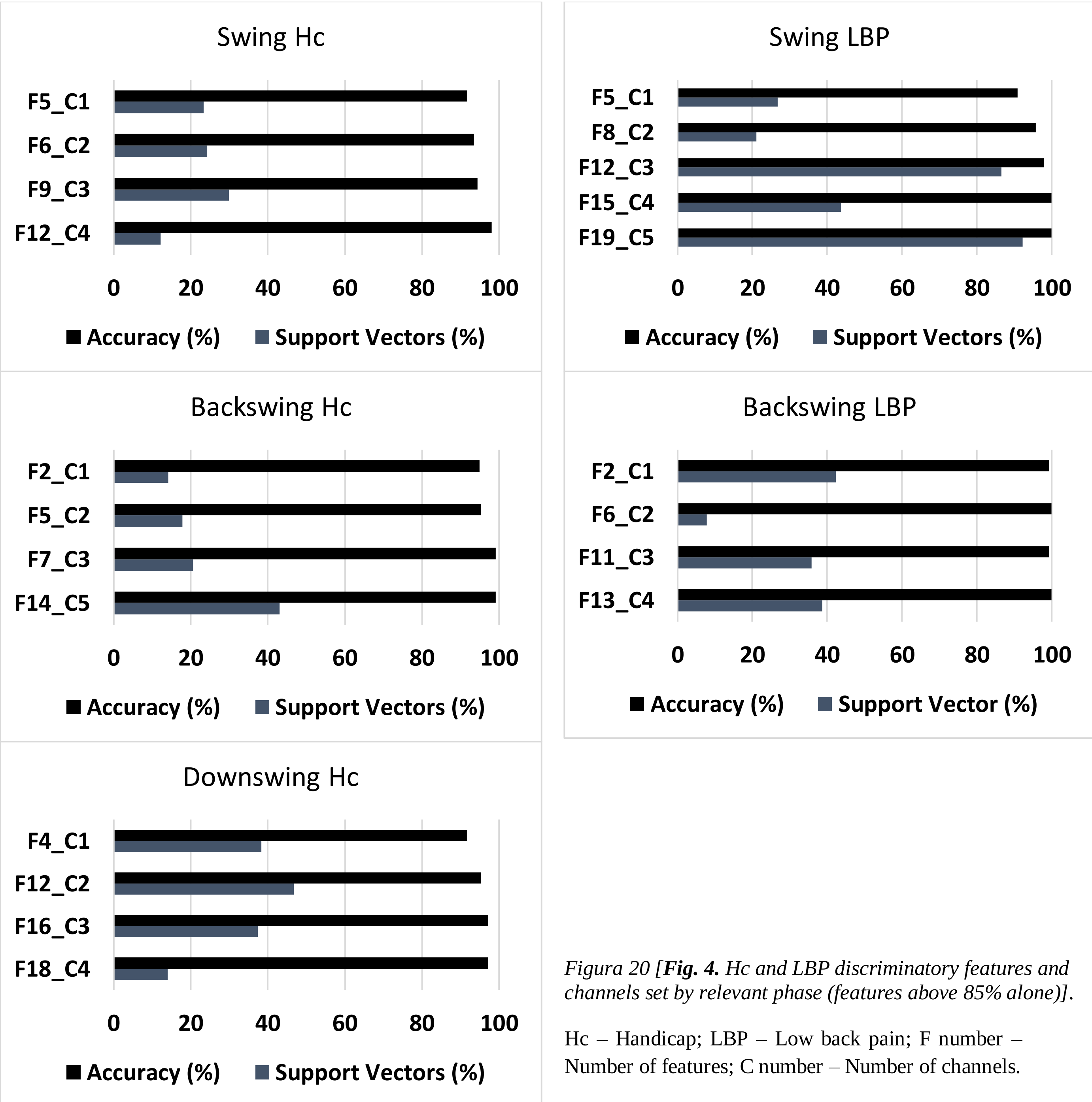


*Figura 20 [**Fig. 4.** Hc and LBP discriminatory features and channels set by relevant phase (features above 85% alone)].*

Hc – Handicap; LBP – Low back pain; F number – Number of features; C number – Number of channels.

LBP discrimination had the LAM in 50% of the models. The DET did not show an elevated contribution when used alone. On the other hand, the DET/RR had a mean value of 66.3%, achieving 90% of the SVM models, namely in the downswing for Hc and backswing for LBP. The DET has been showing good results in studies about motor units synchronization and changes determined by fast transitions of force production (Farina et al., 2002; Filligoi and Felici, 1999). The set of features with RR, DET, DIV means that recurrence density and the information based on diagonal lines are the most important to discriminate Hc and LBP muscle activity. In trails with periodic-chaos/chaos-periodic transitions the measures based on diagonal structures increase when vertical measures shows highly drop (Marwan et al., 2007).

7.4.2. Handicap

The major handicap discriminatory muscles during all swing were the RF, EO and GM from left side, and the right BF. The total swing is composed by a number of different motor actions for the same muscle. Accordingly, the ability of the muscles to discriminate the Hc group when performing various motor actions in same task gives information about different strategies due the skill level. Those muscles contain several EMG information relating muscle and load torques, force-velocity, and force-length relationships (Enoka, 1996) during the swing.

The top of the backswing could be the primary position that justifies these results, since there is an elastic energy transferring to downswing and then maximizing the impact on the ball. An increase in ground reaction force will occur in the trail foot (right foot in right handed golfers) during the backswing, followed by a transferring of weight to the lead foot (left foot in right handed golfers) during the downswing phase (Hume et al., 2005). In the top of backswing, the ground reaction force achieves mean values of 64.5% of body weight in the trail foot and 29% in the lead foot (Chu et al., 2010). The left EO is axial rotator of the trunk to the right during the backswing, stretching the hip and the trunk muscles, which corresponds to an increase of the angle between the pelvis and upper torso segments referred to as X- Factor (Cheetham et al., 2001). Increasing the X-Factor at top of backswing should facilitate a high club head speed at impact. However, the reason behind the power of discrimination of these muscles for all swing could be also related with the activity during the early downswing. The X-Factor in the top of backswing and the maximum that occurs at the early stage of downswing ("X-Factor Stretch") is 11% higher in highly skilled golfers than the less-skilled (Cheetham et al., 2001). This means that LHc golfer's pelvis turns back towards the ball while the right shoulder continues to rotate to the right. So, during the downswing, the left EO activity increases (Ashish et al., 2008), remaining high close to the impact (Ashish et al., 2008; Silva et al., 2013).

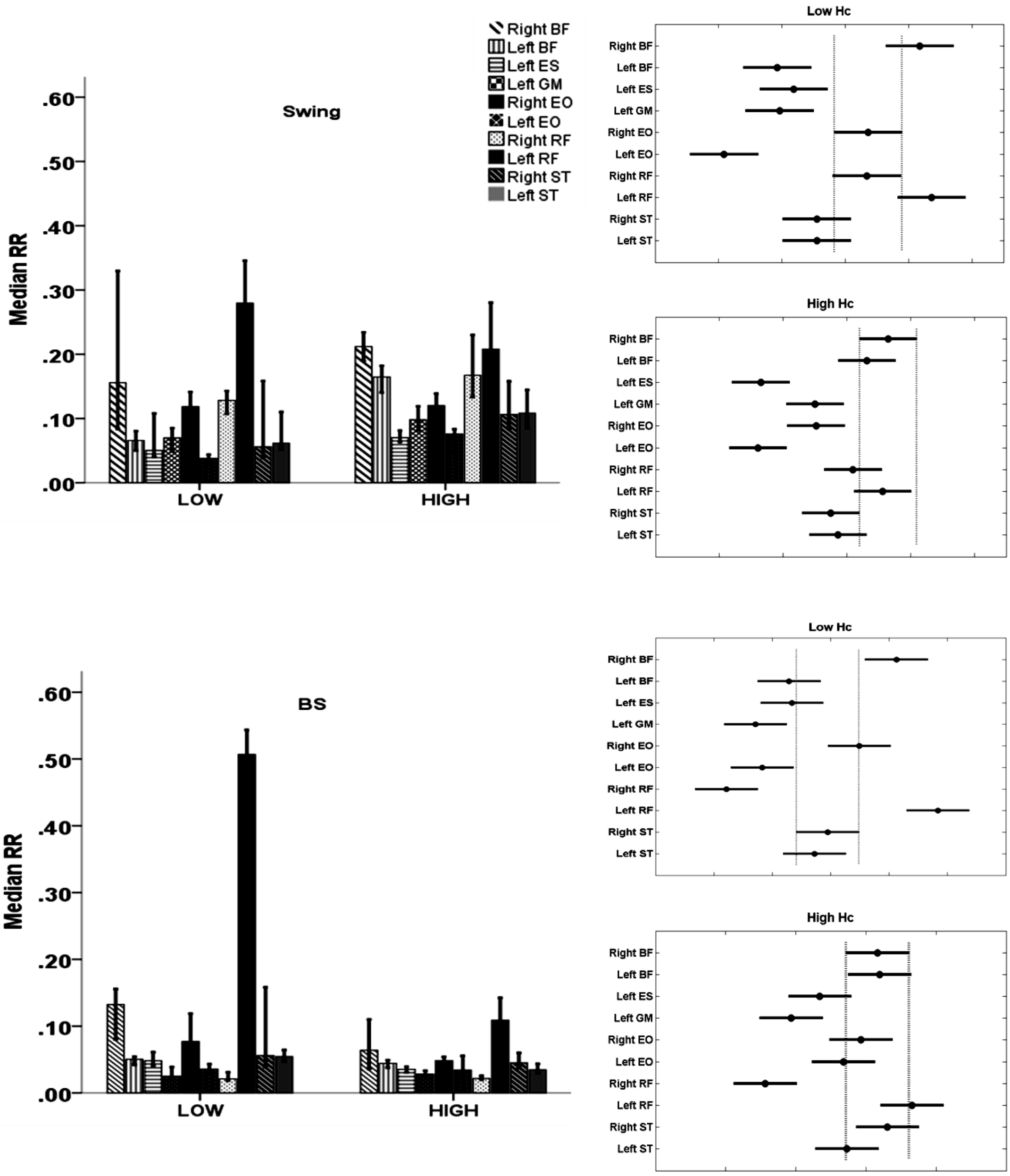


*Figura 21 [**Fig. 5.** Hc recurrence rate in the swing and backswing. Error bars 95% C.I.].*

The backswing and the downswing also show a high capacity to discriminate the Hc, showing the first higher accuracy. When this phase was isolated from the rest of the swing, the right EO and the hamstrings assume relevance and not the EO of the left side. The purpose of backswing is to align and position the golfer's hub center and club head to boost to enhance the accuracy and speed in the downswing (Hume et al., 2005). At the peak of the backswing, LHc golfers exhibit higher left shoulder horizontal adduction, right shoulder external rotation, and trunk rotation than HHc (Zheng et al., 2008). During the downswing, LHc golfers have the largest angular velocities for the club shaft, right elbow and writs extension, as well as, angular displacement in trunk rotation (Zheng et al., 2008). Muscle change in the capacity to discriminate handicap expresses both different nervous strategies due the complexity of the task and the amount of information that combine eccentric and concentric contractions called stretch-shorten cycle.

7.4.3. Low back pain

The backswing is the phase where there is a greater discriminatory power to distinguish between LBP and NLBP golfers, achieving 100% of accuracy. When looking muscles alone, the left BF showed the higher accuracy at 99.3%. The BF arises from the ischial tuberosity and is responsible for the hip extension, knee flexion and participates in pelvic retroversion. Motor strategies program to facilitate these movements during the backswing include left thigh extension, the lead side of right handed golfers. LBP golfers tended to flex their spines more when addressing the ball, showing greater left side bending during the backswing (Lindsay and Horton, 2002). Also, they have less trunk extension strength at 60°/s and left hip adduction strength, as well trunk rotation angle toward the trail side (Lindsay and Horton, 2002; Tsai et al., 2010). During the backswing, left hamstrings, left RF and right BF were the more discriminatory muscles, participating in extension and flexion of the thigh and hip. The neuromuscular strategies are limited by muscle length, since elite golfers with reduced hip flexor length reported more often that golf was affected by LBP (Evans et al., 2005). Golfers with history of LBP have limitations in the lead hip rotation and lumbar spine extension (Vad et al., 2004). However, the downswing did not show the same relevance that swing and backswing, i.e., for an accuracy above 85%, suggesting that for LBP groups, X-Factor could be more crucial than the “X-Factor Stretch”.

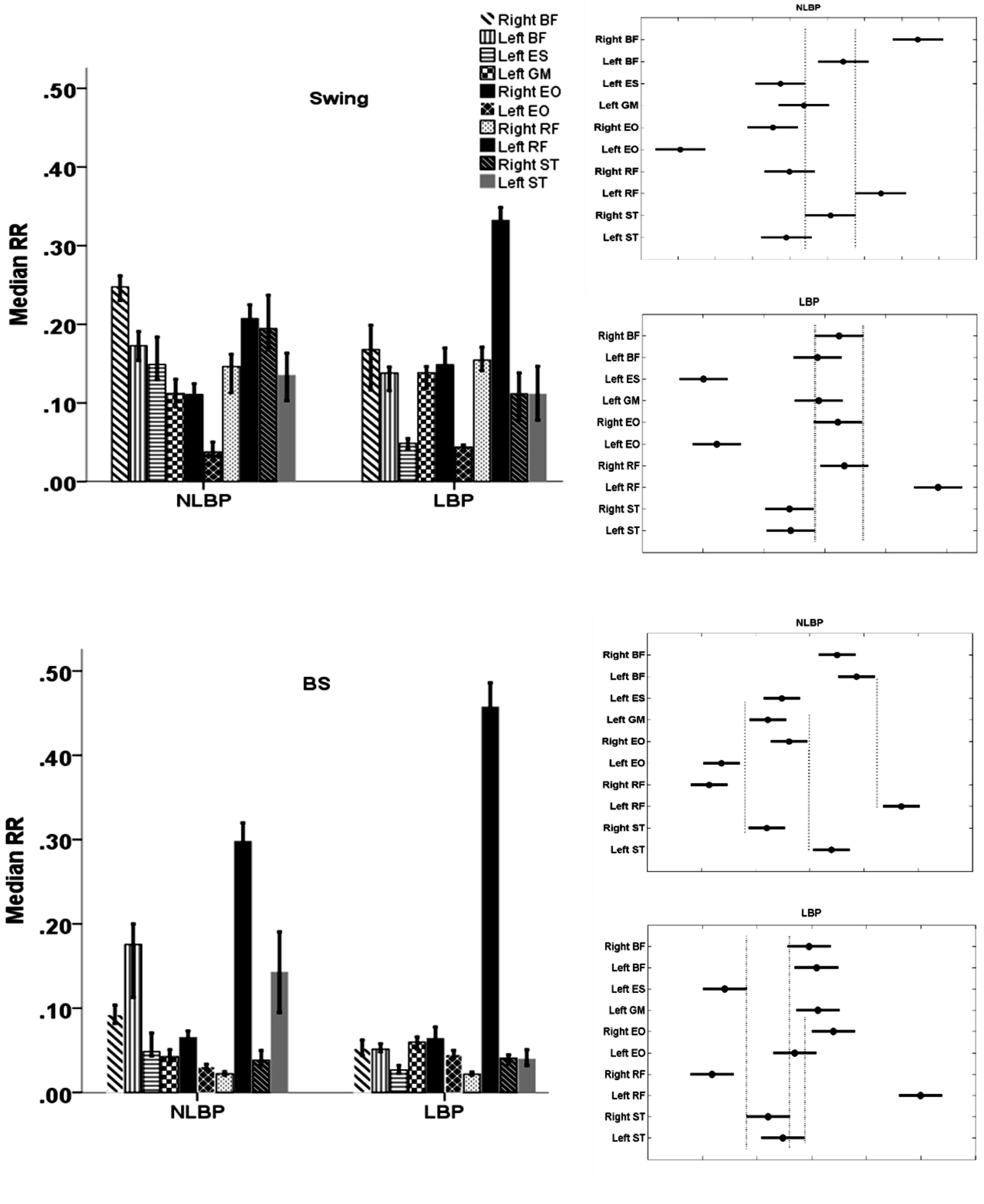


*Figura 22 [**Fig. 6.** LBP recurrence rate in the swing and backswing. Error bars 95% C.I.].*

Contrary to what one would expect due the neuromuscular imbalance of ES in LBP subjects (Renkawitz et al., 2006), left ES did not show a discriminative power comparatively with the trunk rotators and thigh muscles. Cole and Grimshaw, (2008b) found that the LHc LBP golfers present lower ES activity than the LHc NLBP. Also, the

LBP golfers activate ES early than NLBP players (Cole and Grimshaw, 2008a). Besides these results, the right ES presented higher activity than the left side (Cole and Grimshaw, 2008b), namely in the early downswing. However, in the present study, the right ES was not monitored. The crunch factor (product between lateral bending and axial trunk rotation) is another parameter associated to LBP, although no significant differences between LBP with NLBP were found (Cole and Grimshaw, 2013).

As for Hc, the follow-through phase revealed no power to discriminate the two classes (LBP and NLBP), for any muscle, despite the fact that more swing-related injuries were reported (McHardy et al., 2007). Discrimination strategies in neuromuscular coordination were more prevalent during the backswing.

7.5. Conclusions

The recurrence rate and the ratio between recurrence ratio and determinism are the two RQA features that showed higher discriminating power in SVM, for both for handicap and low back pain during the swing. The swing phase with higher discriminant power was the backswing for both handicap and low back pain. The hamstrings muscles, trunk rotators and left rectus femoris showed greater discriminant power for handicap, depending on the swing phase studied. The hamstrings and left rectus femoris showed higher discrimination for low back pain, and the lead side offers better major accuracy classification than the trail side, despite the downswing being the execution phase. Low back pain golfers show different neuromuscular coordination strategies than the golfers without low back pain. None of the muscles showed discriminative power during the follow-through, in both handicap and low back pain golfers.

Recurrence quantification analysis and support vector machines showed a high performance, but the type of features extraction is not consistent between muscles, being muscle dependent for electromyography processing and normalization, as well as embedding parameters used in this study. The authors recommend that further studies examine the weight of different processing types and embedding parameters on the accuracy to discriminate different electromyography signals by groups. It is also recommended to combine kinematics and kinetics with electromyography to understand the relationship of those variations with motion parameters.

7.6. Practical applications

In order to prevent low back pain, exercise and health professionals should include rectus femoris and hamstrings exercises in golfers exercise programs, particularly on the lead lower limb. Exercise professionals should also be aware that backswing is a critical phase for lower back injury episode.

When both exercise and health professionals intend to improve performance, the exercises prescription should comprise rectus femoris and hamstrings from the lead lower limb and external oblique from the trail lower limb. Similar to low back pain prevention, the backswing is the phase that shows higher values of discrimination.

Machine learning and recurrence analysis features show a higher performance while identifying subjects with low back pain via EMG. Therefore, such methods may support and help clinicians to diagnose and exercise professionals to better assess sports performance.

Conflict of Interest

None declared

Acknowledges

The project “Neuromuscular activity in the golf swing with implications for the practice and in the prevention of overuse injuries” was supported by the Portuguese Foundation for Science and Technology (FCT) fund (PTDC/DES/105176/2008).

The authors want to thank to Professor Norman Marwan (PIK-Potsdam) for counseling.

References

Ashish A, Shweta S, Singh SJ. Comparison of lumbar and abdominal muscle activation during two types of golf swing: An EMG analysis. RICYDE. International Journal of Sport Science 2008;IV(12):59–71.

Boser BE, Guyon IM, Vapnik VN. A Training Algorithm for Optimal Margin Classifiers. In: Proceedings of the Fifth Annual Workshop on Computational Learning Theory 1992:144–152.

Burges C. A tutorial on support vector machines for pattern recognition. Data Mining & Knowledge Discovery 1998;2:121–167.

Cabri J, Sousa J, Kots M, Barreiros J. Golf-related injuries: a systematic review. European Journal of Sport Sciences 2009;9(6):353–66.

Castellini C, Smagt P van der. Surface EMG in advanced hand prosthetics. Biological Cybernetics 2009;100(1):35–47.

Castellini C, Smagt Pvan der, Sandini G, Hirzinger G. Surface EMG for force control of mechanical hands. In: IEEE International Conference on Robotics and Automation Pasadena, CA, USA, May 19-23, Pasadena 2008:725–730.

Cheetham PJ, Martin, PE, Mottram RE, Laurent, BFSt. The importance of stretching the "X-Factor" in the downswing of golf : The "X-Factor Stretch". In: Thomas PR, editor. Optimising Performance in Golf. Brisbane, Australia: Australian Academic Press, 2001:192–199.

Chu Y, Sell TC, Lephart SM. The relationship between biomechanical variables and driving performance during the golf swing. Journal of sports Sciences 2010;28(11):1251–9.

Cole M, Grimshaw P. Trunk muscle onset and cessation in golfers with and without low back pain. Journal of Biomechanics 2008a;41:2829–2833.

Cole MH, Grimshaw PN. Electromyography of the trunk and abdominal muscles in golfers with and without low back pain. Journal of Science and Medicine in Sport 2008b;11(2):174–81.

Cole MH, Grimshaw PN. The crunch factor's role in golf-related low back pain. Spine Journal 2013 Oct 12.

Cortes C, Vapnik V. Support-vector networks. Machine Learning. 1995;20(3):273–297.

Del Santo F, Gelli F, Schmied A, Vedel, J-P, Rossi A, Mazzocchio, R. Motor unit synchronous firing as revealed by determinism of surface myoelectric signal. Journal of Neuroscience Methods 2006;155(1):116–21.

Dobrowolski AP, Wierzbowski M, Tomczykiewicz K. Multiresolution MUAPs decomposition and SVM-based analysis in the classification of neuromuscular disorders. Computer Methods and Programs in Biomedicine 2012;107(3):393–403.

Eckmann J-P, Kamphorst OS, Ruelle D. Recurrence Plots of Dynamical Systems. Europhysics Letters 1987;4(9):973–977.

EGA. *Sistema de Handicap EGA 2012-2016*. Epalinges/Lausanne: Edição Federação Portuguesa de Golfe, 2012.

Enoka RM. Eccentric contractions require unique activation strategies by the nervous system. Journal of Applied Physiology 1996;81(6):2339–2346.

Evans K, Refshauge KM, Adams R, Aliprandi L. Predictors of low back pain in young elite golfers: A preliminary study. Physical Therapy in Sport 2005;6(3):122–130.

Farina D, Fattorini L, Felici F, Filligoi G. Nonlinear surface EMG analysis to detect changes of motor unit conduction velocity and synchronization. Journal of Applied Physiology 2002;93(5):1753–63.

Filligoi G, Felici F. Detection of hidden rhythms in surface EMG signals with a non-linear time-series tool. Medical Engineering & Physics 1999;21(6-7):439–48.

Fox E, Lindsay D, Vandervoort A. Musculoskeletal injury questionnaire for senior golfers. In: Science and Golf IV. Proceedings of the World Scientific Congress of Golf. St. Andrews, Scotland: Routledge 2002:88–99.

Gatt CJ, Pavol MJ, Parker RD, Grabiner MD. Three-dimensional knee joint kinetics during a golf swing. Influences of skill level and footwear. American Journal of Sports Medicine 1998; 26(2):285–294.

Gluck GS, Bendo, JA, Spivak JM. The lumbar spine and low back pain in golf: a literature review of swing biomechanics and injury prevention. The Spine Journal 2008;8(5):778–88.

Güler NF, Koçer S. Use of support vector machines and neural network in diagnosis of neuromuscular disorders. Journal of Medical Systems 2005;29(3):271–284.

Gulgin H, Armstrong C, Gribble P. Hip rotational velocities during the full golf swing. Journal of Sports Science and Medicine 2009; 8:296-299.

Hall MA. Correlation-based Feature Selection for Machine Learning. University of Waikato, 1999.

Hermens HJ, Merletti, Freriks B. European Activities on Surface Electromyography. European Activities on Surface Electromyography. In: Proceedings of the first general SENIAM workshop. SENIAM project, 1996.

Horton JF, Lindsay DM, Macintosh BR. Abdominal muscle activation of elite male golfers with chronic low back pain. Medicine and Science in Sports and Exercise 2001;33(10):1647–54.

Hsu C, Chang C, Lin C. A practical guide to support vector classification. 2010; retrieved from https://www.cs.sfu.ca/people/Faculty/teaching/726/spring11/svmguide.pdf

Hume PA, Keogh J, Reid, D. The role of biomechanics in maximising distance and accuracy of golf shots. Sports Medicine 2005;35(5):429–49.

Ikegawa S, Shinohara M, Fukunaga T, Zbilut JP, Webber Jr CL. Nonlinear time-course of lumbar muscle fatigue using recurrence quantifications. Biological Cybernetics 2000;82:373–382.

Lei M, Meng G. Nonlinear Analysis of Surface EMG Signals. In: Lei M, Meng G, editors. Computational Intelligence in Electromyography Analysis - A Perspective on Current Applications and Future Challenges. INTECH 2012:119-174.

Lindsay D, Horton J, Vandervoort A. A review of injury characteristics, aging factors and prevention programmes for the older golfer. Sports Med 2000;30(2):89–103.

Lindsay D, Horton J. Comparison of spine motion in elite golfers with and without low back pain. J Sports Sci 2002;20(8):599–605.

Marwan N, Thiel M, Nowaczyk NR. Nonlinear Processes in Geophysics Cross recurrence plot based synchronization of time series. Nonlinear Process Geophys 2002; 2(9):325–331.

Marwan N, Carmenromano M, Thiel M, Kurths J. Recurrence plots for the analysis of complex systems. Phys Rep. 2007;438(5-6):237–329.

McHardy AJ, Pollard HP, Luo K. Golf-related lower back injuries: an epidemiological survey. Journal of Chiropractic Medicine 2007;6(1):20–6.

Renkawitz T, Boluki D, Grifka J. The association of low back pain, neuromuscular imbalance, and trunk extension strength in athletes. The Spine Journal 2006;6(6):673–83.

Shi J, Zheng Y, Chen X, Xie H. Modeling the relationship between wrist angle and muscle thickness during wrist flexion-extension based on the bone-muscle lever system: a comparison study. Medical Engineering & Physics 2009;31(10):1255–60.

Silva L., Castro MA, Marta S, Almeida J, Vaz J, Cabri J, Pezarat-Correia P. Cross-cultural adaptation and validation of the Portuguese Survey of musculoskeletal conditions, playing characteristics and warm-up patterns of golfers. Submitted to Journal of Back and Musculoskeletal Rehabilitation 2014.

Silva L, Marta S, Vaz J, Fernandes O, Castro MA, Correia-Pezarat P. Trunk muscle activation during golf swing: Baseline and threshold. Journal of Electromyography & Kinesiology 2013;23(5):1174–1182.

Sultornsanee S, Zeid I, Kamarthi S. Classification of Electromyogram Using Recurrence Quantification Analysis. Procedia Computer Science. 2011;6:375–380.

Tavakolan M, Xiao ZG, Menon C. A preliminary investigation assessing the viability of classifying hand postures in seniors. Biomed Eng Online 2011;10(1):79.

Tolambiya A, Thomas E, Chiovetto E, Berret B, Pozzo T. An ensemble analysis of electromyographic activity during whole body pointing with the use of support vector machines. PLoS One. 2011;6(7):e20732.

Tsai Y-S, Sell TC, Smoliga JM, Myers JB, Learman KE, Lephart SM. A comparison of physical characteristics and swing mechanics between golfers with and without a history of low back pain. The Journal of Orthopaedic and Sports Physical Therapy 2010;40(7):430-438.

Vad VB. Low Back Pain in Professional Golfers: The Role of Associated Hip and Low Back Range-of-Motion Deficits. American Journal of Sports Medicine 2004;32(2):494–497.

Vapnik VN. An Overview of Statistical Learning Theory. IEEE Transactions on Neural Networks 1999;10(5):988–999.

Webber Jr CL, Zbilut JP. Dynamical assessment of physiological systems and states using recurrence plot strategies. Journal of Applied Physiology 1994;76(2):965-973.

Zheng N, Barrentine SW, Fleisig GS, Andrews JR, States U. Kinematic analysis of swing in pro and amateur golfers. International Journal of Sports Medicine 2008;29(6):487–93.

Zhao Z, Morstatter F, Sharma S, Alelyani S, Anand A, Liu H. Advancing feature selection research-ASU feature selection repository. Arizona State University, 2010. Retrieved from http://www.public.asu.edu/~zzhao15/papers/TR-10-007%20Zhao,%20Morstatter,%20Sharma,%20Aleyani,%20Anand,%20and%20Liu.pdf

Ziai A, Menon C. Comparison of regression models for estimation of isometric wrist joint torques using surface electromyography. Journal of Neuroengineering and Rehabilitation 2011;8(1):56.

CAPÍTULO VIII

8. Discussão Geral

O objetivo desta dissertação parte de paradigmas na análise do comportamento neuromuscular quanto à quantificação de parâmetros eletromiográficos, nomeadamente no domínio do tempo. O primeiro ponto refere-se à deteção de momentos que podem conter informação relevante na descrição de determinado fenómeno motor, sendo o principal parâmetro o *onset* de ativação muscular. Através da literatura constatou-se uma controvérsia, pois está subjacente a utilização de algoritmos automáticos em detrimento da inspeção visual por esta ser subjetiva, quando a capacidade de análise desses algoritmos tende a ser confirmada por inspeção visual. O segundo ponto refere-se à reprodutibilidade entre estudos, mais comprometida nos métodos automáticos de limiar (Staude et al., 2001). Estes dois aspetos conduzem a consequências que são por si só igualmente paradigmáticas. A primeira consequência refere-se à ênfase em estudos metodológicos, ficando carenciada a aplicação destes métodos em estudos exploratórios. Veja-se o exemplo do golfe em relação a parâmetros temporais (Cole & Grimshaw, 2008a; Horton, Lindsay, & Macintosh, 2001) ao pretender analisar um dos principais constrangimentos na prática da modalidade, a lombalgia. Como a literatura tende por abordar essencialmente a questão metodológica (ex.: Hodges & Bui, 1996; Solnik, Rider, & Steinweg, 2010; Staude et al., 2001; Vaisman, Zariffa, & Popovic, 2010) surge outra lacuna. Os métodos de deteção do *onset* são testados em contrações isométricas em detrimento de ações dinâmicas, quando se pretende inferir sobre ações dinâmicas. Perante tais paradigmas, sendo o sinal uma série temporal, como retirar informação que permita a discussão de diversos fenómenos fisiológicos discriminando constrangimentos motores aliados a determinada habilidade motora dinâmica (neste caso o *swing*)? A resposta a esta questão fornece informação sobre estratégias neuromusculares subjacentes aos programas motores. Esta carência de investigação está bem evidenciada em McGill, Chaimberg, Frost e Fenwick (2010) no estudo do fenómeno do duplo pico em movimentos das artes marciais. Os autores descrevem o fenómeno através da sincronização de fotografias e o registo da atividade elétrica.

Face ao exposto surgiam questões pertinentes como quais os constrangimentos que devem assumir relevância associados à prática de golfe e como quantificá-los. A primeira questão foi facilmente respondida pela literatura referindo o handicap (Cabri et al., 2009; Zheng et al., 2008) e a lombalgia (Cabri et al., 2009; Evans, Refshauge, Adams, & Aliprandi, 2005; Gluck, Bendo, & Spivak, 2008; McHardy, Pollard, & Luo, 2007; McHardy & Pollard, 2005; Tsai et al., 2010) como constrangimentos a estudar. O problema colocava-se então quanto à quantificação destas variáveis. O handicap é um parâmetro bem definido pelos golfistas, pelo que se utilizou o Sistema de Handicap da *European Golf Association* (2012). A quantificação da lombalgia já não seria tão pacífica, pois a mesma pode estar associada às mais diversas causas. Por este motivo, emergiam dois aspetos tidos como relevantes. O primeiro refere-se a quantificar lombalgia de igual forma para todos os praticantes. O segundo é associar a lombalgia à prática de golfe. A incidência ou prevalência da lombalgia pode ser causada tanto no golfe como noutra prática quotidiana. Por sua vez, a prática de golfe pode agravar ou não a sintomatologia de lombalgia. Os estudos que abordam a lombalgia tendem a usar o *McGill Pain Questionnaire* (ex.: Cole & Grimshaw, 2008a; 2008b; Horton et al., 2001), mas desta forma, a sintomatologia não é associada à prática desportiva. Esta questão já tinha sido previamente assinalada, nomeadamente quanto à população golfista sénior (Fox et al., 2002), pelo que surge o primeiro estudo desta dissertação, a validação cultural do respetivo questionário para a população portuguesa, visto este questionário ser originário do Canadá.

8.1. Caracterização e Quantificação de Constrangimentos no Golfe

O primeiro artigo (Capítulo III), apesar de ter como objetivo a validação cultural para a população portuguesa, acaba por retratar a lombalgia e a incidência de lesões. Igualmente, promoveu-se um instrumento que permite caracterizar a população golfista em Portugal, dentro das limitações amostrais inerentes. O ICC apresentou valores superiores a 0.9 para 44.5% das variáveis e superior a 0.7 para 72.5%, muito próximo do original (Fox et al., 2002). Apesar desta similaridade, os portugueses gastam mais tempo com tarefas de aquecimento quando comparados com os canadianos, o que pode justificar o facto da incidência de lesões ter sido inferior no presente estudo quando comparado com outras aplicações do questionário no Canadá (Fox et al., 2002; Palmer, Young, Fox, Lindsay, & Vandervoort, 2003). Outra razão pode estar relacionada com a idade dos sujeitos que integram estes estudos, já que as versões original e a aplicação de Palmer et al. (2003) referem-se a uma população sénior, sendo que na versão portuguesa procedeu-se a um maior leque de idades. Um fator associado a maior prevalência de lesões em idades compreendidas entre os 50 e 65 anos refere-se à persistência na prática por longos períodos de tempo (Cabri et al., 2009). Igualmente, a população mais envelhecida tem maior probabilidade de lesões devido a alterações fisiológicas relevantes que se repercutem no decréscimo de força, da flexibilidade, coordenação e equilíbrio. Além disso, muitos dos golfistas seniores já vivenciaram lesões anteriormente (Palmer et al., 2003). Assim, por mais que na revisão de literatura do Capítulo V tenha sido verificada a confiabilidade no reconhecimento de padrões e similaridade de parâmetros tendo em conta a idade (Tavakolan et al., 2011), convém promover a homogeneidade entre grupos de dados quando comparados face a diferentes constrangimentos. Desta forma, evitam-se as denominadas variáveis *confounding*.

Um dos pontos relevantes deste questionário, para além da caracterização da amostra, refere-se à quantificação de lombalgia. Não seria cordial ou ético solicitar aos golfistas que realizassem o *swing* durante um episódio agudo de lombalgia. Por outro lado, a quantificação de lombalgia era carregada de grande subjetividade, por dois motivos principais. O primeiro corresponde à causa da lombalgia, se derivada da prática da modalidade ou de alguma ocorrência externa à mesma. O segundo motivo prende-se com a influência da prática da modalidade na sintomatologia de lombalgia. Podiam ser recolhidas repetições de sujeitos com lombalgia em que esta seria agravada pelo *swing*, e outras em que, apesar da presença de lombalgia, a prática do golfe teria um efeito de alívio. O questionário ora validado vem resolver esta problemática, pois o que mede é a perceção do sujeito quanto à sintomatologia de lombalgia após realizar uma volta (um jogo) de golfe com 18 buracos (questão 36 do questionário). Desta forma, garante-se que a medida é similar para todos os praticantes. Juntamente com esta questão as restantes desse grupo de perguntas em que *gostaríamos de lhe perguntar sobre as lesões sofridas durante a prática de golfe*, tal como é indicado, permite realizar o histórico de lesões. Na validação do presente estudo ainda foi incluída uma questão que associa o tempo de paragem à lesão (questão 35.1).

Estando garantida a categorização dos golfistas já era possível continuar com o objetivo principal desta tese, a quantificação da informação no domínio do tempo do sinal

EMG em habilidades motoras dinâmicas, discriminando a atividade neuromuscular face a constrangimentos. Colocava-se novamente a questão da quantificação na qual teriam de ser considerados outro tipo de constrangimentos, estes relacionados com o sinal EMG.

8.2. Aprendizagem Automática e a Deteção do *Onset*

Sabendo que as principais influências na determinação de parâmetros temporais, especificamente do *onset* são a linha de base (Hodges & Bui, 1996; Lee et al., 2007; Staude et al., 2001) e o incremento do declive da amplitude do sinal (Allison, 2003), procedeu-se à aplicação de algoritmos de limiar. Este método tem a vantagem de possibilitar a construção de restrições no algoritmo tendo em conta as diferentes fases do movimento. Então, apesar de considerarem o mesmo desvio-padrão, no estudo 2 (Capítulo IV) são consideradas linhas de base diferentes (próprio sinal e repouso entre duas contrações máximas), sendo posteriormente associadas restrições de procura antes do *backswing* e do *downswing*, as fases de preparação e de execução do *swing*, respetivamente. Neste estudo é lançado um novo conceito de *onset*, o *onset peak*. Este parâmetro temporal pode ser definido como qualquer *onse*t que antecede um pico de atividade tido como relevante. Esta análise suporta a informação retirada do sinal suavizado sabendo que excessos ou deficiências neste processamento podem conduzir a desvios (Hodges & Bui, 1996; Van Boxtel et al., 1993), assim como, uma maior precisão na descrição e discussão sobre o fenómeno neuromuscular em estudo.

Uma das inspirações a esta abordagem foi o estudo de McGill et al., (2010). Procura-se descrever momentos de ativação relevantes face às fases do *swing*, procurando fundamentação para uma abordagem fenomenológica recorrendo a um método tido como tradicional. Foi descrita a congruência entre métodos de limiar com referências de base diferentes e a inspeção visual, apesar de não resolver a procura de diferentes fenómenos temporais, nem identificar diferentes estratégias no recrutamento neuromuscular subjacentes a diferentes constrangimentos da tarefa, como o Hc e a lombalgia.

Ambos os artigos que estudaram o *onset* no *swing* usaram um algoritmo de deteção de limiar, tendo focado a análise em músculos do tronco e na comparação de sujeitos com e sem lombalgia. Logo, neste ponto verifica-se uma divergência no limiar utilizado, 7 (Horton et al., 2001) e 1 desvios-padrão (Cole & Grimshaw, 2008a) acima da linha de base. No estudo 2 optou-se pela aplicação de três desvios-padrão fixos em ambos os métodos. Visto a literatura não focar com ênfase o jogador recreativo mediano (Marta et al., 2012), aplicou-se aos sinais recolhidos de sujeitos desta população, o algoritmo de limiar com as restrições evidenciadas no Capítulo IV.

No estudo 2 (Capítulo IV) verifica-se que a congruência entre os três métodos (linha de base da própria repetição, linha de base entre dois MVC, inspeção visual) varia mediante o músculo em causa, não havendo necessariamente alguma tendência de um método em relação ao outro. Esta informação corrobora a pressuposição de que o sinal EMG contém informação específica sobre a função de cada músculo naquela tarefa podendo a mesma ser quantificada. No caso do reto abdominal, o *onset* é similar nos três métodos e bilateralmente (Direito – ICC = 0.958; Esquerdo – ICC = 0.957). Outro músculo com congruência semelhante foi oblíquo externo do lado direito (ICC = 0.906). Porém, para o primeiro caso, o *onset* determinado é o *onset peak* quando este é similar ao *onset burst* (na maioria das repetições) por ter sido quando iniciou atividade relevante. No caso do oblíquo externo foi sempre considerado o *onset burst.* Significa que para cada um destes músculos, por mais relações *onset/offset* que sejam pretendidas para estudo, existem aquelas que são mais relevantes consideradas como padrão (ver figura 5 [figura 2 artigo 2]). Quando focamos a análise na massa comum ou *erector spinae* verifica-se uma dificuldade que contrasta com o indicado por Cole e Grimshaw (2008a), onde a massa comum na região lombar chega a ativar antes do *backswing* em jogadores com dor lombar. No presente estudo, denota-se uma dificuldade em determinar o *onset burst*, pois o músculo já se encontra ativado durante o *address* (início do *backswing*) devido à flexão do tronco nesse momento. Para este grupo muscular (indiferentemente da lateralidade) observa-se um momento *off-onset* cerca de 300 ms antes do impacto que ocorre a ±5% da amplitude do pico máximo. Apesar da baixa congruência entre métodos (Direito – ICC = 0.374; Esquerdo – ICC = 0.277), por todos foi detetado este *onset peak*. A maior ou menor proximidade de valores (em ms) foi influenciada pelo incremento da amplitude do sinal como referido na literatura (Allison, 2003). Neste caso, a utilização de atividade em repouso que não da própria habilidade motora pode ser considerada em algoritmos de limiar, pois uma elevada atividade de linha de base de EMG é comum em tarefas posturais, levando a um atraso na deteção do *onset* comparativamente à inspeção visual (Hodges & Bui, 1996).

Outra contribuição deste estudo deve-se à forma como os limiares costumam ser construídos no que respeita à percentagem do pico máximo, sendo referidos intervalos de 15 a 25% (Hug, 2011). Neste estudo foi realizada a situação inversa, aplicou-se um algoritmo de limiar e posteriormente verificou-se a que percentagem do pico máximo correspondia. Veja-se o caso da massa comum, o músculo que apresentou maior dificuldade na deteção do *onset burst* devido ao baixo rácio sinal-ruído (ou linha de base já em atividade). Uma proporção mais elevada do rácio sinal-ruído conduz a uma melhor deteção e concordância entre métodos, enquanto o inverso conduz à situação contrária (Staude et al., 2001). Como já houve oportunidade de evidenciar, no *onset peak* da massa comum a deteção foi a cerca de 5%, porém, a deteção do *onset burst* já foi realizada acima de 20% do pico máximo. No caso do reto abdominal, verificamos a deteção por volta dos 8% no lado direito e 18% no lado esquerdo. No oblíquo externo entre os 6 e os 12% da amplitude máxima. Face às comparações encontradas na literatura (Staude et al., 2001; Van Boxtel et al., 1993), podemos referir que nem sempre os artigos se referem ao mesmo fenómeno *onset*. Esta indefinição conduz ao desprezar de fenómenos temporais ao longo do sinal e dificulta a reprodutibilidade entre estudos. Quando os músculos numa habilidade motora como o *swing* evidenciam comportamentos padrão e com elevado grau de sincronização como se pode verificar pelo *onset peak* e instante do pico máximo nos vários músculos antes do impacto, este fenómeno não deve ser ignorado. Realça-se a quantificação de padrões em momentos de transição de uma fase de movimento para outra (Cordo et al., 2003) em relação ao comportamento do sinal.

Porém, no estudo 2 recorremos à inspeção visual como os demais estudos, sinal por sinal para testar os resultados, não resolvendo o problema da automatização e reprodutibilidade entre métodos (Jöllenbeck, 2000; Morey-Klapsing et al., 2004; Vaisman et al., 2010). Face à necessidade de detetar instantes relevantes no domínio do tempo e diferenciar estratégias neuromusculares resultantes de um constrangimento surge o pensamento de que o ideal seria o investigador interferir em determinada altura do processo para que o algoritmo soubesse as características do fenómeno a estudar de forma a classificá-lo. Apesar de até então, a determinação do *onset* não ser vista como o problema no reconhecimento de padrões (Oskoei & Hu, 2007) esse pensamento definia a aprendizagem automática supervisionada. Isto é, construir previamente um modelo para que depois a máquina (computador) conheça o fenómeno a estudar, possibilitando a reprodutibilidade entre fenómenos e a sua discussão.

Surge o terceiro estudo, a revisão de literatura onde foi verificado que as SVM apresentavam uma elevada performance em tarefas de classificação comparativamente com outras técnicas matemáticas (Khushaba, Kodagoda, Takruri, & Dissanayake, 2012; Oskoei & Hu, 2008). Desta revisão de literatura são identificadas três grandes categorias de aplicação das SVM a EMG: controlo mio-elétrico, discriminação de desordens neuromusculares e análise cinesiológica (esta última intimamente ligada à primeira em muitos estudos). Estudos sobre controlo mio-elétrico visam reconhecer padrões neuromusculares através de EMG levando uma prótese ativa a realizar tarefas com a maior precisão possível. A maioria dos estudos incide sobre a musculatura do antebraço havendo uma clara preferência por EMG de superfície. Os distúrbios neuromusculares estão associados a falhas das unidades motoras, sendo importante para identificar as causas da doença como deficiências miogénicas e/ou neurogénicas. As formas de onda de cada potencial de ação das unidades motoras (MUAP) devem ser suficientemente discriminativas para permitir a determinação do tipo de patologia. O registro miogénico mostra uma curta duração, baixa amplitude e potencial multifásico devido à atrofia, que correspondem a menos potenciais de membrana. Os MUAPs dos neurogénicos podem ser caracterizados como de alta amplitude de longa duração, mas também multifásico, pois esta patologia leva à perda de axônios e neurónios motores (Dobrowolski et al., 2012). A última categoria obtida refere-se ao reconhecimento de alterações no sinal de EMG que podem ser associados à realização do movimento. Essencialmente, a maioria dos estudos procurava discriminar movimentos com vários canais de EMG, à semelhança do controlo mio-elétrico. Porém, é de realçar a classificação de estratégias neuromusculares em vários níveis de esforço no membro inferior durante a corrida (Stirling et al., 2011), a classificação de movimentos tendo em conta diferentes escalões etários (Tavakolan et al., 2011) e a classificação de movimentos de apontar e alcançar quando sujeitos a constrangimentos motores (Tolambiya et al., 2011). Estes estudos evidenciavam bom prelúdio, mas não foram encontradas aplicações de SVM a diferentes fases do movimento realçando músculos com capacidade discriminatória entre constrangimentos populacionais como o Hc e a lombalgia. Igualmente, não se verificou a aplicação de SVM a EMG com o intuito de determinar o *onset* ou outro parâmetro fisiológico temporal.

Outro ponto fundamental foi perceber formas de retirar informação do sinal, ou seja, as *features* a considerar para aplicar SVM a EMG, com o intuito da classificação pretendida. Os dois pontos fundamentais após o tratamento do sinal (filtros, retificação, suavização, normalização) são a segmentação do sinal e a extração de *features* (Oskoei & Hu, 2008), como pode ser verificado pela Figura 2 da presente dissertação. Teria de considerar a inclusão de momentos tanto estacionários como de transição, quando as técnicas de segmentação no reconhecimento de padrões estão geralmente relacionadas com estados estacionários (Englehart & Hudgins, 2003; Oskoei & Hu, 2007). O procedimento de extração de *features* tinha que ser realizado a partir de vários segmentos do sinal de forma a preservar a sua estrutura e não a partir meramente de amostras individuais, o que podia conduzir a uma perda significativa de informação (Hudgins et al., 1993).

A aplicação das SVM alimentadas por *features* no domínio do tempo apresentou uma elevada performance ao atingir um valor médio de 95% de precisão, sem discriminar o tipo de fenómeno temporal. Foram considerados três conjuntos (não mutuamente disjuntos) de *features* identificados como F2, F4 e F6, consoante o número de *features* e tendo a sua construção como base, a ponderação obtida pelo *Fisher Score* (Duda et al., 2001) e pelo algoritmo *Correlation-based Feature Selection* (Hall, 1999). Os resultados apontam para uma utilização de quatro *features* como a dimensão necessária para uma boa classificação (tendo em conta as *features* no domínio do tempo utilizadas) semelhante ao verificado anteriormente pela literatura (Futamata et al., 2012; Tavakolan et al., 2011), apesar de F2 conter as *features* mais relevantes, o valor absoluto médio e o comprimento do formato da onda. Esta quantificação no domínio do tempo já tem sido referenciada tanto em relação ao valor absoluto médio (Oskoei & Hu, 2008; Tkach, Huang, & Kuiken, 2010) como ao formato do comprimento da onda (Oskoei & Hu, 2008). A desvantagem das *features* no domínio do tempo está associada ao facto de serem calculadas através da amplitude do sinal sendo incluídas interferências (Phinyomark et al., 2009). Porém, comparativamente às *features* no domínio da frequência, as *features* no domínio do tempo tendem a apresentar melhor poder de classificação (Oskoei & Hu, 2008; Phinyomark et al., 2013).

Os músculos que apresentaram maior precisão foram o semitendinoso, vasto interno, reto femoral, vasto externo e oblíquo externo. Para estes músculos, quando utilizados os conjuntos F4 e F6, a precisão subiu para 97.5±0.9% e 97.6±0.9%, respetivamente, quanto à classificação de instantes de tempo como de repouso ou ativação. O valor mínimo obtido foi de 96.1% atingindo uma precisão máxima de 99.04%.

Com exceção do oblíquo externo, todos estes músculos apresentam o *onset burst* coincidente com o *onset peak*, variando apenas consoante a inclinação de crescimento do sinal ou por existir alguma pré ativação. No oblíquo externo esquerdo não foi considerado nenhum pico principal de atividade tendo sido determinado sempre o *onset burst*, como anteriormente no estudo 2 (Silva et al., 2013). O aumento do número de amostras também não evidenciou uma descida significativa na classificação, tendo-se verificado que quanto menor o número de vetores suporte necessário melhor o resultado da classificação. Os músculos que apresentam melhor classificação também são aqueles em que a atividade na linha de base é diminuta tendo um rácio sinal-ruído elevado. O erro de classificação das diferentes *features* consoante o rácio sinal ruído aumenta (Phinyomark et al., 2009).

8.3. Classificação de Constrangimentos: Handicap e Lombalgia

Assume-se que ao identificar a capacidade discriminatória de cada músculo em determinado constrangimento, conhece-se pontos de atividade que serão críticos a nível neuromuscular. Ao analisar o poder discriminatório entre os músculos devido aos constrangimentos obtém-se informação sobre a coordenação neuromuscular. Esta forma de análise pode ser importante na reabilitação e prevenção de lesões associadas à lombalgia, como também, no desenvolvimento de estratégias de treino. Os músculos que apresentam maior poder discriminatório durante o *swing* foram o reto femoral, oblíquo externo e grande glúteo do lado esquerdo, e o bicípite femoral direito. Estes músculos serão aqueles que contêm maior informação quanto à variabilidade entre constrangimentos quanto ao momento de força, relação força-velocidade e relação força-comprimento (Enoka, 1996). O topo do *backswing*, que será o mesmo que mencionar o momento de transição do *backswing* para o *downswing*, surge como a primeira justificação para estes resultados. Existe nesse momento uma transferência de energia elástica para o *downswing* com o intuito de maximizar o impacto com a bola. Durante o *downswing*, irá então ocorrer um aumento da força de reação ao solo no pé direito (golfistas destros) seguido da transferência de peso para o pé de esquerdo (Hume et al., 2005).

O *backswing* foi a fase que apresentou maior relevância discriminatória tanto em relação ao Hc como à lombalgia. Será o mesmo que referir que é na fase de preparação onde encontramos maior diferenciação nas estratégias de recrutamento neuromuscular face a estes constrangimentos. Uma das variáveis mais apontadas que pode explicar a nível da cinemática a importância do que se passa no topo do *backswing* expondo o *downswing* como uma consequência da fase anterior é o *X-factor* e o "*X-factor strecht*". Um maior *X-factor* no topo do *backswing* facilita a transferência de energia através das respetivas cadeias cinéticas resultando no potenciar da velocidade da cabeça do taco no impacto com a bola (Cheetham et al., 2001). Porém, uma estratégia caracteriza os golfistas de baixo Hc que poderá dificultar em estudos cinemáticos a marcação do topo do *backswing* (caso não seja através do taco). Os jogadores de baixo Hc já iniciaram o regresso da pélvis na direção da bola quando os ombros ainda estão a rodar para o lado contrário, realizando maior adução horizontal e rotação externa do ombro esquerdo, assim como, maior capacidade na rotação do tronco (Zheng et al., 2008). Desta forma, o golfista de baixo Hc apresenta maior "*X-factor stretch*" comparativamente ao golfista de alto Hc.

Apesar do *downswing* ter demostrado capacidade discriminatória em relação ao Hc, tal não se verificou face à lombalgia. Durante o *backswing*, os músculos com maior poder discriminatório de lombalgia foram os posteriores da coxa do lado esquerdo, o reto femoral esquerdo e o bicípite femoral direito. Alterações no comprimento destes músculos podem estar associadas a cenários clínicos com sintomatologia de lombalgia (Evans et al., 2005). Contudo, golfistas com lombalgia tendem a flexionar mais o tronco no momento do *address* realizando flexão lateral para o lado esquerdo durante o *backswing*. No que concerne à massa comum, ao contrário do que seria de esperar devido ao desequilíbrio neuromuscular existente em sujeitos com lombalgia (Renkawitz et al., 2006), este grupo muscular não evidenciou poder discriminatório. Cole e Grimshaw (2008b) constataram que os jogadores de baixo Hc e com sintomatologia de lombalgia apresentam menor atividade na massa comum comparativamente aos de baixo handicap e assintomáticos. Seria de esperar também, uma ativação precoce da massa comum em golfistas com lombalgia (Cole & Grimshaw, 2008a).

Paralelamente a este tipo de análise é de salientar a elevada precisão na discriminação de constrangimentos como o Hc e a lombalgia aliando a análise de recorrência às SVM, chegando a um sucesso de 100%, mesmo com apenas duas *features*. Valores a este nível foram apenas verificados na discriminação de desordens neuromusculares (Sultornsanee et al., 2011), tendo em comum a utilização das *features* taxa de recorrência, determinismo e laminaridade. No presente estudo alargamos o leque de *features* oriundas da análise de recorrência, verificando o forte desempenho das *features* taxa de recorrência, determinismo e especificamente o rácio determinismo/taxa de recorrência. Esta informação demonstra que a informação contida nos gráficos de recorrência quantificada pela densidade de recorrência e pelas linhas diagonais contém elevada capacidade de discriminação deste tipo de constrangimentos a nível neuromuscular (Marwan et al., 2007), podendo ainda ser reforçado por outro tipo de informação.

8.4. Limitações

O problema apresentado nesta dissertação parte de limitações associadas à interpretação da informação retirada do sinal EMG, pelo que, para além daquelas já conhecidas e explanadas no enquadramento teórico, este ponto irá incidir também sobre questões inerentes ao processamento dos dados. Será sempre de realçar cuidados a ter com o *crosstalk*, pois acresce de preponderância quando se pretende estabelecer a capacidade discriminatória da atividade EMG de cada músculo face a constrangimentos motores. Motivo este que conduziu a uma maior abrangência deste fenómeno no enquadramento teórico.

Também associado à fase de recolhas está o elevado peso laboratorial que um estudo desta natureza envolve, comprometendo uma elevada adesão de participantes. Tal pode ser verificado pelo número de respondentes ao questionário (Capítulo III) em relação aos que participaram nas recolhas em laboratório, tendo ainda em conta que destes nem todos podem ser posteriormente utilizados. Quando se pretende comparar grupos específicos com diferentes níveis de handicap (baixo e alto) e a sintomatologia de lombalgia, a dimensão da amostra ainda se torna mais reduzida, visto serem considerados valores extremos e desconsiderados os restantes. Considera-se como limitação a dificuldade na aplicação de técnicas amostrais probabilísticas e logo a dimensão da amostra quanto a uma representatividade geográfica.

O processamento e a interpretação de resultados quanto a fenómenos temporais acabou por vincar o desenvolvimento desta dissertação, conduzindo o mesmo à aprendizagem automática. A principal limitação sentida refere-se ao custo computacional que os algoritmos de aprendizagem automática exigem no processamento de elevado número de amostras. O facto das amostras que servem de exemplos ao classificador serem instantes de vários sinais, como no estudo 4 (Capítulo VI), necessita de um consumo de tempo considerável. Porém, este processo tende a ser realizado apenas uma vez para cada construção de modelos, a classificação de novos sinais é relativamente rápida. Ao analisar outros métodos na deteção do *onset* foi possível verificar dificuldades em reproduzir estudos e logo em comparar resultados, assim como, o desprezar de outros fenómenos temporais tidos como relevantes. Ainda, para determinado músculo poderia estar a ser considerado um determinado momento como *onset*, divergindo este quando aplicado o mesmo algoritmo a outro músculo. Nesta dissertação foi realizada a construção de modelos de classificação por fenómeno tendo em conta aquele que é mais relevante para resolver esta problemática. No melhor do nosso conhecimento esta dissertação apresenta trabalho pioneiro na aplicação de aprendizagem automática a fenómenos temporais, sendo necessário estabelecer aperfeiçoamentos na eliminação de erros que aproximem os resultados de uma precissão na classificação de 100%. Estes erros são mais elevados quando o fenómeno a determinar é o *onset peak* com atividade muscular precedente considerável. Não se considera complexa esta eliminação de erros sendo posteriormente sugeridos procedimentos nas recomendações quanto a modelos de classificação. No entanto, deve ser uma limitação a ter presente após classificação do fenómeno temporal.

Apesar de ser possível na aprendizagem automática testar os melhores parâmetros na construção de modelos, a qualidade das *features* utilizadas compromete a classificação. Na análise de recorrência, os valores extraídos dependem dos parâmetros de entrada e de diferentes tipos de processamento (*embedding dimension*, *delay*). Se o ajustamento destes parâmetros fosse realizado individualmente a cada série temporal poderia ser afirmado que a escolha teria sido a ideal para cada sinal, pendendo a limitação de estarem a ser usados diferentes critérios para o conjunto dos sinais. Para evitar esta situação no estudo 5, os sinais foram submetidos a análise prévia por *mutual information* e *false nearest neighbors* sendo posteriormente realizada inspeção visual dos gráficos de recorrência na determinação dos parâmetros que garantissem a uniformização. Torna-se necessário ter presente a compreensão se o método matemático está a responder à questão biológica proposta ou se ao ser considerado determinado fenómeno biológico se está meramente a desregrar a utilização matemática.

Referências

Allison, G. T. (2003). Trunk muscle onset detection technique for EMG signals with ECG artefact. *Journal of Electromyography and Kinesiology*, *13*(3), 209–216. doi:10.1016/S1050-6411(03)00019-1.

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992) A Training algorithm for optimal margin classfiers. *Proceedings of the Fifth Annual Workshop on Computational Learning Theory* (pp. 144–152).

Cabri, J., Sousa, J. P., & Barreiros, J. (2009). Golf-related injuries: A systematic review. *European Journal of Sport Science*, *9*(6), 353–366. doi:10.1080/17461390903009141.

Cheetham, P. J., Martin, P. E., Mottram, R. E., & Laurent, B. F. St. (2001). The importance of stretching the "X-Factor" in the downswing of golf : The "X-Factor Stretch". In P. R. Thomas (Ed.), *Optimising performance in golf* (pp. 192–199). Brisbane, Australia: Australian Academic Press.

Cole, M., & Grimshaw, P. (2008a). Trunk muscle onset and cessation in golfers with and without low back pain. *Journal of Biomechanics*, *41*, 2829–2833. doi:10.1016/j.jbiomech.2008.07.004.

Cole, M. H., & Grimshaw, P. N. (2008b). Electromyography of the trunk and abdominal muscles in golfers with and without low back pain. *Journal of Science and Medicine in Sport / Sports Medicine Australia*, *11*(2), 174–81. doi:10.1016/j.jsams.2007.02.006.

Cordo, P. J., Gurfinkel, V. S., Smith, T. C., Hodges, P. W., Verschueren, S. M., & Brumagne, S. (2003). The situp: complex kinematics and muscle activity in voluntary axial movement. *Journal of Electromyography and Kinesiology, 13*(3), 239–52.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. doi:10.1007/BF00994018.

Dobrowolski, A. P., Wierzbowski, M., & Tomczykiewicz, K. (2012). Multiresolution MUAPs decomposition and SVM-based analysis in the classification of neuromuscular disorders. *Computer Methods and Programs in Biomedicine*, *107*(3), 393–403. doi:10.1016/j.cmpb.2010.12.006.

Duda, R., Hart, P., & Stork, D. (2001). *Pattern classification* (2nd Ed). New York: John Wiley & Sons.

EGA. (2012). *Sistema de handicap EGA 2012-2016*. Epalinges/Lausanne: Edição Federação Portuguesa de Golfe.

Englehart, K., & Hudgins, B. (2003). A robust, real-time control scheme for multifunction myoelectric control. *IEEE Transaction on Biomedical Engineering*, *50*(7), 848–854. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1206493.

Enoka, R. M. (1996). Eccentric contractions require unique activation strategies by the nervous system. *Journal of Applied Physiology*, *81*(6), 2339–2346.

Evans, K., Refshauge, K. M., Adams, R., & Aliprandi, L. (2005). Predictors of low back pain in young elite golfers: A preliminary study. *Physical Therapy in Sport*, *6*(3), 122–130. doi:10.1016/j.ptsp.2005.05.003.

Fox, E., Lindsay, D., & Vandervoort, A. (2002). Musculoskeletal injury questionnaire for senior golfers. *Science and Golf IV. Proceedings of the World Scientific Congress of Golf* (pp. 88–99). St. Andrews, Scotland: Routledge Inc.

Futamata, M., Nagata, K., & Magatani, K. (2012). The evaluation of the discriminant ability of multiclass SVM in a study of hand motion recognition by using SEMG. In *Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society* (Vol. 2012, pp. 5246–9). doi:10.1109/EMBC.2012.6347177.

Gluck, G. S., Bendo, J. A., & Spivak, J. M. (2008). The lumbar spine and low back pain in golf: a literature review of swing biomechanics and injury prevention. *The Spine Journal*, *8*(5), 778–88. doi:10.1016/j.spinee.2007.07.388.

Hall, M. A. (1999). *Correlation-based feature selection for machine learning*. (Doctoral dissertation). University of Waikato.

Hodges, P., & Bui, B. (1996). A comparison of computer-based methods for the determination of onset of muscle contraction using electromyography. *Electroencephalography and Clinical Neurophysiology*, *101*, 511–519.

Horton, J. F., Lindsay, D. M., & Macintosh, B. R. (2001). Abdominal muscle activation of elite male golfers with chronic low back pain. *Medicine and Science in Sports and Exercise*, *33*(10), 1647–1654.

Hudgins, B., Parker, P., & Scott, R. N. (1993). A new strategy for multifunction myoelectric control. *IEEE Transactions on Biomedical Engineering*, *40*(1), 82–94.

Hug, F. (2011). Can muscle coordination be precisely studied by surface electromyography? *Journal of Electromyography and Kinesiology*, *21*(1):1–12.

Hume, P. A, Keogh, J., & Reid, D. (2005). The role of biomechanics in maximising distance and accuracy of golf shots. *Sports Medicine*, *35*(5), 429–249.

Jöllenbeck, T. (2000). Methodological limitations of EMG-based bio-mechanical motion analysis. *ISBS – Conference Proceedings Archive, 18 International Symposium on Biomechanics in Sports*.

Khushaba, R. N., Kodagoda, S., Takruri, M., & Dissanayake, G. (2012). Expert Systems with Applications Toward improved control of prosthetic fingers using surface electromyogram (EMG) signals. *Expert Systems with Applications*, *39*(12), 10731–10738. doi:10.1016/j.eswa.2012.02.192.

Lee, A. S., Cholewicki, J., & Reeves, N. P. (2007). The effect of background muscle activity on computerized detection of sEMG onset and offset. *Journal of Biomechanics*, *40*(15), 3521–3526. doi:10.1016/j.jbiomech.2007.05.012.

Marta, S., Silva, L., Castro, M. A., Pezarat-Correia, P., & Cabri, J. (2012). Electromyography variables during the golf swing: a literature review. *Journal of Electromyography and Kinesiology*, *22*(6), 803–813. doi:10.1016/j.jelekin.2012.04.002.

Marwan, N., Carmenromano, M., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics Reports*, *438*(5-6), 237–329. doi:10.1016/j.physrep.2006.11.001.

McGill, S., Chaimberg, J., Frost, D., & Fenwick, C. (2010). Evidence of a double peak in muscle activation to enhance strike speed and force: an example with elite mixed martial arts fighters. *Journal of Strength and Conditioning Research*, *24*(2), 348–357. doi:10.1519/JSC.0b013e3181cc23d5.Evidence.

McHardy, A., & Pollard, H. (2005). Lower back pain in golfers: A review of the literature. *Journal of Chiropractic Medicine*, *4*(3), 135–143. doi:10.1016/S0899-3467(07)60122-0.

McHardy, A. J., Pollard, H. P., & Luo, K. (2007). Golf-related lower back injuries: An epidemiological survey. *Journal of Chiropractic Medicine, 6*(1), 20–26. doi:10.1016/j.jcme.2007.02.010.

Morey-Klapsing, G., Arampatzis, A., & Brüggemann, G. P. (2004). Choosing EMG parameters: comparison of different onset determination algorithms and EMG integrals in a joint stability study. *Clinical Biomechanics* (Bristol, Avon), *19*(2), 196–201.

Oskoei, A. M., & Hu, H. (2007). Myoelectric control systems—A survey. *Biomedical Signal Processing and Control*, *2*(4), 275–294. doi:10.1016/j.bspc.2007.07.009.

Oskoei, M. A., & Hu, H. (2008). Support vector machine-based classification scheme for myoelectric control applied to upper limb. *IEEE Transaction on Biomedical Engineering*, *55*(8), 1956–1965.

Palmer, J. L., Young, S. D., Fox, E., Lindsay, D. M., & Vandervoort, A. A. (2003) Senior Recreational Golfers: a survey of musculoskeletal conditions, playing characteristics, and warm-up patterns. *Physiotherapy Canada*, *55*(2), 79-85.

Phinyomark, A., Limsakul, C., & Phukpattaranont, P. (2009). A novel feature extraction for robust EMG pattern recognition. *Journal of Computing*, *1*(1), 2151–9617.

Phinyomark, A., Quaine, F., Charbonnier, S., Serviere, C., Tarpin-Bernard, F., & Laurillau, Y. (2013). EMG feature evaluation for improving myoelectric pattern recognition robustness. *Expert Systems with Applications*, *40*(12), 4832–4840. doi:10.1016/j.eswa.2013.02.023.

Renkawitz, T., Boluki, D., & Grifka, J. (2006). The association of low back pain, neuromuscular imbalance, and trunk extension strength in athletes. *The spine journal*, *6*(6), 673–683. doi:10.1016/j.spinee.2006.03.012.

Silva, L., Castro, M. A., Marta, S., Almeida, J., Vaz, J., Cabri, J., & Pezarat-Correia, P. (2014). Cross-cultural adaptation and validation of the Portuguese Survey of musculoskeletal conditions, playing characteristics and warm-up patterns of golfers. Submitted to *Journal of Back and Musculoskeletal Rehabilitation.*

Silva, L., Marta, S., Vaz, J., Fernandes, O., Castro, M. A., & Correia-Pezarat, P. (2013). Trunk muscle activation during golf swing: Baseline and threshold. *Journal of Electromyography and Kinesiology*, *23*(5), 1174–1182. doi:10.1016/j.jelekin.2013.05.007.

Solnik, S., Rider, P., & Steinweg, K. (2010). Teager–Kaiser energy operator signal conditioning improves EMG onset detection. *European journal of Applied Physiology, 2010*(110), 489–498. doi:10.1007/s00421-010-1521-8.

Staude, G., Flachenecker, C., Daumer, M., & Wolf, W. (2001). Onset detection in surface electromyographic signals: A systematic comparison of methods. *Journal on Applied Signal Processing*, *2001*(2), 67–81.

Stirling, L. M., von Tscharner, V., Kugler, P. F., & Nigg, B. M. (2011). Classification of muscle activity based on effort level during constant pace running. *Journal of Electromyography and Kinesiology*, *21*(4), 566–571. doi:10.1016/j.jelekin.2011.02.005.

Sultornsanee, S., Zeid, I., & Kamarthi, S. (2011). Classification of electromyogram using recurrence quantification analysis. *Procedia Computer Science*, *6*, 375–380. doi:10.1016/j.procs.2011.08.069.

Tavakolan, M., Xiao, Z. G., & Menon, C. (2011). A preliminary investigation assessing the viability of classifying hand postures in seniors. *Biomedical Engineering Online*, *10*(1), 79. doi:10.1186/1475-925X-10-79.

Tkach, D., Huang, H., & Kuiken, T. (2010). Study of stability of time-domain features for electromyographic pattern recognition. *Journal of Neuroengineering and Rehabilitation*, *7:21*, 1–13.

Tolambiya, A., Thomas, E., Chiovetto, E., Berret, B., & Pozzo, T. (2011). An ensemble analysis of electromyographic activity during whole body pointing with the use of support vector machines. *PloS one*, *6*(7), e20732. doi:10.1371/journal.pone.0020732.

Tsai, Y.-S., Sell, T. C., Smoliga, J. M., Myers, J. B., Learman, K. E., & Lephart, S. M. (2010). A comparison of physical characteristics and swing mechanics between golfers with and without a history of low back pain. *Journal of Orthopaedic and Sports Physical Therapy*, 40(7), 430–438. doi:10.2519/jospt.2010.3152.

Vaisman, L., Zariffa, J., & Popovic, M. R. (2010). Application of singular spectrum-based change-point analysis to EMG-onset detection. *Journal of Electromyography and Kinesiology*, *20*(4), 750–760. doi:10.1016/j.jelekin.2010.02.010.

Van Boxtel, G. J. M., Geraars, L. H. D., Van Den Berg-Lenssen, M. M. C., & Brunia, C. H. M. (1993). Detection of EMG onset in ERP research. *Psychophysiology*, *30*(4), 405–412.

Zheng, N., Barrentine, S. W., Fleisig, G. S., Andrews, J. R., & States, U. (2008). Kinematic analysis of swing in pro and amateur golfers. *International Journal of Sports Medicine*, 29(6), 487–493. doi:10.1055/s-2007-989229.

CAPÍTULO IX

9. Conclusões Gerais e Recomendações

Neste ponto destacam-se as principais conclusões e recomendações para o futuro incidindo em dois aspetos fundamentais, a deteção do *onse*t e a discriminação neuromuscular face a constrangimentos associados à habilidade motora.

9.1. Características na Deteção do *Onset* pelo Método Limiar e pela Aprendizagem Automática

A performance na deteção do *onset* depende do músculo estudado e do fenómeno temporal considerado (*onset burst*, *onset peak*…) independentemente do método. A definição de *onset* neste estudo foi o início de atividade mais relevante e padrão entre sujeitos nas repetições estudadas. Este procedimento traduziu-se na identificação de três fenómenos de *onset*, os quais podem ser detetados com maior ou menor dificuldade. Qualquer *onset burst*, coincidente com o maior pico de atividade ou não, apresenta elevada capacidade de deteção. Quando detetado um *onset peak* já com atividade precedente, a precisão desce aumentando a dificuldade de deteção e os erros obtidos (apesar de ter sido identificado o *onset* pretendido). A amplitude na linha de base e o incremento do sinal em amplitude influenciam a deteção do *onset* em habilidades motoras complexas independentemente do método.

Os músculos do tronco tendem a aumentar a sua atividade perto do impacto com a bola, ou seja, o instante do pico máximo ou um segundo pico de atividade tendem a ocorrer de forma sincronizada perto desse instante.

A utilização de diferentes tipos de taco não apresenta influência no início de ativação muscular durante o *swing* do golfe e respetiva sequência de coordenação.

Dos músculos do tronco, o reto abdominal apresenta melhor congruência nos métodos utilizados para a determinação do *onset* por não sofrer influência relevante da atividade na linha de base e consequentemente evidenciar um elevado rácio sinal-ruído. A deteção do *onset burst* no oblíquo externo (mais vincado do lado direito) e do *onset* *peak/burst* no reto abdominal bilateralmente apresenta elevada congruência entre métodos.

Apesar de menor congruência entre métodos e maior de dificuldade na deteção do *onset*, mesmo quando usada aprendizagem automática, a massa comum apresenta cerca de 300 ms antes do impacto um momento padrão de *onset peak* que se caracteriza por um período de desligar/ligar da atividade muscular.

Recorrer à aprendizagem automática usando quatro *features* no domínio do tempo num classificador de máquinas de vetor suporte garante uma precisão média de 95% de instantes de tempo correspondentes a repouso ou ativação, não diferenciando o tipo de fenómeno identificado. Caso a diferenciação seja realizada para o *onset burst*, ou para o *onset peak* quando este tende a ser coincidente com o *onset burst*, a precisão sobe para cerca de 98%.

A utilização de quatro *features* no domínio do tempo é suficiente para garantir a precisão previamente mencionada. As *features* utilizadas foram: valor absoluto médio, comprimento do formato da onda, diferença absoluta do desvio-padrão e variância do EMG. Tendo como base este tipo de *features*, outros conjuntos podem ser associados, como de carácter fenomenológico, esperando-se uma melhoria da precisão nesse caso.

O aumento na dimensão da amostra através de mais exemplos de treino acentua as diferenças entre os grupos de *features* quando usados grupos de quatro ou seis *features* comparados com a utilização de apenas duas *features* no domínio do tempo. Os parâmetros custo e *kernel* apresentam diferenças entre a utilização de apenas duas *features*, sendo indiferente o uso de quatro ou seis *features* no domínio do tempo em relação aos parâmetros identificados na *grid-search*.

9.2. Constrangimentos Handicap e Lombalgia

As *features* da análise de recorrência que apresentaram maior ponderação foram a taxa de recorrência, o determinismo e o rácio entre o determinismo e a taxa de recorrência, tanto na discriminação do handicap como lombalgia.

A fase do *swing* que evidenciou maior poder discriminativo foi o *backswing*, tanto para o handicap como para a lombalgia. No handicap tanto o *swing* na totalidade, como o *backswing* e o *downswing* apresentaram capacidade discriminativa, mas para a lombalgia apenas o swing e o *backswing* foram considerados.

A precisão na classificação para o handicap foi cerca de 94% chegando a 98% para no *swing*, 97% no *backswing*, e 95% no *downswing*. Na lombalgia, a precisão de classificação dos modelos contruídos foi de 97% no *swing* e 99 a 100% no *backswing*, não tendo sido considerado o *downswing*. O *backswing* é a fase de preparação e aquela onde se verifica maior informação discriminatória de constrangimentos motores (em relação aos estudados). A fase *follow-through* não evidenciou características discriminatórias.

Os músculos posteriores da coxa, rotadores do tronco (oblíquo externo) e reto femoral esquerdo apresentaram elevado poder discriminativo para o handicap, variando consoante a fase estudada. Os posteriores da coxa e o reto femoral esquerdo apresentam elevado poder discriminatório em relação à sintomatologia de lombalgia após uma volta de golfe. Isolando cada músculo, durante o *backswing* o oblíquo externo do lado direito apresenta a mais elevada precisão (95%), na discriminação do handicap seguido pelo bicípite femoral direito (94%), semitendinoso esquerdo (89%), semitendinoso direito (87%), e reto femoral direito (87%). Quanto à lombalgia, verificou-se a maior precisão no bicípite femoral esquerdo (99%), seguido pelo semitendinoso esquerdo (89%), bicípite femoral do lado direito (87%), e reto femoral do lado esquerdo (86%).

Sujeitos com lombalgia apresentam diferentes estratégias de coordenação neuromuscular comparativamente a golfistas assintomáticos, assim como os diferentes grupos de handicap estudados. Sujeitos com lombalgia apresentem valores de recorrência com maior similaridade entre músculos do que assintomáticos.

9.3. Recomendações e aplicações práticas

Uma das grandes vantagens deste tipo de abordagem refere-se também à descrição do fenómeno a estudar, pelo que, se consideram como producentes os seguintes passos na aplicação da aprendizagem automática:

(1) Descrição e definição operacional do fenómeno temporal a ser detetado (determinado *onset peak*, *onset burst*,…);

(2) Decisão sobre o tipo de segmentação (caso seja incluído), assim como, em caso de segmentação *overlapped* (necessária da deteção do *onset*) definir o valor do atraso (l*ag*);

(3) Fundamentação dos métodos de extração de *features* (próprio sinal, domínio do tempo, fenomenológicas);

(4) Verificação no músculo em estudo sobre qual o momento de transição de classe que deve ser considerado como o *onset* desejado;

(5) Seleção das *features* relevantes e eliminação daquelas que possam ser redundantes, ou então utilização da proposta neste estudo em relação ao conjunto de quatro *features* no domínio do tempo;

(6) Realização de pesquisa de rede (*grid-search*) e determinação do valor custo e parâmetro *kernel* mais vantajoso, ou então, utilização dos domínios expostos no presente estudo. Os mesmos deverão ser aceites caso se verifique uma boa pontuação classificativa;

(7) Teste dos modelos em relação a novas entradas de dados.

Desta forma, recomenda-se a descrição do fenómeno a detetar quando utilizadas máquinas de vetores suporte. Um ponto que poderá assumir relevância é a aplicação das máquinas de vetor suporte multiclasses, onde são incluídos modelos com vários fenómenos temporais. Bastará a classificação binária quando os pontos a classificar retratam meramente a classe repouso ou a classe atividade. Recomenda-se a análise de outras *features* no domínio do tempo ou diferentes *lags*, mas a maior relevância será estudar outro tipo de *features* que possam integrar modelos mistos. Por exemplo, sabendo que o conjunto proposto garante cerca de 95% de precisão, adicionar a este conjunto *features* tempo-frequência e/ou fenomenológicas.

Recomendam-se estudos futuros sobre a ponderação de diferentes tipos de processamento e parâmetros de incorporação (*embedding dimension*, *delay*) sobre a precisão na discriminação de diferentes sinais de eletromiografia por constrangimentos. Também se recomenda combinar variáveis a nível da cinemática e da cinética com eletromiografia, contruindo modelos tanto de classificação como de regressão que sejam explicativos do movimento.

Referências

ACSM (2010). *Acsm's guidelines for exercise testing and prescription*. 8th ed. Baltimore: Lippinkott Williams & Wilkins.

Adlington, G. S. (1996). Proper swing technique and biomechanics of golf. *Clinical Sports Medicine*, *15*, 9–26.

Alkan, A., & Günay, M. (2012). Identification of EMG signals using discriminant analysis and SVM classifier. *Expert Systems with Applications*, *39*(1), 44–47. doi:10.1016/j.eswa.2011.06.043.

Allison, G. T. (2003). Trunk muscle onset detection technique for EMG signals with ECG artefact. *Journal of Electromyography and Kinesiology*, *13*(3), 209–216. doi:10.1016/S1050-6411(03)00019-1.

Ashish, A., Shweta, S., & Singh, S. J. (2008). Comparison of lumbar and abdominal muscle activation during two types of golf swing: An EMG analysis. *RICYDE. International Journal of Sport Science. IV*(12), 59–71. doi:10.5232/ricyde2008.012.04.

Batt, M. E. (1992). A survey of golf injuries in amateur golfers. *British Journal of Sports Medicine*, *26*(1), 63–65.

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, 25:3186-3191.

Bechler, J.R., Jobe, F. W., Pink, M., Perry, J., & Ruwe, P. A. (1995). Electromyographic analysis of the hip and knee during the golf swing. *Clinical Journal of Sport Medicine*, *5*(3), 162–166.

Beck, T. W., Stock, M. S. & Defreitas, J. M. (2011). Paired Pattern Classification of Electromyographic Intensity Patterns During Concentric and Eccentric Muscle Actions. *Clinical Kinesiology*, *65*(4):76–82.

Begg, R. K., Palaniswami, M., Member, S., & Owen, B. (2005). Support vector machines for automated gait classification. *IEEE Transactions on Biomedical Engineering*, *52*(5), 828–838.

Bitzer, S., Smagt, P. van der (2006) Learning EMG control of a robotic hand: Towards Active Prostheses. *Proceedings of the 2006 IEEE International Conference on Robotics and Automation* (pp. 2819–23). Orlando, Florida.

Boschmann, A., & Platzner, M. (2012). Reducing classification accuracy degradation of pattern recognition based myoelectric control caused by electrode shift using a high density electrode array. In: *Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (Vol 2012, pp. 4324–4327). doi:10.1109/EMBC.2012.6346923.

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992) A Training algorithm for optimal margin classfiers. *Proceedings of the Fifth Annual Workshop on Computational Learning Theory* (pp. 144–152).

Bradshaw, E. J., Keogh, J. W., Hume, P. A., Maulder, P. S., Nortje, J., & Marnewick, M. (2009). The effect of biological movement variability on the performance of the golf swing in high- and low-handicapped players. *Research Quarterly for Exercise & Sport*, *80*(2), 185–196.

Bulbulian, R., Ball, K. a, & Seaman, D. R. (2001). The short golf backswing: effects on performance and spinal health implications. *Journal of Manipulative and Physiological Therapeutics*, *24*(9), 569–75. doi:10.1067/mmt.2001.118982.

Burden, A. M., Grimshaw, P. N., & Wallace, E. S. (1998). Hip and shoulder rotations during the golf swing of sub-10 handicap players. *Journal of Sports Sciences*, *16*(2), 165–176. doi:10.1080/026404198366876.

Burges, C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, *2*, 121–167. Retrieved from http://link.springer.com/article/10.1023/A:1009715923555.

Cabri, J., Sousa, J. P., & Barreiros, J. (2009). Golf-related injuries: A systematic review. *European Journal of Sport Science*, *9*(6), 353–366. doi:10.1080/17461390903009141.

Cann, A. P., Vandervoort, A. A., & Lindsay, D. M. (2005). Optimizing the benefits versus risks of golf participation by older people. *Journal of Geriatric Physical Therapy*, *28*(3), 85-92.

Cao, J., Lin, Z., & Huang, G. (2010). Composite function wavelet neural networks with extreme learning machine. *Neurocomputing*, *73*(7-9):1405–1416. doi:10.1016/j.neucom.2009.12.007.

Castellini, C., & Smagt, P. van der. (2009). Surface EMG in advanced hand prosthetics. *Biological Cybernetics*, *100*(1), 35–47. doi:10.1007/s00422-008-0278-1.

Castellini, C., Gruppioni, E., Davalli, A., & Sandini, G. (2009). Fine detection of grasp force and posture by amputees via surface electromyography. *Journal of Physiology, Paris*, *103*(3-5), 255–62. doi:10.1016/j.jphysparis.2009.08.008.

Castellini, C., Smagt, P. van der, Sandini, G., & Hirzinger, G. (2008). Surface EMG for force control of mechanical hands. In *IEEE International Conference on Robotics and Automation Pasadena, CA, USA, May 19-23, 2008* (pp. 725–730). Pasadena. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4543291.

Chang, C., & Lin, C. (2011). LIBSVM: a library for support vector machines. *ACM transactions on intelligent systems and technology*, *2*, 1–39. Retrieved from http://dl.acm.org/citation.cfm?id=1961199.

Cheetham, P. J., Martin, P. E., Mottram, R. E., & Laurent, B. F. St. (2001). The importance of stretching the “X-Factor” in the downswing of golf : The “X-Factor Stretch” In P. R. Thomas (Ed.), *Optimising performance in golf* (pp. 192–199). Brisbane, Australia: Australian Academic Press.

Cheetham, P. J., Rose, G. A., Hinrichs, R. N., Neal, R. J., Mottram, R. E., Hurrion, P. D., & Vint, P. F. U. S. O. C. (2008). Comparison of Kinematic Sequence Parameters between Amateur and Professional Golfers. Chapter 4. In D. Crews & R. Lutz (Eds.), *Science and Golf V: Proceedings of the World Scientific Congress of Golf* (pp. 30–36). Phoenix, Arizona: The LPGA Foundation.

Chen, X., & Wang, Z. J. (2013). Pattern recognition of number gestures based on a wireless surface EMG system. *Biomedical Signal Processing and Control*, *8*(2), 184–192. doi:10.1016/j.bspc.2012.08.005.

Chen, X., Zhu, X., Zhang, D. (2010). A discriminant bispectrum feature for surface electromyogram signal classification. *Medical Engineering & Physics*, *32*(2):126–135. doi:10.1016/j.medengphy.2009.10.016.

Chen, P.-H., Lin, C.-J., & Scholkopf, B. (2005). A tutorial on nu-support vector machines. *Applied Stochastic Models in Business and Industry*, *21*(2), 111–136. doi:10.1002/asmb.537.

Choi, C., & Kim, J. (2011). Synergy matrices to estimate fluid wrist movements by surface electromyography. *Medical Engineering & Physics*, *33*(8), 916–923. doi:10.1016/j.medengphy.2011.02.006.

Christodoulou, C. I., Kaplanis, P. a., Murray, V., Pattichis, M. S., Pattichis, C. S., & Kyriakides, T. (2012). Multi-scale AM–FM analysis for the classification of surface electromyographic signals. *Biomedical Signal Processing and Control*, *7*(3), 265–269. doi:10.1016/j.bspc.2012.01.001.

Chu, Y., Sell, T. C., & Lephart, S. M. (2010). The relationship between biomechanical variables and driving performance during the golf swing. *Journal of Sports Sciences*, *28*(11), 1251–1259. doi:10.1080/02640414.2010.507249.

Cole, M., & Grimshaw, P. (2008a). Trunk muscle onset and cessation in golfers with and without low back pain. *Journal of Biomechanics*, *41*, 2829–2833. doi:10.1016/j.jbiomech.2008.07.004.

Cole, M. H., & Grimshaw, P. N. (2008b). Electromyography of the trunk and abdominal muscles in golfers with and without low back pain. *Journal of Science and Medicine in Sport / Sports Medicine Australia*, *11*(2), 174–81. doi:10.1016/j.jsams.2007.02.006.

Cole, M. H, & Grimshaw, P. N. (2013). The crunch factor's role in golf-related low back pain. *The Spine Journal*, *Oct 12*.

Coleman, S. G. S., & Rankin, A. J. (2005). A three-dimensional examination of the planar nature of the golf swing. *Journal of Sports Sciences*, *23*(3), 227–234. doi:10.1080/02640410410001730179.

Conradsen, I., Beniczky, S., Wolf, P., Kjaer, T. W., Sams, T., & Sorensen, H. B. D. (2012). Automatic multi-modal intelligent seizure acquisition (MISA) system for detection of motor seizures from electromyographic data and motion data. *Computer Methods and Programs in Biomedicine*, *107*(2), 97–110. doi:10.1016/j.cmpb.2011.06.005.

Cordo, P. J., Gurfinkel, V. S., Smith, T. C., Hodges, P. W., Verschueren, S. M., & Brumagne, S. (2003). The situp: complex kinematics and muscle activity in voluntary axial movement. *Journal of Electromyography and Kinesiology,* 13(3), 239–252.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. doi:10.1007/BF00994018.

Cover, T. M. (1965). Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE Transactions on Electronic Computers*, 326–334.

Crawford, B., Miller, K., Shenoy, P., & Rao, R. (2005). Real-time classification of electromyographic signals for robotic control. In *Proceeding of AAAI* (pp. 523–528).

Cristianini, N., & Shawe-Taylor. (2000). *An introduction to support vector machines and other kernel-based learning methods*. Cambridge: Cambridge University Press.

De Luca, C. J. (1997). The use of surface electromyography in biomechanics. *Journal of Applied Biomechanics*, *13*, 135–63.

De Luca, C. J., Kuznetsov, M., Gilmore, L. D., & Roy, S. H. (2012). Inter-electrode spacing of surface EMG sensors: reduction of crosstalk contamination during voluntary contractions. *Journal of Biomechanics*, *45*(3), 555–561. doi:10.1016/j.jbiomech.2011.11.010.

Del Santo, F., Gelli, F., Schmied, a, Vedel, J.-P., Rossi, a, & Mazzocchio, R. (2006). Motor unit synchronous firing as revealed by determinism of surface myoelectric signal. *Journal of Neuroscience Methods*, *155*(1), 116–21. doi:10.1016/j.jneumeth.2006.01.003.

Dobrowolski, A. P., Wierzbowski, M., & Tomczykiewicz, K. (2012). Multiresolution MUAPs decomposition and SVM-based analysis in the classification of neuromuscular disorders. *Computer Methods and Programs in Biomedicine*, *107*(3), 393–403. doi:10.1016/j.cmpb.2010.12.006.

Duda, R., Hart, P., & Stork, D. (2001). *Pattern classification* (2nd Ed). New York: John Wiley & Sons.

Eckmann, J.-P. , Kamphorst, O. S., & Ruelle, D. (1987). Recurrence plots of dynamical systems. *Europhysics. Letters*, *4*(9), 973–977.

EGA. (2012). *Sistema de handicap EGA 2012-2016*. Epalinges/Lausanne: Edição Federação Portuguesa de Golfe.

Egret, C. I., Vincent, O., Weber, J., Dujardin, F. H., & Chollet. D. (2003). Analysis of 3D kinematics concerning three different clubs in golf swing. *International Journal of Sports Medicine, 24*(6):465–470.

Englehart, K., & Hudgins, B. (2003). A robust, real-time control scheme for multifunction myoelectric control. *IEEE Transaction on Biomedical Engineering*, *50*(7), 848–854. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1206493.

Enoka, R. M. (1996). Eccentric contractions require unique activation strategies by the nervous system. *Journal of Applied Physiology, 81*(6), 2339–2346.

Evans, K., Refshauge, K. M., Adams, R., & Aliprandi, L. (2005). Predictors of low back pain in young elite golfers: A preliminary study. *Physical Therapy in Sport*, *6*(3), 122–130. doi:10.1016/j.ptsp.2005.05.003.

Evgeniou, T., Pontil, M., & Poggio, T. (2000). Statistical Learning Theory : A Primer. *International Journal of Computer Vision*,*38*(1):9–13.

Farber, A. J., Smith, J. S., Kvitne, R. S., Mohr, K. J., & Shin, S. S. (2009). Electromyographic analysis of forearm muscles in professional and amateur golfers. *The American Journal of Sports Medicine*, *37*(2), 396–401. doi:10.1177/0363546508325154.

Farina, D., Fattorini, L., Felici, F., & Filligoi, G. (2002). Nonlinear surface EMG analysis to detect changes of motor unit conduction velocity and synchronization. *Journal of Applied Physiology*, *93*(5), 1753–63. doi:10.1152/japplphysiol.00314.2002.

Farina, D., & Merletti, R. (2000). Comparison of algorithms for estimation of EMG variables during voluntary isometric contractions. *Journal of Electromyography and Kinesiology*, 10(5):337–49.

Farina, D., Merletti, R., & Enoka, R. M. (2004). The extraction of neural strategies from the surface EMG. *Journal of Applied Physiology (Bethesda, Md: 1985)*, *96*(4), 1486–1495. doi:10.1152/japplphysiol.01070.2003.

Farmer, S. F., Halliday, D. M., Conway, B. A., Stephens, J. a, & Rosenberg, J. R. (1997). A review of recent applications of cross-correlation methodologies to human motor unit recording. *Journal of Neuroscience Methods*, *74*(2), 175–187.

Farrally, M. R., Cochran, A. J., Crews, D. J., Hurdzan, M. J., Price, R. J., Snow, J.T., & Thomas, P. R. (2003). Golf science research at the beginning of the twenty-first century. *Journal of Sports Science*, *21*(9), 753-765.

Filligoi, G., & Felici, F. (1999). Detection of hidden rhythms in surface EMG signals with a non-linear time-series tool. *Medical Engineering & Physics*, *21*(6-7):439–448.

Finch, C., Sherman, C., James, T., Farrally, M., & Cochran, A. (1999). The epidemiology of golf injuries in Victoria, Australia: evidence from sports medicine clinics and emergency department presentations. In: M. Farrally & A. Cochran (Eds). *Science and Golf III: Proceedings of the World Scientific Congress of Golf* (pp. 73-82). Champaign: Human Kinetics.

Fligge, N., Urbanek, H., Smagt, P. van der (2013) Relation between object properties and EMG during reaching to grasp. Journal of Electromyography and Kinesiology, *23*(2):402–410. doi:10.1016/j.jelekin.2012.10.010.

Fontana, J. M., & Chiu, a. W. L. (2009). Control of Prosthetic Device Using Support Vector Machine Signal Classification Technique. *American Journal of Biomedical Sciences*, *1*(4), 336–343. doi:10.5099/aj090400336.

Fox, E., Lindsay, D., & Vandervoort, A. (2002). Musculoskeletal injury questionnaire for senior golfers. In *Science and Golf IV. Proceedings of the World Scientific Congress of Golf* (pp. 88–99). St. Andrews, Scotland: Routledge Inc.

Futamata, M., Nagata, K., & Magatani, K. (2012). The evaluation of the discriminant ability of multiclass SVM in a study of hand motion recognition by using SEMG. In *Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society* (Vol. 2012, pp. 5246–9). doi:10.1109/EMBC.2012.6347177.

Gatt, C. J., Pavol, M. J., Parker, R. D., & Grabiner, M. D. (1998). Three-dimensional knee joint kinetics during a golf swing. Influences of skill level and footwear. *American Journal of Sports Medicine*, *26*(2), 285–294.

Glazebrook, M., Curwin, S., Islam, M., Kozey, J., & Stanish, W. (1994). Medial epicondylitis. An electromyographic analysis and an investigation of intervention strategies. *American Journal of Sports Medicine*, *22*(5), 674–679.

Gluck, G. S., Bendo, J. A., & Spivak, J. M. (2008). The lumbar spine and low back pain in golf: a literature review of swing biomechanics and injury prevention. *The Spine Journal*, *8*(5), 778–88. doi:10.1016/j.spinee.2007.07.388.

Goker, I., Osman, O., Ozekes, S., Baslo, M. B., Ertas, M., & Ulgen, Y. (2012). Classification of juvenile myoclonic epilepsy data acquired through scanning electromyography with machine learning algorithms. *Journal of Medical Systems*, *36*(5), 2705–2711. doi:10.1007/s10916-011-9746-6.

Gosheger, G., Liem, D., Ludwig, K., Greshake, O., & Winkelmann, W. (2003). Injuries and overuse syndromes in golf. *American Journal of Sports Medicine 31*(3), 438-443

Guillemin, F., Bombardier, C., Beaton, D. (1993). Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *Journal of Clinical Epidemiology, 46*, 1417-1432.

Güler, N. F., & Koçer, S. (2005). Use of support vector machines and neural network in diagnosis of neuromuscular disorders. *Journal of Medical Systems*, *29*(3), 271–284. doi:10.1007/s10916-005-5187-4.

Gulgin, H., Armstrong, C. & Gribble, P. (2009). Hip rotational velocities during the full golf swing. *Journal of Sports Science and Medicine*, *8*, 296-299.

Hall, M. A. (1999). *Correlation-based feature selection for machine learning*. (Doctoral dissertation). University of Waikato.

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorial Quantitative Methods Psychology*, *8*(1), 23–34.

Haykin, S. (2001). *Redes neurais. Princípios e prática*. (2ª ed.). Santana: ARTMED.

Hellström, J., & Tinnmark, F. (2008). The association between stability and swing kinematics of skilled high school golfers. In D. Crews & R. Lutz (Eds.), *Science and Golf V: Proceedings of the World Scientific Congress of Golf* (pp. 37–43). Phoenix, Arizona: The LPGA Foundation.

Herbrich, R. (2002). *Learning kernel classifiers. Theory and algorithms*. Massachusetts: Massachusetts Institute of Technology.

Hermens, H. J., Merletti, & Freriks. B. (1996). European Activities on Surface Electromyography. European Activities on Surface Electromyography, *Proceedings of the First General SENIAM Workshop*. SENIAM project, published by Roessingh Research and Development.

Hirashima, M., Kadota, H., Sakurai, S., Kudo, K., & Ohtsuki, T. J. (2002). Sequential muscle activity and its functional role in the upper extremity and trunk during overarm throwing. *Sports Science*, *20*(4):301–10.

Hodges, P. W. (2001). Changes in motor planning of feedforward postural responses of the trunk muscles in low back pain. *Experimental Brain Research*, *141*(2), 261–266.

Hodges, P. W., & Richardson, C. A. (1996). Inefficient muscular stabilization of the lumbar spine associated with low back pain: a motor control evaluation of transversus abdominis. *Spine*, 21, 2640–2650.

Hodges, P., & Bui, B. (1996). A comparison of computer-based methods for the determination of onset of muscle contraction using electromyography. *Electroencephalography and Clinical Neurophysiology*, *101*, 511–519.

Hofmann, T., Schölkopf, B., & Smola, A. J. (2008). Kernel methods in machine learning. *The Annals of Statistics*, *36*(3), 1171–1220. doi:10.1214/009053607000000677.

Horton, J. F., Lindsay, D. M., & Macintosh, B. R. (2001). Abdominal muscle activation of elite male golfers with chronic low back pain. *Medicine and Science in Sports and Exercise*, *33*(10), 1647–1654.

Hosea, T. M. & Gatt, C. J. Jr. (1996). Back pain in golf. *Clinical Sports Medicine, 15*, 37–53.

Hosea, T. M., Gatt, C. J., Galli, K. M., Langrana, N. A., & Zawadasky, J. P. (1990). Biomechanical analysis of the golfer's back. In Cochrane A. J. *Science and Golf* (pp. 43-48). London: Chapman and Hall.

Hsu, C., Chang, C., & Lin, C. (2010). A practical guide to support vector classification. Retrieved from https://www.cs.sfu.ca/people/Faculty/teaching/726/spring11/svmguide.pdf

Hsu, C.-W. & Lin, C.-J. (2002). A Comparison of methods for multiclass support vector machines. *IEEE Transctions on Neural networks*, *13*(2), 415–425.

Huang, H., Liu, H., Li, N., Jiang, L., Yang, D., Wan, L., … Hirzinger, G. (2012). An anthropomorphic controlled hand prosthesis system. *Journal of Zhejiang University Science C*, *13*(10), 769–780. doi:10.1631/jzus.C1100257.

Huang, H., Zhang, F., Hargrove, L. J., Dou, Z., Rogers, D. R., & Englehart, K. B. (2011). Continuous locomotion-mode identification for prosthetic legs based on neuromuscular–mechanical fusion. *IEEE Transactions on Biomedical Engineering, 58*(10):2867–2875. doi:10.1109/TBME.2011.2161671.Continuous.

Hudgins, B., Parker, P., & Scott, R. N. (1993). A new strategy for multifunction myoelectric control. *IEEE Transactions on Biomedical Engineering*, *40*(1), 82–94.

Hug, F. (2011). Can muscle coordination be precisely studied by surface electromyography? *Journal of Electromyography and Kinesiology*, 21(1):1–12.

Hume, P. A, Keogh, J., & Reid, D. (2005). The role of biomechanics in maximising distance and accuracy of golf shots. *Sports Medicine*, *35*(5), 429–249.

Ikegawa, S., Shinohara, M., Fukunaga, T., Zbilut, J. P., & Webber Jr, C. L. (2000). Nonlinear time-course of lumbar muscle fatigue using recurrence quantifications. *Biological Cybernetics*, *82*, 373–382.

Jobe, F. W., Perry, J., & Pink, M. (1989). Electromyographic shoulder activity in men and women professional golfers. *American Journal of Sports Medicine*, 17(6), 782-787.

Jobe, F., Moynes, D., & Antonelli, D. (1986). Rotator cuff functions during a golf swing. *American Journal of Sports Medicine*, *14*(5), 388–392.

Jöllenbeck, T. (2000). Methodological limitations of EMG-based bio-mechanical motion analysis. *ISBS – Conference Proceedings Archive, 18 International Symposium on Biomechanics in Sports*.

Kao, J., Pink, M., Jobe, F., & Perry, J. EMG analysis of the scapular muscles during a golf swing. *American Journal of Sports Medicine*, *23*(1), 19-23.

Katsis, C. D., Goletsis, Y., Likas, A., Fotiadis, D. I., & Sarmas, I. (2006). A novel method for automated EMG decomposition and MUAP classification. *Artificial Intelligence in Medicine*, *37*(1):55–64. doi:10.1016/j.artmed.2005.09.002.

Kaur, G., Arora, A., & Jain, V. (2009). Multi-class support vector machine classifier in EMG diagnosis. *WSEAS Transactions on Signal Processing*, *5*(12):379–389.

Keogh, J. W. & Hume, P. A. (2012). Evidence for biomechanics and motor learning research improving golf performance. *Sports Biomechanics*, *11*(2), 288-309.

Keogh, J. W., Marnewick, M. C., Maulder, P. S., Nortje, J. P., Hume, P. A., & Bradshaw, E. J. (2009). Are anthropometric, flexibility, muscular strength, and endurance variables related to clubhead velocity in low- and high-handicap golfers? *Journal of Strength & Conditioning Research*, *23*(6), 1841-50.

Khokhar, Z. O., Xiao, Z. G., & Menon, C. (2010). Surface EMG pattern recognition for real-time control of a wrist exoskeleton. *Biomedical Engineering Online*, *9:41*, 1–17. doi:10.1186/1475-925X-9-41.

Khushaba, R. N., Kodagoda, S., Takruri, M., & Dissanayake, G. (2012). Expert Systems with Applications Toward improved control of prosthetic fingers using surface electromyogram (EMG) signals. *Expert Systems with Applications*, *39*(12), 10731–10738. doi:10.1016/j.eswa.2012.02.192.

Khushaba, R. N., Kodagoda, S., Liu, D., & Dissanayake, G. (2013). Muscle computer interfaces for driver distraction reduction. *Computer Methods and Programs in Biomedicine*, *110*(2), 137–149. doi:10.1016/j.cmpb.2012.11.002.

Konrad, P. (2005). *The ABC of EMG: a practical introduction to kinesiological electromyography*. USA: Noraxon Inc.; 2005.

Kuhn, H. W., & Tucker, A. W. (1951). Nonlinear programming. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability* (pp. 481–492).

Kulkarni, S. R., & Harman, G. (2011). Statistical learning theory: a tutorial. *Wiley Interdisciplinary Reviews: Computational Statistics*, *3*(6), 543–556. doi:10.1002/wics.179.

Kumar, D. K., Poosapadi Arjunan, S., & Singh, V. P. (2013). Towards identification of finger flexions using single channel surface electromyography-able bodied and amputee subjects. *Journal of Neuroengineering and Rehabilitation*, *10*(50), 1–7. doi:10.1186/1743-0003-10-50.

Leader 3rd, J. K., Boston, J. R., & Moore, C. A. (1998). A data dependent computer algorithm for the detection of muscle activity onset and offset from EMG recordings. *Electroencephalography and Clinical Neurophysiology*, *109*(2):119–23.

Lee J, Jung MY, Kim SH. (2011). Reliability of spike and turn variables of surface EMG during isometric voluntary contractions of the biceps brachii muscle. *Journal of Electromyography and Kinesiology*, *21*(1):119–27.

Lee, A. S., Cholewicki, J., & Reeves, N. P. (2007). The effect of background muscle activity on computerized detection of sEMG onset and offset. *Journal of Biomechanics*, *40*(15), 3521–3526. doi:10.1016/j.jbiomech.2007.05.012.

Lei, M., & Meng, G. (2012). Nonlinear Analysis of Surface EMG Signals. In M. Lei & G. Meng (Eds.), *Computational Intelligence in Electromyography Analysis - A Perspective on Current Applications and Future Challenges.* (pp. 119–174). INTECH. Retrieved from http://cdn.intechopen.com/pdfs/40114/InTech-Nonlinear_analysis_of_surface_emg_signals.pdf

Li, N., Yang, D., Jiang, L., Liu, H., & Cai, H. (2012). Combined use of FSR sensor array and SVM classifier for finger motion recognition based on pressure distribution map. *Journal of Bionic Engineering*, *9*(1), 39–47.

Li, X., & Aruin, A. (2005). Muscle activity onset time detection using teager-kaiser energy operator. In *Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference* (Vol. 7, pp. 7549–52) Shanghai, China. doi:10.1109/IEMBS.2005.1616259.

Lin, H., & Lin, C. (2003). A study on sigmoid kernels for SVM and the training of non-PSD kernels by SMO-type methods. *Submitted to Neural Computation*, 1–32. Retrieved from http://home.caltech.edu/~htlin/publication/doc/tanh.pdf

Lindsay, D. M., Horton, J. F., & Paley, R. D. (2002). Trunk Motion of Male Professional Golfers Using Two Different Golf Clubs. *Journal of Applied Biomechanics*, *18*, 366–373.

Lindsay, D. M., Horton, J. F., & Vandervoort, A. A. (2000). Review of injury characteristics, aging factors and prevention programmes for the older golfer. *Sports Medicine*, *30*(2): 89-103.

Lindsay, D., & Horton, J. (2002). Comparison of spine motion in elite golfers with and without low back pain. *Journal of Sports Sciences*, *20*(8), 599–605. doi:10.1080/026404102320183158.

Lorena, A. C., & Carvalho, A. C. P. L. F. (2003). *Introdução às Máquinas de Vetores Suporte*. (Relatórios Técnicos do ICMC).

Lorena, A. C., & Carvalho, A. C. P. L. F. De. (2007). Uma introdução às support vector machines. *RITA*, *XIV*(2), 43–67.

Lucas, M.-F., Gaufriau, A., Pascual, S., Doncarli, C., & Farina, D. (2008). Multi-channel surface EMG classification using support vector machines and signal-based wavelet optimization. *Biomedical Signal Processing and Control*, *3*(2), 169–174. doi:10.1016/j.bspc.2007.09.002.

Luxburg, U. Von, & Schölkopf, B. (2009). Statistical learning theory: models, concepts, and results. In D. M. Gabbay, S. Hartmann, & J. Woods (Eds.), *Handbook of the history of logic*. (Vol. 10, pp. 651–706). Elsevier BV.

Marta, S., Silva, L., Castro, M. A., Pezarat-Correia, P., & Cabri, J. (2012). Electromyography variables during the golf swing: a literature review. *Journal of Electromyography and Kinesiology*, *22*(6), 803–813. doi:10.1016/j.jelekin.2012.04.002.

Marta, S., Silva, L., Vaz, J., Bruno, P., & Pezarat-correia, P. (2013). Electromyographic analysis of trunk muscles during the golf swing performed with two different clubs. *International Journal of Sports Science & Coaching*, *8*(4): 779–787.

Marwan, N., Carmenromano, M., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics Reports*, *438*(5-6), 237–329. doi:10.1016/j.physrep.2006.11.001.

Marwan, N., & Kurths, J. (2004). Cross Recurrence Plots and their Applications. In C. V. Benton (Ed.), *Mathematical physics research at the cutting edge* (pp. 101–139). Nova Science Publisher, Inc.

Marwan, N., Thiel, M., & Nowaczyk, N. R. (2002). Cross recurrence plot based synchronization of time series. *Nonlinear Processes in Geophysics*, *2*(9), 325–331.

McCarroll, J. R. (1996). The frequency of golf injuries. *Clinics in Sports Medicine*, *15*(1): 1-7.

McCarroll, M., Retting, A. C., & Shelbourne, K. (1990). Injuries in the amateur golfer. *Physician and Sports Medicine*, *18*, 122-126.

McGill, S. M. (1991). Electromyographic activity of the abdominal and low back musculature during the generation of isometric and dynamic axial trunk torque: implications for lumbar mechanics. *Journal of Orthopaedic Research*, *9*(1), 91–103.

McGill, S. M. (2004). Linking latest knowledge of injury mechanisms and spine function to the prevention of low back disorders. *Journal of Electromyography and Kinesiology, 14*(1), 43–47. doi:10.1016/j.jelekin.2003.09.012.

McGill, S., Chaimberg, J., Frost, D., & Fenwick, C. (2010). Evidence of a double peak in muscle activation to enhance strike speed and force: an example with elite mixed martial arts fighters. *Journal of Strength and Conditioning Research*, *24*(2), 348–357. doi:10.1519/JSC.0b013e3181cc23d5.Evidence.

McGill, S., Juker, D., & Kropf, P. (1996). Appropriately placed surface EMG electrodes reflect deep muscle activity (psoas, quadratus lumborum, abdominal wall) in the lumbar spine. *Journal of Biomechanics*, *29*(11), 1503–1507.

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, *1*(1), 30–46.

McHardy, A. J., Pollard, H. P., & Luo, K. (2007). Golf-related lower back injuries: An epidemiological survey. *Journal of Chiropractic Medicine*, *6*(1), 20–6. doi:10.1016/j.jcme.2007.02.010.

McHardy, A., & Pollard, H. (2005a). Lower back pain in golfers: A review of the literature. *Journal of Chiropractic Medicine*, *4*(3), 135–143. doi:10.1016/S0899-3467(07)60122-0.

McHardy, A., & Pollard, H. (2005b). Muscle activity during the golf swing. *British Journal of Sports Medicine*, *39*(11), 799–804. doi:10.1136/bjsm.2005.020271.

McHardy, A., Pollard, H., & Luo, K. (2006). Golf injuries: a review of the literature. *Sports Medicine*, *36*(2), 171–87.

McLean, J. (1992). Widen the gap. *Golf Magazine*. *Dec*.

McLean, J. (2008). Triple X-factor. *Golf Digest, Jan*.

Merletti, R., & Farina, D. (2008). Surface EMG processing: Introduction to the special issue. *Biomedical Signal Processing and Control*, *3*(2), 115–117. doi:10.1016/j.bspc.2008.01.001

Merlo, A., & Campanini, I. (2010). Technical Aspects of Surface Electromyography for Clinicians, *The Open Rehabilitation Journal*, *3*, 98–109.

Metz, J. P. (1999). Managing golf injuries. Technique and equipment changes to aid treatment. *Physician and Sports Medicine*, *27*(77), 41-58.

Mezzarane, R. A., & Kohn, A. F. (2009). A method to estimate EMG crosstalk between two muscles based on the silent period following an H-reflex. *Medical Engineering & Physics*, *31*(2009), 1331–1336. doi:10.1016/j.medengphy.2009.09.005.

Micera, S., Sabatini, A. M., & Dario, P. (1998). An algorithm for detecting the onset of muscle contraction by EMG signal processing. *Medical Engineering & Physics*, *20*(3), 211–215.

Mogk, J. P. M., & Keir, P. J. (2003). Crosstalk in surface electromyography of the proximal forearm during gripping tasks. *Journal of Electromyography and Kinesiology*, *13*, 63–71. doi:10.1016/S1050-6411(02)00071-8.

Morey-Klapsing, G., Arampatzis, A., & Brüggemann, G. P. (2004). Choosing EMG parameters: comparison of different onset determination algorithms and EMG integrals in a joint stability study. *Clinical Biomechanics* (Bristol, Avon), *19*(2), 196–201.

Müller, K.-R., Mika, S., Rätsch, G., Tsuda, K., & Schölkopf, B. (2001). An introduction to kernel-based learning algorithms. *IEEE Transactions on Neural Networks*, *12*(2), 181–201.

Myers, J., Lephart, S., Tsai, Y.-S., Sell, T., Smoliga, J., & Jolly, J. (2008). The role of upper torso and pelvis rotation in driving performance during the golf swing. *Journal of Sports Sciences*, *26*(2), 181–188. doi:10.1080/02640410701373543.

Natarajan, G. S., Wininger, M., Kim, N. H., & Craelius, W. (2012). Relating biceps EMG to elbow kinematics during self-paced arm flexions. *Medical Engineering & Physics*, *34*(5), 617–624. doi:10.1016/j.medengphy.2011.09.009.

Nesbit, S. M., & McGinnis, R. (2009). Kinematic analyses of the golf swing hub path and its role in golfer/club kinetic transfers. *Journal of Sports Science & Medicine*, *8*(2), 235–346.

Nesbit, S. M., & Serrano, M. (2005). Work and power analysis of the golf swing. *Journal of Sports Science & Medicine*, *4*, 520–533.

Orosco, E. C., Lopez, N. M., & di Sciascio, F. (2013). Bispectrum-based features classification for myoelectric control. *Biomedical Signal Processing and Control*, 8(2):153–168. doi:10.1016/j.bspc.2012.08.008.

Oskoei, A. M., & Hu, H. (2007). Myoelectric control systems—A survey. *Biomedical Signal Processing and Control*, *2*(4), 275–294. doi:10.1016/j.bspc.2007.07.009.

Oskoei, M. A., & Hu, H. (2008). Support vector machine-based classification scheme for myoelectric control applied to upper limb. *IEEE Transaction on Biomedical Engineering*, *55*(8), 1956–1965.

Palmer, J. L., Young, S. D., Fox, E., Lindsay, D. M., & Vandervoort, A. A. (2003) Senior Recreational Golfers: a survey of musculoskeletal conditions, playing characteristics, and warm-up patterns. *Physiotherapy Canada*, *55*(2), 79–85.

Park, S. H., Lee, S. P. (1998). EMG pattern recognition based on artificial intelligence techniques. *IEEE Transactions on Rehabilitation Engineering*, *6*(4):400–405.

Penner, A. R. (2003). The physics of golf. *Reports on Progress in Physics*, *66*(2), 131–171. doi:10.1088/0034-4885/66/2/202.

Pezarat-Correia, P. & Mil-Homens, P. (2004). *A Electromiografia no estudo do movimento humano*. Cruz Quebrada: Edições FMH.

Phinyomark, A., Limsakul, C., & Phukpattaranont, P. (2009). A novel feature extraction for robust EMG pattern recognition. *Journal of Computing*, *1*(1), 2151–9617.

Phinyomark, A., Nuidod, A., Phukpattaranont, P., & Limsakul, C. (2012). Feature extraction and reduction of wavelet transform coefficients for EMG pattern classification. *Electronics and Electrical Engineering*, *122*(6), 27-32. doi:10.5755/j01.eee.122.6.1816.

Phinyomark, A., Phukpattaranont, P., & Limsakul, C. (2012). Feature reduction and selection for EMG signal classification. *Expert Systems with Applications*, *39*(8), 7420–7431. doi:10.1016/j.eswa.2012.01.102.

Phinyomark, A., Quaine, F., Charbonnier, S., Serviere, C., Tarpin-Bernard, F., & Laurillau, Y. (2013). EMG feature evaluation for improving myoelectric pattern recognition robustness. *Expert Systems with Applications*, *40*(12), 4832–4840. doi:10.1016/j.eswa.2013.02.023.

Pink, M., Jobe, F., & Perry, J. (1990). Electromyographic analysis of the shoulder during the golf swing. *American Journal of Sports Medicine*, *18*(2), 137–140.

Pink, M., Perry, J., & Jobe, F. (1993). EMG analysis of the trunk in golfers. *American Journal of Sports Medicine*, *21*(3):385–388.

Portney, L., Watkins, M. (2009). *Foundations of clinical research: Applications to practice*. 3rd ed. Upper Saddle River, New Jersey: Pearson Prentice Hall.

Raez, M. B. I., Hussain, M. S., & Mohd-Yasin, F. (2006). Techniques of EMG signal analysis: detection, processing, classification and applications. *Biological Procedures Online*, *8*(1), 11–35. doi:10.1251/bpo115.

Rahimi, S. Y., Singh, H., Yeh, D. J., Shaver, E. G., Flannery, A. M., & Lee, M. R. (2005). Golf associated head injury in the pediatric population: a common sports injury. *Journal of Neurosurgery*, *102*(2 Suppl), 163-166.

Renkawitz, T., Boluki, D., & Grifka, J. (2006). The association of low back pain, neuromuscular imbalance, and trunk extension strength in athletes. *The spine journal*, *6*(6), 673–683. doi:10.1016/j.spinee.2006.03.012.

Richardson, A. K., Hughes, G., & Mitchell, A. C. S. (2012). Center of Pressure Excursion During the Golf Putting Stroke in Low, Mid and High Handicap Golfers. *International Journal of Golf Science*, *1*, 127–139.

Rojas, R. (1996). *Neural networks: A systematic introduction*. Berlin: Springer-Verlag.

Scheme, E., & Englehart, K. (2011). Electromyogram pattern recognition for control of powered upper-limb prostheses: State of the art and challenges for clinical use. *Journal of Rehabilitation Research and Development*, *48*(6), 643–59. doi:10.1682/JRRD.2010.09.0177.

Schölkopf, B., & Smola, A. (2002). *Learning with kernels*. Massachustts: Massachusetts Institute of Technology. Retrieved from http://dip.sun.ac.za/~hanno/tw796/lesings/mlss06au_scholkopf_lk.pdf

Scholkopf, B., Burges, C., & Vapnik, V. (1995). Extracting Support Data for a Given Task. In U. M. Fayyad & R. Uthurusamy (Eds.), *First International Conference on Knowledge Discovery & Data Mining*. (pp. 252–257). Menlo Park: AAAI Press.

Schölkopf, B., Smola, A. J., Williamson, R. C., & Bartlett, P. L. (2000). New support vector algorithms. *Neural Computation*, *12*, 1207–1245.

Scholkopf, B., Smola, A., & Muller, K. R. (1996). *Nonlinear component analysis as a kernel eigenvalue problem* (pp. 1–18). Tübingen.

Sherman, C. A., & Finch, C. F. (2000). Preventing injuries to competitive and recreational adult golfers: what is the evidence? *Journal of Science and Medicine in Sport / Sports Medicine Australia*, *3*(1), 65–78.

Shi, J., Guo, J.-Y., Hu, S.-X., & Zheng, Y.-P. (2012). Recognition of finger flexion motion from ultrasound image: a feasibility study. *Ultrasound in Medicine and Biology*, *38*(10), 1695–1704. doi:10.1016/j.ultrasmedbio.2012.04.021.

Shi, J., Zheng, Y., Chen, X., & Xie, H. (2009). Modeling the relationship between wrist angle and muscle thickness during wrist flexion-extension based on the bone-muscle lever system: a comparison study. *Medical Engineering & Physics*, *31*(10), 1255–1260. doi:10.1016/j.medengphy.2009.08.003.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass Correlations: uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420-428.

Silfies, S. P., Mehta, R., Smith, S. S., & Karduna, A. R. (2009). Differences in Feedforward Trunk Muscle Activity in Subgroups of Patients with Mechanical Low Back Pain. *YAPMR*, *90*(7), 1159–1169. doi:10.1016/j.apmr.2008.10.033.

Silva, L., Castro, M. A., Marta, S., Almeida, J., Vaz, J., Cabri, J., & Pezarat-Correia, P. (2014). Cross-cultural adaptation and validation of the Portuguese Survey of musculoskeletal conditions, playing characteristics and warm-up patterns of golfers. Submitted to *Journal of Back and Musculoskeletal Rehabilitation.*

Silva, L., Marta, S., Vaz, J., Fernandes, O., Castro, M. A., & Correia-Pezarat, P. (2013). Trunk muscle activation during golf swing: Baseline and threshold. *Journal of Electromyography and Kinesiology*, *23*(5), 1174–1182. doi:10.1016/j.jelekin.2013.05.007.

Smagt, P. van der, Grebenstein, M., Urbanek, H., Fligge, N., Strohmayr, M., Stillfried, G., … Gustus, A. (2009). Robotics of human movements. *Journal of Physiology, Paris*, *103*(3-5), 119–32. doi:10.1016/j.jphysparis.2009.07.009.

Soda, P., Mazzoleni, S., Cavallo, G., Guglielmelli, E., & Iannello, G. (2010) Human movement onset detection from isometric force and torque measurements: a supervised pattern recognition approach. *Artificial Intelligence in Medicine*, *50*(1), 55–61. doi:10.1016/j.artmed.2010.04.008.

Solnik, S., DeVita, P., Rider, P., Long, B., & Hortobágyi, T. (2008). Teager–Kaiser operator improves the accuracy of EMG onset detection independent of signal-to-noise ratio. *Acta of Bioengineering and Biomechanics*, *10*(2), 65–8.

Solnik, S., Rider, P., & Steinweg, K. (2010). Teager–Kaiser energy operator signal conditioning improves EMG onset detection. *European journal of Applied Physiology*, *2010*(110), 489–498. doi:10.1007/s00421-010-1521-8.

Somjarod, M., Tanawat, V., & Weerawat, L. (2011). The analysis of knee Joint movement during golf swing in professional and amateur golfers. *International Science Index. World Academy of Science, Engineering and Technology*, *53*, 495–498.

Soylu, A. R., & Arpinar-Avsar, P. (2010). Detection of surface electromyography recording time interval without muscle fatigue effect for biceps brachii muscle during maximum voluntary contraction. *Journal of Electromyography and Kinesiology*, *20*(4), 773–776.

Staude, G., Flachenecker, C., Daumer, M., & Wolf, W. (2001). Onset detection in surface electromyographic signals: A systematic comparison of methods. *Journal on Applied Signal Processing*, *2001*(2), 67–81.

Stirling, L. M., von Tscharner, V., Kugler, P. F., & Nigg, B. M. (2011). Classification of muscle activity based on effort level during constant pace running. *Journal of Electromyography and Kinesiology*, *21*(4), 566–571. doi:10.1016/j.jelekin.2011.02.005.

Subasi, A. (2012). Medical decision support system for diagnosis of neuromuscular disorders using DWT and fuzzy support vector machines. *Computers in Biology and Medicine*, *42*(8), 806–815. doi:10.1016/j.compbiomed.2012.06.004.

Subasi, A. (2013). Classification of EMG signals using PSO optimized SVM for diagnosis of neuromuscular disorders. *Computers in Biology and Medicine*, *43*(5), 576–86. doi:10.1016/j.compbiomed.2013.01.020.

Suckel, A. (2002). [Sports medicine analysis of golf "swing" and lesions occurring during golf practice]. *Portverletz Sportschaden*, *16*(1), 31-35.

Sugaya, H., Tsuchiya, A., Moriya, H., Morgan, D. A., & Banks, S. A. (1999). Low back injury in elite and professional golfers: An epidemiologic and radiographic study. In A. J. Cochran & M.R. Fairally (Eds.), *Science and Golfl III: Proceedings of the World Scientific Congress of Golf* (pp. 83-91). Champaign, IL: Human Kinetics

Sultornsanee, S., Zeid, I., & Kamarthi, S. (2011). Classification of electromyogram using recurrence quantification analysis. *Procedia Computer Science*, *6*, 375–380. doi:10.1016/j.procs.2011.08.069.

Tavakolan, M., Xiao, Z. G., & Menon, C. (2011). A preliminary investigation assessing the viability of classifying hand postures in seniors. *Biomedical Engineering Online*, *10*(1), 79. doi:10.1186/1475-925X-10-79.

Thériault, G., & Lachance, P. (1998). Golf Injuries. An overview. *Sports Medicine*, *26*(1), 43–57. doi:10.2165/00007256-199826010-00004.

Thiel, M., Romano, M. C., & Kurths, J. (2004). How much information is contained in a recurrence plot? *Physics Letters A*, *330*(5), 343–349. doi:10.1016/j.physleta.2004.07.050.

Tkach, D., Huang, H., & Kuiken, T. (2010). Study of stability of time-domain features for electromyographic pattern recognition. *Journal of Neuroengineering and Rehabilitation*, *7:21*, 1–13.

Tolambiya, A., Thomas, E., Chiovetto, E., Berret, B., & Pozzo, T. (2011). An ensemble analysis of electromyographic activity during whole body pointing with the use of support vector machines. *PloS one*, *6*(7), e20732. doi:10.1371/journal.pone.0020732.

Tsai, Y.-S., Sell, T. C., Smoliga, J. M., Myers, J. B., Learman, K. E., & Lephart, S. M. (2010). A comparison of physical characteristics and swing mechanics between golfers with and without a history of low back pain. *Journal of Orthopaedic and Sports Physical Therapy*, 40(7), 430–438. doi:10.2519/jospt.2010.3152.

Tyler, A., & Karst, G. (2004). Timing of muscle activity during reaching while standing: systematic changes with target distance. *Gait & posture*, *20*, 126–133. doi:10.1016/j.gaitpost.2003.07.001.

Vad, V. B., Bhat, A. L., Basrai, D., Gebeh, A., Aspergren, D. D., & Andrews, J. R. (2004). Low back pain in professional golfers: the role of associated hip and low back range-of motion deficits. *American Journal of Sports Medicine*, *32*(2), 494–497. doi:10.1177/0363546503261729.

Vaisman, L., Zariffa, J., & Popovic, M. R. (2010). Application of singular spectrum-based change-point analysis to EMG-onset detection. *Journal of Electromyography and Kinesiology*, *20*(4), 750–760. doi:10.1016/j.jelekin.2010.02.010.

Van Boxtel, G. J. M., Geraars, L. H. D., Van Den Berg-Lenssen, M. M. C., & Brunia, C. H. M. (1993). Detection of EMG onset in ERP research. *Psychophysiology*, *30*(4), 405–412.

Vannozzi, G., Conforto, S., & Alessio, T. D. (2010) Automatic detection of surface EMG activation timing using a wavelet transform based method. *Journal of Electromyography and Kinesiology*, *20*(4), 767–772. doi:10.1016/j.jelekin.2010.02.007.

Vapnik, V. N. (1999). An overview of statistical learning theory. *IEEE Transactions on Neural Networks*, *10*(5), 988–999.

Watkins, R. G., Uppal, G. S., Perry, J., Pink, M., & Dinsay, J. M. (1996). Dynamic electromyographic analysis of trunk musculature in professional golfers. *The American Journal of Sports Medicine*, *24*(4), 535–538.

Watson, D. S., Mehan, T. J., Smith, G. A., McKenzie, L. B. (2008). Golf cart-related injuries in the U.S. *American Journal of Preventive Medicine*, *35*(1), 55-59.

Webber, C. L. J., & Zbilut, J. P. (2005). Recurrence quantification analysis of nonlinear dynamical systems. In M. A. Riley & G. C. Van Orden (Eds.), *Tutorials in contemporary nonlinear methods for the behavioral sciences* (pp. 26–94). Retrieved 2012, from http://www.nsf.gov/sbe/bcs/pac/nmbs/nmbs.jsp.

Webber, Jr C. L., & Zbilut, J. P. (1994). Dynamical assessment of physiological systems and states using recurrence plot strategies. *Journal of Applied Physiology*, *76*(2):965-973.

Winter, D., Fuglevand, A., & Archer, S. (1994). Crosstalk in surface electromyography: Theoretical and practical estimates. *Journal of Electromyography and Kinesiology*, 4(1), 15–26. doi:10.1016/1050-6411(94)90023-X.Crosstalk.

Wong, Y. M., & Ng, G. Y. (2005). The double peak-to-peak analysis for determining EMG onset of muscle contraction. *Electromyography and Clinical Neurophysiology*, *45*(5), 267–271.

Worsfold, P., Smith, N., & Dyson, R. (2008). Weight transfer time moderation by Golf handicap. Chapter 9. In D. Crews & L. Rafer (Eds.), *Science and Golf V: Proceedings of the World Scientific Congress of Golf* (pp. 67–73). Phoenix, Arizona: The LPGA Foundation.

Xie, H.-B., Wang, Z.-Z., Huang, H., & Qing, C. (2003). Support vector machine in computer aided clinical electromyography. In: *Machine Learning and Cybernetics*, *2003 International Conference on* (Volume 2), 2-5 Nov, (pp. 1106–1108). doi:10.1109/ICMLC.2003.1259649.

Yan, Z., Wang, Z., & Xie, H. (2008). The application of mutual information-based feature selection and fuzzy LS-SVM-based classifier in motion classification. *Computer Methods and Programs in Biomedicine*, *90*, 275–284. doi:10.1016/j.cmpb.2008.01.003.

Yang, D., Zhao, J., Gu, Y., Wang, X., Li, N., Jiang, L., … Zhao, D. (2009). An Anthropomorphic Robot Hand Developed Based on Underactuated Mechanism and Controlled by EMG Signals. *Journal of Bionic Engineering*, *6*(3), 255–263. doi:10.1016/S1672-6529(08)60119-5.

Zbilut, J. P., & Webber Jr, C. L. (2006). Recurrence quantification analysis. In *Wiley Encyclopedia of Biomedical Engineering*. John Wiley & Sons, Inc.

Zecca, M., & Micera, S. (2002). Control of multifunctional prosthetic hands by processing the electromyographic signal. *Critical Reviews, 30*(4-6), 459–485.

Zhao, Z., Morstatter, F., Sharma, S., Alelyani, S., Anand, A., & Liu, H. (2010). Advancing feature selection research-ASU feature selection repository. Arizona State University. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Advancing+Feature+Selection+Research+-+ASU+Feature+Selection+Repository#0

Zheng, N., Barrentine, S. W., Fleisig, G. S., Andrews, J. R., & States, U. (2008). Kinematic analysis of swing in pro and amateur golfers. *International Journal of Sports Medicine*, 29(6), 487–493. doi:10.1055/s-2007-989229.

Zia,i A., & Menon, C. (2011). Comparison of regression models for estimation of isometric wrist joint torques using surface electromyography. *Journal of Neuroengineering and Rehabilitation*, 8(56), 1-12. doi:10.1186/1743-0003-8-56.

Apêndices

1. Consentimento Informado

Consentimento Informado

Tema do Estudo: Análise cinemática, cinética e EMG do *swing* no Golf

Objetivos do Estudo:
Analisar durante o *swing* efetuado com diferentes ferros (4, 7 e *pitch*) através do comportamento cinemático de todo o corpo, o comportamento cinético e o comportamento electromiográfico do tronco e membros inferiores.

Estudar o movimento que ocorre entre os vários segmentos corporais, durante um *swing* bem como as forças exercidas e também atividade muscular do membro inferior. Para tal será necessário colocar "sensores" e marcas em determinadas localizações anatómicas e realizar alguns testes, antes da recolha propriamente dita.

Procedimentos na Recolha:

- Informação ao participante sobre os procedimentos, esclarecimento de dúvidas, preenchimento questionário caracterização.
- Preparação da pele (depilação e limpeza) e colocação dos elétrodos nos músculos dos membros inferiores e tronco: Tibial anterior, Longo peroneal, Gémeos *(músculos das pernas)*, Recto Femoral, Vasto Externo, Vasto Interno, Bicípite Femoral, Semitendinoso *(músculos das coxas)*, Grande nadegueiro *(músculo da nádega)*, Eretor da espinha *(músculo das costas)*, Obliquo externo (músculo da barriga). ***(análise da atividade muscular)***
- Colocação de marcadores reflexivos: Fita com 4 marcadores na cabeça, Ombros, Cotovelo lados, Punho, 7ª vértebra cervical, 12ª vértebra torácica, Pélvis posterior, Pélvis anterior, Anca, Joelho (lateral e medial), Parte anterior da tíbia, Tornozelos (lateral e medial), Calcanhar, Base dedo grande do pé, Base dedo mínimo do pé, Parte superior e inferior do taco. ***(análise do movimento).***
- Colocação sobre o green e a plataforma de forças ***(análise das forças).***

Os elétrodos e marcas reflexivas serão fixos à pele por fita adesiva apropriada.

A tarefa a realizar consistirá na realização do *swing* da forma mais natural possível tentando obter uma eficácia máxima no total de 30 tacadas efetuadas alternadamente com o ferro 4, 7 e *pitch*. O *swing* deverá ser efetuado.

As recolhas terão lugar no laboratório de Comportamento motor da Faculdade de motricidade humana.

Estou disposto(a) a colaborar como voluntário(a) no estudo Análise do Swing com diferentes ferros.

Foi-me explicado pelos investigadores a finalidade deste estudo, bem como os seus objectivos, princípios e procedimentos, sendo que os compreendi na totalidade e os aceito.

Sei também que os meus dados vão ser tratados confidencialmente e permito que sejam usados para a investigação, mantendo-se sempre o meu anonimato e privacidade. De igual forma, sei que sou livre de recusar a participação e posso desistir a qualquer momento, não resultando daqui qualquer consequência.

Assinatura do participante:

____

Data: _____|_____|_____

2. Informação para os Participantes

Objectivo do estudo

Com o intuito de efetuar a prevenção de lesões no Golfe pretendemos fazer uma análise do *swing*, estudando para isso o comportamento cinemático, cinético de todo o corpo e electro miográfico dos membros inferiores. Ou seja, pretendemos estudar o movimento que ocorre entre os vários segmentos corporais, durante um *swing* e também catividade muscular do membro inferior. Para tal será necessário colocar "sensores" e marcas em determinadas localizações anatómicas e realizar alguns testes, antes da recolha propriamente dita que consistirá em 24 *swings* com diferentes ferros (4, 7 e *pitch*)

Sequência dos Procedimentos

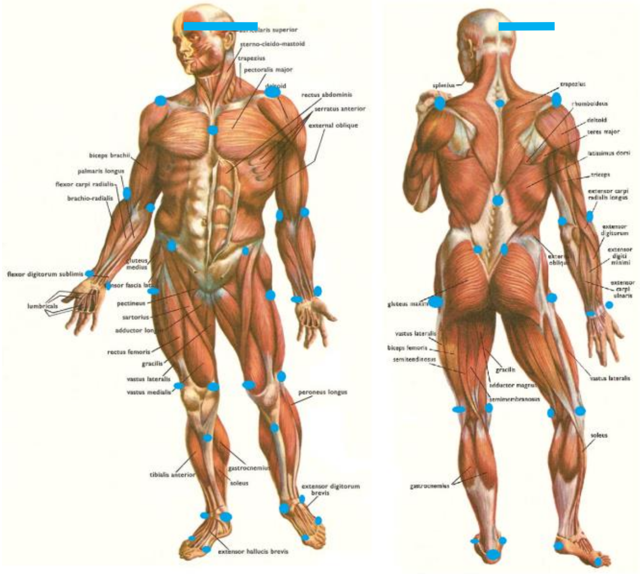
1º Depois do preenchimento de um questionário de caracterização do atleta, começamos por colocar os "sensores" de atividade muscular, que é precedida por depilação e limpeza da pele para que o sinal seja genuíno. Seguidamente precisamos saber qual a força máxima desses músculos para que seja possível efetuar comparações entre a amostra. Para isso vamos fazer-lhe os testes musculares abaixo descritos em que será aplicada resistência para contrariar o seu movimento com máxima força:

Músculo	**Testes de força**
	Sentado
Tibial anterior	Sentado na marquesa flectir o tornozelo trazendo o pé para si ao mesmo tempo que vira as plantas dos pés uma para a outra. Este teste é feito nos 2 pés ao mesmo tempo. Suportar resistência aplicada!
Peroneal longo	Deitado de lado (perna de baixo semiflectida) colocar o pé de cima em bico do pé e depois levá-lo na direcção do tecto. Suportar resistência aplicada!
Vasto medial Recto femoral Vasto lateral	Sentado na marquesa com a perna de fora fazer extensão do joelho Suportar resistência aplicada!
	Deitado de barriga para cima
Oblíquo Externo	Deitado de barriga para cima, mãos atrás da nuca, tentar chegar com o ombro de um lado à anca do lado contrário. Suportar resistência aplicada!
	Em pé
Gémeo medial Gémeo lateral	Em pé sobre um pé apoiado na parede, levantar o calcanhar do chão, mantendo o joelho esticado enquanto com os membros superiores resiste a esse movimento. Mais do que pressionar a ponta do pé deverá tentar levantar o calcanhar. Suportar resistência aplicada!
	Deitado de barriga para baixo
Grande nadegueiro	Deitado de barriga para baixo rodar a anca para fora e levantá-la da marquesa na direção do teto. Suportar resistência aplicada!
Eretor da espinha	Deitado de barriga para baixo mãos debaixo da testa, levantar o tronco da marquesa. Suportar resistência aplicada!

Bicípite femoral	Deitado de barriga para baixo fletir o joelho em teste. Suportar resistência aplicada!
Semitendinoso	

2º Para conhecermos a pressão efetuada em cada um dos pés colocaremos umas palmilhas dentro dos seus sapatos.

3º Seguidamente colocamos os "sensores" para a análise do movimento dos braços, pernas, cabeça e tronco, através da colocação de refletores esféricos nas referências anatómicas abaixo descritas:

1. Fita com 4 marcadores na cabeça 2. Ombros 3. Cotovelo lados 4. Punho 5. 7ª vértebra cervical 6. 12ª vértebra torácica 7. Pélvis posterior 8. Pélvis anterior 9. Anca 10. Joelho (lateral e medial) 11. Parte anterior da tíbia 12. Tornozelos (lateral e medial) 13. Calcanhar 14. Base dedo grande do pé 15. Base dedo mínimo do pé 16. Parte superior e inferior do taco	

4º Iremos para o *green*, onde ligaremos as palmilhas ao aparelho e o deixamos efectuar o seu aquecimento habitual.

5º Por último, para percebermos a amplitude habitual do seu movimento, vamos gravar a sua execução dos seguintes movimentos:

De pé à posição de cócoras flectindo tronco e braços e depois para de pé fazendo o máximo de extensão do tronco e braços
Em Pé, com os joelhos esticados chegar com mãos ou chão (ou até onde conseguir)
Em Pé, inclinação lateral direita e esquerda
Em Pé, com os braços afastados e os cotovelos estendidos fazer rotação máxima dta e esq (rotação global dos tornozelos à coluna)

6º Finalmente procederemos às recolhas do movimento de *swing* no total de 30 tacadas efectuadas alternadamente com o ferro 4, 7 e *pitch*. O *swing* deverá ser efectuado da forma mais natural possível tentando obter uma eficácia máxima.

Pedimos-lhe que traga uns **calções confortáveis**, os seus sapatos de treino e os seus **tacos 4, 7 e *pitch***. Prevemos que a totalidade dos procedimentos possa demorar cerca de 2h30m.
Estamos muito gratos pela sua colaboração e disponibilidade que desde já agradecemos

3. Rotinas Específicas a Conjuntos de Dados

Neste ponto apresentam-se algumas rotinas, não caracterizando o total do processamento realizado, pois tal seria demasiado extensivo. Importa compreender que estas rotinas utilizaram funções auxiliares e/ou determinada toolbox mediante o objetivo do processamento de dados. As rotinas aqui apresentadas despenderam da especificade do tratamento de dados.

```
%*****************************************************************%
%       Análise das Variáveis Temporais: Trunk Golf EMG          %
%*****************************************************************%
%               Mean plus SD with Sliding Window                 %
%          Exemplo para um sujeito com 10 repetições             %
%*****************************************************************%
% Limpar Workspace
clear all;
clc;
close all;
% Defini??o de Par?metros dos Filtros
cut1=490;cut2=10;
cutOff=12;ordem=4;
SR=1000;
% Load dos Ficheiros
suj=input('Introduza o Sujeito:','s');
lado=input('Introduza o Lado - L(left) ou R(right)','s');

if lado=='L'
    side='_1_';
elseif lado=='R'
    side='_2_';
end
ext='.txt';
nome = { [suj side 'Q01' ext]
                [suj side 'Q02' ext]
                [suj side 'Q03' ext]
                [suj side 'Q04' ext]
                [suj side 'Q05' ext]
                [suj side 'P01' ext]
                [suj side 'P02' ext]
                [suj side 'P03' ext]
                [suj side 'P04' ext]
                [suj side 'P05' ext]
                };

comp = length(nome);
% interv=zeros(10,5);

for j = 1:comp
    file = char(nome(j));
    DadosRaw = load(file);

    data = load([suj 'kin' ext]);

    for i = 4:8
        DadosRaw(:,i) = DadosRaw(:,i)-mean(DadosRaw(:,i));
    end

    DadosRaw(:,4:8) = detrend((DadosRaw(:,4:8)*5)/4096);
    RawData = DadosRaw(:,4);
    RawData(:,2:4) = DadosRaw(:,6:8);
    Lgh = length(RawData);

    trigger = DadosRaw(:,2);
    trig = diff(trigger);
    ini = find(trig==max(trig));
    fim = find(trig==min(trig));
    data = data+ini;
```

```
    DadosF = HPassF(RawData,cut1,cut2,SR,Lgh);
    DadosR = abs(DadosF);
    Dados = LPassF(DadosR,cutOff,ordem,SR);

%*****************************************************************%
%                        Cálculo do ONSET                         %
%*****************************************************************%
    % Treshold A - Linha de base em repouso
    baseline=load(['base_' suj ext]);
    tresholdA=zeros(1,4);
    if lado=='L'
        for i=1:4
            tresholdA(:,i)=baseline(1,i)+(baseline(2,i)*3);
        end
    elseif lado=='R'
        for i=1:4
            tresholdA(:,i)=baseline(1,i+4)+(baseline(2,i+4)*3);
        end
    end
        % Treshold B - Actividade EMG na tarefa: mean([-1000 : -500])
    tresholdB=zeros(1,4);
    for i=1:4
        tresholdB(:,i)=mean(Dados(data(j,1)-1000:data(j,1)-500,i))+...
            (std(Dados(data(j,1)-1000:data(j,1)-500,i))*3);
    end

% Determina??o do ONSET
onA1=zeros(4,1);
yA1=zeros(4,1);
onB1=zeros(4,1);
yB1=zeros(4,1);
onA2=zeros(4,1);
yA2=zeros(4,1);
onB2=zeros(4,1);
yB2=zeros(4,1);


onsets1=data(j,1)-150:data(j,2)-150;
onsets2=data(j,2)-150:data(j,6);

for i=1:4
    deriv(:,i) = diff(Dados(:,i));

    % Treshold A MVC
% 1
    for k = 1:length(onsets1)-50
        if (deriv(onsets1(k):onsets1(k+49),i)>0) &&...
                (mean(Dados(onsets1(k):onsets1(k+49),i))>tresholdA(i))
           onA1(i,1)=onsets1(k);
           break
        end
    end
    if onA1(i,1) == 0
        onA1(i,1) = NaN;
        yA1(i,1) = NaN;
    else
    yA1(i,1) = Dados(onA1(i,1),i);
    end
% 2
    for k=1:length(onsets2)-50
        if (deriv(onsets2(k):onsets2(k+49),i)>0) &&...
                (mean(Dados(onsets2(k):onsets2(k+49),i))>tresholdA(i))
           onA2(i,1)=onsets2(k);
           break
        end
    end
    if onA2(i,1) == 0
        onA2(i,1) = NaN;
        yA2(i,1) = NaN;
    else
    yA2(i,1) = Dados(onA2(i,1),i);
    end
    % Treshold B  Shot
% 1
    for k=1:length(onsets1)-50
        if (deriv(onsets1(k):onsets1(k+49),i)>0) &&...
                (mean(Dados(onsets1(k):onsets1(k+49),i))>tresholdB(i))
```

```
            onB1(i,1)=onsets1(k);
            break
         end
     end
     if onB1(i,1)==0
         onB1(i,1) = NaN;
         yB1(i,1)=NaN;
     else
     yB1(i,1)=Dados(onB1(i,1),i);
     end
% 2
   for k=1:length(onsets2)-50
         if (deriv(onsets2(k):onsets2(k+49),i)>0) &&....
                 (mean(Dados(onsets2(k):onsets2(k+49),i))>tresholdB(i))
            onB2(i,1)=onsets2(k);
            break
         end
     end
     if onB2(i,1)==0
         onB2(i,1) = NaN;
         yB2(i,1) = NaN;
     else
     yB2(i,1)=Dados(onB2(i,1),i);
     end
end
clear onsets1 onsets2 deriv

    for i=1:4

    name(1,1:2)=['RA'];name(2,1:2)=['EO'];...
    name(3,1:2)=['GM'];name(4,1:2)=['ES'];
    rep(1,1:3)=['Q01'];rep(2,1:3)=['Q02'];rep(3,1:3)=['Q03'];...
    rep(4,1:3)=['Q04'];rep(5,1:3)=['Q05'];
    rep(6,1:3)=['P01'];rep(7,1:3)=['P02'];rep(8,1:3)=['P03'];...
    rep(9,1:3)=['P04'];rep(10,1:3)=['P05'];


    figure; h=plot(Dados(:,i),'k','LineWidth',2);
    title([name(i,1:2) '_' lado '_' rep(j,1:3)]);
    xlabel('ms','Fontsize', 12)
    ylabel('mV','Fontsize', 12);
    vline(onA1(i),'-.r');
    vline(onB1(i),'-.g');
    hline(tresholdA(i));
    vline(onA2(i),'-r');
    vline(onB2(i),'-g');
    hline(tresholdB(i),'-.g');
    vline(data(j,:),':K')
    titulo=[suj '_' name(i,1:2) '_' lado '_' rep(j,1:3)];
    saveas(h,titulo,'fig');
    saveas(h,titulo,'jpeg');
    end

 for i=1:4
    [vmax(i) imax(i)]=max(Dados(:,i));

     vmax = vmax';
     variaveis(:,1) = onA1 - data(j,4); % onsetA1 to impact
     variaveis(:,2) = onA2 - data(j,4); % onsetA2 to impact
     variaveis(:,3) = onB1 - data(j,4); % onsetB1 to impact
     variaveis(:,4) = onB2 - data(j,4); % onsetB2 to impact
     variaveis(i,5) = imax(i) - data(j,4); % peak maximum to impact
     variaveis(:,6) = onA1 - data(j,1); % onsetA1 to Adress
     variaveis(:,7) = onB1 - data(j,1); % onsetA2 to Adress
     variaveis(:,8) = onA2 - data(j,2); % onsetB1 to DownSwing
     variaveis(:,9) = onB2 - data(j,2); % onsetB2 to DownSwing
     variaveis(i,10) = yA1(i)./vmax(i) *100; % %onsetA1
     variaveis(i,11) = yA2(i)./vmax(i) *100; % %onsetA2
     variaveis(i,12) = yB1(i)./vmax(i) *100; % %onsetB1
     variaveis(i,13) = yB2(i)./vmax(i) *100; % %onsetB2
 end

  car = {'A1_imp','A2_imp','B1_imp','B2_imp','pmax_imp',...
      'A1_BS','B1_BS','A2_DS','B2_DS','vA1','vA2','vB1','vB2'};

    % Exportar Dados para Excel
```

```
if lado=='L'
    folha='Left Muscles';
elseif lado=='R'
    folha='Right Muscles';
end

if j==1
    cell={'A2'};
elseif j==2
    cell={'A6'};
elseif j==3
    cell={'A10'};
elseif j==4
    cell={'A14'};
elseif j==5
    cell={'A18'};
elseif j==6
    cell={'A22'};
elseif j==7
    cell={'A26'};
elseif j==8
    cell={'A30'};
elseif j==9
    cell={'A34'};
elseif j==10
    cell={'A38'};
end

    xlswrite([suj 'onset_JEK'],car,folha,'A1');
    xlswrite([suj 'onset_JEK'],variaveis,folha,char(cell));

end


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%****************************************************************%
%       Preparação de dados antes da aplicação de SVM
%****************************************************************%
%               Utiliza funções adicionais
%                  Extração de features
%****************************************************************%
%
%Limpar janela de comandos%
clear all
close all
clc

%Parâmetros do tratamento de dados%
SR=1000;
dt=1/SR;
cut1=490;
cut2=10;
cutOff=12;
ordem=4;
lghfile=1000;

%Ficheiros a analisar%

p = dir('*.txt');
name = {p.name};
[~, index] = sort(name);
name = name(index);

fases = xlsread('ZABfases3.xlsx');

%%
segm=input('Tipo de segmentação (1) ou (2):  ');
window=input('Janela de amostras:  ');
caixa=1;

%%
tic;

emg = struct('tr',[], 'sg',[]);

for ii=1:length(name)
```

```
   [emg(ii).tr, emg(ii).sg] = trata_emg(name{ii}, cut1, cut2,...
   cutOff, ordem, SR);

end
%%
% Definição das matrizes segmentadas%

m=struct('ma', [], 'mb', [], 'mc', [], 'md', [], 'me', [], 'mf',[],...
    'mg',[], 'mh',[]);

if segm==1

    segmen='disjoint segmentation';

for ii=1:length(name)

    [m(ii).ma, m(ii).mb, m(ii).mc, m(ii).md, m(ii).me, m(ii).mf,...
        m(ii).mg, m(ii).mh,]=seg(emg{ii}.sg, 8, window, segm);

end

    elseif segm==2

    segmen='overlapped segmentation';

lag=input('Qual o lag:   ');


for ii=1:length(name)
    [m(ii).ma, m(ii).mb, m(ii).mc, m(ii).md, m(ii).me, m(ii).mf,...
        m(ii).mg, m(ii).mh,]=seg(emg(ii).sg, 8, window, segm, lag);

end

end
 %%
 %Sincronização com as fases%

 sinc=struct('dif',[],'scn', [], 'col', [], 'n',[], 'sc',[]);

for jj=1:length(name)

    [sinc(jj).dif, sinc(jj).scn] = sincro_H(fases, emg(jj).tr, jj);
        sinc(jj).n = length(emg(jj).tr)./length(m(jj).ma);
        sinc(jj).sc = (sinc(jj).scn)./(sinc(jj).n);

end

save('sincr.mat','sinc');

%%
%EXTRACÇÃO DE FEATURES%

%Preparação para o vector classificação%

x=struct('xa',[],'xb',[],'xc',[],'xd',[],'xe',[],'xf',[],...
    'xg',[], 'xh', []);
y=struct('ya',[],'yb',[],'yc',[],'yd',[],'ye',[],'yf',[],...
    'yg',[], 'yh', []);

%Mean absolute value%


mav=struct('mava',[], 'mavb',[], 'mavc',[], 'mavd',[], 'mave',[],...
    'mavf',[], 'mavg',[], 'mavh',[]);

mav_crt=struct('mav_a', [], 'mav_b', [], 'mav_c', [],...
    'mav_d', [], 'mav_e', [], 'mav_f', [],...
    'mav_g', [],'mav_h', []);


for ii=1:length(name)

   mav(ii).mava = feat(m(ii).ma, 1, window);
   mav(ii).mavb = feat(m(ii).mb, 1, window);
   mav(ii).mavc = feat(m(ii).mc, 1, window);
```

```
   mav(ii).mavd = feat(m(ii).md, 1, window);
   mav(ii).mave = feat(m(ii).me, 1, window);
   mav(ii).mavf = feat(m(ii).mf, 1, window);
   mav(ii).mavg = feat(m(ii).mg, 1, window);
   mav(ii).mavh = feat(m(ii).mh, 1, window);

   [mav_crt(ii).mav_a, x(ii).xa, y(ii).ya, sinc(ii).nf] = ...
       plotafeatgin(mav(ii).mava,sinc(ii).sc, name{ii}, lghfile, 1, 1);
   [mav_crt(ii).mav_b, x(ii).xb, y(ii).yb] = plotafeatgin(mav(ii).mavb,...
       sinc(ii).sc, name{ii}, lghfile, 1, 2);
   [mav_crt(ii).mav_c, x(ii).xc, y(ii).yc] = plotafeatgin(mav(ii).mavc,...
       sinc(ii).sc, name{ii}, lghfile, 1, 3);
   [mav_crt(ii).mav_d, x(ii).xd, y(ii).yd] = plotafeatgin(mav(ii).mavd,...
       sinc(ii).sc, name{ii}, lghfile, 1, 4);
   [mav_crt(ii).mav_e, x(ii).xe, y(ii).ye] = plotafeatgin(mav(ii).mave,...
       sinc(ii).sc, name{ii}, lghfile, 1, 5);
   [mav_crt(ii).mav_f, x(ii).xf, y(ii).yf] = plotafeatgin(mav(ii).mavf,...
       sinc(ii).sc, name{ii}, lghfile, 1, 6);
   [mav_crt(ii).mav_g, x(ii).xg, y(ii).yg] = plotafeatgin(mav(ii).mavg,...
       sinc(ii).sc, name{ii}, lghfile, 1, 7);
   [mav_crt(ii).mav_h, x(ii).xh, y(ii).yh] = plotafeatgin(mav(ii).mavh,...
       sinc(ii).sc, name{ii}, lghfile, 1, 8);

   close all

end

Mmava = [mav_crt(1:length(name)).mav_a]';
Mmavb = [mav_crt(1:length(name)).mav_b]';
Mmavc = [mav_crt(1:length(name)).mav_c]';
Mmavd = [mav_crt(1:length(name)).mav_d]';
Mmave = [mav_crt(1:length(name)).mav_e]';
Mmavf = [mav_crt(1:length(name)).mav_f]';
Mmavg = [mav_crt(1:length(name)).mav_g]';
Mmavh = [mav_crt(1:length(name)).mav_h]';

Mmava = Mmava./max(Mmava);
Mmavb = Mmavb./max(Mmavb);
Mmavc = Mmavc./max(Mmavc);
Mmavd = Mmavd./max(Mmavd);
Mmave = Mmave./max(Mmave);
Mmavf = Mmavf./max(Mmavf);
Mmavg = Mmavg./max(Mmavg);
Mmavh = Mmavh./max(Mmavh);


close all

%%
%Wave form length %

wl=struct('wla',[], 'wlb',[], 'wlc',[], 'wld',[], 'wle',[],...
    'wlf',[], 'wlg',[], 'wlh',[]);

wl_crt=struct('wl_a', [], 'wl_b', [], 'wl_c', [],...
    'wl_d', [], 'wl_e', [], 'wl_f', [],...
    'wl_g', [],'wl_h', []);


for ii=1:length(name)

   wl(ii).wla = feat(m(ii).ma, 3, window);
   wl(ii).wlb = feat(m(ii).mb, 3, window);
   wl(ii).wlc = feat(m(ii).mc, 3, window);
   wl(ii).wld = feat(m(ii).md, 3, window);
   wl(ii).wle = feat(m(ii).me, 3, window);
   wl(ii).wlf = feat(m(ii).mf, 3, window);
   wl(ii).wlg = feat(m(ii).mg, 3, window);
   wl(ii).wlh = feat(m(ii).mh, 3, window);

   [wl_crt(ii).wl_a, sinc(ii).nf] = plotafeat(wl(ii).wla,...
       sinc(ii).sc, name{ii}, lghfile, 3, 1);
   wl_crt(ii).wl_b = plotafeat(wl(ii).wlb, sinc(ii).sc, name{ii},...
       lghfile, 3, 2);
   wl_crt(ii).wl_c = plotafeat(wl(ii).wlc, sinc(ii).sc, name{ii},...
       lghfile, 3, 3);
   wl_crt(ii).wl_d = plotafeat(wl(ii).wld, sinc(ii).sc, name{ii},...
```

```
        lghfile, 3, 4);
    wl_crt(ii).wl_e = plotafeat(wl(ii).wle, sinc(ii).sc, name{ii},...
        lghfile, 3, 5);
    wl_crt(ii).wl_f = plotafeat(wl(ii).wlf, sinc(ii).sc, name{ii},...
        lghfile, 3, 6);
    wl_crt(ii).wl_g = plotafeat(wl(ii).wlg, sinc(ii).sc, name{ii},...
        lghfile, 3, 7);
    wl_crt(ii).wl_h = plotafeat(wl(ii).wlh, sinc(ii).sc, name{ii},...
        lghfile, 3, 8);

    close all

end

Mwla = [wl_crt(1:length(name)).wl_a]';
Mwlb = [wl_crt(1:length(name)).wl_b]';
Mwlc = [wl_crt(1:length(name)).wl_c]';
Mwld = [wl_crt(1:length(name)).wl_d]';
Mwle = [wl_crt(1:length(name)).wl_e]';
Mwlf = [wl_crt(1:length(name)).wl_f]';
Mwlg = [wl_crt(1:length(name)).wl_g]';
Mwlh = [wl_crt(1:length(name)).wl_h]';

Mwla = Mwla./max(Mwla);
Mwlb = Mwlb./max(Mwlb);
Mwlc = Mwlc./max(Mwlc);
Mwld = Mwld./max(Mwld);
Mwle = Mwle./max(Mwle);
Mwlf = Mwlf./max(Mwlf);
Mwlg = Mwlg./max(Mwlg);
Mwlh = Mwlh./max(Mwlh);

%%
%Difference absolute standart deviation value%

dsd=struct('dsda',[], 'dsdb',[], 'dsdc',[], 'dsdd',[], 'dsde',[],...
    'dsdf',[], 'dsdg',[], 'dsdh',[]);

dsd_crt=struct('dsd_a', [], 'dsd_b', [], 'dsd_c', [],...
    'dsd_d', [], 'dsd_e', [], 'dsd_f', [],...
    'dsd_g', [],'dsd_h', []);


for ii=1:length(name)

    dsd(ii).dsda = feat(m(ii).ma, 4, window);
    dsd(ii).dsdb = feat(m(ii).mb, 4, window);
    dsd(ii).dsdc = feat(m(ii).mc, 4, window);
    dsd(ii).dsdd = feat(m(ii).md, 4, window);
    dsd(ii).dsde = feat(m(ii).me, 4, window);
    dsd(ii).dsdf = feat(m(ii).mf, 4, window);
    dsd(ii).dsdg = feat(m(ii).mg, 4, window);
    dsd(ii).dsdh = feat(m(ii).mh, 4, window);

    [dsd_crt(ii).dsd_a, sinc(ii).nf] = plotafeat(dsd(ii).dsda,...
        sinc(ii).sc, name{ii}, lghfile, 4, 1);
    dsd_crt(ii).dsd_b = plotafeat(dsd(ii).dsdb, sinc(ii).sc, name{ii},...
        lghfile, 4, 2);
    dsd_crt(ii).dsd_c = plotafeat(dsd(ii).dsdc, sinc(ii).sc, name{ii},...
        lghfile, 4, 3);
    dsd_crt(ii).dsd_d = plotafeat(dsd(ii).dsdd, sinc(ii).sc, name{ii},...
        lghfile, 4, 4);
    dsd_crt(ii).dsd_e = plotafeat(dsd(ii).dsde, sinc(ii).sc, name{ii},...
        lghfile, 4, 5);
    dsd_crt(ii).dsd_f = plotafeat(dsd(ii).dsdf, sinc(ii).sc, name{ii},...
        lghfile, 4, 6);
    dsd_crt(ii).dsd_g = plotafeat(dsd(ii).dsdg, sinc(ii).sc, name{ii},...
        lghfile, 4, 7);
    dsd_crt(ii).dsd_h = plotafeat(dsd(ii).dsdh, sinc(ii).sc, name{ii},...
        lghfile, 4, 8);

    close all

end

Mdsda = [dsd_crt(1:length(name)).dsd_a]';
Mdsdb = [dsd_crt(1:length(name)).dsd_b]';
```

```
Mdsdc = [dsd_crt(1:length(name)).dsd_c]';
Mdsdd = [dsd_crt(1:length(name)).dsd_d]';
Mdsde = [dsd_crt(1:length(name)).dsd_e]';
Mdsdf = [dsd_crt(1:length(name)).dsd_f]';
Mdsdg = [dsd_crt(1:length(name)).dsd_g]';
Mdsdh = [dsd_crt(1:length(name)).dsd_h]';

Mdsda = Mdsda./max(Mdsda);
Mdsdb = Mdsdb./max(Mdsdb);
Mdsdc = Mdsdc./max(Mdsdc);
Mdsdd = Mdsdd./max(Mdsdd);
Mdsde = Mdsde./max(Mdsde);
Mdsdf = Mdsdf./max(Mdsdf);
Mdsdg = Mdsdg./max(Mdsdg);
Mdsdh = Mdsdh./max(Mdsdh);

%%
%%variancia N-1%%

var=struct('vara',[], 'varb',[], 'varc',[], 'vard',[], 'vare',[],...
    'varf',[], 'varg',[], 'varh',[]);

var_crt=struct('var_a', [], 'var_b', [], 'var_c', [],...
    'var_d', [], 'var_e', [], 'var_f', [],...
    'var_g', [],'var_h', []);


for ii=1:length(name)

   var(ii).vara = feat(m(ii).ma, 5, window);
   var(ii).varb = feat(m(ii).mb, 5, window);
   var(ii).varc = feat(m(ii).mc, 5, window);
   var(ii).vard = feat(m(ii).md, 5, window);
   var(ii).vare = feat(m(ii).me, 5, window);
   var(ii).varf = feat(m(ii).mf, 5, window);
   var(ii).varg = feat(m(ii).mg, 5, window);
   var(ii).varh = feat(m(ii).mh, 5, window);

   [var_crt(ii).var_a, sinc(ii).nf] = plotafeat(var(ii).vara,...
       sinc(ii).sc, name{ii}, lghfile, 5, 1);
   var_crt(ii).var_b = plotafeat(var(ii).varb, sinc(ii).sc, name{ii},...
       lghfile, 5, 2);
   var_crt(ii).var_c = plotafeat(var(ii).varc, sinc(ii).sc, name{ii},...
       lghfile, 5, 3);
   var_crt(ii).var_d = plotafeat(var(ii).vard, sinc(ii).sc, name{ii},...
       lghfile, 5, 4);
   var_crt(ii).var_e = plotafeat(var(ii).vare, sinc(ii).sc, name{ii},...
       lghfile, 5, 5);
   var_crt(ii).var_f = plotafeat(var(ii).varf, sinc(ii).sc, name{ii},...
       lghfile, 5, 6);
   var_crt(ii).var_g = plotafeat(var(ii).varg, sinc(ii).sc, name{ii},...
       lghfile, 5, 7);
   var_crt(ii).var_h = plotafeat(var(ii).varh, sinc(ii).sc, name{ii},...
       lghfile, 5, 8);

   close all

end

Mvara = [var_crt(1:length(name)).var_a]';
Mvarb = [var_crt(1:length(name)).var_b]';
Mvarc = [var_crt(1:length(name)).var_c]';
Mvard = [var_crt(1:length(name)).var_d]';
Mvare = [var_crt(1:length(name)).var_e]';
Mvarf = [var_crt(1:length(name)).var_f]';
Mvarg = [var_crt(1:length(name)).var_g]';
Mvarh = [var_crt(1:length(name)).var_h]';

Mvara = Mvara./max(Mvara);
Mvarb = Mvarb./max(Mvarb);
Mvarc = Mvarc./max(Mvarc);
Mvard = Mvard./max(Mvard);
Mvare = Mvare./max(Mvare);
Mvarf = Mvarf./max(Mvarf);
Mvarg = Mvarg./max(Mvarg);
Mvarh = Mvarh./max(Mvarh);
```

```
%%
%Integrated of EMG%
iemg=struct('iemga',[], 'iemgb',[], 'iemgc',[], 'iemgd',[], 'iemge',[],...
    'iemgf',[], 'iemgg',[], 'iemgh',[]);

iemg_crt=struct('iemg_a', [], 'iemg_b', [], 'iemg_c', [],...
    'iemg_d', [], 'iemg_e', [], 'iemg_f', [],...
    'iemg_g', [],'iemg_h', []);


for ii=1:length(name)

   iemg(ii).iemga = feat(m(ii).ma, 5, window);
   iemg(ii).iemgb = feat(m(ii).mb, 5, window);
   iemg(ii).iemgc = feat(m(ii).mc, 5, window);
   iemg(ii).iemgd = feat(m(ii).md, 5, window);
   iemg(ii).iemge = feat(m(ii).me, 5, window);
   iemg(ii).iemgf = feat(m(ii).mf, 5, window);
   iemg(ii).iemgg = feat(m(ii).mg, 5, window);
   iemg(ii).iemgh = feat(m(ii).mh, 5, window);

   [iemg_crt(ii).iemg_a, sinc(ii).nf] = plotafeat(iemg(ii).iemga,...
       sinc(ii).sc, name{ii}, lghfile, 5, 1);
   iemg_crt(ii).iemg_b = plotafeat(iemg(ii).iemgb, sinc(ii).sc, name{ii},...
       lghfile, 5, 2);
   iemg_crt(ii).iemg_c = plotafeat(iemg(ii).iemgc, sinc(ii).sc, name{ii},...
       lghfile, 5, 3);
   iemg_crt(ii).iemg_d = plotafeat(iemg(ii).iemgd, sinc(ii).sc, name{ii},...
       lghfile, 5, 4);
   iemg_crt(ii).iemg_e = plotafeat(iemg(ii).iemge, sinc(ii).sc, name{ii},...
       lghfile, 5, 5);
   iemg_crt(ii).iemg_f = plotafeat(iemg(ii).iemgf, sinc(ii).sc, name{ii},...
       lghfile, 5, 6);
   iemg_crt(ii).iemg_g = plotafeat(iemg(ii).iemgg, sinc(ii).sc, name{ii},...
       lghfile, 5, 7);
   iemg_crt(ii).iemg_h = plotafeat(iemg(ii).iemgh, sinc(ii).sc, name{ii},...
       lghfile, 5, 8);

   close all

end

Miemga = [iemg_crt(1:length(name)).iemg_a]';
Miemgb = [iemg_crt(1:length(name)).iemg_b]';
Miemgc = [iemg_crt(1:length(name)).iemg_c]';
Miemgd = [iemg_crt(1:length(name)).iemg_d]';
Miemge = [iemg_crt(1:length(name)).iemg_e]';
Miemgf = [iemg_crt(1:length(name)).iemg_f]';
Miemgg = [iemg_crt(1:length(name)).iemg_g]';
Miemgh = [iemg_crt(1:length(name)).iemg_h]';

Miemga = Miemga./max(Miemga);
Miemgb = Miemgb./max(Miemgb);
Miemgc = Miemgc./max(Miemgc);
Miemgd = Miemgd./max(Miemgd);
Miemge = Miemge./max(Miemge);
Miemgf = Miemgf./max(Miemgf);
Miemgg = Miemgg./max(Miemgg);
Miemgh = Miemgh./max(Miemgh);

%%
%logaritmic detector%

log=struct('loga',[], 'logb',[], 'logc',[], 'logd',[], 'loge',[],...
    'logf',[], 'logg',[], 'logh',[]);

log_crt=struct('log_a', [], 'log_b', [], 'log_c', [],...
    'log_d', [], 'log_e', [], 'log_f', [],...
    'log_g', [],'log_h', []);


for ii=1:length(name)

   log(ii).loga = feat(m(ii).ma, 5, window);
   log(ii).logb = feat(m(ii).mb, 5, window);
   log(ii).logc = feat(m(ii).mc, 5, window);
   log(ii).logd = feat(m(ii).md, 5, window);
```

```
   log(ii).loge = feat(m(ii).me, 5, window);
   log(ii).logf = feat(m(ii).mf, 5, window);
   log(ii).logg = feat(m(ii).mg, 5, window);
   log(ii).logh = feat(m(ii).mh, 5, window);

   [log_crt(ii).log_a, sinc(ii).nf] = plotafeat(log(ii).loga,...
       sinc(ii).sc, name{ii}, lghfile, 5, 1);
   log_crt(ii).log_b = plotafeat(log(ii).logb, sinc(ii).sc, name{ii},...
       lghfile, 5, 2);
   log_crt(ii).log_c = plotafeat(log(ii).logc, sinc(ii).sc, name{ii},...
       lghfile, 5, 3);
   log_crt(ii).log_d = plotafeat(log(ii).logd, sinc(ii).sc, name{ii},...
       lghfile, 5, 4);
   log_crt(ii).log_e = plotafeat(log(ii).loge, sinc(ii).sc, name{ii},...
       lghfile, 5, 5);
   log_crt(ii).log_f = plotafeat(log(ii).logf, sinc(ii).sc, name{ii},...
       lghfile, 5, 6);
   log_crt(ii).log_g = plotafeat(log(ii).logg, sinc(ii).sc, name{ii},...
       lghfile, 5, 7);
   log_crt(ii).log_h = plotafeat(log(ii).logh, sinc(ii).sc, name{ii},...
       lghfile, 5, 8);

   close all

end

Mloga = [log_crt(1:length(name)).log_a]';
Mlogb = [log_crt(1:length(name)).log_b]';
Mlogc = [log_crt(1:length(name)).log_c]';
Mlogd = [log_crt(1:length(name)).log_d]';
Mloge = [log_crt(1:length(name)).log_e]';
Mlogf = [log_crt(1:length(name)).log_f]';
Mlogg = [log_crt(1:length(name)).log_g]';
Mlogh = [log_crt(1:length(name)).log_h]';

Mloga = Mloga./max(Mloga);
Mlogb = Mlogb./max(Mlogb);
Mlogc = Mlogc./max(Mlogc);
Mlogd = Mlogd./max(Mlogd);
Mloge = Mloge./max(Mloge);
Mlogf = Mlogf./max(Mlogf);
Mlogg = Mlogg./max(Mlogg);
Mlogh = Mlogh./max(Mlogh);

%%
%Vector Classificação%

cly=struct('cla',[], 'clb',[], 'clc',[], 'cld', [], 'cle', [], 'clf',[],...
    'clg',[], 'clh', []);

 for jj=1:length(name)

    len=length(mav_crt(jj).mav_a);

    cly(jj).cla=ones(1,len);

    for ii=1:len

    if  ii <= x(jj).xa

    cly(jj).cla(1,ii) = -1;

    else

         cly(jj).cla(1,ii) = 1;

    end
    end

 end

 for jj=1:length(name)

    len=length(mav_crt(jj).mav_b);

    cly(jj).clb=ones(1,len);
```

```
   for ii=1:len
   if  ii <= x(jj).xb
   cly(jj).clb(1,ii) = -1;
   else
        cly(jj).clb(1,ii) = 1;
   end
   end
end
for jj=1:length(name)
   len=length(mav_crt(jj).mav_c);
   cly(jj).clc=ones(1,len);
   for ii=1:len
   if  ii <= x(jj).xc
   cly(jj).clc(1,ii) = -1;
   else
        cly(jj).clc(1,ii) = 1;
   end
   end
end
for jj=1:length(name)
   len=length(mav_crt(jj).mav_d);
   cly(jj).cld=ones(1,len);
   for ii=1:len
   if  ii <= x(jj).xd
   cly(jj).cld(1,ii) = -1;
   else
        cly(jj).cld(1,ii) = 1;
   end
   end
end
for jj=1:length(name)
   len=length(mav_crt(jj).mav_e);
   cly(jj).cle=ones(1,len);
   for ii=1:len
   if  ii <= x(jj).xe
   cly(jj).cle(1,ii) = -1;
   else
        cly(jj).cle(1,ii) = 1;
   end
   end
end
```

```
 for jj=1:length(name)

    len=length(mav_crt(jj).mav_f);

    cly(jj).clf=ones(1,len);

    for ii=1:len

    if  ii <= x(jj).xf

    cly(jj).clf(1,ii) = -1;

    else

         cly(jj).clf(1,ii) = 1;

    end
    end

 end

 for jj=1:length(name)

    len=length(mav_crt(jj).mav_g);

    cly(jj).clg=ones(1,len);

    for ii=1:len

    if  ii <= x(jj).xg

    cly(jj).clg(1,ii) = -1;

    else

         cly(jj).clg(1,ii) = 1;

    end
    end

 end

 for jj=1:length(name)

    len=length(mav_crt(jj).mav_h);

    cly(jj).clh=ones(1,len);

    for ii=1:len

    if  ii <= x(jj).xh

    cly(jj).clh(1,ii) = -1;

    else

         cly(jj).clh(1,ii) = 1;

    end
    end

 end
 Mclassa=[cly(1:length(name)).cla]';
  Mclassb=[cly(1:length(name)).clb]';
   Mclassc=[cly(1:length(name)).clc]';
    Mclassd=[cly(1:length(name)).cld]';
     Mclasse=[cly(1:length(name)).cle]';
      Mclassf=[cly(1:length(name)).clf]';
       Mclassg=[cly(1:length(name)).clg]';
        Mclassh=[cly(1:length(name)).clh]';
 save('onsetVI.mat','x');
%%
%MÚSCULO A%

MMa=[Mclassa Mmava Mwla Mdsda Mvara Miemga Mloga];
save('MMa.mat','Mclassa', 'Mmava', 'Mwla', 'Mdsda','Mvara',...
    'Miemga', 'Mloga');
```

```
%MÚSCULO B%

MMb=[Mclassb Mmavb Mwlb Mdsdb Mvarb Miemgb Mlogb];
save('MMb.mat','Mclassb', 'Mmavb', 'Mwlb', 'Mdsdb','Mvarb',...
    'Miemgb', 'Mlogb');

%MÚSCULO C%

MMc=[Mclassc Mmavc Mwlc Mdsdc Mvarc Miemgc Mlogc];
save('MMc.mat','Mclassc', 'Mmavc', 'Mwlc', 'Mdsdc','Mvarc',...
    'Miemgc', 'Mlogc');


%MÚSCULO D%

MMd=[Mclassd Mmavd Mwld Mdsdd Mvard Miemgd Mlogd];
save('MMd.mat','Mclassd', 'Mmavd', 'Mwld', 'Mdsdd','Mvard',...
    'Miemgd', 'Mlogd');

%MÚSCULO E%

MMe=[Mclasse Mmave Mwle Mdsde Mvare Miemge Mloge];
save('MMe.mat','Mclasse', 'Mmave', 'Mwle', 'Mdsde','Mvare',...
    'Miemge', 'Mloge');

%MÚSCULO F%

MMf=[Mclassf Mmavf Mwlf Mdsdf Mvarf Miemgf Mlogf];
save('MMf.mat','Mclassf', 'Mmavf', 'Mwlf', 'Mdsdf','Mvarf',...
    'Miemgf', 'Mlogf');

%MÚSCULO G%

MMg=[Mclassg Mmavg Mwlg Mdsdg Mvarg Miemgg Mlogg];
save('MMg.mat','Mclassg', 'Mmavg', 'Mwlg', 'Mdsdg','Mvarg',...
    'Miemgg', 'Mlogg');

%MÚSCULO H%

MMh=[Mclassh Mmavh Mwlh Mdsdh Mvarh Miemgh Mlogh];
save('MMh.mat','Mclassh', 'Mmavh', 'Mwlh', 'Mdsdh','Mvarh',...
    'Miemgh', 'Mlogh');

toc
%%
site = 'http://www.fmh.utl.pt/labcmotor/';
title = 'Laboratório de Comportamento Motor - Faculdade de Motricidade Humana';
fprintf('<a href = "%s">%s</a>\n',site,title);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function [Acc bestsigma bestC] = BestParamRBF(TrainSample, TrainLabel,...
    sigma, C, nfolds)

%k-fold crossavalidation
% Explicação em http://www.youtube.com/watch?v=4GB5HVw0CuM%

ind = crossvalind('kfold', TrainLabel, nfolds);
opIT = optimset('maxIter',10000000);
accM = zeros(length(C),length(sigma));

for jj=1:length(C)
    jj

for kk=1:length(sigma)
    kk

for ii=1:nfolds
    ii

    TestfoldSample = TrainSample(ind==ii,:);
    TrainfoldSample = TrainSample(ind~=ii,:);
    TrainfoldLabel = TrainLabel(ind~=ii,:);
```

```
        svmStruct = svmtrain(TrainfoldSample, TrainfoldLabel,...
            'method','SMO','options',opIT,...
            'showplot', false,'kernel_function', 'rbf', 'rbf_sigma',...
            sigma(kk), 'boxconstraint', C(jj));

    outLabel(ind==ii,1) = svmclassify(svmStruct, TestfoldSample,...
        'showplot', false);

end

    accM(jj,kk) = sum(grp2idx(outLabel)==grp2idx(TrainLabel))/...
        length(TrainLabel);

end

end

Acc=accM*100;

[Mcol, Icol] = max(accM, [],1);
[~, Irow] = max(Mcol,[],2);
bestsigma = sigma(Irow);
bestC = C(Icol(Irow));

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%                                                                       %
% Rotina Grid-search dos parametros RBF-C para SVM não linear           %
%             Grava os modelos para poderem ser usados.                 %
%                                                                       %
% O valor de p refere-se ao número de dados teste que serão retirados %
% do total amostral.                                                    %
%                                                                       %
% Necessita trabalhar em conjunto com a função BestParamRBF()           %
% "xdata" deve ser introduzido como um vector linha com várias colunas%
% como exemplo [features_1 features_2 ... features_n]                   %
%                                                                       %
% group corresponde ao vector coluna de classificação {-1,1}            %
%                                                                       %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%
clear all
close all
clc

%%
fil=input('Ficheiro para classificação:  ','s');
ext='.mat';
load(char([fil ext]));
xdata=input('Introduza a matriz de features:  ');
group=input('Introduza a matriz label:  ');

tic
%%
p=0.2;
[Train, Test] = crossvalind('HoldOut', group, p);

TrainSample = xdata(Train,:);
TrainLabel = group(Train,1);
TestSample = xdata(Test,:);
TestLabel = group(Test,1);

nfolds=5;

sigma = 2.^(-5:0.25:-4);
C = 2.^(7.75:0.25:8.75);

opIT = optimset('maxIter',10000000);

[AA, bestsigma, bestC] = BestParamRBF(TrainSample, TrainLabel,...
    sigma, C, nfolds);
%%
figure;
```

```
grafico = plot(AA);

%%
%Salvar modelos estrutura%

StructRBF = svmtrain(TrainSample, TrainLabel,'method','SMO',...
        'options',opIT,'kernel_function', 'rbf', 'rbf_sigma',...
        bestsigma, 'boxconstraint', bestC,'showplot', false);

  vLabelRBF = svmclassify(StructRBF, TestSample,...
      'showplot', false);

save('RBFSVM_2fF6_5x5_Thc.mat', '-struct','StructRBF');
save('dados_2fF6_5x5_Thc.mat','AA','bestsigma','bestC','sigma','C','vLabelRBF');
%%
toc

 site = 'http://www.fmh.utl.pt/labcmotor/';
    title = 'Laboratório de Comportamento Motor - FMH';
        fprintf('<a href = "%s">%s</a>\n',site,title);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Rotina de tratamento dos sinais EMG para classificar grupos segundo o
% Handicap e a incidência de lombalgia após a realização de 18 buracos,
% tendo em conta a fase de preparação (backswing, downswing e
% follow-through. A classificação é realizada por Máquinas de Vector
% Suporte (Support Vector Machines).
%
%      1ª Parte tratamento do sinal EMG por inteiro e por fases para
%      posteriormente ser processada a análise de recurrencia.
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% As repetições de cada sujeito são inseridas extraindo features através de
% recurrence quantification analysis (RAQ) com segunda rotina e após
% escalonamento no tempo.
% Este exemplo está preparado em concordância com os sujeitos de LBP devido
% aos procedimentos de normalização em amplitude.
%
% Necessita de funções paralelas
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
tic
%%
% Limpar o Workspace%
clear all
close all
clc

%%
%Parâmetros para o tratamento de dados%
SR=1000;
dt=1/SR;
cut1=490;
cut2=10;
cutOff=12;
ordem=4;
lghfile=1000;
nfolds = 5;

%%
% Importar dados %
n = dir('*.txt');
name = {n.name};
[~, index] = sort(name);
name = name(index);
comp = length(name);

%%
% Label p/ SVM %
labelHc = ones(comp,1);
labelHc(1:102) = 1;
labelHc(103:end) = -1;
save('label.mat','labelHc');
```

```
%%
%%%%%%%%%%% 1ª PARTE: TRATAMENTO EMG PARA PROCESSAMENTO RQA %%%%%%%%%%%%%%
% Cortar a parte do sinal que interessa (já inclui o tratamento do sinal)%
% através da função trata_emgR. O sinal é tratado, filtrado, rectificado e%
% normalizado 0-1 %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
emg = struct('tr',[], 'sign',[],'ini',[], 'fim',[],'data',[]);

for ii = 1:comp

   [emg(ii).tr, ~, emg(ii).sign] = trata_emgR(name{ii}, cut1, cut2, SR);
    emg(ii).tr = diff(emg(ii).tr);
        emg(ii).ini = find(emg(ii).tr==max(emg(ii).tr));
        emg(ii).fim = find(emg(ii).tr==min(emg(ii).tr));
           emg(ii).data(:,1:8) = emg(ii).sign(emg(ii).ini:emg(ii).fim,1:8);

end

%%
% Determinar o máximo das repetições%
maxi = struct('a',[],'b',[],'c',[],'d', [],'e',[],'f',[],'g',[],'h',[]);

f1 = 14; %input('Last file suj1:  ');
f2 = 30; %input('Last file suj2:  ');
f3 = 46; %input('Last file suj3:  ');
f4 = 58; %input('Last file suj4:  ');
f5 = 74; %input('Last file suj5:  ');
f6 = 89; %input('Last file suj6:  ');
f7 = 102; %input('Last file suj7:  ');
f8 = 115; %input('Last file suj8:  ');
f9 = 131; %input('Last file suj9:  ');
f10 = 146; %input('Last file suj10:  ');
f11 = 163; %input('Last file suj11:  ');

%%
% Músculo A %
for i = 1:f1
    maxi(i).a = max(emg(i).data(:,1));
end
for i = f1+1:f2
    maxi(i).a = max(emg(i).data(:,1));
end
for i = f2+1:f3
    maxi(i).a = max(emg(i).data(:,1));
end
for i = f3+1:f4
    maxi(i).a = max(emg(i).data(:,1));
end
for i = f4+1:f5
    maxi(i).a = max(emg(i).data(:,1));
end
for i = f5+1:f6
    maxi(i).a = max(emg(i).data(:,1));
end
for i = f6+1:f7
    maxi(i).a = max(emg(i).data(:,1));
end
for i = f7+1:f8
    maxi(i).a = max(emg(i).data(:,1));
end
for i = f8+1:f9
    maxi(i).a = max(emg(i).data(:,1));
end
for i = f9+1:f10
    maxi(i).a = max(emg(i).data(:,1));
end
for i = f10+1:f11
    maxi(i).a = max(emg(i).data(:,1));
end
for i = f11+1:comp
    maxi(i).a = max(emg(i).data(:,1));
end

maxi_a = zeros(comp,1);

for k = 1:comp
```

```
    maxi_a(k,1) = maxi(k).a;
end

%%
% Músculo B %
for i = 1:f1
    maxi(i).b = max(emg(i).data(:,2));
end
for i = f1+1:f2
    maxi(i).b = max(emg(i).data(:,2));
end
for i = f2+1:f3
    maxi(i).b = max(emg(i).data(:,2));
end
for i = f3+1:f4
    maxi(i).b = max(emg(i).data(:,2));
end
for i = f4+1:f5
    maxi(i).b = max(emg(i).data(:,2));
end
for i = f5+1:f6
    maxi(i).b = max(emg(i).data(:,2));
end
for i = f6+1:f7
    maxi(i).b = max(emg(i).data(:,2));
end
for i = f7+1:f8
    maxi(i).b = max(emg(i).data(:,2));
end
for i = f8+1:f9
    maxi(i).b = max(emg(i).data(:,2));
end
for i = f9+1:f10
    maxi(i).b = max(emg(i).data(:,2));
end
for i = f10+1:f11
    maxi(i).b = max(emg(i).data(:,2));
end
for i = f11+1:comp
    maxi(i).b = max(emg(i).data(:,2));
end

maxi_b = zeros(comp,1);

for k = 1:comp

    maxi_b(k,1) = maxi(k).b;
end

%%
% Músculo C %
for i = 1:f1
    maxi(i).c = max(emg(i).data(:,3));
end
for i = f1+1:f2
    maxi(i).c = max(emg(i).data(:,3));
end
for i = f2+1:f3
    maxi(i).c = max(emg(i).data(:,3));
end
for i = f3+1:f4
    maxi(i).c = max(emg(i).data(:,3));
end
for i = f4+1:f5
    maxi(i).c = max(emg(i).data(:,3));
end
for i = f5+1:f6
    maxi(i).c = max(emg(i).data(:,3));
end
for i = f6+1:f7
    maxi(i).c = max(emg(i).data(:,3));
end
for i = f7+1:f8
    maxi(i).c = max(emg(i).data(:,3));
end
for i = f8+1:f9
```

```
    maxi(i).c = max(emg(i).data(:,3));
end
for i = f9+1:f10
    maxi(i).c = max(emg(i).data(:,3));
end
for i = f10+1:f11
    maxi(i).c = max(emg(i).data(:,3));
end
for i = f11+1:comp
    maxi(i).c = max(emg(i).data(:,3));
end

maxi_c = zeros(comp,1);

for k = 1:comp

    maxi_c(k,1) = maxi(k).c;
end

%%
% Músculo D %
for i = 1:f1
    maxi(i).d = max(emg(i).data(:,4));
end
for i = f1+1:f2
    maxi(i).d = max(emg(i).data(:,4));
end
for i = f2+1:f3
    maxi(i).d = max(emg(i).data(:,4));
end
for i = f3+1:f4
    maxi(i).d = max(emg(i).data(:,4));
end
for i = f4+1:f5
    maxi(i).d = max(emg(i).data(:,4));
end
for i = f5+1:f6
    maxi(i).d = max(emg(i).data(:,4));
end
for i = f6+1:f7
    maxi(i).d = max(emg(i).data(:,4));
end
for i = f7+1:f8
    maxi(i).d = max(emg(i).data(:,4));
end
for i = f8+1:f9
    maxi(i).d = max(emg(i).data(:,4));
end
for i = f9+1:f10
    maxi(i).d = max(emg(i).data(:,4));
end
for i = f10+1:f11
    maxi(i).d = max(emg(i).data(:,4));
end
for i = f11+1:comp
    maxi(i).d = max(emg(i).data(:,4));
end

maxi_d= zeros(comp,1);

for k = 1:comp

    maxi_d(k,1) = maxi(k).d;
end

%%
% Músculo E %
for i = 1:f1
    maxi(i).e = max(emg(i).data(:,5));
end
for i = f1+1:f2
    maxi(i).e = max(emg(i).data(:,5));
end

for i = f2+1:f3
    maxi(i).e = max(emg(i).data(:,5));
end
```

```
for i = f3+1:f4
    maxi(i).e = max(emg(i).data(:,5));
end
for i = f4+1:f5
    maxi(i).e = max(emg(i).data(:,5));
end
for i = f5+1:f6
    maxi(i).e = max(emg(i).data(:,5));
end
for i = f6+1:f7
    maxi(i).e = max(emg(i).data(:,5));
end
for i = f7+1:f8
    maxi(i).e = max(emg(i).data(:,5));
end
for i = f8+1:f9
    maxi(i).e = max(emg(i).data(:,5));
end
for i = f9+1:f10
    maxi(i).e = max(emg(i).data(:,5));
end
for i = f10+1:f11
    maxi(i).e = max(emg(i).data(:,5));
end
for i = f11+1:comp
    maxi(i).e = max(emg(i).data(:,5));
end

maxi_e = zeros(comp,1);

for k = 1:comp

    maxi_e(k,1) = maxi(k).e;
end

%%
% Músculo F %
for i = 1:f1
    maxi(i).f = max(emg(i).data(:,6));
end
for i = f1+1:f2
    maxi(i).f = max(emg(i).data(:,6));
end
for i = f2+1:f3
    maxi(i).f = max(emg(i).data(:,6));
end
for i = f3+1:f4
    maxi(i).f = max(emg(i).data(:,6));
end
for i = f4+1:f5
    maxi(i).f = max(emg(i).data(:,6));
end
for i = f5+1:f6
    maxi(i).f = max(emg(i).data(:,6));
end
for i = f6+1:f7
    maxi(i).f = max(emg(i).data(:,6));
end
for i = f7+1:f8
    maxi(i).f = max(emg(i).data(:,6));
end
for i = f8+1:f9
    maxi(i).f = max(emg(i).data(:,6));
end
for i = f9+1:f10
    maxi(i).f = max(emg(i).data(:,6));
end
for i = f10+1:f11
    maxi(i).f = max(emg(i).data(:,6));
end
for i = f11+1:comp
    maxi(i).f = max(emg(i).data(:,6));
end

maxi_f = zeros(comp,1);
for k = 1:comp
```

```
    maxi_f(k,1) = maxi(k).f;
end

%%
% Músculo G %
for i = 1:f1
    maxi(i).g = max(emg(i).data(:,7));
end
for i = f1+1:f2
    maxi(i).g = max(emg(i).data(:,7));
end
for i = f2+1:f3
    maxi(i).g = max(emg(i).data(:,7));
end
for i = f3+1:f4
    maxi(i).g = max(emg(i).data(:,7));
end
for i = f4+1:f5
    maxi(i).g = max(emg(i).data(:,7));
end
for i = f5+1:f6
    maxi(i).g = max(emg(i).data(:,7));
end
for i = f6+1:f7
    maxi(i).g = max(emg(i).data(:,7));
end
for i = f7+1:f8
    maxi(i).g = max(emg(i).data(:,7));
end
for i = f8+1:f9
    maxi(i).g = max(emg(i).data(:,7));
end
for i = f9+1:f10
    maxi(i).g = max(emg(i).data(:,7));
end
 for i = f10+1:f11
    maxi(i).g = max(emg(i).data(:,7));
end
for i = f11+1:comp
    maxi(i).g = max(emg(i).data(:,7));
end
maxi_g = zeros(comp,1);

for k = 1:comp

    maxi_g(k,1) = maxi(k).g;
end

%%
% Músculo H %
for i = 1:f1
    maxi(i).h = max(emg(i).data(:,8));
end
for i = f1+1:f2
    maxi(i).h = max(emg(i).data(:,8));
end
for i = f2+1:f3
    maxi(i).h = max(emg(i).data(:,8));
end
for i = f3+1:f4
    maxi(i).h = max(emg(i).data(:,8));
end
for i = f4+1:f5
    maxi(i).h = max(emg(i).data(:,8));
end
for i = f5+1:f6
    maxi(i).h = max(emg(i).data(:,8));
end
for i = f6+1:f7
    maxi(i).h = max(emg(i).data(:,8));
end
for i = f7+1:f8
    maxi(i).h = max(emg(i).data(:,8));
end
for i = f8+1:f9
    maxi(i).h = max(emg(i).data(:,8));
end
```

```
for i = f9+1:f10
    maxi(i).h = max(emg(i).data(:,8));
end
for i = f10+1:f11
    maxi(i).h = max(emg(i).data(:,8));
end
for i = f11+1:comp
    maxi(i).h = max(emg(i).data(:,8));
end

maxi_h= zeros(comp,1);

for k = 1:comp

    maxi_h(k,1) = maxi(k).h;
end

%%
maxi = [maxi_a, maxi_b, maxi_c, maxi_d, maxi_e, maxi_f, maxi_g, maxi_h];
max1 = max(maxi(1:f1,:));
max2 = max(maxi(f1+1:f2,:));
max3 = max(maxi(f2+1:f3,:));
max4 = max(maxi(f3+1:f4,:));
max5 = max(maxi(f4+1:f5,:));
max6 = max(maxi(f5+1:f6,:));
max7 = max(maxi(f6+1:f7,:));
max8 = max(maxi(f7+1:f8,:));
max9 = max(maxi(f8+1:f9,:));
max10 = max(maxi(f9+1:f10,:));
max11 = max(maxi(f10+1:f11,:));
max12 = max(maxi(f11+1:comp,:));

%%
% Sincronizar os dados com a divisão cinemática das fases %
fases = xlsread('Ztotal_fases.xlsx');
fases = fases.*10;

sinc = struct('difs',[],'fasesS',[]);

for ii = 1:comp

    sinc(ii).difs = length(emg(ii).data) - fases(ii,8);
        sinc(ii).fasesS = sinc(ii).difs + fases(ii,2:7);
end

dados = struct('swing',[],'bs',[],'ds',[],'ft',[]);

for ii = 1:comp

    dados(ii).swing(:,1:8) = emg(ii).data(sinc(ii).fasesS(1):...
        sinc(ii).fasesS(6),1:8);
    dados(ii).bs(:,1:8) = emg(ii).data(sinc(ii).fasesS(1):...
        sinc(ii).fasesS(2),1:8);
    dados(ii).ds(:,1:8) = emg(ii).data(sinc(ii).fasesS(2):...
        sinc(ii).fasesS(4),1:8);
    dados(ii).ft(:,1:8) = emg(ii).data(sinc(ii).fasesS(4):...
        sinc(ii).fasesS(6),1:8);
end

%%
% Normalização %

dadosN = struct('swing',[],'bs',[],'ds',[],'ft',[]);
%1%
for ii = 1:f1
    for jj = 1:8
    dadosN(ii).swing(:,jj) = dados(ii).swing(:,jj)./max1(jj);
    dadosN(ii).bs(:,jj) = dados(ii).bs(:,jj)./max1(jj);
    dadosN(ii).ds(:,jj) = dados(ii).ds(:,jj)./max1(jj);
    dadosN(ii).ft(:,jj) = dados(ii).ft(:,jj)./max1(jj);
    end
end

%2%
for ii = f1+1:f2
    for jj = 1:8
    dadosN(ii).swing(:,jj) = dados(ii).swing(:,jj)./max2(jj);
```

```
    dadosN(ii).bs(:,jj) = dados(ii).bs(:,jj)./max2(jj);
    dadosN(ii).ds(:,jj) = dados(ii).ds(:,jj)./max2(jj);
    dadosN(ii).ft(:,jj) = dados(ii).ft(:,jj)./max2(jj);
    end
end

%3%
for ii = f2+1:f3
    for jj = 1:8
    dadosN(ii).swing(:,jj) = dados(ii).swing(:,jj)./max3(jj);
    dadosN(ii).bs(:,jj) = dados(ii).bs(:,jj)./max3(jj);
    dadosN(ii).ds(:,jj) = dados(ii).ds(:,jj)./max3(jj);
    dadosN(ii).ft(:,jj) = dados(ii).ft(:,jj)./max3(jj);
    end
end

%4%
for ii = f3+1:f4
    for jj = 1:8
    dadosN(ii).swing(:,jj) = dados(ii).swing(:,jj)./max4(jj);
    dadosN(ii).bs(:,jj) = dados(ii).bs(:,jj)./max4(jj);
    dadosN(ii).ds(:,jj) = dados(ii).ds(:,jj)./max4(jj);
    dadosN(ii).ft(:,jj) = dados(ii).ft(:,jj)./max4(jj);
    end
end

%5%
for ii = f4+1:f5
    for jj = 1:8
    dadosN(ii).swing(:,jj) = dados(ii).swing(:,jj)./max5(jj);
    dadosN(ii).bs(:,jj) = dados(ii).bs(:,jj)./max5(jj);
    dadosN(ii).ds(:,jj) = dados(ii).ds(:,jj)./max5(jj);
    dadosN(ii).ft(:,jj) = dados(ii).ft(:,jj)./max5(jj);
    end
end

%6%
for ii = f5+1:f6
    for jj = 1:8
    dadosN(ii).swing(:,jj) = dados(ii).swing(:,jj)./max6(jj);
    dadosN(ii).bs(:,jj) = dados(ii).bs(:,jj)./max6(jj);
    dadosN(ii).ds(:,jj) = dados(ii).ds(:,jj)./max6(jj);
    dadosN(ii).ft(:,jj) = dados(ii).ft(:,jj)./max6(jj);
    end
end

%7%
for ii = f6+1:f7
    for jj = 1:8
    dadosN(ii).swing(:,jj) = dados(ii).swing(:,jj)./max7(jj);
    dadosN(ii).bs(:,jj) = dados(ii).bs(:,jj)./max7(jj);
    dadosN(ii).ds(:,jj) = dados(ii).ds(:,jj)./max7(jj);
    dadosN(ii).ft(:,jj) = dados(ii).ft(:,jj)./max7(jj);
    end
end

%8%
for ii = f7+1:f8
    for jj = 1:8
    dadosN(ii).swing(:,jj) = dados(ii).swing(:,jj)./max8(jj);
    dadosN(ii).bs(:,jj) = dados(ii).bs(:,jj)./max8(jj);
    dadosN(ii).ds(:,jj) = dados(ii).ds(:,jj)./max8(jj);
    dadosN(ii).ft(:,jj) = dados(ii).ft(:,jj)./max8(jj);
    end
end

%9%
for ii = f8+1:f9
    for jj = 1:8
    dadosN(ii).swing(:,jj) = dados(ii).swing(:,jj)./max9(jj);
    dadosN(ii).bs(:,jj) = dados(ii).bs(:,jj)./max9(jj);
    dadosN(ii).ds(:,jj) = dados(ii).ds(:,jj)./max9(jj);
    dadosN(ii).ft(:,jj) = dados(ii).ft(:,jj)./max9(jj);
    end
end

%10%
```

```
for ii = f9+1:f10
    for jj = 1:8
    dadosN(ii).swing(:,jj) = dados(ii).swing(:,jj)./max10(jj);
    dadosN(ii).bs(:,jj) = dados(ii).bs(:,jj)./max10(jj);
    dadosN(ii).ds(:,jj) = dados(ii).ds(:,jj)./max10(jj);
    dadosN(ii).ft(:,jj) = dados(ii).ft(:,jj)./max10(jj);
    end
end

%11%
for ii = f10+1:f11
    for jj = 1:8
    dadosN(ii).swing(:,jj) = dados(ii).swing(:,jj)./max11(jj);
    dadosN(ii).bs(:,jj) = dados(ii).bs(:,jj)./max11(jj);
    dadosN(ii).ds(:,jj) = dados(ii).ds(:,jj)./max11(jj);
    dadosN(ii).ft(:,jj) = dados(ii).ft(:,jj)./max11(jj);
    end
end

%12%
for ii = f11+1:comp
    for jj = 1:8
    dadosN(ii).swing(:,jj) = dados(ii).swing(:,jj)./max12(jj);
    dadosN(ii).bs(:,jj) = dados(ii).bs(:,jj)./max12(jj);
    dadosN(ii).ds(:,jj) = dados(ii).ds(:,jj)./max12(jj);
    dadosN(ii).ft(:,jj) = dados(ii).ft(:,jj)./max12(jj);
    end
end
%%
% Arrumar e guarade dados por músculo para poterior RQA %

swing = struct('a',[],'b',[],'c',[],'d', [],'e',[],'f',[],'g',[],'h',[]);

    for ii = 1:comp

    swing(ii).a = dadosN(ii).swing(:,1);
    swing(ii).b = dadosN(ii).swing(:,2);
    swing(ii).c = dadosN(ii).swing(:,3);
    swing(ii).d = dadosN(ii).swing(:,4);
    swing(ii).e = dadosN(ii).swing(:,5);
    swing(ii).f = dadosN(ii).swing(:,6);
    swing(ii).g = dadosN(ii).swing(:,7);
    swing(ii).h = dadosN(ii).swing(:,8);

    end
%%
    bs = struct('a',[],'b',[],'c',[],'d', [],'e',[],'f',[],'g',[],'h',[]);

    for ii = 1:comp

    bs(ii).a = dadosN(ii).bs(:,1);
    bs(ii).b = dadosN(ii).bs(:,2);
    bs(ii).c = dadosN(ii).bs(:,3);
    bs(ii).d = dadosN(ii).bs(:,4);
    bs(ii).e = dadosN(ii).bs(:,5);
    bs(ii).f = dadosN(ii).bs(:,6);
    bs(ii).g = dadosN(ii).bs(:,7);
    bs(ii).h = dadosN(ii).bs(:,8);

    end
%%

ds = struct('a',[],'b',[],'c',[],'d', [],'e',[],'f',[],'g',[],'h',[]);

    for ii = 1:comp

    ds(ii).a = dadosN(ii).ds(:,1);
    ds(ii).b = dadosN(ii).ds(:,2);
    ds(ii).c = dadosN(ii).ds(:,3);
    ds(ii).d = dadosN(ii).ds(:,4);
    ds(ii).e = dadosN(ii).ds(:,5);
    ds(ii).f = dadosN(ii).ds(:,6);
    ds(ii).g = dadosN(ii).ds(:,7);
    ds(ii).h = dadosN(ii).ds(:,8);

    end
%%
```

```
ft = struct('a',[],'b',[],'c',[],'d', [],'e',[],'f',[],'g',[],'h',[]);

    for ii = 1:comp

    ft(ii).a = dadosN(ii).ft(:,1);
    ft(ii).b = dadosN(ii).ft(:,2);
    ft(ii).c = dadosN(ii).ft(:,3);
    ft(ii).d = dadosN(ii).ft(:,4);
    ft(ii).e = dadosN(ii).ft(:,5);
    ft(ii).f = dadosN(ii).ft(:,6);
    ft(ii).g = dadosN(ii).ft(:,7);
    ft(ii).h = dadosN(ii).ft(:,8);

    end
 toc
%%
% Gravar dados para posterior análise %

    save('dados2EMG_RN.mat','swing','bs','ds','ft');

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%            Extracção e Ponderação de features RQA              %
%                                                                 %
% Deve ser aplicada a dados após realizado ajustamento no tempo e%
% isolamento a cada músculo e fase.                               %
%                                                                 %
% Necessita de instalação da Toolbox 5.17 (R28.20; PIK Potsdam)  %
%                                                                 %
%          http://tocsy.pik-potsdam.de/CRPtoolbox/               %
%                                                                 %
%Selecção de features através necessita de instalação de carregar%
%ficheiro load_fspackage.m. Informações em:                      %
%http://featureselection.asu.edu/featureselection_techreport.pdf %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%
tau = input('Delay:  ');
comp = input('Comp:  ');
half = input('half:  ');
label = labelHc;
label(half:end,:)=2;
serie = input('Série Temporal [1]swing [2]BS [3]DS [4]FT:  ');

% feat(:,1) - Recurrence rate.
% feat(:,2) - Determinism.
% feat(:,3) - DIV.
% feat(:,4) - Entropy of diagonal length.
% feat(:,5 )- Laminarity.
% feat(:,6) - Trapping time.
% feat(:,7 )- Ratio DET/RR.
% feat(:,8) - Ratio LAM/DET
%%

if serie == 1

    rqa_sw = struct('ts',[],'fnn',[],'mat',[],'m',[],'emax',[],'e',[]);

for i=1:comp

    rqa_sw(i).fnn = fnn(zscore(SW(i).ts),20,tau);

end

for k=1:comp

    rqa_sw(k).mat = find(rqa_sw(k).fnn == 0);
    rqa_sw(k).m = rqa_sw(k).mat(1);

end

for i=1:comp

    [rqa_sw(i).emax, rqa_sw(i).e] = pss(zscore(SW(i).ts),rqa_sw(i).m,...
        tau);
    rqa_sw(i).e = rqa_sw(i).e.*0.1;
```

```
end

for ii = 1:comp
    ii
    rqa_sw(ii).ts =  crqa(SW(ii).ts,rqa_sw(ii).m,tau,...
       rqa_sw(ii).e ,[],'euclidean','nogui');

end

featT_swing = zeros(comp,13);

for i = 1:comp

    featT_swing(i,:)= rqa_sw(i).ts;

end

feat_swing = [featT_swing(:,1), featT_swing(:,2), 1./featT_swing(:,4),...
    featT_swing(:,5),featT_swing(:,6), featT_swing(:,7),...
    featT_swing(:,2)./featT_swing(:,1),...
    featT_swing(:,6)./featT_swing(:,2)];

F1=(feat_swing(:,1)/max(feat_swing(:,1)));
F2=(feat_swing(:,2)/max(feat_swing(:,2)));
F3=(feat_swing(:,3)/max(feat_swing(:,3)));
F4=(feat_swing(:,4)/max(feat_swing(:,4)));
F5=(feat_swing(:,5)/max(feat_swing(:,5)));
F6=(feat_swing(:,6)/max(feat_swing(:,6)));
F7=(feat_swing(:,7)/max(feat_swing(:,7)));
F8=(feat_swing(:,8)/max(feat_swing(:,8)));

F_SW = [F1 F2 F3 F4 F5 F6 F7 F8];
f = fsFisher(F_SW,label);
cf = fsCFS(F_SW,label);
fc = fsFCBF(F_SW,label);

save('Musculo_SW_ts.mat','rqa_sw','feat_swing','featT_swing','F_SW',...
    'f','cf','fc');

elseif serie == 2

    rqa_bs = struct('ts',[],'fnn',[],'mat',[],'m',[],'emax',[],'e',[]);

for i=1:comp

    rqa_bs(i).fnn = fnn(zscore(BS(i).ts),20,tau);

end

for k=1:comp

    rqa_bs(k).mat = find(rqa_bs(k).fnn == 0);
    rqa_bs(k).m = rqa_bs(k).mat(1);

end

for i=1:comp

    [rqa_bs(i).emax, rqa_bs(i).e] = pss(zscore(BS(i).ts),rqa_bs(i).m,...
        tau);
    rqa_bs(i).e = rqa_bs(i).e.*0.1;

end

for ii = 1:comp
    ii
    rqa_bs(ii).ts =  crqa(BS(ii).ts,rqa_bs(ii).m,tau,...
       rqa_bs(ii).e ,[],'euclidean','nogui');

end

featT_bs = zeros(comp,13);

for i = 1:comp

    featT_bs(i,:)= rqa_bs(i).ts;
```

```
end

feat_bs = [featT_bs(:,1), featT_bs(:,2), 1./featT_bs(:,4),...
    featT_bs(:,5),featT_bs(:,6), featT_bs(:,7),...
    featT_bs(:,2)./featT_bs(:,1),...
    featT_bs(:,6)./featT_bs(:,2)];

F1=(feat_bs(:,1)/max(feat_bs(:,1)));
F2=(feat_bs(:,2)/max(feat_bs(:,2)));
F3=(feat_bs(:,3)/max(feat_bs(:,3)));
F4=(feat_bs(:,4)/max(feat_bs(:,4)));
F5=(feat_bs(:,5)/max(feat_bs(:,5)));
F6=(feat_bs(:,6)/max(feat_bs(:,6)));
F7=(feat_bs(:,7)/max(feat_bs(:,7)));
F8=(feat_bs(:,8)/max(feat_bs(:,8)));

F_BS = [F1 F2 F3 F4 F5 F6 F7 F8];
f = fsFisher(F_BS,label);
cf = fsCFS(F_BS,label);
fc = fsFCBF(F_BS,label);

save('Musculo_BS_ts.mat','rqa_bs','feat_bs','featT_bs','F_BS',...
    'f','cf','fc');

elseif serie == 3

    rqa_ds = struct('ts',[],'fnn',[],'mat',[],'m',[],'emax',[],'e',[]);

for i=1:comp

    rqa_ds(i).fnn = fnn(zscore(DS(i).ts),30,tau);

end

for k=1:comp

    rqa_ds(k).mat = find(rqa_ds(k).fnn == 0);
    rqa_ds(k).m = rqa_ds(k).mat(1);

end

for i=1:comp

    [rqa_ds(i).emax, rqa_ds(i).e] = pss(zscore(DS(i).ts),rqa_ds(i).m,...
        tau);
    rqa_ds(i).e = rqa_ds(i).e.*0.1;

end

for ii = 1:comp
    ii
    rqa_ds(ii).ts =  crqa(DS(ii).ts,rqa_ds(ii).m,tau,...
       rqa_ds(ii).e ,[],'euclidean','nogui');

end

featT_ds = zeros(comp,13);

for i = 1:comp

    featT_ds(i,:)= rqa_ds(i).ts;

end
feat_ds = [featT_ds(:,1), featT_ds(:,2), 1./featT_ds(:,4),...
    featT_ds(:,5),featT_ds(:,6), featT_ds(:,7),...
    featT_ds(:,2)./featT_ds(:,1),...
    featT_ds(:,6)./featT_ds(:,2)];

F1=(feat_ds(:,1)/max(feat_ds(:,1)));
F2=(feat_ds(:,2)/max(feat_ds(:,2)));
F3=(feat_ds(:,3)/max(feat_ds(:,3)));
F4=(feat_ds(:,4)/max(feat_ds(:,4)));
F5=(feat_ds(:,5)/max(feat_ds(:,5)));
F6=(feat_ds(:,6)/max(feat_ds(:,6)));
F7=(feat_ds(:,7)/max(feat_ds(:,7)));
F8=(feat_ds(:,8)/max(feat_ds(:,8)));
```

```
F_DS = [F1 F2 F3 F4 F5 F6 F7 F8];
f = fsFisher(F_DS,label);
cf = fsCFS(F_DS,label);
fc = fsFCBF(F_DS,label);

save('Musculo_DS_ts.mat','rqa_ds','feat_ds','featT_ds','F_DS',...
    'f','cf','fc');

elseif serie == 4

    rqa_FT = struct('ts',[],'fnn',[],'mat',[],'m',[],'emax',[],'e',[]);

for i=1:comp

    rqa_ft(i).fnn = fnn(zscore(FT(i).ts),20,tau);

end

for k=1:comp

    rqa_ft(k).mat = find(rqa_ft(k).fnn == 0);
    rqa_ft(k).m = rqa_ft(k).mat(1);

end

for i=1:comp

    [rqa_ft(i).emax, rqa_ft(i).e] = pss(zscore(FT(i).ts),rqa_ft(i).m,...
        tau);
    rqa_ft(i).e = rqa_ft(i).e.*0.1;

end

for ii = 1:comp
    ii
    rqa_ft(ii).ts =  crqa(FT(ii).ts,rqa_ft(ii).m,tau,...
       rqa_ft(ii).e ,[],'euclidean','nogui');

end

featT_ft = zeros(comp,13);

for i = 1:comp

    featT_ft(i,:)= rqa_ft(i).ts;

end

feat_ft = [featT_ft(:,1), featT_ft(:,2), 1./featT_ft(:,4),...
    featT_ft(:,5),featT_ft(:,6), featT_ft(:,7),...
    featT_ft(:,2)./featT_ft(:,1),...
    featT_ft(:,6)./featT_ft(:,2)];

F1=(feat_ft(:,1)/max(feat_ft(:,1)));
F2=(feat_ft(:,2)/max(feat_ft(:,2)));
F3=(feat_ft(:,3)/max(feat_ft(:,3)));
F4=(feat_ft(:,4)/max(feat_ft(:,4)));
F5=(feat_ft(:,5)/max(feat_ft(:,5)));
F6=(feat_ft(:,6)/max(feat_ft(:,6)));
F7=(feat_ft(:,7)/max(feat_ft(:,7)));
F8=(feat_ft(:,8)/max(feat_ft(:,8)));

F_FT = [F1 F2 F3 F4 F5 F6 F7 F8];
f = fsFisher(F_FT,label);
cf = fsCFS(F_FT,label);
fc = fsFCBF(F_FT,label);

save('Musculo_FT_ts.mat','rqa_ft','feat_ft','featT_ft','F_FT',...
    'f','cf','fc');
end
```

Faculty of Human Kinetics
TECHNICAL UNIVERSITY OF LISBON

ETHICS COUNCIL

[illegible]
Prof. Doutor Pedro Teixeira (Presidente)
[illegible]
[illegible]
[illegible]
Prof. Doutor José Alves Diniz
[illegible]
[illegible]
[illegible]
[illegible]

To: Whom It May Concern

Date: November 27, 2012

Research Project: *Neuromuscular Activity in the Golf Swing*

CEFMH Approval Number: 7/2012
Recommendations: None

This Council has reviewed the project indicated above and declares that it is in accordance with Portuguese and international guidelines for scientific research involving human beings, including the 2008 Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects, and the 1997 Convention on Human Rights and Biomedicine (the "Oviedo Convention").

The President of the Ethics Council

Pedro J. Teixeira, Ph.D.

[illegible]