Situation Modeling and Visual Analytics for Decision Support in Sports

Anders Dahlbom[1,2] and Maria Riveiro[1]

[1]*Informatics Research Centre, University of Skövde, P.O. Box 408, SE-541 28 Skövde, Sweden*

[2]*Swedish Winter Sports Research Centre, Mid Sweden University, Östersund, Sweden*

{*anders.dahlbom, maria.riveiro*}*@his.se*

Keywords: Sports, Decision Support, Situation Modeling, Visual Analytics, Information Fusion.

Abstract: High performance is the goal in most sporting activities, for elite athletes as well as for recreational practitioners, and the process of measuring, evaluating and improving performance is one fundamental aspect to why people engage in sports. This is a complex process which possibly involves analyzing large amounts of heterogeneous data in order to apply actions that change important properties for improved outcome. The number of computer based decision support systems in the context of data analysis for performance improvement is scarce. In this position paper we briefly review the literature, and we propose the use of information fusion, situation modeling and visual analytics as suitable tools for supporting decision makers, ranging from recreational to elite, in the process of performance evaluation.

1 INTRODUCTION

Sports have in various forms engaged people throughout history. In its essence, sports is about objectively determining a winner in some sporting event through quantification using measures such as time, length, height and score. The goal and motivation for an athlete is thus to achieve high performance with respect to these measures, in order to perform better than the opponents. The subjective experience of carrying out sports is however also of high importance in order to keep athletes, and recreational practitioners, motivated and engaged.

The development of sports technology has historically been driven by research in e.g. biomechanics and physiology aimed at improving the performance of elite athletes. This involves efforts with respect to a multitude of factors, e.g. physiological performance, technical skill, cardiovascular capacity, muscle strength, tactical decision making and mental focus. To this end, many tools and systems have been developed, e.g. video analysis and motion capture technologies, heart rate monitors, and automatic systems for measuring and monitoring oxygen uptake. However, only a few of the developed systems are mobile and usable outdoors. Moreover, the interest in carrying out sports is also high among novice from recreational and health perspectives, and many of the traditional systems are not generally accessible from these perspectives.

Technology and tools related to sports, recreational activities and healthcare represent rapidly growing areas. There are many applications for smartphones, e.g. Runkeeper, Nike+, as well as equipment such as sports watches with GPS and heart rate monitors. These tools are used by numerous practitioners around the world and are often connected to social media, enabling for data to be shared. Although there is great potential for data analysis and decision support in this domain, there seems to be a lack of tools which includes advanced data analysis capabilities, for supporting performance evaluation and improvement, for both elite and recreational activities. There is a need for rich models that can be personalized and used for aiding performance improvement.

We believe that techniques used in information fusion, situation modeling and visual analytics can be used as a suitable foundation for constructing decision support tools for (1) capturing the complete performance evaluation process, (2) understanding, constructing and maintaining the underlying models and (3) for interpreting and understanding data and information. In itself, interaction with models and e.g. virtual coaching may also lead to enhanced experience.

The rest of this position paper is organized as follows. Section 2 briefly reviews literature regarding the analysis of sports data. Section 3 presents and motivates our future research path towards decision support within sports. Finally, section 4 discusses the this path and outlines our future work.

2 SPORTS AND INFORMATION TECHNOLOGY

Although computer science and sports have more than 50 years together, initially focusing on sports information (documentation) and later on sports informatics (processing), it was not until a few decades ago that computer science emerged as an important interdisciplinary partner to sports science (Baca, 2006). Perl (2006) observes a change about 20 years ago, where the spectrum of computer science in sports changed. Perl attributes this change to advances in computer science coupled to hardware, software, communication and multimedia. Baca also roughly identifies these four aspects as the main areas of research related to computer science in sports: (1) data acquisition, processing and analysis, (2) modeling and simulation, (3) data bases and expert systems, and (4) multimedia, presentation and virtual reality. From a data analysis perspective, we distinguish four different themes in the following: (1) classification and prediction, (2), event and activity detection in sports video data (3) collection and fusion of sensor data, and (4) performance improvement guidance.

2.1 Classification and prediction

Offline classification and prediction can typically be approached using two perspectives: data-driven and knowledge-driven. The former approach is closely tied to data mining and machine learning, where techniques such as artificial neural networks and decision trees are constructed from historical data and then used as models for classifying new data samples. The latter approach is coupled with knowledge engineering and expert systems, where models are created by eliciting knowledge from experts. A common theme in offline predictive analysis is that the focus is on offline use, i.e. a set of recorded data is analyzed in isolation from the actual athletes in order to predict or classify some aspect of interest.

Within sports, e.g. artificial neural networks have been applied to the talent identification problem in team sports, where the aim is to identify promising players/athletes that can be drafted for a team (McCullagh, 2010). In the study, a number of parameters, e.g. physiological, medical, psychological and anthropometric, are used to classify and predict promising talents within the Australian football league. Expert systems have also been applied on the problem of identifying potential sport talents. For example, Papić et al. (2011) have used a fuzzy logic approach to identify potential talent among children with respect to e.g. basketball, gymnastics, swimming, and tennis, based on knowledge from kinesiology experts such as functional, motorical and morphological features. Moreover, expert systems have also been applied for assessing the potential and quality of basketball players (Dežman et al., 2001).

Artificial neural networks have also been applied for predicting the performance/outcome of individual athletes in e.g. swimming (Edelmann-Nusser et al., 2002). In the study, neural networks were constructed using a number of parameters related to training sessions and the predictive variable was the performance of the athlete at contest. Predictive modeling has also been applied to predict the outcome of matches in team sports. In e.g. O'Donoghue (2012) multiple linear regression was used for predicting the outcome of rugby matches based on a set of variables such as difference in world ranking between teams, difference in distance traveled to the match, and difference in recovery days from the previous match. Moreover, machine learning techniques have also been used for e.g. classifying skill of individual players in table tennis using C4.5, naïve Bayes and random forests (Maeda et al., 2012) and for classifying subjects in running/walking using artificial neural networks and support vector machines (Fischer et al., 2011).

2.2 Event and Activity Detection

Event and activity detection in sports video data is another area which has received much focus, where a common theme is discovery and recognition of patterns. Much work has been carried out coupled to the analysis and identification of distinct events in sports video data, e.g. the use of dynamic Bayesian networks for play-break detection in soccer games (Wang et al., 2005) and the use of tracking techniques and semantic extraction for editorial content creation (Xu et al., 2009). These have the common theme of supporting media creation. Similar techniques have however also been used for supporting coaches during actual events, e.g. the use of probabilistic techniques for capturing the key notions of games (Beetz et al., 2009), and the use of simple statistical measures on time-series data for analyzing player performance in team sports to detect key turning points in player performance (Lühr and Lazarescu, 2007).

2.3 Sensor Fusion

Sensor fusion and collection is another interesting application of computer science in sports. This is closely connected to information fusion and the combination of data and information from multiple sources for improved estimation and prediction.

Flanagan (2009) states that a significant challenge for e.g. coaches working with alpine skiing is to accurately detect speed and trajectory of athletes throughout courses during training and competition. Recently, fusion and technologies such as GPS, inertial sensors and high-speed cameras have been applied to this task, however, mostly in cross-country skiing (Flanagan, 2009). Brodie et al. (2008) have investigated a fusion motion capture system for collection of data that can be used for optimizing ski racing. The system fuses data from several different types of sensors (GPS, accelerometers, gyroscopes and magnetometers) to establish a precise global track of the motion of a downhill skier. The information that is fused and collected can be used for e.g. post action analysis to identify and analyze paths. Brodie argues that the usage of such fusion systems is much more suitable compared to the use of classical optical systems which are only able to focus on limited portions of individual runs, compared to the analysis of complete runs. Similarly, Kruger and Edelmann-Nusser (2010) compares a full body inertial system with optical video based systems and concludes that full-body inertial systems are beneficial since the classical systems typically are restricted to analyzing actions within specific volumes. Besides winter sports, wearable systems have also been investigated in connection to running and stride parameter running (Cheng et al., 2010). Similarly, advantages include easy deployment and the possibility of collecting data from multiple runs and over long distances, compared to classical optical based systems which are more difficult to deploy and which have limited viewing volumes.

2.4 Performance Improvement

Performance improvement guidance is perhaps the least investigated area of computer science within sports. Already in 2006, Bartlett argued for the potential of using expert systems and machine learning techniques in sports biomechanics analysis for improvement of performance. Bartlett (2006) however concluded that the usage so far is low. Owusu (2007) presents a general model, recognize critique recommend, which can be used for performance improvement in sports. Owusu discusses the use of neural networks and expert systems for recognition and critique, but concludes that very little has been investigated on these topics and especially on the final step of providing recommendations for improved performance. With the recent wide spread use of mobile platforms, excellent opportunities for feedback arise, which can be closely connected to individual training sessions. Kranz et al. (2013) investigates the use of smart phones for automatic recognition, assessment and feedback during indoor exercises. Novatchkov et al. (2011) presents a framework for mobile coaching based on a client-server architecture, where coaches in near real-time can analyze data collected from athletes offsite and still be able to provide feedback. Similarly, Tampier et al. (2012) also investigates a framework based on the client-server architecture for providing marathon runners with feedback within the context of mobile coaching. Tampier et al. however also investigates the use of a simple model to simulate load and performance, which can be used for automatic feedback to e.g. avoid overload and underperformance. Kirby (2009) discusses the use of real-time measurements, comparison with models and feedback in relation to alpine skiing. The potential benefits of real-time feedback is highlighted by Kirby, who argues that an athlete that receives online feedback can make immediate corrections, compared to post-action review with corrections applied at the next training session.

3 MODELING AND DECISION SUPPORT

As pointed out earlier, measuring and improving performance, and enhancing experience, are two fundamental concept related to why people engage in sports. Measuring and improving performance, also referred to as performance evaluation, is an essential part of sports and it is concerned with domain modeling and evaluation with respect to such models (Owusu, 2007). Owusu suggests the process of recognize, critique and recommend, separating the linear process of performance evaluation from the specific domain dependent model. The recognize step involves identifying the particulars of the actions that are carried out, the critique, or diagnosis, step aims at identifying faults with respect to expected or optimal performance, and the recommend step involves suggesting actions for improving performance. With respect to automation and decision support, a fair amount of work has been carried out related to the critique step, while work on recognition so far mainly has focused on low-level recognition (data level), leaving high-level recognition largely unexplored. Moreover, work on the recommend step is also unexplored within the sports domain. Although work has been carried out on each of these steps, individually, there is a general lack of work on automation and decision support that captures the complete process of performance evaluation (Owusu, 2007).

3.1 Information Fusion

Techniques for data and information fusion have been identied as key enablers for providing decision support when large amounts of heterogeneous data need to be analyzed (Liggins et al., 2009). In information fusion (IF), functions for building and using models and analyzing data are typically separated into six levels (Liggins et al., 2009): (level 0) feature assessment (estimating and predicting feature states), (level 1) object assessment (fusing sensor data to obtain reliable and accurate estimations of an entities attributes), (level 2) situation assessment (estimating entity-to-entity and entity-to-environment relations), (level 3) impact assessment (estimating utility/cost of feature, object, or situation states, including future effects of planned or predicted actions), (level 4) process refinement (meta-processes concerned with monitoring and optimizing the overall fusion process) and (level 5) cognitive renement (cognitive aids and human computer interaction).

This conceptual division of a complex fusion process in low and high level functions matches the tasks that need to be resolved to design and build computer based systems that provide performance evaluation support. For example, in a running context techniques used in level 0 and 1 are needed to establish biometric attributes and tracking objects, while the relation of these objects to the environment (e.g. terrain characteristics and weather) or to other objects is a situation assessment problem. The inferences done for establishing the consequences or effects of certain actions, e.g. increasing the rhythm, are based on the application of impact assessment techniques, while process refinement involves multi-objective optimization methods that monitor the overall performance.

IF provides a unifying framework that can be used for approaching the issues involved in constructing models of performance for athletes and recreational practitioners. This framework, as well as IF methods and techniques have been successfully applied in many different areas, such as military command and control, computer security, transportation, etc. However, they have not still been used extensively in the area of sports and performance evaluation.

3.2 Situation Modeling

The domain specific models employed in performance evaluation should capture relations between situations and performance, or more specifically, between situation types and performance, where situations are instances of situation types. Depending on the level of abstraction, e.g. from technique to tactic, situations and situation types may be represented using feature state, object state or distinct actions/tasks that were carried out. To provide rich support, the models should preferably surpass the use of simple and individual measures (e.g. positions, directions, speeds at individual points in time) to instead model more complex behaviors and causal relations that can be used for increasing performance. They may also require the combination of data not only from individual athletes, but also from the environment and from other athletes. This would address the need for new methods and tools that reflect the complexity and holistic nature of performance in sport identified by Balagué and Torrents (2005).

Models of situations and situation types can be built using both knowledge- and data-driven approaches. Optimally, these approaches are combined in such a way that domain knowledge is used to create a priori models which can be refined in relation to analyzes of available data. The reverse is of course also possible, to extract models from data, which can be used for discovering novel information that can be exploited or refined by experts. A hybrid approach would combine the strengths of information fusion, data mining and knowledge engineering, and promote the use of models and systems that can be adapted.

In order to cover all three steps of the performance evaluation process, from mapping the present situation to its situation type, to suggesting corrections for reaching situations that are instances of optimal situation types, it is important that distinct situation types/performance pairs are related to each other. It is not enough to only be able to classify situations (e.g. novices and experts), but also to understand which parameters to influence in order to change behavior (improving performance). It is thus important that the models allow for prediction of future outcomes based on the current state and future actions. Ideally, the models should, when possible, be used in real-time by athletes, who would experience adequate feedback about how to change the current behavior to achieve optimal performance. This puts the research at the state of the art:

> A well-known problem and most popular in training preparation is finding an accurate stimulus in an appropriate time for each athlete. In efficiency sports solving this problem is the key to success. [...] It is almost impossible to find studies that combine computer science and pin point this topic. [...] Unfortunately none of [the suggested approaches] works, and even try to adopt machine learning algorithms to support trainers decisions. (Mezyk and Unold, 2010, p. 1499)

3.3 Visual Analytics

In order to provide interactive, personalized and immersive interfaces to deal with both data and models, visual analytics can be used as a design framework. Visual analytics (VA) is defined as analytical reasoning supported by highly interactive visual interfaces (Thomas and Cook, 2005) and it strives to facilitate the analytical reasoning process by creating software that maximizes the human capacity to perceive, understand, and learn from large, complex and dynamic data and situations.

VA advocates not only for an interactive presentation of data and outcomes, but also for the design of transparent user modeling and adaptation modules guided by the user, in our case, coaches and experts. Thus, future virtual coaching software can reach the flexibility that Owusu (2007) claims necessary in this domain. VA provides as well techniques that support interactive and user-friendly environments that present relevant information such as summaries, actionable steps, performance enhancement processes, statistics about outcomes, etc.; allowing engaging experiences that leads to more active participation by the users.

3.4 A combined view

Figure 1 presents an illustration of the use of visual analytics and situation modeling for providing decision support in the performance evaluation process. The various users depicted in the model (coaches, athletes, recreational practitioner) assess and modify the data, models and goals through an interactive visual interface. This also includes and combines the knowledge discovery and knowledge engineering processes. On the left, the characterization of the performance evaluation process outlined by Owusu (2007) uses the data, the models and the goals as input for providing recommendations. We have omitted the distinct information fusion levels in the model to not obscure the view. If included they would mainly be mapped onto the left and center of the model, i.e. directly onto the performance evaluation process (including the data, domain specific models, and goals).

The model (figure 1) can be used in a forensic fashion, to identify relations between different factors and the outcome, for increased performance of individual athletes. It can also be used in real-time for comparing the present behavior of individual athletes with patterns of other athletes or with some optimum behavior. More interestingly, the model could be used for analyzing future impacts of the present behavior (e.g., for predicting the outcome of different actions).

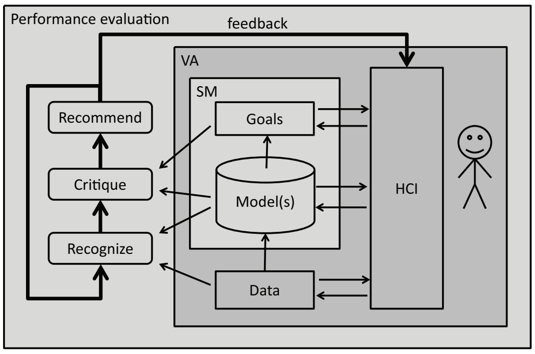

Figure 1: A combined view of situation modeling and visual analytics for performance evaluation.

4 DISCUSSION

Modern technical solutions make it possible to equip people and objects in the environment with sensors that can transmit data to the cloud. This allows for large amounts of data to be collected, which could be used in forming a large number of e-services. Moreover, data can be also be shared between devices and between services, as well as be connected to social media enabling for individuals to share experiences. By exploiting modern infrastructure, more extensive analyzes and decision support (including immersive visualizations and feedback mechanisms) can be provided directly through mobiles devices, e.g. smart phones. The opportunities for providing both elite athletes as well as recreational practitioners with advanced support, e.g. feedback on physiology, technique, tactic and strategy are enourmous and also provide many research challenges. This model could also be generalized and applied in other areas where computer based decision support is needed for user modeling and performance improvement, such as e.g. in learning environments.

In this position paper we have briefly reviewed the present situation with respect to data analysis and performance evaluation in sports. We outline how future computer based decision support systems can be built using information fusion, situation modeling and visual analytics. Our present work focuses on investigating these concepts in the area of golf. Our more long-term goals is to carry out investigations connected to other sports, as well as in other domains.

REFERENCES

Baca, A. (2006). Computer science in sport: An overview of history, present fields and future applications (part i). *Int. J. of Computer Science in Sport*, 4(1):25–34.

Balagué, N. and Torrents, C. (2005). Thinking before computing: changing approaches in sports performance. *Int. J. of Computer Science in Sport*, 4(2):5–13.

Bartlett, R. (2006). Artificial intelligence in sports biomechanics: New dawn of false hope? *J. of Sports Science and Medicine*, 5:474–479.

Beetz, M., von Hoyningen-Huene, N., Kirchlechner, B., Gedikli, S., Siles, F., Durus, M., and Lames, M. (2009). Aspogamo: Automated sports games analysis models. *Int. J. of Computer Science in Sport*, 8(1):4–21.

Brodie, M., Walmsley, A., and Page, W. (2008). Fusion motion capture: a prototype system using inertial measurement units and gps for the biomechanical analysis of ski racing. *Sports Technology*, 1(1):17–28.

Cheng, L., Kuntze, G., Tan, H., Nguyen, D., Roskilly, K., Lowe, J., Bezodis, I., N., Austin, T., Hailes, S., Kerwin, D., G., Wilson, A., and Kalra, D. (2010). Practical sensing for sprint parameter monitoring. In *Proc. of the 7th Annual IEEE Comm. Society Conf. on Sensor Mesh and Ad Hoc Communications and Networks*.

Dežman, B., Trninić, S., and Dizdar, D. (2001). Models of expert system and decision-making systems for efficient assessment of potential and actual quality of basketball players. *Kinesiology*, 33(2):207–215.

Edelmann-Nusser, J., Hohmann, A., and Henneberg, B. (2002). Modeling and prediction of competitive performance in swimming upon neural networks. *European J. of Sport Science*, 2(2):1–10.

Fischer, A., Do, M., Stein, T., Asfour, T., Dillman, R., and Schwameder, H. (2011). Recognition of individual kinematic patterns during walking and running - a comparison of artificial neural networks and support vector machines. *Int. J. of Computer Science in Sport*, 10(1):63–67.

Flanagan, T. (2009). Alpine skiing technology: faster, higher, stronger. *Sports Technology*, 2(1-2):5–6.

Kirby, R. (2009). Development of a real-time performance measurement and feedback system for alpine skiers. *Sports Technology*, 2(1-2):43–52.

Kranz, M., Möller, A., Hammerla, N., Diewald, S., Plötz, T., Olivier, P., and Roalter, L. (2013). The mobile fitness coach: Towards individualized skill assessment using personalized mobile devices. *Pervasive and Mobile Computing*, 9(2):203–215.

Kruger, A. and Edelmann-Nusser, J. (2010). Application of a full body inertial measurement system in alpine skiing: A comparison with an optical video based system. *J. of Applied Biomechanics*, 26:516–521.

Liggins, M., Hall, D., and Llinas, J., editors (2009). *Handbook of Multisensor Data Fusion: Theory and Practice*. CRC Press.

Lühr, S. and Lazarescu, M. (2007). A visual data analysis tool for sport player performance benchmarking, comparison and change detection. In *Proc. of the 19th IEEE Int. Conf. on Tools with Artificial Intelligence*.

Maeda, T., Fujii, M., and Hayashi, I. (2012). Time series data analysis for sport skill. In *Proc. of the 12th Int. Conf. on Intelligent Systems Design and Applications*, pages 392–397.

McCullagh, J. (2010). Data mining in sport: A neural network approach. *Int. J. of Sports Science and Engineering*, 4(3):131–138.

Mezyk, E. and Unold, O. (2010). Machine learning approach to model sport training. *Computers in Human Behavior*, 27:1499–1506.

Novatchkov, H., Bichler, S., Tampier, M., and Kornfeind, P. (2011). Real-time training and coaching methods based on ubiquitous technologies - an illustration of a mobile coaching framework. *Int. J. of Computer Science in Sport*, 10(1):36–50.

O'Donoghue, P. (2012). The assumptions strike back! a comparison of prediction models for the 2011 rugby world cup. *Int. J. of Computer Science in Sport*, 11(2):29–40.

Owusu, G. (2007). Ai and computer-based methods in performance evaluation of sporting feats: an overview. *Artificial Intelligence Review*, 27(1):57–70.

Papić, V., Rogulj, N., and Pleština, V. (2011). Expert systems for identification of sport talents: Idea, implementation and results. In Vizureanu, editor, *Expert Systems for Human, Materials and Automation*, pages 3–16. InTech.

Perl, J. (2006). A computer science in sport: An overview of present fields and future applications (part ii). *Int. J. of Computer Science in Sport*, 4(1):36–45.

Tampier, M., Endler, S., Novatchkov, H., Baca, A., and Perl, J. (2012). Development of an intelligent real-time feedback system. *Int. J. of Computer Science in Sport*, 11(3):58–64.

Thomas, J. J. and Cook, K. A. (2005). *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. National Visualization and Analytics Ctr.

Wang, F., Ma, Y.-F., Zhang, H.-J., and Li, J.-T. (2005). A generic framework for semantic sports video analysis using dynamic bayesian networks. In *Proc. of the 11th Int. Multimedia Modelling Conf.*, pages 115–122.

Xu, C., Cheng, J., Zhang, Y., Zhang, Y., and Lu, H. (2009). Sports video analysis: Semantics extraction, editorial content creation and adaptation. *J. of Multimedia*, 4(2):69–79.