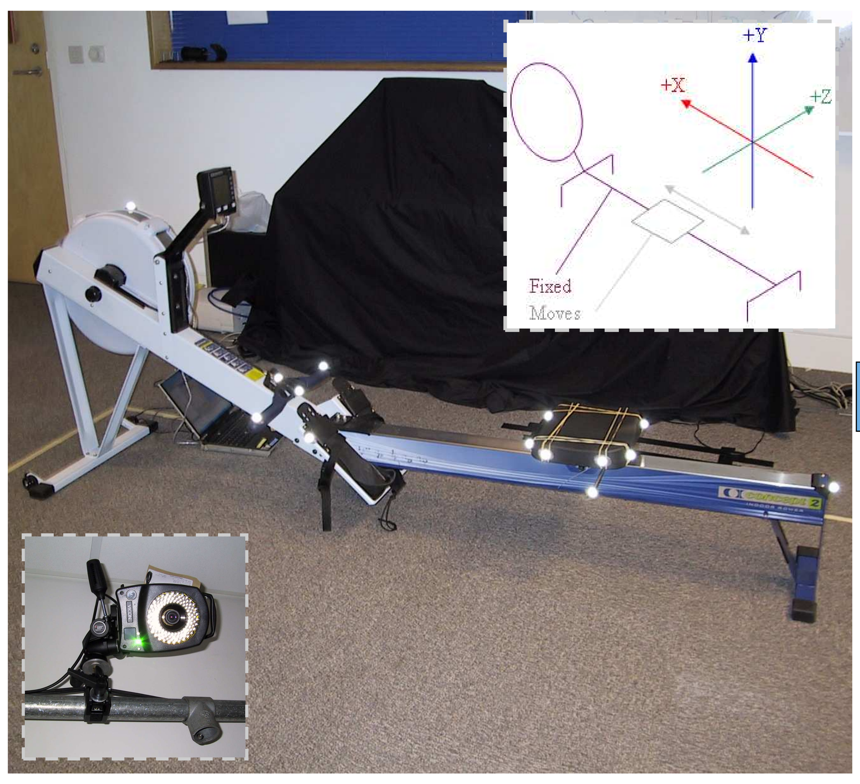
Automatic Coaching of Rowing Technique

Introduction

It is possible to automatically judge the quality of rowing technique exhibited in a performance, with or without having previously seen the athlete, to an accuracy greater than 93% or 67% respectively. This is found to be possible for certain individual aspects of technique, without needing to input any explanations of them or rules, from a coach. Using only the motion of the handle of an ergometer measured using a high fidelity motion capture system, rowers can receive objective feedback when a human coach is unavailable, which helps to ensure they exercise safely and maintain an optimal technique. Coaches' time is valuable and they will get fatigued or blinded to changes in an individual after a while. Their judgements are not easily explained mathematically and tend to be qualitative. A system that can achieve binary classification over the quality of performances having only been shown examples of different techniques has potential to help the sporting community achieve excellence as well as feed into wider reaching healthcare applications.

Feature extraction and modelling



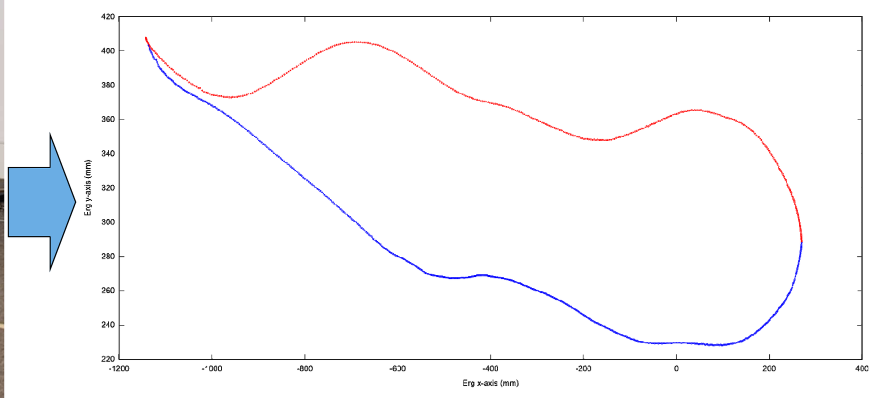


Fig. 1. Ergometer in VICON stage augmented with retro-reflective markers (left); VICON camera (bottom-left); Erg-coordinate system (top-left), Side projection of ergometer's handle trajectory (right)

The handle's trajectory is recorded using a VICON motion capture system at 200Hz, occlusions are linearly interpolated and it is projected into a coordinate system relative to the ergometer (see Fig.1). Once segmented into single strokes, features based on the tempo-spatial structure of the trajectory in 3D space are calculated for each stroke, (see Fig. 2.) ensuring they are invariant to the average speed:

Abstract features including the trajectory's length, height and distance are computed. Moments of the side-projection are calculated using just the shape (λ^{11}, λ^{12}, λ^{21}, λ^{02}, λ^{20}) and both shape and speed (μ^{11}, μ^{12}, μ^{21}, μ^{02}, μ^{20}).

$$\lambda^{pq} = \sum (h_x(s)^p \, h_y(s)^q) \,, \qquad \mu^{pq} = \sum (h_x(s)^p \, h_y(s)^q \, \psi(s))$$

Each sample, s, has coordinates ($h_x(s)$, $h_y(s)$) and instantaneous mean-subtracted speed $\psi(s)$.

Physical Performance features include the variance of the samples from the best fit line when projected into the X-Z plane (wobble) and a measurement of the smoothness:

Speed smoothness:

1. Using the instantaneous speed signal,
2. Subtract the mean,
3. Low pass filter at 3Hz,
4. Double-differentiate w.r.t. time to accentuate peaks,
5. Sum the resulting samples.

Shape smoothness:

1. Using the signal formed from the distances between each consecutive sample and the centroid of the trajectory,
2. Low pass filter at 6 Hz,
3. Calculate the absolute values of the double-differential w.r.t. time,
4. Count the number of samples with resulting values above a threshold of 0.4 ms^{-2}.

Domain specific features include the ratio of the time spent on the drive and recovery stages of the stroke and the angles between the best fit line in the X-Z plane and the X-axis for the drive and recovery sections of the trajectory.

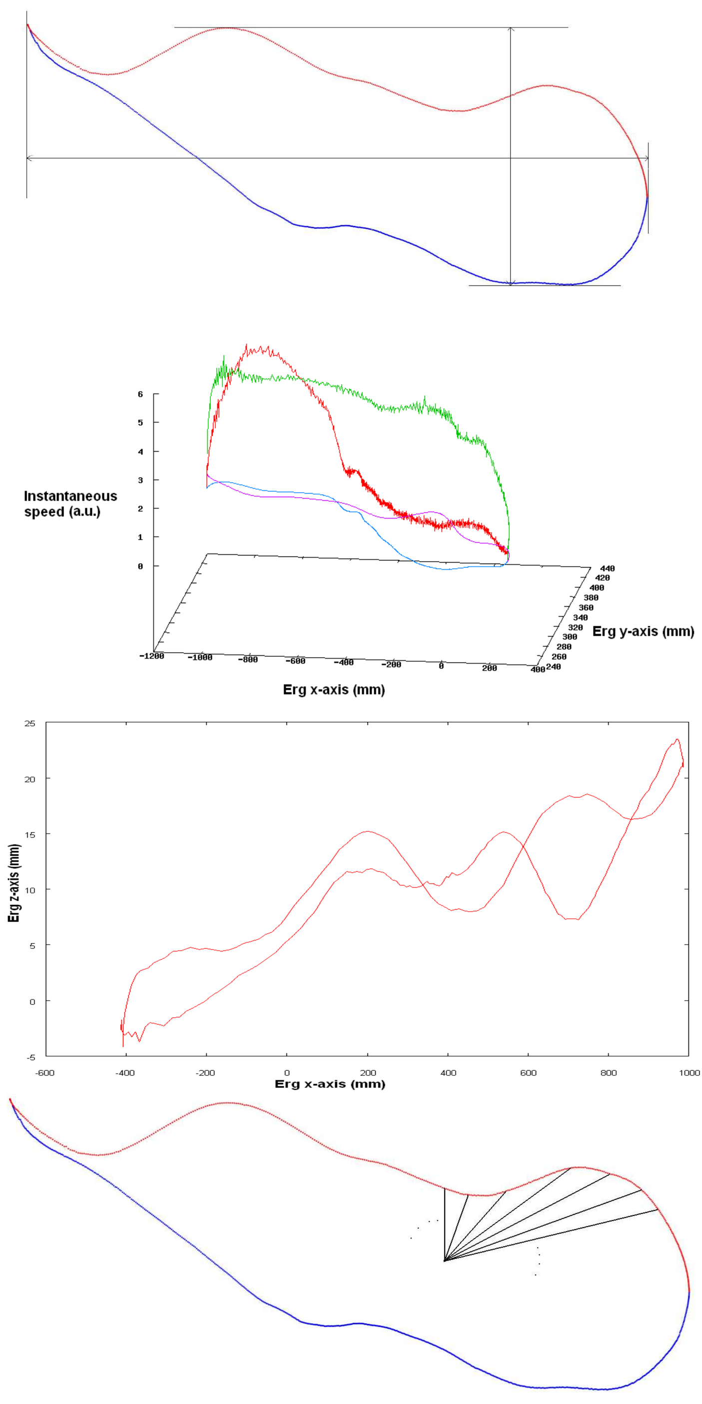


Fig. 2. in descending order: (1) Side projection of the handle trajectory, recovery is blue & drive is red; *(2)* Instantaneous speed of the handle trajectory shown with corresponding side-projection; *(3)* Projection of handle trajectory into X-Z plane; *(4)* Demonstration of calculating the signal for shape smoothness.

Each feature's values are normalised to roughly between 0 and 1 and highly negatively correlated features are negated. Good and bad strokes are labelled as 1 and 0 respectively by an expert coach using conventional video of the performances. The system scores each strokes, s, using a biased, weighted linear combination of the features:

$$l_m^{\,s} = w_0 + \sum^{f \varepsilon F} (\varphi_f^{\,s} . w_f)$$

The weights are learnt from a training set of example strokes by 2 methods: The Moore-Penrose pseudo-inverse of the feature matrix $\boldsymbol{\varphi}$ gives an exact solution to the equation:

$$\begin{pmatrix} 1 \ \varphi_0^{\,0} \dots \varphi_{20}^{\,0} \\ . \quad . \quad . \\ 1 \ \varphi_0^{\,S} \dots \varphi_{20}^{\,S} \end{pmatrix} \begin{pmatrix} w_0 \\ w_1 \\ . \\ w_{20} \end{pmatrix} = \begin{pmatrix} l_e^{\,0} \\ . \\ l_e^{\,S} \end{pmatrix}$$

Gradient descent is used up to 750 iterations to minimise the error function: $Err = \sum^{s \varepsilon S} (l_e^{\,s} - l_m^{\,s})^2$

Sensitivity analysis is used to find the optimum feature set and number of iterations (when applicable) during cross-validation of the models using all the strokes available for the specific experiment.

Results

The system was validated using real-world coaching scenarios. Six novice, male rowers in their mid-twenties, between 60kg and 90kg and with very little or no rowing experience were used. They were not initially fatigued and rowed at a comfortable rate in an uncontrived manner.

Each rower was given a basic explanation of how to row and gave an initial performance. A coaching process was repeated until fatigue:

1. Coach evaluates last performance and teaches the rower how to improve an aspect.
2. Rower gives another performance and coach helps them to maintain the improved technique for the increasing number of corrected aspects.

Each performance lasted for approximately 30 strokes and the corrected aspects were maintained for at least 95% of the performance.

Coaches can concentrate on correcting a single aspect for an individual, which can be investigated using 2 consecutive performances. Coaches also often choose to draw an athlete's attention to a floor in an arbitrary aspect, no matter how they are performing other aspects. The system's accuracy in this situation is evaluated using all an individual's performances, giving a more extensive population of strokes that may be harder to generalise over (see Fig.3.) Being able to evaluate unseen rowers requires training across multiple athletes. Fig. 4. shows results when 1 good and 1 bad performance were used from multiple rowers.

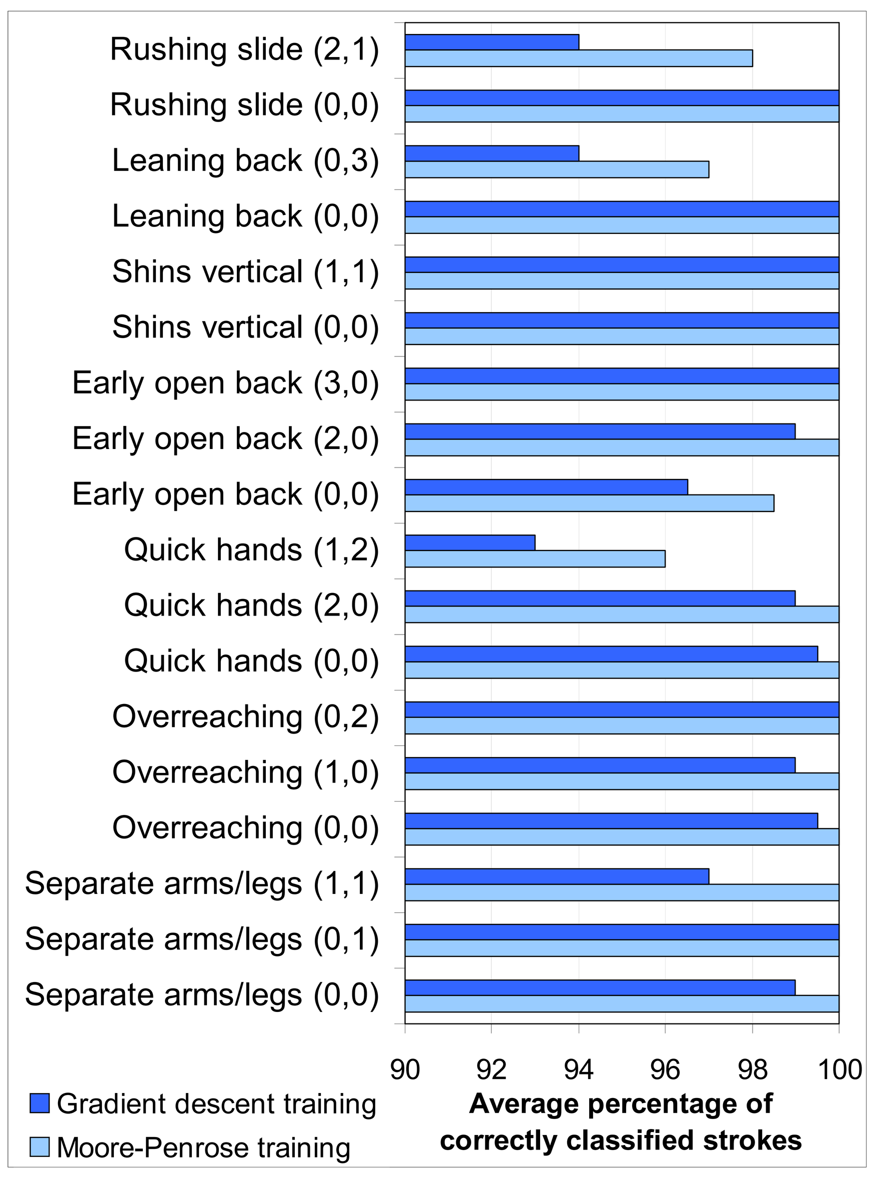


Fig. 3. Classification accuracy for individual performances, of quality of individual aspects of technique, showing number of other aspects changing when judged aspect is (bad, good)

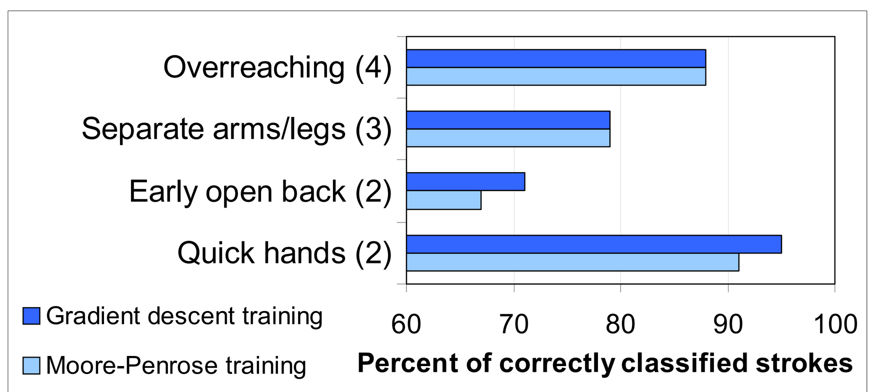


Fig. 4. Classification accuracy across given number of performers, for quality of individual aspects of technique.

Summary

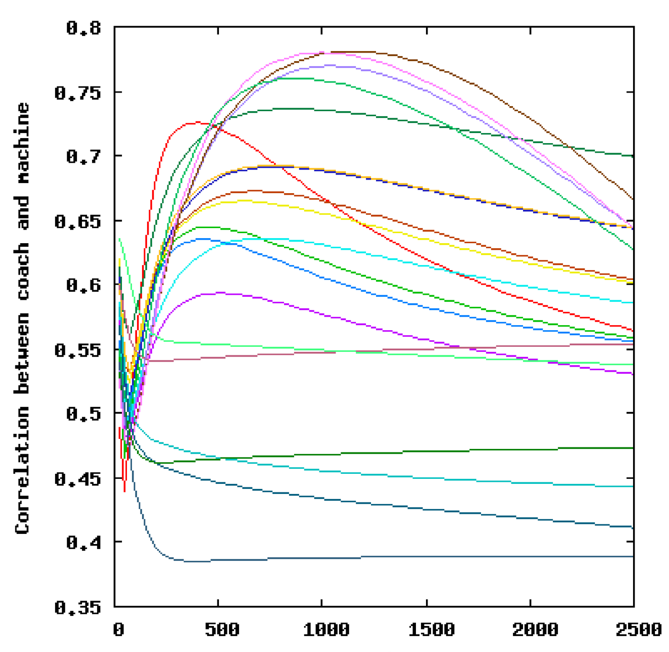


Fig. 5. The exact solutions for the weights can result in over-fitting the model for some feature sets.

Shape and speed moments λ^{02}, λ^{20} μ^{02} and μ^{20} were the only four used in at least 90% of the final feature sets for both algorithms and may capture fundamental characteristics of sporting technique.

For individual performers, encouragingly low misclassification is achieved and gradient descent takes too long to find the optimal, exact solution. Across different performers the misclassification is a little worse, but there is large inter-variation between athletes. It is relatively low for larger training sets. Gradient descent gives better results than finding the exact solution as it ensures no over-fitting of the models occurs (see Fig. 5.)

Work is continuing by developing a system to collect a much larger amount of data for characterising the process further.

Simon Fothergill, Robert Harle, Sean Holden

UNIVERSITY OF CAMBRIDGE

Computer Laboratory

S+SSPR 2008, 4-6 December, Orlando, Florida, U.S.A.

Acknowledgements to Professor Andy Hopper and EPSRC, UK