Enhancing Viewer Engagement Using Biomechanical Analysis of Sport

Robert Dawes, Bruce Weir, Chris Pike, Paul Golds, Mark Mann, Martin Nicholson

BRITISH BROADCASTING CORPORATION
Enhancing Viewer Engagement Using Biomechanical Analysis of Sport

Robert Dawes, Bruce Weir, Chris Pike, Paul Golds, Mark Mann, Martin Nicholson

Abstract

The audience for television sport has high expectations in the analysis that forms part of that coverage. In a competitive broadcasting environment there is always a need to develop new features to engage the audience. This is particularly true of sports with relatively small audiences such as athletics that will only get large viewing figures during occasional big events such as the Olympics. In this paper we describe the results of our recent work in the field of biomechanics. This field of science is a key part of the training regime of almost all athletes and sportsmen and women. By making use of the tools and techniques of this field we have developed systems for both the next generation of television analysis systems and for distribution via the web to put the tools in the hands of the audience. These tools aim to offer a new level of insight and explanation to the audience – including those viewers who may rarely watch the sports in question – and so increase their engagement with the coverage. The web tool illustrates some of the possibilities that new forms of digital media content offer the viewer for direct interaction with video.

This document was originally published in the proceedings of the NEM Summit, Istanbul, 16-18 October 2012.

The slides used in this presentation are included in the appendix.

Additional key words: sport, analysis, biomechanics, Flash, augmented reality, image processing, computer vision
White Papers are distributed freely on request.

Authorisation of the Chief Scientist or General Manager is required for publication.
Enhancing Viewer Engagement Using Biomechanical Analysis of Sport

Robert Dawes\textsuperscript{1}, Bruce Weir\textsuperscript{2}, Chris Pike\textsuperscript{2}, Paul Golds\textsuperscript{2}, Mark Mann\textsuperscript{2}, Martin Nicholson\textsuperscript{1}

\textsuperscript{1}BBC Research & Development, London, UK; \textsuperscript{2}BBC Research & Development, Salford, UK

E-mail: <firstname.lastname>@bbc.co.uk

Abstract: The audience for television sport have high expectations in the analysis that forms part of that coverage. In a competitive broadcasting environment there is always a need to develop new features to engage the audience. This is particularly true of sports with relatively small audiences such as athletics that will only get large viewing figures during occasional big events such as the Olympics. In this paper we describe the results of our recent work in the field of biomechanics. This field of science is a key part of the training regime of almost all athletes and sportsmen and women. By making use of the tools and techniques of this field we have developed systems for both the next generation of television analysis systems and for distribution via the web to put the tools in the hands of the audience. These tools aim to offer a new level of insight and explanation to the audience – including those viewers who may rarely watch the sports in question – and so increase their engagement with the coverage. The web tool illustrates some of the possibilities that new forms of digital media content offer the viewer for direct interaction with video.

Keywords: sport, analysis, biomechanics, Flash, augmented reality, image processing, computer vision

1 INTRODUCTION

One of the challenges of ensuring that sports coverage is as engaging and involving for the audience as possible is the need to help them understand the sport they are watching. This is particularly true of sports where most of the audience will have limited experience of participating in the sport to any level of quality, and so may not be able to appreciate the level of skill or ability they are watching. One method of addressing this problem is to use a commentator or pundit - commonly a former professional from the sport in question - who uses their expertise and experience to explain the situation to the viewer. In a team sport this will usually consist of an explanation of the tactics in use, but for more individual sports such as athletics the explanation is more personal and may concentrate on the specific actions of an athlete, explaining the technique or effort required to perform them. It is this sort of explanation that our work aims to aid, by producing tools to help the pundit to explain the actions of athletes and to help the viewer relate to them.

Pundits are often provided with tools they can use to annotate the action, drawing on the video to illustrate the point they are making. These tools may produce simple flat drawings on the screen or use sophisticated systems to ensure that the drawings appear in the correct perspective, as if painted onto the field of play. However, we are looking beyond this passive annotation and developing tools that actively analyse the scene and extract or help to extract information about the performance.

We are also looking to give the viewer access to the tools and techniques currently only available to the studio pundit. By delivering video and data via the web we can create applications where the viewer can interact with the sport and give themselves a more involved experience of the event.

2 BACKGROUND

To know what is useful to extract from athletic sequences we have investigated the tools and techniques used by the athletes and their coaches and trainers. Increasingly athletes make use of sports scientists to help improve their performance and many of those scientists work in the field of biomechanics. Biomechanics is the application of mechanical principles to living organisms, examining the internal and external forces acting on them and the effects produced by these forces. It is a large and varied scientific field that combines the disciplines of biology and engineering mechanics and utilises the tools of physics, mathematics, and engineering to study everything from the molecular level up to the effect of gravity on entire skeletons.

Within classical mechanics there are two related fields, kinematics and kinetics:

Kinematics – The study of bodies in motion without regard for the causes of motion.

Kinetics – The study of the causes of motion.

Kinematics observes the quantities of motion such as position, velocity and acceleration both linear and angular, such as the angles of joints and the acceleration of a limb. Kinetics studies forces and moments of force and their characteristics such as work, energy, power and momentum.

Analyses of mechanical systems can be split into two categories, forward and inverse dynamics:

Forward Dynamics – prediction of the motion of bodies (kinematics) from forces and moments of force (kinetics).

Inverse Dynamics – prediction of forces and moments of force (kinetics) from the motion of bodies (kinematics) and their inertial properties.
In biomechanics, forward dynamics is often concerned with simulation of movements using a sequence of muscle actions as an input into a musculoskeletal model. The modelling is generally verified by comparison with a recorded real movement. Once a satisfactory correlation with the real world has been obtained, numerical methods can be utilised to search for variations in the sequence of actions that can obtain better sporting results. It can be used to discover what is within the range of human ability and verify our opinions on how movements are achieved.

As a very simple example biomechanists have built mathematical models to simulate the 100 metres sprint, profiling how an athlete accelerates, reaches a top speed and attempts to maintain it. These are verified by comparison to real races where split times have been recorded at regular intervals along the course. The parameters of the model can be altered so it fits the race profile of the athlete. The parameters can then be manipulated to see which aspects of his or her race the athlete should concentrate on in order to improve his or her performance.

Inverse dynamics will normally involve measurement of movement using, for example, a marker based motion capture system, which is combined with the inertial properties of the bodies, to calculate the internal forces and powers, and direct measurement of external forces, such as ground reaction force which is measured using a force plate. These measurements are solved using a regression function to obtain values for the forces and moments involved in the system. This allows high-level biomechanical analysis of the real-world movement. An example of these techniques might be to use the motion capture system to record a long jumper taking off from a force plate. All this data can be used to model the forces and exact positions of the athlete’s body parts allowing for a great deal of further analysis.

While much of this data will only be of interest to the athletes and coaches there is still a wealth of information that might be interesting to viewers at home to help them understand the events they are watching. For example, measuring the stride frequency and length of athletes in a 100 metre sprint can demonstrate what type of athlete they are: a tall long legged athlete who takes few strides and is slow to reach top speed or a smaller, quicker paced athlete who can get more strides in, but lacks the higher top speed, or perhaps somewhere in between. In the long and triple jumps the trajectory of the athlete can be modelled as a projectile, using the centre of mass as the location. The centre of mass begins at around waist height, then, as the athlete tucks in around it, ends up at ground level as he or she lands in the pit. Because the landing height is lower than that at take-off, the optimal take-off angle is less than 45 degrees, so they are able to retain more of their horizontal momentum. We can examine the take-off angle of an athlete to see how near they are to achieving their particular optimum angle.

The tools used by coaches and athletes to extract and record data of this sort typically make use of sensors or markers placed on the athlete or in the environment. However, we wish to analyse competitive events where such methods cannot be used easily because they interfere with the proceedings. This, combined with the logistical difficulties that come from working with sporting events that might be taking place all over the globe, means that we are effectively restricted to just working with the broadcast video of an event.

Some existing tools can work entirely with images. Examples include Dartfish’s products which are used in both the sports science and broadcasting spheres [1]. However, when used with just broadcast video they are generally restricted to producing solely visual effects, while we hope to make use of and gain knowledge about the scene.

### 3 PREVIOUS WORK

We have previously developed tools for augmenting real scenes with annotations that appear to be “painted” into the pitch or arena [2]. For football, rugby and other team sports with reasonably standard pitch markings our existing system uses these markings with known real world positions to determine the position and pose of the camera. It then tracks the movement of the lines to determine how the camera was moving. As the camera moves the graphics are moved such that they appear to be fixed to the same part of the real world. More recent work has developed the system further such that it can track arbitrary points rather than just lines [3]. This allows for similar graphical effects but it can be applied in a greater number of environments, including less regular environments such as athletics grounds.

[Figure 1: Measurement on a Rugby Field]

It is a natural extension of this work to try and gain more information about the performance of the athletes rather than just overlay annotations. This extra data can be presented to the viewer to give them another level of information about the event they are watching. Indeed there is a predecessor of sorts in the existing graphics systems. The camera calibration allows measurements between two points of known position, such as two points on a football pitch. This is often used to measure the distance involved in an incident or activity such as how far a player has run or how far from goal a free kick or conversion is being taken. An example of this facility is shown in Figure 1. Our tools aim to extend and develop this idea.
4 ANALYSIS TOOLS

We have developed a series of standalone tools to analyse different aspects of athletic or sporting performance. These tools operate on video or image sequences, processing them to extract additional data from the scene. This data can then be presented to the viewer to offer insights into the event they are watching.

4.1 Calibration

Some of these tools require information about the position and pose of the camera – much like the graphics drawing tools described above. In order to get hold of this calibration information we first process the video sequences to calculate where the camera is and how it moves. This process can be performed using live video or offline from file. Unlike the graphics drawing tools these analysis tools are mostly not intended to be used live so the process described here is the offline version. However the live video process is very similar.

The video sequence is treated as a sequence of separate images, each of which will be accompanied by a description of the camera position and pose. The system is first calibrated on a single frame. This process requires the coordinates of known points in the scene to be identified manually. With this information the position and initial pose of the camera is computed using an iterative optimisation process to minimize the squared reprojection error of the annotated locations into the image, following the approach in [2].

In Figure 2 known real world locations have been annotated and then highlighted in yellow. In this case these are the far and near edges of the pit and the 15m and 19m lines.

Once calibrated on a single image, a KLT-based tracker is used to track areas of rich texture from frame to frame. The camera position is assumed to be stationary and the movement of the texture patches is used to determine the changing pose of the camera throughout the sequence. This produces camera pose and position data for all of the images. The images and accompanying data can then be used by the other tools.

Figure 2: Known positions marked on an image

4.2 Stride Detection

We have developed a tool that can extract the positions in the real world of the feet of a running athlete. This can be used to automatically extract the stride frequency and length of a runner or the positions of the first two phases of a triple jump.

It uses a motion compensated temporal median filter to build up a background image for the scene with the athlete removed. We make use of the previously extracted frame-by-frame camera calibration data in order to perform the motion compensation.

A second background image is then generated using a filter with a smaller temporal window. The size is chosen such that the temporarily stationary feet of the athlete “burn” into the background. A difference is then taken between the two background images resulting in a mask of possible locations for the foot. The calibration is used to find where these possible locations occur in real world 3D coordinates (making the assumption that stationary feet lie on the ground) and the most likely option is chosen as the foot. If there was no stationary foot in that frame then there will be no suitable candidate in the mask. This process is conducted over the whole sequence and a series of footsteps are extracted. This data is then available for further analysis or for presenting to the viewer.

Figure 3: Original image (top left), background image (top right), background with visible foot (bottom left), difference between the two backgrounds (bottom right)

Figure 4: Step positions annotated onto the video sequence
4.3 Body Modelling

![Figure 5: Athlete with body parts marked and derived centre of mass](image)

In several events, particularly sprints and jumps, the athlete will only move in a single plane as they run down the track. We can combine this assumption with the camera calibration information to work out body positions in three dimensions. An operator can hand annotate a video sequence of an athlete, labelling the 2D positions of the body parts – a process known as “digitisation” by sports scientists. We can then calculate a line of sight from the camera towards this point and discover where it intersects the known plane of motion. For example, we may assume an athlete is running down the centre of his or her lane and that the head, neck and base of the spine can all be found along this plane while the limbs are in planes offset to the left and right accordingly. This plane gives one dimension, while the point of intersection provides us with the other two. Once we have positions for all the body parts we can display them within a virtual environment or we can place them into a mathematical model of the body to try and extract extra information such as the centres of mass of the various body parts and the whole body. The pundit can then make use of this information in his or her analysis or the data could be presented directly to the viewer.

![Figure 6: Visualisation of body positions in a 3D environment](image)

4.4 Cadence Detection

This tool detects the position of a bicycle in a scene by using a Hough transform to locate the wheels. It then looks for the position of the cyclist’s feet – segmenting them from the background using their colour - and tracks them as he or she pedals. From this information a value for cadence (i.e. the pedalling speed) can be extracted. This can inform the pundit or viewer about when a cyclist is accelerating, or in combination with the gear ratio how much power is being produced.

![Figure 7: Annotated bicycle wheel and foot with extracted movement.](image)

4.5 Diving

In the sport of diving points are awarded by a panel of judges. To the unskilled eye of the viewer it may often be unclear why one dive scored better than another. As the diver enters the water he or she aims to be as upright as possible and to minimise the splash which would result in a non straight and vertical entry:

“The entry into the water shall in all cases be vertical, not twisted, with the body straight, the feet together, and the toes pointed.”[4]

The system measures the size of the splash and the angle of entry. It segments the largely white splash from the largely blue background and then measures the size of the resulting mask in order to get a figure for the splash. The angle is detected by segmenting the diver from the blue background and then fitting a line down the length of the extracted object.

These measurements give viewers some insight into why a dive might receive the score it does and offers them the opportunity to compare one dive with another. It is useful to have the angle value for a few frames as the diver enters the water. This can help communicate the speed and intricacy of motion involved in a dive.

![Figure 8: Diver entering the water with visualisation of the splash size (height of bar on the right) and angle of entry of the last 4 frames.](image)
5 WEB BASED AUGMENTED REALITY

5.1 Overview

We have developed a Flash application to allow the viewer to interact with footage of sporting events and help them to get more involved in the action. By rendering a 3D scene on top of a background video sequence, the real footage can be ‘augmented’ with virtual objects in a similar manner to the broadcast graphics tools mentioned above. However, because the rendering is performed client-side by the Flash plug-in, as shown in Figure 9, we can offer an engaging interactive experience to the viewer where they can affect the augmented graphics themselves.

Figure 9: Virtual 3D objects added to a real scene at the client side. During early tests a rolling wireframe sphere was used to represent the virtual athlete.

5.2 Augmenting Graphics

The web application for client-side interaction with the biomechanics data was built using Flash and the Away3D ActionScript library[5]. Away3D is an open-source 3D graphics library and we use it to render a 3D scene on top of a background video sequence. Since we are trying to insert virtual objects into the scene so that they look like that are present in the real environment, the virtual elements must not drift relative to the real objects visible in the background video. In order to make this work, the frame-by-frame, camera-pose calibration data described above must be made available to the 3D renderer as each video frame is updated.

Some sequences are less suitable for the tracking process used to generate the calibration data than others, for example where there are very few distinct background features visible. The resulting camera data can sometimes have faults where the virtual camera wobbles or moves sharply. This is significant problem with a graphics drawing system such as this application. However, the offline nature of the process means there is significant time to make adjustments to cover the faults - a process that is not possible when tracking is used live. One method of fixing faults is to interpolate over bad frames between two known good frames.

The video sequences were converted into FLV format[6] with the video encoding undertaken by the x264 library[8]. The camera pose calibration data was embedded into the resulting file as frame-by-frame ‘ScriptData’ tags [7].

time-stamped for ‘presentation’ with the same timestamp as the video frame to which they correspond. This camera pose data is made available to the Flash application during video playback via a handler method which is triggered whenever a ScriptData tag is encountered. The handler extracts the embedded camera pose data and uses it to control the pose of an Away3D Camera3D object modified to generate the correct transform for converting the 3D model coordinates into the screen coordinates when provided with a camera pose, field of view and aspect ratio. This method also initiates the render pass which ensures that the video updates and 3D overlay updates happen simultaneously. The 3D model elements are positioned in the scene graph such that they are rendered after the video image, this means that they will always appear in the foreground. In this version of the application, there is no alpha masking of the video, so virtual objects cannot appear behind objects in the video.

5.3 Final Application

The application we have developed using this approach allows users to compare their own sporting performance against that of professional athletes in a novel way. The sports included in the application are the 100m sprint, long jump, high jump and triple jump. The user enters details of their own sporting performance, such as their best long jump, or fastest 100m sprint then the video footage of the sporting event is augmented with an avatar of the user competing in the same event. Examples can be seen in Figure 10 and Figure 11. Other athletics events with a simple performance metric (time, distance, etc.) would be straightforward to add, but events such as synchronised swimming or beach volleyball would be much harder to simulate.

The user can first personalise the experience by entering their height and weight. If they wish they can also use a webcam to take an image of their face which is then texture mapped onto the avatar, or they can select the face of a sporting celebrity to represent them. The body of the avatar is textured with the colours from the 2012 British Olympic team.

Figure 10: The user clears the bar in the high jump, matching the women’s world record.

A degree of bespoke animation is required to integrate the avatar with the event. For example, the location of the bar...
in the high jump event needs to be known in the coordinate system of the foreground 3D model so that the avatar’s run up and jump can be correctly positioned, or the orientation of the camera pose with respect to the 100m running track needs to be understood to make sure that the calculated motion vector of the avatar matches the direction of the other competitors in the race. The video footage is also augmented with markers showing world record distances, or the dimensions of well-known real-world objects such as double-decker buses. As well as seeing themselves compete, these real-world objects help the user relate to what they are seeing.

In order to ensure realistic positioning over the course of the whole race the movement of the avatar in the 100m sprint is controlled by a model based on Tibshirani’s extension of the Hill-Keller model [9]. The equations of motion are

\[
D(t) = kt - \frac{1}{2} c \pi^2 + e^{\tau/c}(e^{-\tau/c} - 1)
\]

where \( k = f\tau + \tau^2 c \)

In this model \( f \) represents the acceleration force of the athlete. We calculate this by taking the user’s chosen finish time for the 100m and make assumptions for \( c \) and \( \tau \), respectively representing the athlete’s muscular endurance and a broader measure of flexibility, leg turnover rate, anaerobic response etc [10]. Our assumptions use a combination of typical figures for professional athletes [9] with some variation based on the personal attributes of the user entered earlier on. For example, a taller, heavier athlete gets a higher \( \tau \) value, which gives him or her lower acceleration. \( c \) varies depending on the chosen finish time for the sprint, the longer the sprint took the lower the value, implying less energy is being drained due to a lower running speed.

**Figure 11: A personalised avatar taking part in a 100m race**

Once \( f \) has been calculated the equations of motion can be used to find the location of the user’s avatar throughout the race. We can also provide details of his or her velocity as the race progresses. This is provided in miles per hour, a measure of speed that viewers will be most familiar with, helping them to relate to and understand what they are watching. While this model is of course not a completely accurate representation of how the user might perform, it is at least indicative of relative performance and helps to demonstrate to the user the level of ability inherent in professional competition.

### 6 CONCLUSIONS

With relatively simple tools we are able to present to the viewer an extra level of detail in the events they are watching. For those with a particular interest in a sport this gives them the extra detail that the current coverage may lack. The extra analysis may also give a more mainstream viewer an insight into a sport that has never occurred to them before. This may well encourage a greater interest in a particular sport and engagement in its coverage. This is particularly true of the web application that offers a fun and accessible way to learn about a sport and to look at it in a new way.

In addition the augmented reality application described here only begins to scratch the surface of the possibilities offered by delivering analysis tools via the web. Tools that were previously only the domain of the television pundit may soon be put in the hands of the viewers. This development can be seen as part of the wider trend of giving the power to the audience. They expect to be able to choose, interact and play around with the media they consume and this application offers them just that.

### References


[7] Ibid. p74


Appendix
The following slides are taken from the presentation of this paper at the NEM Summit, Istanbul, 16-18 October 2012. The videos have been removed and the image quality reduced to allow for online distribution. The originals are available on request.
Enhancing Viewer Engagement Using Biomechanical Analysis of Sport

Robert Dawes, BBC Research & Development
Sports Broadcasting

• Sports Broadcasting
• Biomechanics
• Our Background
• Biomechanics Prototypes
  – Stride Detection
  – Body Modelling
  – Cycling Cadence
  – Diving analysis
  – Web application
• Conclusions
Sports Broadcasting

Pundits are used to explain the intricacies of a sport to the viewers.

Some sports make good use of technology in their coverage.
Biomechanics

- Offers new opportunities for insight and analysis
- Some typical data:

![100m Final Speeds chart](chart.png)
Biomechanics

Number of strides

Step Frequency

Step Length

Reaction Time

- Bolt
- Gay
- Powell
- Bailey
- Thompson

© BBC MMXII
Our Background

- Sports analysis tools for sports including football, rugby and athletics
Our Background

- Line and feature tracking are used to provide camera calibration
Our Background

- Biomechanics offers a potential next stage of development

- The camera calibration is a base on which new tools can be built.
  - Some simple measurement tools are already in place:
Prototypes

Stride Detection

Diving Analysis

Web Application

Cadence Measurement

Body Modelling
Stride Detection

- A clean background is generated using a motion compensated temporal media filter
Stride Detection

- A smaller filter window causes the temporarily stationary foot to “burn” into the background
Stride Detection

- The foot should be the main difference between the two backgrounds
Stride Detection
Body Modelling

Camera calibration is combined with assumptions about movement to extract 3D positions of body parts from 2D annotation.

Positions are placed in a biomechanical model of the human body and extra data is extracted.
Body Modelling

- The data can then be used and visualised in various ways.
Cycling Cadence Measurement

- Simple tool to find the speed the cyclist is pedalling
Diving Analysis

- A move from qualitative to quantitative analysis.
- Segment images and measure
Diving Analysis

- Used by the BBC and NBC during the London Olympics
Web Application

- Using Flash to allow client side rendering of augmented reality graphics in the browser
Web Application

Calculate camera calibration data frame by frame

Video sequence

Video with embedded camera data

Delivered to the browser

Extract camera data frame by frame

Render graphics frame by frame

Display
Web Application

- We created a video player where the viewer can put themselves into the action.
Web Application

- [http://www.bbc.co.uk/sport/0/olympics/19024989](http://www.bbc.co.uk/sport/0/olympics/19024989)
- 30,000 users since the start of the Olympics
Conclusions

• Many more possible tools making use of the field of biomechanics.

• They offer the viewer a deeper understanding of an event.

• Help to attract new audiences.

• The metadata produced and technologies such as the AR application open up potential new forms of content for the future.
Thanks

• Co-authors and colleagues involved in these projects: Bruce Weir, Chris Pike, Paul Golds, Mark Mann, Martin Nicholson

• Commercial licensee for the diving work: Piero Team at Red Bee Media

• Assistance with the web application: BBC Future Media
Thanks

robert.dawes@bbc.co.uk

Questions?