International Journal of Computer Science in Sport

Volume 14/2015/Edition 1

TABLE OF CONTENTS

Arnold Baca
Editorial 3

RESEARCH PAPERS

Gravenhorst, F., Muaremi, A., Draper, C., Galloway, M. & Tröster,G.
Identifying Unique Biomechanical Fingerprints for Rowers and Correlations with Boat Speed – A Data-driven Approach for Rowing Performance Analysis… 4

Quintana-Duque, J.-C., Dahmen, T. & Saupe, D.
Estimation of Torque Variation from Pedal Motion in Cycling……..... 34

Tamaki, S. & Saito, H.
Reconstructing the 3D Trajectory of a Ball with Unsynchronized Cameras…….. 51

PROJECT REPORTS

Fernando, T. & Pinidiyaarachchi, U. A. J.
A hybrid algorithm for player arm biomechanics evaluation in outdoor sporting activities…..……..……..... 69

Hirotsu, N., Osawa, K. & Miyaji, C.
Calculation of Probability of Winning and Number of Games Played for Various Tournament Formats of the World Baseball Classic…….. 87

Editorial

Arnold Baca

Department of Biomechanics, Kinesiology and Applied Computer Science,

ZSU, University of Vienna

Dear readers:

Welcome to the summer 2015 issue of the **International Journal of Computer Science in Sport (IJCSS).**

The issue contains three research papers and two project reports.

Gravenhorst et al. introduce new performance indicators to quantify rowing technique. Subsequently, the authors provide two data-driven approaches to identify those indicators which make individual rower's technique unique and those indicators which are correlated with the boat's speed most.

Quintana et al. developed a method to compute the variations in the torque applied to the pedals during crank rotation in cycling.

Tamaki and **Saito** propose a method that reconstructs the 3D trajectory of a ball from unsynchronized cameras. The method consists of ball detection, camera calibration, and trajectory reconstruction.

In their project report **Fernando** and **Pinidiyaarachchi** provide an algorithm to model player arm movements in outdoor sporting activities. The algorithm uses a trained cascade object classifier, an optical flow algorithm and an Active Shape Model.

The project report by **Hirotsu et al.** investigates different tournament formats of the World Baseball Classic from the viewpoint of the probability of winning the tournament and the probability distribution of the number of games played by the same teams.

If you have any questions, comments, suggestions and points of criticism, please send them to me.

Arnold Baca, Editor in Chief
University of Vienna, arnold.baca@univie.ac.at

Identifying Unique Biomechanical Fingerprints for Rowers and Correlations with Boat Speed – A Data-driven Approach for Rowing Performance Analysis

Gravenhorst, F.[1], Muaremi, A.[1], Draper, C.[2], Galloway, M.[2], Tröster, G.[1]

[1]Wearable Computing Lab., Institute of Electronics, ETH Zurich, Switzerland

[2]Movement Science Department, Australian Institute of Sport, Canberra, Australia

Abstract

Finding the best fit of rowers for a crew boat is a challenging task. Each rower has a unique technique and the ability to adapt this to a crew varies from person to person. Currently, subjective evaluations and qualitative measures are the main methods used to try to put the fastest crew together. To make the process more accurate and objective we introduce 177 performance metrics to quantify some of the measureable aspects of rowing technique. We then present two data-driven approaches to select the most relevant features that 1) make individual rower's technique unique and 2) correlate most strongly to boat speed. The first approach uses sequential forward feature selection to identify the features that are most discriminative for individual rowers in crew boats. These features make the unique biomechanical fingerprint of each rower. We recorded a dataset with four world-class female rowers racing in double sculls in different crew combinations. We identified the "Finish Slip" as the most discriminative feature. A rower identification classifier based solely on this feature scored an accuracy of 74.6%. Applying one or two additional features this accuracy improved to 90.7% or 95.6% respectively. In a second approach we proposed linear regression analysis to identify the features that most strongly correlate to boat speed. For the given dataset, a subset of five performance metrics proved sufficient to build a linear model that predicts the boat speed with a root mean square error of less than 0.087 m/s.

KEYWORDS: ROWING, PERFORMANCE ANALYSIS, CLASSIFICATION, MACHINE LEARNING

Introduction

Motivation

Rowing is one of the oldest Olympic disciplines. To perform at the top level rowers must be physically strong and have excellent technique. Strength can be measured in off water setups with rowing ergometers and individual technical skills are often assessed on the water through competitions in small boats. In the case of sculling (two oars per person) tests are usually carried out in single sculls (a one person rowing boat). Different athletes often have different ideal single sculling technique as they have different skill levels, body proportions and

anatomy (Altenburg et al., 2008). As such even athletes who perform at the top level are likely to have slightly different technique from one another.

In the Olympic regatta rowers compete not only in singles but also in crew boats of up to eight rowers. Apart from requiring a good strength and endurance base, rowers in crew boats must be able to synchronize their technique and timing with each other in order to achieve top results (Christov et al., 1988; Hill, 2002; Wing & Woodburn, 1995). It is challenging to determine who the strongest and simultaneously most compatible athletes for a crew boat are.

Results show that crews made up of the best single scullers are often beaten by crews made up of rowers with worse individual performances. As Daniel Topolski, one of the most successful coaches of the annual Oxford-Cambridge boat race stated, "[t]he sum of a crew is greater than its parts" (Robinson & Topolski, 2013). This has become a well-established saying and highlights that making a crew boat successful requires more than a group of individually good rowers.

The most common approach for finding ideal rowing crews within a pool of athletes is through test races. Coaches put crews together and organize races to determine which crew in which seating order performs best. It is usually not possible or practical to test all possible combinations due to time restrictions and the difficulty of ensuring comparable conditions between so many test races. Key challenges include changing weather conditions, differences in athletes' required recovery time and ensuring that athletes perform to their best in each race.

Instead of testing all possible combinations, coaches currently decide which crews to test based on subjective evaluations and personal experience. The success of such an approach is highly based on the coach's experience. This process is intransparent and often leaves the non-selected athletes with unanswered questions and a feeling that they were potentially unfairly overlooked.

A more systematic selection process would be to measure the biomechanical parameters of rowers and compute the features that describe their technique in order to determine which ones fit best together in terms of technique style and synchronicity. There are many possible metrics that could be extracted and some are already measured in leading high-performance rowing centers. One of the main challenges is to identify the most important of these available features. They should be descriptive and of key significance for a rower's technique, meaning they remain different between individuals even when put together in a crew. These features should also be relevant for crew boats' performances, meaning they should correlate with the boat speed.

Rowing Basics

The goal in rowing is to move the boat as fast as possible from start to finish, usually over the Olympic distance of 2000m. The boat travels backwards with the rower's back to the direction of movement. There are two sub-types of rowing: sweep-oar, in which each rower holds one oar and rotates either to the left or right side; and sculling, in which each rower holds two oars making symmetric movements with the left and right oar. The rower sits on a sliding seat allowing the body to move forwards and backwards and enabling the rower to further extend the stroke length. The rowing movement is cyclic and consists of two phases, the drive phase and the recovery phase. For the sake of simplicity this work focuses on sculling, however most of the methods can be applied to sweep rowing as well.

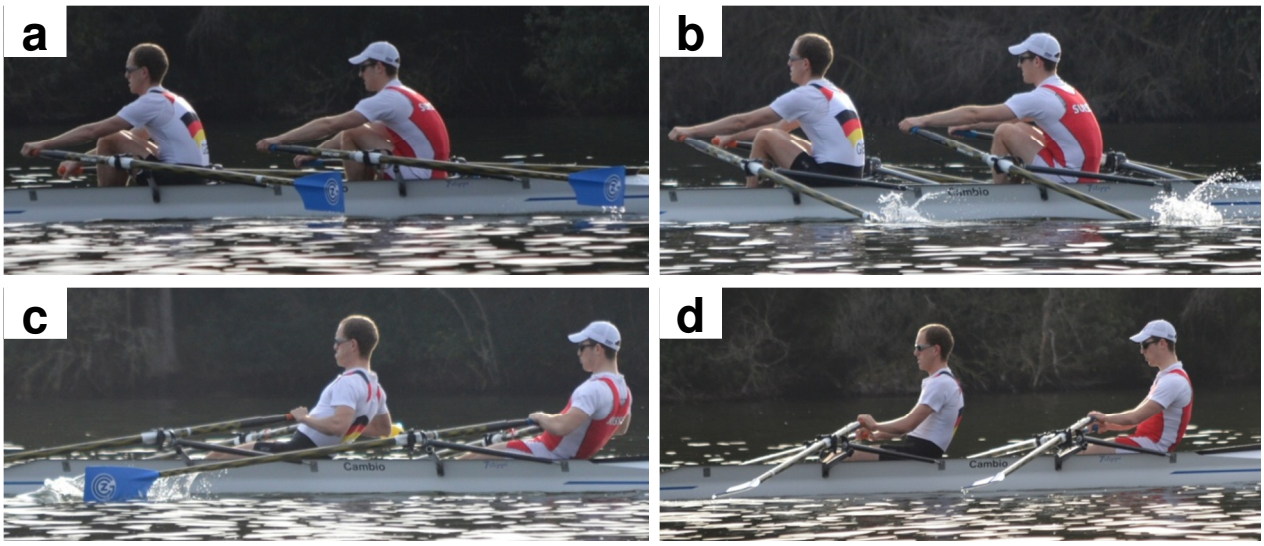


Figure 1. Basic rowing stroke: The boat is moved backwards (from left to right). At the catch position (a) the blades are placed into the water (b). The boat is accelerated by pushing back with the legs and bending the arms until the rower reaches the finish position (c). The blades are extracted from the water and feathered (turned parallel to the water). Then the rower reaches forward, bending their legs and extending the arms to prepare for the next catch position (a). Then the cyclic movement continues with the next stroke.

The drive phase starts in the forward most position, called the catch position (Figure 1a). The legs are bent, so that the shins are perpendicular to the water, the sliding seat is as close to the stern of the boat as possible, the upper body and shoulders are in front of the hips, and the arms and hands are fully extended, reaching out for maximal length. The blades are then placed into the water (Figure 1b). They are then driven through the water and the boat is accelerated by pushing back with the legs, moving the seat towards the bow, bending the arms and taking the upper body back so that the shoulders are just behind the hips, while the back remains relatively straight. At the back most position, called the finish position (Figure 1c), the blades are extracted from the water and the recovery phase begins. During this phase, the rower prepares for the next stroke, moving the blade above the water to the catch position again. To minimize air drag and to increase boat stability, the blade is turned and moved into a feathered position above the water (Figure 1d). Finally, the blade has to be turned again so that it is at a right angle to the water, this is known as a squared blade. The cycle then starts again with the next drive phase (Figure 1a).

This stroke is repeated over and over again. A standard base training rate is 18-20 strokes per minute, while in 2km races it is generally between 32-37 strokes per minute. During the start, finish sprint and other strategic points in a race, the stroke rate can reach up to 43 strokes per minute.

The more force a rower applies the faster the boat is accelerated. However, muscle mass increases a rower's weight, and the heavier a boat is the greater the drag factor. Therefore, the potential gain of boat speed through strength is limited and this is where rowing technique becomes essential. With more efficient technique the rower can manage to increase the boat speed with constant strength (and body mass). Not only does good technique maximize acceleration, it also minimizes deceleration (FISA, 2011). Deceleration of the boat occurs mainly due to the forces applied to the footstretcher as the rower comes towards the catch. The

deceleration is also caused by unstable boat movements, unsmooth oar movements and delays in placing the blade at the catch position.

In crew boats the ideal rowing technique is even more complex, since the individual rowers' techniques have to fit with the other crew members'. According to Soper et al. this "ideal fitting" includes several aspects and similar force profile characteristics are the most important ones for successful scullers in crew boats (Soper & Hume, 2004). Our work on the biomechanical rower fingerprint is based on the assumption that rowers in crew boats should move as synchronously together as possible. For the case of sculling, this is supported by many studies (Christov et al., 1988; Hill, 2002; Wing & Woodburn, 1995). However, for sweep rowing some authors argue that slightly opposite styles could also complement each other in a positive way (Fahrig & Witte, 2007; Smith & Draper, 2002).

Related Work

Rowers are aiming to improve their technique to avoid injuries and to improve boat speed. In most cases, they rely on human coaches who accompany trainings and give feedback. Besides traditional tools such as stop watches and high speed video cameras, there are multiple sensor-based approaches proposed and some are already available on the market. An overview of sensors for instrumented rowing boats is presented in Table 1.

Table 1. Sensors used for rowing in on-water setups, adapted from (Tessendorf et al., 2011)

Sensor	Location	Description
Magnet	Seat	Reed switch to count strokes (Nielsen-Kellermann, 2014)
Impeller	Boat	Measure distance and speed relative to the water (Kleshnev, 2010; Smith & Hopkins, 2012)
GPS	Boat	Measure distance and speed relative to the shore (Smith & Hopkins, 2012)
Accelerometer	Boat	Measure stroke rate and interpolate boat movement (Groh et al., 2014)
Gyroscope	Oars	Measure oar angles (Sabatini & Genovese, 2006)
	Boat	Measure boat stability (Gravenhorst et al., 2011; Wagner et al., 1993)
Potentiometer	Oarlock	Measure horizontal oar angle (Fritsch, 2005)
Force sensor	Oarlocks	Measure force applied to oars (Sinclair et al., 2009)
	Foot stretcher	Measure force applied on the foot stretcher (Smith & Loschner, 2002a)
Strain gauge	Oars	Measure bending force of oars (Nozaki et al., 1993)
Inertial measurement units	Oars, boat	Measure oar and boat orientations and movements (Gravenhorst et al., 2014a; Tessendorf et al., 2011)

Experienced rowing biomechanists are required to interpret the acquired data and generate benefits from it. Finding methods to automatically analyse the data and provide appropriate

feedback for the coaches and athletes is an active field in research (Baca & Kornfeind, 2006; Gravenhorst et al., 2014; Smith & Loschner, 2002a; Tessendorf et al., 2011).

Although rowers have their own styles (Claessens et al., 2002; Claessens et al., 2005; Smith & Spinks, 1995; Tessendorf et al., 2011) coaches expect them to adjust their technique to achieve a common, efficient and synchronous stroke when placed in crew boats. However, as Korndle et al. show, in practice rowers only manage to adapt some aspects of their technique for the crew. Regardless of the crew or boat class rowers are put in, they maintain their individual "signature" force angle profile, (Korndle & Lippens, 1988). Through a more systematic analysis of a group of female rowers Galloway et al. were able to support this finding. They also found that some rowers' movement patterns were more dominant than others in influencing the total boat speed (Galloway & Draper, 2011).

Other studies explore the relative influence of rowers' fitness, strength, physiological constitution, body measures, technique and boat and oar settings on boat speed (Cosgrove et al., 1999; Hill et al., 2003; Kaya et al., 1995; Kleshnev, 2002; Loschner & Smith, 1999; Smith & Loschner, 2002a; Smith & Loschner, 2002b).

Millward developed a model considering the fluid mechanics of the oar and verified it with rowing performance data (Millward, 1987). He identified the shape of the rowing force curve and the proportion of recovery time in the total stroke as important factors for the boat speed. Sanderson and Martindale developed an equation to describe the boat speed as a function of the movement of the rower's center of mass and the applied force (Sanderson & Martindale, 1986). To increase the efficiency, the authors suggest building lighter boats, increase the blade area of the rowing oar and finding an ideal stroke rate depending on the rower's body mass.

Medical and physiological determinants for boat speed are explored in clinical studies. Baguet et al. found that supplementation of β-alanine is highly effective in increasing the performance of elite rowers. After a 7-week study his control group improved an average of 4.3 seconds more than the placebo group (Baguet et al., 2010). Ingham et al. measured the oxygen intake during ergometer rowing. The best correlation to the resulting performance was the applied power during maximum oxygen consumption (Ingham et al., 2002).

Secher and Vaage present a mathematical model for forecasting the racing times of male and female rowers depending on body mass. They found that heavyweight rowers had a 2.6% advantage in comparison to lightweight rowers. This value was supported by on-water results (Secher & Vaage, 1983).

To sum up, most of the published approaches that present dependencies and influences of input factors on the boat speed are based on biomechanical models, trying to cover the causal dependencies and interactions as accurately as possible. These deductive approaches are advantageous in the sense that models can be built based on them and validated using theoretical knowledge and simulations. They can be generalized from because assumptions and limitations are usually known. The main disadvantage is the complexity to which these models can grow if they try to represent reality as completely as possible. These models usually require dozens of input parameters and some of them can hardly be measured or estimated.

Data-driven approaches like regression models have the potential to address this drawback. They are built in an inductive way and based on the input parameters that can be measured in real-life settings. For example, Perl & Baca introduced the application of neural networks to performance analysis in rowing and managed to identify instabilities in the movement patterns (Perl & Baca, 2003). The main drawback of these inductive methods compared to approaches

based on physical models is the missing proof of causality. In fact, they identify correlations solely based on available data, this can be a sign for causal dependencies but not a proof. More data and domain-specific knowledge is needed to confirm dependencies, identify limitations and to explore how the model can be generalized, for example for other boat classes.

Contribution

More and more sensors and measurement systems that can be used to obtain quantitative data of rower's performances on the water are becoming available. To make this available data beneficial for rowers and coaches, this work mainly focuses on the post-processing part and introduces methods for meaningful data analysis.

We extend the state of the art in the following respects:

Performance Metrics. To describe the rowing technique quantitatively, we introduce 37 boat-specific, 28 oar-specific and 112 crew-specific features. These are the base of our data-driven performance analysis.

Biomechanical Fingerprints. We propose and compare three different machine-learning methods to identify the most discriminative features for a group of rowers. We demonstrate how this approach can support coaches towards a more systematic approach for finding the best-fitting team for a crew boat.

Boat Speed Model. We suggest a linear model to describe the dependency between boat- and oar-specific performance metrics as input and the boat speed as output parameter. The number of necessary coefficients and their values are identified with a step-wise linear regression analysis.

Proof of Concept. We carry out an experiment with four elite rowers performing races in different crew combinations within a time period of six days. Data has been recorded with commercial sensor systems. In a post-processing step, performance metrics were calculated, and our proposed methods are applied and results discussed.

Paper Organization

The second section explains how meaningful performance metrics can be extracted from raw sensor data and introduces oar-, boat- and crew-specific metrics. The third section, "Experiment Setup" introduces the measurement system we used to instrument the rowing boats and the experiment design applied to collect data from different crew combinations. The following two sections explain our methods and the results we obtained using the data from our experiment. The section "Analysis and Discussion" analyses the results and outlines possible interpretations of them. Limitations are discussed in the next section and the final section draws conclusions and provides an outlook.

Performance Metrics for Individual Rowers and Crews

As outlined in the introduction, the general idea of good rowing technique is well-established (FISA, 2011). For most technical aspects of rowing, there are rough guidelines rather than exact quantifiable and measureable targets. However, to enable comparison of rowers and crew combinations through data-driven approaches we need quantitative measures, which can be obtained with mobile on-boat sensor systems. It is important that the selected set of measures contain as little redundancy as possible. In the first of the following two sub-sections we introduce performance metrics that can be applied to all boat types. These ones include:

- Boat-specific features: These features describe the movements of the boat. These features have identical values for all crew members and quantify the effects of the crew as a whole.
- Oar-specific features: These features describe the movement of an individual rower's oar. In crew boats these features usually have different values for each person.

In the second sub-section we present additional features that are only available for crew boats. They quantify the crew synchronicity based on the oar-specific features of all crew members. All features are calculated once per stroke.

Performance Metrics for all Boat Types

In collaboration with Olympic-level rowing coaches and leading rowing biomechanists we define 37 boat-specific features and 28 oar-specific features.

Boat-Specific Features

Boat-specific features mainly describe the boat behavior such as its acceleration, speed and instability. Boat instability is quantified by deviations of the boat orientation in three dimensions. The corresponding angles are visualized in Figure 2. The complete list of the proposed boat-specific features can be found in Table 7 (Appendix).

Oar-Specific Features

Oar-specific features describe the movement of the rowing oar, such as the amplitude of movement, the timing and the force applied. For the sake of simplicity this work, which focuses on sculling, only evaluates the rowers' bowside oars. The definitions of the most relevant features are shown in Figure 3. Similarly to the boat-specific features, the oar-specific features are also stroke-based, with one set of features describing one rowing stroke. They are calculated separately for each crew-member. The complete list of proposed oar-specific features can be found in Table 8 (Appendix).

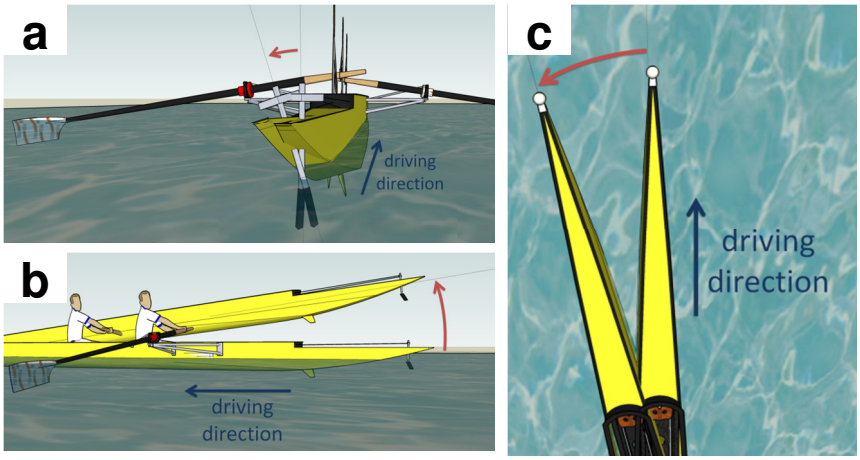


Figure 2. Definition of boat orientation angles used to describe the boat instabilities in three dimensions: (a) roll angle, (b) pitch angle and (c) yaw angle.

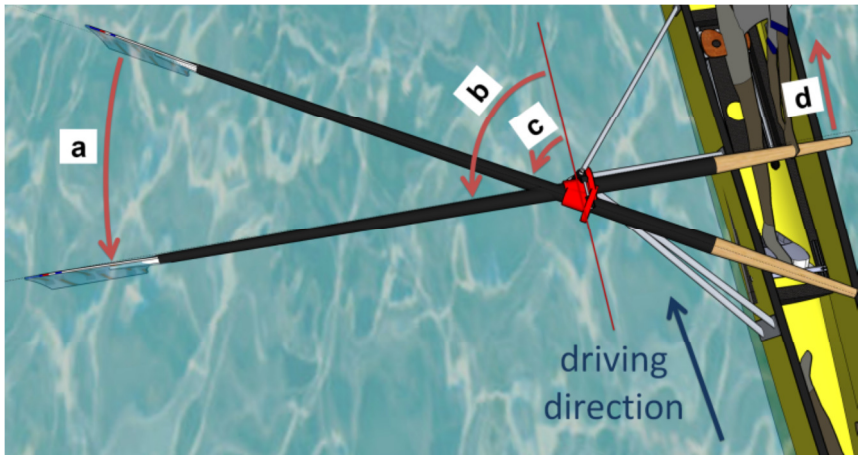


Figure 3. Definitions of most important oar-specific features describing the oar movement: (c) describes the angle at the start of the rowing stroke, the catch position, (b) describes the angle at the end of the rowing stroke, the finish position, (a) describes the swept angle during the stroke, the geometric stroke length, (d) indicates the direction in which the propulsive handle force is measured.

Additional Performance Metrics for Crew Boats: Synchronicity Measures

The first step computes the oar- and boat-specific features introduced in the previous sub-section. In the second step we add additional features that describe the crew's interaction and synchronicity. For a crew boat with N rowers and k oar-specific features per rower, we define (N+2)k crew-specific features. The full set of proposed crew-specific features is described in Table 2.

Table 2. Descriptions of crew-specific features. These features describe the crew interaction and synchronicity in crew boats.

Names of features	**Description**	**Number of features**
Diff	The difference of each rower's features compared to the rower in the stroke position is calculated	(N-1)k
Min	The minimum value of each oar-specific feature within the crew is calculated	k
Max	The maximum value of each oar-specific feature within the crew is calculated	k
Mean	The mean value of all oar-specific features within the crew is calculated	k

Experiment Setup

Mobile On-Boat Measurement System

The minimum measurement setup to compute all proposed performance metrics consists of angle and force sensors on each rowing oar or gate as well as a GPS, accelerometer and gyroscope measurement module mounted on the boat. We decided to combine two

commercially available systems:

- Boat-mounted sensor: A MinimaxX module (Catapult Sports, Australia) is attached to the boat's stern. It measures GPS, 3-axes accelerometer and 3-axes gyroscope data with a sample frequency of 100Hz.
- Oarlock integrated sensor: The PowerLine Rowing Instrumentation system (Peach Innovations Ltd., United Kingdom) is an instrumented oarlock that enables force and angle measurements with a sampling frequency of 50Hz.

The overall setup is depicted in Figure 4. Both devices work independently and save the data to internal memory. After the experiment both memories are read out and a semi-automated method is used to synchronize and merge the data.

Figure 4. The overall measurement setup for a double scull rowing boat is shown in (a), the boat sensor is attached to the stern of the boat (1), the two rowing oar sensors are integrated into the oarlocks (2). Close-ups of the boat module and oarlock module are shown in (b) and (c).

Test Races for Data Collection

For the data recording we recruited four Olympic-level female rowers (A, B, C and D) and equipped two double sculls with the described mobile measurement system. During a 6-day rowing camp, the athletes performed a race over 2000m every other day. The days in between were used for training and recovery (Table 3). This way, all six possible crew combinations were tried out and in total 1459 rowing strokes were recorded during the races.

Table 3. During the 6-day data collection all double combination of the four rowers (A, B, C, D) were measured in race conditions

	Day 1	**Day 2**	**Day 3**	**Day 4**	**Day 5**	**Day 6**
Boat 1	Recovery and training	B, A	Recovery and training	B, D	Recovery and training	C, B
Boat 2		C, D		C, A		D, A

Feature Calculation

The complete feature calculation processing chain is depicted in Figure 5. The recorded continuous data was segmented into strokes, using a peak-detection algorithm on the boat acceleration data. For each of the 1459 recorded rowing stroke cycles, we computed 37 boat-specific features and 56 oar-specific features (28 for each rower), totaling 93 unfiltered features per stroke.

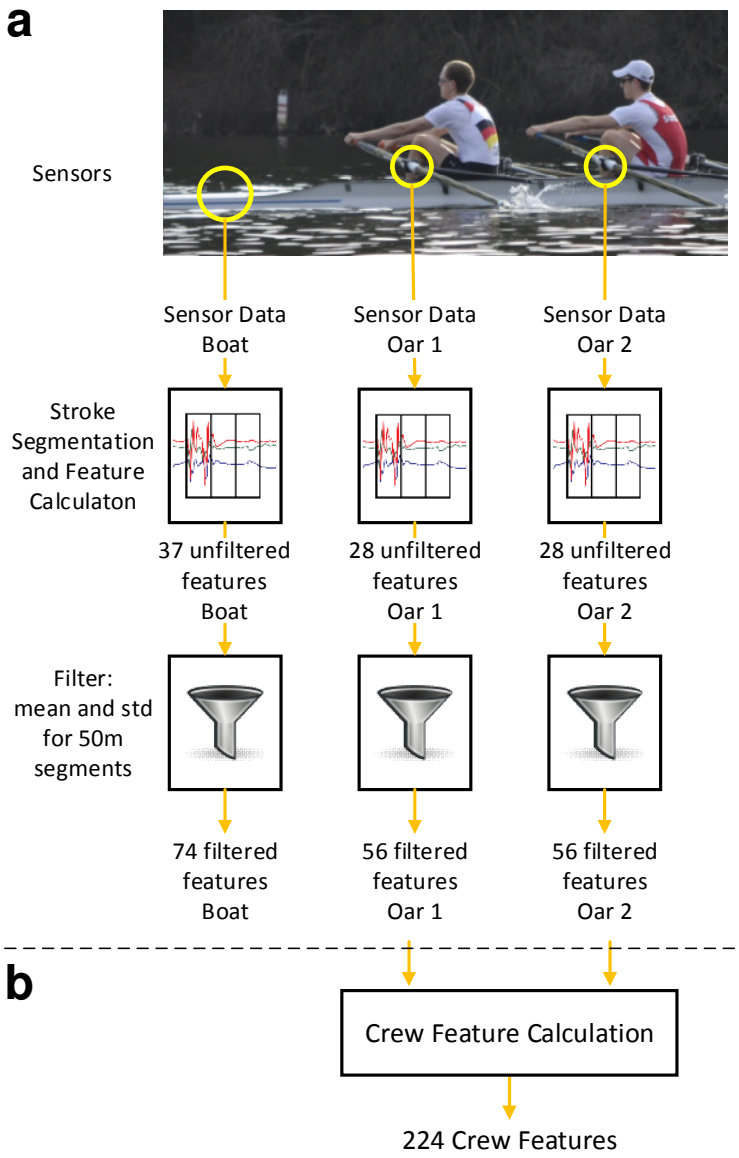


Figure 5. Feature calculation processing chain for double scull. (a) describes the calculation of the filtered boat and oar features, (b) describes the continuation for calculating the crew features.

To filter for noise effects and outliers, we segmented the races in approximately 50-meter intervals and combined strokes within these segments. For each segment we computed the average and the standard deviation for each feature of the strokes within the segment. In total, for all six races, we received 248 race segments, each with 74 filtered boat features and two times 56 filtered oar features (Figure 5a). Based on the 56 filtered oar-specific features (28 averages and 28 standard deviations) for each rower, we computed 224 crew-specific features (Figure 5b). The final database consists of 248 race segments, each with 410 features (74 boat-specific, 56 oar-specific for each rower, 224 crew-specific). In the following analysis, we only consider these filtered features and we use the following naming convention: the original feature names (according to Table 7 and Table 8) refer to the averaged values. The standard deviations are named with the prefix "Std-" following the original features' names. The crew-specific features use the additional prefixes "Diff-", "Min-", "Max-" and "Mean-", as taken

from the definitions in Table 2.

Methods

The data analysis was divided into two parts (Figure 6). In the first part we analyzed which features were the most discriminative for individual rowers even when put together in crew boats. The identified set of features is most suitable for identifying differences in rowing styles, it is the unique biomechanical fingerprint for rowers.

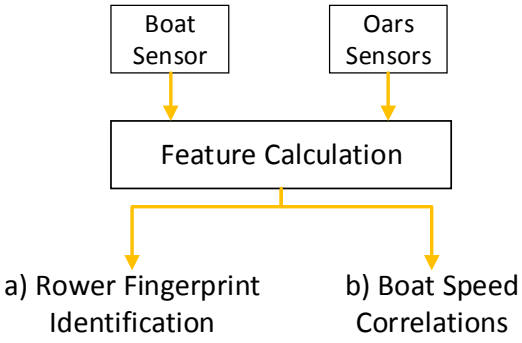


Figure 6. Data-driven support for crew selection consists of two analyses, (a) identifies which features are most discriminative for individual rowers and (b) ranks the features considering their impact on the boat speed.

The second part of the data analysis identified which biomechanical features of crew boats correlated most strongly with boat speed and therefore require particular attention when evaluating how well a crew fits together.

Biomechanical Fingerprint Identification: Wrapper-Based Feature Selection

Problem description and requirements

As described in the previous section, we generated 74 boat-specific and 56 oar-specific features. In this section we want to identify which features out of these 130 proposed features make up the rower's biomechanical fingerprint. We define the following requirements these fingerprint features should fulfil:

1) Uniqueness: The selected feature subset should be most discriminative for each rower. This means by knowing these features, the rower can be identified. The selected features are the ones in which any two rowers of our dataset are most different from each other.
2) Constant: The selected features do not depend on the crew partner. For each rower, the values of the selected features stay within a specific and individual range, even when put together with other rowers.

Our approach

The problem of finding the most discriminative features out of a given pool of features is well-known in the machine learning community. The overview of our approach is depicted in Figure 7. The input data are the instances for the 130 considered features. Since each of the 248 race segments contains two sets of oar-specific features, one for each of the two athletes in the boat, there are in total 496 data instances available. The first step of our iterative process

(Figure 7) is to generate a subset of features according to a search strategy. Then, this subset is evaluated according to an optimization criterion. These two steps are repeated over and over again until a stop criterion is achieved and the feature subset with the best evaluation value is outputted. The stop criterion can be a certain threshold of the evaluation value or the end of the implemented search strategy. In our work, the latter is the case.

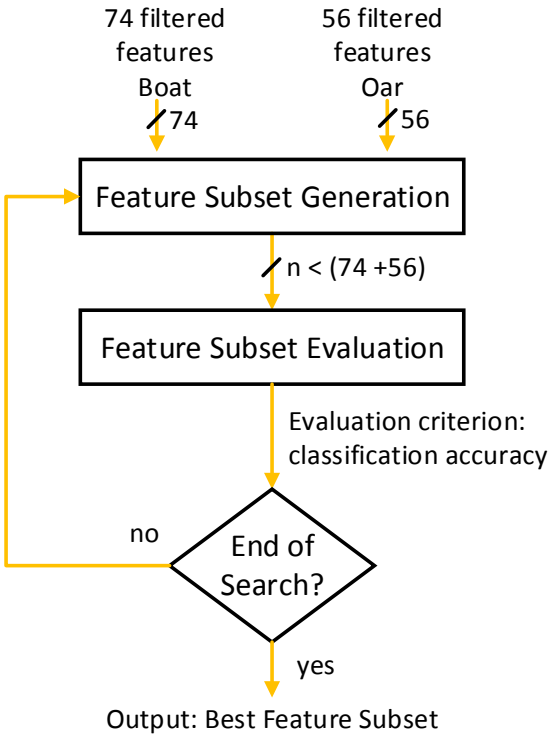


Figure 7. Feature selection procedure. Subsets of features are generated and evaluated. The subset with the best evaluation value is the final output. Adapted from (Qin et al., 2009).

Feature subset generation. A full test of all possible subsets of features would require testing $2^{130} \approx 10^{39}$ possibilities. To reduce the number of required iterations, we implemented sequential forward feature selection as search strategy. In a first step, all possible subsets containing only one single feature are evaluated. The second step considers the best subset from the first step and evaluates all possibilities of extending this subset with a second feature. The extended subset with the best evaluation score makes it through to the third step and so on. In each step, the winning subset of the previous step is extended by one additional feature. This search strategy ends as soon as a predefined threshold accuracy is reached, a predefined number of features is selected or after all features are selected. In our case the maximal number of tested subsets is $131 \cdot \frac{130}{2} = 8515$. Compared to the full search, this search strategy saves computational resources while proving good results in many applications (Jain & Zongker, 1997), however it does not assure to find the theoretically best possible result.

Feature subset evaluation. To evaluate a given subset of features there are filter and wrapper approaches. Filter methods are based on statistical characteristics; they evaluate the given subset for example according to its dependencies, relevance or redundancy (Peng et al., 2005). Wrapper methods evaluate the given data according to the accuracy a classifier can correctly classify it. Wrapper methods are usually computationally more expensive than filter approaches. Feature subsets selected by wrappers are optimized for the specific classifier used

during the selection process. In this work we used wrappers, because they usually outperform filters when it comes to prediction accuracy using the classifiers they are optimized for (Zhu et al., 2007). A schematic overview of our wrapper-based feature subset evaluation is shown in Figure 8. We used 3-fold cross-validation, which means the evaluation is done in three parallel processes. Each process uses two thirds of the 496 input data sets to train the classifier (i.e. build the model), the other third is used to test the model and calculate the classification accuracy. The mean value of all three accuracies is the evaluation output for the given feature subset.

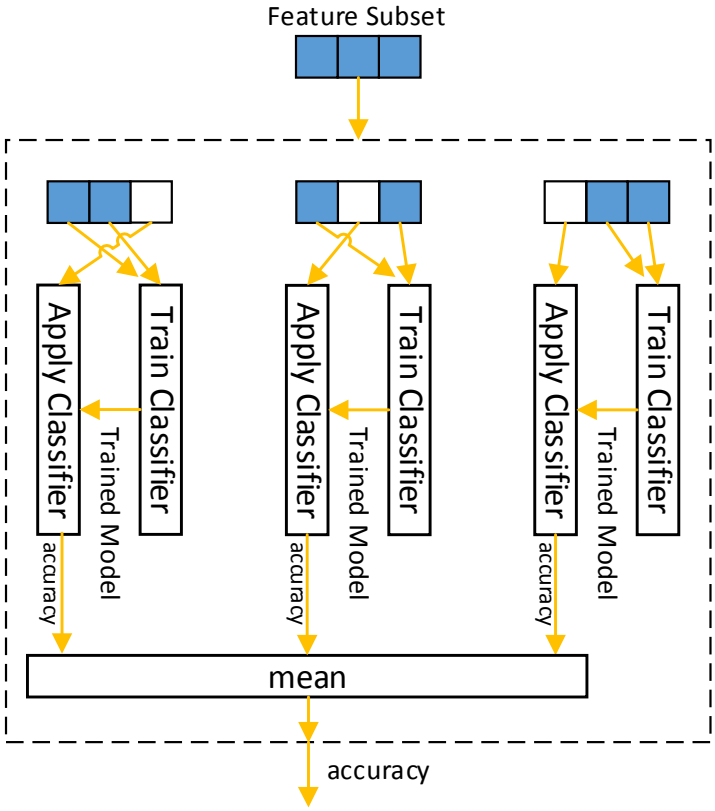


Figure 8. Feature subset evaluation with classifier (wrapper approach) and three-fold cross-validation.

Classifiers. In general, classifiers are used to identify categories based on the input data. In our case, the input data is the subset of feature values which was calculated for the race segments. The 'category' which has to be inferred from this input data is the athlete ID. Working with classifiers involves two steps: Building the model and applying it. The first step, also known as training phase, considers input datasets as training data and outputs the model. The second step applies the model to the test datasets. For each test dataset, the model outputs a predicted athlete ID, which is compared to the actual athlete ID. The accuracy is calculated as the share of correctly predicted athletes:

$$accuracy = \frac{\#correctly\ predicted\ athlete\ IDs}{\#total\ number\ of\ predictions} \cdot 100\%$$

In this work we use three state of the art classifier algorithms:

k-nearest neighbors (kNN): This is a very basic and transparent classifier. For each classification result it is easy to trace back which learning datasets are responsible for the result (Altman, 1992). We chose k=4.

Support Vector Machines (SVM): This is the most popular classification algorithm, scoring best results for most applications. The main disadvantages are the danger of overfitting and the computationally expensive process when building the model (Cortes & Vapnik, 1995). We used SVM with radial basis functions.

Random forest (RF): A high number of decision-trees are generated. The majority vote of these sub-classifiers determine the classifiers output. This way, even complex cluster boundaries can be represented while overfitting is avoided (Montillo, 2009). We used N=300 trees.

Correlations with Boat Speed: Linear Regression Analysis

Problem description and requirements

We want to find out which of the proposed features most strongly correlate with the boat speed. This subset of features should 1) contain as few features as possible and at the same time it should 2) enable the prediction of the boat speed as accurately as possible.

Besides the identification of the most strongly correlated features, we want to determine how strong and in which direction (positively or negatively correlated) the dependencies are.

Our approach

The feature selection and classifier approaches presented in the previous sub-section are designed to predict qualitative and discrete labels such as the athlete IDs. In contrast, the boat speed is a continuous quantitative output, thus we need another approach to model and predict it. We decided to use linear regression analysis because the resulting model defines transparently which features are correlated; it also calculates weight factors to describe the strength and direction. According to Zou et al. these weight factors are more suitable to assess the strength of the relationships in the data than correlation coefficients (Zou et al., 2003). The overall process is depicted in Figure 9 and is described in the following paragraphs.

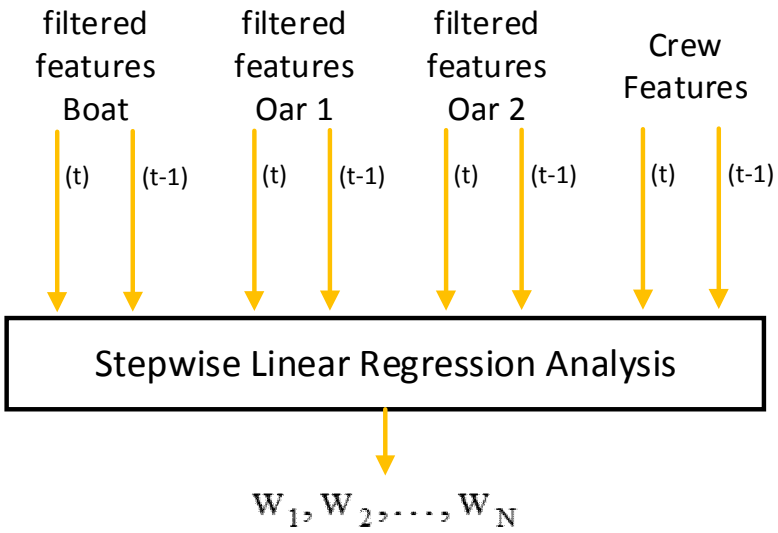


Figure 9. Feature subset selection with linear regression. Input values are the features of the current (t) and the previous stroke (t-1). Output values are the features and corresponding weight factors that are used to describe the boat speed.

We expect that the boat speed of race segment i does not only depend on the feature values of the same race segment, but also on the previous race segment $i-1$. For this reason, the feature values of each previous 50m race segment were used to increase the dimensionality of the input feature space. These features are named with the additional prefix "Last-". On the other hand, we excluded all features which are calculated based on boat velocity (e.g. "Distance per stroke") because their dependency on the boat speed is obvious and therefore not interesting. All features are scaled to ensure values between -1 and 1.

As ground truth value for the boat speed we used the speed output of the boat-mounted sensor. This is calculated based on GPS speed and boat accelerometer data as proposed by Davey et al. (Davey et al., 2010). For each race segment i, the average v_i of this measured velocity was calculated and summarized in vector $\boldsymbol{v} = (v_i)$.

For a given (sub)set of N features, each column of the data matrix $\boldsymbol{X} \in \mathbb{R}^{KxN}$ represent one feature, the rows are the $K = 242$ instances of the features. We assume a linear relationship between feature values and the predicted boat speed $\boldsymbol{v_p}$ (output):

$$\boldsymbol{v_p} = \boldsymbol{X} \cdot \boldsymbol{w}$$

The weight factors $\boldsymbol{w}$ are calculated using the least squares approach (Legendre, 1805; Wolberg, 2006):

$$\boldsymbol{w} = (\boldsymbol{X}^T \cdot \boldsymbol{X})^{-1} \cdot \boldsymbol{X}^T \cdot \boldsymbol{v}$$

This method minimizes the root mean square error (RMSE) between the predicted velocity and the measured velocity (Armstrong & Collopy, 1992):

$$RMSE = K^{-0.5} \cdot \|\boldsymbol{e}\|_2 \quad \text{with} \quad \boldsymbol{e} = \boldsymbol{v} - \boldsymbol{v_p}$$

We wanted to generate a ranking of the features which are most strongly correlated with the boat speed. For this task, we implemented a stepwise linear regression, which increases the number of features N following a sequential forward selection strategy, similar to the

presented wrapper approach described before and depicted in Figure 7. The evaluation criterion is the RMSE.

To ensure the ranking and the calculated weight factors to be as representative as possible, we considered the full set of $K = 242$ instances of the data (no cross-validation) for this feature selection/ranking step. However, for the sake of statistical correctness, the RMSE values which are used for further calculations and are given in the result tables are calculated using 10-fold cross-validation.

With the generated ranking we can determine the minimal number of features N needed to achieve a 'good fit', which our collaborating coaches and biomechanists defined as a RMSE of less than 35% of the measured velocity's standard deviation.

Results

Biomechanical Fingerprint Identification

The five top-ranked features and the corresponding average accuracies (see section 0) are listed in Table 4. These accuracies are also visualized in Figure 10.

Table 4. Rower's biomechanical fingerprint: Top five most discriminative features for identifying individual rowers. The rankings are obtained by sequential forward feature selection with three different classifiers: k-Nearest-Neighbor (kNN), Random Forest (RF) and Support Vector Machines (SVM). The percentages indicate the average achieved accuracy of each classifier when using only the top-ranked feature, the top two features, the top three features, the top four or top five features.

Rank	kNN		RF		SVM	
1	Ang Finish Slip	74.6 %	Ang Finish Slip	67.3 %	Ang Finish Slip	63.5 %
2	Ang Drive Accel Point	90.7 %	Ang Drive Accel Point	89.3 %	Yaw Recov Yaw Range	86.7 %
3	Handle Vel Drive Max	95.6 %	RFD Peak	94.0 %	Ang Drive Accel Point	91.7 %
4	Std-Ang Catch	96.6 %	Handle Dist	96.6 %	RFD Peak	95.6 %
5	Ang Catch	97.4 %	t Max Force	97.2 %	t Drive	96.2 %

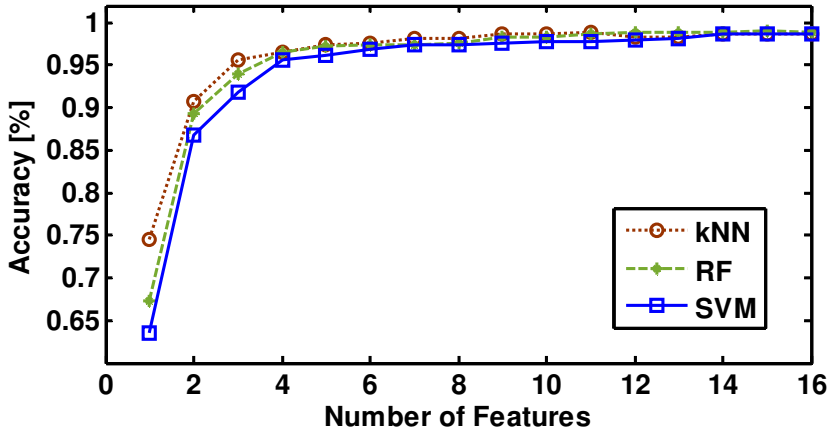


Figure 10. Visualization of how accurately a rower can be identified with a defined biomechanical fingerprint. The more features that are allowed in the fingerprint, the more accurately the rower can be identified. The results are plotted for three different classifiers (kNN, RF, SVM).

The feature "Ang Finish Slip" is ranked as the most discriminative feature in all tested approaches. The feature "Ang Drive Accel Point" is ranked as second or third in all tested approaches. The distributions of the values of these two features and some of their statistical properties are illustrated in Figure 11.

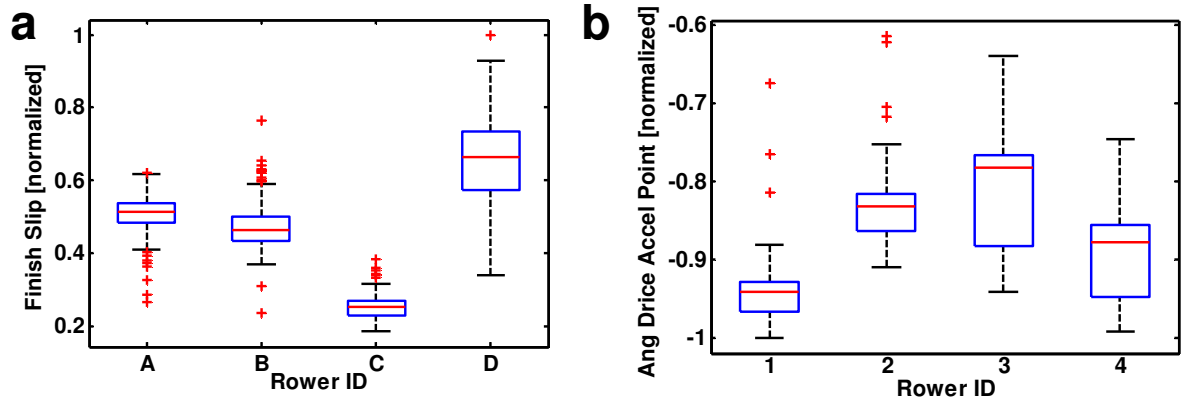


Figure 11. The boxplot (a) shows the statistical parameters of the "Finish Slip", the most discriminative feature for rower identification. The boxplot (b) shows the distribution of the feature "Ang Drive Accel Point" which is ranked second or third, depending on the used classifier.

Correlations with Boat Speed: Linear Regression Analysis

For the total database of $K = 242$ race segments, we received 79 different boat speed values, ranging between 4.37m/s and 5.6m/s. The mean boat speed is 4.75m/s and the standard deviation is 0.25m/s. The distribution of the boat speed values is shown in Figure 12.

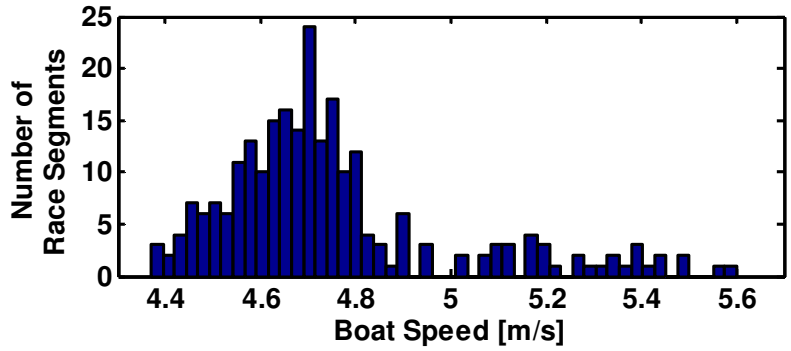


Figure 12. This histogram shows the distribution of boat speed values in our recorded dataset

The dependency between the number of features and the resulting RMSE is visualized in Figure 13. The graphic shows that $N = 5$ features are needed to achieve the defined goal regarding the resulting RMSE. The selected features for this case and the corresponding weight factors are summarized in

Table 5. A visual comparison of the measured and the predicted boat speed based on this model is shown in Figure 14.

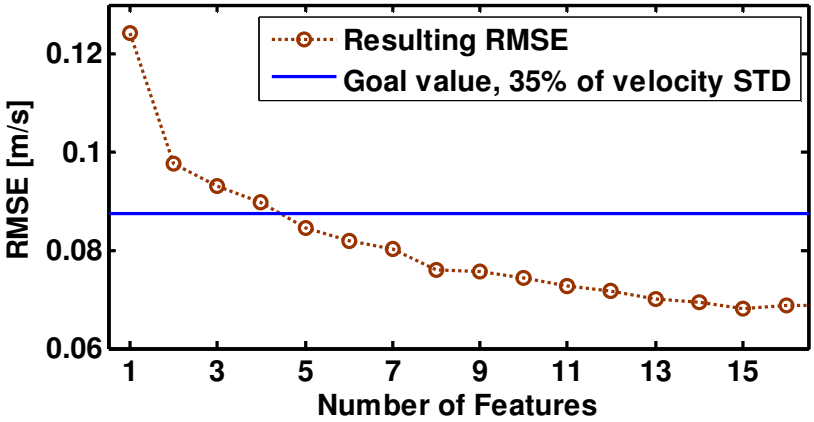


Figure 13. Dependency between number of features used for the linear regression and the resulting root mean square errors (RMSE) between predicted and measured boat speed.

Table 5. Top five features and their corresponding weights for linear regression model for boat speed prediction. The last column indicates the root mean square error (RMSE) value when using only the first, the first two, three, four or all five features in the model.

Rank (i)	**Feature (x_i)**	**Weight (w_i)**	**RMSE**
1	Max-Handle Vel Drive Max: Maximal oar handle velocity during drive phase, maximal value of both rowers during the current 50m race segment	2.79	0.124
2	Last-Min-Handle Vel Recov Ave: Average oar handle velocity during the recovery phase, minimal value of	-1.06	0.098

	both rowers during the previous 50m race segment		
3	Last-Std-Accy Recov Min: Minimal value of transversal boat acceleration during recovery phase, standard deviation of these values during the previous 50m race segment	0.15	0.093
4	Last-Max-Power Drive Ave: Average power applied by the rower to the oar handle in propulsive direction, the maximal value of all strokes during the previous 50m segment from both rowers.	0.56	0.090
5	Last-Mean-Handle Vel Drive Ave: Average oar handle velocity during drive phase, mean value of both rowers during the previous 50m race segment	1.17	0.085

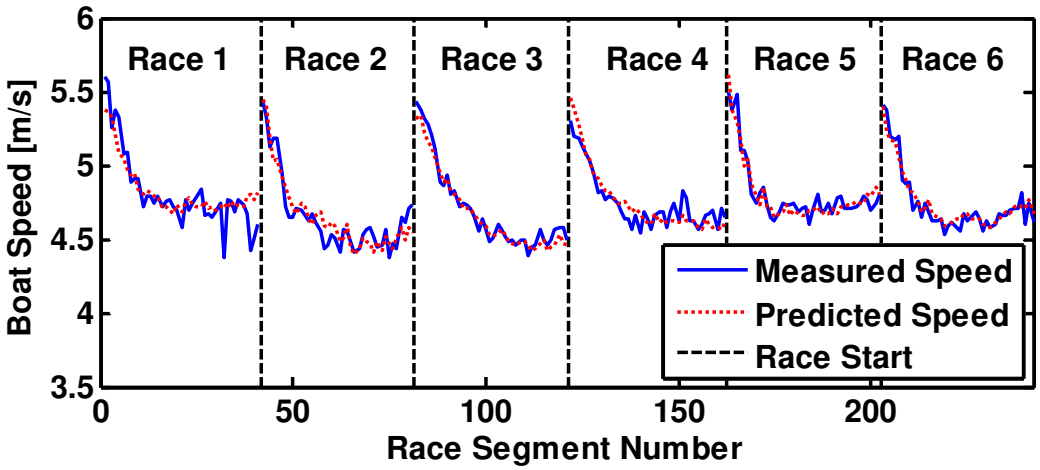


Figure 14. Comparison between predicted and measured boat speed, complete data set consisting of 6 races by different crews.

Analysis and Discussion

Biomechanical Fingerprint Identification

With the help of machine learning approaches we showed that our proposed features are discriminative enough to identify the individual rowers in our experiment. Out of all 130 features, the "Finish Slip" is the most discriminative feature for all three tested feature selection methods. By knowing this single value, the corresponding rower can be identified with an average accuracy of 74.6%. Using the top three features, the classification accuracy scores 95.6%. None of the boat-specific features ranked within the top-five features to discriminate a rower. Rower's individual characteristics can be found primarily in their oar movement rather than their impact on the boat drive or stability.

Figure 15a depicts the values of the two most discriminative features. One point in the plot corresponds to one rower during one race segment. The four colors indicate the athlete identities (A, B, C or D). The six different shapes of the point markers correspond to crew combinations. For example, squares are available in red color and in black color. These points correspond to the race segments in which rower A (red) and rower C (black) were rowing the

double together.

Figure 15b includes covariance error ellipses (Jackson, 2005; Spruyt, 2014). These ellipses represent the two-dimensional 80% confidence intervals in which the features for each rower can be found. It shows that each rower occupies a dedicated area and forms an individual cluster which only partly overlaps with other rowers' clusters. For each rower, the values of the two features are variable but they stay within individual ranges. This illustrates the discriminative nature of these two features for each rower.

Within each of these four clusters three sub-clusters can be identified. For rower D (green), the sub-clusters are marked with green stars, green diamonds and green circles. These sub-clusters are highlighted with the corresponding 80% confidence ellipses in Figure 15c. Each of these sub-clusters corresponds to one crew combination the rower D was part of. The size of the sub-clusters illustrate the rower's consistency: The bigger the area of one of these sub-clusters is, the more inconsistent the rowing technique of the corresponding rower was when racing in the corresponding crew combination (Perl & Baca, 2003). For example, the technique of rower D was most consistent when rowing together with rower C. Besides looking at the size of the sub-clusters, we can also find information considering their positions and overlaps. This distribution is a measure for the rower's adaptability: If the sub-clusters of one rower are close to each other or even overlapping, this indicates a dominant rowing technique or limited adaptation capabilities of the rower. Rowers with sub-clusters which are apart from each other show that they adjusted their technique under the influence of the other rower in the crew. The sub-clusters from rower A are the least spread-out ones. This rower kept her technique most constant, also when rowing together with other crew partners.

For each rower, the corresponding sub-clusters are mainly spread out on the y-axis. This means the main adjustments were made concerning the "Ang Drive Accel Point" feature, only rower D managed to vary the "Finish Slip" parameter as well.

Clusters from rower B and D have the largest overlap (Figure 15b). This means these two rowers have the capability to perform similar rowing technique according to the two considered features. However, when these two rowers are sitting together in the boat, they do not use this technique, the two corresponding sub-clusters (blue diamonds and green diamonds in Figure 15a) are not overlapping.

Clusters from rower A and B as well have an overlap in Figure 15b. In contrast to the previous example, these two rowers not only have the capability to perform similar technique, they also actually apply this common technique when they row together. This is visualized by the overlap of the corresponding two sub-clusters (blue crosses and red crosses) in Figure 15a.

Table 6. Summary of achieved 2000m race times (in minutes) and corresponding cluster size (in pixels) for each crew combination. Cluster sizes are calculated from Figure 15d, they are a measure on how different the rowing technique of both crew partners is.

Race 1			**Race 2**			**Race 3**		
Crew	**Time**	**Cluster size**	**Crew**	**Time**	**Cluster size**	**Crew**	**Time**	**Cluster size**
BA	6:57.51	8.055	BD	7:12.98	29.984	CB	7:03.13	17.500
CD	7:00.06	20.073	CA	7:08.87	15.610	DA	7:05.16	22.023

In Figure 15d each pair of sub-clusters is merged and their corresponding confidence ellipses are illustrated. Each of these six clusters visualizes one crew combination and includes all strokes of both rowers sitting in the boat. When compared to the others, the cluster BA is the most compact one. This again illustrates, that these two rowers are most similar and consistent over time concerning the two considered features. This combination also achieved the fastest time over the 2000m race distance. The size of each of the six clusters is calculated and compared to the achieved race time in Table 6. In each race, the winning boat also scored a smaller cluster size.

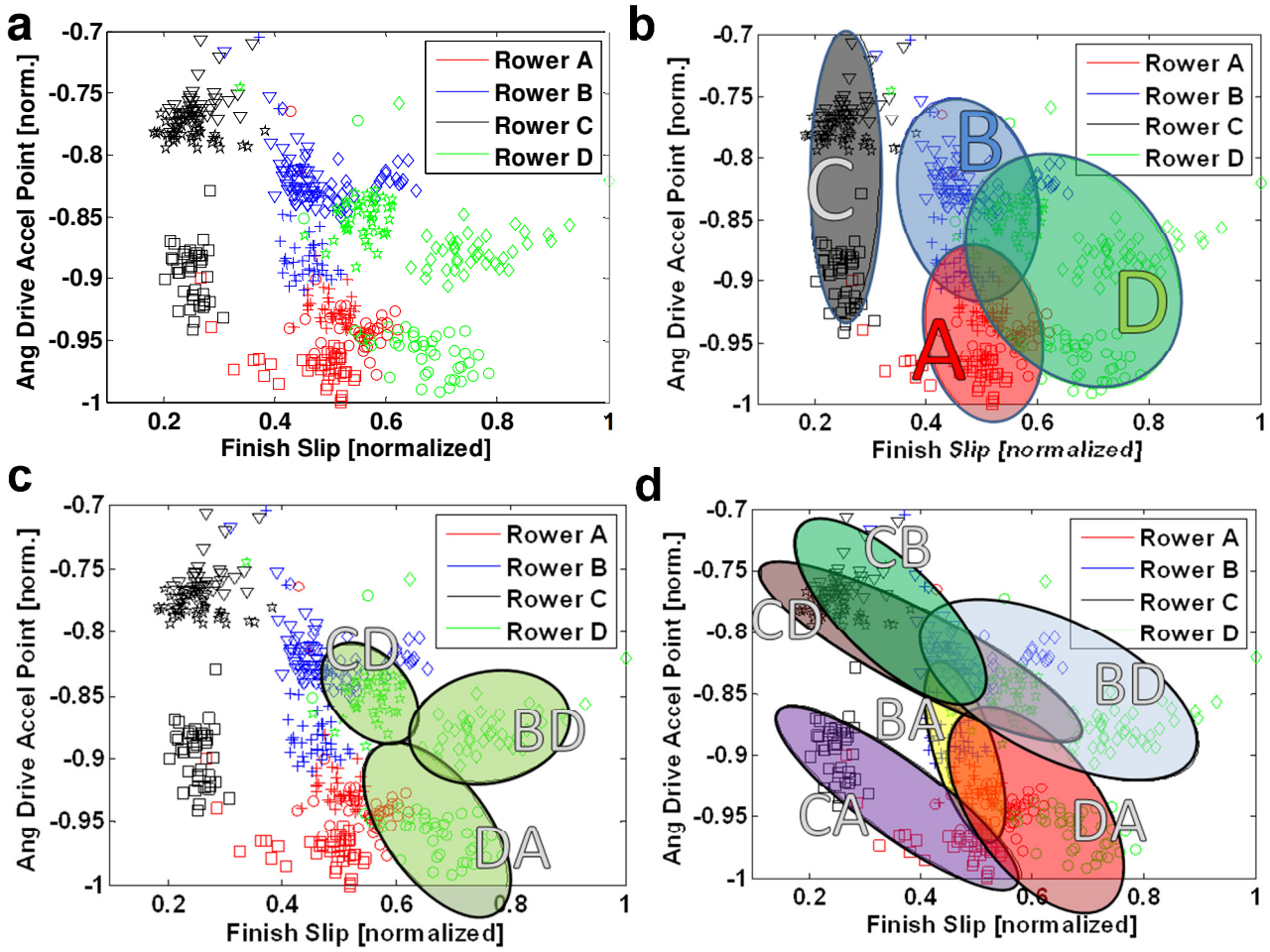


Figure 15. The scatter plot (a) shows the distribution of the two most discriminative features. One marker corresponds to averaged features of one rower during one race segment. Different colors indicate different rowers. Same marker shapes indicate same crew combinations. (b) indicates the 80% confidence ellipses forming clusters of the individual rowers. (c) marks three sub-clusters of rower D. Each sub-cluster corresponds to rower D's technique when rowing in three different crews. Each cluster in (d) correspond to one crew and considers all strokes from both crew partners. The more compact the clusters are, the more similar the two crew partners are rowing.

Correlations with Boat Speed

The proposed linear model needs $N = 5$ features as input variables to be able to predict the

boat speed with a root mean square error of less than 35% of the standard deviation of the measured boat speed. In Figure 14 it can be seen that the highest error usually occurs during the start phase of the race, where rowers' technique is usually different compared to the rest of the race. Measured velocity peaks during the middle phase of the race are most probably due to changed environment conditions such as wind or bumpy water. These external influences are not considered in our regression model and therefore the prediction does not include these peaks.

We used the identified correlations as well as their weights as starting points for discussions with rowing biomechanists to discover and explain potential causal dependencies:

Four out of the top five features (Table 5, features 2-5) are metrics that were measured during the previous race segment. This indicates that the average boat speed during one 50m race segment correlates highly to what happened during the previous 50m race segment. This suggests that there is a reaction time of several meters until changes in the oar or boat movement impact the boat speed. This inertia can be explained by the crew and boat mass of 160kg to 180kg.

All five selected features are boat-specific (Table 5, feature 3) or crew-specific (Table 5, features 1, 2, 4 and 5). None of the individual oar-specific features falls within the most correlated features. The experiment was not able to determine whether a particular seat in the boat, in this case bow or stern, correlates more strongly to the boat speed than the other.

The maximal value the handle speed reaches during the drive phase of the stroke is the most strongly correlated feature to boat speed (Table 5, feature 1). Assuming the blade is fully in the water, a higher handle velocity leads to more boat acceleration and therefore the positive correlation factor with the boat speed is consistent with coaching expectations. Similar explanations apply for the power applied to the oar handle (Table 5, feature 4) and the average handle velocity (Table 5, feature 5). These two features are the fourth- and fifth-ranked features and also correlated positively with the boat speed.

The second-ranked feature is based on the average oar handle velocity during the recovery phase (Table 5, feature 2). The correlation factor of this feature is negative, meaning the model predicts the faster the handles are moved during the recovery phase, the slower the boat speed. This is in line with coaching literature (Soper & Hume, 2004), which teaches that the rower should move slowly to the catch position relative to the speed of the drive phase. The slower movement is especially important at the end of the recovery phase in order to enable a smooth transition to the next drive phase.

The third-ranked feature is the standard deviation of the transversal boat acceleration during the recovery phase (Table 5, feature 3). This feature is a measure for the boat instability. The correlation factor is positive. However, the data-driven method cannot decide whether a) higher instability causes higher boat speed, or b) higher boat speed causes higher instability. Biomechanical models and coaching literature suggest that option b) is the most likely (Altenburg et al., 2008; Fritsch, 2005; Gravenhorst et al., 2011; Sinclair et al., 2009). Higher boat speed leads to more unintentional boat movements, which increases drag factor.

The standard deviation of transversal boat acceleration (Table 5, feature 3) is the only boat-specific feature in the top-five features, the others are all crew-specific. According to Loschner et al. variations of boat orientations, which is another indicator of instability, can largely be explained by different rowing styles, skills and experience levels (Loschner et al., 2000). Thus, this boat-specific feature can probably be substituted by a combination of more fine-grained

oar- or crew-specific features that more specifically explain the causes of the boat instabilities.

Limitations

We are aware that the described approach has several limitations.

Obtrusiveness. The boat measurement system is designed to be as unobtrusive as possible, however the instrumented boat is not physically identical to the usual uninstrumented one, and therefore it cannot be fully assured that the boat behavior is unchanged. Additionally, rowers might be psychologically impacted by knowing that their movements are being recorded and consequently they might row differently. However, similar sensor setups have been found unobtrusive in related studies (Tessendorf et al., 2011).

Measurement errors. We have not validated the error involved by the used measurement systems. According to the manufacturers, the force sensor is accurate to 2% of full scale (1500N), the oar angle sensor's accuracy is better than 0.5°, the boat speed and accelerometers are accurate to 2%. However there might be additional inaccuracies caused by mounting errors and misalignments.

Generalization. The number of rowers and measurements we used is limited and the measured group of rowers is not representative. Thus, the specific results cannot be generalized. To answer analogical questions for other rowers, the presented methods have to be applied to these rowers' data. The goal of this work is to introduce generalizable methods, rather than to provide statistically relevant data for a representative group of rowers.

Conclusion

Each rower has a different rowing technique and different capability to adjust their technique to be compatible for a crew boat. In collaboration with elite-level rowing coaches and biomechanists we introduce quantitative performance metrics that describe crucial parts of rowing technique and which can be measured using unobtrusive mobile sensor systems in rowing boats. We implemented three different classifiers and performed sequential forward feature selection to identify the features that are most unique for each rower. These features make up the rower's biomechanical fingerprint and are relevant for identifying the best-fitting rowers for crews. We collected data from four world-class female athletes while they raced against each other in different crew combinations. We applied the described method to this dataset. The "Finish Slip" feature, which describes the rower's efficiency at the end of the rowing stroke, turned out to be the most discriminative feature. Our proposed k-nearest neighbor classifier outperformed the random forest and support vector machine classifiers in terms of rower identification accuracy. It was able to identify 74.6% of the rowers correctly solely based on this single feature which requires only oar sensor modules. Applying one or two additional features this accuracy improved to 90.7% or 95.6% respectively, however these features require an additional boat sensor to acquire boat accelerations. The two rowers with the best similarities regarding the two most discriminative features also scored the best time in comparison to all the other combinations within the group and later received an Olympic medal.

In the second part of this work, we showed how a linear regression model can be used to identify correlations between rowing features and boat speed. Our goal was to achieve a root mean square error below 0.087 m/s which is 35% of the boat velocity's standard deviation. By gradually increasing the number of features, we found that five features were sufficient for the

given dataset to reach this goal. The results of this data-driven approach suggest that boat velocity is mainly correlated to oar-specific input features and that there is a time delay of about 50 meters before input variances fully effect the output. The corresponding weight factors for these five input parameters were determined with the least-squares optimization algorithm. We also outlined that the statistical dependencies we found were consistent with the experiences of rowing biomechanists and coaches, and we discussed potential causal relationships for these dependencies.

Outlook

In future studies we plan to increase the number of sensors to extend the number of available performance metrics. Although the number of features presented in this work is sufficient for achieving classification accuracies of >95%, a broader set of features could offer new insights. This way, the results of the proposed data-driven approaches can be more fine-tuned and features describing boat-movements can be broken down into actual causes. For example instead of features describing the overall oar movement, measurements of the leg, upper body and arm movements can be included.

Further studies will also add more data to the database and this way makes the results statistically more relevant and easier to generalize from. Specifically, we are interested in applying and adapting the methods for male rowers, bigger crew boats, lightweight rowers and sweep rowing. Additionally, time-dependency can be considered in order to account for anomalies due to sprint phases during races or different degrees of fatigue during training.

The presented crew selection example is based on the assumption that crew members should row as synchronously as possible. Although our recorded speed data supports this assumption for sculling, it would be interesting to investigate if and how a crew could benefit from complimentary rowing movements in sweep rowing.

We would like to further explore whether the position that a rower is seated in within a crew influences his/her biomechanical fingerprint. Based on qualitative experiences in related works, we expect rowers' behavior and their effect on the boat to differ depending on the position they sit in within the crew.

Further research concerning data-driven boat speed dependencies is ongoing. Instead of linear models with calculated features as input, we want to test non-linear models and/or use raw data as input.

Finally, the vision is to combine both presented methods, the biomechanical fingerprint identification and the boat speed correlation, to find an overall crew efficiency measure that identifies the rowers within a group that best fit together and generate the features that most increases the boat speed.

Acknowledgment

The authors thank all rowers who participated in the studies and pre-studies, and the collaborating coaches for their valuable time and feedback. Further thanks go to Rosa Brown from www.topproofreading.com for editing and proofreading this work.

References

Altenburg, D., Mattes, K. & Steinacker, J. (2008). *Handbuch Rudertraining*. Limpert.

Altman, N. S. (1992). An introduction to kernel and nearest-neighbor nonparametric regression, *The American Statistician, 46*(3), 175–185.

Armstrong, J. S. & Collopy, F. (1992). Error measures for generalizing about forecasting methods: Empirical comparisons, *International Journal of Forecasting, 8*(1), 69–80.

Baca, A. & Kornfeind, P. (2006). Rapid feedback systems for elite sports training, *Pervasive Computing, IEEE, 5*(4), 70–76.

Baguet, A., Bourgois, J., Vanhee, L., Achten, E. & Derave, W. (2010). Important role of muscle carnosine in rowing performance, *Journal of Applied Physiology, 109*(4), 1096–1101.

Christov, R., Christov, R., Zdravkov, N. & Nolte, V. (1988). Selection and testing system based on biomechanical studies in racing boats and on rowing ergometer, *FISA Coaches Conference*.

Claessens, A., Bourgois, J., Pintens, K., Lefevre, J., Van Renterghem, B., Philippaerts, R., Loos, R., Janssens, M., Thomis, M. & Vrijens, J. (2002). Body composition and somatotype characteristics of elite female junior rowers in relation to competition level, rowing style and boat type., *Human Biology Budapestiensis, 27,* 159–165.

Claessens, A. L., Bourgois, J., Van Aken, K., Van der Auwera, R., Philippaerts, R., Thomis, M., Vrijens, J., Loos, R. & Lefevre, J. (2005). Body proportions of elite male junior rowers in relation to competition level, rowing style and boat type, *Kineziologija, 37*(2), 123–132.

Cortes, C. & Vapnik, V. (1995). Support-vector networks, *Machine learning, 20*(3), 273–297.

Cosgrove, M., Wilson, J., Watt, D. & Grant, S. (1999). The relationship between selected physiological variables of rowers and rowing performance as determined by a 2000 m ergometer test, *Journal of Sports Sciences, 17*(11), 845–852.

Davey, N., Grenfell, R., James, D., Mackintosh, C. & Zhang, K. (2010). Monitoring sports. US Patent 7,715,982.

Fahrig, S. & Witte, K. (2007). Koordination der interaktion im ruderzweier ohne steuermann, *dim, 1,* 4.

FISA (2011). *Rule Book*. http://www.worldrowing.com/fisa/resources/rule-books.

Fritsch, W. (2005). *Das grosse Buch vom Rennrudern*, Meyer & Meyer.

Galloway, M. & Draper, C. (2011). The dominance or adaptability of female rowers in double sculls, *Proceedings of the 8th Australasian Biomechanics Conference*.

Gravenhorst, F., Muaremi, A., Kottmann, F., Sigrist, R., Gerig, N., Draper, C. & Tröster, G. (2014a). Strap and row: Rowing technique analysis based on inertial measurement units implemented in mobile phones, *International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP).*

Gravenhorst, F., Tessendorf, B., Arnrich, B., Codoni, C. & Tröster, G. (2011). Analyzing rowing crews in different rowing boats based on angular velocity measurements with gyroscopes, *Proceedings of the International Symposium on Computer Science in Sport (IACSS 2011).*

Gravenhorst, F., Thiem, C., Tessendorf, B., Adelsberger, R., Arnrich, B., Draper, C., Smith, R. M. & Tröster, G. (2014). Sonicseat: design and evaluation of a seat position tracker based on ultrasonic sound measurements for rowing technique analysis, *Springer Journal of Ambient Intelligence and Humanized Computing* pp. 1–10.

Groh, B. H., Reinfelder, S. J., Streicher, M. N., Taraben, A. & Eskofier, B. M. (2014). Movement prediction in rowing using a dynamic time warping based stroke detection,

Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2014 IEEE Ninth International Conference on, IEEE, pp. 1–6.

Hill, D. W., Alain, C. & Kennedy, M. D. (2003). Modeling the relationship between velocity and time to fatigue in rowing, *Medicine and Science in Sports and Exercise, 35*(12), 2098–2105.

Hill, H. (2002). Dynamics of coordination within elite rowing crews: evidence from force pattern analysis, *Journal of Sports Sciences, 20*(2), 101–117.

Ingham, S., Whyte, G., Jones, K. & Nevill, A. (2002). Determinants of 2,000 m rowing ergometer performance in elite rowers, *European Journal of Applied Physiology, 88*(3), 243–246.

Jackson, J. E. (2005). *A user's guide to principal components*, Vol. 587, John Wiley & Sons.

Jain, A. & Zongker, D. (1997). Feature selection: Evaluation, application, and small sample performance, *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 19*(2), 153–158.

Kaya, M., Minamitani, H., Hase, K. & Yamazaki, N. (1995). Motion analysis of optimal rowing form by using biomechanical model, *Engineering in Medicine and Biology Society, 1995., IEEE 17th Annual Conference*, Vol. 2, IEEE, pp. 1281–1282.

Kleshnev, V. (2002). Moving the rowers: biomechanical background, *Australian Rowing, Carine, WA* 25: 16–19.

Kleshnev, V. (2010). Boat acceleration, temporal structure of the stroke cycle, and effectiveness in rowing, *Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology, 224*(1), 63–74.

Korndle, H. & Lippens, V. (1988). Do rowers have a particular 'footwriting', *Biomechanics in sport. London: Institution of Mechanical Engineers* pp. 7–11.

Legendre, A. M. (1805). *Nouvelles méthodes pour la détermination des orbites des comètes*, F. Didot.

Loschner, C. & Smith, R. (1999). The relationship between seat movement and boat acceleration during sculling, *ISBS-Conference Proceedings Archive*, Vol. 1.

Loschner, C., Smith, R. & Galloway, M. (2000). Intra-stroke boat orientation during single sculling, *ISBS-Conference Proceedings Archive*, Vol. 1.

Millward, A. (1987). A study of the forces exerted by an oarsman and the effect on boat speed, *Journal of Sports Sciences, 5*(2), 93–103.

Montillo, A. A. (2009). Random forests, *Lecture in Statistical Foundations of Data Analysis* .

Nielsen-Kellermann (2014). Retrieved October 11, 2014. http://www.nkhome.com/

Nozaki, D., Kawakam, Y., Fukunaga, T. & Mlyashtta, M. (1993). Mechanical efficiency of rowing a single scull, *Scandinavian Journal of Medicine & Science in Sport,s 3*(4), 251–255.

Peng, H., Long, F. & Ding, C. (2005). Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy, *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 27*(8), 1226–1238.

Perl, J. & Baca, A. (2003). Application of neural networks to analyze performance in sports, *Proceedings of the 8th Annual Congress of the European College of Sport Science, Salzburg: ECSS*, Vol. 342.

Qin, J., Sun, X., Xiang, X. & Niu, C. (2009). Principal feature selection and fusion method for image steganalysis, *Journal of Electronic Imaging, 18*(3), 033009–033009.

Robinson, P. & Topolski, D. (2013). *True blue: the Oxford boat race mutiny*, Random House.

Sabatini, A. M. & Genovese, V. (2006). Gyroscopic measurements of the rower's oar pitch angle, *Proceedings of the 7th WSEAS International Conference on Automation &*

Information, ICAI'06, World Scientific and Engineering Academy and Society (WSEAS), Stevens Point, Wisconsin, USA, pp. 1–6. http://dl.acm.org/citation.cfm?id=1362937.1362938

Sanderson, B. & Martindale, W. (1986). Towards optimizing rowing technique., *Medicine and Science in Sports and Exercise, 18*(4), 454–468.

Secher, N. & Vaage, O. (1983). Rowing performance, a mathematical model based on analysis of body dimensions as exemplified by body weight, *European Journal of Applied Physiology and Occupational Physiology, 52*(1), 88–93.

Sinclair, P., Greene, A. & Smith, R. (2009). The effects of horizontal and vertical forces on single scull boat orientation while rowing, *ISBS-Conference Proceedings Archive*, Vol. 1.

Smith, R. & Draper, C. (2002). Quantitative characteristics of coxless pair-oar rowing, *ISBS-Conference Proceedings Archive*, Vol. 1.

Smith, R. & Loschner, C. (2002a). Biomechanics feedback for rowing, *Journal of Sports Sciences, 20*(10), 783–791.

Smith, R. & Loschner, C. (2002b). Net power production and performance at different stroke rates and abilities during pair-oar rowing, *International Research in Sports Biomechanics* p. 279.

Smith, R. M. & Spinks, W. L. (1995). Discriminant analysis of biomechanical differences between novice, good and elite rowers, *Journal of Sports Sciences, 13*(5), 377–385.

Smith, T. B. & Hopkins, W. G. (2012). Measures of rowing performance, *Sports Medicine, 42*(4), 343–358.

Soper, C. & Hume, P. (2004). Towards an ideal rowing technique for performance: the contributions from biomechanics, *Sports Medicine, 34*(12), 825–848.

Spruyt, V. (2014). How to draw a covariance error ellipse?, http://www.visiondummy.com/-2014/04/draw-error-ellipse-representing-covariance-matrix/.

Tessendorf, B., Gravenhorst, F., Arnrich, B. & Tröster, G. (2011). An imu-based sensor network to continuously monitor rowing technique on the water, *Proceedings of the Seventh International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP 2011)*, IEEE press.

Wagner, J., Bartmus, U. & De Marees, H. (1993). Three-axes gyro system quantifying the specific balance of rowing, *International Journal of Sports Medicine, 14*, 35–35.

Wing, A. M. & Woodburn, C. (1995). The coordination and consistency of rowers in a racing eight, *Journal of Sports Sciences*, *13*(3), 187–197.

Wolberg, J. (2006). *Data analysis using the method of least squares*, Springer.

Zhu, Z., Ong, Y.-S. & Dash, M. (2007). Wrapper–filter feature selection algorithm using a memetic framework, *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, 37*(1), 70–76.

Zou, K. H., Tuncali, K. & Silverman, S. G. (2003). Correlation and simple linear regression 1, *Radiology, 227*(3), 617–628.

Appendix

Table 7. Descriptions of boat-specific features. The column "Sensors" mentions one possible set of sensors that can deliver the necessary raw data to calculate the appropriate feature.

Names of features	Description	Sensors
Stroke Rate	Number of strokes per minute	Boat-acceleration
Vel Max Vel Min Vel Ave Vel Range	Maximum, minimum, average and range of boat velocity during one stroke	GPS, boat acceleration
Dist/Stk	Distance in meters the boat travelled during one stroke	GPS, boat acceleration
Vel Catch Vel Finish	Boat velocity at the beginning (catch position) and the end (finish position) of the drive phase	GPS, boat acceleration
t Catch Lost Vel Catch Lost	Time interval between reaching the catch position and placing the blade into the water; and amount of velocity the boat lost during this delay.	GPS, boat acceleration
Acc Drive Min Acc Drive Max	Minimal and maximal boat acceleration (propulsive direction) during the drive phase. The same features are extracted for the other two acceleration axes (transversal and vertical acceleration).	boat acceleration
Acc Recov Peak	Value of highest boat acceleration peak (propulsive direction) during recovery phase	boat acceleration
t Recov Decel	Length of time interval the boat is decelerating for during recovery phase	boat acceleration
Accy Recov Min Accy Recov Max Accz Recov Min Accz Recov Max	Minimal and maximal boat acceleration (transversal and vertical acceleration) during the recovery phase.	boat acceleration
Pitch Min Pitch Max Pitch Range	Minimum, maximum and range of boat pitch angle (up/down movement of bow ball) during one stroke	boat acceleration and gyroscope
Roll Drive Min	Minimum, maximum and range of boat roll angle (boat tipping to left or right) during drive and	boat acceleration

Roll Drive Max	recovery phase	and gyroscope
Roll Drive Range	The same features are extracted for the boat yaw angle (making a turn to bow or stroke side).	
Roll Recov Min		
Roll Recov Max		
Roll Recov Range		

Table 8. Descriptions of oar-specific features. These features are individual for each rower. The column "Sensors" mentions one possible set of sensors that can deliver the necessary raw data to calculate the appropriate feature.

Names of features	**Description**	**Sensors**
t Drive t Recovery t Stroke Drive:Total	Duration (in seconds) of drive phase, recovery phase and total stroke. From that the ratio (in percent) between drive and total stroke duration is also calculated.	Angle sensor at gate
Ang Catch Ang Finish	Absolute angle of oar relative to the boat at the beginning and end of the drive phase (catch and finish position).	Angle sensor at gate
Ang Stroke Length Handle Dist	Swept oar angle during the drive phase. The angle value (Ang Stroke Length) is also converted to meters (Handle Dist).	Angle sensor at gate
F Max Gate F Mean Gate	Force (propulsive direction) at gate. The same two features are extracted for the force at the oar handle.	Force sensor at gate
RFD Ave RFD Peak	Rate of Force Development (average and peak value): Slope of the force curve at the beginning of the stroke.	Angle and force sensor at gate
t Max Force	Time interval from the beginning of the stroke until the maximum oar force is attained.	Angle and force sensor at gate
Handle Vel Drive Ave Handle Vel Drive Max	Average and maximal handle velocity during the drive phase.	Angle sensor at gate
Handle Vel Recov Ave	Average handle velocity during the recovery phase.	Angle sensor at gate

Ang Catch Slip Ang Finish Slip Ang Effective Stk	Both slip values are measures of the rower's inefficiency at the beginning (catch) and the end (finish) of the stroke respectively. Ang Catch Slip is the catch oar angle (Ang Catch) minus the oar angle when the gate force reaches threshold value (50N) at the beginning of the drive phase. Ang Finish Slip is the finish oar angle (Ang Finish) minus the oar angle when the gate force reaches threshold (10N) approaching the end of the stroke. The total stroke length (Ang Stroke Length) minus both slip values results in the effective stroke length (Ang Effective Stk).	Angle and force sensor at gate
Ang Max Force	Oar angle during drive phase at which the maximum oar force is applied.	Angle and force sensor at gate
Ang Recov Decel Point	Oar angle during recovery phase when boat starts to decelerate.	Angle and force sensor at gate, boat acceleration
Ang Drive Accel Point	Oar angle during drive phase at which the boat accelerates the most.	Angle and force sensor at gate, boat acceleration
Power Handle Ave Work Handle	Average power (propulsive direction) applied by the rower to the oar handle during one stroke. The stroke duration and work is calculated with this value.	Angle and force sensor at gate
Power Drive Ave	Average power (propulsive direction) applied by the rower to the oar handle during drive phase.	Angle and force sensor at gate
Power Handle Tot	Average power (propulsive and transversal direction) applied by the rower to the oar handle during one stroke.	Angle and force sensor at gate

Estimation of Torque Variation from Pedal Motion in Cycling

Quintana-Duque, J.-C., Dahmen, T., Saupe, D.

Department of Computer and Information Science, University of Konstanz, Germany

Abstract

In cycling, the pedalling technique is determined mostly by variations in the torque applied to the pedals during crank rotation. We developed and validated a method to compute these variations from the pedalling motion using an ergometer. The torque at the pedal is the sum of the torques needed to overcome all resistive forces and the torque required for any changes of angular momentum of the ergometer flywheel. This last torque is proportional to the angular acceleration of the crank. For an ergometer with almost constant brake torque, we may assume that variations in the pedal force can be extracted from the pedal motion alone. The key problem is to reliably estimate the angular pedal acceleration from noisy 3D motion capture (MoCap) or 2D video data. We projected the positional data onto a least squares fitting circle, then filtered the resulting angular time sequence by local polynomial regression. Finally, we solved the torque equilibrium equation for the pedal torque. For validation of the method, we used direct pedal torque measurement. In our experiments, pedal brake forces ranged between 100 and 250 N, and cadences of 60, 80, and 100 rpm were used. The pedal torque results from MoCap were better than from video. The results from video were close to MoCap results when a correction of the marker position was applied.

KEYWORDS: CRANK ACCELERATION, PEDALLING MOTION, NET TORQUE VARIATION, FORCE VARIATION

Introduction

Cycling is the result of the interaction between the cyclist, the bicycle, and the environmental constraints. The right selection of the variables involved in adjusting a bicycle (e.g., handle bar position, seat height, crank length, etc.), the correct body position on the bicycle, and a good pedalling technique are necessary to prevent cycling injuries and to optimize the force distribution during the pedal stroke.

Applying correctly oriented forces to the pedal (i.e., pedalling technique) is a major component of skilled performance on the bicycle. Although there is no agreement on the characteristics of the optimal pedalling technique, any desired pedalling technique can be learned and trained by indoor ergometer cycling. A training session in the lab is based on the analysis and feedback of the distribution of the applied pedal forces during each revolution of the crank. In many cases, the torque variation is enough for the training of pedalling techniques and the correction of force asymmetry between the legs. See some examples in Böhm, Siebert, and Walsh (2008) and Faria (2009).

The determination of pedal forces is fundamental to analyse cycling performance from a biomechanical point of view. Sensors implemented in the pedal have been validated in the literature for measuring force in one dimension up to three dimensions based on strain gauges or piezoelectric elements. For a list of sensors see Mornieux, Zameziati, Mutter, Bonnefoy, and Belli (2006) and Stapelfeldt, Mornieux, Oberheim, Belli, and Gollhofer (2007). A direct measure of the torque (resp., tangential force applied to the pedal that rotates the crank) requires expensive sensors, but indirect ways of calculating the torque variations are possible and presented in this contribution.

The study of pedal forces in cycling is of interest for several applications. For example, some related topics are: pedalling biomechanics (Christensen et al., 2000; Hug, Turpin, Guével, & Dorel, 2010; Kautz & Hull, 1993), limb coordination (Bini, Diefenthaeler, & Mota, 2010; Jirsa, Fink, Foo, & Kelso, 2000), human motion modeling (Franz Höchtl, Harald Böhm, & Veit Senner, 2010; Hull, Kautz, & Beard, 1991), detection and correction of asymmetry (Carpes, Rossato, Faria, & Bolli Mota, 2007; Sanderson, 1990; Smak, Neptune, & Hull, 1999), evaluation of body performance given a pedalling technique (Cannon, Kolkhorst, & Cipriani, 2007; Ettema & Loraas, 2009), cadence and workload effects on pedalling technique (Black, 1994; Rossato, Bini, Carpes, Diefenthaeler, & Moro, 2008; Stapelfeldt, Mornieux, & Gollhofer, 2006) and the influence of pedalling technique on muscular efficiency (Theurel, Crepin, Foissac, & Temprado, 2011).

In order to calculate the angular acceleration from the pedalling motion, the angular position of a point representing the angular position of the crank is differentiated twice. This is not a trivial task, because the differentiation process is noise amplifying by its very nature, (Ovaska & Valiviita, 1998). For this reason, a bandwidth differentiator with special characteristics is required for this calculation. We tested the Savitsky-Golay filter (Savitzky & Golay, 1964) from which we obtained the filtered components required for the calculation of the second derivative of the angular position of the crank. In addition, we observed the differences in our results when a correction of the marker position was applied for MoCap and video data. The physical relation between the angular acceleration of the crank and the net torque applied to the pedals is derived, allowing us to compute the net torque from our motion data. The comparison of these results with directly measured crank torque (using an SRM Torque Box[1]) provides the validation of our method.

Methods

The variation in the angular acceleration is proportional to the variation in the applied torque when the pedal brake force is constant, which can be achieved using ergometers. We propose to derive the variation in the net torque from measurements of pedal motion that can be made in the laboratory using motion-capturing or plain commercial cameras. The reconstruction of the position and the orientation of the pedal was done in two ways: 1) recording video of two LEDs, and 2) by motion capture using two active infrared markers. See the experimental setup in Figure 1.

[1] http://www.srm.de/products/torque-analysis-system/

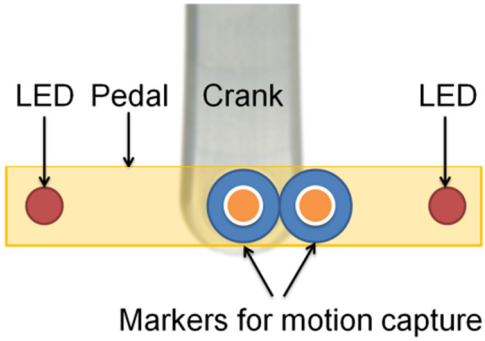


Figure 1 Pedal motion was recorded using video (through two LEDs) and motion capture (through two active infrared markers) in order to reconstruct the position and the orientation of the pedal. For motion capture one marker was placed on the pivot point of the pedal rotation and the other marker was placed to the right side of the first marker. For video the LEDs were placed on both sides of the pedal such that the midpoint between LEDs was aligned with the pivot point.

Calculation of the second derivative from positional data

We assumed that the angular position of the pivot point of rotation of the pedal and the crank angle θ with respect to the center of rotation of the crank are the same. If $(x(t), y(t))$ denote the marker coordinates of a point representing the angular position of the crank, with the origin of the coordinate system placed at the center of the crank rotation, then crank angle is given by $\theta(t) = \tan^{-1}(y(t)/x(t))$. The second derivative of θ (i.e., angular acceleration) is calculated using the chain rule for derivatives

$$\ddot{\theta} = \frac{d^2}{dt^2}\tan^{-1}\frac{y}{x} = \frac{(\ddot{y}x - y\,\ddot{x})(x^2 + y^2)}{(x^2 + y^2)^2} - \frac{(\dot{y}x - \dot{x}y)(2x\dot{x} + 2y\dot{y})}{(x^2 + y^2)^2}. \qquad (1)$$

In our application, the Savitsky-Golay filter is applied separately to the x- and y-coordinates of the pedal motion to obtain smoothed data and its first and second derivatives. This filter is briefly reviewed in the next section.

Savitzky-Golay smoothing filter

The recordings of the pedal position with motion capture data (MoCap) and video data contain noise. This noise, viewed in the frequency domain, is amplified in the calculation of the second derivative by a factor of $4\pi^2 f^2$ with f being the frequency. In order to increase the signal-to-noise ratio without greatly distorting the signal, we used the Savitzky-Golay filter (Savitzky & Golay, 1964), also known as polynomial smoothing (Hamming, 1989) or as least-squares smoothing filters (Schafer, 2011).

The Savitzky-Golay filter is a method for data smoothing based on an approximation by a local least-squares polynomial. This filter reduces the noise while maintaining the shape and height of the waveform peaks (e.g., their relative widths and heights), (Press, Teukolsky, Vetterling, & Flannery, 2007). This filter is a generalization of the FIR averaging filter. It can preserve the high frequency content of the desired signal, at the expense of not removing as much noise as the averager, (Orfanidis, 2010).

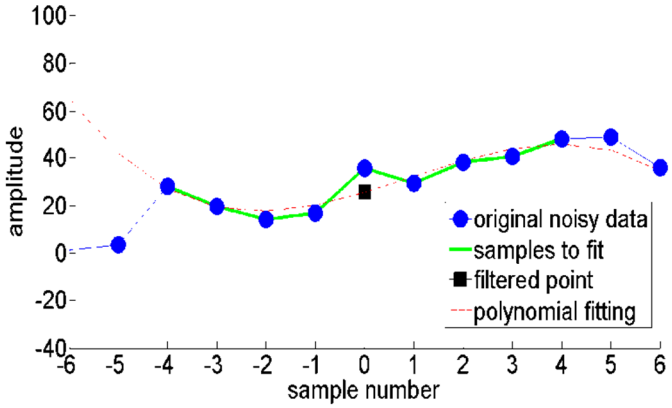


Figure 2 Example of polynomial smoothing. The smoothed output value (black square) is obtained by evaluating a polynomial of order N = 3 (red line) fitted on a group of 2M + 1 samples (on the green line), here with half-length M = 4, at n = 0.

The basic idea of this filter is to fit a polynomial to a set of consecutive input samples and then evaluate the resulting polynomial at the central point. Figure 2 shows an example of polynomial smoothing for the central point $n = 0$. The input $x[k]$, with $k \in \mathbb{Z}$, is a sequence of discrete points (solid blue dots). A window (i.e., points on the green line) with a window size of $2M + 1$ data points, with $M \geq 1$, is used to calculate a least-squares polynomial fit (here $2M + 1 = 9$). A polynomial $p[k]$ of degree $N = 3$ (red dashed line) is fitted to the data $x[k]$, with $k = n - M, \ldots, n + M$ and the smoothed output value (black square) is obtained by evaluating $p[k]$ for $k = n$. We obtain the coefficients a_i of the polynomial p of order N,

$$p[k] = \sum_{i=0}^{N} a_i k^i \,, \tag{2}$$

that minimizes the mean-squared approximation error ϵ_n for the group of input samples centered on n,

$$\epsilon_n = \sum_{m=-M}^{M} (p[n+m] - x[n+m])^2. \tag{3}$$

It can be shown that this is equivalent to discrete convolution with a fixed impulse response, (Schafer, 2011). Savitzky and Golay (1964) published tables of filter coefficients for combined smoothing and differentiation. These tables are given for two parameters: the half-length of the fitting window, M, and the order of the fitted polynomial, N. The dependence of the cutoff frequency f_c on N and M is given by

$$f_c = \frac{N+1}{3.2M - 4.6} \tag{4}$$

for $M \geq 25$ and $N \leq M$. Sometimes the same cutoff frequency can be achieved using different combinations of N and M.

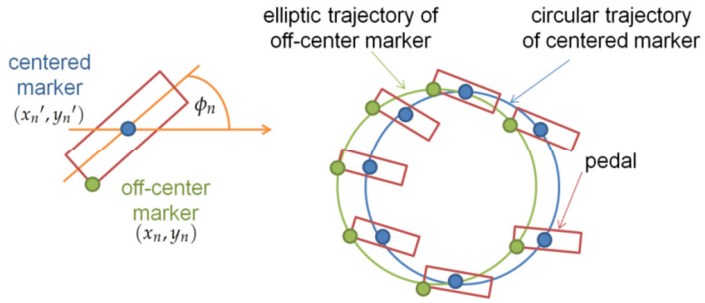


Figure 3 An off-center marker on the pedal gives a trajectory that is not a circle when the pedal orientation is not constant.

Correction of the marker position

The crank rotation during pedaling motion ideally takes place in a two dimensional plane. We expect that any marker on the crank, except the pivot of crank rotation, describes a perfect (sampled) circular trajectory. The angular position of the crank can be obtained from the trajectory of a point representing the angular position of the crank. In the same way, the angular position can be calculated from the trajectory described by a marker on the pedal, provided that the marker is on the pivot of pedal rotation. Otherwise the marker trajectory turns into an approximate ellipse due to simultaneous rotations of pedal and crank. See Figure 3.

However, although the marker position is on the pivot of pedal rotation, the crank and pedal rotations in the real world describe an elliptical trajectory because the bearings at the pedals and in the bottom bracket may have small play and the crank sets are slightly elastic.

The eccentricity of an ellipse fitted to the marker trajectory characterizes the amount by which the marker trajectory deviates from a circle. The eccentricity e is defined as $e = \sqrt{1 - b^2/a^2}$ with a and b denoting the length of the major and minor axes, respectively.

The correction $(\Delta x, \Delta y)$ of the 2D position of a marker on the pedal can be calculated if the pedal orientation ϕ_n with respect to the horizontal axis is known. The correction of the marker position is given by

$$\begin{pmatrix} x_n' \\ y_n' \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \end{pmatrix} + \begin{pmatrix} \cos\phi_n & -\sin\phi_n \\ \sin\phi_n & \cos\phi_n \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix}, \quad (5)$$

where x_n and y_n are the old coordinates, x_n' and y_n' are the new coordinates after correction. The correction vector $(\Delta x, \Delta y)$ is defined such that: either 1) a prescribed target eccentricity of an ellipse fitted to the corrected positional data is achieved (in following called „prescribed eccentricity“) or 2) the RMS distance between the fitted ellipse and the corrected data is minimal among all possible corrections and their respective fitted ellipses (in following called „minimum error eccentricity M.E.“).

Related concepts: force and torque

The total force $\mathbf{F}_{\text{total}}$ applied to the pedal is the sum of all vector forces, see Figure 4, produced by the contractions and extensions of the leg and hip muscles which can be decomposed into tangential and radial forces, $\mathbf{F}_{\text{tan}}$ and $\mathbf{F}_{\text{rad}}$, respectively (Equation 6). The force $\mathbf{F}_{\text{tan}}$ is tangential to the crank rotation and the radial force $\mathbf{F}_{\text{rad}}$ is parallel to the crank. Only $\mathbf{F}_{\text{tan}}$ contributes to the crank rotation,

$$\mathbf{F}_{\text{total}} = \mathbf{F}_{\text{tan}} + \mathbf{F}_{\text{rad}}. \quad (6)$$

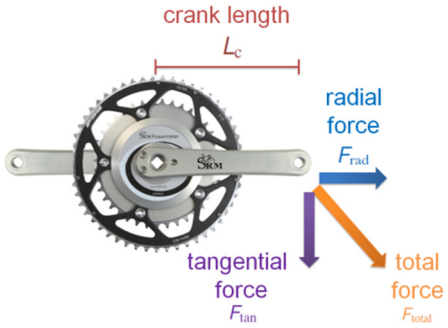


Figure 4 The total force applied to the pedal, $\mathbf{F}_{\text{total}}$, is the sum of two perpendicular forces: the tangential $\mathbf{F}_{\text{tan}}$ and the radial force $\mathbf{F}_{\text{rad}}$.

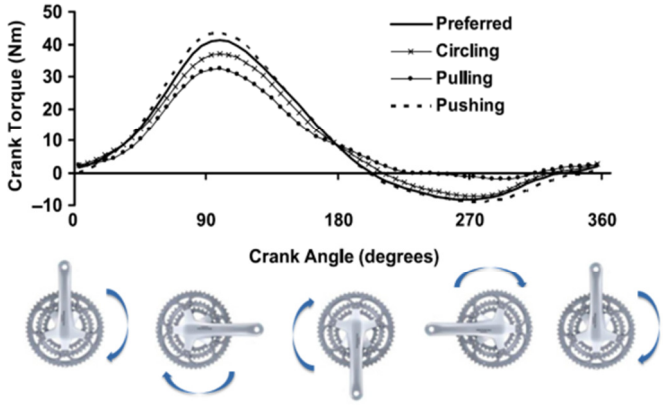


Figure 5 Right leg torque profiles using different pedalling styles at 90 rpm and 200 W. Figure adapted from Korff, Romer, Mayhew, and Martin (2007). Four pedalling styles are shown: individually preferred pedalling technique (preferred), pedalling emphasizing the transition phases through top dead center at 0° and bottom dead center of the crank cycle at 180° (circling), emphasizing an active pull during the upstroke of the crank cycle (pulling), and emphasizing the pushing action during the downstroke of the crank cycle (pushing).

The torque describes the effect of a force on the rotational motion of the pedal pivot point about the axis on the bearing. Mathematically, the torque is the cross product of the lever-arm length vector, $\mathbf{L_c}$, and the force $\mathbf{F}_{\text{tan}}$ acting on the end of the lever-arm. Here, the lever-arm length is equal to the crank length L_c. Thus, the magnitude of the torque is given by

$$\|\boldsymbol{\tau}\| = \|\mathbf{L}_c \times \mathbf{F}_{\text{tan}}\| = L_c F_{\text{tan}}. \quad (7)$$

Figure 5 shows the torque patterns of one pedal for different pedalling styles, where a positive peak around 90° (downstroke) and a negative peak around 270° (upstroke) can be seen for all styles. The measurements were made in Korff, Romer, Mayhew, and Martin (2007) using a custom made force pedal with two triaxial piezoelectric force sensors, where 0° is highest position of the crank rotation.

The net torque τ_{net} is the sum of the individual torques τ_{left} and τ_{right} applied to the left and right pedals:

$$\tau_{\text{net}} = \tau_{\text{left}} + \tau_{\text{right}}. \quad (8)$$

Figure 6 shows an example of the net torque curve with some typical features to describe it.

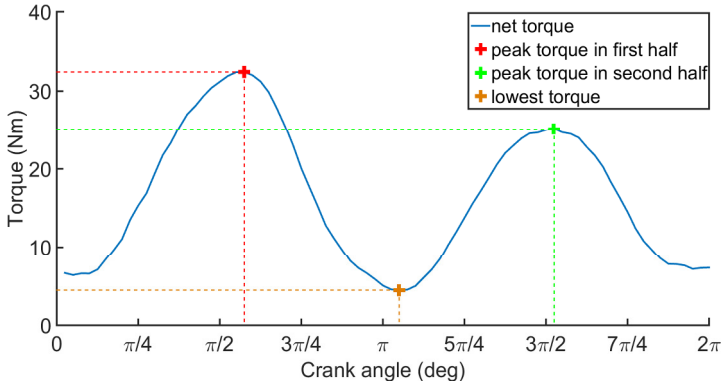


Figure 6 Mean of net torque τ_{net} for one pedal revolution with 60 rpm and 100 N ergometer brake force (i.e., 17.6 Nm).

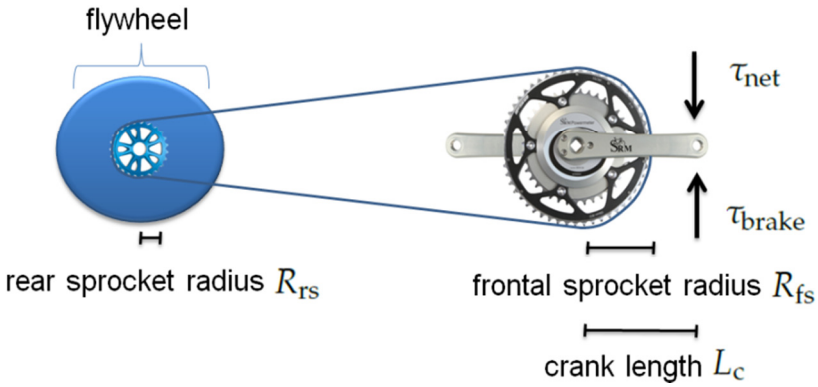


Figure 7 Basic model for the deduction of the torque equation.

For example, different peak values indicate an asymmetry due to emphasizing an unequal pushing action during the downstroke or due to a problem with pulling during the upstroke.

During the pedalling motion, two main torques act on the crank at the same time: the applied net torque τ_{net} and the brake torque τ_{brake}. τ_{net} is the sum of the individual pedal torques (Equation 8) and the brake pedal torque is the sum of all torques produced by the different forces against the cycling motion, e.g., the rolling and aerial resistive forces of a bicycle on the road, or the forces produced by an eddy-current brake in an ergometer in the lab.

Figure 7 shows a simple model of the torques acting during pedalling. In this model, the inertial mass of the cyclist and the bicycle is (partly) realised by the flywheel in the rear part of an ergometer. Assuming that the rear and frontal sprockets are connected with a chain which is not elastic and does not slip, and that the frictional forces between the chain and the sprockets are small, the physical relation between the torques τ_{net} and τ_{brake}, and the crank angular acceleration α_{crank}, is given by Equation 9. In this equation, I_{flywheel} and I_{crank} represent the moments of inertia of the flywheel and the crank, $R_{\text{fs}}/R_{\text{rs}}$ is the ratio of the radii of the chain wheel and the rear sprocket, i.e., the gear ratio, and L_{c} is the length of the crank,

$$\tau_{\text{net}} = \left(I_{\text{crank}} + \left(\frac{R_{\text{fs}}}{R_{\text{rs}}} \right)^2 I_{\text{flywheel}} \right) \cdot \alpha_{\text{crank}} + \tau_{\text{brake}} . \qquad (9)$$

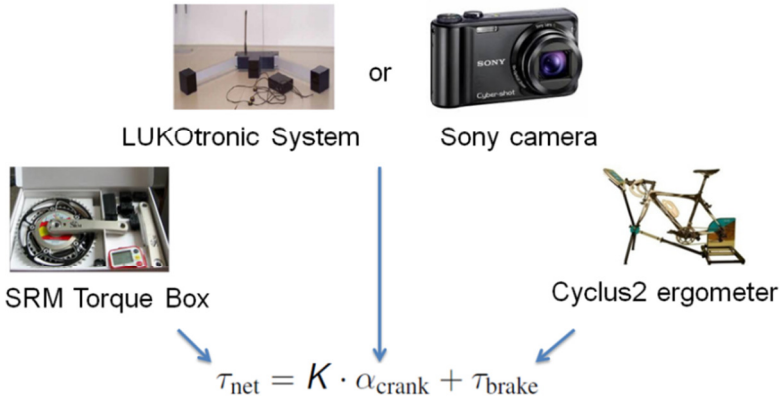


Figure 8 Equipment used for the experiment and their relation with the equation to validate. The angular crank acceleration was calculated from the data captured by the LUKOtronic System and the Sony camera.

Thus, the net torque is an (affine) linear function of the crank angular acceleration,

$$\tau_{\text{net}} = K \cdot \alpha_{\text{crank}} + \tau_{\text{brake}}, \quad (10)$$

with

$$K = \left(I_{\text{crank}} + \left(\frac{R_{\text{fs}}}{R_{\text{rs}}} \right)^2 I_{\text{flywheel}} \right). \quad (11)$$

Torque validation

Our goal was to determine to what extent net torques computed using Equation 9, based on numerically approximated crank acceleration from pedal motion measurements, are close to the directly measured (assumed to be true) net torques. For this purpose, we recorded the data simultaneously obtained from the motion capture system (i.e., MoCap data), the video camera recorder (i.e., video data), the bicycle ergometer, and the torque sensor, see Figure 8.

For the validation, we assumed that the angular crank acceleration α_{crank} is equal to the second derivative of angular position of the pivot of rotation of the pedal θ. To calculate the right-hand side of Equation 9, we used a crank length $L_{\text{c}} = 176$ mm, and moments of inertia

$I_{\text{flywheel}} = 0.6576\,\text{kg} \cdot \text{m}^2$ and $I_{\text{crank}} = 0.02\text{kg} \cdot \text{m}^2$. The gear ratio $R_{\text{fs}}/R_{\text{rs}}$ was calculated directly from the number of teeth on the gears in the gear train with $R_{\text{fs}} = 50$ and $R_{\text{rs}} = 13$. The inertia of the crank is composed of the inertia of the pedals, of the crank arm, of the chainrings, and of the SRM Power Meter. Each component can be approximated by a primitive geometric form rotating around the center of the crank axis (pedal - point mass, crank arm - solid cylinder, chainrings - rings, SRM - solid disc). We measured the weights and the sizes of each component and computed the individual moments of inertia. Their sum yields the total inertia of the crank $I_{\text{crank}} = 0.02\,\text{kg} \cdot \text{m}^2$, see more details in Dahmen and Saupe (2011). The moment of inertia of the flywheel was found empirically under the assumption that the friction force is an affine function of the velocity, which was verified through a fitting procedure. For this, we pedaled to accelerate the flywheel until it had a high speed. Then, we stopped pedalling and we counted the time until the flywheel did not move anymore.

The pipeline for calculating angular acceleration from positional data of the pedal is shown in Figure 9. The input is either the 3D MoCap data projected onto a two-dimensional plane using Principal Component Analysis (PCA), described in detail by Smith (2002), or the 2D

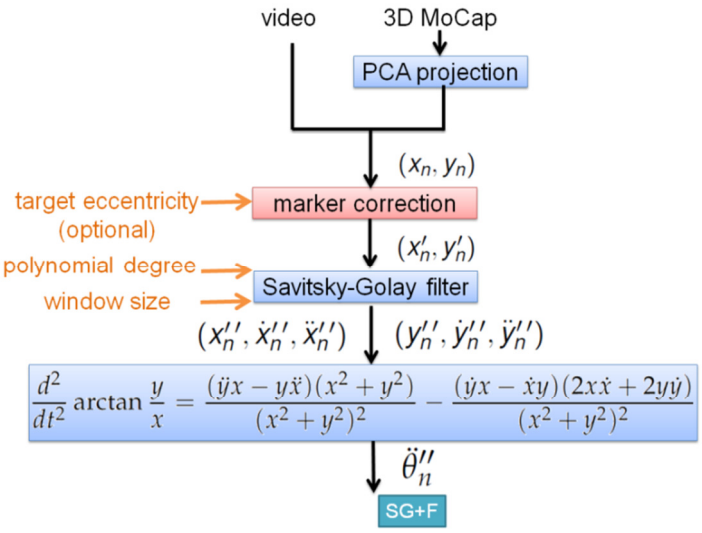


Figure 9 Pipeline for calculating angular crank acceleration from positional data. Parameters are given in orange. We assumed that the angular acceleration of the crank α_{crank} is equal to $\ddot{\theta}_n$.

positional video data. The correction of the marker position, explained in page 38, is a part of the preprocessing step (red module in Figure 9). We compared the results without this correction and with it based on two different methods, i.e., minimising the RMS distance between a fitted ellipse and the corrected data among all possible corrections and fitted ellipses, and with prescribed eccentricities. We tested the approach SG+F shown in Figure 9 consisted of the filtering and the calculation the zero, first and second derivatives of x and y coordinate sequences using the corresponding Savitzky-Golay filters for each case, and with this data then calculating the second derivative of θ using Equation 1.

Finally, the effects of correction of the marker position and filtering with different parameters were compared using the signal to noise ratio (SNR), where the true signal was assumed to be the SRM Torque Box data, and the noise was assumed to be the difference between the SRM Torque Box data and the right side of Equation 9.

Experimental setup

Data acquisition and processing

We recorded the pedalling motion using simultaneously a commercial camera (Sony Optical Steadyshot DSC–H55) and a motion capture system (LUKOtronic-Steinbichler Optotechnik GmbH, Neubeuern, Germany). The camera recorded videos with a sampling frequency of 29.97 Hz and a resolution of 1280×720 pixels. The video camera was placed at a distance of 270 cm from the bicycle.

The motion capture system consisted of a beam with three integrated infrared cameras. As a result of the fixed positions of the cameras within the beam, the system was pre-calibrated and did not require any separate calibration procedure. The motion capture system provided three-dimensional positional data of the infrared active markers. The sampling rate was 240 Hz. The positional accuracy provided by LUKOtronic for the distance used in our experiments was 1– 3 mm. Experimentally, the positional accuracy was 1.5 mm with a precision of 0.9 mm calculated from two markers fixed on the pedal during 180 crank rotations.

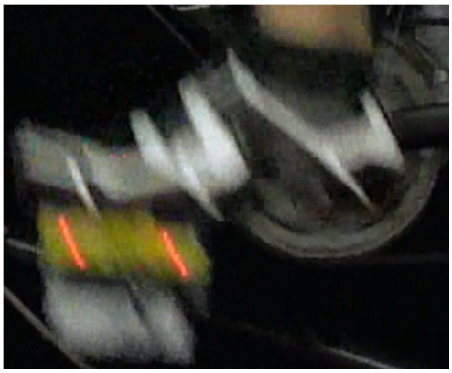


Figure 10 Example of LED trajectories as red streaks on yellow pedal in one video frame.

Using the motion capture system we recorded the position of two markers as shown in Figure 1. One marker was placed on the pivot point of the pedal rotation (i.e., a point representing the angular position of the crank during the motion of the pedal) and the other marker was placed on the right side next to the first marker. During the recording, we used three reference markers attached to the bicycle frame in order to define the coordinate system, which was re-calculated for each measurement. This has the advantage that oscillations and deflections of the bicycle frame due to the pedalling motion and bike sway are taken into account during the recording of the pedal motion. The pedal orientation for the motion data was calculated from the positional data of both markers with respect to the horizontal axis. We projected the three-dimensional coordinates of the marker position to a two-dimensional plane using the Principal Component Analysis (PCA). Then, we rotated the coordinate system using the information of additional measurements of the pedal in the lowest position.

Using the video camera we recorded two red LEDs placed on both sides of the pedal. See Figure 1. Each LED in a frame was captured as a streak due to blurring by the rapid movement of the pedal during the exposure time of the camera. See Figure 10 for an example. The LEDs were placed with enough space between them to avoid a possible streak overlapping. The coordinates of the heads of each streak were used in each frame as positional data of LEDs. The pivot point of the pedal was calculated as the average of both LED positions and the pedal orientation was calculated from the angle of the line between both LEDs with respect to the horizontal axis of the frame. For the detection of the streaks, each frame was converted from RGB format (i.e., red, green, and blue channels) to HSV format (i.e., hue H, saturation S, and value V). Then, a threshold on the value V was applied to find the blobs corresponding to the LED trajectories.

To control the pedal brake force, we used the Cyclus2 ergometer (RBM Elektronik-Automation GmbH, Leipzig, Germany). During the ergometer operation, the flywheel on the rear part is supplied with kinetic energy which maintains the angular momentum. The angular speed of the flywheel is decelerated by means of an eddy current brake. The eddy current brake guarantees a non-slipping transmission of the braking resistance. Operating the Cyclus2 in pedal force mode, a constant pedal brake force ($\pm 5\%$ error) is imposed.

Our SRM Torque Box (Schoberer Rad Messtechnik, Welldorf, Germany) gave an instantaneous torque signal with a sampling rate of 200 Hz. This torque corresponds to the net torque, as explained in Equation 8. Furthermore, when the crank has completed one pedal revolution (i.e., when the crank has crossed the sensor of the SRM Torque Box attached to the bicycle frame), this event was reported. SRM claims an accuracy of 2% for power and torque measurements.

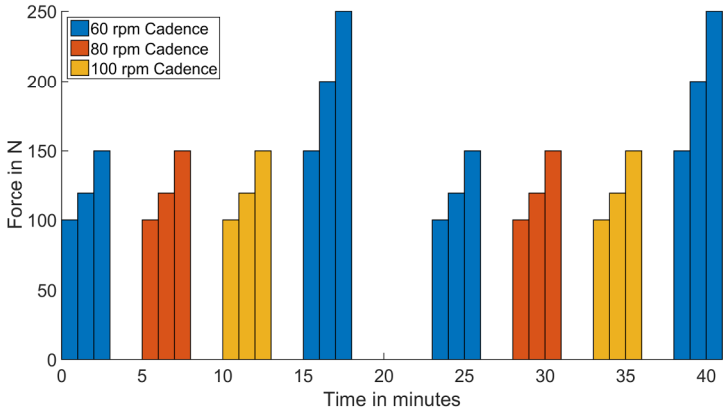


Figure 11 Experimental protocol. Eight test of three minutes each are performed with breaks of 2 resp. 5 minutes in between.

The laboratory conditions were optimized for recording the pedal motion: we darkened the room (i.e., we covered the window to avoid sunlight), used low infrared emitting light bulbs, and a special carpet to avoid infrared reflections from the ground.

Before the validation of the physical relation between the angular acceleration of the crank and the tangential force applied to the pedals, using Equation 9, we preprocessed the data obtained from all devices (see Figure 8). First, we applied linear interpolation when there was loss of data or outliers. Then, we resampled all data to 200 Hz. The data from all devices were registered using the SRM Torque Box data as reference. We used the information of the angular position of the crank when it crossed the sensor of SRM Torque Box attached to the bicycle frame in order to align the MoCap data and video with the SRM Torque Box data. The SRM Torque Box data and the Cyclus2 data were aligned using the time stamps provided by our system.

The correction of marker positions was performed using a quasi-Newton strategy minimizing either the difference between the prescribed eccentricity and the eccentricity of an ellipse fitted on the corrected positional data or the RMS distance between a fitted ellipse and the corrected data among all possible corrections and fitted ellipses. The ellipse fitting is performed by minimizing the squared sum of orthogonal distances from the points to the fitted ellipse described in Gander, Golub, and Strebel (1994).

Test design

Ten cyclists participated in this study (male, 31.45 ± 9.9 years). Each participant was asked to ride at a fixed cadence using his preferred technique and cycling shoes with cleats locked onto the pedal interface. A continuous feedback of the cadence was given on a projected video image, positioned in front of the cyclist. Before the measurements cyclists performed a warm up session of 5 minutes at a power output of 100 – 140 W.

Each subject rode the bike simulator performing eight tests of three minutes each. In each test the cadence was held constant and the brake force increased every minute. For the first three tests the cadences were 60 rpm, 80 rpm, and 100 rpm, and for each test the brake forces were 100 N, 120 N, and 150 N (i.e., brake torques of 17.6 Nm, 21.1 Nm and 26.4 Nm). For the fourth test, the cadence was again 60 rpm but with forces of 150 N, 200 N, and 250 N (i.e., brake torques of 26.4 Nm, 35.2 Nm and 44.0 Nm). This last test was done to check the effects of large forces on the SNR results which can vary due to crank deformation and brake force

fluctuation. All four tests were repeated once after a pause of 5 minutes. See Figure 11 for an illustration.

Results and discussion

Our goal was to determine to what extent net torques computed using Equation 9, based on numerically approximated crank acceleration from pedal motion measurements, are close to the directly measured (assumed to be true) net torques. For an example of the resulting torques, see Figure 12. We considered the following parameters of the Savitzky-Golay filter:

polynomial degrees $(2, 3)$ and window sizes $(10n + 1, n = 3, \dots, 22)$. We applied the Savitzky-Golay filter to the resampled data at 200 Hz. We also considered the results with and without the correction of marker positions, i.e., minimising the error between the fitted ellipse and corrected positional data (i.e., the minimum error eccentricity method) and with prescribed eccentricities $(e = 0.05i, i = 0, \dots, 6)$. See page 38 for a description of the methods for correction of marker position. Thus, we applied altogether for each test 360 different combinations of parameters and compared their SNR results.

Table 1 shows the average of the best SNR results for MoCap and video of all tests with and without marker correction among all parameter combinations. The SNR results with a marker correction and eccentricity 0 (i.e., a perfect circle) were the lowest for both MoCap and video data (see Table 1, column 4). This confirms that the crank and the bearings are elastic to a small degree and allow for some play.

Table 2 shows the average results for each test without marker correction and with marker correction based on M.E. for each combination of force and cadence. The marker correction based on M.E. improved the SNR results of the video data. These results were expected because in our video data recordings the position of the midpoint between LEDs cannot be precisely at the pivot point. The correction of the marker position based on M.E. did not improve the SNR results of the MoCap data. This indicates that the active infrared marker of the motion capture device was well placed on the pedal pivot point.

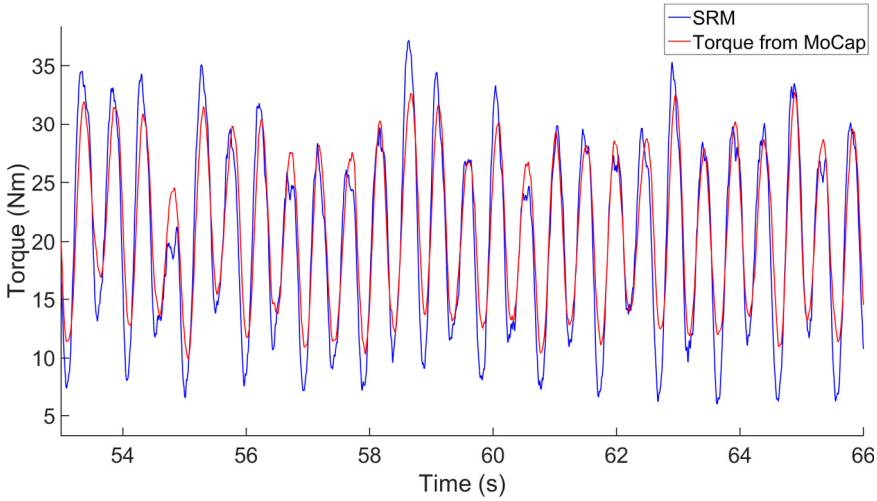


Figure 12 Result after alignment of SRM Torque Box data and MoCap data for the validation of the torque equation for the test with 60 rpm and 120 N (21.1 Nm) using the SG+F approach with window size 121, polynomial degree 2, and marker correction based on M.E.

Table 1 Average SNR for MoCap and video. Columns correspond to the type of input data, the SNR results without marker correction, the SNR results with M.E., and the results with different prescribed eccentricities for marker correction, respectively.

type	SNR	SNR e =M.E.	SNR $e = 0$	SNR $e = 0.15$	SNR $e = 0.2$	SNR $e = 0.25$	SNR $e = 0.3$
MoCap	14.82±1.86	14.67±1.75	9.74±1.97	13.28±2.04	14.74±1.94	15.12±1.96	14.80±2.11
video	10.81±2.11	12.30±2.47	9.61±2.18	11.60±1.84	13.42±1.75	14.02±1.78	14.18±2.11

Table 2 Average SNR of each test. The table shows the results for each combination of force and cadence. Columns correspond to brake force $\mathbf{F}_{\mathrm{brake}}$, cadence, the SNR results without marker correction and the results with marker correction based on M.E., respectively.

test		MoCap		video	
$\mathbf{F}_{\mathrm{brake}}$ N	cadence rpm	SNR $e = 0$	SNR e =M.E.	SNR $e = 0$	SNR e =M.E.
100	60	15.22±1.24	15.27±1.27	11.22±1.90	13.13±2.13
100	80	13.55±1.29	13.34±0.89	10.24±1.41	11.03±1.65
100	100	12.39±0.98	12.30±1.00	9.74±1.75	10.90±2.04
120	60	16.28±1.46	16.22±1.40	12.18±2.57	14.01±2.89
120	80	14.71±1.18	14.49±0.80	12.03±2.43	12.27±2.22
120	100	13.88±1.44	13.95±1.37	10.31±1.89	12.08±2.68
150	60	17.16±1.31	16.88±1.21	10.94±1.99	13.63±2.86
150	80	15.20±1.24	14.85±1.30	11.34±2.10	12.42±1.86
150	100	14.96±1.70	14.69±1.41	9.29±1.54	11.20±2.33
200	60	17.10±2.06	17.19±1.78	11.85±2.33	15.18±3.51
250	60	17.55±2.77	17.30±2.26	10.76±1.63	14.79±2.67

Table 3 Parameters for the best average SNR of the SG+F approach with marker correction based in two different methods. Columns correspond to type of data, brake force $\mathbf{F}_{\mathrm{brake}}$, cadence, polynomial degree and window size of the Savitsky-Golay filter employed for the method with prescribed eccentricities e, polynomial degree and window size of the Savitsky-Golay filter employed for the M.E. method, respectively.

			prescribed eccentricity			M.E.	
type	$\mathbf{F}_{\mathrm{brake}}$ N	cadence rpm	polynomial degree	window size	eccentricity	polynomial degree	window size
MoCap	100–150	60–100	2	121	0.25	2	91
MoCap	200	60	2	121	0.3	2	81
MoCap	250	60	3	191	0.2	3	191
video	100–150	60–100	2	121	0.25	2	81
video	200	60	3	171	0.25	3	201
video	250	60	3	181	0.2	3	201

The SNR results using the marker correction based on M.E. were lower than those obtained when fitting to ellipse with prescribed eccentricities. For example, the marker correction based on prescribed eccentricity $e = 0.25$ improved the SNR results for video ($+3.2$ dB) and slightly for MoCap ($+0.3$ dB), see Table 1. Figure 13 shows the effects on SNR results with different combinations of parameters of Savitsky-Golay filters and prescribed eccentricities for the tests with forces 100 N -150 N. The best results for both MoCap and video data were obtained with prescribed eccentricity $e = 0.25$, window size 121, and polynomial degree 2. For forces larger than 150 N another set of parameters provided the best results. These parameters are given in Table 3.

The deformation of trajectory of the point representing the angular position of the crank for the bicycle is difficult to obtain but possible, e.g., with torsional strain gauges at the crank. However, the strain gauges are the main component of several commercial torque meters that would allow direct toque measurements. In any case, the marker correction based on M.E. can improve the SNR results for the proposed video-based indirect torque measurements.

The differences between the torque based on the angular acceleration (obtained from the MoCap and video data) and the measured torque with SRM Torque Box might have arisen from the eddy current brake of the bicycle simulator. This brake is less than ideal, so that fluctuations in the pedal brake force could occur during our test. Furthermore, we assumed that the chain, the crank and the pedal are completely inelastic and non-slipping, and that the frictional forces between the chain and the gears is negligible compared with the pedal brake force produced by the eddy current brake of the Cyclus2 ergometer.

In addition, for the calculation of the crank acceleration it was assumed that the motion of the marker on the pedal pivot point lies in a two dimensional plane. This plane can be easily calculated from MoCap data projecting the 3D data onto 2D by using Principal Component Analysis (PCA) but with video data an additional reference to align the camera plane with the pedal motion plane is necessary to have better results. The positional data extracted from each video frame corresponded to the 2D projection of the crank motion plane to the camera plane. If these planes are not aligned, the projection of the circular trajectory of the crank motion to the camera plane results in an oval trajectory. Such misalignment of both planes may happen during the pedalling motion due to bike sway. This undesirable distortion could be minimized in each frame by a perspective correction based on some reference marker points on the bike frame.

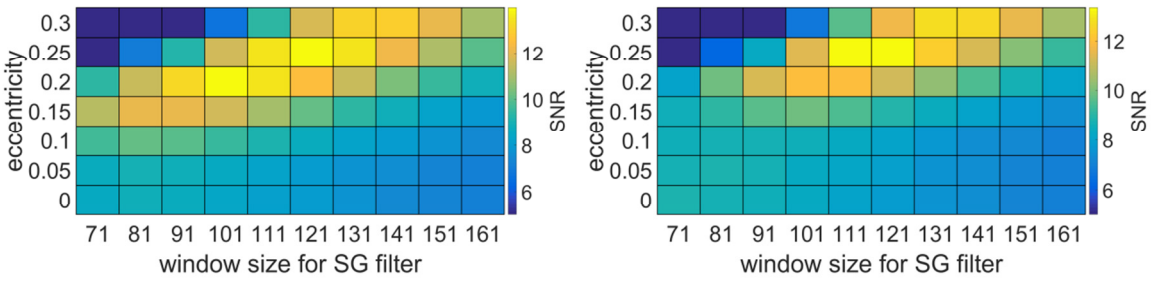


Figure 13 Mean SNR result of MoCap data (left) and video data (right) with correction of the marker position and different window sizes for the SG filter with degree 2 polynomials in the SG+F approach.

Conclusions

The physical relation between the angular acceleration of the crank and the tangential force applied to the pedals was validated with measured data of the net torque, brake force, video and MoCap data. These variables are related by an affine linear equation. The variation of the angular acceleration is proportional to the force variation when the pedal brake force is constant.

The proposed methods provided a valid calculation of the variation of the pedal acceleration (i.e., the torque variation). Our results show that optical motion tracking of the crank rotation can be used to estimate the tangential force variation applied to the pedal. The results obtained using a commercial video camera were close to the results achieved with a motion capture system, when a correction of the marker position was applied. Thus, an expensive optical device is not necessary to estimate the torque variation. We proposed values of the parameters for the relaxation and the Savitsky-Golay filter that can be used for pedal brake forces ranging between 100 and 250 N with cadences between 60 and 100 rpm for MoCap and video data. The training of a particular pedalling technique can be performed by providing athletes real-time performance feedback based on our proposed calculation of pedal torque variation. The only prerequisite is the availability of an ergometer delivering an adjustable constant pedal brake force, two LEDs on the pedal and a video camera.

For future research we intend to further improve our results with high speed camcorders (e.g., 59.94 fps) with a high image resolution. Furthermore, we plan to compensate the distortion due to bike sway performing a perspective correction with parameters calculated from the distortion of a reference attached to the bike frame.

Acknowledgments

This work was partially supported by the DFG Research Training Group GRK-1042 ”Explorative Analysis and Visualization of Large Information Spaces” and by the DFG project SA 449/12-1 “Powerbike: Model-based optimal control for cycling”[2] .

References

Bini, R. R., Diefenthaeler, F., & Mota, C. B. (2010). Fatigue effects on the coordinative pattern during cycling: Kinetics and kinematics evaluation. *Journal of Electromyography and Kinesiology*, *20*(1), 102–107.

Black, A. H. (1994). *The effect of steady rate exercise on the pattern of force production of the lower limbs in cycling*: University of British Columbia, Canada.

Böhm, H., Siebert, S., & Walsh, M. (2008). Effects of short-term training using SmartCranks on cycle work distribution and power output during cycling. *European Journal of Applied Physiology*, *103*(2), 225–232.

Cannon, D. T., Kolkhorst, F. W., & Cipriani, D. J. (2007). Effect of pedaling technique on muscle activity and cycling efficiency. *European Journal of Applied Physiology*, *99*(6), 659–664.

[2] https://www.informatik.uni-konstanz.de/saupe/forschung/laufende-projekte/powerbike/

Carpes, F. P., Rossato, M., Faria, I. E., & Bolli Mota, C. (2007). Bilateral pedaling asymmetry during a simulated 40-km cycling time-trial. *Journal of Sports Medicine and Physical Fitness*, *47*(1), 51–57.

Christensen, L., Johannsen, P., Sinkjær, T., Petersen, N., Pyndt, H. S., & Nielsen, J. B. (2000). Cerebral activation during bicycle movements in man. *Experimental Brain Research*, *135*(1), 66–72.

Dahmen, T., & Saupe, D. (2011). Calibration of a Power-Speed-Model for Road Cycling Using Real Power and Height Data. *International Journal of Computer Science in Sport*, *10*(2), 18–36.

Ettema, G., & Loraas, H. W. (2009). Efficiency in cycling: A review. *European Journal of Applied Physiology*, *106*(1), 1–14.

Faria, M. (2009). Recent advances in specific training for cycling. *International Sport Medicine Journal*, *10*(1), 16–32.

Höchtl, F., Böhm, H., & Senner, V. (2010). Prediction of energy efficient pedal forces in cycling using musculoskeletal simulation models. *Procedia Engineering*, *2*(2), 3211–3215.

Gander, W., Golub, G. H., & Strebel, R. (1994). Least-squares fitting of circles and ellipses. *BIT Numerical Mathematics*, *34*(4), 558–578.

Hamming, R. W. (1989). *Digital Filters*. *Dover Civil and Mechanical Engineering Series*: Dover Publications.

Hug, F., Turpin, N. A., Guével, A., & Dorel, S. (2010). Is interindividual variability of EMG patterns in trained cyclists related to different muscle synergies? *Journal of Applied Physiology*, *108*(6), 1727–1736.

Hull, M. L., Kautz, S., & Beard, A. (1991). An angular velocity profile in cycling derived from mechanical energy analysis. *Journal of Biomechanics*, *24*(7), 577–586.

Jirsa, V. K., Fink, P., Foo, P., & Kelso, J. A. (2000). Parametric stabilization of biological coordination: A theoretical model. *Journal of Biological Physics*, *26*(2), 85–112.

Kautz, S. A., & Hull, M. L. (1993). A theoretical basis for interpreting the force applied to the pedal in cycling. *Journal of Biomechanics*, *26*(2), 155–165.

Korff, T., Romer, L. M., Mayhew, I., & Martin, J. C. (2007). Effect of pedaling technique on mechanical effectiveness and efficiency in cyclists. *Medicine and Science in Sports and Exercise*, *39*(6), 991–995.

Mornieux, G., Zameziati, K., Mutter, E., Bonnefoy, R., & Belli, A. (2006). A cycle ergometer mounted on a standard force platform for three-dimensional pedal forces measurement during cycling. *Journal of Biomechanics*, *39*(7), 1296–1303.

Orfanidis, S. J. (2010). *Introduction to Signal Processing*. *Prentice Hall international editions*: Prentice Hall.

Ovaska, S. J., & Valiviita, S. (1998). Angular acceleration measurement: A review. *IEEE Transactions on Instrumentation and Measurement*, *47*(5), 1211–1217.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (2007). *Numerical Recipes 3rd Edition: The Art of Scientific Computing* (3rd ed.): Cambridge University Press.

Rossato, M., Bini, R. R., Carpes, F. P., Diefenthaeler, F., & Moro, A. (2008). Cadence and workload effects on pedaling technique of well-trained cyclists. *International Journal of Sports Medicine*, *29*(9), 746–752.

Sanderson, D. J. (1990). The influence of cadence and power output on asymmetry of force application during steady-rate cycling. *Journal of Human Movement Studies*, *19*, 1–9.

Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, *36*(8), 1627–1639. doi:10.1021/ac60214a047

Schafer, R. W. (2011). What is a Savitzky-Golay filter? *Signal Processing Magazine, IEEE*, *28*(4), 111–117.

Smak, W., Neptune, R. R., & Hull, M. L. (1999). The influence of pedaling rate on bilateral asymmetry in cycling. *Journal of Biomechanics*, *32*(9), 899–906.

Stapelfeldt, B., Mornieux, G., & Gollhofer, A. (2006). Wirkung von Feedback-Training im Radsport auf physiologische und biomechanische Parameter. *BISp-Jahrbuch Forschungsförderung*, *7*, 175–179.

Stapelfeldt, B., Mornieux, G., Oberheim, R., Belli, A., & Gollhofer, A. (2007). Development and evaluation of a new bicycle instrument for measurements of pedal forces and power output in cycling. *International Journal of Sports Medicine*, *28*(4), 326–332.

Theurel, J., Crepin, M., Foissac, M., & Temprado, J. J. (2011). Effects of different pedalling techniques on muscle fatigue and mechanical efficiency during prolonged cycling. *Scandinavian Journal of Medicine & Science in Sports*, *22*(6), 714–721.

Reconstructing the 3D Trajectory of a Ball with Unsynchronized Cameras

Tamaki, S., Saito, H.

Graduate School of Science and Technology, Keio University

Abstract

The 3D trajectory of the ball is one of the most important performance indicators in ball games. However, it often cannot be measured because of the requirements: synchronized cameras and control points within the playing field. Thus, this paper proposes a method that reconstructs the 3D trajectory of a ball with unsynchronized cameras. The proposed method consists of ball detection, camera calibration, and trajectory reconstruction. At first, ball candidates are detected on the basis of their appearance. Balls are then extracted from the candidates on the basis of their motion. After the extraction, ball trajectories are reconstructed in image spaces. The corresponding points in two views are estimated on the basis of the ball trajectory in image spaces and the temporal offset between cameras, which is supposed to be recorded with a few erroneous frames. The matrix including the geometrical relationship between cameras, namely the fundamental matrix, is then computed from the corresponding points. The estimated one is inaccurate due to the error of the temporal offset. The key feature of this method is to optimize the temporal offset and fundamental matrix simultaneously. After the optimization, the geometrical relationship between cameras is computed from the matrix. Balls are extracted again by using the temporal offsets and the geometrical relationship between cameras. A ball trajectory is finally reconstructed as connected trajectories that are separated at collisions. It is experimentally demonstrated that the proposed method accurately calibrates cameras and successfully reconstructs the 3D trajectory of the ball.

KEYWORDS: UNSYNCHRONIZED CAMERA, CALIBRATION, TRAJECTORY, TABLE TENNIS, BALL

Introduction

Ball trajectory is an important performance indicator in ball games. For example, baseball pitchers throw a ball in various trajectories to confuse batters and make it difficult to hit the ball. These kinds of skills and tactics can be analyzed if we know the ball trajectories. In net sports, such as tennis or table tennis, the positions of impacts and bounces are essential to analyze the spatial tactics, and those data can be extracted from ball trajectories. The velocity of the ball can be computed from a trajectory if it was reconstructed as a temporal function. If the 3D trajectory of a ball is easy to reconstruct, and if we can reconstruct it in a lot of matches, performance analysis in ball games will be improved and the temporal and spatial understandings of the ball games will be advanced.

However, ball trajectory is, unfortunately, difficult to reconstruct in practical scenarios. The common method of 3D measurement requires synchronized cameras. We need cameras that

can be synchronized to other cameras and a signal generator and must connect them with cables. In addition to that, the control points, *i.e.* the points whose 3D coordinates are known, are required within the playing field to compute the geometrical relationship between cameras (hereinafter referred to as "geometrical calibration"). There are many scenarios where those conditions are not met in sports competitions and training. Although the 3D trajectory of a ball would be helpful, it usually cannot be measured due to the above requirements.

Antoku, Kashima, Sato, & Watanabe (2013) measured the trajectories of a tennis ball and the position of sportspeople in image space with a camera and recorded the stroke style, *i.e.* forehand or backhand, automatically. Kitahara & Uchida (2013) measured the positions of bounces and players in the reference frame on the court with two unsynchronized camera. The 2D positional data of a ball and players were adopted in these cases to make the system easy to use. However, these approaches might be inaccurate when the optical axis of a camera is close to parallel to the court plane, which is inevitable. The curvedness of the trajectory cannot be analyzed as long as 2D data is used. The method for sports analysis should be able to reconstruct the 3D trajectory of a ball.

Takanohashi, Manabe, Yasumuro, Imura, & Chihara (2007) measured 2D ball trajectories by using motion blur and reconstructed the 3D shape of ball trajectory by a shape-from-silhouette method, *i.e.* projecting the image segments to the 3D space from multiple viewpoints and computing the intersection of them. The 3D trajectory of a ball can be reconstructed even though the cameras are unsynchronized. Their method, however, is unstable outside of an ideal environment because they made motion blur intentionally by lengthening exposure time and the ball merges into the background easily and sometimes becomes invisible. They did not propose any methods that solve these problem, so their method might be inappropriate to employ.

Shimizu, Fujiyoshi, Nagasaka, Takahashi, & Iwahori (2009) computed the intersection of the epipolar line (see "Projective Geometry") and the ball trajectory in image spaces as a pseudo corresponding point between different viewpoints and computed 3D coordinates of balls from them. In their method, the 3D coordinates of a ball can be computed even if the cameras are unsynchronized, panned, and tilted after geometrical calibration. The accuracy of their method, however, reduces when the ball trajectory and the epipolar line become close to parallel. This often happens when we shoot a match of net sports because cameras are usually put horizontally on one side of the court to reduce the spatial occlusion between players and a ball. Thus, their method is inappropriate to employ to analyze net sports.

Matsumoto, Sato, & Sakaue (2010) focused on the geometrical relationship of multiple unsynchronized affine cameras in the frequency domain and demonstrated that the points in 3D space can be reconstructed by reconstructing the point in the frequency domain even if there is no corresponding point in images. It is, however, inappropriate to approximate cameras as affine when the cameras are placed close to the playing field, which often happens. Kakumu, Sakaue, Sato, Ishimaru, & Imanishi (2013) demonstrated that the theory proposed by Matsumoto et al. can be applied to the projective camera when the optical axes of the cameras are parallel. From the viewpoint of the accuracy, cameras should be aligned so that their optical axes are orthogonal. Therefore, these methods are inappropriate for 3D measurements in sports.

Zhou & Tao (2003) estimated temporal offset between cameras (hereinafter referred to as "temporal calibration") using a set of four corresponding points in different viewpoints. Pseudo synchronized image pairs were synthesized on the basis of the estimated temporal

offset and optimal flow. This method shows an approach to estimate temporal offsets between unsynchronized cameras and seems useful for the 3D measurement in sports. However, this method approximates objects' motion as linear, which will lead to the wrong result when the object is a traveling ball because it can move fast and curve. In addition to that, they did not provide a solution for the geometrical calibration.

Noguchi & Kato (2006) developed an approach for geometrical and temporal calibration using a marker. In their method, the trajectories of the marker are reconstructed at first. Corresponding points are then estimated on the basis of the roughly estimated temporal offset between cameras. Finally, temporal offset and geometrical relationship between cameras are optimized simultaneously on the basis of the error of the fundamental matrix (see "Projective Geometry"). If a ball could replace a marker, this method would be useful for the 3D measurement of ball trajectory. However, a ball cannot simply replace a marker. In their method, a marker needs to move slowly because they linearly interpolate the trajectory of the markers between two consecutive frames. This limitation is inappropriate when the target is a ball. In addition to that, they did not deal with false detections because a marker can be detected robustly. To apply this method to the 3D measurement of ball trajectory, the above problems need to be solved.

In this paper, a method is proposed that reconstructs the 3D trajectory of a ball with unsynchronized cameras. In this method, balls are detected on the basis of their appearance and motion at first. Cameras are then geometrically and temporally calibrated by using the detected balls. The camera calibration is based on Noguchi and Kato's method but the ball trajectory estimation and the error evaluation algorithms are modified for dealing with ball. Balls are detected again by using the temporal offset and the geometrical relationship between cameras. At the end, the 3D trajectory of a ball is reconstructed. As we can see from the description above, the proposed method is a framework that contains the entire process of ball trajectory reconstruction with unsynchronized cameras. This is the frst research which realizes camera calibration using balls and ball trajectory reconstruction without control points or camera synchronization. The 3D trajectory of the ball will be easy to reconstruct because the calibrated cameras and control points within playing fields will no longer be required due to the proposed method. Although table tennis is focused on in this paper, the proposed method can be applied to all sports where a ball travels in the air, such as tennis and volleyball.

The paper consists of the following sections. "Projective Geometry" briefly describes projective geometry. "Methods" describes the proposed method in detail. "Experiment" experimentally evaluates the proposed method. Finally, "Conclusion" concludes the paper.

Projective Geometry

This section describes the basic theories of projective geometry on which the proposed method is based. More detailed information about projective geometry is described by Hartley & Zisserman (2003).

Camera Model

The camera modeled by perspective projection is called a projective camera. A projective camera can accurately approximate actual cameras and has been broadly used for 3D measurements. Figure 1(a) shows an imaging model of the projective camera. The relationship between the 3D coordinate $\boldsymbol{X}$ and the 2D coordinates in image space $\boldsymbol{x}$ is defined as the following equations.

$$\tilde{\boldsymbol{x}} = \boldsymbol{P}\tilde{\boldsymbol{X}}. \tag{1}$$

$$\mathbf{P} = \mathbf{K}[\mathbf{R}\,|\mathbf{t}]$$

$$= \begin{bmatrix} a_x & s & x_0 \\ & a_y & y_0 \\ & & 1 \end{bmatrix} \boldsymbol{R}\left[\boldsymbol{I} \mid -\tilde{\boldsymbol{C}}\right]. \tag{2}$$

where a_x is the scale factor in the x-coordinate direction, a_y is the scale factor in the y-coordinate direction, s is the skew, and x_0 and y_0 are the coordinates of the principal point. The above five parameters are called intrinsic parameters, which depend on the extrinsic environment, namely the position and the pose of the camera. In contrast to intrinsic parameters, camera position $\boldsymbol{C}$ and camera pose $\boldsymbol{R}$ are called extrinsic parameters. If the intrinsic parameters and extrinsic parameters of two cameras are known, 3D coordinates of corresponding points can be reconstructed by triangulation (Figure 1(b)).

Epipolar Geometry and Fundamental Matrix

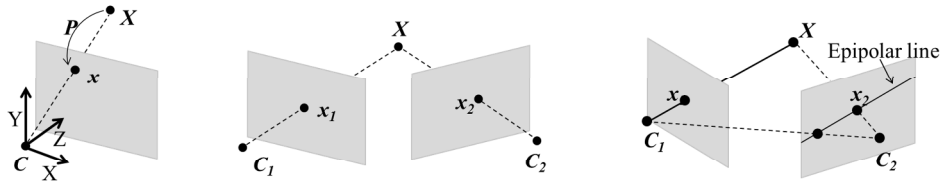
Let us denote the $\boldsymbol{X}$ in the two different image spaces as $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$, and the homogeneous coordinates of $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$ are $\tilde{\boldsymbol{x}}_1$ and $\tilde{\boldsymbol{x}}_2$. Their relationship is defined by the following equation.

$$\tilde{\boldsymbol{x}}_2{}^{\boldsymbol{T}}\boldsymbol{F}\tilde{\boldsymbol{x}}_1 = \boldsymbol{0}. \tag{3}$$

The matrix $\boldsymbol{F}$ contains the intrinsic parameters and relative geometrical relationship of two cameras. The matrix is called the "Fundamental Matrix". The line, which is called the "Epipolar Line", can be projected to the other image by the following equation (Figure 1(c)).

$$\tilde{\boldsymbol{l}} = \boldsymbol{F}\tilde{\boldsymbol{x}}_1. \tag{4}$$

The epipolar line is the set of points related to $\boldsymbol{x_1}$. If the intrinsic parameters are known, the geometrical relationship can be extracted from the fundamental matrix without the scale of translation. Computing the fundamental matrix, therefore, is sometimes regarded as geometrical calibration without scale. The scale of the translation can be computed from the line segments whose length in the real world is known.



(a) Camera model (b) Triangulation (c) Epipolar line

Figure 1. Projective geometry

Methods

Outline

Figure 2 shows the flow of the method proposed in this paper. The start and the end times of rallies are supposed to be known. The trajectory of a rally is reconstructed in the method. First, ball candidates are detected in two view images that are separately referred to as image 1 and

image 2. The cameras that took images 1 and 2 are referred to as cameras 1 and 2 hereinafter. If the two cameras are not calibrated temporally and geometrically, they are calibrated by using detected ball candidates. In the calibration, balls are extracted from the candidates on the basis of the motion. Ball trajectories are reconstructed after the calibration. In order to enhance the detectionrate, balls are extracted again using the geometrical and the temporal relationship between camras. The intrinsic camera parameters and the coefficient of lens distortion are supposed to be computed by the method proposed by Zhang (2000), which is easy to use because it is implemented as calibrateCamera in OpenCV, which is a popular software library in computer vision, or Matlab. The temporal offset between cameras is supposed to be roughly estimated by step-by-step playback. The coordinates of four corners of a table are supposed to be recorded manually and used to estimate the vertical direction in 3D space, calculate the actual length of a pixel, detect bounces, and convert the positions of the bounces and impacts in the coordination system on the court.

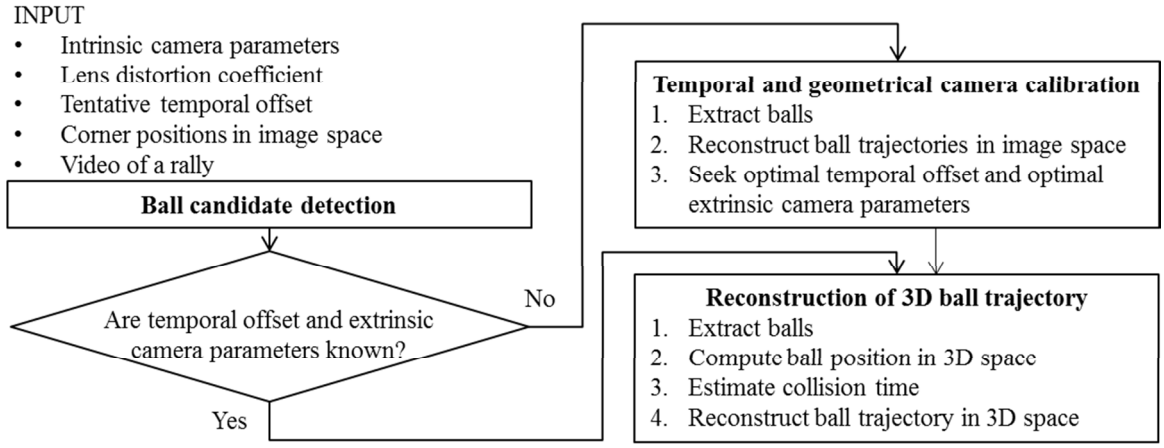


Figure 2. Flowchart of the proposed method

Ball Candidate Detection

The segments of moving objects are extracted by frame differentiation. First, we obtain the difference image between the images at time t and time t-s and the difference image between the images at time t and time t+s, where s denotes the arbitrary interval for difference. The logical conjunction of the each pixel in two difference images is then calculated. The value is true where the part of a moving segment is at time t. Figure 3 shows the segments of an impact scene extracted by frame differentiation, where s was 3 and the camera's frame rate was 60 fps. Because a ball travels faster than many other objects in the image, extracted segments can be reduced by setting a small difference interval.

Next, the segments that are similar to the ball were extracted from the moving objects on the basis of their color, circularity and size. The circularity C of a segment whose boundary length is l and the area is S is defined as follows:

$$C = 4\pi S/l^2. \tag{5}$$

The following thresholds are required in this process.

- Lower bound of the intensity [0-1.0]: L_I
- Lower bound of the circularity [0-1.0]: L_C
- Lower bound of the area [pixel^2]: L_A

- Upper bound of the are [pixel^2]: U_A

The detected segments are the ball candidates. Unfortunately, no methods can determine the true ball only on the basis of the appearance in a single view. Another process is necessary for extracting a true ball. The method for extraction is described in the following subsections because different approaches are used before and after the camera calibration.

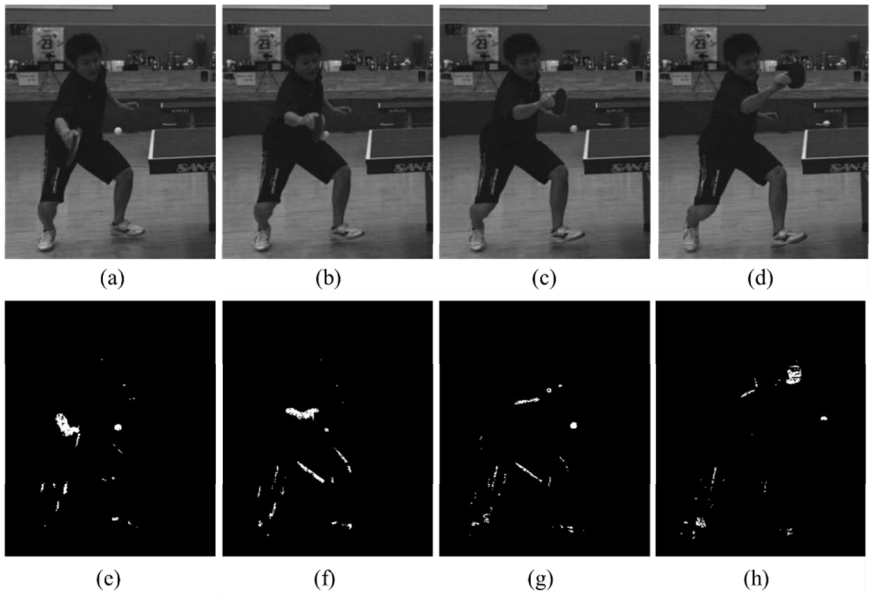


Figure 3. (a)-(d) are input images captured in 60 fps. (e)-(h) are results of frame differentiation with 3ms interval.

Temporal and Geometrical Camera Calibration

Ball Extraction

The candidates that move like a ball traveling in the air are extracted from the ball candidates. The motion is differentiated by the velocity, the difference from the adjacent velocities, and the angle from the adjacent traveling directions. The following thresholds are required in this process.

- Lower bound of the velocity [m/s]: L_V
- Upper bound of the velocity [m/s]: U_V
- Upper bound of the change rate between two consecutive travel distances [%]: U_D
- Upper bound of the angle between two consecutive travel directions [rad]: U_R

Note that L_V and U_V are converted into the velocity in the image coordination system L_v and U_v, respectively, on the basis of the actual length of a pixel, which is supposed to be computed from the corners of the court. Let us denote the position of a ball candidate in the *j*-th frame as $\boldsymbol{p_j}$ and the velocity vector from $\boldsymbol{p_j}$ to $\boldsymbol{p_{j+1}}$ as $\boldsymbol{v_{j,j+1}}$. $\boldsymbol{p_j}$ and $\boldsymbol{p_{j+1}}$ can be the same ball if the norm of $\boldsymbol{v_{j,j+1}}$ is greater than L_v and less than U_v. $\boldsymbol{p_j}$, $\boldsymbol{p_{j+1}}$, and $\boldsymbol{p_{j+2}}$ can be the same ball if the norm of $\boldsymbol{v_{j+1,j+2}}$ is greater than L_v and less than U_v and the change rate between $\boldsymbol{v_{j,j+1}}$ and $\boldsymbol{v_{j+1,j+2}}$ is less than U_D and the angle between them is less than U_R. If the above conditions are met up to $\boldsymbol{p_{j+4}}$, the consecutive five candidates are extracted as the true ball.

The true balls, however, will not be extracted perfectly even if this strong constraint is applied. A small amount of false detections can be extracted as well. This is the limitation of this

method and many others (Huang, Llach, & Zhang, 2008; Ishii, Kitahara, Kameda, Ohta, 2007; Yu, Sim, & Wang, 2004; Chen, & Zhang, 2006). We think it is too difficult to extract only the true ball from images. For that reason, the contamination of a small amount of false detections is assumed in the following processes.

Reconstruction of 2D ball trajectory

The ball trajectory in the image spaces is reconstructed by third-degree spline interpolation with the balls obtained so far. The trajectory is reconstructed only in the period of time when the position of the ball was recorded consecutively. Spline, or any other polynomial expression, cannot accurately represent abrupt change of the motion, such as bounces. This will not be a problem for camera calibration because the existence of outliers is initially supposed and they are handled appropriately in the following processes.

Estimation of temporal offset and extrinsic parameters

The temporal offset between cameras is supposed to be roughly estimated, as mentioned before. Let us denote the initial temporal offset between cameras 1 and 2 as τ_0 and the time of the *j*-th frame captured by camera *i* as t_j^i. If τ_0 is truth, t_j^1 equals $t_j^2 - \tau_0$. The ball position at t_j^i can be estimated only if a trajectory was reconstructed in the image *i* around t_j^i. This fact leads to the temporal offset providing many corresponding points between two views, so the fundamental matrix $\boldsymbol{F_0}$ can be computed (Hartley & Zisserman, 2003). The relative geometrical relationship between cameras can be extracted from the fundamental matrix. However, $\boldsymbol{F_0}$ is inaccurate due to the error of τ_0. Let us denote the point in a *j*-th frame captured by camera *i* as $\boldsymbol{p_j^i}$, the point corresponding to $\boldsymbol{p_j^1}$ as $\boldsymbol{p_j^{2'}}$, the homogeneous coordinates of $\boldsymbol{p}$ as $\widetilde{\boldsymbol{p}}$, the function that calculates the distance between a point and a line as d, and the function that calculates the median of a series of N elements as Med_i^N. The error of the fundamental matrix $\boldsymbol{F}$ is defined as the following equation:

$$\mathrm{E}(\boldsymbol{F}) = \mathrm{Med}_i^N \left(\mathrm{d}\left(\boldsymbol{p_j^1}, \boldsymbol{F}\boldsymbol{p_j^{2'}}\right) + \mathrm{d}\left(\boldsymbol{p_j^{2'}}, \boldsymbol{F^T}\boldsymbol{p_j^1}\right) \right). \quad (6)$$

$\boldsymbol{F}\boldsymbol{p_j^{2'}}$ is the epipolar line related to $\boldsymbol{p_j^{2'}}$ in image 1, and $\boldsymbol{F^T}\boldsymbol{p_j^1}$ is the epipolar line related to $\boldsymbol{p_j^1}$. E(F) is the median of the series of distances between an epipolar line and its corresponding point. This function returns a smaller value when a better fundamental matrix is input. Note that this function can be regarded as the error of a temporal offset because a fundamental matrix has a one-to-one relationship with a temporal offset. Figure 4 shows the value change of Equation 6 caused by temporal offset. This figure shows a better temporal offset leads to a smaller value. The best temporal offset, therefore, can be determined by seeking the temporal offset that makes Equation 6 smallest. However, some random fluctuationsoccur. If the temporal offset is far away from the truth, the geometrical correctness of corresponding points behaves almost randomly. and the error fluctuates strongly. Even if the temporal offset is close to the correct value, the fluctuation can occur because of the unstableness of the computation of the fundamental matrix. Fundamental matrix is calculated from corresponding points estimated using the temporal offset. Outliers are eliminated by RANSAC, and the fundamental matrix is computed from eight points that are randomly chosen from the corresponding points. This computation can return inaccurate results because the fundamental matrix can be degenerated if the corresponding points are not distributed uniformly in 3D space. The extraction of corresponding points and the computation of the fundamental matrix, therefore, are repeated $\mathrm{N_i}$ times to make sure some accurate results are obtained. The fundamental matrix

is finally determined by selecting the one that returns the smallest reprojection error. In this method, a fixed step approach is used for the seeking to avoid converging on a local minimum and achieve the constant accuracy within a specific bound.

The method described in this subsection is an extension of the method proposed by Noguchi & Kato (2006). One of the differences is the method for reconstruction of ball trajectory. Because a slowly moving marker is used in Noguchi & Kato's method, they reconstruct the trajectory linearly. On the other hand, a spline curve is used in the proposed method to reconstruct the curvy trajectory of a table tennis ball. Another difference is the error function. Noguchi & Kato employed the mean for computing the representative value of the error distribution. In the proposed method, the median is employed instead to deal with outliers because false detections can be included in the balls. The median is much more robust than the mean against the contamination of outliers. The proposed method will calibrate cameras more robustly than Noguchi & Kato's method when the detected balls contain false detections.

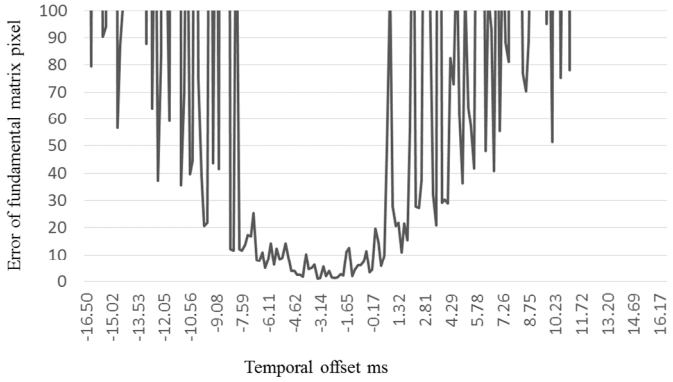


Figure 4. Error of fundamental matrix. The ground truth of temporal offset is -3.0 ms.

Reconstruction of the 3D Trajectory of a Ball

Ball Extraction

Balls are extracted again from the ball candidates by using epipolar lines. The following thresholds are required in this process.

- Upper bound of the distance between a epipolar line and a ball [m]: U_E
- Upper bound of the distance between a trajectory and its supports [m]: U_S

Note that U_E is converted into the distance in image space U_e before it is used. The pairs of candidates that satisfy the following conditions are extracted (Figure 5(a)): (1) one of which is captured at the first frame before $t_j^1 + \tau$, (2) the other is captured at the first frame after $t_j^1 + \tau$, (3) the distance between them is less than the product of U_v and the period of time between adjacent frames, i.e. maximum distance a ball can travel. The candidates of $\boldsymbol{p}_{t_j^1+\tau}^{\mathbf{1}}$ are then estimated linearly. The circles in Figure 5(a) denote the estimated candidates. The epipolar line related to $\boldsymbol{p}_{t_j^2}^{\mathbf{2}}$ is computed after that. The candidates whose distance from the epipolar line is less than U_e are extracted. The linearly estimated candidate, its two source candidates, and the source candidate of epipolar line are stored and passed through the first filter. Many false positives are eliminated through the process. Because the linearly estimated candidates are used only for the ball extraction, their errors do not influence the final measurement accuracy

at all.

The 3D positions of the corresponding points, which were obtained previously, are calculated by triangulation. Because most false balls are far away from the true balls in 3D space, true balls are extracted in 3D space by the algorithm proposed by Yan, Kostin, Christmas, & Kittler (2006). In their method, the constant acceleration model fits to the ball candidates in every h_w frame. An integer greater than three is assigned to h_w and the fitting step is assigned to $h_w/3$ in our case. The fitting needs to be a window function like the above because the constant acceleration model will be inaccurate when the approximation time is very long. For the fitting, the candidates are $\boldsymbol{p_1}, \boldsymbol{p_2}$, and $\boldsymbol{p_3}$, the distances between two of which are less than the maximum distance a ball can travel. Let us denote the times of the candidates as t_1, t_2, and t_3, the difference between t_1 and t_2 as Δt_{21}, the distance between t_2 and t_3 as Δt_{32}, the acceleration as $\boldsymbol{a}$, and the velocity at t_1 as $\boldsymbol{v_1}$. The parameter of the constant acceleration model can be obtained by the following equations.

$$\boldsymbol{v_1} = \frac{\boldsymbol{p_2} - \boldsymbol{p_1}}{\Delta t_{21}} - \frac{\Delta t_{21}\boldsymbol{a}}{2}. \tag{7}$$

$$\boldsymbol{a} = 2\frac{\Delta t_{21}(\boldsymbol{p_3} - \boldsymbol{p_2}) - \Delta t_{32}(\boldsymbol{p_2} - \boldsymbol{p_1})}{\Delta t_{21}\Delta t_{32}(\Delta t_{21} + \Delta t_{32})} \tag{8}$$

The model may be inaccurate when it is reconstructed from successive candidates. This problem is solved by iteratively optimizing the trajectory model by using three points, *i.e.* the earliest, the latest, and the middle ones in inliers, which were consistent with the trajectory. This iterative optimization stops when the three points do not change or the sum of errors starts to increase. As we can see from Figure 6, the distance between the balls and the trajectory estimated from the three successive points (Figure 6(a)) is longer than that of the trajectory estimated after three iterations of the optimization process (Figure 6(b)). The above fitting is repeatedly done to the candidates in h_w frames. The set of parameters that returns the minimum distance between candidates is selected, and the candidates within U_S from the trajectory are selected as true balls. The trajectory in image spaces is reconstructed again by spline interpolation with the extracted balls. The 3D coordinates of the balls are then computed

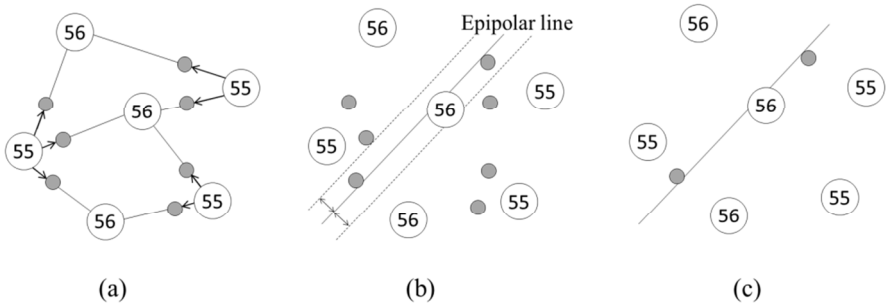


Figure 5. Ball extraction from ball candidates with epipolar line. White circles denote ball candidates. The number of each candidate is the frame number at which the candidate was detected. Gray circles denote estimated ball candidates. (a) Connect two candidates whose distance is less than threshold and estimate ball position at t + linearly. (b) Obtain epipolar line from the ball imaged in the other camera and detect candidates that are close to the epipolar line. (c) Detected candidates are extracted.

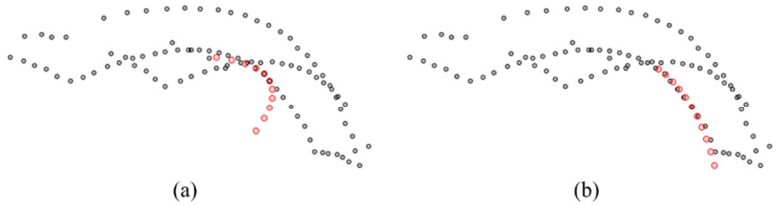


Figure 6. Connect two trajectories. Numbers denote times when the ball was detected. Distance between two trajectories becomes shortest at 58. (a) denotes "connectable". (b) denotes "mergeable". (c) denotes "unconnectable.

by triangulation with the corresponding points estimated from the trajectory in image spaces and the temporal offset between cameras.

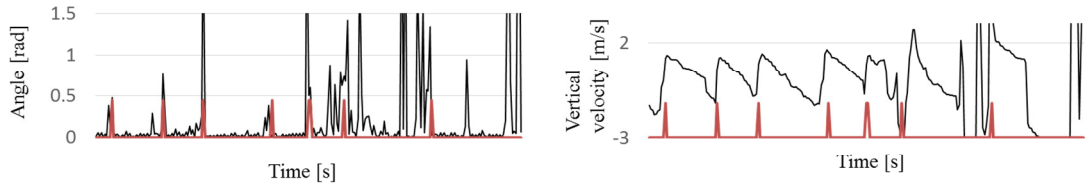
Collision Estimation

The velocity of each previously extracted ball is calculated by Equation 7. Collisions are estimated on the basis of the fluctuation of the ball velocity. Note that the trajectory was not reconstructed accurately around the collisions because of the limitation of polynomial fitting including spline. The trajectory is separated at every collision so as to reconstruct ball trajectory around the collisions accurately.

Figure 7 shows the change of vertical velocity and angle of two adjacent velocity of a ball when it collided with other objects. This figure indicates collisions flip the sign of the vertical velocity or enlarge the angle of two adjacent velocities. The following threshold is required for the collision estimation.

- Upper bound of the angle of two adjacent velocity vector [rad]: U_P

False collisions can be estimated because of the error of the ball velocities, which were calculated from the ball positions with measurement error. The false collisions are eliminated in the following process.



(a) Vertical velocity and collisions (b) Angle of two consecutive velocities and collisions

Figure 7. Change of ball motion when it collided with other objects. Pulses of red line denote collisions.

Trajectory Reconstruction

First, the trajectories between every two successive collisions are reconstructed by spline interpolation and resampled so that all sampling intervals are consistent. After that, the ball positions are smoothed by a Savitzky-Gollay filter (Savitzky & Golay, 1964), and then the trajectory is reconstructed again. Although the intervals of balls are not originally consistent when they were captured with unsynchronized cameras, the above process enables a common smoothing algorithm to be applied to them.

Next, multiple trajectories are connected. "Connect" means to assign the start and end times of each trajectory so that a ball position can be computed uniquely at anytime between the start and end times of the rally. Let us denote the former trajectory as T and the latter one as T'.

Then their start and end times are t_S, t_e, t'_s, and t'_e, respectively. The start time of the first trajectory is the start time of the rally, and the end time of the last trajectory is the end time of the rally. That means the problem is to determine t_e and t'_s for every two successive trajectories. The separated time of two trajectories is thought to be when a collision happens, and the relationship between them is shown in Figure 8(a). There, however, can be false collisions. Let us consider the two trajectories shown in Figure 8(b). The traveling directions between two trajectories are identical, and the ball positions at the separated time are close together. Therefore, the two trajectories should be merged. Next, let us consider the two trajectories shown in Figure 8(c). The ball positions at the separated time are far apart. The figures show three kinds of the relationship between two trajectories: connectable, mergeable, and unconnectable. Mergeable is a special case of connectable. The following thresholds are required to define the relationship between two trajectories.

- Upper bound of the distance between balls at a collision [m]: U_I
- Upper bound of the angle between velocity vectors at a collision [rad]: U_M
- Upper bound of the temporal interval between two trajectories [s]: U_T

Let us denote the time when the distance between two trajectories takes the smallest value as t_i, the time of the newest ball used for reconstructing T as t_f, and the time of the oldest ball used for reconstructing T' as t'_b. The relationship between two trajectories is classified as unconnectable if the following two conditions are met: (1) the difference between $\mathrm{t_f}$ and $\mathrm{t'_b}$ is greater than U_T, and (2) the distance between T and T' at t_i is greater than U_I. The relationship between two trajectories is classified as connectable, not mergeable, if the angle between two trajectories at t_i is greater than U_M, otherwise it is classified as mergeable. If the two trajectories are connectable, t_i is assigned to t_e and t'_s. If the two trajectories are mergable, the new trajectory is reconstructed with the balls used for reconstructing T and T', t_s is assigned to the start time of the new one, and the end time is determined when it is connected to the next trajectory. If the two trajectories are unconnectable, the mean of t_f and t'_b is assigned to t_e and t'_s. The flow of the trajectory connection is as follows;

Step 1 Extract the pairs of trajectories that are connectable and mergeable

Step 2 Extract the pair whose length of time is longest

Step 3 Eliminate trajectories that temporally conflict with the extracted ones

Step 4 Extract the pair whose length of time is longest except for the eliminated ones

Step 5 Repeat Steps 3-4 until no trajectory is extracted

Step 6 Merge mergaeble pairs of trajectories

The ball trajectory of a rally is reconstructed through the above steps as the connected trajectories, and this is the final output of the method.

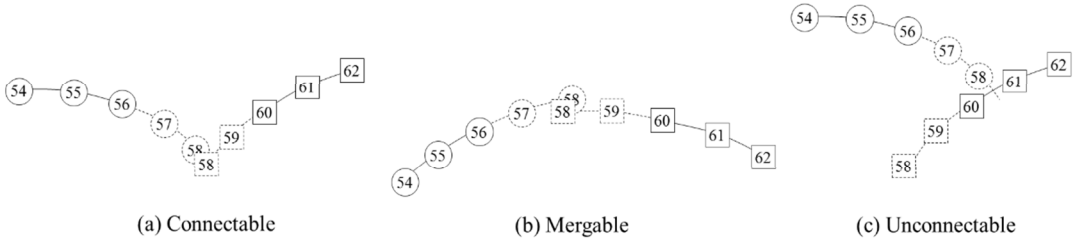
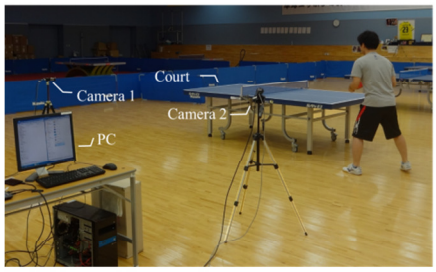


Figure 8. Connect two trajectories. Numbers denote times when the ball was captured. Distance between two trajectories becomes smallest at 58.

Experiment

Conditions

Two matches were played by experienced table tennis players. Different cameras were used and different placements were chosen for both matches. The two matches are separately referred to as Match 1 and Match 2 hereinafter. Target rallies are 35 and 10 rallies randomly chosen from Matches 1 and 2, respectively. Two Grasshopper 3s (Point Grey Research Inc.) were used for taking videos of Match 1. A 3 ms delay was set in the shutter timing of the one of the two cameras, *i.e.* the true temporal offset was -3 ms. The resolution and frame rate of the cameras were 1920 × 1080 and 60fps. Figure 9(a) shows the placement of the cameras, and Figure 10(a) and 10(b) show sample images captured by the cameras in Match 1. Two Lumix GH3s (Panasonic co.) were used for taking videos of Match 2. The true temporal offset was not known because these cameras do not have the functionality to use an external trigger for shuttering. The resolution and frame rate of the cameras were 1920 × 1080 and 60fps. Figure 9(b) shows the placement of the cameras, and Figure 10(c) and 10(d) show sample images captured by the cameras in Match 2. Balls were detected in a downsampled images whose resolution was 960 × 540, and ball positions were converted into the original image coordinates after detection. If the distance between a detected ball and the manually digitized position was greater than 15 pixels, the ball was classified as false detection. The threshold, 15 pixels, was the maximum radius of the ball in image spaces. Temporal and geometrical camera calibration was done with the first 10 rallies by the proposed method and by Noguchi & Kato's method. The step used for seeking temporal offset was 0.0165 ms. To evaluate geometrical calibration, the extrinsic parameters of cameras were computed with 18 control points and optimization by SBA (Lourakis & Argyros, 2009). The reprojection error was calculated with various poses of a checker board because its corners' positions can be obtained accurately. The performance of temporal and geometrical calibration was evaluated from the results of Match 1 because the true temporal offset between cameras was unknown in Match 2. The results of Match 2 were used for evaluating the performance of ball detection in a different environment from Match 1.



(a) Match 1

(a) Match 2

Figure 9. Placement of cameras.

(a) Camera 1 of Match 1 (b) Camera 2 of Match 1 (c) Camera 1 of Match 2 (d) Camera 2 of Match 2

Figure 10. Sample of input images. The brightness is modified to improve the visibility.

The thresholds of the proposed method were set as follows in Match 1: L_I: 0.24, L_C: 0.3, L_A: 13, U_A: 531, U_V: 25, L_V: 1, U_D: 0.3, U_R: 0.524, U_E: 0.025, U_S: 0.15, U_P: 0.209, U_I: 0.015, U_M: 0.785, and U_T: 1. In Match 2, 0.12 was set to L_I because the videos of Match 1 were brighter than those of Match 2. Other parameters were not changed from those of Match 1. Optimal L_C, L_A, U_A, U_E, and U_S are changed if the resolutions are different. Optimal $\mathrm{U_D}$ is changed if the relative camera angle to the table tennis court is different, although the placement of the two cameras will be quite similar in other shoots because the angle of two cameras should be close to orthogonal and cameras should shoot the match from the sidelines of the table tennis court to avoid the spatial occlusion between players and a ball. U_V, L_V, U_R, U_P, U_I, U_M, and U_T are static as long as the target sport is table tennis.

Results

Ball Detection and Extraction Before Camera Calibration

The overall number of frames and the number of frames where a ball exists are shown in Table 1. The results of the ball candidate detection are shown in Table 2, in which the numbers denote the ratio of the number of each kind of results to the number of frames where a ball exists.

Table 1. The number of frames. #frames stands for overall number of frames. #ball-frames stands for the number of frames where a ball exists

	Match 1		**Match 2**	
	View 1	View 2	View 1	View 2
#frames	3422	3379	1412	1397
#ball-frames	3312	3239	1199	1251

Table 2. Results of the appearance based ball detection. Detected stands for the frames where the true ball was included in the ball candidate, and other frames are Missed.

	Match 1		**Match 2**	
Indicator	View 1	View 2	View 1	View 2
Detected [%]	96.0	92.2	92.4	91.3
Missed [%]	4.0	7.8	7.6	8.7

The reasons for misdetection were as follows: spatial occlusion with players: 66.2% (398/601), spatial occlusion with the net: 19.8% (119/601), background color similar to the ball: 11.3% (68/601), and being static in the image: 2.7% (16/601). A ball becomes static when it travels along the optical axis of the camera or reaches the vertex of a parabola and becomes too slow. At that time, the ball is judged as background through frame differentiation, and misdetection occurs. The results of the ball extraction based on the motion, namely before camera calibration, are shown in Table 3. More than 97% of the extracted balls were true. The ratios of the detected balls whose 3D coordinates can be computed to the number of frames where a ball exists are shown in Table 3 as “Detection rate 3D”.

Table 3. Results of motion based ball extraction in image spaces. TP stands for true positive, FP stands for false positive, TN stands for true negative, and FN stands for false negative. "Detection rate 3D" stands for the ratios of the balls whose 3D coordinates can be computed to the number of frames where a ball exists.

	Match 1		**Match 2**	
Indicator	View 1	View 2	View 1	View 2
TP [%]	82.7	70.7	62.2	64.9
FP [%]	0.3	0.1	0.1	1.6
TN [%]	3.2	4.1	15.1	10.5
FN [%]	13.8	25.1	22.6	23.1
Detection rate [%]	85.5	73.8	73.2	72.4
Detection rate 3D [%]	61.6	66.4	48.7	46.1
Precision [%]	99.7	99.9	99.8	97.6

Temporal and Geometrical Camera Calibration

The results of the temporal and geometrical calibration are shown in Table 4.

Table 4. Results of temporal and geometrical camera calibration

	Nogucni & Kato	**Proposed method**
Error of temporal offset [ms]	1.95	0.04
Error of optical axis [degree]	3.26	1.07
Error of camera center [m]	0.27	0.03
Reprojection error [pixel]	2.71	1.31

Ball Extraction After Camera Calibration

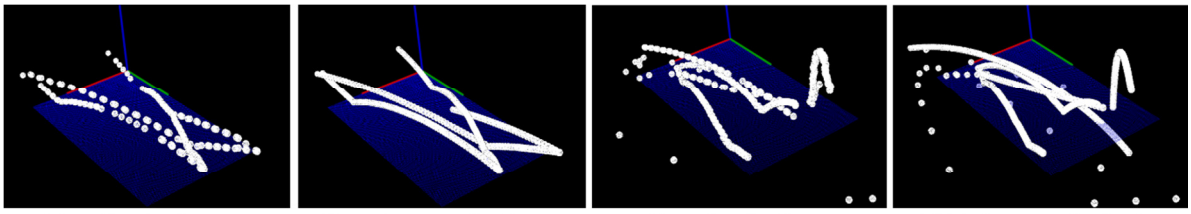
The results of the ball extraction using the temporal offset and geometrical relationship between cameras, namely after camera calibration, are shown in Table 5. The ratios of the detected balls whose 3D coordinates can be computed to the number of frames where a ball exists are shown in Table 5 as "Detection rate 3D".

Table 5. Results of ball extraction using temporal offset and fundamental matrix. TP stands for true positive, FP stands for false positive, TN stands for true negative, and FN stands for false negative. "Detection rate 3D" stands for the ratios of the balls whose 3D coordinates can be computed to the number of frames where a ball exists.

	Match 1		**Match 2**	
Indicator	View 1	View 2	View 1	View 2
TP [%]	85.4	85.0	66.7	69.8
FP [%]	0.8	1.9	3.9	2.1
TN [%]	3.2	4.1	15.1	10.5
FN [%]	10.6	9.0	14.3	17.6
Detection rate [%]	88.2	88.6	78.6	77.9
Detection rate 3D [%]	79.9	85.2	72.4	73.9
Precision [%]	99.1	97.8	94.5	97.0

Reconstruction of 3D ball trajectory

A sample of the successfully reconstructed trajectories is shown in Figure 11(b), and the balls used for the reconstruction are shown in Figure 11(a). Additionally, a sample of the unsuccessfully reconstructed trajectories is shown in Figure 11(d), and the ball positions used for the reconstruction are shown in Figure 11(c). The ratios of the balls in image space, namely the detected ball segments, covered by the reconstructed trajectory were 93.9% and 95.6% in Match 1 and 80.6% and 81.8% in Match 2.



(a) Balls of rally 1 (b) Trajectory of rally 1 (c) Balls of rally 2 (d) Trajectory of rally 2

Figure 11. Reconstruction of trajectories from balls. (a) was used for reconstructing (b). (c) was used for reconstructing (d).

Discussion

Ball Detection and Extraction Before Camera Calibration

The biggest reason for misdetection was spatial occlusion with players. This result suggests that the cameras should be appropriately placed to avoid this. True balls were able to be detected with a high probability. Although false detections are also contained in the detection results, true balls could be extracted with high precision in the extraction process. The results of the ball extraction (Table 3) indicate that this method is appropriate to extract balls for camera calibration because of the high precision. However, the detection rate is not sufficiently good for reconstruction of ball trajectory. In the proposed method, therefore, another method is used for trajectory reconstruction.

Temporal and Geometrical Camera Calibration

The error of the temporal offset estimated by the proposed method is 2% of that estimated purely by Noguchi & Kato's method. All errors of the geometrical calibration are also much smaller. These results demonstrate that the modification proposed in this paper is helpful to improve the accuracy of the temporal and geometrical calibration. The reprojection error of 1.31 pixels equals about 3.2 mm in the real world. This accuracy is sufficiently good for performance analysis in sports.

Ball Extraction After Camera Calibration

The detection rate is 18-27% higher than that of the previous extraction. Although precisions are 0.6 -5.3% smaller than that of the previous extraction, this difference is relatively small. This result indicates that the robustness can be improved by taking advantage of the proposed method and that the temporal offset and the geometrical relationship between cameras can be known.

The reasons 20% or 27% extractions were mis-extractions need to be considered. One of the biggest reasons is the misdetection at the ball detection. About 4-9% of balls were not detected as shown in Table 2. Another main reason is that a ball was not imaged in one of the two cameras, which occurred in 7.9% and 6.7% of frames in Match 1 and 9.3% and 12.5% of frames in Match 2. Unfortunately, these kinds of misdetections cannot be avoided with only two cameras. Spatial occlusion with players, nets, or other objects can happen wherever two cameras are placed. A color similar to the ball's can be included in the background because this cannot be controlled in many cases. The angle of two optical axes cannot be small in terms of the measurement accuracy; 90° is the best. When the angle of two optical axes is big, a ball is often not imaged in one of the two cameras. Although the two-camera case is discussed in this paper, the proposed method can be applied to multi-camera camera case by calibrating each pair of cameras one by one. The number of cameras should be increased so as to make a ball visible for two or more cameras if a higher detection rate is required.

Reconstruction of 3D ball trajectory

The ratios of the balls in image spaces covered by the reconstructed trajectory were higher than the ratios of the balls whose 3D coordinates can be computed. That means the trajectory well interpolated in the period of time when balls were missed. For example, the 3D positions of the balls around a bounce were not computed in the sample in Figure 11(a) (see the top right of the table). The trajectory interpolated the term well as shown in Figure 11(b). Balls can be missed for reasons described above. These results indicate the proposed method deals with the problem of misdetection if the adjacent trajectories are successfully reconstructed. Some trajectories were, however, reconstructed unsuccessfully as shown in Figure 11(d). The trajectory right after the bounce could not be reconstructed, while the wrong trajectory, such as that penetrating the table, was reconstructed. Moreover, false detections remain in 3D space, and when false positives existed, a completely wrong trajectory was reconstructed. The longest trajectory in a specific period of time is extracted in 3D space in the proposed method. Thus, if a ball is not detected during several frames, other objects, such as a player's foot, can be extracted as a ball. In this experiment, false detections remain in 1-4% frames. Therefore, the number of cameras should be increased when trajectories need to be reconstructed more robustly.

Conclusion

A method was proposed that reconstructs the 3D trajectory of a ball with unsynchronized cameras. In the method, balls are detected on the basis of their intensity, area, circularity, and motion. Cameras are temporally and geometrically calibrated by using the ball trajectories in image spaces. Balls are detected again by using the temporal offset and geometrical relationship between cameras. It was experimentally demonstrated that the temporal offset and geometrical relationship between cameras can be estimated accurately and the ball trajectory can be reconstructed robustly. The next challenge will be to develop the software for actual use. More specifically, the following two challenges remain as future work: (1) develop software that can handle videos from three or more viewpoints, and (2) develop an algorithm for extracting the useful data for sports analysis from trajectories. Using the proposed method is expected to advance the spatial and temporal understanding of sports.

Acknowledgments

A part of this work was supported by Kozuki foundation, and another part was supported by National Institute of Information and Communications Technology (NICT), Japan. I would also like to express my gratitude to Haruhiko Ikebukuro of the Japan Sports Council for his cooperation in the experiment.

References

Antoku, C., Kashima, M., Sato, K., & Watanabe, M. (2013). Research for automatic tennis play recognition and recording based on motion analysis. *The Special Interest Group Technical Reports of IPSJ*, *2013*(14), 1-6.

Brutzer, S., Hoferlin, B., & Heideman, G. (2011). Evaluation of background subtraction techniques for video surveillance. *Proc. Computer Vision and Pattern Recognition*, 1937-1944.

Chen, W., & Zhang, Y.-J. (2006). Tracking ball and players with applications to highlight ranking of broadcasting table tennis video. *Proc. Compu-tational Engineering in Systems Applications*, 1896-1903.

Guo, J.-M., Hsia, C.-H., Liu, Y.-F., Shih M.-H., Chang, C.-H., & Wu, J.-Y. (2013). Fast background subtraction based on a multi-layer codebook model for moving object detection. *Circuits and Sys-tems for Video Technology*, *23*(10), 1809-1821.

Hartley, R. & Zisserman, A. (2003). *Multiple view geometry in computer vision*. UK:Cambridge Univ Press.

Hu, S., Zaojun, F., De, X., & Min, T. (2013). Trajectory prediction of spinning ball based on fuzzy filtering and local modeling for robotic ping-pong player. *IEEE Transactions on Instrumentation and Measurement*, *62*(11), 2890-2900.

Huang, Y., Llach, J., & Zhang, C. (2008). A method of small object detection and tracking based on particle filters. *Proc. 19th International Conference on Pattern Recognition*, 1-4.

Ishii, N., Kitahara, I., Kameda, Y., & Ohta, Y. (2007). A position estimation of a soccer ball by using two viewpoint videos. *Proc. the IEICE General Conference*, 157.

Kakumu, Y., Sakaue, F., Sato, J., Ishimaru, K., & Imanishi, M. (2013) High frequency 3D reconstruction from unsynchronized multiple cameras. *Proc. British Machine Vision Conference*, 43.1-43.10.

Kitahara, S., & Uchida, O. (2013). A proposal on automatic analysis method of tennis play using movies of tennis match. *ITE Technical Report*, *37*(35), 51-54.

Lourakis, M. I. A., & Argyros, A. A. (2009). SBA: A software package for generic sparse bundle adjustment. *ACM Transactions on Math. Software*, *36*(1-30)

Matsumoto, H., Sato, J., & Sakue, F. (2010). Multiview constraints in frequency space and camera calibration from unsynchronized images. *Proc. Computer Vision and Pattern Recognition*, 1601-1608.

Noguchi, M., & Kato, T. (2006). Geometric and timing calibration for unsynchronized cameras using trajectories of a moving marker. *IPSJ Transactions on Computer Vision and Image Media*, *47*(SIG 5(CVIM 13)), 59-68.

Savitzky, A. & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, *36*(8), 1627-1639.

Shimizu, S., Fujiyoshi, H., Nagasaka, Y., Takahashi, T., & Iwahori, Y. (2009). A method of 3D position estimation using asynchronous multiple pan-tilt cameras. *Journal of JSPE*, *75*(2), 284-289.

Takanohashi, K., Manabe, Y., Yasumuro, Y., Imura, M., & Chihara, K. (2007). Measurement of 3D ball trajectory using motion blur. *IPSJ Transactions on Computer Vision and Image Media*, *48*(SIG 1(CVIM 17)), 35-47.

Yan, F., Kostin, A., Christmas, W., & Kittler, J. (2006). A novel data association algorithm for object tracking in clutter with application to tennis video analysis. *Proc. Computer Vision and Pattern Recognition*, 634-641.

Yu, X., Sim, C. H., Wang, J. R., & Cheong, L. F. (2004). A trajectory-based ball detection and tracking algorithm in broadcast tennis video. *Proc. International Conference on Image Processing*, 1049-1052.

Zhou, C., & Tao, H. (2003). Dynamic depth recovery from unsynchronized video streams. *Proc. Computer Vision and Pattern Recognition*, *2*, II-351-358.

Zhang, Z. (2000). A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(11), 1330-1334.

A hybrid algorithm for player arm biomechanics evaluation in outdoor sporting activities

Fernando, T., Pinidiyaarachchi, U. A. J.

Department of Statistics and Computer Science, Faulty of Science, University of Peradeniya, Sri Lanka

Abstract

Despite the large amount of studies conducted in the field of human pose estimation and tracking in sports, currently there is a lack of a system which is capable to track player arm movements in real time. Such a system can assist the players to master the right techniques with guaranteed optimal performance of the player. In this paper we propose a robust algorithm to model player arm movements in outdoor sporting activities. The system uses trained cascade object classifier to predict a region of interest for arm in the monocular input video sequence. Optical flow algorithm is employed to extract the motion in that region. Arm region in resultant binary image is later classified using Active Shape Model. The algorithm is tested and validated using several experiments for tracking ball delivery process in cricket as well as for tracking service in volleyball and tennis. The algorithm is capable of classifying and tracking player arm movements with more than 80 percent accuracy, irrespective of the position and background complexities that the real gaming conditions offer.

KEYWORDS: MOTION ANALYSIS, JOINT KINEMATICS, ARM TRACKING, SPORTS BIOMECHANICS, VIDEO SELF CALIBRATION.

Introduction

A sport is an activity involving physical exertion and skill in which an individual or team competes against another. It is also used for entertainment as well as for improvement of physical, social and psychological aspects of life. Not only professional athletes but any individual involved in sports requires mastering the correct techniques to achieve his or her goals.

For an example in cricket, the process of delivering the ball last less than five seconds even when the bowler is a spinner. Therefore it is extremely difficult to master the correct techniques by monitoring without any tools. As the competiveness of the game increases, countries tend to move on to computer technology which enables to obtain information on athlete's bio-mechanics techniques which can be difficult to evaluate using naked eye. Such model will not only help the player to master the accurate technique to deliver the ball more precisely but also to check whether the player is chucking the ball. These facts are applicable when considering sports such as tennis and volleyball. The arm movement of the player is directly related with the angel and velocity of the ball in service.

As stated by Yu et al. (2005), developed countries spend large amounts of money on advanced motion capturing systems to obtain accurate details of body position and elbow movements of their players. The utilization of videos, to watch players delivery or serve, has been proven a

useful tool for a coach to detect and correct the players techniques or important mistakes. If the player is able to obtain a model of his or her delivery action or the arm movement in the service, it will be helpful in determining how to apply the force and the acceleration needed more accurately and more efficiently .

Most approaches to computer vision based motion capture techniques make use of special markers attached to the athlete. As described by Shotton et al. (2011), the drawback of marker based motion capturing is that it requires subjects to wear special suits with markers or to have the markers attached to their clothing or body. It is not practical to apply markers in sports like cricket, volleyball and tennis due to the fact that they are outdoor sports as well as it might disturb the players. Additionally, if the marker slips or becomes blocked by another object, the tracking process will be inaccurate. Therefore the need for a low cost, marker-less motion capturing systems with high accuracy is inevitable especially for developing regions of the world.

With the introduction of low cost three dimensional cameras such as Microsoft Kinet and Asus Xtion PRO into the market the researches focuses in using such devices to track the movements in the player's Bio mechanical model in real time. But such approaches have lot of constraints due to the lack of frame rate in the device. As per the specification in the device, it is only accurate for limited depth range (for Asus Xtion PRO the value is 2 meters). This approach cannot be used in most of the outdoor sports. These cameras capture colour and depth data using a RGB camera, a monochrome camera and reflective infra-red (IR) camera. In outdoor sporting environment can have additional IR radiation due to other factors such as illumination of sun or other artificial light sources. These factors will introduce noise in the depth measurement; hence the bio mechanical model will be inaccurate.

We propose a marker less motion capturing method to model the arm movements in outdoor sporting activities. The proposed algorithm combines current state-of-the-art motion capturing and pattern recognition techniques in order to come up with a model to capture the variant nature in player arm movements. A trained cascade object classifier is used to detect the face area of the athlete. A Region Of Interest (ROI) for the arm region is defined using the head height as a parameter. An Active Shape Model (ASM) is trained to fit in to the contour of the arm region within the defined ROI. The method is well tested and applicability is proven not only for training purposes but also for real game scenarios as well.

The remainder of this paper is organized as follows: in Section 2, we present the materials and methods used in the study. The proposed method is validated with obtained experimental results in Section 3. Finally, conclusions are drawn in Section 4.

Methods

I. Input-video

The video capturing device should have a higher frame rate and higher shutter speed.

A higher frame rate is chosen to cope with the fast reflexes of the professional athletes. The camera should be mounted to a tripod and kept still in video capturing process. This constraint is imposed to simplify the motion estimation. A separate experiment was conducted to determine the minimum frame rate required. Videos of delivery and service process of national level athletes were taken at 25, 30, 35, 50, 100 and 300 frames per second. It is concluded that the minimum frame rate required for tracking the arm movement of professional athletes in

cricket, tennis and volleyball is 50 frames per second. If a lower frame rate is chosen, some intermediate data will be lost; hence the generated motion model will be erroneous. With a similar kind of experiment, in which we obtained videos in 1000, 2000, 2500, 3000 and 5000 shutter speeds, it is observed that the optimal shutters speed for input video stream is 5000. In the frames of the videos from lower shutter speeds, a blurriness was observed in the arm region.

II. Background subtraction

When considering the real gaming environments, background subtraction is one of the most challenging problems. Audience as well as the other players in the background are all in motion. In order to isolate the motion of the player form the complex dynamic background optical flow is employed. Motion vectors generated by the player who is delivering the ball or serving the ball are larger than the motion vectors generated by other motions in the background. Though it simplifies the motion estimation process, the introduction of optical flow introduces some assumptions such as smoothness of the motion and constant illumination in the neighbourhood, which hardly holds in real gaming environments. According to Bruhn et al. (2005) optical flow requires generation of temporal and special derivatives, in which large fluctuations on these derivatives can be occurred by small noise in the signal. Local optical flow methods like Lucas–Kanade (1981) method tries to overcome this problem by optimise some local energy-like expression, and global strategies such as the method of Horn and Schunck (1981) which attempt to minimise a global energy functional. In a study conducted by Galvin et al (1998), they have compared these two classical methods in the presence of Gaussian noise. Their conclusion was that the global approach of Horn and Schunck is more sensitive to noise than the local Lucas–Kanade method. Considering the above stated constraints Lucas–Kanade method is employed in order to generate optical flow. As the results may be corrupted by periodic noise, filtering and segmentation are required to perform more precise foreground estimation. The motion vectors are thresholded to remove small motions in the background. Morphological operations are then performed to enhance the edge strength of the detected objects. Closing, with short line segments in diagonal directions as the structuring element, is used to connect closely related disconnected components. Area opening is used to remove the small disconnected pixel patches in the segmented image.

III. Player localization and defining a region of interest for the hand.

In order to minimize the noise in the extracted arm region, a region of interest for the player's arm is defined. In the study by Jarzem and Gledhill (1993) a correlation of 0.989 between arm span and height of a person was observed. They have considered the height of 119 normal individuals aged 0.5-56 year. In their study, gender differences as well as the height difference between different nationalities were considered. Using the classical ratio that an average person is generally 7-and-a-half heads tall, it can be inferred that length of arm region of a person is approximately 4 times of his or her heads tall.

To locate the players head tall, face recognition techniques are employed. The typical Viola & Johns (2003) face recognition works only for straight face profiles. The method does not work for tilt heads or side faces. But during the delivery of the ball in cricket or during the service in tennis or volleyball the player's head is unlikely to be straight all the time. Therefore a separate cascade object classifier will be trained to locate the face area of the players. Cascade object classifiers for face detection is first introduced by Viola and Johns and later Wu B., et al (2004) enhanced the classifier to adapts to the distribution of weights on the training set.

Cascade object classifier is a collection of weak classifiers. These simple classifiers are called decision stumps. Each stage is trained using a technique called boosting. According to Solar et al. (2006), boosting provides the ability to train a highly accurate classifier by taking a weighted average of the decisions made by the weak learners. The hierarchical design of the classifiers which is obtained using the stages enhances the efficiency of the face detection task. It rejects the negative samples as fast as possible by assuming that the vast majority of windows do not contain the object of interest. Conversely, true positives are rare, and it passes the positive windows to the next levels.

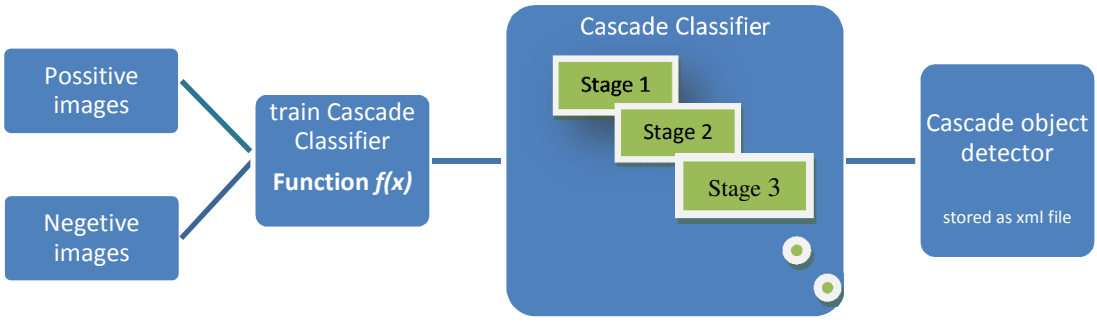


Figure 1: Cascade object classifier training process

Figure 1 elaborates the cascade object classifier training process. At each stage, the classifier labels the region defined by the current location of the sliding window as either positive or negative. Positive indicates an object was found and negative indicates no object. If the label is negative, the classification of this region is complete, and the detector slides the window to the next location. If the label is positive, the classifier passes the region to the next stage. The detector reports an object found at the current window location when the final stage classifies the region as positive (i.e. in this particular application as a face region).

The cascade object classifier is trained in 45 cascade stages with 200 positive samples and 400 negative samples which achieves the negative sample factor of 2. Using the above configurations 0.0051 overall false alarm rate is achieved.

The classifier will detect the face area of the player who is delivering or serving, as well as the face area of all the people in that particular frame. This problem is solved by further analyzing the delivery and serving processes. In delivery of the ball or during the service process only the arm of the player performing that activity will be over his or her shoulder. Therefore the proposed method is designed to detect the players arm when it comes over the shoulder. This simplifies the tracking procedure and doesn't impose any constraint on obtained data. Angular momentum, serve or delivery speed, delivery angel are gained from the moment arm comes over the shoulder but not from the rotations prior to it. Figure 2 show the effectiveness of the above approach in a real game scenario. The detected face regions and region of interest for arm are displayed with yellow colour bounding boxes in the figure. Though the classifier detects the face region of the umpire it doesn't define a region of interest for the arm area of the umpire because there is no significant component within that bounding box (i.e. no arm region over his shoulder).

Figure 2 (a): Cricket

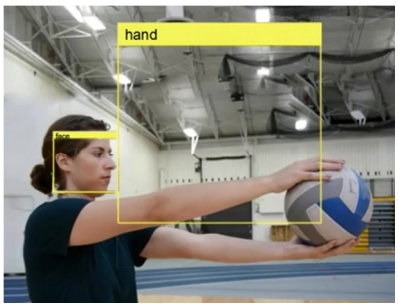


Figure 2 (b): Volleyball

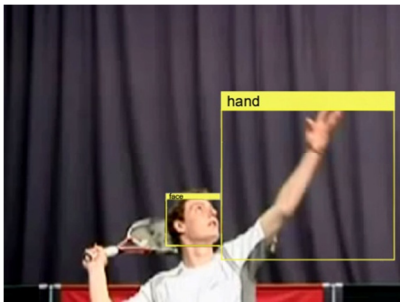


Figure 2 (c): Tennis

Figure 2: Face detection and hand region estimation

IV. Ball detection and back tracking

This module is specifically applied to remove complexities in player arm tracking in cricket. When considering the ball delivery process in cricket, there can be several rotations of arm over shoulder prior to the delivery of the ball. In order to obtain the frames which are taken just before the delivery of the ball a special back tracking algorithm is designed. For example in cricket, bowlers hand may come several times over his or her shoulder before the delivery. The proposed system overcomes this problem by back tracking the released ball to the point of delivery. A Circle Hough transform (CHT) proposed by Duda et al. (1972) combined with colour image segmentation is employed here. The key idea in CHT is similar as in line detection is to extract edges first and then for each edge pixel accumulate votes in an appropriate subset of parametric space using all possible circles passing through it. As described by Sýkora et al. (2008) the significant peaks in such a 3D histogram determine centers and radii of salient circles in the image. Separate colour streams are being processed for different ball colours. For an example in order to track the red cricket ball, the red channel of the RGB frame is subtracted from the gray frame, which is obtained by performing RGB to gray conversion of the same frame. After thresholding the result to reduce noise and remove holes, small regions are connected to their sufficiently large neighbours. In the next step CHT is applied on the segmented image. The estimation of circle radius is vital for detection of the released ball. It is set with a close match to the radius of the actual ball and with less degree of freedom therefore if the ball is partly occluded, it want be detected. After locating the released ball the frames are backtracked to the point where the defined ROI for the arm region becomes empty. Connected component analysis is used to obtain a list of regions and their neighbouring relations. Ones the largest blob inside the defined ROI disappears, it can be concluded that the player's arm hasn't come over the shoulder yet. Frames from this point onwards to the frame in which Hough circle detection comes true first, are selected for further processing in steps to follow. Figure 3 shows a particular implementation where tracked released ball is indicated with a circle.`

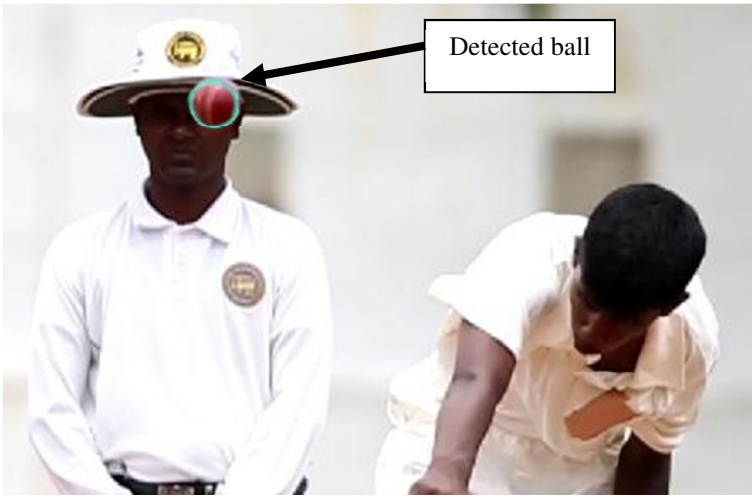


Figure 3: Ball tracking

V. Hand segmentation and tracking

When tracking non rigid objects such as human arm which has a greater degree of freedom for rotation, two dimensional deformable objects are ideal. They use the contour representation to describe an object in the image. Koschan et al. (2002) have used an Active Shape Model to segment the human silhouette from the video frames as it is one of the best-suited approaches in terms of both accuracy and efficiency among the other methods that fall under the class of deformable models. It uses prior knowledge about the shape of the object for the segmentation process. More specifically, ASM-based tracking algorithms consist of four steps. (i) Assignment of landmark points, (ii) PCA, (iii) model fitting, and (iv) local structure modelling. In the training set, predefined number of landmark points are selected manually. In the testing phase the transformations needed for the alignment of the trained contour model into the target object in the testing frame, are determined in an iterative process. In order to eliminate the exhaustive search to find a suitable initialization point, we use the region of interest predicted for the arm region in step iii. The search process will be limited to the objects within this region.

i. Assignment of landmark points

For the purpose of generating the contour model of the object, suitable landmark points are selected from the given frame of input video. In order to achieve a greater degree of accuracy, the landmark points should be continuously available throughout the frames. In particular, for a 2 dimension training image, the selected n landmark points can be represented as,

$$x = [x_1 \ , \dots , x_n \ , y_1 \ , \dots , y_n]^T \quad (1)$$

In this study 24 landmark points are selected from 100 training frames obtained from 35 different players. The landmark points are selected manually on the edge images obtained from background subtraction process.

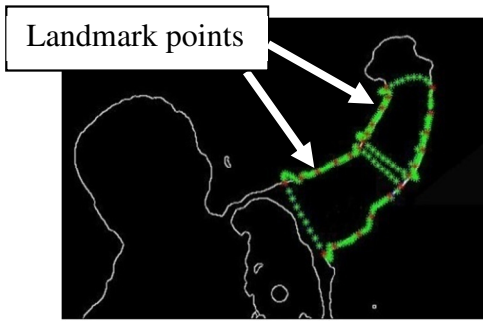


Figure 4: 24 manually selected landmark points

Figure 4 shows the arm model with 24 manually selected landmark points on the initial contour. The landmark points are indicated in red and subsequent points in green.

ii. Principal component analysis

Figure 5 elaborates a sample of twelve different configurations that can be generated in the previous phase. The figures are not normalised for scale variations. It is evident that with the available degree of freedom in motion several different arm models can be created.

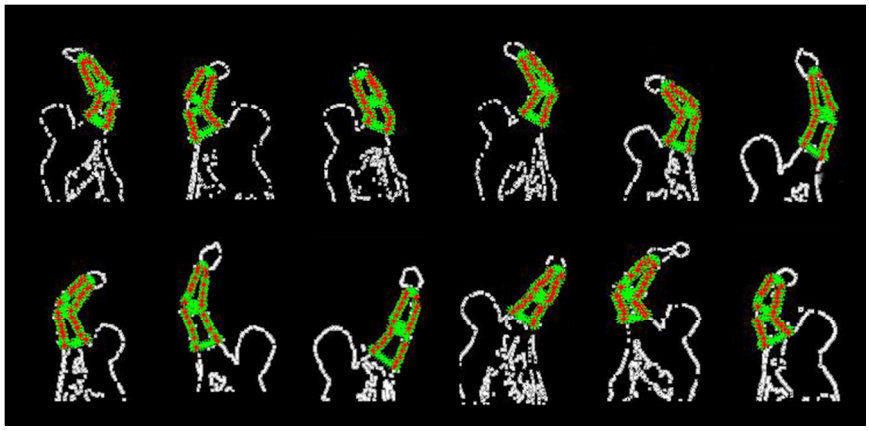


Figure 5: Different landmark points configurations

The trained shape in the previous step can be modelled with reduced number of parameters using PCA technique. The main objective of PCA is to perform dimensionality reduction while preserving as much of the randomness in the high-dimensional space as possible. Hence each shape in the training set which is in the 2n-dimensional space can be modelled with a lesser number of parameters.

Suppose we have *m* shapes in the training set, presented by x_i, for i=1;...,m. The PCA algorithm is as follows.

Compute the mean of the *m* sample shapes in the training set

$$\overline{x} = \frac{1}{m}\sum_{i=1}^{m} x_1 \tag{2}$$

Compute the covariance matrix of the training set

$$S = \frac{1}{m}\sum_{i=1}^{m}(x_i - \bar{x})(x_i - \bar{x})^T \quad (3)$$

Construct the matrix

$$\Phi = [\phi_1|\phi_2 \ldots|\phi_q] \quad (4)$$

Where ϕ_j=1....,q represent eigenvectors of S corresponding to the q largest eigenvalues.

Given ϕ and $\bar{x}$, each shape can be approximated as

$$x_i \approx \bar{x} + \Phi b_i \quad (5)$$

where

$$b_i = \phi^T(x_i - \bar{x}) \quad (6)$$

iii. Model fitting

$$E = (y - Mx)^T W(y - Mx) \quad (7)$$

The modelled shape in the coordinate x is matched with a new shape in the coordinate y by minimising the equation 2. The weighting matrix W is inversely proportional to M, which represents the geometric transformation necessary to overlap the model with the new frame.

As the result of the searching procedure along profiles, the optimal displacement of a landmark point is obtained. The combination of optimally updated landmark points generates a new shape in the image coordinate frame, y.

iv. Modelling a local structure

The nearest sample shape that fits the model can be obtained by minimizing the Mahalanobis distance between the sample and the mean of the model.

VI. Construction of 2D model of the arm

ASM based tracking process results in locating the changes of the arm model over time. This two dimensional data will be plotted and presented to the users as the output.

It is a vital factor in training to monitor the changes of the technique of a player over time. Also when considering cricket, one can monitor whether the player is deliberately chucking the ball or whether it is his or her natural action. The proposed method enables users to compare two or more video clips of the same player, which will result in plotting of several different bowling actions in same grid. This introduces a need for calibration of the videos. As elaborated in section II the background subtraction algorithm assumes a static camera. The only constraint that it imposes is that camera should be still during the video capturing process. Hence a tripod should be used. But after the video capture is completed, the camera can be moved and the calibration algorithm which we propose in this section will allow the users to analyze two or more such videos with different angels, provided that there are some overlapping parts in the frames.

Dwarakanath et al. (2012) have proposed a method for video self calibration based on image feature detected by Scale Invariant Feature Transform (SIFT) which was fisrt introduced by Lowe D. (2004). More precisely in their method, the systems automatically match the features between camera images using SIFT, which are then used to perform the calibration. The

feature points detected by SIFT are assigned a weight value based on the stability of the feature detection. In order to estimate the essential matrix the SIFT feature points which have the highest stability are chosen. Essential matrix E is derived using normalized 8-point algorithm which is proposed by Hartley et al.(1997). The algorithm assumes the rotation matrix and the translation matrix as identity matrix and zero matrix respectively. Relative rotation R and translation t of the second camera of the camera pair represents the camera pose, and are related to essential matrix as E=[t]XR, where [t]X is a skew-symmetric matrix [20],

$$[t]_x = \begin{bmatrix} 0 & t_x & -t_z \\ -t_x & 0 & t_y \\ t_z & -t_y & 0 \end{bmatrix}$$

The above essential matrix can be used for determining both the relative position and orientation between the cameras and the 3D position of corresponding image points.

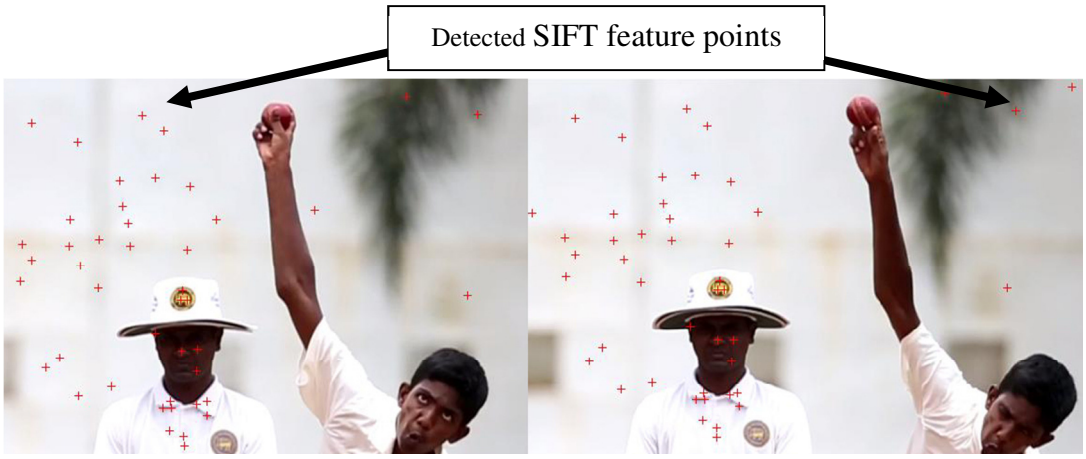


Figure 6: SIFT point selection for video self calibration

Figure 6 shows first frames of two videos of same bowler bowling in two different overs. The locations of the selected SIFT points are marked in red. Intersection operator will provide the common feature points for both cases and later utilized by the calibration algorithm.

Results

In this section, we present the quantitative and qualitative results of our tracking approach. To generalize our approach we have obtained data from different camera types, in different frame rates, in different gaming conditions. We have used Cannon 5D Mark III (5D) camera with 800mm lens, Sony HVR-S270E (Sony I) and Sony NX5 Full HD camera (Sony II). Some practice session videos were obtained from an online resource. (Analyzing volleyball serve by Abby Oakland, Northern State University, Basic tennis serve technique by David Leyod tennis school. url: www.introsprot.com)

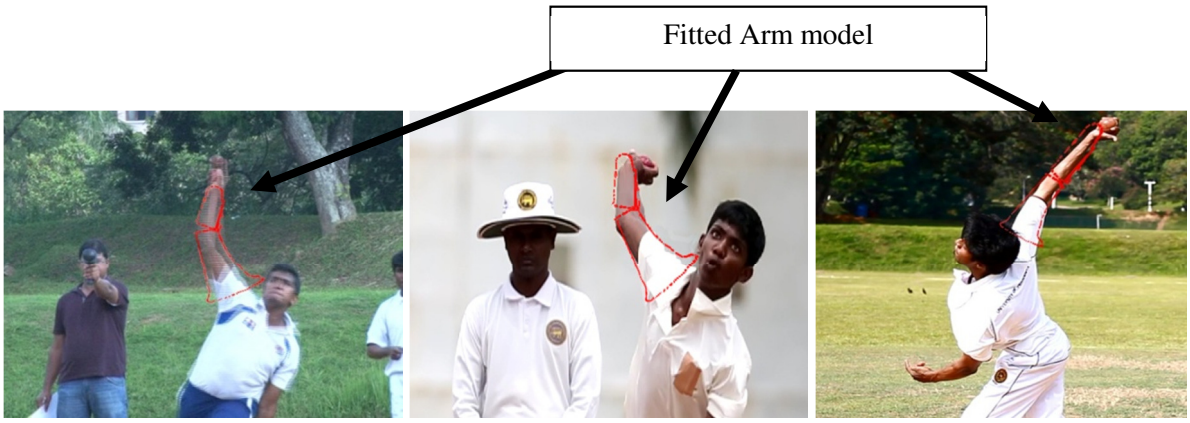
The sequences cover different challenges of object tracking. The details regarding frame rates, number of videos and resolutions are elaborated in Table 1.

Table 1: Details of the video dataset

Game	Gaming condition	Camera Model	Frame rate	# of videos	Resolution
Cricket	Practice session	5D	50 fps	72	1024x800
	Real Game	5D	50 fps	65	1024x800
	Practice session	Sony I	100fps	80	600x480
	Real Game	Sony I	100fps	40	600x480
	Practice session	Sony II	60 fps	30	720x480
	Real Game	Sony II	60 fps	22	720x480
	Practice session	Internet	30 fps	05	320x240
	Real Game	Internet	30 fps	02	320x240
Volleyball	Practice session	5D	50 fps	32	1024x800
	Real Game	5D	50 fps	25	1024x800
	Practice session	Sony I	100fps	11	600x480
	Real Game	Sony I	100fps	15	600x480
	Practice session	Sony II	60 fps	20	720x480
	Real Game	Sony II	60 fps	10	720x480
	Practice session	Internet	30 fps	12	320x240
	Real Game	Internet	30 fps	15	320x240
Tennis	Practice session	5D	50 fps	27	1024x800
	Real Game	5D	50 fps	24	1024x800
	Practice session	Sony I	100fps	26	600x480
	Real Game	Sony I	100fps	13	600x480
	Practice session	Sony II	60 fps	44	720x480
	Real Game	Sony II	60 fps	52	720x480
	Practice session	Internet	30 fps	09	320x240
	Real Game	Internet	30 fps	10	320x240

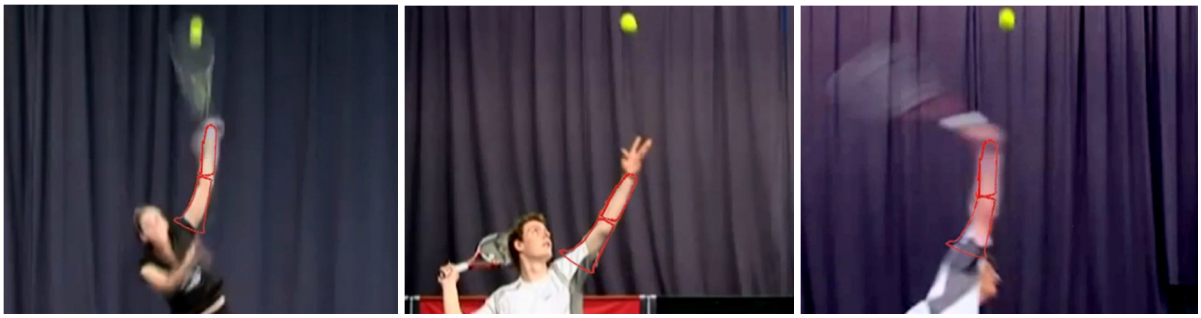
Figure 7: Data capturing in actual game situation

I. Qualitative results

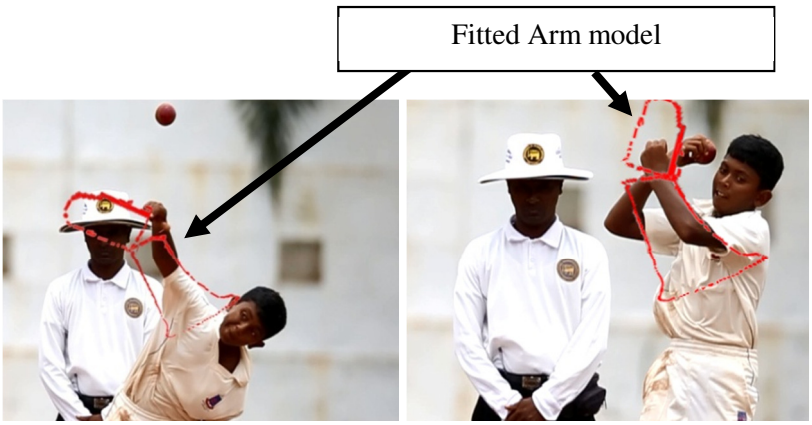


(a)

(b)



(c)



(d)

Figure 8: Qualitative results (a): In cricket, (b): In volleyball, (c):In tennis, (d): Miss classifications.

Figure 8 shows example inferences of our algorithm. The fitted arm model is indicated in red. After fitting the model in binary image, the model is super imposed in respective RGB frame. Regardless of the large variations in the dataset considering body and camera pose, depth in scene, cropping, and body size and shape (e.g. school athlete vs. heavy adult) a higher accuracy in both arm classification and tracking was achieved. The bottom row shows some misclassifications and failures due to subtle changes such as the crossed arms. Despite the fact that no temporal or kinematic constraints are being used for any of our computation, the results show higher accuracy in classification and tracking with lesser amount of jitter.

II. Quantitative results

We compared our approach to two related well known tracking methods. The first tracking method is the general Continuously Adaptive Mean Shift (CAMSHIFT) algorithm which was first introduced by Bradski et al. in (1998). The next algorithm, proposed by Klein et al. (2010) is an extended version of CAMSHIFT that enhances the general algorithm using Kalman filter. The input parameters of the Kalman filter, respectively, the position of the object in the image at time k, the size of the object and the width and length of the search window.

We conducted three experiments to evaluate the accuracy of the proposed algorithm. In the first two experiments we compared the algorithm against the above two algorithms in different gaming conditions. In the first experiment we analysed the tracking process under real gaming conditions in cricket, volleyball and tennis. Arm tracking process in practice sessions was evaluated in the second experiment. In order to evaluate the different setups, we ran each of the above experiment 100 times and evaluated following metrics.

(1) Overlap join-ratio percentage

$$\text{overlap join} - \text{ratio percentage} = \frac{G \cap T}{G \cup T} \times 100\%$$

Where G and T are the areas of the bounding boxes of the idealized configuration that is given by ground truth of the arm configuration and the tracking hypothesis respectively. Here as the measurement of the area we consider the number of pixels in each set. The ground truth is obtained from manual selection of the arm region. In order to minimize the human error in manual tracking process, we have used 3 different personals and assigned them the videos randomly.

(2) Centre Location Error

This evaluation matrix measures how much the evaluated centre of the arm region matches with the actual centre. The Euclidian distance, in pixels, between the actual centre pixel and the evaluated centre pixel in the arm bounding box is calculated as the Centre Location Error.

(3) Average false-Alarm rate.

How often an incorrect arm region is predicted by a particular algorithm is evaluated using this evaluation matrix.

$$\text{Average false alarm Rate} = \frac{\#\text{ of incorrect arm region detections}}{\#\text{ of trials } \times \#\text{ of arm regions in the video sequence}}$$

Finally in the third experiment we have evaluated the accuracy of the generated bio mechanical model against the ground truth. Hence the accuracy of the joint position in the generated bio mechanical model is evaluated.

i. Experiment 1: Real Game scenario

In this experiment we evaluated the three tracking algorithm against the real game scenario data set. For a fair comparison CAMSHIFT tracker and Extended CAMSHIFT tracker are initialized on the first frame using the actual location of the arm. The evaluated matrices are shown in the Table 2. In sequences A, B, D and F our proposed method has outperformed the other two algorithms. Only in sequence C the Extended CAMSHIFT tracker has slightly performed better than the proposed method due to the problem of player hand overlapping with face of the umpire. We can easily overcome such problems by slightly adjusting the camera position.

Table 2: Evaluation matrices for videos in real game environments

Seq.	# Fr.	Average overlap join-ratio %			Average Centre Location Error			Average False-Alarm rate.		
		CAMSHIFT	Extended CAMSHIFT	Proposed	CAMSHIFT	Extended CAMSHIFT	Proposed	CAMSHIFT	Extended CAMSHIFT	Proposed
A	51	50.3	71.7	87.7	5.9	3.1	1.2	0.370	0.290	0.016
B	32	44.4	84.5	95.8	8.6	3.6	1.9	0.210	0.027	0.520
C	71	50.7	73.1	72.8	7.2	5.1	6.9	0.630	0.075	0.024
D	45	77.3	80.3	86.0	7.5	7.2	5.3	0.170	0.041	0.020
E	80	69.2	80.4	85.0	5.0	2.5	4.7	0.191	0.014	0.082
F	62	61.7	76.4	83.1	6.1	4.4	3.0	0.139	0.199	0.027
avg		**58.9**	**77.7**	**85.1**	**6.7**	**4.3**	**3.8**	**0.285**	**0.108**	**0.115**

ii. Experiment 2: Practice sessions

The accuracy of the tracking algorithms in practice session situations are compared in this experiment. Similar to experiment 1, we manually initialised the CAMSHIFT tracker and Extended CAMSHIFT tracker on the first frame using the actual location of the arm. The evaluated matrices are shown in the Table 3. In all the sequences the proposed tracker has outperformed other two algorithms.

Table 3: Evaluation matrices for videos in practice sessions

Seq.	# Fr.	Average overlap join-ratio %			Average Centre Location Error			Average False-Alarm rate.		
		CAMSHIFT	Extended CAMSHIFT	Proposed	CAMSHIFT	Extended CAMSHIFT	Proposed	CAMSHIFT	Extended CAMSHIFT	Proposed
G	46	57.5	60.3	80.4	7.8	5.0	2.7	0.470	0.210	0.011
H	95	65.9	64.2	83.5	5.3	3.9	1.0	0.310	0.240	0.098
I	62	50.7	52.7	76.6	2.6	2.2	0.8	0.830	0.430	0.062
J	40	46.8	59.3	86.0	1.3	1.1	0.3	0.190	0.140	0.069
H	21	42.4	74.7	89.4	1.2	0.4	0.1	0.191	0.133	0.053
K	63	53.6	77.9	79.0	4.1	3.5	2.6	0.370	0.173	0.067
avg		**52.8**	**64.8**	**82.5**	**3.7**	**2.6**	**1.3**	**0.394**	**0.221**	**0.060**

iii. Experiment 3: Accuracy of the generated bio mechanical model.

Ten videos from each category were randomly picked and they were fed to the algorithm. To generate the bio mechanical model we consider wrist, elbow and shoulder joints. These three joints were manually tracked in the original video. The position of each joint in each frame in ground truth is compared against the joint position in the model. The Euclidian distance, in pixels, between the actual joint position and the inferred position is evaluated as the precision error. According to the results in Table 4, it can be seen that the generated bio mechanical model has a higher degree of precision. Figure 9 visualises a frame from the cricket data set and its respective bio mechanical model.

Table 4: Average precision error of the generated bio mechanical model

Seq.	# Fr.	Average precision error %		
		Wrist	Elbow	Shoulder
1	70	2.40	4.16	6.93
2	50	2.93	3.47	5.75
3	34	6.87	2.68	4.28
4	47	4.71	5.29	5.71
5	79	2.29	4.22	4.28
6	44	1.94	4.65	1.22
7	21	3.08	3.19	5.42
8	46	1.08	4.71	4.81
9	51	2.68	5.39	6.56
10	71	4.15	3.04	2.25
avg		**3.21**	**4.08**	**4.72**

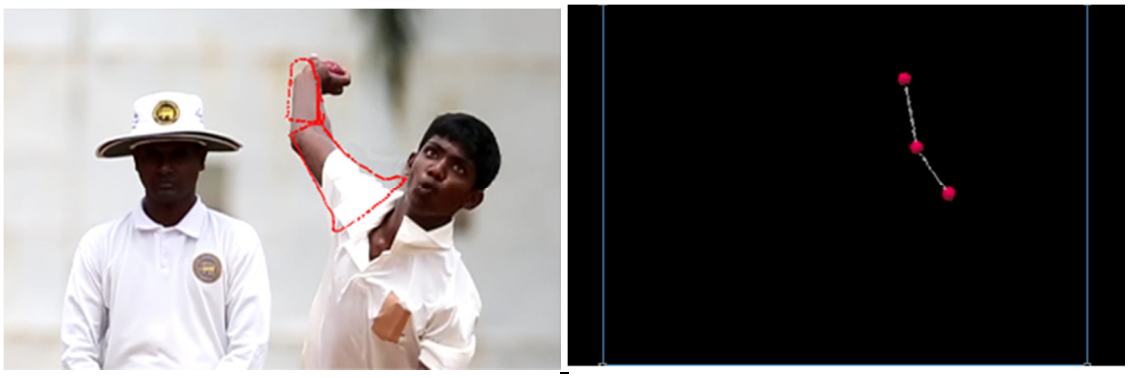


Figure 9: A frame from the cricket data set (Left) and its respective bio mechanical model (Right)

Conclusion

We have presented an effective method to track player arm motion during sporting activities. The algorithm was tested for tracking the delivery of the ball in cricket, and during the service in tennis and volleyball. We have validated the proposed algorithm against a large dataset which is generalized in terms of resolution and frame rates. The video sequences cover different challenges of object tracking such as object occluding and illumination inconsistencies.

Irrespective of the object size, position and the background complexities, the proposed

approach finds the optimal matching position and tracks the arm movement in the sporting activity. The quantitative comparison shows that the method performs equally or better to current state-of-the-art methods in object tracking. The effectiveness of proposed algorithm to track fast reflexes of professional athletes is vindicated by its higher degree of accuracy. For each of the three evaluation matrices the proposed apporach has out perforemed currently available tracking methods. We have observed over 80 percent accuracy when considering overlap join-ratio percentage. The mere score of 3.8 pixels, for average center location error and 0.115 for average false alarm rate elaborates how precisely our algorithm can track the arm movements even under complex real gaming conditions. Furthermore, this method exhibits more consistent results over lighting variations, stroke variations as well as gaming condition variations.

The novelty in our approach is the hybrid model to track the player arm movement. Cascade object classifier is learned to discriminate the target player face area from background and is utilized to estimate a region of interest for the arm area. This hybrid approach together with Active Shape Model was able to classify and track player arm movements and achieve cooperatively higher accuracy. Additionally, with the third experiment we have validated that the generated bio mechanical model has a higher degree of precision and hence the model can be directly applied for applications such as player delivery angel validations in cricket or serve speed detection in tennis.

According to best of our knowledge this is the only available method which is capable to track player arm movements in Cricket. The algorithm requires the arm to be over the shoulder of the athlete in order to be tracked. This doesn't impose any constraints, because vital factors such as serve, delivery speed or delivery angel are measured from the moment arm comes over the shoulder. But, if necessary, subsequent movements can be tracked easily by employing a method like particle filter, after detecting the initial arm region using this algorithm. Hence this algorithm is self-sufficient for tracking player arm movements in outdoor sporting activities.

In the future work, we aim to modify the presented ASM model to a hierarchical model and to accompany different resolution levels in order to reduce the number of iterations in model fitting process.

References

Bradski, G. (1998). Computer Vision Face Tracking For Use in a Perceptual User Interface. *Intel Technology Journal, 2*(1),705-740.

Bruhn, A., Weickert, J., & Schnörr, C. (2005). Lucas/Kanade Meets Horn/Schunck: Combining Local and Global Optical Flow Methods. *International Journal of Computer Vision, 61*(3), 211–231.

Choppin, S., & Wheat, J. (2013). The potential of the Microsoft Kinect in sports analysis and biomechanics. *Journal of Sports Technology, 6*(2), 53-54.

Duda, R., & Hart, P. (1972). Use of the Hough transformation to detect lines and curves in pictures. *Communications of the ACM, 15*(1), 11-15.

Dwarakanath, D., Eichhorn, A., Griwodz, C., & Halvorsen, P. (2012). Faster and more Accurate Feature-based Calibration for Widely Spaced Camera Pairs. *Proc of Digital Information and Communication Technology and its Applications, DICTAP 2012: Bangkok, Thailand, 16-18 May 2012,(pp. 87 - 92).* UTCC: University of the Thai Chamber of Commerce.

Galvin, B., McCane, B., Novins, K., Mason, D., & Mills, S. (1998). Recovering motion fields: An analysis of eight optical flow algorithms. *In Proc of British Machine Vision Conference, Southampton, BMVC 1998: UK, 12-14 September 1998, (pp. 43-77).* BMVC :British Machine Vision Association.

Hartley, R. (1997). In defense of the eight-point algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 19*(6), 580-593.

Horn, B., & Schunck, B. (1981). Determining optical flow. *Artificial Intelligence, 17*, 185–203.

Jarzem, P., & Gledhill, R. (1993). Predicting height from arm measurements. *Journal of paediatrics orthopaedics, 13*(6), 761-766.

Jones ,M., & Viola, P. (2003). Fast Multi-view Face Detection. *Proc of IEEE Computer Vision and Pattern Recognition,CVPR 2003: Madison, Wisconsin 16-22 June 2003,(pp.* 674 - 679*).* IEEE computer society.

Klein, D.,& Schulz, D. (2010). Adaptive real-time video-tracking for arbitrary objects. *Proc of Intelligent Robots and Systems ,IROS 2010: Taipei, Taiwan, 18-22 October, 2010 (pp. 772–777).* IEEE computer society.

Koschan,.A, Kang, S, & Paik, J. (2003).Color active shape models for tracking non-rigid objects. *Pattern Recognition Letters, 24,* 1751–1765.

Koschan,.A, & Kang, S. (2002). Video Object Tracking Based On Extended Active Shape Models With Color Information. *Proc of First European Conference on Colour in Graphics, Imaging, and Vision, CGIV 2002: Poitiers, France 2 -5 April , 2002 (pp. 126-131*). USA: Society for Imaging Science and Technology.

Lowe, D. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision, 60(2),* 91-110.

Lucas, B., & Kanade, T. (1981). An iterative image registration technique with an application to stereo vision. *Proc of Seventh International Joint Conference on Artificial Intelligence*, *IJCAI 1981: University of British Colombia, Vancouver, B.C, Canada, 24-28 August 1981, (pp. 674–679).* USA: The International Joint Conferences on Artificial Intelligence, Inc.

Rosenfeld, A., & Kak, A. (1982*). Digital Picture Processing*, vol. 1. Academic Press, Orlando, USA.

Shirgeri, S. & Naik, P. (2013). Design and development of Optical flow based Moving Object Detection and Tracking (OMODT) System. *International Journal of Computational Engineering Research*, *3*(4), 65-72.

Shotton, J., Fitzgibbon, T., Mat Cook, M., & Sharp T. (2011). Real-Time Human Pose Recognition in Parts from Single Depth Images. *Proc of IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2011: Colorado Springs, CO, 21-23 June 2011, (pp. 1297-1304).* USA: University of Colorado.

Singh, A., Kumar, A., Chavali, K., & Harish, D. (2012). Use of arm-span and foot length for estimation of height of the person. *Journal of Punjab Acad Forensic Med Toxicol*, *12*(2), 87-91.

Solar, J., & Verschae, R. (2006). Object detection using cascades of boosted classifiers. *EVIC: Chile*,*15-17 December, 2006, (pp. 123-127).* IEEE Computational Intelligence Society.

Sýkora, D., & Sedlácek, D. (2008). Real-time Color Ball Tracking for Augmented Reality. *Proc of Eurographics* Symposium *on Virtual Environments, EGVE 2008: Eindhoven, The Netherlands, 12-15 September 2008, (pp. 1-8).*EU: European Association of Virtual Reality and Augmented Reality.

WU, B., AI, H., Huang, C., & Lao, S. (2004). Fast Rotation Invariant Multi-view Face Detection based on Real Adaboost. *International conference on Face and Gesture Recognition, FGR 2004: Seoul, Korea, May 17-19, 2004,(pp. 79–84*). IEEE Computer Society.

www.asus.com/Multimedia/Xtion_PRO_LIVE/ 31 December, 2013

Yu, X., & Farin, D. (2005). Current and Emerging Topics in Sports Video Processing. *Proc of IEEE International Conference on Multimedia and Expo, ICME 2005: Amsterdam, The Netherlands, 6-8 July, 2005, (pp. 526 - 529).* IEEE Communications Society.

Calculation of Probability of Winning and Number of Games Played for Various Tournament Formats of the World Baseball Classic

Hirotsu, N.[1], Osawa, K.[2], Miyaji, C.[2]

[1]Juntendo University, Graduate School of Health and Sports Science

[2]Japan Sport Council, Japan Institute of Sports Sciences

Abstract

The third World Baseball Classic (WBC) was held in March 2013. In this tournament, 16 teams play in Round 1 under a round-robin (RR) format and 8 teams which advanced to Round 2 play under a modified double-elimination (MDE) format. This 2013 WBC format is compared with the formats such as the past two WBCs held in 2006 and 2009, from the aspect of the probability of winning the tournament and the probability distribution of the number of games played by the same teams. We make the comparison by changing the relative strength of teams, and demonstrate the difference between the tournament formats.

KEYWORDS: BASEBALL, DOUBLE-ELIMINATION, ROUND-ROBIN, TOURNAMENT, WBC

Introduction

The World Baseball Classic (WBC) is an international baseball competition. The main tournament of the third WBC was held in March 2013, and the Dominican Republic won the tournament. The basic structure of the main tournament consists of three rounds. In Round 1, 16 teams were divided into 4 pools, each of which consists of 4 teams, and competed in each pool. The top two teams from each of the four pools advanced to Round 2. The top two teams from each of the subsequent two pools advanced to single-elimination (SE) in Finals.

This basic structure has not been changed since the first WBC held in 2006, but the tournament format of each pool is different between the WBCs. In the 2006 WBC, a round-robin (RR) format was employed in Rounds 1 and 2. Under this format, there were ties in 2 out of 6 pools. (We here call “tie” when three teams resulted in the same win-loss record).

In the 2009 WBC, the RR format was replaced by a modified double-elimination (MDE) format, in which any ties do not occur. However, the same two teams faced each other often in the tournament under the MDE format, such that Japan faced South Korea in 5 out of 9 games which Japan played throughout the tournament. Further, Game 6 in MDE is thought to be virtually redundant because Game 6 is played between the top two teams which have already been decided to advance to the next round. Actually, the manager of South Korea commented that they didn’t put great meaning to winning or losing (Schwarz, 2009) and good pitchers were not introduced to Game 6 intentionally (Nakamura, 2009). Although the matches in the next round are set up following the result of Game 6 and this somehow affects the final result

of the tournament consequently, the existence of Game 6 hardly affects the probability of winning the tournament as shown later in this paper.

In the 2013 WBC, the MDE format was replaced by a RR format only for Round 1. In the process of discussion to decide the format of the 2013 tournament, the format deleting Game 6 from the MDE format was suggested (Okada, 2010). We here denote it as modified modified-double-elimination format (MMDE), and also look at it for the comparison of the formats.

In this paper, we present the formulation for calculation of probabilities on the above formats, and compare between the four tournament formats (i.e. the 2013 tournament format, the past two WBCs' formats, and the case of employing the MMDE format), from the aspect of the probability of winning the tournament and the probability distribution of the number of games played by the same teams. We make the comparison by changing the relative strength of teams, using the Bradley-Terry model (Bradley & Terry, 1952) for setting the probability of winning a game between two teams, and demonstrate the difference between these tournament formats. We focus on the main tournament in this paper, although a qualifying round played by 12 teams was first introduced before the main tournament in the 2013 WBC.

By the way, a variety of sport tournament structures are well studied by a lot of researchers. Designing sporting contest is one of the major issues (e.g. Szymanski, 2003; O'Donoghue, 2005), and the impact of seeding in SE format has well been analyzed (Hwang, 1982; Marchand, 2002; Monks & Husch, 2009; Schwenk, 2000). For the impact of seeding under the different type of tournament formats, Scarf & Yusof (2011) and Scarf, Yusof, & Bilbao (2009) study the seeding policy in terms of RR and SE and hybrids of them, which are used in soccer World Cup Finals or other tournaments, and propose the tournament metrics to measure the success of the tournament. In terms of modelling a major soccer tournament, Koning, Koolhaas, Renes, & Ridder (2003) propose a practical model to identify the team that is most likely to win a tournament. The design of play-off system of football or baseball is also studied in connection to the seeding problem (Martine & Troendle, 1999; Carlin & Stern, 1999, Annis & Wu, 2006).

In terms of efficiency for selecting the strong teams in comparison between the SE and the RR format (e.g. David, 1959; Appleton, 1995; McGarry & Schutz, 1997), there is the following tendency: Under the RR format, the stronger team is more likely to win the tournament, but the number of total games in the tournament is larger than in SE, in which the stronger team is less likely to win the tournament than in the RR format. In terms of the double-elimination (DE) format, McGarry & Schutz (1997) indicate that the DE format has the well balanced feature in that it yields selection of stronger teams with a moderate total number of games.

With regards to the calculation for the DE format, several studies provide the calculation methods (Glenn, 1960; Ladwig & Schwertman, 1992; Edwards, 1996; McGarry & Schutz, 1997), although they did not consider multi-round structures such as the WBC tournament. Stanton & Williams (2013) study the DE format considering it as two SE formats which consist of the winner bracket and the loser bracket, and investigate the design of the two brackets.

We here use the method along the lines of Ladwig & Schwertman (1992) to calculate the probability of winning the tournament and the probability distribution of the number of games in terms of the multi-round structure including the RR format and/or the MDE format in the first two rounds and the SE format in the finals. We make a comparison of the four formats by changing the relative strength of teams, along the lines of Marchand (2002) who uses the Jackson's (1993) model.

In this paper, we describe the formats employed in the WBCs in Section 2, and then we present the formulation for calculation of probabilities on these formats in Section 3. After showing how to set a probability of winning a game in Section 4, we illustrate and discuss the difference between these formats in Section 5 and conclude in Section 6.

Tournament format of the WBCs

In this section, we present the format of the main tournament of the WBCs. The format consists of three rounds. In Round 1, 16 teams are divided into 4 pools (A, B, C and D) each of which consists of 4 teams, and the top two teams in each pool advance to Round 2. In Round 2, the 4 teams from Pools A and B and the 4 teams from Pools C and D compete in Pools 1 and 2, respectively. The top two teams in each pool of Round 2 advances a single-elimination in Finals. The 4 teams cross over for the semifinals, with the winner of each pool playing against the runner-up from the other pool. Figure 1 shows the main structure of the tournament, which is consistent between the WBCs, but the format in each pool is different between the WBCs.

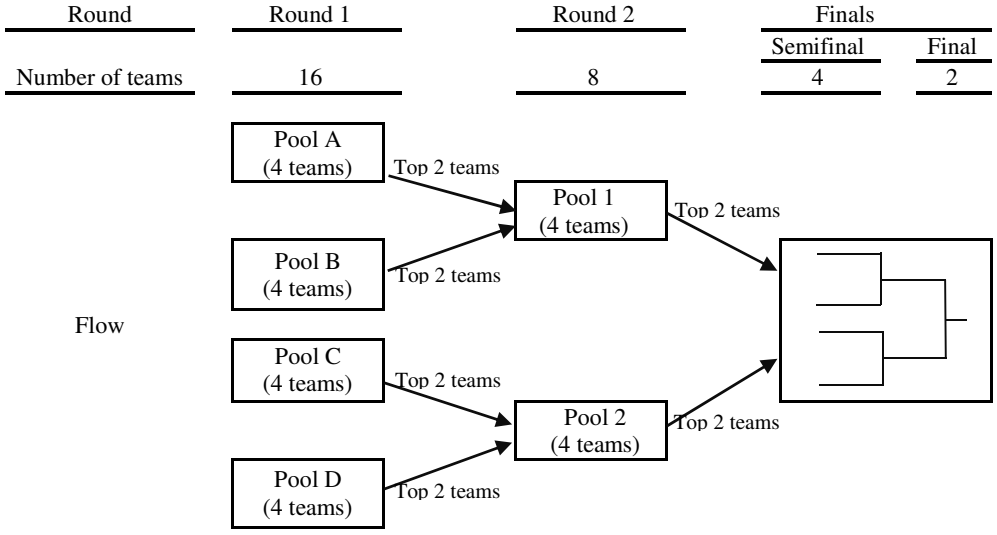
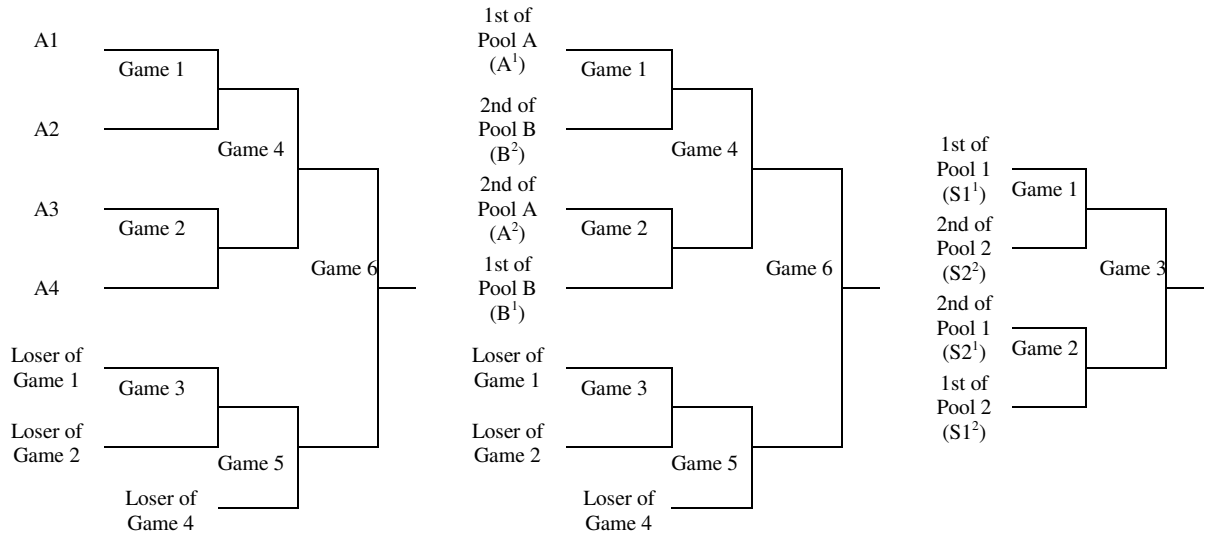


Figure 1. Main structure of the tournament format of the WBCs.

In the 2006 tournament, the RR format was employed in Rounds 1 and 2. That is, each team plays other three teams in a pool once. Teams are ranked by the winning percentage in each round, and the top two teams in each pool advance to the next round. (If ties occur, the top two teams allowing the fewest runs per nine innings in head-to-head games between the tied teams were qualified to the next round).

In the 2009 tournament, the RR format was replaced by the MDE format. This MDE format is illustrated in Figure 2. Figure 2 (a) and (b) shows the draw of Pool A of Round 1 and the draw of Pool 1 of Round 2, respectively, as an example. Figure 2 (c) shows the draw of the Finals.



(a) Pool A of Round 1 (b) Pool 1 of Round 2 (c) Finals

Figure 2. Modified double-elimination format (Rounds 1 and 2) and single-elimination (Finals) in WBC 2009

In the 2013 tournament, the MDE format was replaced by the RR format only for Round 1. Together with this replacement to the RR format, the tie breaking rule was modified a little. (i.e. not only allowing runs per nine innings but also scoring runs per nine innings in head-to-head games between the tied teams are evaluated for advancing to Round 2).

We summarize the feature of the four formats in Table 1. The 2013 tournament format is indicated by RR&MDE, and other three formats (the past two classics and the MMDE format) are indicated by RR^2, MDE^2 and $MMDE^2$, respectively. We note that the total number of games in $MMDE^2$ is 6 games less because of the absence of Game 6 in each pool.

Table 1. Comparison between the tournament formats

Abbreviation		RR^2	MDE^2	RR&MDE	$MMDE^2$
WBC No.		1	2	3	-
Format	Round 1	RR	MDE	RR	MMDE
	Round 2	RR	MDE	MDE	MMDE
Total number of games		39	39	39	33
Max. number of played by same teams		3	5	4	3
Ties		Yes	No	Yes	No

In terms of the main structure of the tournament, it is controversial to maintain the general structure shown in Figure 1 (e.g. http://en.wikipedia.org/wiki/2006_World_Baseball_Classic). Other competitions such as the soccer World Cup have a structure not to make teams play each other more than twice. Although there are the pros and cons of this general structure, we here calculate the probability based on this general structure.

Formulation for calculation of probabilities

In this section, we present a formulation for calculation of probabilities in terms of the RR

format and the MDE format. In this paper, we assume that the result of the game is mutually independent throughout the tournament.

Round-robin format within a round

We first look at a formulation for calculation of probabilities in terms of the RR format of a pool within a round. Let $P(i,j)$ be the probability of team i winning a game against team j. Using this notation, for example, an event that "Team A1 wins 3 games, Team A2 wins 2 and loses 1 game, Team A3 wins 1 and loses 2 games, Team A4 loses 3 games" occurs in the probability of

$$P(\mathrm{A}1,\mathrm{A}2)P(\mathrm{A}1,\mathrm{A}3)P(\mathrm{A}1,\mathrm{A}4)P(\mathrm{A}2,\mathrm{A}3)P(\mathrm{A}2,\mathrm{A}4)P(\mathrm{A}3,\mathrm{A}4)\,. \qquad (1)$$

Here, the total number of possible win-loss events in a pool is 2^6=64. These events are categorized into four patterns shown in Table 2. We can calculate the probability of each of the four patterns occurring by just counting the number of events in each category. In Table 2, ties occur in Win-loss patterns 2 and 3. In Win-loss pattern 2, three teams are tied with the record of 1 win and 2 losses, and these teams are ranked as 2nd, 3rd and 4th, following the tiebreaking rule. In Win-loss pattern 3, three teams are also tied with 2 wins and 1 loss, and ranked as 1st, 2nd and 3rd by the tiebreaking rule.

Table 2. Win-loss pattern of the round-robin format

	Win-loss pattern			
	1	2	3	4
1st	3-0	3-0	2-1	2-1
2nd	2-1	1-2	2-1	2-1
3rd	1-2	1-2	2-1	1-2
4th	0-3	1-2	0-3	1-2

Modified double-elimination format within a round

We move to look at the formulation for calculation of probabilities in terms of the MDE format of a pool within a round. To obtain this probability, we look at A1 in Figure 2(a). As there are 6 games in the draw of Figure 2(a), the total number of possible win-loss events is 2^6=64. As there are 4 teams in a pool, a quarter of the 64 events corresponds to the events that A1 is ranked as the 1st in this pool. That is, there are 16 (=64/4) win-loss events for A1 to be ranked as the 1st in Pool A of Round 1. We arrange these 16 events according to the 2nd ranked team in this pool, as shown in Table 3.

Table 3. Probability of a win-loss event occurring and the number of games between two teams when A1 becomes the 1st in the draw shown in Figure 2 (a) in the MDE format

No.	Win-loss event				Number of games							Remark
	1st		2nd		Probability	A1-A2	A1-A3	A1-A4	A2-A3	A2-A4	A3-A4	
1	A1	$W_1W_4W_6$	A2	$L_1W_3W_5L_6$	$P_{1T}(A1,A2)_1$	2	1	0	1	1	1	A3 wins in Game 2
2					$P_{1T}(A1,A2)_2$	2	0	1	1	1	1	A4 wins in Game 2
3		$L_1W_3W_5W_6$	A2	$W_1W_4L_6$	$P_{1T}(A1,A2)_3$	2	1	1	1	0	1	A3 wins in Game 2
4					$P_{1T}(A1,A2)_4$	2	1	1	0	1	1	A4 wins in Game 2
5	A1	$W_1W_4W_6$	A3	$W_2L_4W_5L_6$	$P_{1T}(A1,A3)_1$	1	2	0	1	1	1	A2 wins in Game 3
6					$P_{1T}(A1,A3)_2$	1	2	0	0	1	2	A4 wins in Game 3
7			A3	$L_2W_3W_5L_6$	$P_{1T}(A1,A3)_3$	1	1	1	1	0	2	
8		$W_1L_4W_5W_6$	A3	$W_2W_4L_6$	$P_{1T}(A1,A3)_4$	2	2	0	0	1	1	A2 wins in Game 3
9					$P_{1T}(A1,A3)_5$	1	2	1	0	1	1	A4 wins in Game 3
10		$L_1W_3W_5W_6$	A3	$L_2W_3W_5L_6$	$P_{1T}(A1,A3)_6$	2	1	1	1	0	1	
11	A1	$W_1W_4W_6$	A4	$W_2L_4W_5L_6$	$P_{1T}(A1,A4)_1$	1	0	2	1	1	1	A2 wins in Game 3
12					$P_{1T}(A1,A4)_2$	1	0	2	1	0	2	A3 wins in Game 3
13			A4	$L_2W_3W_5L_6$	$P_{1T}(A1,A4)_3$	1	1	1	0	1	2	
14		$W_1L_4W_5W_6$	A4	$W_2W_4L_6$	$P_{1T}(A1,A4)_4$	2	0	2	1	0	1	A2 wins in Game 3
15					$P_{1T}(A1,A4)_5$	1	1	2	1	0	1	A3 wins in Game 3
16		$L_1W_3W_5W_6$	A4	$L_2W_3W_5L_6$	$P_{1T}(A1,A4)_6$	2	1	1	0	1	1	

W_k and L_k in Table 3 represent a win and a loss of the regarding team in game *k*, respectively. For example, the win-loss event that A1 is the 1st and A2 is the 2nd in this pool appears in No.1 - 4 rows in Table 3, and if A1 achieves 3 wins ($W_1W_4W_6$), then A2 has to result in a 2 wins and 2 losses ($L_1W_3W_5L_6$) as shown in No.1 - 2 rows. In this event, there are two cases which correspond to the different result of Game 2, that is, either A3 or A4 wins in Game 2. Looking at No.1 row, this event occurs when "A1 wins A2 in Game 1, A3 wins A4 in Game 2, A2 wins A4 in Game 3, A1 wins A3 in Game 4, A2 wins A3 in Game 5 and A1 wins A2 in Game 6" in probability of

$$P_{1T}(\mathrm{A1},\mathrm{A2})_1 = P(\mathrm{A1},\mathrm{A2})P(\mathrm{A3},\mathrm{A4})P(\mathrm{A2},\mathrm{A4})P(\mathrm{A1},\mathrm{A3})P(\mathrm{A2},\mathrm{A3})P(\mathrm{A1},\mathrm{A2}). \quad (2)$$

By summing up the probabilities regarding to the cases of "A1 is the 1st and A2 is the 2nd" appearing in No.1 - 4 rows, the probability that A1 and A2 become the 1st and the 2nd in Round 1 can be obtained as

$$P_{1T}(\mathrm{A1},\mathrm{A2}) = \sum_m P_{1T}(\mathrm{A1},\mathrm{A2})_m . \quad (3)$$

(The subscript *m* corresponds to the subscript of the notation appearing in the column "Probability" in Table 3). In the same manner, not only P_{1T}(A1,A3) and P_{1T}(A1,A4) but also other teams such as A2 to be the 1st in this pool can be calculated.

Here, note that the 4 teams should be randomly assigned in the draw in Pool A. So, the above calculations are conducted for the 3 different assignments, and then averaged out. That is, we

consider the 3 different draws for assigning 4 teams in the MDE format in a pool such that A1 faces A2, A3 or A4 as the first game of A1, and the 3 different draws occur in 1/3 each.

As shown in Table 3, the number of games between two teams in the pool is fixed, according to its win-loss event. For example, if A1 is the 1st and A2 is the 2nd with the event of $W_1W_4W_6$ and $L_1W_3W_5L_6$, A1 has to play A2 twice in the pool.

Using the number of games played by two teams, we can calculate the probability distribution of the number of games between two teams. Here, let $Q_1(i,j,n_1|\mathrm{A1,A2})$ be the probability that team i plays against team j (i、 $j \in \{\mathrm{A1,A2,A3,A4}\}$) n_1 times in Round 1, under the condition that A1 and A2 results in the 1st and the 2nd in the pool, respectively. We calculate this using the probabilities shown in Table 3 such that

$$Q_1(\mathrm{A1,A3},1|\mathrm{A1,A2}) = P_{1\mathrm{T}}(\mathrm{A1,A2})_1 + P_{1\mathrm{T}}(\mathrm{A1,A2})_3 + P_{1\mathrm{T}}(\mathrm{A1,A2})_4 . \quad (4)$$

Here, (4) corresponds to the probability that A1 and A3 play once in the pool, and this probability can be obtained by summing up the probabilities for resulting in the win-loss events corresponding to the number of games "1" appearing in No.1 - 4 rows in the column of "A1 - A3" in Table 3.

Calculation between the rounds

So far, we looked at the format in a round, actually in Round 1. In order to formulate the multi-round structure, we should identify the 8 teams advanced from Round 1. Here, we denote the 1st and the 2nd team in Pool A and Pool B as A^1, $\mathrm{A}^2 \in \{\mathrm{A1,A2,A3,A4}\}$ and B^1, $\mathrm{B}^2 \in \{\mathrm{B1,B2,B3,B4}\}$, respectively.

Using this notation, in terms of the RR format in Round 2, we can calculate the probability of each of the four patterns occurring shown in Table 2 just by looking at A^1, A^2, B^1 and B^2. When a tie occurs, the advanced teams are chosen from tie teams in equal probability.

In terms of the MDE format in Round 2, the calculation becomes a little complicated because the 4 teams are assigned into the draw of Pool 1 according to the rank in pool A and B as shown in Figure 2(b). As A^1 and A^2 are chosen from the 4 teams in Pool A in $4\times3 = 12$ different ways, and so as to B^1 and B^2, there are $12\times12=144$ different ways in total for the 4 teams to be assigned in the draw of Pool 1. By calculating each of these 144 different ways, we can obtain such probability that A1 advances to Finals from Pool 1.

In Finals, we should calculate in the SE format shown in Figure 2(c) using 4 teams advanced from Round 2, both in the RR and MDE formats. Here, we denote the 1st and the 2nd team in Pool 1 and Pool 2 as $\mathrm{S1}^1$, $\mathrm{S1}^2 \in \{\mathrm{A1,A2,A3,A4,B1,B2,B3,B4}\}$ and $\mathrm{S2}^1$, $\mathrm{S2}^2 \in \{\mathrm{C1,C2,C3,C4,D1,D2,D3,D4}\}$, respectively. Using this notation, the draw of semifinals is fixed according to the rank in Pools 1 and 2. As $\mathrm{S1}^1$, $\mathrm{S1}^2$ are chosen from the 8 teams in $8\times7 = 56$ different ways, and so as to $\mathrm{S2}^1$, $\mathrm{S2}^2$, there are $56\times56=3{,}136$ different ways in total for the 8 teams to be assigned in the draw of semifinals. By calculating for these 3,136 different ways, we can obtain such probability that A1 wins the tournament.

In order to calculate the probability of the total number of games between team i and j occurring throughout the tournament in the MDE format, we need to sum up the number of games from Round 1 to Finals. In practice, we refer to Table 4, which shows the patterns of the number of games between two teams from Round 1 to Finals, and calculate the probability of the total number of games occurring.

Table 4. Patterns of the number of games between two teams in the MDE2 format

Round	Times	Number of games between two teams																										
Round 1	n_1	2	2	2	2	1	2	2	1	1	1	-	2	2	1	1	-	-	-	1	1	-	-	-	0	-	-	-
Round 2	n_2	2	2	2	1	2	1	1	2	2	1	2	0	-	1	1	2	1	2	0	-	1	1	-	-	0	-	-
Finals	n_F	1	0	-	1	1	0	-	0	-	1	1	-	-	0	-	0	1	-	-	-	0	-	0	-	-	0	-
Total	n	5	4	4	4	4	3	3	3	3	3	3	2	2	2	2	2	2	2	1	1	1	1	1	0	0	0	0

As shown in Table 4, in order for two teams to face each other 5 times they should face twice in Round 1, and twice in Round 2, and once in Finals. We here denote this as "2 2 1". In order to face 4 times, there are four patterns such that "2 2 0", "2 2 -", "2 1 1" and "1 2 1", as shown in Table 4. The notation " - " denotes "not-faced" because of being assigned in a different pool, or one of two teams failing to advance to the round. The probability of each pattern occurring is calculated by taking account of the conditional probability to advance to Round 2 or Finals. For example, in order to calculate the probability that "2 2 1" occurs, we look at the following event: A1 and A2 face each other twice (n_1=2) in Pool A under the condition that A1 and A2 become the 1st and the 2nd of the pool in Round 1. A1 and A2 then face each other twice (n_2=2) in Pool 1 under the condition that A1 and A2 become the 1st and the 2nd of Pool 1 in Round 2. A1 and A2 finally face each other once (n_F=1) in Finals. This event occurs in probability which is given by the product of the above probabilities, such that,

$$Q_1(\mathrm{A1},\mathrm{A2},2\,|\,\mathrm{A1},\mathrm{A2})\times Q_2(\mathrm{A1},\mathrm{A2},2\,|\,\mathrm{A1},\mathrm{A2})\times Q_F(\mathrm{A1},\mathrm{A2},1). \quad (5)$$

By conducting the similar calculation for all patterns shown in Table 4, we can obtain the probability distribution of the number of games between two teams throughout the tournament.

Setting of strength of teams

Until now, we present the formulation for calculation of probabilities, but we have not referred to concrete values of the probability P(i,j). In actual calculation we need to set these probabilities as concrete values. In this paper, the Bradley-Terry model (1952) is used for setting the probability of team i winning the game against team j. That is, we calculate

$$P(i,j) = \pi_i \,/\, (\pi_i + \pi_j) = 1 \,/\, (1 + \pi_j / \pi_i), \quad (6)$$

where π_i and π_j represent the strength of team i and j, respectively.

Here, we set π_i following the Jackson's (1993) model, in which $\pi_i = r_i^{\alpha}$, where, r_i is the relative ranking of team i such that r_i=1 when team i is top-ranked. α is a parameter to set the dispersion of the relative strength. We change the value of α from 0 to 1.5 as Marchand (2002) calculated the probabilities of a player winning a standard knockout tournament and a random knockout tournament. In teams of the standard knockout tournament, Marchand assigned 16 teams to consist in matching higher ranked teams with lower ranked teams. To compare with his result, we set up "Case S" in which the match-ups of first round are basically same structure as his standard knockout tournament. That is, we assign 16 teams into 4 pools in "Case S" as shown in Figure 3. Note that the numbers in Figure 3 indicate the teams as its relative rankings. Without loss of generality, we show the result by setting A1 be the strongest team and arrange the relative strengths such that $\pi_{A1} > \pi_{A2} > \pi_{A3} > \pi_{A4}$ in descending order later in this paper. (In "Case S", the relative ranking of A1, A2, A3 and A4 are 1, 8, 9 and 16, respectively).

In terms of the random knockout tournament of Marchand, it consists in selecting the initial structure totally at random. He calculated the probabilities of a top-seed player winning a tournament for the random knockout tournament by simulating 1 million draws according to a uniform distribution and estimated each probability by averaging them out. However, we compare the four formats from the aspect of not only the probability of winning the tournament but also the number of games between two teams. For calculating these probabilities, we have developed our programing code with Excel VBA, and it takes several minutes to complete the calculation of one structure (considering all different ways of advancement of the 16 teams mentioned in Section 3) using our computer (CPU: Intel Celeron M (1.0GHz)). As it would take more than decades to calculate the probability distribution of the number of games between two teams based on a large number of random sampling in a similar manner to Marchand, we set up just three cases, a standard ("Case S") and two other special cases "Case F" and "Case U", as shown in Figure 3 for comparing the four formats, and see the effect of the selection of the initial structure on the result.

Here, in "Case F", each pool is fairly balanced such that the 16 teams are assigned by its relative ranking in its order, although it is not perfectly fair because Pool A is relatively stronger than other Pools. In "Case U", the assignment of teams to each pool is definitely unbalanced such that the top 4 teams are in Pool A. "Case F" and especially "Case U" are unlikely to occur in practice, but we set them as extreme cases to see the extent of variation of the effect of initial structure on the probabilities. In this sense, most of all possible assignments of 16 teams will result in the range between Cases S and U in practice.

Case S	Case F	Case U
1,8,9,16	1,5,9,13	1,2,3,4
2,7,10,15	2,6,10,14	5,6,7,8
3,6,11,14	3,7,11,15	9,10,11,12
4,5,12,13	4,8,12,16	13,14,15,16

Pool A
Pool B
Pool C
Pool D

Figure 3. Three cases of the assignment of 16 teams to each Pool

We further note that it is better to randomly assign teams to the 4 pools because the number of game faced by two teams are significantly depends on which pool these teams are assigned. So, we take into account three different initial structure for assigning 4 pools shown by the arrows in Figure 3, such that {1,2,3,4} to Pool A, {5,6,7,8} to Pool C, {9,10,11,12} to Pool B, and {13,14,15,16} to Pool D as one of the three initial structures.

Furthermore, we also calculate the case of the knockout tournament of 16 teams for comparison. In terms of Case S, it is basically the same case as the standard knockout tournament of Marchand (2002). In terms of Cases F and U, we assigned the 4 pools to the four separate blocks in the initial draw of the knockout tournament, such as assigning {1,5,9,13}, {2,6,10,14}, {3,7,11,15} and {4,8,15,16} separately in the initial draw.

Calculation results

In this section, we demonstrate the calculation results between the four formats based on the above three cases of assignment of 16 teams. By changing the value of α, we have calculated such the probability that A1 wins the tournament, or the probability distribution of the number

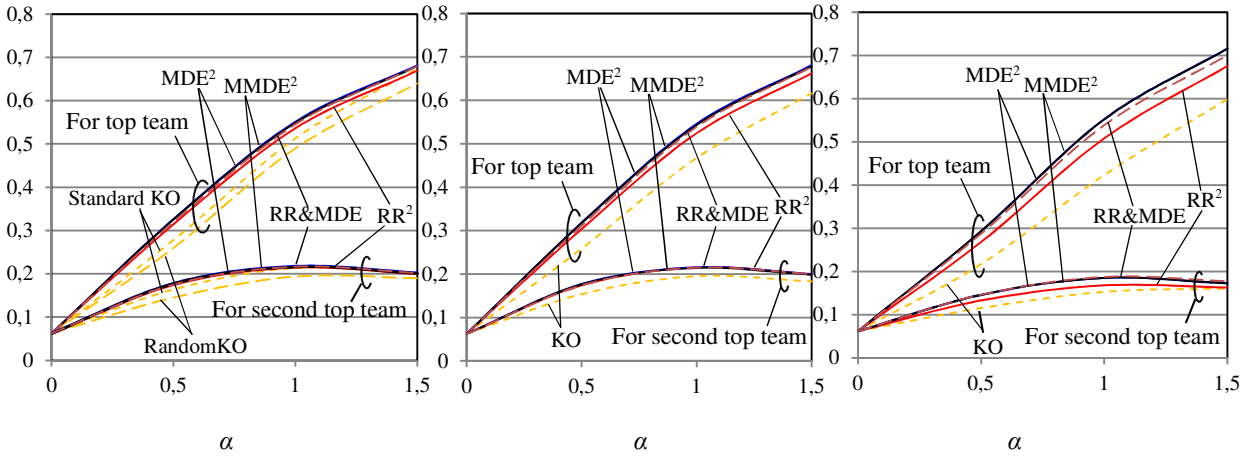
of games between A1 (here setting A1 be the strongest team) and others. We also present the detail about the occurrence of ties in the RR format.

Comparison between the tournament formats in terms of probability of winning the tournament

Figure 4 shows the comparisons of the probabilities of the top two teams winning the tournament in Cases S, F and U. For comparing them with the knockout tournament, we also draw the curve of the standard and the random knockout from Marchand (2002), as indicated by "standard KO" and "random KO", respectively in Figure 4 (a). In Cases F and U, we also show the calculation result of the case of the knockout tournament of 16 teams, as indicated by "KO" in Figures 4 (b) and 4 (c).

As shown in Figure 4, in terms of the probability of the top team winning all of the four formats are higher than that of the knockout tournaments throughout the range of α=0～1.5. The difference of the probabilities in the four formats is small, especially the probabilities of the MDE^2 and $MMDE^2$ formats are almost the same, although the RR^2 format of Case U in Figure 4 (c) is a little different from the other formats.

Not surprisingly, the probabilities of the second top team winning the tournament are definitely smaller than these of the top team. They have the similar tendency with regard to the difference of the probabilities between the four formats.



(a) Case S (b) Case F (c) Case U

Figure 4. Comparisons of the probabilities of the top two teams winning the tournament.

In order roughly to estimate the extent of the difference between the probabilities, we just statistically check the distribution of these winning probabilities of the four formats, by considering random distributions of teams' relative strengths. In practice, we generated normal random numbers ε ～ N(0,σ2) and added to πj in expression (6) for all teams. When σ is set to be 10% of relative strength, for instance, the standard deviation of the probability of top team winning is around 0.03～0.05 for all cases in the range of around α=1.0～1.5. Assuming that sample size n=25 (WBC is assumed to be held 25 times in the next hundred years), significant difference of the probabilities of winning will be roughly estimated to around 0.02～0.03 using two sample t-test with significance level of 0.05. Using this rough estimate in Figure 4, we can find the difference between RR2 and other formats in Case U, but for almost all of the other cases, we cannot find statistical differences between the four formats. (When σ is large, it

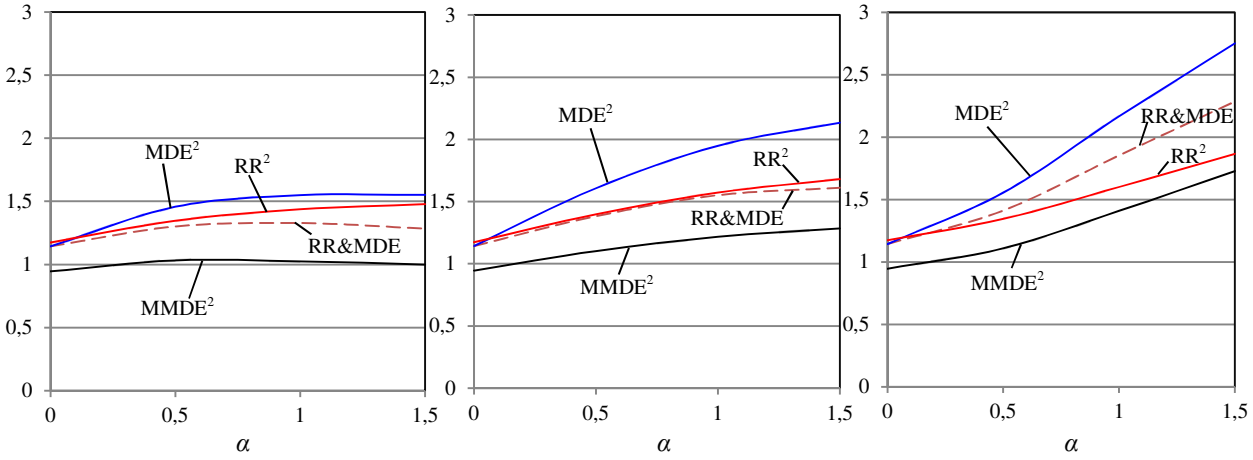
becomes more difficult to distinguish the difference between the four formats in terms of the probability of winning the tournament, although the sample size will affect the estimation of statistical significance.)

Calculation result on the number of games between two teams

We now look at the probability distribution of the number of games between two teams. The calculation result of the expected number of games faced between the top two teams within Pool A (i.e. A1 and A2) is shown in Figure 5. Here, we can see that the profiles of the expected number of games are different between Cases S, F and U. This is caused by an effect of the seeding structure. That is, in Case S, two top teams are not faced so much, but in Case U, definitely the top two teams have to face often. The result of Case F is roughly in the range between Cases S and U.

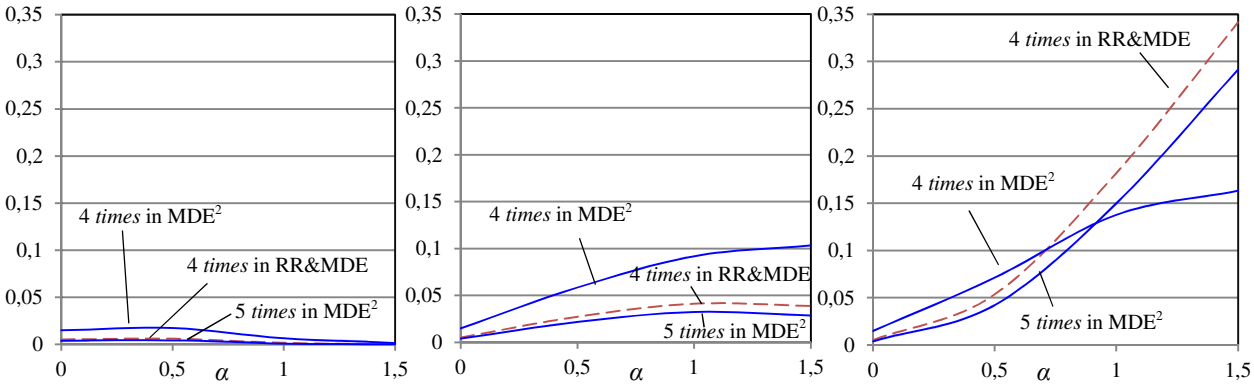
In terms of the difference between the four formats in each case, MDE^2 has the highest expected number of games between A1 and A2 than other formats. Especially in Case U, this tendency is distinct. Further, it would be interesting that $MMDE^2$ has clearly lower expected number, not depending on these three Cases.

In order to look at the detail of this tendency, the probabilities of facing 4 or 5 times in the tournament between A1 and A2 are shown in Figure 6. We can confirm that the seeding structure will definitely affect the number of games faced by two teams. That is, in "Case S" there are almost no differences between the four formats. Even MDE^2 has a small probability for A1 and A2 to face 4 or 5 games. However, in "Case U" A1 and A2 face 5 times in the probability of around 0.3 when α is around 1.5.



(a) Case S (b) Case F (c) Case U

Figure 5. Comparisons of the expected number of games faced between the top two teams within Pool A (i.e. A1 and A2).



(a) Case S (b) Case F (c) Case U

Figure 6. Comparisons of the probability of facing 4 or 5 times in the tournament between the top two teams within Pool A (i.e. A1 and A2).

Calculation result on occurrence of ties in round-robin format

In terms of the occurrence of ties in the RR format, the probabilities of ties occurring with the same win-loss record in a pool are shown in Figure 7. Here, we calculated this probability with regard to the 12 different assignments of 4 teams such as {1,8,9,16} and {2,7,10,15} in a pool shown in Figure 3. The probabilities of ties occurring are more than 0.15 for all of the 12 cases. From this result, the fact that ties occurred in 2 out of 6 pools in WBC 2006 and in 1 out of 4 pools in WBC 2013 is not curious. We do not discuss the tie breaking rule in detail here, but when the RR format is included in the tournament, as the probabilities of ties occurring are not small, the effect of the tie breaking rule will not be small on the selection of the top two teams in a pool.

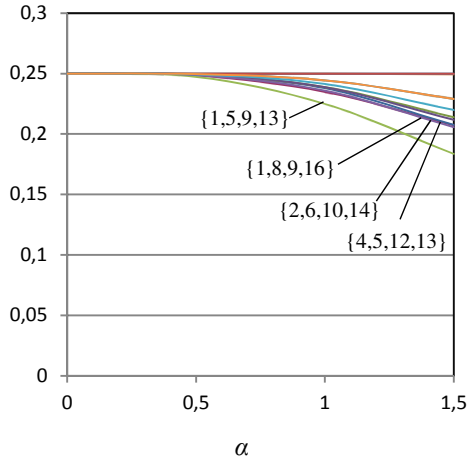


Figure 7. Probability of ties occurring in a pool in a round-robin format

Discussion

From the above calculation results, we here discuss our findings with respect to the probability of winning and the number of games played in the tournament.

In terms of the probability of winning, other literature such as McGarry & Schutz (1997) calculated it in such formats as SE, RR and DE. We calculated it in terms of four other

formats, and found that the difference of the probabilities of the top two teams winning are small, especially between MDE^2 and $MMDE^2$, as shown in Figure 4. Thus, the existence of Game 6 in MDE^2 hardly affects the probability of winning the tournament. Considering the small difference of the probability of winning and the total number of games in the tournament, $MMDE^2$ format would be more efficient than other three formats.

We also looked at the effect of the assignment of 16 teams to 4 pools on the probability of winning using Cases S, F and U. As the seeding favors stronger teams in general (Scarf & Yusof (2011), Monks & Husch (2009) and Scarf, Yusof, & Bilbao (2009)), the balanced assignment, Case S, would be naturally imagined to be preferable for the top team. However, when tournaments are unbalanced because the top two teams are assigned to the same pool, as in Case U, the probability of the top team winning the tournament is greater than in Case S in terms of MDE^2, $MMDE^2$, and RR&&MDE, in the area where α is large, as shown in Figure 4. That is, somewhat surprising, Case U seems to be preferable for the top team rather than Case S, when strength of teams are disparate. This phenomenon could be explained in part by the number of games between the top two teams. By comparing the calculation results of Cases S and U, the probability of facing 4 or 5 times in the tournament between the top two teams in Case U is definitely larger than in Case S, as shown in Figure 6, especially in the area where α is large. That is, the increase of the number of games between the top two teams seems to positively effect on differentiating the two teams' probability of winning the tournament.

In terms of the probability of the number of games, other literature such as Glenn (1960) calculated the expected total number of games in such formats as SE, RR and DE. We here calculated the probability distribution of the number of games by the same two teams. According to our calculation, it is not unusual for two teams to face 4 or 5 times in MDE^2 format, as shown in Figure 6(c). Thus, events such Japan facing South Korea 5 times in the 2009 WBC under MDE^2 format will plausibly occur, especially when two strong teams are assigned to the same pool.

In terms of reducing the expected number of games played by the same two teams, we have quantitatively confirmed that $MMDE^2$ is better than RR&MDE and RR^2, and much better than MDE^2, as shown in Figure 5. In this sense, $MMDE^2$ is balanced for the same two particular teams which do not face often.

Thus, $MMDE^2$ seems to be efficient and balanced from the standpoint of selecting the top teams avoiding 4 or 5 time same match-ups throughout the tournament. If the total number, 33, of games is acceptable for conducting the WBC, $MMDE^2$ should be recommended. Otherwise, if up to 4 games played by same teams is acceptable, the current RR&MDE format could be a reasonable selection from the four formats. In the use of the RR format, the tie breaking rule should be carefully designed for the selection of top two teams in a pool.

Conclusions

In this paper, we have presented the formulation for calculation of probabilities on the main tournament format of the WBC, and compare the 2013 tournament format to the past two tournament formats and the case of employing the MMDE format, from the aspect of the probability of winning the tournament and the probability distribution of the number of games played by the same teams, by changing the relative strength of teams. We have illustrated the difference between these four formats.

As a calculation result, the difference of the probability of winning the tournament between the

four formats is not large. However, the expected number of games and the probability distribution of the number of games between two teams depend on the formats or the assignment of 16 teams to 4 pools. In general, we could conclude that $MMDE^2$ format seems to be more efficient than other formats, and $MMDE^2$ is also better than other formats in terms of the expected number of games played by same two teams. Thus, if the total number of games is acceptable for conducting the WBC, $MMDE^2$ could be recommended. Otherwise, if up to 4 games played by same teams is acceptable, the current RR&MDE format could be a reasonable selection. Anyway, as the design of tournament should be considered from a variety of factors, this study should be extended including another aspect in evaluation of the tournament format. As a further study, we could statistically evaluate the strength of teams using empirical data, and apply statistical models to study a design of the WBC type tournament structure.

Acknowledgments

This work was partly supported by Grant-in-aid for Scientific Research (C) of Japan [#21510159] and [#26350434].

References

Annis, D.H. & Wu, S.S. (2006). A comparison of potential playoff systems for NCAA I-A Football. *The American Statistician*, *60*(2), 151-157.

Appleton, D.R. (1995), May the best man win? *The Statistician*, *44*(4), 529-538.

Carlin, B.P. & Stern, H.S. (1999). Designing a college football playoff system. *Chance*, *12*(3), 21-26.

Bradley, R.A & Terry, M.E. (1952). Rank analysis of incomplete block designs: I. The method of paired comparisons. *Biometrika*, *39*(3/4), 324-345.

David, H.A. (1959). Tournament and paired comparisons. *Biometrika*, *46*(1/2), 139-149.

Edwards, C.T. (1996). Double-elimination tournaments: counting and calculating. *The American Statistician*, *50*(1), 27-33.

Glenn, W.A. (1960). A comparison of the effectiveness of tournaments. *Biometrika*, *47*(3/4), 253-262.

Hwang, F.K. (1982). New concepts in seeding knockout tournaments. *The American Mathematical Monthly*, *89*(4), 235-239.

Jackson, D.A. (1993). Independent trials are a model for disaster. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, *42*(1), 211-220.

Koning, R.H., Koolhaas, M. Renes, G. & Ridder, G. (2003). A simulation model for football championships. *European Journal of Operational Research*, *148*(2), 268–276.

Ladwig, J.A. & Schwertman, N.C. (1992). Using probability and statistics to analyze tournament competitions. *Chance*, *5*(3-4), 49-35.

Marchand, E. (2002). On the comparison between standard and random knockout tournaments. *The Statistician*, *51*(2), 169–178.

Martin, D. E. K. & Troendle, J.F. (1999). Paired comparison models applied to the design of the Major League baseball play-offs. *Journal of Applied Statistics*, *26*(1), 69-80.

McGarry, T. & Schutz, R.W. (1997). Efficacy of traditional sports tournament structures. *Journal of the Operational Research Society*, *48*(1), 65-74.

Monks, J & Husch, J. (2009). The impact of seeding, home continent, and hosting on FIFA World Cup results. *Journal of Sports Economics*, *10*(4), 391-408.

Nakamura, K. (2009). 'Game 6 Japan vs Korea'. *Sportiva*, 25 April, p.25. (in Japanese)

O'Donoghue, P. (2005). The role of simulation in sports tournament design for game sport. *International Journal of Computer Science in Sport*, *4*(1),14-27.

Okada, T. (2010) '28 teams participate in 2013 World Baseball Classic'. *Yomiuri Shimbun (Evening edition)*, 6 October, p.2. (in Japanese)

Scarf, P., Yusof, M.M. & Bilbao, M. (2009). A numerical study of designs for sporting contests. *European Journal of Operational Research*, *198*(1), 190–198.

Scarf, P. & Yusof, M.M. (2011). A numerical study of tournament structure and seeding policy for the soccer World Cup Finals. *Statistica Neerlandica*, *65*(1), 43–57.

Schwarz, A (2010) 'Japan defeats Korea, setting semifinal games', *New York Times*, 20 March < http://www.nytimes.com/2009/03/21/sports/baseball/21wbcjapan.
html>

Schwenk, A.J. (2000). What is the correct way to seed a knockout tournament? *The American Mathematical Monthly*, *107*(2), 140–150.

Stanton, I. & Williams, V.V. (2013). The structure, efficacy, and manipulation of double-elimination tournaments. *Journal of Quantitative Analysis in Sports*, *9*(4), 319–335.

Szymanski, S. (2003). The economic design of sporting contests. *Journal of Economic Literature*, *41*(4), 1137–1187.