

Accuracy of Pattern Detection Methods in the Performance of Golf Putting

Micael S. Couceiro, Gonçalo Dias, Rui Mendes & Duarte Araújo

To cite this article: Micael S. Couceiro, Gonçalo Dias, Rui Mendes & Duarte Araújo (2013) Accuracy of Pattern Detection Methods in the Performance of Golf Putting, Journal of Motor Behavior, 45:1, 37-53, DOI: [10.1080/00222895.2012.740100](https://doi.org/10.1080/00222895.2012.740100)

To link to this article: <https://doi.org/10.1080/00222895.2012.740100>



Published online: 13 Feb 2013.



Submit your article to this journal [↗](#)



Article views: 861



View related articles [↗](#)



Citing articles: 4 View citing articles [↗](#)

RESEARCH ARTICLE

Accuracy of Pattern Detection Methods in the Performance of Golf Putting

Micael S. Couceiro¹, Gonçalo Dias², Rui Mendes³, Duarte Araújo^{4,5}

¹RoboCorp, Department of Electrical Engineering, Engineering Institute of Coimbra (ISEC), Portugal. ²RoboCorp, Faculty of Sport Sciences and Physical Education (CIDAF), University of Coimbra, Portugal. ³Coimbra College of Education and Interdisciplinary Centre for the Study of Human Performance (CIPER), Technical University of Lisbon, Portugal. ⁴Faculty of Human Kinetics, Technical University of Lisbon, Portugal. ⁵Interdisciplinary Centre for the Study of Human Performance (CIPER), Technical University of Lisbon, Portugal.

ABSTRACT. The authors present a comparison of the classification accuracy of 5 pattern detection methods in the performance of golf putting. The detection of the position of the golf club was performed using a computer vision technique followed by the estimation algorithm Darwinian particle swarm optimization to obtain a kinematical model of each trial. The estimated parameters of the models were subsequently used as sample of five classification algorithms: (a) linear discriminant analysis, (b) quadratic discriminant analysis, (c) naive Bayes with normal distribution, (d) naive Bayes with kernel smoothing density estimate, and (e) least squares support vector machines. Beyond testing the performance of each classification method, it was also possible to identify a putting signature that characterized each golf player. It may be concluded that these methods can be applied to the study of coordination and motor control on the putting performance, allowing for the analysis of the intra- and interpersonal variability of motor behavior in performance contexts.

Keywords: classification, detection, golf putting, individualized kinematic pattern

Golf has become an increasingly popular sport all over the world. However, there is a gap in the literature on further studying the motor skills of golfers, especially involving the performance of the golf putting (Pelz, 1989, 2000). For instance, what are the coordination and control requirements to perform golf putting? Importantly, the putt represents about 43% of the strokes in a golf game (Alexander & Kern, 2005). The skill to perform a putt is severely constrained by each person profile and characteristics (Jonassen & Grabowski, 1993). Also, similar to other motor skills, individual performance on the putt results in a unique pattern (i.e., signature) that characterizes each player (e.g., Araújo, Davids, & Hristovski, 2006; Davids, Button, & Bennett, 2008; Hristovski, Davids, Araújo, & Passos, 2011; Phillips, Davids, Renshaw, & Portus, 2010; Schöllhorn, Mayer-Kress, Newell, & Michelbrink, 2008).

In order to better understand the concepts of signature or fingerprint of players in putting performance and the differences within and between individuals, it is important to clarify the notion of variability. Variability, inherent to the human movement systems, emerges as an essential feature of biological organisms (Davids, Bennett, & Newell, 2006; Newell & Corcos, 1993). Sometimes random variability is confounded with functional variability (Newell, Deutsch, Sosnoff, & Mayer-Kress, 2006). This confusion is a problem, because it dismisses as insignificant this feature of motor behavior, and it treats variability operationally as a standard

deviation in distributional statistics (Newell & Corcos, 1993). However, variability is not only the magnitude of variations, it has structure. The traditional approach to motor variability as noise has been that of additive white Gaussian noise rather than that of multiple timescales of variability and their multiplicative time-dependent influences (Newell et al., 2006). Random variability or white noise is a very small background component in the structure of variability that one may find in biological systems (Newell et al., 2006). The inherent noisiness of the motor system results in variability being omnipresent and unavoidable in all high human performance. Rarely will a person execute exactly the same strategy to perform a given action, such as golf putting (e.g., Button, McLeod, Sanders, & Coleman, 2003). Although some regularities may occur, factors such as the exact path of the rotational movement of torso and limbs always vary across responses. There are several reasons this variability occurs (Riley, Kuznetsov, & Bonnette, 2011). One of such reasons is that it seems to be the case that bodies are naturally noisy, such that there are irreducible sources of movement variability originating, perhaps, in uncontrollable, noisy neural firing that in turn originates from even lower level sources. As a consequence of neural noise, muscular forces required to generate movement are also constantly changing, introducing variability at the kinematic level (Riley et al., 2011). Second, human movement is vastly underconstrained (Newell, Slobounov, Slobounova, & Molenaar, 1997). This is so because there is a high dimensionality of the body, consequently there are many solutions that permit one to satisfy the basic biomechanical requirement of golf putting. A person might exploit any solution for one (back)swing and another for the next (down)swing, or even change solutions midstream during a single downswing, and different people might tend to exploit different solutions for reasons related to their own biomechanics or idiosyncratic histories. Third, neuromotor control is degenerate (Davids, Araújo, Button, & Renshaw, 2007; Edelman & Gally, 2001; Mayer-Kress, Liu, & Newell, 2006), where structurally different components can produce the same functional outcome, such as when the same motor outcome is achieved through different combinations of muscle activations or kinematics. Finally, the conditions under which golf putting is controlled

Correspondence address: Micael S. Couceiro, Department of Electrical Engineering, Rua Pedro Nunes, Coimbra 3030, Portugal.
e-mail: micael@isec.pt

are typically variable. This means that sometimes different solutions are required to match the extant conditions, so golf putting must be highly flexible to accommodate the many possible scenarios the body might face. Such flexibility might naturally be accompanied by high variability, because a system that does not vary may not be able to flexibly adapt to changing circumstances. The many implications of variability for understanding the control of movement have been widely recognized (e.g., Davids et al., 2006; Harbourne & Stergiou, 2009; Latash, Scholz, & Schoner, 2002; Newell & Corcos, 1993; Newell & Slifkin, 1998; Riley & Turvey, 2002). Importantly, the revelation of the structure of variability in motor behavior, with all the technological advances that it requires, has recently shifted the attention of the research spotlight away from examining the behavioral tendencies of groups of individuals and toward the analysis of individual performance (Button, Davids, & Schollhorn, 2006).

The last decades have seen technological advances and new analysis tools that allowed for comprehensive changes in the way patterns of human movement are analyzed and interpreted. As patterns in movement are being analyzed in more depth, researchers are identifying subtle individualities, or signature patterns of movement, in even very simple tasks (Beek, Rikkert, & van Wieringen, 1996; Button, Bennett, & Davids, 1998; Port, Lee, Dassoenville, & Georgopoulos, 1997). In this sense, Kelso (1995) argued that pooled group data have limited value, in that the instantiation of laws and principles of learning are specific to the individual.

In this article we propose that individual analysis of kinematics are preferable to the traditional pooling of group data if an individual is interested in unraveling and specifically enhancing a given motor performance, in this case golf putting. On this basis, the aim of this study was to compare classification methods, which are useful to classify the athletes' performance and their intra- and interindividual signatures in golf putting.

This emphasis on performance analyses has consequences for practice, because reference to enhancement processes in the motor behavior literature may be associated with individual movement outcomes and not with an idealized, hypothetical optimal pattern of movement that always achieves the same outcome. The consequence for practical approaches requires the implementation of a variety of techniques and exercises that contribute to the self-organization of the motor system in training or competition context (Davids et al., 2008; Schöllhorn, 2000; Schöllhorn, et al., 2008). These practical approaches use variability as a mechanism for forcing individual performers to adapt their movement to interacting (individual, task, and environment) constraints unique to each situation. This is particularly apparent within sport, where such factors change frequently and unexpectedly (Button et al., 2006). Therefore, the methods presented in this study to analyze the golf putting cannot be regarded as closed solutions to describe the performance of athletes in a standardized manner. Hence, by studying a representative task of golf performance, such as golf putting, we also take into

account other variables that may influence players' performance, such as the effect of fatigue that is likely to alter the pattern of movement (i.e., particularly when the athlete carried out the putting gesture in different practice conditions and facing different environmental constraints; Newell, 1986).

That being said, we argue that the coordination and control of movements are dynamic processes that change from athlete to athlete (i.e., as a signature or fingerprint of the athlete). Therefore, this study may contribute to characterize the intraindividual differences (i.e., athlete's own behavior) and interindividual (i.e., between athletes) that result from putting performance. It is also noteworthy that this work presents new solutions for the analysis of the motor behavior of athletes that goes beyond the distributional statistical results (e.g., mean, standard deviation, coefficient of variation). Several researchers (e.g., Coello, Delay, Nougier, & Orliaguet, 2000; Delay, Nougier, Orliaguet, & Coello, 1997; Karlsen, Smith, & Nilsson, 2008) that studied putting based their analysis on process-tracing measurements such as position, velocity, and acceleration (Figure 1) of limbs and body segments during movement. Usually, these studies search for expert-novice differences in the variability (or stability) of the execution of this movement. However, only a few studies analyzed process variables such as the position, the velocity or the acceleration (linear or angular) of the golf club, more specifically the putter, during putt execution (e.g., Hume, Keogh, & Reid, 2005; Nesbit, Hartzell, Nalevanko, & Starr, 1996) as we aimed to do in this study. Also, to our knowledge, automatic tracking of the putter's trajectory on the green or in laboratory is not common. Most works (Gehrig, Lepetit, & Fua, 2003; Urtasun, Fleet, & Fua, 2005) are based on manual tracking where the user needs to follow the putter frame by frame thus constructing the full motion path of the putting by hand. This method can be tedious and impracticable for a large number of samples.

In line with this, the primary purpose of the present work was to analyze the golf putting performed by expert players, using an algorithm that automatically detects stationary and dynamic objects, such as the putter. This algorithm is used to obtain a point cloud (i.e., set of points of the Cartesian coordinates of the putter's head) during the execution. Consequently, many process variables, such as the trajectory function, can be obtained by estimation of a sinusoidal model to fit the cloud points, using the Darwinian particle swarm optimization (DPSO) method. The DPSO method was previously compared with others, such as the gradient descent and the downhill simplex, and experimental results confirmed the superior performance of the DPSO with a smaller mean squared error (*MSE*; Couceiro, Luz, Figueiredo, Ferreira, & Dias, 2010; Dias et al., 2009). Subsequently, we could perform our main contribution: to compare the classification accuracy of five pattern estimation methods in the performance of golf putting. To that end, the parameters of the kinematic model (i.e., amplitudes, frequencies and phases of the sinusoids) of each expert player's putting trial may be classified in

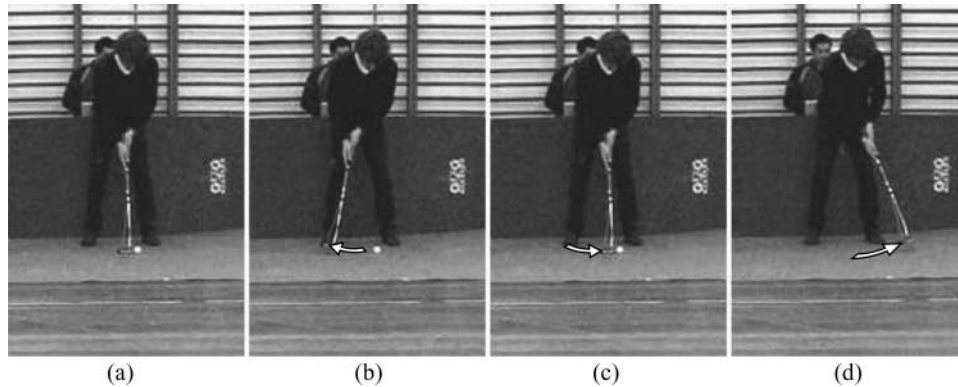


FIGURE 1. Putter movement action parameter analysis: (a) initial stage, (b) backswing, (c) downswing and ball impact, (d) follow-through.

order to identify possible links between different executions of the same player, thus extracting putting signatures for every player. However, little is known about what are the better classification techniques for these measurements. Therefore, we compared five different classification techniques: (a) linear discriminant analysis (LDA), (b) quadratic discriminant analysis (QDA), (c) naive Bayes with normal (Gaussian) distribution (NV), (d) naive Bayes with kernel smoothing density estimate (NVK), and (e) least squares support vector machines with radial basis function kernel (LS-SVM).

The LDA and QDA are the most basic parametric methods and will be used as a reference because they are presented in all books about pattern recognition (Bishop, 2006). The naive Bayes technique is a simple probabilistic classifier based on the Bayesian theorem and is particularly suited when the dimensionality of the inputs is high. Besides outperforming more sophisticated and complex classification methods, the NV can also be improved using a kernel smoothing density estimation. This method has been largely used in text classification but, more recently, it was used to classify human activities (Altun, Barshan, & Tunçel, 2010). Finally, the SVM method has been successfully applied to a wide range of pattern recognition and classification problems in human movement (Janssen et al., 2011). It normally provides better prediction on unseen test data and a unique optimal solution for training problems while containing fewer parameters when compared with other methods.

Similarly, other classification methods have been used in human movement analysis, especially in the identification of individual patterns (Schöllhorn, 2000; Schöllhorn et al., 2008; Schöllhorn, Nigg, Stefanyshyn, & Liu, 2002). In this context, Janssen et al. (2011) analyzed the properties and characteristics of human movement encompassing several motor skills (e.g., walking and running). To this end, according to these authors, using pattern recognition techniques such as SVM and self-organizing maps (SOM) and classification methods based on neural networks (Bishop, 2006;

Rennie, 2001) allows the quantification of the fatigue level of a given muscle group, the recognition of walking patterns, and the determination of the energy consumption that involves physical exercise in the context of training or competition.

Before introducing the previously mentioned five classification methods, in the next section we present a description and discussion of the detection and estimation algorithms used in this study.

Method

Detection

Participants were six expert golf players, with a handicap between five and 15, more than 10 years of golf practice, and inclusion in the national championship of golf. The players were 32 ± 10 years old, volunteer, male, and right-handed. The adopted task under examination was the golf putting, meaning the strike of a ball (Titleist; model Pro V1, Fairhaven, MA) with a putter (Putter Jumbo Black Beauty; size 35; standard, United Kingdom) on a horizontal and still surface, placed on the ground over a ramp. The players performed 30 trials at 2 m from the hole.

The following detection algorithm was used in order to detect the putter's head, signaled by a red marker, accordingly with the red-green-blue (RGB) range values defined (Figure 2b):

1. Select the region of interest (Figure 2b, around the place where is the red marker), within the entire frame (Figure 2a), to be analyzed.
2. Analyze the region of interest of the current frame, searching for pixels with the RGB values within the defined RGB range (RGB ranges depicted in Figure 2).
3. If a pixel under condition 2 is found, then verify if it is inside a blob with at least eight pixels within the same RGB range.

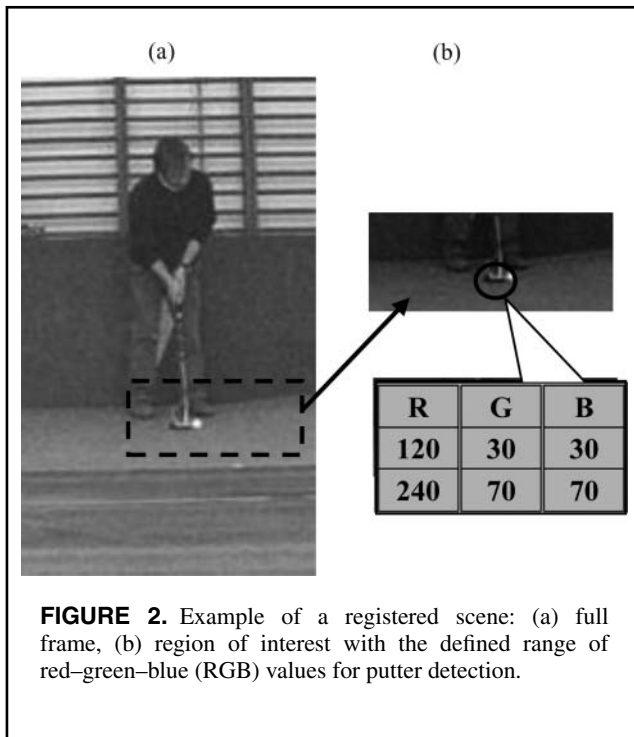


FIGURE 2. Example of a registered scene: (a) full frame, (b) region of interest with the defined range of red–green–blue (RGB) values for putter detection.

The analysis of the video for each trial resulted in a vector that included the object position in the corresponding frame (pixel/frame; i.e., trajectory). Optionally, we could do the following:

4. Convert the pixel/frame value of the object in metric units (e.g., m/s).

This algorithm was computationally efficient because it relied in simple image processing techniques and ensured satisfactory results on both true positives (TP) detection and reduced processing time.

The digital camera used was a Casio Exilim/High Speed EX-FH25 with 210 Hz, at a resolution of 480×360 pixels with a focal length of lens of 26 mm, and provided a considerable depth of field. Therefore, a reference in the same plane of the analyzed gesture was necessary to perform the conversion into $m \cdot s^{-1}$. This reference was the putt’s metallic part with a length of 585 mm.

The graph presented in Figure 3 shows an example of the detected position of a golf club, in the horizontal plane, during the putting execution of an expert player.

In Figure 3 it can be seen that the detection algorithm’s output presented some lacking data. This happened when the conditions of the second step of the algorithm were not met. In such cases, the detection was skipped during the corresponding time instant to avoid the inclusion of errors. In order to classify the point cloud, linear and nonlinear estimation techniques were compared to fit the acquired points of the cloud to a sinusoidal function, thus obtaining a mathematical model to describe the putter’s position during the execution of the play (cf. Couceiro, Luz et al., 2010; Dias

et al., 2009). In previous studies (Couceiro, Luz et al., 2010; Dias et al., 2009), the DPSO method presented a superior performance when compared to other classical techniques, thus resulting in a smaller *MSE* when describing the putter’s head horizontal trajectory. Hence, next section briefly summarizes the DPSO estimation technique.

Estimation

Particle swarm optimization (PSO; Carlisle & Dozier, 2001; Kennedy & Eberhart, 1995) is a well-known and consolidated population of particles (i.e., local solutions) based on a stochastic optimization technique. In PSO, the candidate solutions are called particles. These particles travel through the search space to find an optimal solution, by interacting and sharing information with neighbor particles, namely their individual best solution (local best) and computing the neighborhood best. Also, in each step of the procedure, the global best solution obtained in the entire swarm is updated. Using all of this information, the particles realize the locations of the search space where success was obtained, and are guided by these successes.

In each step of the algorithm, a fitness function is used to evaluate the particle success. To model the swarm, each particle n moves in a multidimensional space according to position (x_n) and velocity (v_n) values that are highly dependent on the local best (\tilde{x}_n), neighborhood best (\tilde{n}_n), and global best (\tilde{g}_n) information:

$$v_n = wv_n + \rho_1 r_1 (\tilde{g}_n - x_n) + \rho_2 r_2 (\tilde{x}_n - x_n) + \rho_3 r_3 (\tilde{n}_n - x_n). \quad (1)$$

$$x_n = x_n + v_n. \quad (2)$$

The coefficients w , ρ_1 , ρ_2 , and ρ_3 assign weights to the inertial influence, the global best, the local best, and the neighborhood best when determining the new velocity, respectively. Typically, the inertial influence is set to a value slightly lower than 1. ρ_1 , ρ_2 and ρ_3 are constant integer values, which represent the local solution of a particle and the global solution of the population. However, different results can be obtained by assigning different influences for each component. For example, several works do not consider the neighborhood best and ρ_3 is set to zero (Carlisle & Dozier, 2001). Depending on the application and the characteristics of the problem, tuning these parameters properly will lead to better results. The parameters r_1 , r_2 and r_3 are random vectors, with each component generally a uniform random number between 0 and 1. They are used for multiplying a new random component per velocity dimension, rather than multiplying the same component with each particle’s velocity dimension revealing a stochastic explorative effect that enhances the ability to avoid local solutions.

In the beginning, the particles’ velocities are set to zero and their position is randomly set within the boundaries of

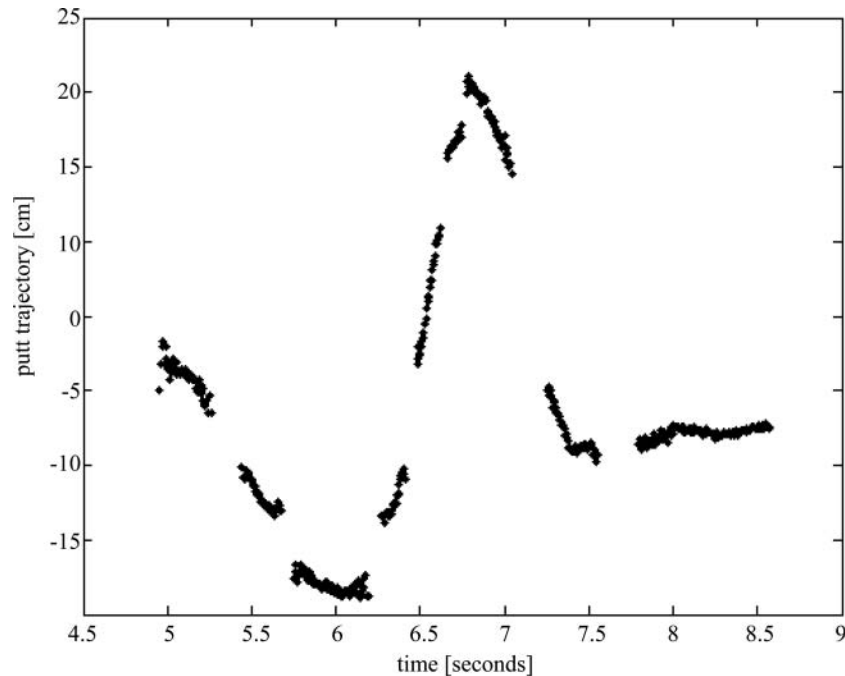


FIGURE 3. Example of a point cloud obtained with the detection algorithm.

the search space. The local, neighborhood and global bests are initialized with the worst possible values, taking in consideration the problem nature. For instance, if dealing with a minimization problem, the local, neighborhood, and global bests will be initialized with a value that tends to positive infinity.

There are other few parameters that need to be adjusted: (a) population size, to optimize in order to get overall good solutions (i.e., optimal or near-optimal solution) in acceptable time (i.e., faster than optimal methods); and (b) stopping criteria, which could be a predefined number of iterations, after which results are no better.

PSO reveals an effect of implicit communication between particles (similar to broadcasting) by updating neighborhood and global information, which affects the velocity and consequent position of particles. However, in order to improve the efficiency of the algorithm, many variations of the PSO algorithm have been presented (Bonilla-Petriciolet & Segovia-Hernández, 2010) since its first description (Kennedy & Eberhart, 1995). In this work, we focused on an algorithm that incorporates techniques of evolutionary approaches to the PSO: The Darwinian Particle Swarm Optimization (Tillett, Rao, Sahin, & Rao, 2005).

Despite the similarities between the PSO and genetic algorithms (GAs; e.g., randomly generated population, fitness function evaluation, population update, search for optimality with random techniques, not guaranteeing success), PSO does not use genetic operators such as crossover and mutation. Thus, it is not considered an evolutionary technique.

On the other hand, the DPSO extends the PSO, in that it determines if natural selection (i.e., Darwinian principle of survival of the fittest) can enhance the ability of the PSO algorithm to escape from local optima. This is made by running many simultaneous parallel PSO algorithms, each one a different swarm, on the same test problem, and a simple selection mechanism is applied. When a search tends to a local optimum, the search in that area is simply discarded and another area is searched instead.

At each step (see Figure 4), swarms that get better are rewarded (by extending the particle's life or by spawning a

Main Program Loop

For each swarm in the collection
Evolve the swarm (Evolve Swarm Algorithm: right)
Allow the swarm to spawn
Delete "failed" swarms

Evolve Swarm Algorithm

For each particle in the swarm
Update Particles' Fitness
Update Particles' Best
Move Particle
If swarm gets better
Reward swarm: spawn particle: Extend swarm life
If swarm has not improved
Punish swarm: possibly delete particle: reduce swarm life

FIGURE 4. Darwinian particle swarm optimization algorithm.

new descendent) and swarms that stagnate are punished (by reducing swarm's life or by deleting particles). To examine the general state of each swarm, the fitness of all particles was evaluated and the neighborhood and individual best positions of each of the particles were updated. If a new global solution was found, a new particle was spawned. A particle was deleted if the swarm failed to find a fitter state in a defined number of steps.

Some simple rules were followed to delete a swarm, delete particles, and spawn a new swarm and a new particle. First, when the swarm population falls below a minimum bound, the swarm is deleted. Second, The worst performing particle in the swarm is deleted when a maximum threshold number of steps (search counter) without improving the fitness function are reached. After the deletion of the particle, instead of being set to zero, the counter is reset to a value, SC_C approaching the threshold number, SC_C^{\max} , according to

$$SC_C(N_{kill}) = SC_C^{\max} \cdot 1 - \left[\frac{1}{N_{kill} + 1} \right]. \quad (3)$$

with N_{kill} being the number of particles deleted from the swarm over a period in which there was no improvement in fitness. Third, to spawn a new swarm, a swarm must not have any particle ever deleted and the maximum number of swarms must not be exceeded. Still, the new swarm is only created with a probability of $p = f/NS$, with f a random number in $[0,1]$ and NS the number of swarms. This factor avoided the creation of newer swarms when there were large numbers of swarms in existence. The parent swarm was unaffected and half of the parent's particles were selected at random for the child swarm and half of the particles of a random member of the swarm collection were also selected. If the swarm initial population number was not obtained, the rest of the particles were randomly initialized and added to the new swarm. Fourth, a particle is spawned whenever a swarm achieves a new global best and the maximum defined population of a swarm has not been reached.

Similar to the PSO, a few parameters also needed to be adjusted to run the algorithm efficiently: (a) social and cognitive coefficients of, v_n , (b) initial swarm population, (c) maximum and minimum swarm population, (d) initial number of swarms, (e) maximum and minimum number of swarms, (f) stagnancy threshold, and (g) maximum and minimum velocity value.

These parameters are further explained in Tillet et al. (2005), and because they are not needed to understand the main aims of this work, they are not described here.

By analysis of the shape of various point clouds given by the detection algorithm, it was clear that to model the horizontal putter's trajectory in time, a sinusoidal function ought to be used because it can be described by a pendulum-like movement. However, a function composed by only one sinusoid was not precise enough to describe the movement, as it is clear in f_1 of Figure 5, which results, in this case, in a *MSE* of 2.6568 units. This happened because the amplitude, angular frequency, and phase of the descending half-wave, which corresponded to the player's backswing and downswing, was usually different than the ascending half-wave, which corresponded to the ball's impact and follow-through. These disparities could not be represented using solely one sinusoid wave. Therefore, to obtain a more precise model a sum of sinusoid waves were employed. However, a compromise between precision and complexity of the problem was assumed, because each sinusoid adds three more dimensions to the estimation problem (amplitude, angular frequency, and phase of the corresponding sine wave). In order not to let the complexity of the problem grow inappropriately, a function composed by the sum of three sinusoids was used (f_3 of Figure 5), due to its precision, with a *MSE* of 0.6926, when compared to using solely a sum of two sinusoids, with a *MSE* of 0.7124 (f_2 of Figure 5).

Consequently, having the estimation function defined as a sum of three sine waves, each of the three parameters of each wave needed to be estimated, resulting in a nine-dimension estimation problem, already performed in previous studies

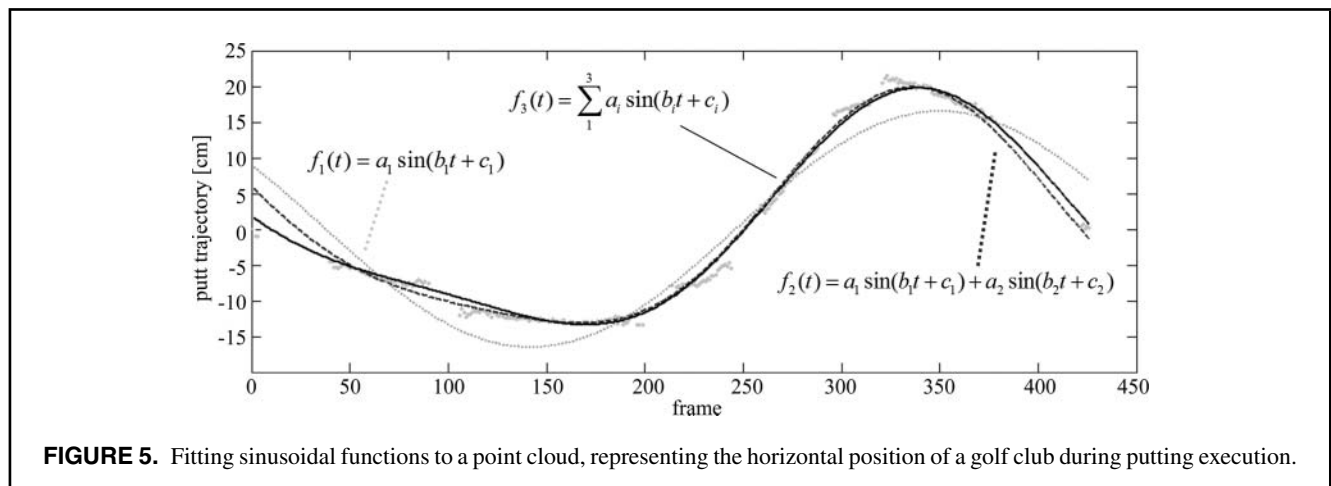


FIGURE 5. Fitting sinusoidal functions to a point cloud, representing the horizontal position of a golf club during putting execution.

(Couceiro, Dias, Luz, Figueiredo, & Ferreira, 2010; Couceiro et al., 2012). These studies attempted to minimize the mean squared estimation error for every experiment, in order to obtain a precise function that described the trajectory of the putter's head during putting execution.

$$f(t) = a_1 \sin(b_1 t + c_1) + a_2 \sin(b_2 t + c_2) + a_3 \sin(b_3 t + c_3). \quad (4)$$

Hence, Equation 4 allows a mathematical description the kinematic model of a putting execution. Although the parameters of the kinematic model (i.e., amplitudes, frequencies, and phases of the sinusoids) may exhibit different values for the same player in different executions, it may be possible to identify a signature inherent to the player. Therefore, five classification algorithms were reviewed and tested in order to identify a putting signature for each of the six players based on the parameters of the kinematic model. For performing this, we analyzed the parameters that represented key points in their respective point cloud for each trial.

Classification

In this section we present a review of the five classification algorithms used in this work. In particular, the LDA and QDA, the NV and NVK, and the LS-SVM.

The confusion matrix and the area under the receiver-operating characteristic curve (AUC) are some of the most well-known methods to evaluate the performance of the classification algorithms. This is why we used them in this work.

A confusion matrix is a matrix containing information about actual and predicted classifications done by a classification system (Kaladhar, Nageswara Rao, & Ramesh Naidu Rajana, 2010). The confusion matrix lists errors and successes in the test set. The main diagonal represents the correctly classified samples while the other elements of the matrix correspond to samples that were incorrectly classified.

The receiver-operating characteristic (ROC) is a technique used to visualize, evaluate, organize, and select classifiers based on their performance (Figure 6). The ROC graphs can show the line between the true- and the false-positive rate of the classifiers (Fawcett, 2006; for further details about both methods, see Chan & Lee, 2002; Freitas, Carvalho, Oliveira, Aires, & Sabourin, 2007).

To compare classifiers it is necessary to reduce the curve to a scalar value. A common method for this is to calculate the AUC. The AUC is a way to measure classifiers performance. Because the AUC is a part of the area of the unit square, its value always varies between 0 and 1. An AUC value of 1 represents a perfect test while the AUC value of 0.5 represents a weak or worthless test.

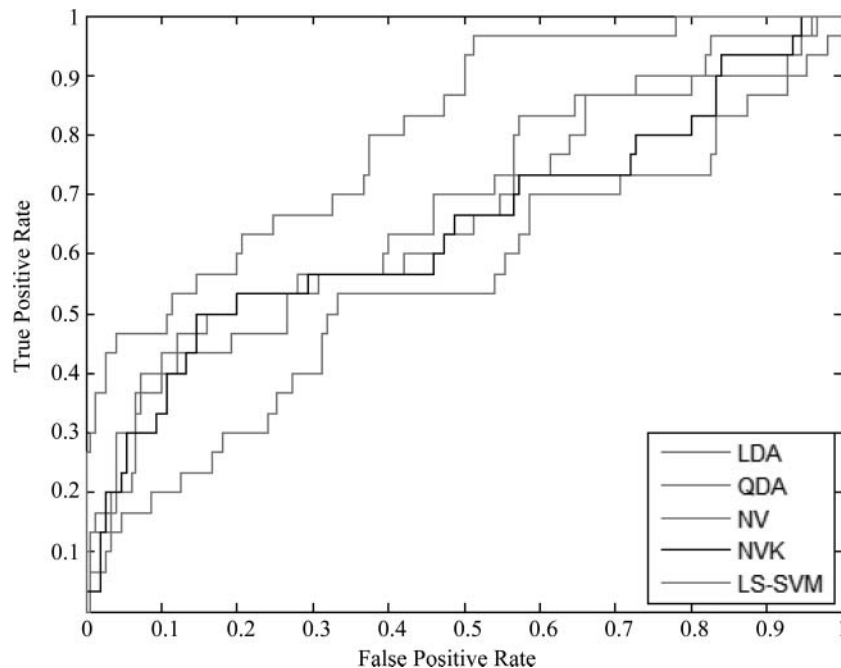


FIGURE 6. Example of receiver-operating characteristic (ROC) curves for the third sine wave (amplitude $[a_3]$ vs. angular frequency $[b_3]$) of the fifth player (class 5). The best possible prediction method would yield a point in the upper left corner (i.e., coordinate $[0,1]$, of the ROC space, hence the least squares support vector machines (LS-SVM) is the classifier that presents a superior performance for this trial. LDA = linear discriminant analysis; QDA = quadratic discriminant analysis; NV = naive Bayes with normal distribution; NVK = naive Bayes with kernel smoothing density estimate.

LDA

The LDA is one of the most commonly used techniques for data classification and dimensionality reduction in statistics, pattern recognition and machine learning (Bishop, 2006). This technique easily handles the case where the within-class frequencies are unequal and its performances have been examined on randomly generated test data (Martínez & Kak, 2001). LDA maximizes the ratio of between-class variance to the within-class variance in any particular data set thereby guaranteeing maximal separability.

LDA is closely related to logistic regression, which also attempt to express one dependent variable as a combination of other features or measurements (Bishop, 2006). LDA looks for linear combinations of variables that best explain the data, by explicitly attempting to model the difference between the classes of data. Other methods are preferable in applications where it is not reasonable to assume that the independent variables are normally distributed, which is a fundamental assumption of the LDA method. Hence, a difference between independent variables and dependent variables must be made.

LDA assumes that the conditional probability density functions (PDF) follow a normal distribution for all classes. In practice, the class means and covariances are usually not known and are estimated from the training set used (e.g., using the maximum likelihood estimate or the maximum a posteriori estimate). Also, LDA assumes that all classes have a common covariance matrix and all covariances have full rank, which is called the homoscedastic assumption. The discriminant function is given by testing the probability that a sample x is contained in one class or another:

$$\delta_k(x) = x^T \sum^{-1} \mu_k - \frac{1}{2} \mu_k^T \sum^{-1} \mu_k + \log(\pi_k). \quad (5)$$

wherein Σ is the covariance matrix common to all classes, μ_k and π_k is the mean and the prior probability of class k , respectively. The value of x in each $\delta_k(x)$ is calculated and the classification of x is the class k that yields the largest value.

The resulting classifier implies that the decision boundary between pairs of classes is linear and a hyperplane when using more than two classes. This is verified through the comparison of classes using the log ratio.

In geometrical terms, it is clear that an input observation, in a given class if the multidimensional space observation point is located on a certain side of a hyperplane, perpendicular to the normal to the discriminant hyperplane (Figure 7).

LDA finds other applications in areas such as face recognition, marketing or financial prediction (for more details on the implementation of the method, see Hastie, Tibshirani, & Friedman, 2009; Lee, Roan, Smith, & Lockhart, 2009).

QDA

Although they differ in their derivation, QDA is similar to LDA (Roushanzamir, Valafar, & Valafar, 1999). The essential difference between them is in the way the linear function

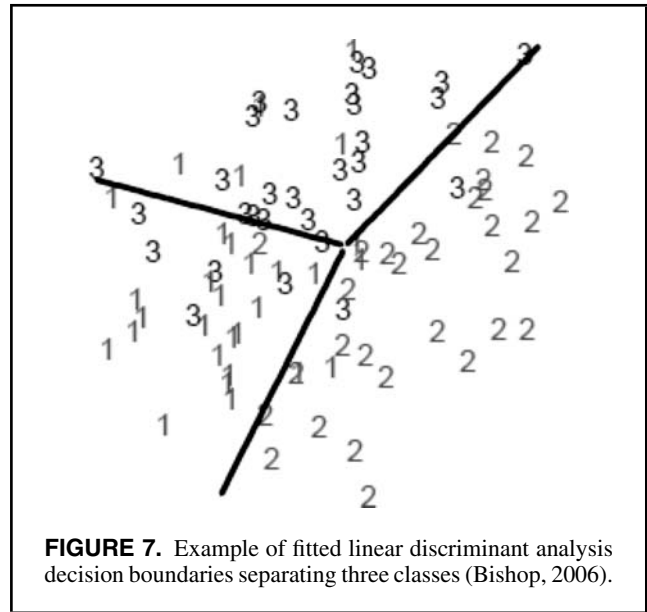


FIGURE 7. Example of fitted linear discriminant analysis decision boundaries separating three classes (Bishop, 2006).

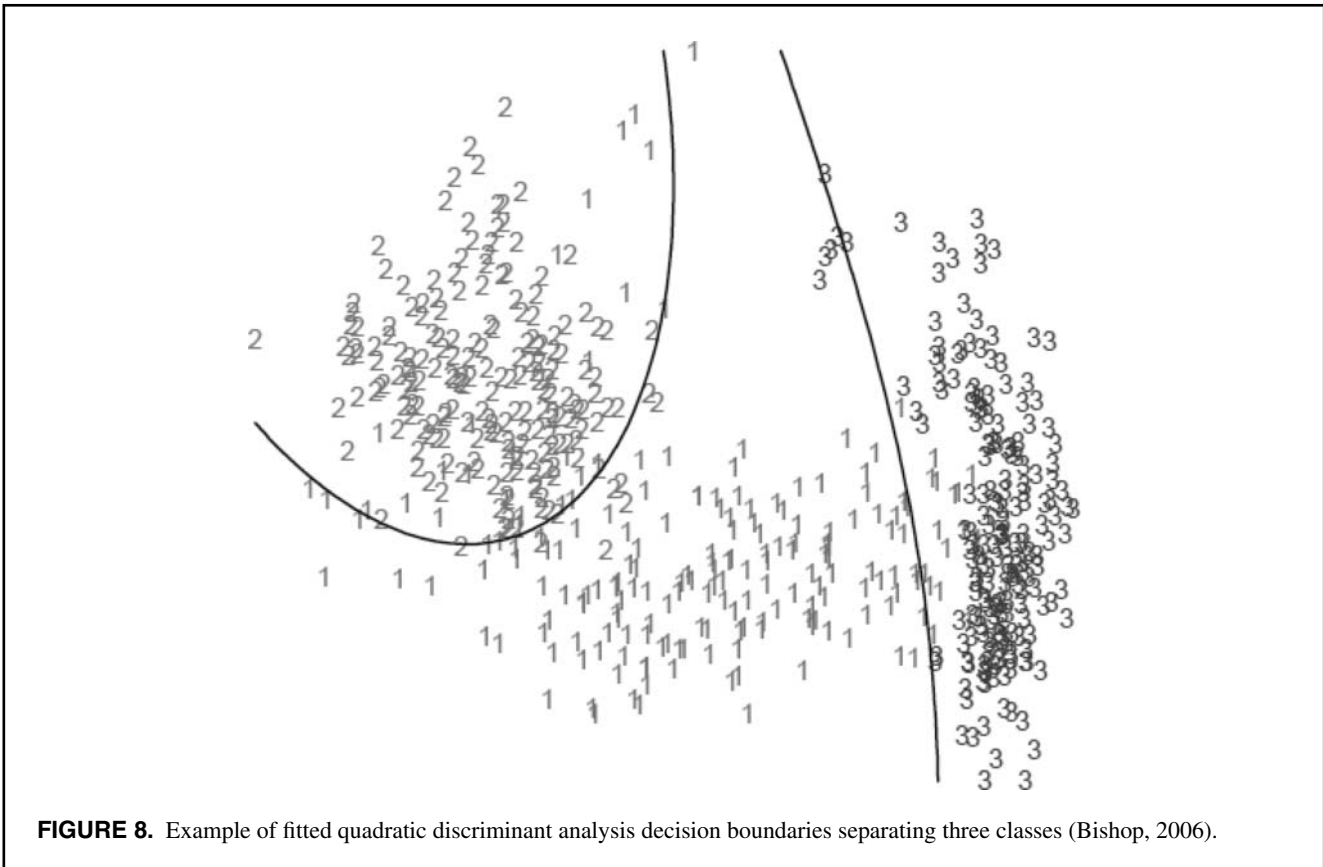
is fit to the training data. Also very popular, QDA separates measurements of classes of objects or events with a boundary between each pair of classes described by a quadratic equation (Figure 8).

Normal distributions are assumed, but each class can have a different covariance matrix. Thus, the separate covariance matrices must be estimated for each class, which means there are much more parameters than in LDA, that increase with the number of dimensions. Because the decision boundaries are functions of the parameters of the densities, counting the number of parameters must be carefully performed. When the homoscedastic assumption is true, the best possible test for the hypothesis that a given measurement is from a given class is the likelihood ratio test, similarly to the LDA. The discriminant function is given by testing the probability that a sample x is contained in one class or another:

$$\delta_k(x) = -\frac{1}{2} \log |\Sigma_k| - \frac{1}{2} (x - \mu_k)^T \Sigma_k^{-1} (x - \mu_k) + \log P(C = k). \quad (6)$$

where Σ_k is the covariance matrix for class k , and μ_k is the mean of class k . After calculating each covariance matrix, estimating the mean, μ_k , and calculating $P(C = k)$, the classification of new samples is accomplished by calculating their discriminant function value for each class.

The less rigid model underlying QDA may better approximate the solution in a real classification problem (compared with LDA). While the allowance of curved decision boundaries can lead to reduced bias in the estimation of the (unknown) optimal decision boundaries, having to estimate variance and covariance values with less data can lead to increased variance in the estimation of the optimal boundaries.



The literature shows that both LDA and QDA perform well on large and diverse set of classification tasks (Bishop, 2006).

NV

NV is one of the most efficient and effective inductive learning algorithms for machine learning and data mining (Zhang, 2005). The NV classifier is designed for use when features are independent of each another within each class, which is a rather unrealistic assumption that is almost always violated in real-world applications. However, it has surprisingly good performance, performing well in practice even when that independence assumption is not valid. Furthermore, this assumption dramatically simplifies the estimation. The individual class-conditional marginal densities can be estimated separately; also if the variables are discrete, then an appropriate histogram estimate can be used.

Assuming conditional independence among X_i s (vectors of observed random variables), Bayes rules is given by

$$P(Y = y_k | X_1 \dots X_n) = \frac{P(Y = y_k) \prod_i P(X_i | Y = y_k)}{\sum_j P(Y = y_j) \prod_i P(X_i | Y = y_j)}. \quad (7)$$

Hence, the NV classifier selects the class Y^{new} with maximum discriminant function for $X_{new} = \langle X_1, \dots, X_n \rangle$:

$$Y^{new} \leftarrow \arg \max_{y_k} P(Y = y_k) \prod_i P(X_i^{new} | Y = y_k). \quad (8)$$

If there is a continuous X_i , a common approach is to assume that $P(X_i | Y = y_k)$ follows a normal (Gaussian) distribution:

$$P(X_i = x | Y = y_k) = \frac{1}{\sqrt{2\pi\sigma_{ik}^2}} e^{-\frac{1}{2}\left(\frac{x-\mu_{ik}}{\sigma_{ik}}\right)^2}. \quad (9)$$

And the classification becomes:

$$Y^{new} \leftarrow \arg \max_{y_k} \pi_k \prod_i \mathcal{N}(X_i^{new}; \mu_{ik}, \sigma_{ik}). \quad (10)$$

in which π_k is estimated for each value of y_k by $\pi_k \equiv P(Y = y_k)$. Also, for each attribute X_i , it is necessary to estimate the class conditional mean μ_{ik} and variance σ_{ik} .

A well-known limitation of NV is in the case of binary features (Rish, 2001), where it can only learn linear discriminant functions, and thus it is always suboptimal for nonlinearly separable concepts. Nonetheless, NV has proven effective in many practical applications, including text classification,

medical diagnosis, and systems performance management (Domingos & Pazzani, 1997; Wolpert, 2007). Moreover, it is efficient using memory space and in terms of time complexity (Zhang, 2005; Zhang, Ling, & Zhao, 2000).

NVK

Kernel smoothing density estimation is an unsupervised learning procedure, which historically precedes kernel regression. This method fits a different but simple model separately at each query point x_0 , using only observations close to the target point, in a way that the resulting estimated function becomes smooth. This is accomplished via a weighting function or kernel $K_\lambda(x_0; x_i)$ that assigns a weight to x_i based on its distance from x_0 (Hastie et al., 2009). Kernel methods use weights that decrease smoothly to zero with distance from the target point. In high-dimensional spaces the distance kernels are modified to emphasize some variable more than others.

The kernel estimate is probably the most widely used nonparametric density estimation method (Hastie et al., 2009). Kernels K_λ are typically indexed by a parameter λ that controls the width of the neighborhood. This results in a memory-based method that requires little or no training at all where all the work gets done at evaluation time. The only parameter that needs to be determined from the training data is λ . The model is the entire training data set. Generally the kernel function is given by

$$K_\lambda(x_0, x) = PDF\left(\frac{|x - x_0|}{h_\lambda(x_0)}\right) \quad (11)$$

where, h_λ is a sequence of positive number usually called the bandwidth or smoothing parameter. Three popular PDFs used are Epanechnikov, tri-cube, and Gaussian distribution. A comparison of these kernels for local smoothing is presented in Figure 9. Bayesian classification and decision making are based on probabilities that a given set of measurements come from objects belonging to a certain class (probability theory). Statistical methods based on class conditional

probability density functions of features, are suitable in diverse classification tasks (Rahimizadeh, Marhaban, Kamil, & Ismail, 2009). Estimated PDFs have been used for classification utilizing Bayes’s formula. The classification can be done based on the probability density function, instead of estimating posterior probability using NV. The attempt is to estimate the underlying density function from the training data, and the idea is that the more data in a region, the larger is the density function.

Kernel smoothing density estimation leads naturally to a simple family of procedures for nonparametric density estimates for classification in a straightforward fashion using Bayes’s theorem (Hastie et al., 2009).

LS-SVM

SVM is a powerful methodology for solving problems in nonlinear classification, function estimation, and density estimation that has also led to many developments in kernel-based methods in general (Scholkopf & Smola, 2002; Vapnik, 1998). This method solves convex optimization problems, typically by quadratic programming. The LS-SVM is a reformulation to the standard SVMs (Suykens & Vandewalle, 1999), which was recently proposed. In fact, when the data points are linearly independent, LS-SVM is equivalent to hard marginal SVM (Ye & Xiong, 2007). LS-SVM involves the equality constraints only. Hence, the solution is obtained by solving a system of linear equations.

SVM models are similar to multilayer perceptron neural networks (Scholkopf & Smola, 2002). However, using a kernel function, SVMs are an alternative training method for polynomial, radial basis function and multilayer perceptron classifiers in which the weights of the network are found by solving a quadratic programming problem with linear constraints, rather than by solving a nonconvex, unconstrained minimization problem as in standard neural network training. Furthermore, rather than fitting nonlinear curves to the data, SVM handles this by using the kernel function to map the

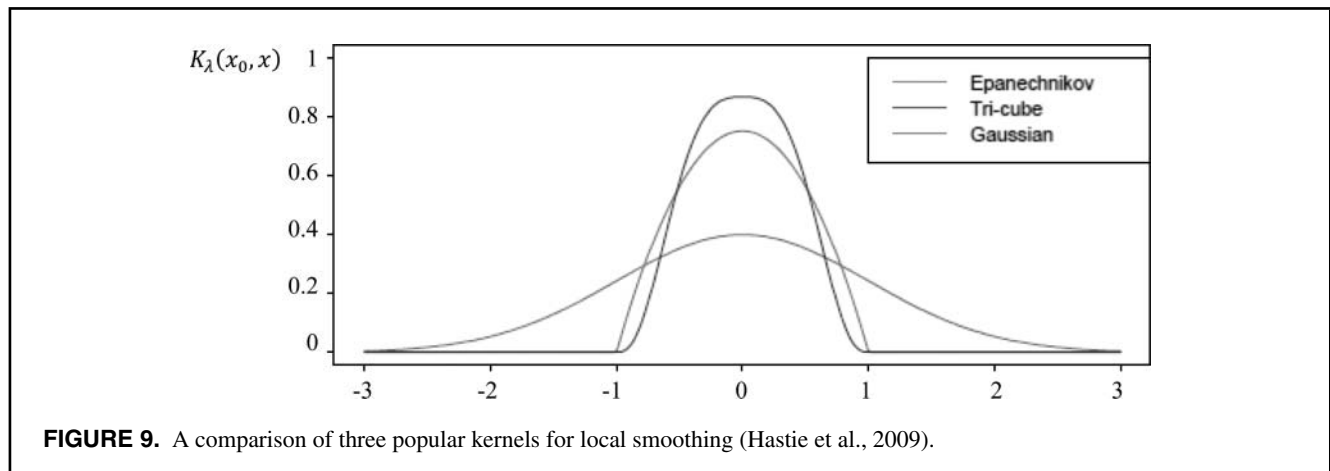


FIGURE 9. A comparison of three popular kernels for local smoothing (Hastie et al., 2009).

data into a different space where a hyperplane can be used to do the separation.

Many kernel mapping functions can be used but only a few have been found to work well in a wide variety of applications. The radial basis function (RBF) is one of the most used and recommended kernel mapping functions in human movement studies and others (Goudelis, Karpouzis, & Kollias, 2011; Hastie et al., 2009).

According to Hastie et al. (2009), kernel methods achieve flexibility by fitting simple models in a region local to the target point x_0 . Localization is achieved via a weighting kernel K_λ , and individual observations receive weights $K_\lambda(x_0; x_i)$. Radial basis functions combine these ideas, by treating the kernel functions $K_\lambda(\mu; x)$ as base functions, where each basis element is indexed by a location and a scale parameter (μ_m and λ_m , respectively). Thus, radial basis functions are symmetric p-dimensional kernels located at particular centroids:

$$f_\theta(x) = \sum_{m=1} M_{\lambda_m}(\mu_m, x)\theta_m \quad (12)$$

The centroids μ_m and scales λ_m have to be determined. A usual choice for the probability density functions is the standard Gaussian density function. There are also several approaches for learning the parameters μ_m , λ_m and θ_m . For example, a popular method is estimating θ_m , given μ_m and λ_m by a simple least squares problem. Often the kernel parameters μ_m and λ_m are chosen in an unsupervised way using the X distribution alone. One of the methods is to fit a Gaussian mixture density model to the training x_i , which provides both the centers μ_m and the scales λ_m . Figure 10 shows an example of Gaussian radial basis function kernels with scale parameter $\lambda = 1$ and centered at 5 centroids, which were chosen at random.

The LS-SVM classifier was implemented using the Least Squares – Support Vector Machines MatLab Toolbox (Pelckmans et al., 2002).

Results

Experimental results are presented in two subsections: (a) we present a comparative case study on the classification accuracy among the five methods for the detection of signatures in the performance of the golf putting and (b) based on these obtained results, we extract the individual putting signatures using the best classification method.

The parameters that present each player's trial are the ones from the mathematical model of the putting previously presented in Equation 4.

Comparison of Classification Methods

At this stage, intensive Matlab simulation was performed using the detection algorithm and the DPSO as an estimation technique with the earlier defined parameters to obtain the putter's motion function that describes 30 putt executions of six different expert participants (classes), in a total of 180 trials.

After calculating all the estimation parameters, the five classification methods previously described were used in order to identify the signature of each player.

Figures 11–13 and Table 1 depict the confusion matrix and the AUC of the five classifiers.

In the Figures 11–13, the LS-SVM shows a better classification accuracy because it presents a higher percentage of TP in almost all situations closely followed by the NVK method. However, in order to confirm the superiority of the LS-SVM over the NVK, the AUC of each player's trials was determined for the five classification methods. In order to allow a straightforward comparison of the five classifiers, next table depicts the average value of the AUC highlighting the maximum value for each player's trial.

Table 1 provides evidence about the superiority of the LS-SVM classifier, which shows better results in the majority of the trials closely followed by the NVK classifier.

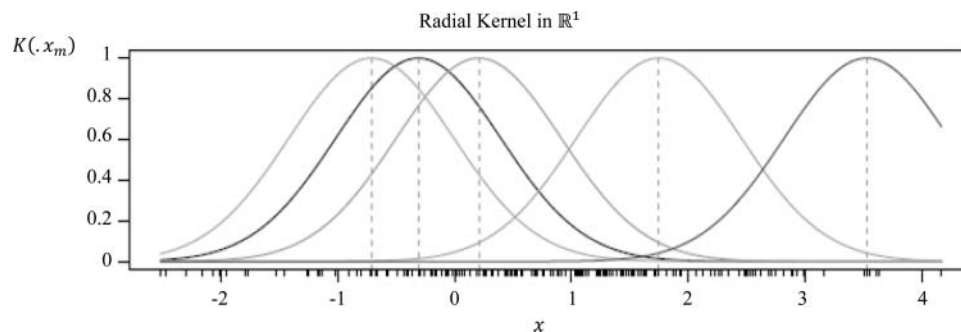


FIGURE 10. An example of a Gaussian radial basis function kernel centered at 5 centroids chosen at random and scale parameter $\lambda = 1$ (Hastie et al., 2009).

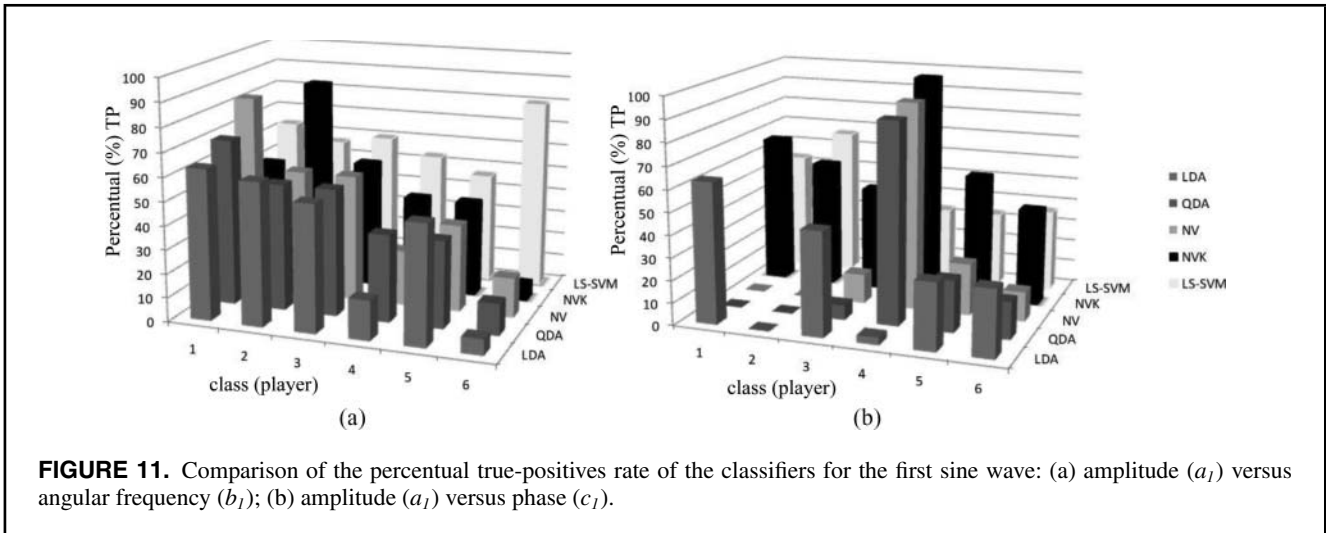


FIGURE 11. Comparison of the percental true-positives rate of the classifiers for the first sine wave: (a) amplitude (a_1) versus angular frequency (b_1); (b) amplitude (a_1) versus phase (c_1).

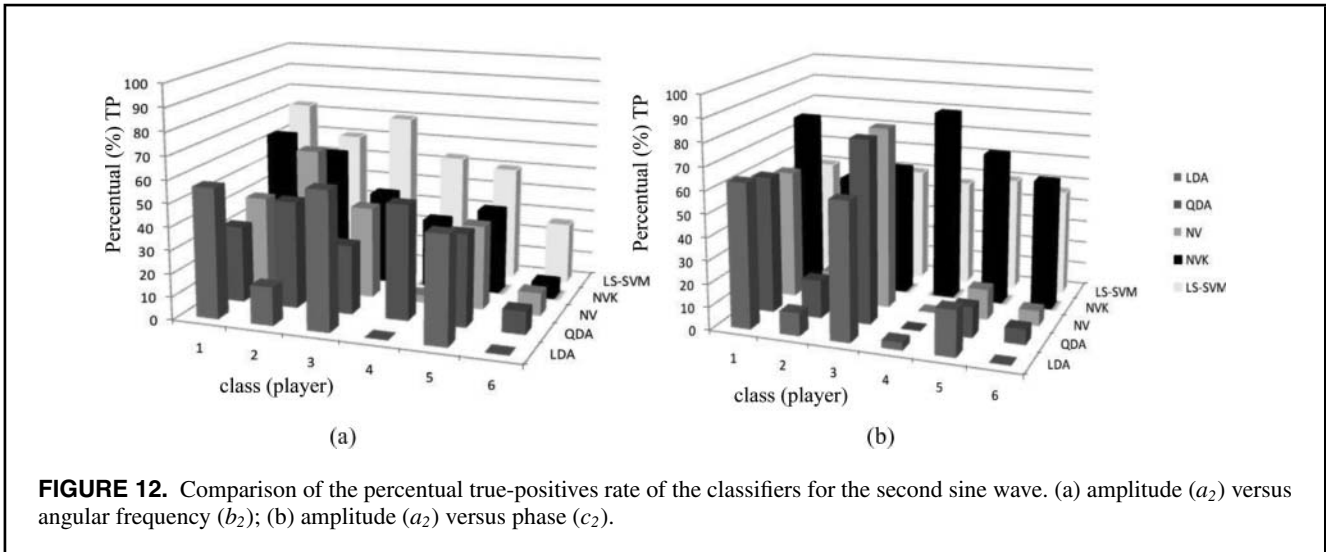


FIGURE 12. Comparison of the percental true-positives rate of the classifiers for the second sine wave. (a) amplitude (a_2) versus angular frequency (b_2); (b) amplitude (a_2) versus phase (c_2).

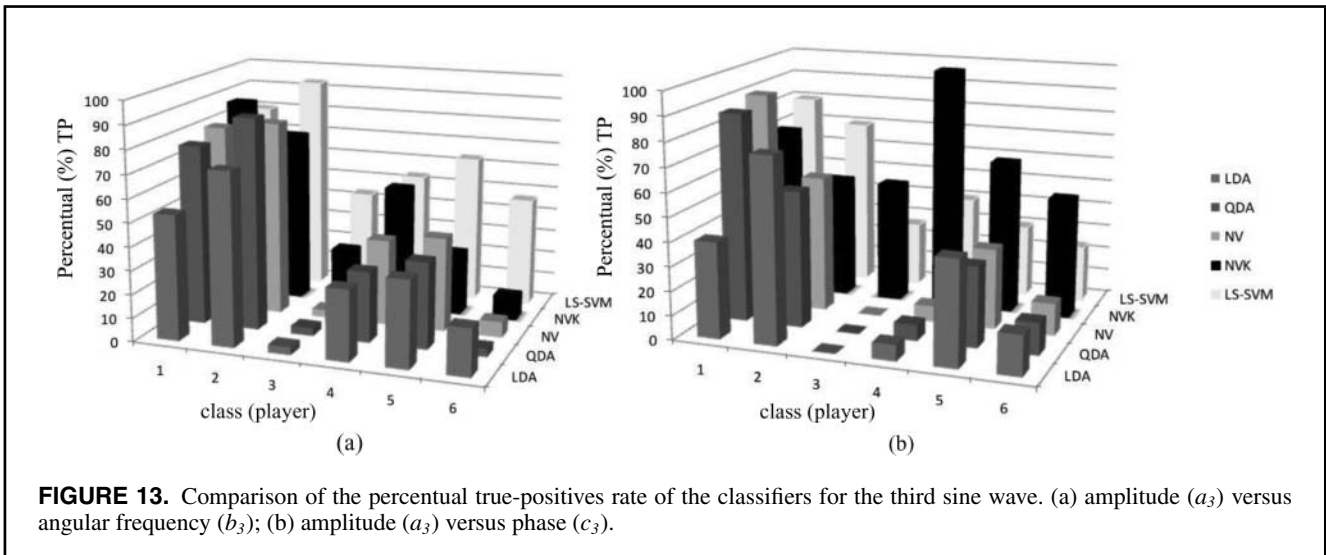


FIGURE 13. Comparison of the percental true-positives rate of the classifiers for the third sine wave. (a) amplitude (a_3) versus angular frequency (b_3); (b) amplitude (a_3) versus phase (c_3).

TABLE 1. Average Value of the Area Under the Receiver-Operating Characteristic Curve on the Five Classification Methods

Class	LDA	QDA	NV	NVK	SVM
1	0.619	0.601	0.671	0.680	0.744
2	0.650	0.623	0.692	0.685	0.737
3	0.566	0.582	0.634	0.761	0.734
4	0.507	0.585	0.574	0.675	0.690
5	0.622	0.651	0.692	0.766	0.797
6	0.493	0.602	0.650	0.718	0.745

Note. Bolded text indicates the higher area under the receiver-operating characteristic curve value. LDA = linear discriminant analysis; QDA = quadratic discriminant analysis; NV = naive Bayes with normal distribution; NVK = naive Bayes with kernel smoothing density estimate; SVM = support vector machines.

TABLE 2. Correspondent Player-to-Color. (Color table available online).

Player 1	Player 2	Player 3	Player 4	Player 5	Player 6
----------	----------	----------	----------	----------	----------

However, as expected, intersections (misclassified data) between the regions in the parameter’s space were common. This was due to the players’ expertise resulting in a similar value of AUC of approximately 70%.

These results showed that classification algorithms could describe the motor performance of individual players, considering them as a unique signature for each player. The scattering data also bears that it was possible to categorize different ways to perform the golf putt.

Although data didn’t clarify if a player had a better or worse performance compared to others, it was clear that there are individual profiles that are unique to each player. That is, a greater or smaller dispersion of data (colored area) may indicate whether a player was more consistent and regular than others along the motor execution. However, more research should be conducted in order to consolidate the results obtained in this work, and thus avoiding falling into a classification error.

For instance, player 6 showed a more consistent and unique putt gesture because it presented regions with a smaller area when compared to other players. This means that player 6 is more regular than the other players. The opposite is also true for players 4 and 5. The cyan and purple regions are scattered and usually form disconnected regions.

As it can be seen in the results, all players displayed unique putting characteristics. Similarities could be easily identified between them; however, the same could be stated about differences. These differences allowed for the extraction of a signature for each player, confirming that a given putting stroke had a higher probability of belonging to a specific player.

Extracting the Individual Putting Signature

The previous results compared the performance of the five classifiers used to identify the signature of each player. According to the analysis of the confusion matrix and the ROC, the LS-SVM was the one that presented an overall better performance.

On Table 2, each color corresponds to a player (class) following the colors specified by the built-in MatLab Colormaps lines (Table 2). The colored areas correspond to the relation between the parameters of the kinematical model previously defined in Equation 4 for the 30 trials performed by each player.

It is shown in Figures 14–16 that each player’s putting ability can be defined by six pairs of parameters related to the three sinusoids of the estimation model. Different players have different regions defined in the parameter’s space, which in turn, indicates a difference in the game style of each player.

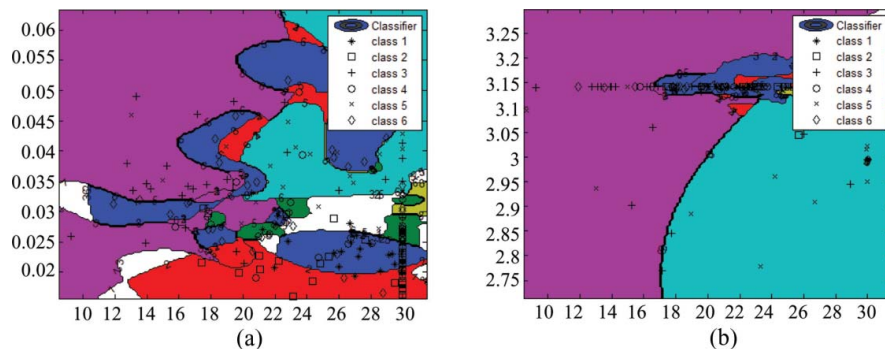


FIGURE 14. Analysis of the first sine wave–least squares support vector machines classification of all six players. (a) amplitude (a_1) versus angular frequency (b_1); (b) amplitude (a_1) versus phase (c_1). (Color figure available online).

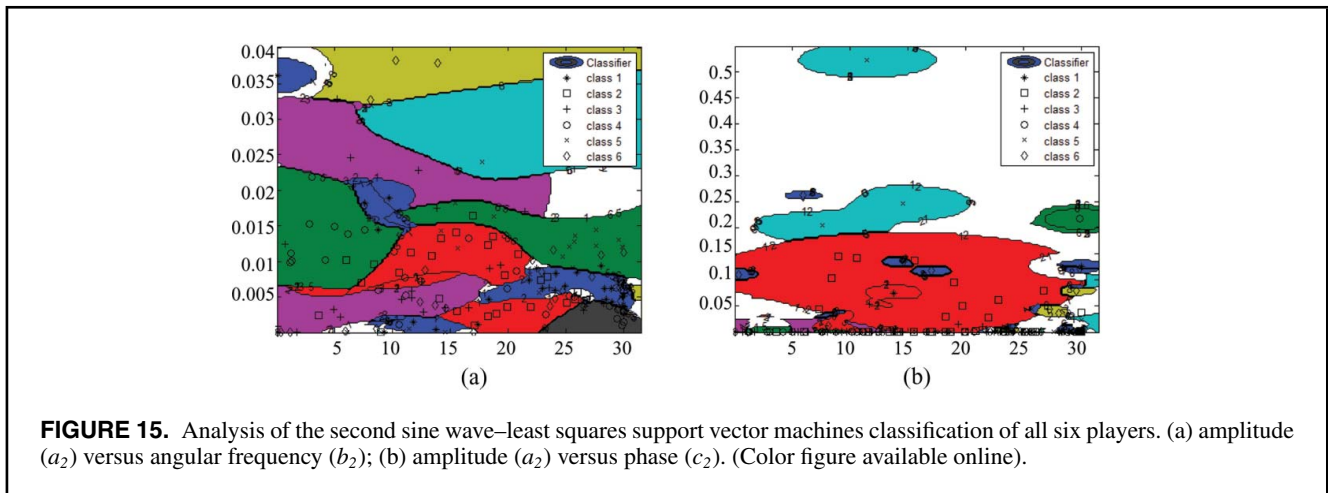


FIGURE 15. Analysis of the second sine wave-least squares support vector machines classification of all six players. (a) amplitude (a_2) versus angular frequency (b_2); (b) amplitude (a_2) versus phase (c_2). (Color figure available online).

Discussion

Experimental results of the putting performance of six expert golf players clearly show, when considering different trials of every player, that each one of them had a typical and distinct game style, a signature, confirming our expectancies.

The presented system for data classification is functional and allows retrieving a series of different types of information simultaneously. In this work, the putter's detection was carried out using a simple and computationally efficient computer vision algorithm. Also, a study of the nonlinear estimation technique DPSO was conducted in order to extract a sinusoidal function to model the trajectory of the putter's head in time. Therefore, five classification methods were used and compared in order to identify the putting signature of each player. Results confirmed the superior performance of the LS-SVM method.

Scrutinizing this movement solely based on detection and estimation methods, which are necessary to understand the process variables of motor performance associated with a

complex and dynamic gesture such as the putt (e.g., velocity, acceleration, position of the putter during players' execution) may not be enough. According with the coordination dynamics (Kelso, 1995) of each individual, we should also understand the development of new coordinative patterns and describe the human performance as a individual, unique, creative process where every athlete discovers new and different solutions to problems in motor movement (Hirstovski et al., 2011).

The present work was particularly important with regard to the study of movements and motor skills training, going beyond the analysis of the measures associated with the product and the response magnitude (e.g., position of the ball over the hole or quantification of the final result). We argue that it is important to address the motor execution process variables when facing constraints (Newell, 1986), and how the behavior of individuals changes and evolves over time (Davids, Button, & Bennett, 2008). To this end it would be worth considering a study with golf novices.

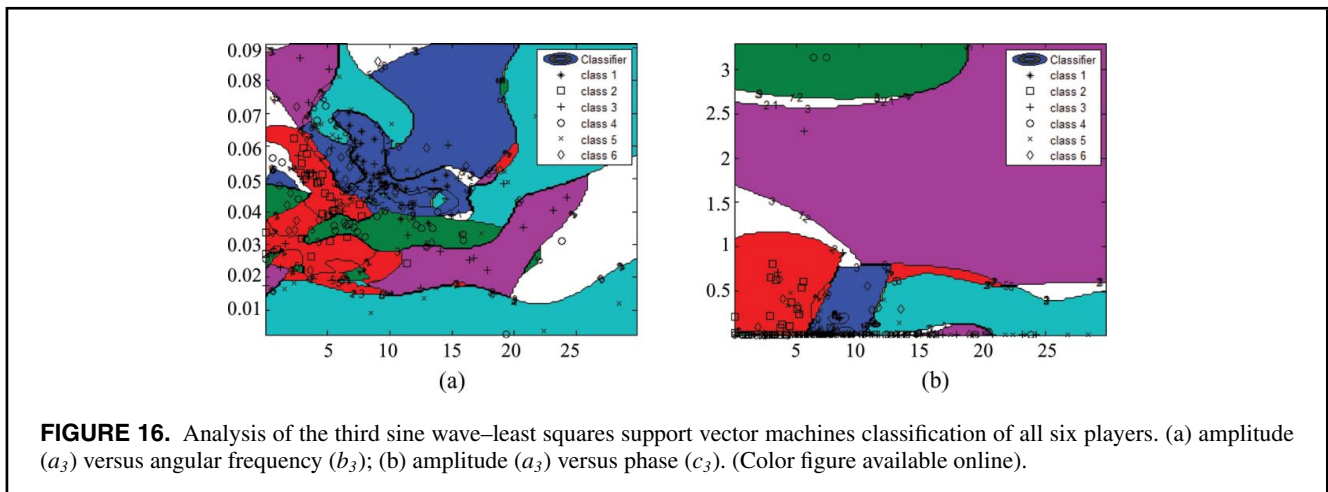


FIGURE 16. Analysis of the third sine wave-least squares support vector machines classification of all six players. (a) amplitude (a_3) versus angular frequency (b_3); (b) amplitude (a_3) versus phase (c_3). (Color figure available online).

The findings from this study also point toward the implementation of new methods for the performance analysis of the putting, where the individual analysis of kinematics are preferable to the traditional pooling of group data. In this line of thought, the analysis of motor behavior profiles of athletes based on the measurement of individual kinematic strategies can help in better understanding the relevant changes resulting from the interaction between the athlete's characteristics and the surroundings where the task is performed.

Regarding the functional concept of variability and the contribution that the classification of signatures may have in understanding the motor control of golf putting, it is considered important to further investigate the structure of the variability and influence that this movement may have in the performance of athletes with different skill levels (Newell & Corcos, 1993; Newell, Deutsch, Sosnoff, & Mayer-Kress, 2006). In this sense, given the overemphasis on current research on motor behavior on pooling of group data, it is also important to focus in the analysis of the individual performance on this type of motor skills (Button, Davids, & Schöllhorn, 2006; Davids et al., 2008).

Given neuromotor degeneracy (Davids et al., 2007), it is plausible that a golfer may present several solutions to solve a particular motor problem in a given performance context (Pelz, 2000). This means that, sometimes, in a golf game, different motor solutions are necessary to appropriately adapt to the variable conditions where the performance occurs (cf. Newell & Corcos, 1993; Davids et al., 2006; Harbourne & Stergiou, 2009; Latash et al., 2002; Newell & Slifkin 1998; Riley & Turvey, 2002).

Moreover, it is certain that technological advances and new tools that recently emerged to human movement pattern analysis can make a strong contribution to the study of golf putting (Pelz, 2000). Thus, we believe that this work may contribute to a deeper analysis of the human motor behavior and performance (Beek et al., 1996; Kelso, 1995; Port et al., 1997) in other sports movements (Button et al., 1998; Davids et al., 2008). We also point out that the performance analysis of the putting can benefit from other methods of performance classification similar to those presented in this study.

Hereupon, extending the obtained results to other movements or motor skills, we believe that similar studies could eventually be performed to analyze and extract signatures in the context of other sports such as hand motion on swimming, boxing, tennis, or baseball or even other body parts in everyday human actions.

ACKNOWLEDGMENTS

This work was supported in part by *RoboCorp* at the Engineering Institute of Coimbra (ISEC) and a PhD scholarship (SFRH/BD/73382/2010) granted to the first author by the Portuguese Foundation for Science and Technology (FCT).

REFERENCES

Alexander, D. L., & Kern, W. (2005). Drive for show and putt for dough. *Journal of Sports Economics*, 6, 46–60.

- Altun, K., Barshan, B., & Tunçel, O. (2010). Comparative study on classifying human activities with miniature inertial and magnetic sensors. *Pattern Recognition*, 43, 3605–3620.
- Araújo, D., Davids, K., & Hristovski, R. (2006). The ecological dynamics of decision making in sport. *Psychology of Sport and Exercise*, 7, 653–676.
- Beek, P. J., Rikkert, W., & van Wieringen, P. C. (1996). Limit cycle properties of rhythmic forearm movements. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 1077–1093.
- Bishop, C. M. (2006). *Pattern recognition and machine learning (information science and statistics)*. New York, NY: Springer-Verlag.
- Bonilla-Petriciolet, A., & Segovia-Hernández, J. G. (2010). A comparative study of particle swarm optimization and its variants for phase stability and equilibrium calculations in multicomponent reactive and non-reactive systems. *Fluid Phase Equilibria*, 289, 110–121.
- Button, C., Bennett, S. J., & Davids, K. (1998). Coordination dynamics of rhythmical and discrete prehension: Implications for the scanning procedure and individual differences. *Human Movement Science*, 17, 801–820.
- Button, C., Davids, K., & Schöllhorn, W. (2006). Coordination profiling of movement systems. In K. Davids, S. J. Bennett, & K. Newell (Eds.), *Movement system variability* (pp. 133–152). Champaign, IL: Human Kinetics.
- Button, C., Macleod, M., Sanders, R., & Coleman, S. (2003). Examining movement variability in the throwing action at different skill levels. *Research Quarterly in Exercise and Sport*, 74, 257–269.
- Carlisle, A., & Dozier, G. (2001). An off-the-shelf PSO. *Proceedings of the Particle Swarm Optimization Workshop*, 1–6.
- Chan, K., & Lee, T. W. (2002). Comparison of machine learning and traditional classifiers in glaucoma diagnosis. *IEEE Transactions on Biomedical Engineering*, 49, 963–974.
- Coello, Y., Delay, D., Nougier, V., & Orliaguët, J. P. (2000). Temporal control of impact movement: The “time from departure” control hypothesis in golf putting. *International Journal of Sport Psychology*, 31, 24–46.
- Couceiro, M. S., Dias, G., Luz, J. M. A., Figueiredo, C. M., & Ferreira, N. M. F. (2010). *Parameter estimation for a mathematical model of the golf putting*. Paper presented at WACI 10 – ISEC.IPC, Coimbra, Portugal.
- Couceiro, M. S., Luz, J. M., Figueiredo, C. M., Ferreira, N. M. F., & Dias, G. (2010). Parameter estimation for a mathematical model of the golf putting. In V. M. Marques, C. S. Pereira, & A. Madureira (Eds.), *Proceedings of WACI. ISEC.IPC* (pp. 1–8). Coimbra, Portugal: ISEC.IPC.
- Couceiro, M. S., Portugal, D., Gonçalves, N., Rocha, R., Luz, J. M. A., Figueiredo, C. M., & Dias, G. (2012, June). A methodology for detection and estimation in the analysis of the golf putting. *Pattern Analysis and Applications*, 1–16. DOI: 10.1007/s10044-012-0276-8
- Davids, K., Araújo, D., Button, C., & Renshaw, I. (2007). Degenerate brains, indeterminate behavior and representative tasks: Implications for experimental design in sport psychology research. In G. Tenenbaum, & B. Eklund (Eds.), *Handbook of sport psychology* (3rd ed., pp. 224–244). Hoboken, NJ: Wiley.
- Davids, K., Button, C., & Bennett, S. (2008). *Dynamics of skill acquisition: A constraints-led approach*. Champaign, IL: Human Kinetics.
- Davids, K., Bennett, S. J., & Newell, K. M. (2006). *Movement system variability*. Champaign, IL: Human Kinetics.
- Delay, D., Nougier, V., Orliaguët, J. P., & Coello, Y. (1997). Movement control in golf putting. *Human Movement Science*, 16, 597–619.

- Dias, G., Luz, M., Couceiro, M., Figueiredo, C., Fernandes, O., Iglésias, P., et al. (2009). *Human motor behavior and engineering: Contributions and interactions in the study of putt*. Paper presented at WACI'09. ISEC.IPC, Coimbra, Portugal.
- Domingos, P., & Pazzani, M. (1997). On the optimality of the simple Bayesian classifier under zero-one loss. *Machine Learning*, 29, 103–130.
- Edelman, G., & Gally, J. (2001). Degeneracy and complexity in biological systems. *Proceedings of the National Academy of Science*, 98, 13763–13768.
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27, 861–874.
- Freitas, C. O. A., Carvalho, J. M., Oliveira, J. J., Aires, S. B. K., & Sabourin, R. (2007). Confusion matrix disagreement for multiple classifiers. *Progress in Pattern Recognition, Image Analysis and Applications. Lecture Notes in Computer Science*, 4756, 387–396.
- Gehrig, N., Lepetit, V., & Fua, P. (2003, September). *Visual golf club tracking for enhanced swing analysis*. Paper presented at the British Machine Vision Conference, Norwich, UK.
- Goudelis, G., Karpouzis, K., & Kollias, S. (2011). *Robust human action recognition using history trace templates*. Paper presented at the 12th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), Delft, the Netherlands.
- Harbourne, R. T., & Stergiou, N. (2009). Movement variability and the use of nonlinear tools: Principles to guide physical therapist practice. *Journal of Neurologic Physical Therapy*, 89, 267–82.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. New York, NY: Springer.
- Hristovski, R., Davids, K., Araújo, D., & Passos, P. (2011). Constraints-induced emergence of functional novelty in complex neurobiological systems: A basis for creativity in sport. *Nonlinear Dynamics, Psychology, and Life Sciences*, 15, 175–206.
- Hume, P. A., Keogh, J., & Reid, D. (2005). The role of biomechanics in maximising distance and accuracy of golf shots. *Sports Medicine*, 35, 429–49.
- Janssen, D., Schöllhorn, W. I., Newell, K. M., Jäger, J. M., Rost, F., & Vehof, K. (2011). Diagnosing fatigue in gait patterns by support vector machines and self-organising maps. *Human Movement Science*, 30, 966–975.
- Jonassen, D. H., & Grabowski, B. L. (1993). *Handbook of individual differences, learning and instruction*. Hillsdale, NJ: Erlbaum.
- Kaladhar, D. S. V. G. K., Nageswara Rao, P. V., & Ramesh Naidu Rajana, B. L. V. (2010). Confusion matrix analysis for evaluation of speech on Parkinson disease using Weka and MatLab. *International Journal of Engineering Science and Technology*, 2, 2733–2736.
- Karlsen, J., Smith, G., & Nilsson, J. (2008). The stroke has only a minor influence on direction consistency in golf putting among elite players. *Journal of Sports Sciences*, 26, 243–250.
- Kelso, J. A. S. (1995). *Dynamics patterns: The self-organization of brain and behavior*. Cambridge MA: MIT Press.
- Kennedy, J., & Eberhart, R. C. (1995). Particle swarm optimization. *Proceedings of the IEEE International Conference on Neural Networks, Perth, Australia, IEEE Service Center*, 12–13.
- Latash, M. L., Scholz, J. P., & Schoner, G. (2002). Motor control strategies revealed in the structure of motor variability. *Exercise and Sport Sciences Reviews*, 30, 26–31.
- Lee, M., Roan, M., Smith, B., & Lockhart, T. E. (2009). Gait analysis to classify external load conditions using linear discriminant analysis. *Human Movement Science*, 28, 226–235.
- Martínez, A. M., & Kak, A. C. (2001). PCA versus LDA. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23, 228–233.
- Mayer-Kress, G., Liu, Y.-T., & Newell, K. M. (2006). Complex systems and human movement. *Complexity*, 12, 40–51.
- Nesbit, S. M., Hartzell, T. A., Nalevanko, J. C., & Starr, R. M. (1996). A discussion of iron golf club inertia tensors and their effects on the golfer. *Journal of Applied Biomechanics*, 12, 449–469.
- Newell, K. M. (1986). Constraints on the development of coordination. In M. G. Wade & H. T. A. Whiting (Eds.), *Motor development in children: Aspects of coordination and control* (pp. 341–360). Boston, MA: Martinus Nijhoff.
- Newell, K. M., & Corcos, D. M. (1993). Issues in variability and motor control. In K. M. Newell & D. M. Corcos (Eds.), *Variability and motor control* (pp. 1–12). Champaign, Illinois: Human Kinetics.
- Newell, K. M., Deutsch, K., Sosnoff, J., & Mayer-Kress, G. (2006). Variability in motor output as noise: A default and erroneous proposition? In K. Davids, S. Bennett, & K. Newell (Eds.), *Movement system variability* (pp. 3–24). Champaign, IL: Human Kinetics.
- Newell, K. M., & Slifkin, A. B. (1998). The nature of movement variability. In J. P. Piek (Ed.), *Motor behavior and human skill: A multidisciplinary perspective* (pp. 143–164). Champaign, IL: Human Kinetics.
- Newell, K. M., Slobounov, S. M., Slobounova, E., & Molenaar, P. C. M. (1997). Short-term non stationarity and the development of postural control. *Gait and Posture*, 6, 56–62.
- Pelckmans, K., Suykens, J. A. K., Gestel, T. V., Brabantar, J. D., Lukas, L., Hamers, B., & Vandewalle, J. (2002). *LS-SVMlab: A MATLAB/C toolbox for least squares support vector machines* (Internal Report 02-44). Leuven-Heverlee, Belgium: ESAT-SISTA.
- Pelz, D. (1989). *Putt like the pros*. New York, NY: HarperCollins.
- Pelz, D. (2000). *Putting bible: The complete guide to mastering the green*. New York, NY: Doubleday.
- Phillips, E., Davids, K., Renshaw, I., & Portus, M. (2010). Expert performance in sport and the dynamics of talent development. *Sports Medicine*, 40, 271–283.
- Port, N. L., Lee, D., Dassonville, P., & Georgopoulos, P. (1997). Manual interception of moving targets. I. Performance and movement initiation. *Experimental Brain Research*, 116, 406–420.
- Rahimizadeh, H., Marhaban, M. H., Kamil, R. M., & Ismail, N. B. (2009). Color image segmentation based on Bayesian theorem and kernel density estimation. *European Journal of Scientific Research*, 26, 430–436.
- Rennie, J. D. (2001). *Improving multi-class text classification with naive bayes*. Master's thesis, Massachusetts Institute of Technology.
- Riley, M. A., Kuznetsov, N., & Bonnette, S. (2011). State-, parameter-, and graph-dynamics: Constraints and the distillation of postural control systems. *Science & Motricité*, 74, 5–18.
- Riley, M. A., & Turvey, M. T. (2002). Variability and determinism in motor behavior. *Journal of Motor Behavior*, 34, 99–125.
- Rish, I. (2001). An empirical study of the naive Bayes classifier. In *Proceedings of IJCAI-01 workshop on Empirical Methods in AI* (pp. 41–46). Available at <http://ijcai.org/Past%20Proceedings/IJCAI-2001/content/content.htm>
- Roushanzamir, S., Valafar, H., & Valafar, F. (1999). A comparative study of linear and quadratic discriminant classifier techniques for variable selection: A case study in predicting the effectiveness of hydroxyurea treatment of sickle cell anemia. In *Proceedings of the International Joint Conference on Neural Networks (IJCNN '99)* (pp. 3627–3631). Washington, DC: International Joint Conference on Neural Networks.
- Scholkopf, B., & Smola, A. (2002). *Learning with kernels*. Cambridge, MA: MIT Press.
- Schöllhorn, W. I. (2000). Applications of systems dynamic principles to technique and strength training. *Acta Academiae Olympicae Estoniae*, 8, 67–85.

- Schöllhorn, W. I., Mayer-Kress, G., Newell, K. M., & Michelbrink, M. (2008). Time scales of adaptive behavior and motor learning in the presence of stochastic perturbations. *Human Movement Science*, 28, 319–333.
- Schöllhorn, W. I., Nigg, B. M., Stefanyshyn, D. J., & Liu, W. (2002). Identification of individual walking patterns using time discrete and time continuous data sets. *Gait and Posture*, 15, 180–186.
- Suykens, J. A. K., & Vandewalle, J. (1999). Least squares support vector machine classifiers. *Neural Processing Letters*, 9, 293–300.
- Tillett, T., Rao, T. M., Sahin, F., & Rao, R. (2005). Darwinian particle swarm optimization. *Proceedings of the 2nd Indian International Conference on Artificial Intelligence*, 1474–1487.
- Urtasun, R., Fleet, D. J., & Fua, P. (2005). Monocular 3-d tracking of the golf swing. *IEEE Conference on Computer Vision and Pattern Recognition*, 2, 932–938.
- Vapnik, V. (1998). *Statistical learning theory*. New York, NY: Wiley.
- Wolpert, D. M. (2007). Probabilistic models in human sensorimotor control. *Human Movement Science*, 26, 511–524.
- Ye, J., & Xiong, T. (2007). SVM versus least squares SVM. *Proceedings of the Eleventh International Conference on Artificial Intelligence and Statistics (AISTATS 2007)*, 640–647.
- Zhang, H. (2005). Exploring conditions for the optimality of naive Bayes. *International Journal of Pattern Recognition and Artificial Intelligence*, 19, 183–198.
- Zhang, H., Ling, C. X., & Zhao, Z. (2000). *The learnability of naive Bayes*. In H. Hamilton & Q. Yang (Eds.), *Advances in artificial intelligence (proceedings of CAI2000)* (pp. 432–441). New York, NY: Springer.

Received May 10, 2012
 Revised September 17, 2012
 Accepted October 10, 2012