International Journal of Computer Science in Sport

Volume 8/Edition 2

ISSN 1684-4769

TABLE OF CONTENTS

Editorial

Arnold Baca

Department of Biomechanics, Kinesiology and Applied Computer Science,

ZSU, University of Vienna

Dear readers:

Welcome to the winter 2009 issue of the **International Journal of Computer Science in Sport (IJCSS)**.

The current issue is subdivided into 2 parts, including seven original papers.

The first part ("Special Edition"), starting with an editorial by **Christoph Igel** and **Larry Katz** on the selected topic "Multimedia, the World Wide Web and eLearning in Sport", contains five papers.

Two additional articles are appended thereafter in the second part.

The paper by **Daniel Link** and **Martin Lames** overviews the historical background of the interdisciplinary area of sport informatics and gives a prognosis on its future development up to the year 2020.

Richard G. Lomax discusses a method for clustering NHL players in different forward and defensive groups. The classification of the ice hockey players is based on the use of a specific computer simulator.

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

Best wishes for 2010!

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

PART 1

SPECIAL EDITION

Multimedia, the World Wide Web
and eLearning in Sport

Editorial

Christoph Igel & Larry Katz

[1] *Competence Centers VISU, Saarland University, Germany*

[2] *Faculty of Kinesiology, The University of Calgary, Canada*

Dear readers:

In June of 2009, we issued a call for papers on the topic of “Multimedia, the World Wide Web and elearning in Sport for a Special Issue of the **International Journal of Computer Science in Sport (IJCSS).** We were pleased to accept five papers for this issue that explored innovative approaches to improve learning and performance.

Kazumoto Tanaka has developed a unique approach to training athletes using a mixed reality technology to provide immediate feedback to Karate students. Tanaka explorers the possibility of creating an environment that allows the trainee to visualize his or her performance in 3D thereby focusing on the visual feedback rather than the virtual environment. This is a novel approach which needs to be tested with larger sample sizes, but is sufficiently advanced to warrant inclusion in this special issue.

Roland Leser, **Johannes Uhlig** and **Manfred Uhlig** have taken a multimedia approach to teaching and learning soccer tactics. Using multimedia to teach is not new, but the concept of the coaches interactive participation in the learning activity followed by detailed analysis of the performance is distinctive and merits a spot in this special edition of the IJCSS.

Global Positioning Systems are powerful tools for those involved in sports and numerous applications are currently available for those involved in sport. What makes the work by **Eskofier** and **Melzer** so exciting is the integration of the GPS with Google earth and activity monitors (e.g. heart rate) to create a valuable public domain resource. This program is an adaptable and modularized tool that can provide detailed and understandable multimedia data to the user. Hopefully, it will spark much discussion and activity.

Judgment in sport is always controversial. Referees are highly visible and under attack by most competitors, coaches and spectators. **Shang-Juh Kao**, **Chia-Chun Wu** and **Chin-Fa Chen** have undertaken a multifaceted, multimedia approach to design, evaluate and test a system that can effectively and reliably make judgments in volleyball situations. Integrating various technologies and sport vision make this approach highly creative and has great potential for other sports.

Lastly, the article by **Jenny Vincent**, **Pro Stergiou** and **Larry Katz** provides an overview of databases in sport and the potential that they bring for evidence based decision-making. Understanding the value of incorporating multimedia objects and performance data into online multimedia databases for online learning, performance analysis and training are timely and serious considerations.

We hope you find this special issue of the International Journal of Computer Science in Sport informative and inspiring.

Sincerely,

Christoph Igel, PhD, Vice president of media and technology of the German Association of Sport Science

And

Larry Katz, PhD, Professor and Director, Sport Technology Research Laboratory, Faculty of Kinesiology, The University of Calgary, Calgary, Canada.

Virtual Training System Using Visual Feedback for Sport Skill Learning

Kazumoto Tanaka

Department of Information and Systems Engineering, School of Engineering,

Kinki University, Japan

Abstract

Recently, the number of studies on virtual environments for supporting sport skill learning has been growing. Most of the studies have focused on improving the quality of the environment created by virtual reality technology. In contrast, this study focused on the visual feedback of trainees' performances. When a trainee moves within a virtual environment shown on a display device, the trainee imagines the pictures 3-dimensionally. However, since the spatial relation between the trainee and the virtual environment cannot be filmed 3-dimensionally, it is a challenge to get visual information about the trainee's performance fed back to the trainee effectively. This paper proposes a method which makes it possible to visualize the performance 3-dimensionally and feed back pictures using Mixed Reality technology. A virtual Karate system using the method and an experiment on the system are also described.

KEYWORDS: VIRTUAL ENVIRONMENT, SPORT SKILL LEARNING, MIXED REALITY, VISUAL FEEDBACK

Introduction

Over the past decade, a number of studies have been conducted on interactive computer games in sports. Almost all of the game systems have been studied for entertainment purposes, but the number of studies supporting sport skill learning has been recently growing (Katz, Parker, Tyreman, & Levy, 2008).

Researchers have previously studied a virtual batting environment in which a trainee practiced batting using a pitcher and a ball created by computer graphics (Komura, Kuroda, & Shinagawa, 2002). The system could sense the motion of the user's bat and evaluate the contact between the bat and the ball. The evaluation result was fed back to the trainee so that the trainee could improve batting skills. By creating the 3-dimensional motion of the virtual pitcher based on motion capture data from well-known real pitchers, the realism of the virtual experience was improved.

Haptic devices have been also applied to virtual sport environments. A study on virtual boxing where a trainee sparred with a virtual boxer showed that it is possible to feed back punching-impact information using haptic devices (Hasegawa, Ishikawa, Hashimoto, Salvati, et al., 2006). Furthermore, the researchers succeeded in creating natural gaze reactions based on a model of human gaze motions, further improving the fidelity of the interaction between the trainee and virtual boxer (Mitake, Hasegawa, Koike & Sato, 2007).

As described above, most of the past studies have focused on the improvement of fidelity. Realistic reactions from the virtual environment provide appropriate feedback for a trainee.

Such feedback is very important for sport skill improvement.
In general, feedback can be divided broadly into two categories: *Intrinsic Feedback* that a trainee perceives directly (e.g., reaction force when batting) and the *Extrinsic Feedback* that a trainee gets from information from a third person or by artificial means. The latter can be further divided into two categories (Schmidt, 1991). One is the *knowledge of results* (KR); for example, the informing of trainees of their time in 100 meters race. The other is the *knowledge of performance* (KP) that is extracted from such as motion performance video.
Past studies, as described earlier, involved the use of Intrinsic Feedback generated by a haptic device or KR feedback such as the evaluation results of batting. In contrast, this paper proposes a novel visual feedback method as an effective KP feedback. The novelty of this method is in how it realizes visual feedback as an interactive performance with virtual objects (e.g., virtual opponent player, virtual ball).
When a trainee acts in a virtual environment shown on a display device, the trainee imagines the pictures 3-dimensionally. However, since the spatial relation between the trainee and virtual objects cannot be filmed 3-dimensionally, it is a challenge to get visual information about the trainee's performance fed back to the trainee effectively. The proposed method in this paper makes it possible to visualize the performance 3-dimensionally and feed back the pictures using *Mixed Reality* (MR) technology. A virtual Karate system using the method and an experiment on the system are also described in the paper.

Visual Feedback Method in Virtual Training

How does the method visualize the performance?

To visualize a trainee's performance with virtual objects, the method employs MR technology. MR technology superimposes the virtual objects on pictures which the trainee is filmed in. It is also possible to superimpose the trainee image on the virtual environment pictures. MR refers to the merging of real and virtual worlds to produce new environments and visualisations where physical and digital objects co-exist and interact in real time (Mixed Reality from Wikipedia). Using MR technology, visualization of the performance is carried out according to the procedure described below (refer to Figure 1).

Step 1: A marker indicating the world coordinate system is placed on the floor where a trainee performs a motion task prepared for the purpose of learning skills. The trainee and a video camera are then placed at suitable positions and directions in the coordinate system (see Figure 1; Up Left). Next, by using the picture of the marker taken with the camera, the method is able to precisely calculate the camera position and posture expressed by the coordinate system (Kato & Billinghurst, 1999).

Step 2: When the virtual world coordinate system is associated with the real world coordinate system, virtual cameras A and B are placed at the points that correspond to the trainee's viewpoint and the real camera location, respectively (see Figure 1; Down Left). Next, the virtual objects in the virtual world are placed at the initial positions in accordance with the motion task.

Step 3: According to the motion task, the virtual objects are moved to the next position after Δt passes.

Step 4: The method takes two pictures of the virtual objects by virtual cameras A and B respectively and simultaneously captures a picture of the trainee with the real camera.

Step 5: The picture taken by virtual camera A is shown on a display device for the trainee.

Step 6: The method merges the pictures taken by virtual camera B and the real camera (see Figure 1; Right). The mixed picture is added to the image sequence for the feedback.

Step 7: If it is the appropriate time described in the next paragraph, the image sequence is shown to the trainee as visual feedback. Next, if the motion task is not completed, the method measures the position of the trainee using the picture taken by the real camera and changes the view point of virtual camera A according to the measured trainee's head position, and then returns to step 3.

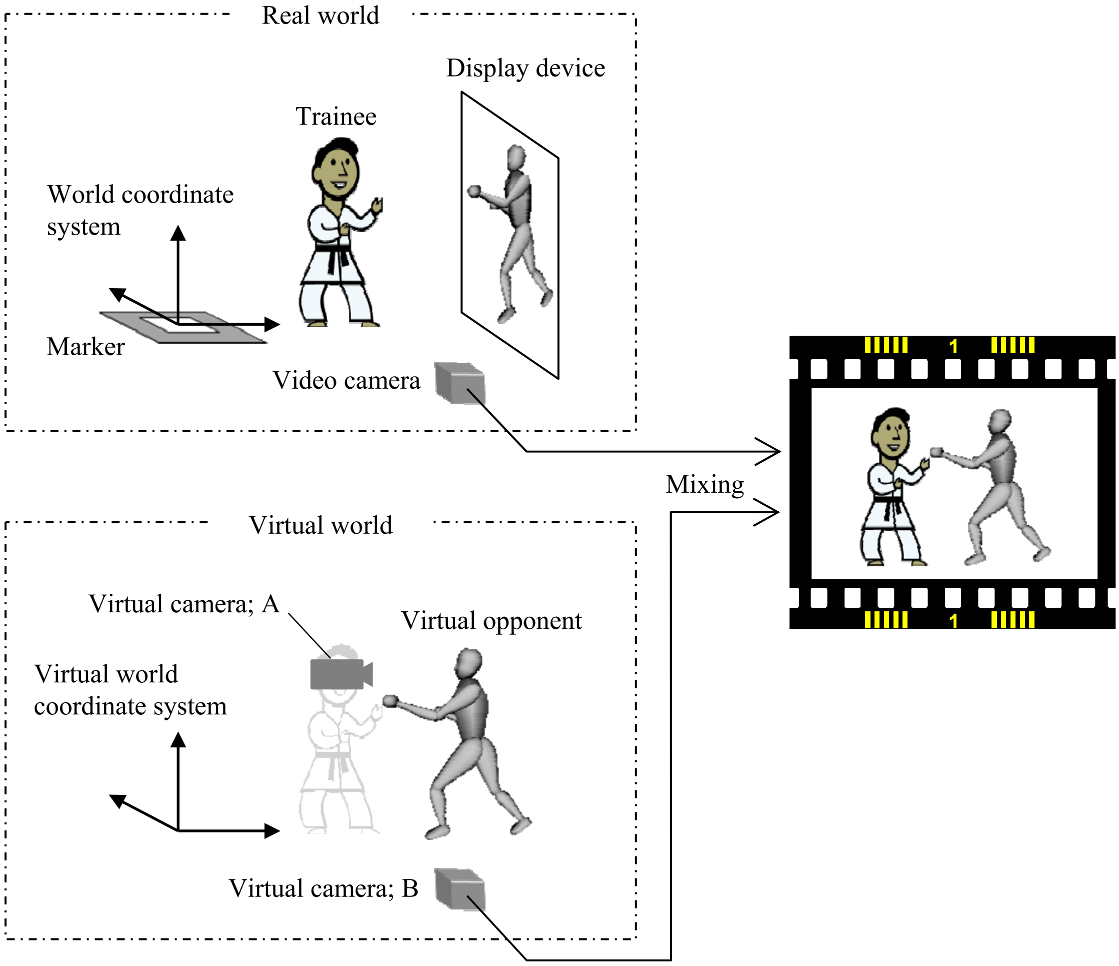

Figure 1. The method for creating 3-dimensional performance image.

If multiple cameras are placed in both worlds and are utilized in the same way as the real camera and virtual camera B in the procedure, the method makes it possible to observe the performance from multiple points of view. It will enhance the integrity of the 3-dimensional visual feedback. Moreover, virtual camera A is suitable as a stereo camera. As a result, it allows projection of the image in a binocular fashion to the display, thereby enabling the trainee to obtain a more realistic three-dimensional image.

When should the method create and feed back the mixed image?

In the research field of physical skill learning, it is said that meta-cognition of one's movements and perception promotes the discovery of unheeded variables in one's own body and acquisition of physical skills (Suwa, 2008). Visual feedback would contribute to this meta-cognition. On the other hand, it is known that visual feedback that includes information

unrelated to skill acquisition excessively obstructs skill learning (Schmidt, 1991). Therefore, the content of the feedback should be limited accordingly. The method accomplishes this by creating contents for feedback only for periods where the trainee performs motions that are strongly related with the skills to be learned. For example, in virtual batting training, the limited period would be defined as the period between the time at which the virtual ball is thrown to the time at which the ball is hit by the bat.
Taking the problem of short-term memory into consideration, it is imperative that feedback is provided immediately. Thus, at the time at which the period ends, the method immediately suspends interactions between the trainee and virtual objects and feeds back the visualized performance of the period.

Virtual Karate Training system

In this study, a virtual training system was built to verify that the visual feedback method based on MR technology has an effect on sport skill improvement. The study focused on a traditional Karate game called *Kumite* that is played by two competitors under non-contact rules (attacks that hit the opponent are prohibited). In the *Kumite* game, bodily contact between the two competitors occurs infrequently, so it is possible for the system to avoid the problem of how to employ haptic devices and the problem of occlusion occurring in the mixing process in MR.
In order to avoid the problems completely, the motion of a basic punching technique in Karate was selected as the skill to be learned, and the system was built for learning the technique. The basic technique is one of several techniques classified as *Sen-no-Joudan-Tsuki*, and is that a player momentarily punches toward an opponent’s face, in accordance with non-contact rules, at the same time the opponent steps forward and reaches rapidly. It is not necessary to take into account any contact and any occlusion between the two players.

System structure

Figure 2 shows the structure of the virtual Karate training system. The system employs a computer system, a display device, a video camera, and a wireless accelerometer sensor device. The computer system is composed of a motion recognition unit, a CG control unit, and a training control unit. The system requires the trainee to hold the accelerometer sensor device in her/his hand for measuring the timing and intensity of her/his punch.

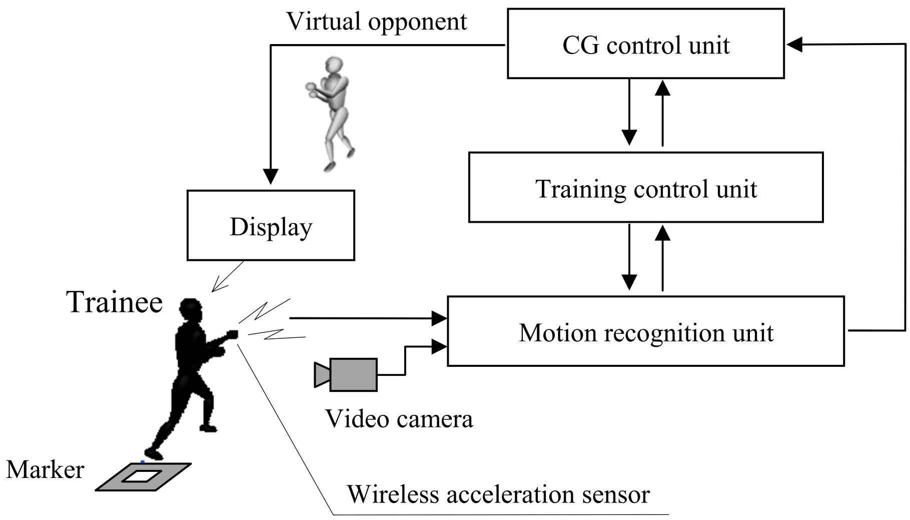

Figure 2. Structure of the virtual Karate training system.

The motion recognition unit receives motion data from the accelerometer sensor device and the video camera, and then extracts the trainee's position and timing information of the punch. The extracted information is sent to both the CG control unit and the training control unit. The motion recognition unit also sends the captured pictures of the trainee's motion to the training control unit.

The CG control unit places virtual camera A (see Figure 1) at the position corresponding to the trainee's head and creates pictures of the virtual opponent's motion as viewed by the virtual camera. The pictures are sent to the display device. Simultaneously, the unit also creates pictures of the virtual opponent's motion as viewed by virtual camera B (see Figure 1) and sends them to the training control unit.

The training control unit generates a motion sequence of the virtual opponent based on the motion task and sends the sequential data to the CG control unit. The generation method is described in detail in the next paragraph. Next, the unit controls the synchronization between the motion captured on the motion recognition unit and the CG created on the CG control unit and thus can precisely create a mixed image. The mixed image is sent to the display device through the CG control unit at the appropriate time described in the former section. The unit also feeds back evaluation results of the trainee's punches as KR feedback. The evaluation is based on the comparison between the timing of the negative peak of acceleration of the trainee's fist and the displacement of the trainee (see Figure 3) when the distance between the trainee's fist and the face of the virtual opponent is less than an appropriately decided reference value, and the peak value of the acceleration is more than another appropriately decided reference value. The negative peak indicates the time at which the punch is rapidly stopped, namely, the time at which the punch is finished. In addition, while providing feedback, the unit suspends creating CG taken by the virtual cameras. Furthermore, in the case of a weak punch, the wrong distance between a punch and an opponent, or a badly timed punch, the unit feeds back error information. The error will be described in detail in the experiment section.

Pictures of a trainee taken by the real camera and a mixed image are shown in Figures 4 and 5, respectively.

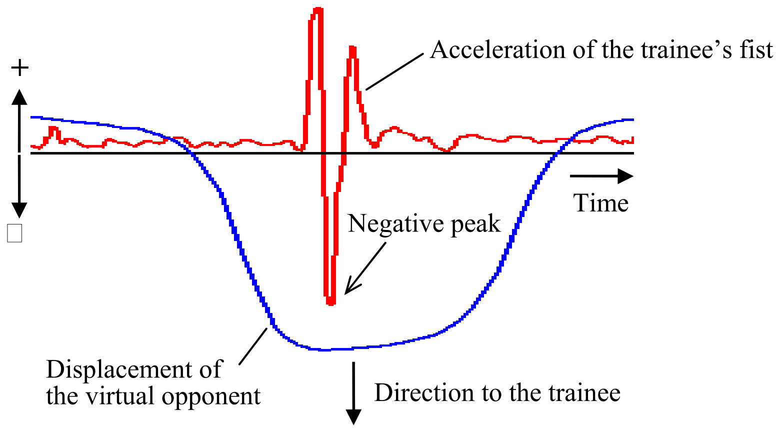

Figure 3. A sample of measurement results of both an acceleration of a trainee's fist and displacement of the virtual opponent.

Figure 4. A picture of a trainee's performance. The trainee wears a head-mounted display where the virtual opponent is shown on.

Figure 5. A mixed picture for the feedback.

Motion sequence of the virtual opponent

The technique selected for verifying the system performance was one of the *Sen-no-Joudan-Tsuki* technique as mentioned previously. In general, trainees practice the technique in the manner described below.

i) While a trainee is standing in a defence posture, the opponent steps forward or backward freely against the trainee.

ii) The trainee tries to punch in the direction of the opponent's face only when the opponent steps forward and reaches rapidly.

In regard to the technique, trainees are not allowed to move except when they try to punch, so there is no opportunity for the opponent to react to the trainees' movements. Thus, for building a simple system, the system does not require a function that changes the virtual opponent's behaviour dynamically with changes of the trainee's position. The system generates the motion sequence by means of stochastically combining four basic motions which are *jumping with defence posture*, *forward steps*, *backward steps*, and *feint forward steps*. These motions were established as the basic motions according to observation of actual training in the technique. In addition, CG animation data of each basic motion was created by motion capturing actual Karate players in action.

The system changes the motion state stochastically within the four motion states around the state corresponding to the *jumping with defence posture*. Figure 6 shows the motion state transition. The symbol $p_{(i)}$ (i=0,1,2,3; $\Sigma p_{(i)}$=1.0) and the values next to the arrows indicate transition probabilities.

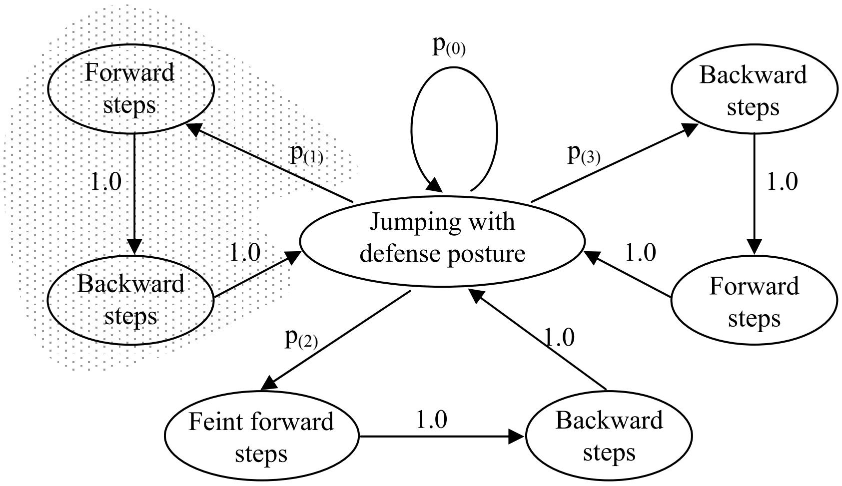

Figure 6. The motion state transition diagram of the virtual opponent.

The timing of the feedback

The period in which a trainee should perform the technique begins when the *jumping with defence posture* transitions into the *forward steps* and ends when the *backward steps* transitions into the *jumping with defence posture* (see the half-tone dot meshing part in Figure 6). Thus, the system was built to be able to carry out the feedback at the end of the period. Since the training control unit of the system is able to grasp the state transitions, it is capable of visualizing the period successfully.

Related Works

Yang and Kim (2002) proposed a virtual training system that was able to superimpose a virtual trainer on pictures of a trainee. The trainee could observe the superimposed pictures in real-time from a first person point of view using a head-mounted display and was requested to follow the demonstrated motion of the trainer as closely as possible. Owing to the observation, trainees were able to receive feedback showing the difference between their motion and the demonstrated motion visually, thus enabling them to learn the motion skills.

Ishii, Hatayama, Seki, et al. (2005) developed a computer-based training system for learning motor skills involved in Kyudo (Japanese traditional archery). The system captured a trainee's motion and created a 3D animation that reproduced the captured motion. Furthermore, the system allowed the trainees to attach comments regarding the perceived motor sense to the corresponding frames in the animation. The trainee was able to observe the animation with the comments and try to improve learning of the motor skills. Although the system could not provide immediate visual feedback, the system's use of animation facilitated the discovery of new physical variables.

A system proposed by Kwon and Gross (2005) was able to extract, from a movie of the trainee performing a martial arts technique, the sequence of frames which captured the target motion. The system was also able to place marks on the frames enabling the trainee to notice significant motions. Thus, the trainee could learn the significant motions by observing the movie.

The studies described above were conducted on visual feedback of trainees' motions. The aim

of these systems is to let trainees learn best forms and motions, namely, to support closed-skill learning. In contrast, this study focuses on visualizing the spatial relation between trainees and virtual objects and is, thus, suited for supporting open-skill learning.
As for recent studies on interactive computer sport games, Pronost, Multon, Li, et al. (2008) proposed a virtual Kung-fu system that simply realized interactions of punch and kick between a trainee and a virtual fighter that was controlled by a Kung-fu trainer's order. The system was able to retrieve motion data that fitted the action order from a motion database and create 3D-CG effectively.
A number of studies have been conducted on such human-virtual interactions, but all virtual humans are on 2-dimensional displays. It is not possible to visualize any spatial relations without introducing a third dimension.

Experiment

The purpose of the experiment was to verify the effectiveness of visual feedback with MR on the virtual Karate training system. Two participants in the experiment, who were considered virtually physically equivalent and had no experience in Karate, were 21 year old male university students. They practiced the basic technique on the system, where one (call him subject A) was given visual feedback while the other (subject B) was not given feedback. Evaluation results of their punch (described below) were fed back to them as KR feedback. The motion transition probabilities were set as p(0)=0.7, p(1)=p(2)=p(3)=0.1.
The evaluation of the punch was made as follows.
In the case of an error:

The error was identified when one of the following four conditions was met.

1) The negative peak of the acceleration was less than 80% of the maximum acceleration of the subject.
2) The distance between the fist and the opponent's face was more than 10 pixels (image size: 640*480).
3) The overlap between the fist and the opponent's face was more than 10 pixels (violation of the non-contact rule).
4) The timing of the punch falls within any of the periods except the period from the opponent's *forward steps* to *backward steps* (the half-tone dot meshing part in Figure 6).

In this case, the system added one point per error to an error counter.

In the case of a favourable outcome:

Having divided the period from the *forward steps* to the *backward steps* into three sub-periods (see Figure 7), where L1 was from the beginning of the *forward steps* to the middle point of the displacement, L3 was the sub-period of the return from the middle point and the remaining sub-period was L2, scores of 2, 3 and 1 points were given when the acceleration peak landed in the L1, L2, and L3 ranges, respectively.

The subjects were required to do the virtual training for 10 minutes at every training session and practiced 1 to 4 times a day almost every day until the target result, where the error count was less than two points and the average punch score was more than 2.5 points, was achieved three consecutive times. The visual feedback was evaluated in terms of the frequency of practice that was needed to reach the target result.
They were also asked to reflect on their own perception and movements during the training, verbalizing what they noticed in the reflection, and recording them at the end of every training session. Recent studies on physical skills have revealed that such meta-cognitive verbalization promotes attention to body variables, and there is a causal relationship between the increase in

the number of words about body parts and skill improvement (Suwa, 2008). Based on this knowledge, the experiment was designed to check the influence of visual feedback on the attention to body variables.

The equipment used for the system was as follows:

PC: AMD Opteron(tm) 2.99GHz, 4.0GB memory, GeForce8800 graphics board.

Acceleration sensor device: WAA-001, Wireless Technologies Inc.

Video camera: Lu075C-IO (640*480, 60fps), Lumenera Corp.

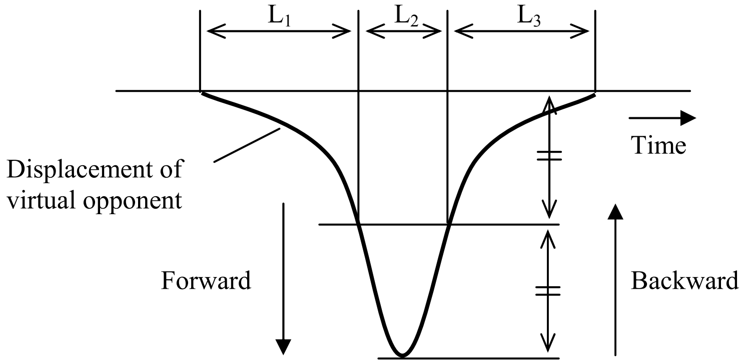

Figure 7. Sub-periods for the evaluation of timing of punches.

Results

Figure 8 shows how the score and the number of words of subject A, each of which was normalized so that the maximum value is one, change until the target result. The horizontal axis of the graph is the serial numbers corresponding to the practice sessions and the vertical axis is the normalized number. Figure 9 shows the results of subject B in the same manner. The frequency of practice done until the target result and the average number of words that concerned body parts and appeared in meta-cognitive verbalization are shown in Table 1.

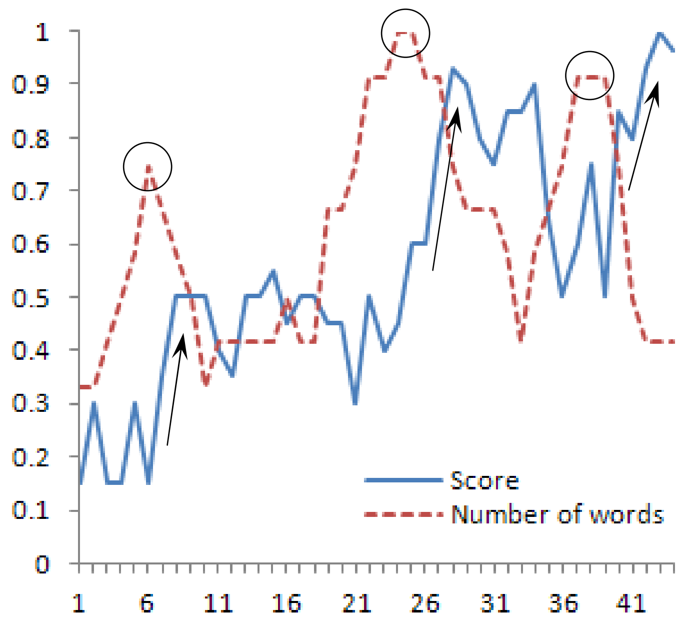

Figure 8. Changes of the score and the number of words of subject A.

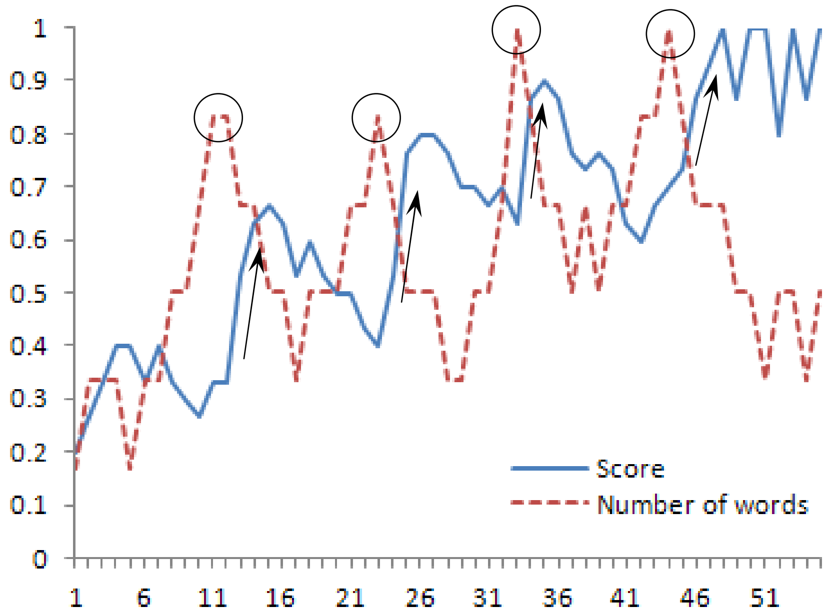

Figure 9. Changes of the score and the number of words of subject B.

Table 1. Frequency of practice and average of the number of words.

Subjects	A	B
Frequency of practice	44	55
Number of words (Average)	7.4	3.4

Discussion

In figures 8 and 9, the comparison between the score and the number of words shows that the score tended to rise (see arrows in the figures) immediately after the number of words had a local peak (see circles in the figures) for both subjects. Moreover, the attention to body variables settled to a low level during the final period when the subjects had mastered the skill. These phenomena match the features in the skill improvement process in real sports trainings that has been revealed by the studies on meta-cognitive verbalization (Suwa, 2008).

The features described above are common and seen in the skill improvement processes of both subjects. However, Table 1 shows that the subject who received visual feedback noticed more body variables during meta-cognition (about 2.2 times more than when there is no visual feedback) and reached the target result faster (20% decrease of the frequency of practice compared with the case without the feedback).

It is important to further examine the effect of visual feedback in a case where subject B receives feedback while preparing another motion task (i.e., another technique). However, such experiments were not conducted because it was thought that the skills acquired by the subjects from the experiment influenced the mastering of other techniques. Future work will involve further verification with higher statistical accuracy by increasing the number of subjects.

It has been known for a long time that visual feedback is effective in skill improvement in the real practice of sports. Similarly, it can be said that visual feedback with MR technology is also effective in virtual training.

Summary

This paper proposed a visual feedback method based on MR technology in an attempt to improve skill with a virtual training system. According to the experimental results, where a shift in attention to body variables was observed, and evaluation scores using a Karate technique, we were able to confirm that visual feedback with MR technology promotes skill improvement. Future work will involve further verification of the method with higher statistical accuracy using a larger number of subjects.

Acknowledgement

This work was supported by a Grant-in-Aid for Scientific Research (c) 20500856 from the Japan Society for the Promotion of Science.

References

Hasegawa, S., Ishikawa, T., Hashimoto, N., et al. (2006). Human Scale Haptic Interaction with A Reactive Virtual Human in A Realtime Physics Simulator, ACM Computers in Entertainment, Vol.4, No.3, Article No. 9.

Ishii, K., Hatayama, J., Seki, K., et al. (2005). Image Training Assist for Motor Skill Learning, Proc. of the 2005 IEEE/ASME International Conference on Advanced Intelligent Mechatronics, pp. 1342-1347.

Kato, H. and M. Billinghurst (1999). Marker Tracking and HMD Calibration for a video-based Augmented Reality Conferencing System, Proc. of the 2nd International Workshop on Augmented Reality, pp. 85-94.

Katz, L., Parker, J., Tyreman, H. and Levy, R. (2008). Virtual Reality *in* Computers in Sport, ed. Dabnichki, P. and Baca, A., WIT PRESS, Southampton, Boston.

Komura, T., Kuroda, A. and Shinagawa, Y. (2002). NiceMeetVR: facing professional baseball pitchers in the virtual batting cage, Proc. of the 2002 ACM Symposium on Applied Computing, pp. 1060-1065.

Kwon, D. Y. and Gross, M. (2005). Combining Body Sensors and Visual Sensors for Motion Training, Proc. of the 2005 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology, pp. 94-101.

Mitake, H., Hasegawa, S., Koike, Y. and Sato, M. (2007). Reactive Virtual Human with Bottom-up and Top-down Visual Attention for Gaze Generation in Realtime Interactions, Proc. of Virtual Reality Conference 2007, pp. 211-214.

Mixed Reality from http://en.wikipedia.org/wiki/Mixed_reality

Pronost, N., Multon, F., Li, Q., et al. (2008). Interactive Animation of Virtual Characters: Application to Virtual Kung-Fu Fighting, Proc. of International Conference on Cyberworlds 2008, pp. 276-283.

Schmidt, R. A. (1991). Motor Learning and Performance: from Principles to Practice, Human Kinetics, Pub., Champaign, Illinois.

Suwa, M. (2008). A Cognitive Model of Acquiring Embodied Expertise Through Meta-cognitive Verbalization, Transaction of the Japanese Society for Artificial Intelligence, Vol. 23, No. 3, pp. 141-150.

Yang, U. and Kim, G. J. (2002). Implementation and Evaluation of "Just Follow Me": An Immersive, VR-Based, Motion-Training System, PRESENCE, MIT Press Journals, Vol.11, No.3, pp. 304-323.

Development of an Application for Learning and Teaching Soccer Tactics

Roland Leser, Johannes Uhlig, Manfred Uhlig

Department of Biomechanics, Kinesiology and Computer Science, Faculty of Sport Science, University of Vienna, Austria

Abstract

"Pervasive computing" in the sense of "permeating", but also in the sense of supporting human day-to-day living with computer applications has long entered the field of teaching and learning. In the area of education, developments of this kind often come up under the keyword "e-learning". Preparing teaching and learning materials based on multi-media and using interactive elements and interactivity seem especially promising when complex, dynamic, and real situations should be presented. A good example therefore, is the subject area of sports. The authors describe an e-learning system, which was developed to teach and for learning selected fields of soccer tactics based on the design-based research strategy.

KEYWORDS: E-LEARNING, MULTIMEDIA/HYPERMEDIA SYSTEMS, SOCCER, TACTICS, TEACHING/LEARNING STRATEGIES

Introduction

Digitalizing of modern world brings about a rise in new possibilities of technology-based teaching and learning, which in the meantime are widely in use. Shoham and Perry (2009) for example refer to organisation-wide technological changes that have infiltrated higher education. Over the past years, tendencies toward computer or internet based educational systems could increasingly be observed in the field of sports, which for example is evidenced by a large publicly funded project in Germany (Sturm, 2008).

The so-called new media are showing their strengths in teaching and learning specifically when traditional teaching materials (book, static means of communication) fail. Probably the best example is found in subject areas that contain dynamic phenomenological and real situations. In their meta analysis, Hoffler and Leutner (2007) confirm that dynamic visualisations are far more effective in such cases. Those findings are used to communicate specific contents in sport/competitive games such as presenting and simulating tactical behaviour (Senner, 2008).

The article at hand addresses a problem coming from soccer practice respectively from soccer coaches' education. In soccer group and team tactical capabilities are very important. Nowadays, almost all teams playing at an elevated level of performance use zone defence or zone coverage as basic tactical concept. But in contrast to man-marking, which was the basic concept in earlier times, zone defence is very difficult to convey, as attention must at any given time be directed towards several different aspects (ball, opponent, team-mates) and must dynamically adapt to every game situation. This paper first describes content and process of a project, which was conducted by the authors on this subject and which focused on the possibilities of new media or rather animation and multi-media based educational materials. Subsequently, the topic of tactical abilities in soccer was addressed to prove the didactical considerations of the subsequent presented e-learning application. This article will be concluded by presenting the evaluation of the system that was introduced together with its resulting perspectives.

The Project "Multimedia-supported Tactical Learning in Soccer"

The authors of this article conducted in the years 2005 to 2007 an e-learning project dealing with the topic of how to implement forms of new media to improve learning processes in soccer practice. In particular, a concept was designed for learning and conveying zone defence, which is supposed to be one of the most important tactical basic elements in soccer. Furthermore, this concept was consistently implemented in the e-learning application T-A-P (Tactics-Animation-Programme), which is explained in chapter *The e-learning application T-A-P*. In this section, first of all the motivation and objectives for conducting the project are discussed, followed by explanations of the further course of the project and the therein realised research strategy.

Project Motivation and Objective

The motivation to deal with the present thematic was the question of how modern soccer tactics can be conveyed as efficiently as possible. This question was often asked of the authors, who are partially active soccer coaching teachers themselves, by coaches and coaching trainers, as the new organisation forms in soccer are much more difficult to teach than earlier forms. It was decided to develop a multimedia-supported tool for the teaching and learning of soccer tactics, paying special attention to zone defence because the practitioners were also having difficulty learning and teaching the dynamic aspects of zone defence with old media (book, static presentation forms), This tool should also be suitable to be used as a learning medium and accompanying teaching training.

Starting Investigations

A main part of the project was the elaboration of exercise and game forms, which highlight a predefined methodical-didactical teaching path for zone coverage. A decisive question in connection with the multi-medial preparation of these exercise and game forms is the modality of presentation. After discussions among the authorship and target group, the exclusive illustration per video was promptly ruled out. Video sequences (especially from actual games) rarely demonstrate tactical principles in a perfect form. Moreover, the viewing perspective is often not very useful, because one rarely has the proper top view of the action. For this reason, other media or possible combinations were considered and to assure empirical validity a study was conducted (Leser, 2005):

A comparatively complex pass exercise was chosen to assess the efficiency of three instructional alternatives (exercise instruction either just verbal, additionally with the help of a static graphical illustration, or with animated pictures added). The exercise (Figure 1) starts at two diagonal corners (A, D) and with two players passing the ball to the central running stations synchronously (1). Each of the five running stations B to F is engaged with one player, at the positions A and D stand two players at the beginning. After the player at running station B receives the ball from A, he passes it to C and follows the ball by running to position C. In the meantime the player from station A runs to B (2), receives there a ball from position F (3) and passes this ball directly to the player who runs from position F to A (4). After that the player who is at station B now receives the ball from A (5), stops it, turns into the direction of position C, passes the ball (6) and follows it (7). Arrived there, he receives the ball from B (8) and passes it to position E (9) and so on.

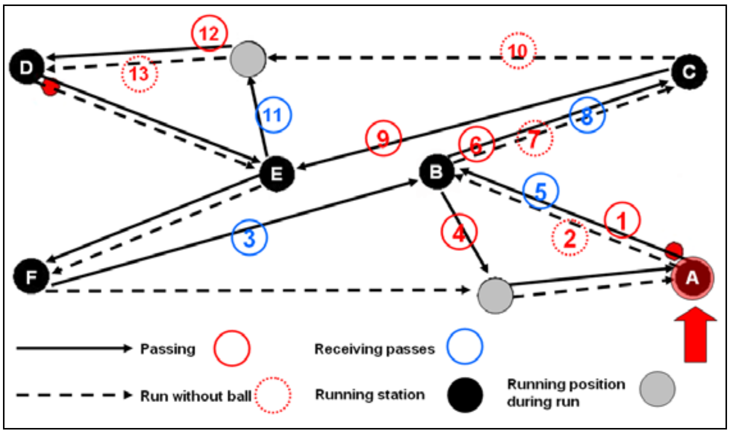

Figure 1: Graphical illustration of the pass exercise with six main running positions. The numbers represent the sequence of action from the perspective of the player starting from running position A.

The course of the experiment was that way, that seven players acted like a well-rehearsed team – who were able to perform the exercise without any problems. One missing player was integrated into the team. This player got either verbal instructions about the pass exercise by the trainer, or was additionally instructed by graphic illustration or the help of an animation. The exercise was repeated for as long as it took to perform it without any problems and the test subjects played without any errors.
The progress of the exercise was video recorded, digitalized and computer-assisted analysed. The quantification of the length of time it took to reach "proficiency" in the pass exercise, was carried out by the number of the subjects' running positions. The determination of the moment, when the exercise was deemed as mastered, was defined in a criteria catalogue (Interrater-reliability: 81%). Overall, the experiment was carried out on 20 subjects per test group (verbal, illustrational, animation). Special attention was given on performance homogeneity of the athletes.
Summarised, the study produced similar results as the ones found by Lewalter (2003) in connection with instructional methods in dynamic behaviour aspects:
It shows a difference in the average length of time it takes to master the exercise between the group which received verbal instructions only as well as the group instructed by images and the one instructed with the support of images and animation (Figure 2). However, statistically significant ($p<0.05$) is only the difference between the players who were instructed only verbally and those which received animations as support.

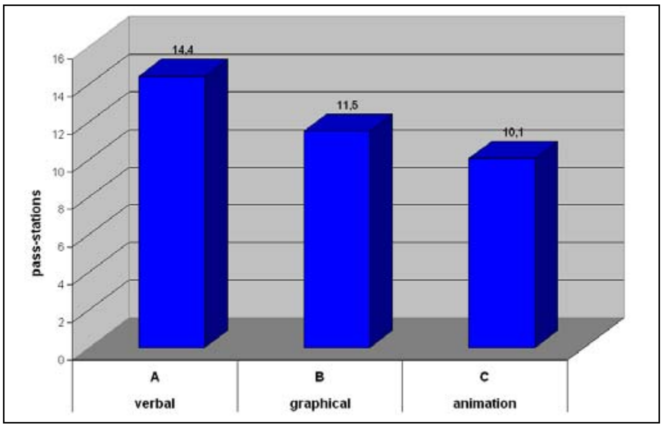

Figure 2: Number of pass stations necessary to master the exercise in the three groups.

It should be noted that instructions with animations tend to be better than instructions with static images. However, the explanation with moving pictures took considerably more time ($p<0.05$), because the players wanted to watch the exercise repeatedly.
Based on the investigational results, we can conclude, that in teaching sports game tactic animations should be implemented especially when the time factor is not important (lecture room prior to game/training etc.). However, images/graphic illustrations are better suited, if the available time is limited (during training sessions, breaks, timeouts during competitions etc.).

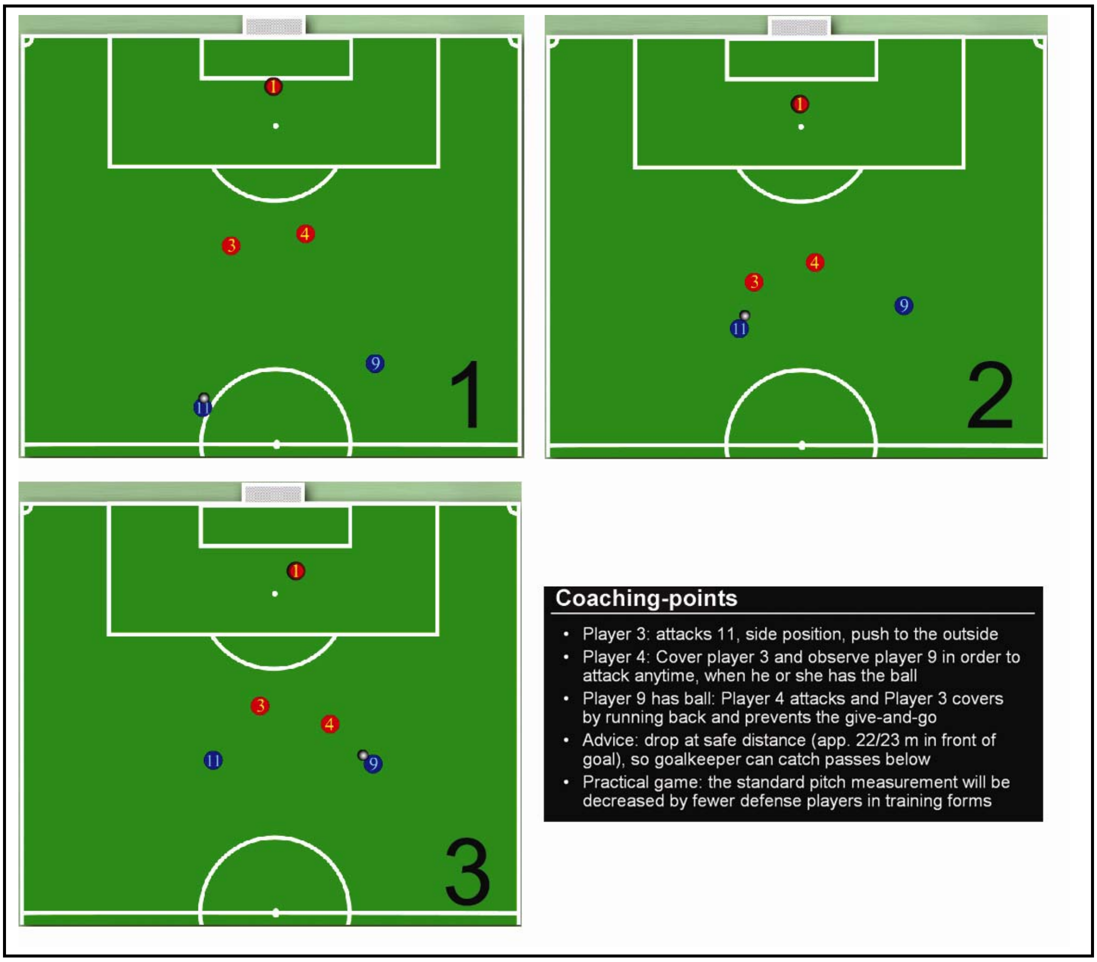

Figure 3: Exercise form “Defence by two“ with corresponding coaching advices.

The findings suggest, that the proper medium for the e-learning application with the above stated objectives are animations. The degree of abstraction of the images in motion is hereby a significant aspect under consideration (Wouters, Paas, & Merrienboer, 2008). Although, today's computer technology allows an already high reality 3-D animation, these do not always yield the same desired effects in learning situations. Presenting sportive movements demonstrates that a medium degree of abstraction shows a high degree of recognition in the trainee. The authors also transferred these results to the illustration of tactical behaviour patterns by showing the objects (player, ball, etc.) highly abstract, but the dynamical aspects (timing, synchronism etc.) rather realistically. Figure 3 shows an elementary example of an exercise form for beginners with the corresponding coaching advices by means of three freeze images of the animation.

Research Strategy

The continued course of the project was shaped by closely cooperating with the target audience, for which the e-learning application was intended. This was expressed by developing programme elements together and continuously evaluating completed project phases. As mentioned above, the project was created from practical problems and during the entire implementation; it was strongly intertwined with sports. Without doubt, some partial areas of the project were "typical scientific" such as implementing the discussed study of presentation media. Nevertheless, the software development was strongly setting-oriented and thus is in a certain way a contrast to traditional research approaches. Taking a closer look at the approach during project realisation or rather the phases of software development, one notices certain similarities with a relatively new research method in the field of educational technology, namely the design-based research approach. Design-based research deals less with decisions about truth or falseness or examining truth in a way classical scientific methods look at it. It is more about usefulness, matching certain conditions, and future potentials (Bowler & Large, 2008). The necessity to develop design-based research methods is more often found when dealing with complex problems in teaching and learning processes. Teaching and learning is not based on trivial actions-reactions or cause-effect principles. Many times, they consist of a very complicated phenomenon, which cannot simply be packaged into classical experiments and into clearly structured sets of variables. Design-based research is an attempt to leave the purely quantitative level of investigation. It tries to increase the possibility to reach usable results through a qualitative approach. In particular, the creative developmental process (in this case, it concerns the development of an e-learning software) should use the experience of participants. The term "design" refers to the relationship of e-education to systematic development and design in creating teaching and learning software (Barab & Squire, 2004, p. 3). Decisive factors are primarily experiences made during this process.

Tactical Capabilities in Soccer

For a long time training to enhance the complex performance for the team ball sport of soccer only concentrated on the concept of biological adaptation on the basis of the process of supercompensation. In doing so the basal conditions for physical and energetic fitness have been the centre of attention, fully having the orientation, adaptation and therefore increase in performance of the organism in mind (Hohmann, Lames, & Letzelter, 2002).
However, soccer is highly tactical determined, which is why the mentioned adaptation principles can no longer alone be considered for performance evaluation. Rather more significant is the teaching and experiencing of tactics, although this has been neglected or even ignored, especially concerning the cognitive and psychological process of information

and self-organisation. Hoffmann's model for cognitive learning anticipative behavioural control (1993) initiates prospective expectations about how soccer players may learn tactical action alone by playing. This way of learning, which happens by chance and in passing with regard to content is also called incidental learning of tactics and should be offered beginners when starting their training (Hohmann, Lames, & Letzelter, 2002). The player expects certain results of his or her action before he or she solves a specific situation (for example dribbling with an opponent in a man-to-man situation). If the result of individual action is positive (e.g., successful dribble) the player then receives the „inner response" of having done the right thing. If the result is negative (i.e., the player loses possession of the ball in the man-to-man situation), he or she realises that the taken action is not apt to solve the situation successfully and therefore will take different actions in the future. The more experienced in playing the children get, the broader their "tactical knowledge" becomes on how complicated situations in play can be solved.

This imminently begs the question if soccer should rather be taught in a game oriented way, i.e. in an integral manner, or predominantly in an analytic-elementary manner and which of both ways leads to the goal more quickly and more effectively (Roth, 2005). Undoubtedly this question cannot be solved without bearing in mind the age and performance level of the player and – according to the present state of research – it's not that there are no two ways about it but the question should be settled combining both. Implicit and explicit learning of tactics – methodically consisting in playing and training – are in accordance with a combination of cognitive and constructivistic teaching methods without any doubt legitimate.

While the theory of cognitivism adopts the positive approach of problem solution by using internal information processing models, i.e. the soccer game situation is solved by means of explicit tactics and tactical instructions, the constructivism model assumes that the player finds creative and effective solutions to specific situations by means of implicit training methods (variable styles of play) based on individual experience (Schwetz, 2001; Memmert & Roth, 2007).

The E-learning Application T-A-P

T-A-P stands for "Tactics-Animation-Programme". It is an e-learning application that was developed within the project "Multimedia-supported Tactical Learning in Soccer". The following section gives an overview on the application and the underlying theoretical considerations.

Considerations of Content

Regarding to the content, the e-learning application is roughly divided into a theoretical part and two blocks with exercise and game forms. In the theoretical part, the structural basics of the game of soccer, the most important principles of zone defence and the concept of the game systems are treated and practical training instructions then provided. A comprehensive glossary ensures that the terms used are also comprehensible for beginners and soccer laymen.

The core of the programme is formed by about 100 exercise and game forms, which are organised in a total of ten categories in the two areas of the defence phase and the offence phase (table 1).

These categories base on the didactical-methodical teaching path of the tactical education in soccer (Peter, 2003). The defence, starting with two defenders, is trained with an increasing number of players up to group and team pressing (e.g., midfield pressing). During the offence phase, one can differentiate up to five team-tactical offensive tactics, which may be used accentuated depending on the own team's strength and considerations on particulars of the opposing team. Each of the five categories of defence and offence phase form the basic orientation after which training and game forms were designed in detail in T-A-P. Moreover, in the defensive area they are coordinated with each other in a way that they follow the teaching path of zone coverage. Furthermore, these exercises and games present a basic repertoire of tactical training types, which can be developed into a comprehensive “catalogue of tactics” by using specific variations.

Table 1: Areas and categories of exercise and game forms.

Defence	Offence
defence with two players	continuoulsy structured attack
defence with three players	opening an attack
defence with four players	wing play
defence with more than four players	playing down the centre
pressing	counter attack

The sophisticated exercises basically treat all important principles of zone defence, but also contain ideal and typical guides for actions in specific situations (e.g. behaviour when a man up/man down, play at the wings, possibilities after winning the ball following a corner by the opponent, and much more) and principles of behaviour for each team position. The latter always refers to behaviour in the currently most common international system of play – the 4:4:2 with diamond (Figure 4).

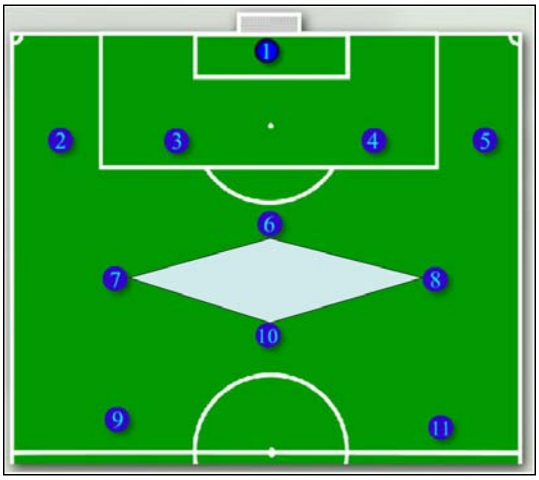

Figure 4: Game system 4:4:2 with midfield-diamond and typical player numbers identifying the playing positions.

Didactic-Methodical Conception

The Tactics-Animation-Programme (T-A-P) rests upon a game- and action-taking-oriented concept, assuming that the tactical capacities are activated primarily by styles of play whereas – as far as tactics are concerned – especially the questions “what has to occur“ and “when does it have to occur during the game“ are of great significance. It is only of secondary importance “how“ to do something from a technical point of view (see also e.g. Werner, Thorpe, & Bunker, 1996; Rink, French, & Tjeerdsma, 1996; Griffin, Mitchell, & Oslin, 1997).

Beginners learn especially in an incidental and implicit way and create their own strategies by solving many diverse situations according to the in mentioned model (chapter *Tactical capabilities in soccer*) of anticipative behavioural control (Hoffmann, 1993), by continuous reinforcing of successful action. “Free“ play without interaction through the trainer predominates (see also Werner & Almond, 1990) and will gradually be replaced by tactical play and by interventions as the players get older and more experienced (explicit teaching and learning of tactics). Besides the observation of tactical principles in soccer (Uhlig & Uhlig, 2000), the creative and surprising momentum as well as the use of variable solutions to specific game situations must be trained and paid attention to (see also Memmert & Roth, 2007; Memmert & Perl, 2009).

For didactical-methodological reasons the spatial conditions for styles of play can be organised in a way that a desired behaviour is provoked (e.g., play in restricted areas). For instance, zones at the side are marked for wing-play so that the player can move without being attacked. Players take turns performing diverse styles – whereas tactical aims have to be met – and practising sequential drills. This explicit teaching in the form of technical-tactical drills is guided by instructions (coaching-points) and helps cognitive processing of knowledge. It represents some promising possibilities in finding solutions. One should keep in mind that variable, diverse strategies should be tried as soon as possible (choose where to run to, pass techniques, completion etc.) in order to challenge adaptation capability conform to new and unexpected situations, in order to evoke individual and collective self-organisational process (see also Schöllhorn, 1999). These special tactics also aim at the use of appropriate, creative and variable strategies. Technical-tactical drills will then be integrated in the play and harmonised with the factors pressure (opponent, field, time), physical stress and high precision in order to train adequate timing.

To gain a better understanding the reader will now find a short description of a specific problem and its solution by means of T-A-P:

A tried and tested method of attacking in modern soccer is the wing play, as especially the centre has to be defended, in other words the direct line to the goal zone has to be covered. By doing this, naturally enough, the boundaries remain partly unmarked and the attacking team can take advantage of this. The problem of attacking on the wing is bringing the ball into the centre in front of the goal of the opponent team. T-A-P includes four types of training with group-tactical elements like third man running, crossing (crossover runs) and playing in a one-two, which represent tactical methods in order to start with promising wing play – at first without tackling influence of the opponent (Figure 5). Interacting with T-A-P means that these types of training can be carried out in different speed and may be stopped in order to illustrate the essential coaching-points of the move and in order to trigger mental representations of promising movements. The coach pre-plans the directions for the players to run and pass for concrete practice and they are discussed with him or her before and only then realised within a so called “fixed move“. This may be called explicit learning of tactics. In detail complicated techniques (e.g., direct attack, good pass from the flanks, volleys) are hereby practised in multi-faceted training situations. The trained course of play will in the following be integrated in types of game (with opponents), at first with concrete targets and then freely, i.e. without external intervention. T-A-P all in all enables six types of play for effective winging. If one looks at these types of play in detail the didactical-methodological process from “pre-set situation training“ in wing areas towards “free (wing-)play“ is apparent.

Figure 5: T-A-P-Screenshot showing an exercise for wing play (explicit form of teaching).

Evaluation

For quality assurance and development, formative and summative methods were used to evaluate the T-A-P e-learning application. They were based on the following evaluation design:

- The formative evaluation method was conducted during the project and consisted of one of the authors' field notes. On the other hand, short key questions requesting opinions about the programme were asked during the education of trainers in approximately three-month intervals. These measures were applied mainly during the application's developmental phase. However, they were also integrated during the planning stage.
- For the summative evaluation method a questionnaire and closed questions were used after the programme was completed. Therefore, it constituted a necessary feedback during the implementation phase or rather about the effects of the software application.

Formative Evaluation

Especially suitable in this setting proved that part of the target group (tactic coaches and students) was integrated into the author group or rather into concept development. Thus, one was able to consider their feedback continuously. Besides these participative elements for quality development (Burnett, Brunstrom, & Nilsson, 2003), the programme was evaluated using the formative method. In intervals of several weeks, players and trainers in training were asked to provide an assessment. This was accomplished by using open key questions (such as: Which content should be covered by the application? Is the content you desired properly treated? If not, what would you like to improve? etc.). Those questions were discussed by participants in trainer education during open discussions, after those had the chance to test prototypes of T-A-P sufficiently. Keywords of these discussion contents were written down by course managers and subsequently discussed with the authors.

Summative Evaluation

At the end, 52 users were evaluated using the summative method. A total of 50 men and two women were surveyed. The average age was 37 years. Users used T-A-P on average five times. The average time of a session was between 20 and 30 minutes. Prerequisite was that the surveyed population had to be at least one year trainer for an adult team and they had to have used T-A-P in their current training as learning medium. Furthermore, they also had to evaluate the software concerning its potential as teaching medium in their training praxis.
The evaluation was conducted by a questionnaire following Wiemeyer (2004). The surveyed group answered demographic questions (age, sex) and questions about using the software (frequency and duration) using a four-level assessment scale (1 to 4 – very good/good/bad/worse; applies very much/…/does not apply at all, etc.) for a total of 40 additional questions to six items. Figure 6 shows the items evaluated together with their assessment. The average grade for all items was 1.42.

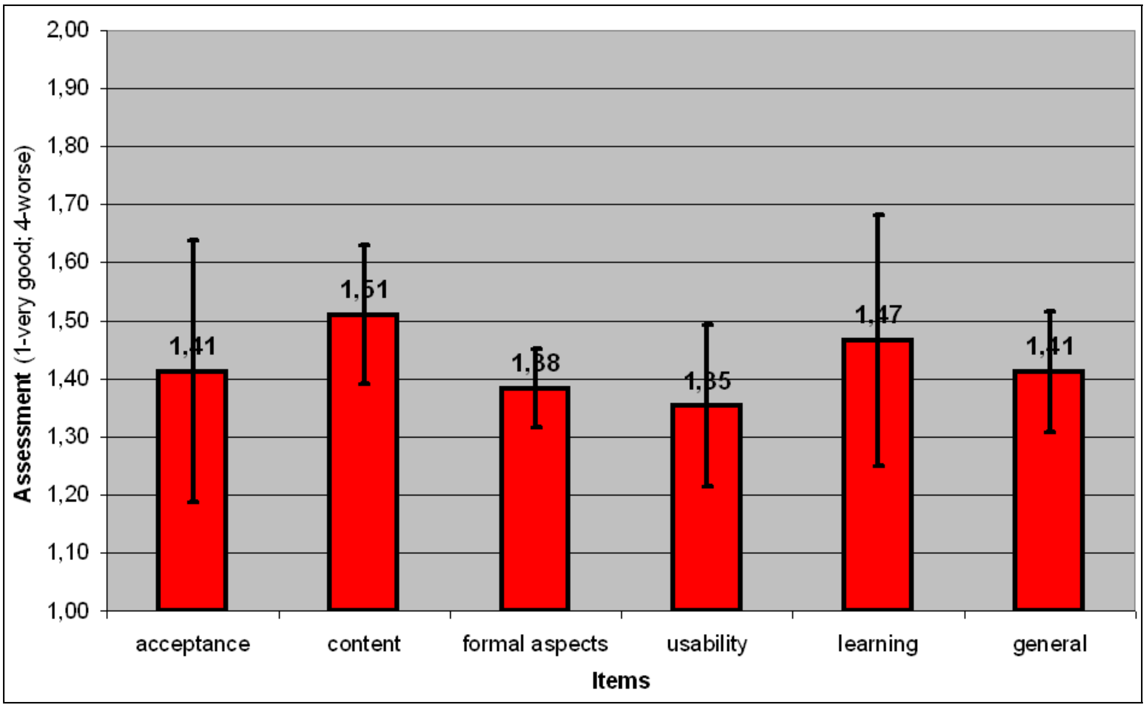

Figure 6: Summarised result of the summative evaluation via query.

Results and Consequences

The formative evaluation method gave decisive impulses for target-specific development of the e-learning application according to design-based research. In particular, these quality assurance measures have made a positive impact on the overall evaluation of the application (summative evaluation).
Regarding the application’s objective, one can emphasize that questions concerning the programme’s suitability for self-study was judged highly positive. Most surveyed said that working with this programme helps to understand better the tactical areas to be covered. The surveyed saw the programme’s weaknesses in T-A-P that it does not include a knowledge test, context between animation and real game scenes was missing, and there was no possibility to export the exercises in writing in order to take them to the practical training to help remember them. Furthermore, we were often asked for more interactive possibilities in discussions about animation. The authors sought to eliminate these defects by expanding the programme:

Questions concerning tactical theory and behaviours in certain game situations were catalogued. Answering randomly chosen questions should provide the user feedback to his or her knowledge or rather precise notes about which areas of the e-learning application he or she should work on more intensively. Furthermore, video sequences were cut from real games, which can be retrieved in various exercise categories to provide a more practical context. In the final expansion, animation of exercise and game forms were prepared in a way that they can be printed compact on one page together with coaching points and therefore are suitable as note sheet for practical training.

Discussion and Perspectives

Taking up again the idea of design-based research introduced in chapter *Research Strategy* and mirror this thought in the developmental phases of T-A-P, it results in the following congruencies:

- the initial idea for this software design was derived directly from concrete problems of real situations (sports)
- the project pursued objectives of science and of practical use,
- basic considerations and questions about structure were taken from existing knowledge inventory and empirical knowledge,
- the developed e-learning software was systematically evaluated and the results were utilised to advance system development (redesign),
- new concepts created by the design process (study about choice of medium, preparing exercise-oriented learning) have good transfer potential to other contexts.

In this research strategy, the formative evaluation method and its immediate consequences for the further course of the project were most significant among all other stated measures. The success of these measures again can be assessed using the results of the final evaluation by the summative method. Most points, which were objectionable to the target group during the evaluation, could be counteracted with expansions. However, a significant criticism was that while one can control interactively the training animations of T-A-P, one is not able to interact. This gave reason to add one module to the e-learning application. In principal, T-A-P is designed as programme by which one can teach and learn zone coverage in soccer with exercise and game forms in methodological successive stages. The exercise and game forms are dynamically presented as animations. However, they do not provide any possibility for the user to interact by providing input-output, action-reaction functions as possible under the web 2.0 logic. According to Senner (2008), add-on functions could be possible for certain game situations, which are presented in selected animations. This would allow integrating the demanded interaction. One possibility would be that the user can freely move or place objects (players) of the animation (game situation) at certain or any instants during the exercise. For example, this could allow observing the schematic reaction of a defence formation depending on the movement of the user-controlled striker.

The authors believe that for future use of this learning medium, the possibilities to combine this application with existing methods of systematically observing sports games (Leser, 2006), is especially promising. Here the developed application could be embedded in the concept of intervention (Lames & Hansen, 2003). Any weaknesses seen during the game's analysis in the individual, group and team area could be reflected with T-A-P. To accomplish this, the tactical defects, which were determined, would have to be filtered from the entire catalogue, which would present the most important issues to rectify the mistakes observed.

This would allow players to understand tactical measures properly. This effect could be increased by juxtaposing video scenes showing the mistakes.
After the above stated, the e-learning application T-A-P can surely be termed an innovation under modern understanding. Although, there are similar applications in various types of sports, but the combination of didactical model, methodological learning, used medium and targeted implementation as practical tool realised in T-A-P is new according to the authors' view and takes a unique position as of today.
The questions, which remain unanswered, are whether this application will establish itself in day-to-day soccer training and whether it will be useful or change soccer. Those questions cannot be answered definitely, because the programme in relationship to the size of the target population has only been in use for a short time. One can only refer to the results of the evaluation and the very positive feedback provided by trainers using this application for the first time.

References

Barab, S., & Squire, K. (2004). Design-Based Research: Putting a Stake in the Ground. *The Journal of the Learning Sciences 13*, 1-14.

Bowler, L., & Large, A. (2008). Design-based research for LIS. *Library & Iinformation Science Research 30*, 39-46.

Burnett, R., Brunstrom, A., & Nilsson, A. (2003). *Perspectives on Multimedia: Communication, Media and Information Technology*. Chichester, Hoboken, NJ: Wiley.

Griffin, L., Mitchell, S., & Oslin, J. (1997). *Teaching sports concepts and skills. A tactical games approach*. Champaign: Human Kinetics.

Hoeffler, T., & Leutner, D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction 17*, 722-738.

Hoffmann, J. (1993). *Vorhersage und Erkenntnis*. Göttingen: Hogrefe.

Hohmann, A., Lames, M., & Letzelter, M. (2002). *Einführung in die Trainingswissenschaft*. Wiebelsheim: Limpert.

Lames, M., & Hansen, G. (2003). Computer Science for Top Level Team Sports. *International Journal of Computer Science in Sport 2*, 57-72.

Leser, R. (2005). Zur Effizienz verschiedener Präsentationsmedien bei der Vermittlung von Sportspieltaktik. In S. Würth, S. Panzer, J. Krug, & D. Alfermann (Eds.), *Sport in Europa* (p. 357). Hamburg: Czwalina.

Leser, R. (2006). Applied video and game analysis in professional soccer. In H. Dancs, M. Hughes, & P. O'Donoghue (Eds.), *Proceedings of the 7th World Congress of Performance Analysis of Sport*. Szombathely, Hungary.

Lewalter, D. (2003). Cognitive strategies for learning from static and dynamic visuals. *Learning and Instruction 13*, 177-189.

Memmert, D., & Roth, K. (2007). The Effects of Non-Specific and Specific Concepts on Tactical Creativity in Team Ball Sports. *Journal of Sports Sciences 25*, 1423-1432.

Memmert, D. & Perl, J. (2009). Game Creativity Analysis by Means of Neural Networks. *Journal of Sports Sciences 27*, 139-149.

Peter, R. (2003). *Fußball von morgen. Modernes Verteidigen*. Münster: Philippka.

Rink, J., French, K., & Tjeerdsma, B. (1996). Foundations for the learning and instruction of sport and games. *Journal of teaching in Physical Education 15*, 399-417.

Roth, K. (2005). Sportspiel-Vermittlung. In A. Hohmann, M. Kolb, & K. Roth (Eds.), *Handbuch Sportspiel* (pp. 290-308). Schorndorf: Hofmann.

Schöllhorn, W. (1999). Individualität – ein vernachlässigter Parameter? *Leistungssport 29*(2), 5-12.

Schwetz, H. (2001). Neues Lernen für die Informationsgesellschaft. In H. Schwetz, M. Zeyringer, & A. Reiter (Eds.), *Konstruktives Lernen mit neuen Medien. Beiträge zu einer konstruktivistischen Mediendidaktik* (pp. 35-52). Innsbruck et al.: Studienverlag.

Senner, S. (2008). *Modellierung und Simulation des Abwehrverhaltens im Hallenhandball*. Dissertation, University of Augsburg (submitted).

Shoham, S., & Perry, M. (2009). Knowledge management as a mechanism for technological and organizational change management in Israeli universities. *Higher Education 57*, 227-246.

Sturm, R. (2008). *Internetbasiertes Wissensmanagement in Sportwissenschaft und Sport*. Dissertation, University of the Saarland.

Uhlig, M., & Uhlig, J. (2000). Taktische Prinzipien im Fußball. *Leistungssport 30*(2), 18-23.

Werner, P., & Almond, L. (1990). Models of games education. *Journal of Physical Education, Recreation and Dance 61*(4), 23-27.

Werner, P., Thorpe, R., & Bunker, D. (1996). Teaching games for understanding: Evolution of a model. *Journal of Physical Education, Recreation and Dance 67*(1), 28-33.

Wiemeyer, J. (2004). Analyse der Qualität der multimedialen Lernumgebung „Bio-Prinz“. In M. Stempfhuber (Ed.), *Sharing Knowledge: Scientific Communication: 9th Conference of the IuK-Initiative of the Scientific Association in Germany* (pp. 283-292). Bonn: GESIS.

Wouters, P., Paas, F., & Merrienboer, J. (2008). How to optimize learning from animated models: A review of guidelines based on cognitive load. *Review of Educational Research 78*, 645-675.

GERVIT - A Tool for Sports Data Vizualization Employing Google Earth

Björn Eskofier[1] & Daniel Melzer[2]

[1]Faculty of Kinesiology, University of Calgary, Canada

[2]Institute of Pattern Recognition, University of Erlangen-Nuremberg, Germany

Abstract

We developed a software tool for runners and other sportsmen that is capable of visualizing recorded Global Position System (GPS) data of sports activity. If available, heart rate (HR) as well as other meta data can be visualized along with the position. The goal of the work was to create public domain software under the general public license (GPL) for post-training-analysis of running sessions and the possibility to exchange the visualized result with others via the Internet. The software and its sources are available for download.

The tool was realized using Google Earth and its Keyhole Markup Language (KML), which allows to visualize the data directly in Google Earth and to pass visualized data to others. Given the input data, our software creates a KML file. Besides the running path it is also possible to visualize speed and heart rate as well as other information such as elapsed time, average speed, average heart rate, running distance and altitude. Data from different sources can be automatically registered in order to improve visualization.

KEYWORDS: SPORT DATA VISUALIZATION, GPS INTEGRATION, GOOGLE EARTH, POST TRAINING ANALISIS, DATA REGISTRATION

Introduction

Multimedia preparation of data collected in sports gains more and more importance, see for example Papaioannou et al. (2004) or Tong et al. (2005). This is also true for training data collected by running athletes and even amateur runners, who use devices such as GPS and Heart Rate Monitors (HRM) in order to collect information such as speed, heart-rate and position. Not only runners, but also sportsmen from other disciplines like alpine skiing or biking collect such information (Almer & Stelzel, 2002). Evaluation and visualization of this data can be useful in order to optimize the performance of athletes, see Saupe et al. (2007). While professional runners are often trained by experienced trainers, this is not necessarily the case for amateur or recreational runners, which could benefit from the supervision of experienced individuals as well. A web-based presentation can help to solve this problem.

The biggest challenge lies in the visualization of the GPS data. In order to do this, maps and terrain data are needed. A handy and free solution for these problems is Google Earth and its Keyhole Markup Language (Google, 2009). Besides maps generated from satellite images, KML files can be passed easily to other people through the Internet. Thus we developed a software tool which takes a GPS file in the National Marine Electronics Association (NMEA)-0183 format (Langley, 1995) and optionally a HRM file as input and creates a KML file output, which can be directly loaded into Google Earth and shared with others. We

specifically used information from Polar Inc. datasets (Polar, 2006), however, the use of our software is not restricted to these. The HRM data is not mandatory and can be left out. No HR representation will be possible then, but speed and the running path are still available. The tool is available online, please follow http://www5.informatik.uni-erlangen.de/our-team/eskofier-bjoern/projects/gervit. It has been used by several athletes in endurance sports who gave consistent positive feedback.
The intention was to create a program which is available to everybody and relies on freely available sources under the GPL. We also plan to provide an internet-based platform for sharing the created files in the future. In the rest of the paper, we will provide a short overview of previous work on the topic of visualizing running data, followed by further details on the features and different parts of our software.

Previous Work

Google Earth has quite heavily been employed in several recent scientific publications, see for example Lisle (2006) and Butler (2006). However, to our best of knowledge, it has never been employed for a personal sports data visualization tool.
Saupe et al. (2007) also used Google Earth for the visualization of physiological parameters as well as information about endurance sport training activities on a large high resolution display. In contrast to our work, the emphasis was on visualization on a large scale system, while our software aims at visualization on small scale systems which are available to everybody.
Eskofier et al. (2008) visualized data which was acquired with mobile phones including the perceived state of fatigue. The speed and running path of the runners were visualized in Google Earth. Our work takes this idea further and allows visualization of heart rate and other information, such as average speed and heart rate, running distance etc. as well. We also provide a software tool for sportsman for their everyday use.

Methods

Google Earth and KML

Our goal was a free interface which provides maps and terrain data, as well as the possibility to visualize data and exchange it with others. This is fulfilled by the Google Earth software and its Keyhole Markup Language (KML). Google Earth is a program which shows a 3D globe with maps and geographic data created from satellite images. Besides that, it also includes features such as visualizing several types of paths or to include point placemarks, which offer additional information for designated spots on the map. These can be placed directly onto the map, either by the user himself or via a KML file.
KML is an XML-based language used to describe geographic annotation as well as visualization for 2D maps or 3D earth browsers, see Google (2009). For our implementation we especially made use of placemarks, one of the most widely used features of Google Earth. These allow us to place information points on the map, as well as to visualize paths covered by an athlete. In order to create such an information point, the simplest type of placemarks is used, which can be created with the following code:

```
<kml xmlns="http://earth.google.com/kml/2.1">
      <Placemark>
            <name>Waypoint 1</name>
            <description> <![CDATA[
                  actual speed: 10.646 km/h <br>
                  average speed: 10.524 km/h <br>
```

```
                  actual heart rate: 167 bpm <br>
                  average heart rate: 165.39 bpm <br>
                  distance: 1.4018km <br>
                  altitude: 56.7 m <br>
                  elapsed time: 00:10:59]]>
            </description>
            <Point>
                  <coordinates>
                        -122.685031666667,45.5544,10.6463542857143
                  </coordinates>
            </Point>
      </Placemark>
</kml>
```

These placemarks are defined by the <Point> tag which contains the coordinates. Additional information like the name of the placemark and a description can be specified in the associated tags. They appear as yellow pins on the map and are labeled as "waypoint" in our implementation and placed along the running path. When the user clicks on them, a message box appears and offers additional information such as average speed and heart rate, actual speed and heart rate, distance, altitude and elapsed time. Moreover, there is also a "start" placemark at the beginning of the running path containing the date, duration, weight of the runner, maximum and resting heart rate, vo2max and a description. A "run summary" placemark is placed at the end of the running path and contains average speed and heart rate, total distance and elapsed time and altitude difference.
For visualizing the running path, the <Point> tag needs to be replaced by a <Line String> tag. The code for creating a path looks as follows:

```
<kml xmlns="http://earth.google.com/kml/2.1">
      <Placemark>
            <visibility>0</visibility>
            <Style>
                  <LineStyle>
                        <color>6f00ff00</color>
                        <width>4</width>
                  </LineStyle>
                  <PolyStyle>
                        <color>6f00ff00</color>
                  </PolyStyle>
            </Style>
            <name>part1</name>
            <LineString>
                  <extrude>1</extrude>
                  <tessellate>1</tessellate>
                  <altitudeMode>relativeToGround</altitudeMode>
                  <coordinates>
                        -122.693968333333,45.5600533333333,91.26
                        -122.693968333333,45.5600533333333,92.04
                        -122.693968333333,45.5600533333333,92.82
                        -122.693968333333,45.5600533333333,93.6
                  </coordinates>
            </LineString>
      </Placemark>
</kml>
```

The <extrude> tag extends the path line to the ground, creating a band, while the <tesselate> tag breaks it into smaller chunks. For specifying the color of the band the

`<color>` tag is used. Of great importance is the `<altitude mode>` tag that specifies how the height of the band is interpreted. We chose the option "relative to ground". If the height was interpreted relative to the sea level, it could happen that the band is below the terrain.

Implementation Details

Below we will give an overview of the features of our software, followed by the details of our implementation. A representation of our current GUI design is given in Figure 1.

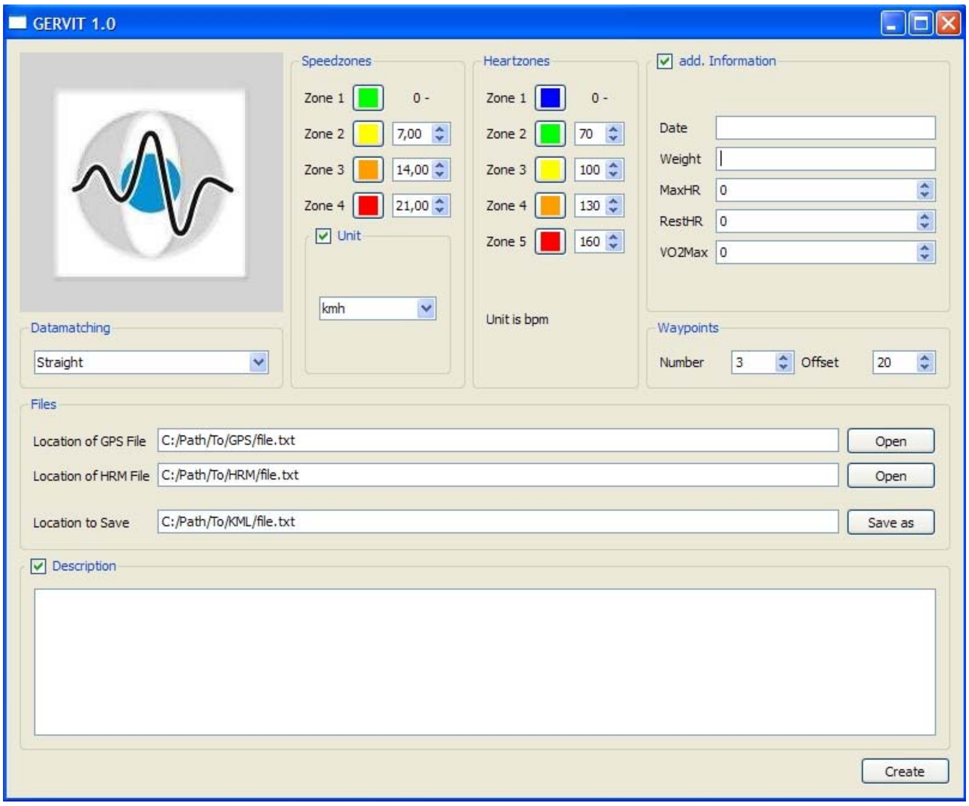

Figure 1. The GUI allows setting up the speed and heart rate zones. These are used for assigning colors, which can be set up as well, to speed and heart rate values when color coded. Other options are the type of data matching, the number of waypoints, the unit for speed values, additional information and a description.

Software Features

The individual speed and heart rate zones and their associated colors can be set in the GUI as can be seen in Figure 1. The visualized KML output is displayed in Figure 2. Here, it is also visible that different representation options can be selected in Google Earth's places view. The options are specified in the following.

- GPS only: This option shows the running path, the speed encoded as the height of the band and the actual speed zone color coded.
- HRM only: This option shows the running path, the heart rate encoded as the height of the band and the actual heart rate zone color coded.
- GPS+HRM colored: This option shows the running path, the speed encoded as the height of the band and the actual heart rate zone color coded.

- HRM+GPS colored: This option shows the running path, the heart rate encoded as the height of the band and the actual speed zone color coded.

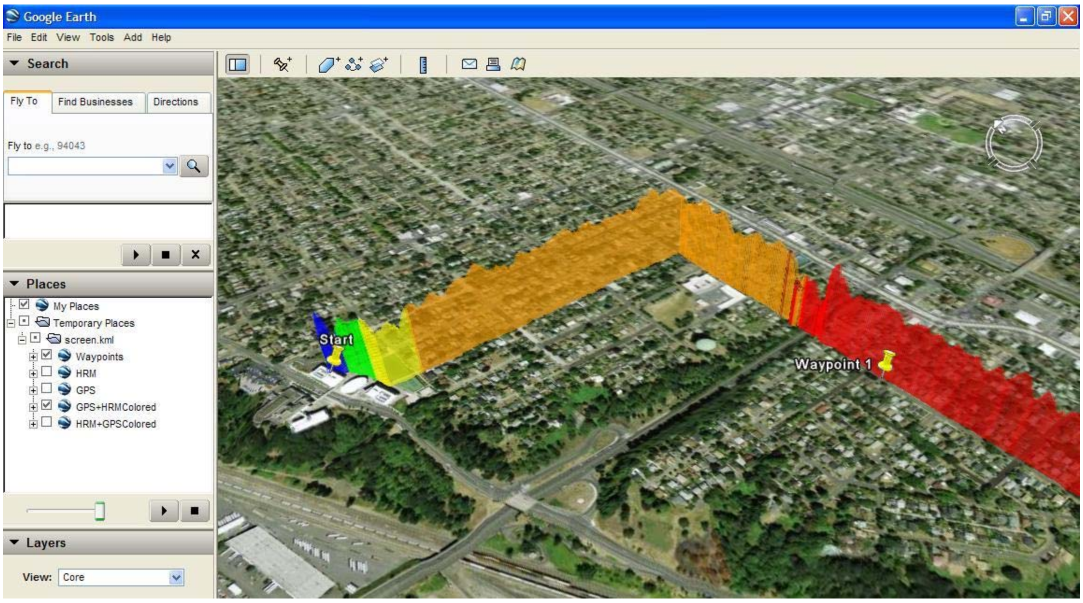

Figure 2. Visualization of running speed as height and heart rate as color of the band. Blue and green represent a low heart rate, yellow and orange a moderate heart rate and red means a high heart rate.

Another feature is the above-mentioned waypoints along the running path. By clicking on these, the user can check the actual and average speed, the actual and average heart rate, the running distance, the altitude and the elapsed time. The number of total waypoints can be set up in the GUI, see Figure1. Additionally two special waypoints are set up at the beginning and end of the running path. These are the "start" waypoint which displays the duration of session, weight of runner, maximum and resting heart rate, vo2max and a description. This information can be either entered in the GUI or is read from the HRM file. The other point is the "run summary" waypoint which contains average speed and heart rate, total distance, altitude difference and total elapsed time.

Software Structure

The software had to be able to read and parse GPS and HRM input files, process this data and create a KML output file. Our class hierarchy is presented in Figure 3. If the two parsers get adapted, other input file formats can be easily handled as well.

GPS and HRM Parser

Here, the data from the NMEA ASCII and the Polar HRM files, which both come in plain text, are parsed and stored. From the GPS file longitude, latitude, altitude, as well as actual speed and actual time are saved for each position. The stored data from the HRM file contains heart rate and additionally the maximum heart rate, actual speed, resting heart rate, vo2max, starting time, length of the session, weight of the runner and the date.

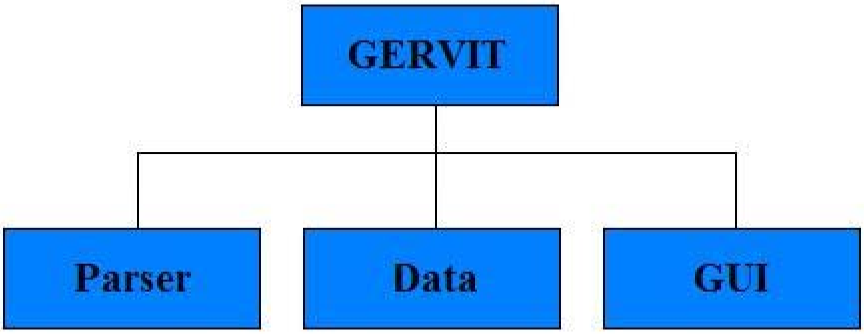

Figure 3. Class hierarchy of our software.

Data Class

After the GPS and heart rate data were parsed, they are passed to the data class for further processing. The available speed data first gets filtered with a mean filter. This is necessary because of a precision offset when recording GPS data, which results in leaps in the speed signal from the GPS device. The mean filter calculates an average value from the actual speed value and its neighbor values. Our neighborhood size was three. The formula for the filter, where f[x] describes the old and g[x] the new value, is:

$$g[x] = \frac{\sum_{i=-3}^{3} f[x+1]}{7}$$

In the case that HRM data is available as well and needs to get visualized we encounter two different problems. Usually the HRM device has a lower sampling rate than the GPS device, which results in having a different amount of values for the GPS and HRM data. Thus, after having filtered the GPS speed signal, we use a linear interpolation for the heart rate and speed signal from the HRM device in order to calculate the missing values. The other problem results from having two different devices, which don't necessarily start recording data at the same time. This results in a time offset between the heart rate and position signal, so that the proper mapping between coordinates and heart rate has to be found first. In case both devices have a correct time signal, the data can be easily matched by time information. This means GPS and HRM data with identical time are mapped to each other. If no correct time signal is available, we offer two additional solutions. The first is matching the data by cross-correlation, which offers a measure for the similarity of two signals. In this case, we use to our advantage that both the HRM device and the GPS device can record speed data. The correlation of two signals will be highest when both signals are aligned along the time axis. So in order to get the correct time offset, one would shift the HRM speed signal and calculate the cross correlation for every shift. The offset for the shift with the highest correlation value is selected as the correct one. In our case we made use of the normalized cross correlation which produces correlation values between zero and one. he formula, where $f_G[t]$ describes the discrete GPS speed signal and $f_H[t]$ the shifted HRM speed signal, is:

$$\frac{\sum_t f_G[t] \cdot f_H[t]}{\sqrt{\sum_t (f_G[t])^2} \sqrt{\sum_t (f_H[t])^2}}$$

The second option is a straight forward one. Here it is assumed that the GPS and HRM device started recording at the same time, so the data gets directly mapped to each other. We simply read this information from the respective files. In the next step the KML file is created. Here, additional information for the waypoints, such as average speed and heart rate, elapsed time and running distance is calculated. The different speed and heart rate zones are determined and finally the coordinates for the running path are written.

Graphical User Interface

Because of our requirement for a free software solution, the GUI was developed with Trolltech QT 4.4 under GNU Public License. QT is a cross-platform toolkit mainly used for developing GUIs and is available for several programming languages such as C++ and Java. Thus our software can be easily compiled on different platforms and operating systems such as Windows, Linux and MacOS X.

Conclusion and Further Work

We designed and implemented a tool for visualizing collected GPS and HRM data from running sessions directly in Google Earth, allowing to make use of multimedia in training evaluation. In order to realize this, we used the Keyhole Markup Language. Our software can visualize the running path and actual speed and heart rate, but also offers information on average speed and heart rate, running distance, elapsed time and others. It allows a web-based presentation and sharing running data with others over the Internet in order to improve and optimize training performance. The usage of free components makes our software available to everybody and besides the GPS and optional HRM devices nothing else is needed. The tool is available online on the following address: http://www5.informatik.uni-erlangen.de/our-team/eskofier-bjoern/projects/gervit. It has already been used by athletes, who reported that they especially liked the possibility to share the visualized information.

Our software also offers a high adaptability through its modular design. As a result it would be possible to adapt the parsers in order to visualize other file formats or other kinds of data, for example respiration rate.

Acknowledgements

The authors would like to thank the adidas innovation team (*ait.*) for the financial and technical support of this research.

References

Almer, A & Stelzl, H (2002). Multimedia Visualisation of Geo-Information for Tourism Regions based on Remote Sensing Data, *ISPRS Congress Ottawa*, 8-12 July 2002.

Butler, D (2006). Virtual Globes: The Web-Wide World. *Nature,* 439, 776-777.

Google (2009). KML Reference. http://code.google.com/apis/kml/documentation/kmlreference.html, Google Inc., June 2009.

Langley, R. B. (1995) NMEA 0183: AGPS Receiver Interface Standard, *GPS World*, 54–57.

Lisle, R. J. (2006). Google Earth: a new geological resource. *Geology Today*, (22):1, 29-32.

Papaioannou, E., Karpouzis, K., De Cuetos, P., Karagianis, V., Guillemot, H., Demiris, A. & Ioannidis, N. (2004). Melisa - a distributed multimedia system for multiplatform interactive sports content broadcasting. *Proceedings of the 30th Euromicro Conference*, Rennes, September 1-3, 2004, 222-229.

Polar (2006). Polar HRM2 File Format Description. *Technical Report available online from* www.polar.fi/files/Polar_HRM_file%20format.pdf, March 2006.

Saupe, D., Lüchtenberg, D., Röder, M. & Federolf, C. (2007). Analysis and Visualization of Space-Time Variant Parameters in Endurance Sport Training. *Proceedings of the 6th International Symposium on Computer Science in Sports (IACSS 2007)*, Calgary, June 3-6, 2007.

Tong, X., Liu, Q., Duan, L., Lu, H., Xu, C. & Tian, Q. (2005). A unified framework for semantic shot representation of sports video. *Proceedings of the 7th ACM SIGMM international workshop on Multimedia information retrieval table of contents*, New York, NY, November 10-11, 2005, 127-134.

Computer-Assisted Evaluation System for Volleyball Referee's Executive Judgment

Shang-Juh Kao[1], Chia-Chun Wu[1], Chin-Fa Chen[2]

[1]Department of Computer Science and Engineering,

[2]Office of Physical Education and Sport,

National Chung Hsing University

No. 250, Kuo Kuang Road, Taichung 402, Taiwan

Abstract

The purpose of this study is to develop an evaluation system for the executive judgment of volleyball referees. We present computer-assisted techniques to digitalize volleyball games, make judgments, and generate evaluation results. Through the help of senior volleyball referees, we first make classification to the recorded video volleyball games according to various foul situations and apply VideoStudio for test data editing. Through the interface of the computer-assisted evaluation system, times and correctness of judgments are recorded, and corresponding scores are prompted immediately after a test is over. The feasibility of this system is validated, as ten qualified volleyball referees of level A, B, and C are chosen for test participators. Statistical analyses of confidence and correlation are also included. According to the Pearson's coefficient analysis, the results from three co-related test sets show that the computer-assisted evaluation system for volleyball referee's executive judgment is highly reliable in both test-retest form and alternate form. In addition to serve for the volleyball referees evaluation, the system can also be used as a training purpose to improve the performance of judgment technique.

KEYWORDS: JUDGMENT, PERFORMANCE, PERCEPTUAL SKILL, VOLLEYBALL REFEREE, TRAINING

Introduction

Referees play an important role in the formal sport competition, sometimes it is more critical than what players and coaches play (Helsen & Bultynck, 2004; Johnston, 2008; Reilly & Gregson, 2006). Incorrect decisions made by referees would affect the final outcome of the game (Helsen, Gilis, & Weston, 2006, 2007). For example, during an American National Football League (NFL) post-season game in 2005, a decision-making error made by a referee led to the loss of New York Giants, the team that should have been the winner. Although NFL apologized to the Giants afterwards, the error had been made and the fact that the Giants had lost the game could not be changed. The importance of the referees can also be realized in the event that all referees in the Taiwanese national baseball league were dismissed and re-selected after the final game in 2004. Since wrong judgments can affect players' emotion and result in poor performance, conflicts, or even affect the result of a game, how to guarantee referees to make correct judgments deserves a further study.

Even though the judgment by referees is decisive in the competition, we are still using traditional ways in referee's training, such as orally explaining the rules of how to make a judgment, reviewing and discussing previous judgments, or practicing in unofficial games. Usually, an official referee license is given once a tester, or student referee, has passed a written test and some practice testing. Due to the huge population of trainees and limited time in training, it is seldom for a student referee to get a chance to practice in an unofficial game. Even when a student referee gets the chance, the game may not be strict because it is unofficial. Even worse, a student referee may not have chance to make a call of controversial judgment. Lacking of practicing experiences during the training course causes some licensed referees being incapable of making judgment in rigorous games later on. This ability in making correct judgments in live games is especially important for the referees in basketball, volleyball, handball, and soccer because they must make judgment decisions in a very short period of time. By traditional procedures, it is extremely difficult to qualify a referee to the level of well-trained.

Since the technique of sports vision has been adopted to improve perceptual motor performance for many years, as surveyed in Knudson & Kluka (1997), most researches were in eyerobics visual skill training (MacLeod, 1991; MacLeod & Hansen, 1989) and dynamic visual acuity (Ishigaki & Miyao, 1994; Long, 1994). According to previous studies (Adolphe, Vickers, & Laplante, 1997; Franks & Hanvey, 1997; Tayler, Burwitz, & Davids, 1994), the performances of athletes were improved after perceptual training with videos. In the study conducted by Tayler, et al. (1994), beginners of badminton players significantly improved in predicting the position of serving a ball after the visual training. Also, Adolphe, et al. (1997) reported that professional volleyball players performed better at visual concentration and ball-tracking ability after systematic training. Moreover, Franks & Hanvey (1997) presented that a goal keeper's ability in predicting the path of the ball during the penalty kick of a soccer game could be significantly increased after a visual training. There also existed literature on how vision is used in improving athletic performance, such as in the areas of baseball (Burroughs, 1984), basketball (Vickers, 1996), and soccer (Cohn & Chaplik, 1991). A overview of artificial intelligence (AI) applications such as speech recognition, computer vision and decision making has been introduced and presented by Bartlett (2004) (Nevill, Atkinson, & Hughes, 2008). Unfortunately, there is less study on how it can be applied to improve official judgment ability, especially for volleyball's referees.

In the earlier study (Chen, Liao, & Wang, 2005), it was also noticed that the Volleyball Referee Executive Analogue Computer System (VREACS) could efficiently record the judgments of referees' execution for a volleyball game to evaluate and improve referees' perceptual skills. Along with the rapid development of computer technology, it is possible to design a computer system to evaluate trainee's perceptual skill as well as to provide athletes an important reference. Through the visual training with analogized official games, in this study, we will investigate the perceptual skills, accuracy, and efficiency of volleyball referees' judgment. By pre-collection of video clips of official volleyball games, a computer-assisted system is developed and is used to evaluate and train volleyball referees. With this system, volleyball referees should be able to practice their judgment skills to reduce the possibility of "missing the call".

Methods

System Goal and System Design

The main goal of the Computer-Assisted Evaluation System (CAES) is to build a variety of condensed game realities for Volleyball Referee's Executive Judgment test to improve his judgment skill. The system can also play the role of qualification and evaluation platform for volleyball referees. There are three phases in the system design, as listed in the followings:

(1) Building the test data base which includes the collection of various official volleyball games.

(2) System development. The procedures of the development include video editing, scenarios classification, and testing.

(3) System evaluation, especially in the system's reliability, consistency, and efficiency.

Data Collection

The test data base is taken through the collection of video clips of on-going referees' judgments under various situations from 42 official volleyball games. The source video clips were selected from the games of 2006 Enterprise League Match-Men (from July 2006 to January 2007, Taiwan), 2006 University Federative Preliminary Contest-Men/Women (December 2006, Taiwan), 2006 Asian Challenge Cup-Men (August 2006, Taiwan), and World Grand Prix-Women (August 2006, Taiwan).

In a volleyball game, the first referee has to conduct his/her duty on top of a stage in a range of less than one metre square, and therefore, can only make limited visual and body movements. To avoid video cameras being blocked by the first referee during a play, we set up three video cameras (top, top-left, and top-right) behind the second referee. The second referee stands on the opposite side of the court facing to the first referee. All three cameras have been recording the ball games at the same time. The interest of this study is the judgment tasks of the first referee, who is responsible for the most and final decision. Accordingly, nine judgment situations are especially collected: failed serve, ace, shank, kill or spike, opponent's touch-out, opponent's valid block, failed attack-ball in the net, near-net players' faults, and regular hitting faults. The audio data of these videos were intentionally deleted to not affect the judgments making for the evaluation purpose.

System Development

The CAES for Volleyball Referee's Executive Judgment was implemented under Intel Pentium 4 2.0 GHz CPU with 512 MB RAM and Windows XP professional system. Microsoft Visual Basic .Net 2005 was used for coding the system with Microsoft SQL Server 2005 being adopted to save users' accounts as well as the historical data of training and testing. Both Ulead VideoStudio 11 and Ulead PhotoImpact 12 were dedicated for video tapes and image editing. The Microsoft Windows Media Player Decoder was used to decode digital video and audio streams into analog video signals. Once the data has been collected, the detailed test development proceeds as follows.

a) *Establishing the test database*: The content of video test data were reviewed and determined by three professional referees including two international referees and one level A national referee. Accordingly, these video clips were then classified and saved separately. The proper timings for whistling, proper calls, and correct gestures are pre-determined by the three professional referees.

b) *Determining the proper gestures*: When one enters into the system, as shown in Figure 1, he is able to choose the subject and category to be tested. The system provides all necessary gestures for end users to enter through mouse clicking. The first gesture determines which side should get the ball as shown in Figure 2. The clicking of the first gesture also indicates the time of whistling to be called.

After the first gesture, the play goes on and possible thirteen faulty occasions with corresponding gestures are also appeared on the screen (as shown in Figure 3). The system allows one of pictures to be clicked to indicate that an appropriate judgment has being made by the referee. Finally, the “OK” should be pressed to finish one round of testing. According to the previous study of Chen, et al. (2005), latency of two seconds is acceptable between the event occurrence and the time of whistling. Therefore, if the gesture clinking time is not within two seconds of the event occurring time, which has been pre-determined and recorded, the system will automatically treat it as a failed judgment or a misjudged call. Also, if the first gesture was wrong, the second gesture would be considered meaningless and be treated as an invalid action.

Figure 1. Main Page

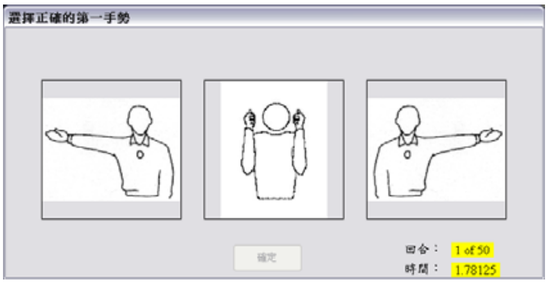

Figure 2. The First Gestures

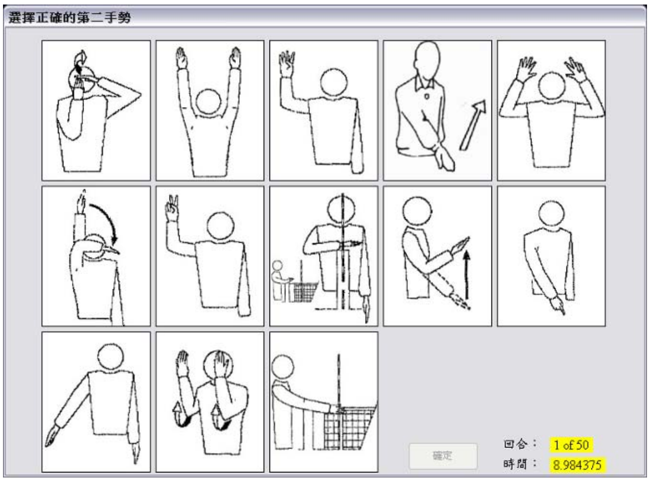

Figure 3. The Second Gestures

Table 1. The Timings of Making Gestures

Account	Level	Total Grades	Test Set	Type	Segment No.	Whistle Time	First Gesture	Second Gesture	Take Time01	Take Time02
0910563xxx	C	85	1	3	309	0.3	3	1	1.1	1.8
0910563xxx	C	85	1	11	113	0.5	3	11	0.8	1.6
0910563xxx	C	85	1	11	1,306	0.5	3	12	4.6	3.4
0910563xxx	C	85	1	7	988	0.7	1	11	0.8	0.9
0910563xxx	C	85	1	6	297	0.5	1	11	0.9	0.8
0910563xxx	C	85	1	1	16	0.5	1	13	1.2	1.6
0910563xxx	C	85	1	9	871	0.5	1	11	0.7	0.9
0910563xxx	C	85	1	9	280	0.7	3	11	0.6	0.5
0910563xxx	C	85	1	11	478	0.8	1	12	2.4	1.7
0910563xxx	C	85	1	1	703	0.7	1	13	0.7	4.1
0910563xxx	C	85	1	8	403	1.7	3	6	0.8	2.1
0910563xxx	C	85	1	3	1,388	0.7	1	10	0.8	1.5
0910563xxx	C	85	1	11	351	0.6	1	1	1.6	3.6
0910563xxx	C	85	1	6	239	0.8	1	12	1.5	3.7
0910563xxx	C	85	1	4	1,350	0.8	3	12	1.8	1.7
0910563xxx	C	85	1	5	1,957	1.0	3	11	0.8	2.3
0910563xxx	C	85	1	3	42	0.4	1	1	1.0	1.5
0910563xxx	C	85	1	7	1,921	0.6	3	11	0.8	1.2
0910563xxx	C	85	1	4	194	0.8	1	12	10	4.5
0910563xxx	C	85	1	10	1,141	0.8	3	11	0.9	2.8
0910563xxx	C	85	1	7	388	1.0	3	11	0.6	0.9
0910563xxx	C	85	1	3	1,370	0.4	3	1	0.8	3.0
0910563xxx	C	85	1	5	1,811	0.9	1	11	1.5	1.3
0910563xxx	C	85	1	10	983	1.3	1	13	0.5	2.8

c) *Completing a test*: To encounter various situations as in real games, different categories of occurrences are chosen according to different level of requirement. Among all categories, the seven previous mentioned categories appear most frequently. They are service error, attack out of the boundary, attack in the boundary, touch out, touch out after blocking, and attack blocked by the net. Since each set of volleyball game usually requires about fifty rounds of judgments, fifty rounds of plays are included in a complete test. Among the fifty rounds of test, forty of them are statically chosen and the rest 10 rounds are randomly chosen from the data base.

In each test, the timings of making both the first and the second gestures, or say having clicks on the pictures, are all recorded and compared to the correct timings. A sample timings table can be seen in Table 1. One point is given when the first gesture is correct. Two points were given when both the first and second gestures are correct. The complete scores of a set are one hundred. After the test, the tester has a chance to repeat those rounds that were failed in his previous judgment. Hence, the system can facilitate with the improvement in judgment skill in addition to play the role of testifying the qualification.

System Evaluation

Verification of the CAES for Volleyball Referee's Executive Judgment involves randomly choosing thirty national referees as testers. Each individual tester has three rounds of practicing test prior to the real test. Once each candidate referee has been familiar with the system, a complete test started and all timings of mouse-clicking were recorded for evaluation. Two evaluation concerns are of interest: consistency and efficiency.

a) *Consistency test*: Ten minutes after the first set of test, an identical test subject to different order is presented as the second test. The scores from these two sets are compared using Pearson's Product Moment Correlation (Derrick, Bates, & Dufek, 1994; Moore, 2006), and the resolutions are analyzed. Similarly, ten minutes after the second set of test, a third test of same level of difficulty but different to the first test is created for the tester. The scores of the third test are compared to that of the first test, again applying the method of Pearson's Product Moment Correlation.

b) *Efficiency test*: The efficiency of the CAES for Volleyball Referee's Executive Judgment is examined with known group difference method (Thomas & Nelson, 1990, 2001). The examination based on the first set of test. The differences between different levels of testers are then analyzed with Analysis of Variance (ANOVA) (Einspruch, 1998; Grimm, 1993; Runyon & Haber, 1988) using in Statistical Package for the Social Sciences (SPSS) (Nie, Hull, Jenkins, Steinbrenner, & Bent, 1975) Version 10.0 for Windows. Theoretically, if the system is effective, significant performance results should be derived among the referees of different levels.

Results and Discussion

To verify the feasibility of CAES for volleyball, we applied the system with 1896 rounds of volleyball games, which include different levels of competitions. Nine main call-situations are included for the purpose of judgment training: failed serve, ace, shank, kill or spike, opponent's touch-out, opponent's valid block, failed attack-ball in the net, near-net players' faults, and regular hitting faults. The time resolution is set to millisecond, that is, the referee has to make a call in a number of milliseconds whenever a certain event occurs.

With the method described above, the CAES provides testers/trainers a friendly on-line test and records the results in an Excel file for further analysis. Ten volleyball referees of each level (A, B, and C) were chosen as the test participators. That is to say there are thirty testers of various levels in total. Three sets of tests were presented to each participator. The first test set contains fifty rounds of play situations. The second set has the same fifty rounds as the first one, but were arranged in different order. The third set has different fifty rounds of plays but with the same level of choices. The results, which were originally stored as Excel files, were transferred into SPSS. The software of SPSS Version 10.0 was used for statistically analysis by adopting the method of ANOVA. By applying ANOVA, the outcomes are analyzed as sample test scores with different levels being shown in Table 2. In addition to the final test outcomes, the CAES can also provide the testers another chance to re-examine the misjudged rounds.

Table 2. The Basic Scores of Testers with Different Levels

Set	Level	N	Mean Scores	Standard Deviation	Standard Error	95% Confidence Interval		Min.	Max.
						Lower	Upper		
1st set	A	10	89.60	3.06	0.97	87.41	91.79	84.00	94.00
	B	10	87.60	3.31	1.05	85.23	89.97	85.00	94.00
	C	10	78.10	4.23	1.34	75.08	81.12	73.00	84.00
	Total	30	85.10	6.16	1.12	82.80	87.40	73.00	94.00
2nd set	A	10	89.70	2.75	0.87	87.73	91.67	85.00	94.00
	B	10	87.10	2.69	0.85	85.18	89.02	84.00	91.00
	C	10	82.00	3.27	1.03	79.66	84.34	76.00	88.00
	Total	30	86.27	4.30	0.78	84.66	87.87	76.00	94.00
3rd set	A	10	86.90	3.51	1.11	84.39	89.41	81.00	92.00
	B	10	83.80	2.25	0.71	82.19	85.41	81.00	88.00
	C	10	78.00	3.37	1.06	75.59	80.41	73.00	83.00
	Total	30	82.90	4.79	0.88	81.11	84.69	73.00	92.00

Table 3. ANOVA Analysis of Different Testers' Scores

Set	Source	SS	DF	MS	F
1	Between	755.00	2	377.50	29.66*
	Within	343.70	27	12.73	
	Total	1098.70	29		
2	Between	306.87	2	153.43	18.09*
	Within	229.00	27	8.48	
	Total	535.87	29		
3	Between	408.20	2	204.10	21.32*
	Within	258.50	27	9.57	
	Total	666.70	29		

*$p < 0.05$

From the analysis of

Table 3, we can conclude that significant differences exist between referees of different levels. The data are then analyzed using Scheffe's method (Boik, 1979; Shapiro, 2003) in SPSS Version 10.0 for Windows, and the results are shown in Table 4. The Scheffe's test shows that the cross comparisons among level A, B, and C. Level C is clearly apart from the other levels.

To see the resolution between the scores in different sets of tests, Pearson's Product Moment Correlation is applied, and the results are listed in

Table 5.According to the Pearson's coefficient analysis (as shown in

Table 5), the first set of test is significantly related to both the second and the third sets, showing that the CAES for Volleyball Referee's Executive Judgment is highly reliable in both test-retest and alternate form. According to
Table 3 and Table 4, the system demonstrates that the scores of testers in level A and B are significantly higher than those in level C; however, there are no significant difference between the testers of level A and B. The inability in distinguishing these two levels of testers is due to either the limited sample size in this test or inappropriate placing the testers in their levels. What is the exact cause is of interest in the future study. Nevertheless, the CAES can faithfully reflect the differences among different level of testers.

Table 4. The Group Comparisons of Scores with Different Levels of Testers

Set	Level		Mean Differences	Standard Error	Significance	95% Confidence Interval	
						Lower	Upper
1	A	B	2.00	1.60	0.46	-2.13	6.13
		C	11.50	1.60	0.00	7.37	15.63
	B	A	-2.00	1.60	0.46	-6.13	2.13
		C	9.50	1.60	0.00	5.37	13.63
	C	A	-11.50	1.60	0.00	-15.63	-7.37
		B	-9.50	1.60	0.00	-13.63	-5.37
2	A	B	2.60	1.30	0.18	-0.77	5.97
		C	7.70	1.30	0.00	4.36	11.07
	B	A	-2.60	1.30	0.18	-5.97	0.77
		C	5.10	1.30	0.00	1.73	8.47
	C	A	-7.70	1.30	0.00	-11.07	-4.33
		B	-5.10	1.30	0.00	-8.47	-1.73
3	A	B	3.10	1.38	0.12	-0.48	6.68
		C	8.90	1.38	0.00	5.32	12.48
	B	A	-3.10	1.38	0.12	-6.68	0.48
		C	5.80	1.38	0.00	2.22	9.38
	C	A	-8.90	1.38	0.00	-12.48	-5.32
		B	-5.80	1.38	0.00	-9.38	-2.22

Table 5. The Pearson Analysis of the Resolution among Different Test Sets

	N	1st set	2nd set	3rd set
1st set	30	1.00	0.64*	0.70*
2nd set	30	0.64*	1.00	
3rd set	30	0.70*		1.00

*$p < 0.05$

Conclusions

Concluding Remarks

This paper presents a Computer-Assisted Evaluation System (CAES) for volleyball referee's executive judgment. Other than in visual analysis of volleyball skills (Wilkinson, 1992; Zetou, Tzetzis, Vernadakis, & Kioumourtzoglou, 2002) as for volleyball players, this system can be used to evaluate the qualification of a volleyball referee as well as to enhance a volleyball referee's judgment skills. There are three phases in the system design: establishing a test data, system development, and system evaluation. By collecting the play videos of more than 40 official volleyball games, customizing into pieces of plays, which certain type of judgments is required in each piece, the test data is then classified into various level of judgment techniques according to the determination of three professional referees. To encounter various situations as in real games, different categories of occurrences are chosen accordingly in each fifty rounds of examinations. In each test, forty out of fifty rounds of plays are statically determined, the other ten rounds of plays are randomly chosen from the test data base. The timing information of every judgment is recorded and compared to the correct, pre-defined times. And, the results are analyzed using the Pearson's Product Moment Correlation. Both consistency and efficiency tests are also included in the paper. Finally, an application of the proposal system is presented.

Future Enhancements

We have demonstrated the proposed CAES and its feasibility of both the qualification test and the judgment skill enhancement for volleyball referees. However, at least three features can be made in the future to allow the system to be universally adopted: (1) extending the test data base to include the variety of video contents, that is to enrich the test database, (2) putting the system online, such as in a peer-to-peer network environment, so that the applicability of the system can be increased, and (3) applying the system for training oriented purpose could also be a feasible application. In addition to the usefulness for volleyball referees, the similar approach can be applied to other kinds of sports.

References

Adolphe, R. M., Vickers, J. N., & Laplante, G. (1997). The effects of training visual attention on gaze behavior and accuracy: A pilot study. *International Journal of Sport Vision*, *4*(1), 28-33.

Bartlett, R. M. (2004). Artificial intelligence in performance analysis. *International Journal of Performance Analysis in Sport*, *4*(2), 4–19.

Boik, R. J. (1979). The rationale of scheffe's method and the simultaneous test procedure. *Educational and Psychological Measurement*, *39*(1), 49-56.

Burroughs, W. A. (1984). Visual simulation training of baseball batters. *International Journal of Sport Psychology*, *15*(2), 117-126.

Chen, C. F., Liao, C. M., & Wang, Y. T. (2005). The development of evaluation and training system of sport referee's perceptual skills: example of volleyball referee. *Proceedings the Taipei University Sports Federation Conference on Physical Education and Sports*, pp. 467-473.Taiwan.

Cohn, T. E., & Chaplik, D. D. (1991). Visual training in soccer. *Perceptual and Motor Skills*, *72*(2), 1238.

Derrick, T. R., Bates, B. T., & Dufek, J. S. (1994). Evaluation of time-series data sets using the Pearson product-moment correlation coefficient. *Medicine & Science in Sports & Exercise*, *26*(7), 919-928.

Einspruch, E. L. (1998). *An introductory guide to SPSS for windows* (2nd ed.). Thousand Oaks, CA: Sage Publications, Inc.

Franks, I. M., & Hanvey, T. (1997). Cues for goalkeepers: High-tech methods used to measure penalty shot response. *Soccer Journal*, *42*, 30-33.

Grimm, L. G. (1993). *Statistical applications for the behavioral sciences*. New York: John Wiley & Sons.

Helsen, W., & Bultynck, J. B. (2004). Physical and perceptual-cognitive demands of top-class refereeing in association football. *Journal of Sports Sciences*, *22*(2), 179-189.

Helsen, W., Gilis, B., & Weston, M. (2006). Errors in judging "offside" in association football: Test of the optical error versus the perceptual flash-lag hypothesis. *Journal of Sports Sciences*, *24*(5), 521-528.

Helsen, W., Gilis, B., & Weston, M. (2007). Helsen, Gilis, and Weston (2006) do not err in questioning the optical error hypothesis as the only major account for explaining offside decision-making errors. *Journal of Sports Sciences*, *25*(9), 991-994.

Ishigaki, H., & Miyao, M. (1994). Implications for dynamic visual acuity with changes in age and sex. *Perceptual and Motor Skills*, *78*, 363-369.

Johnston, R. (2008). On referee bias, crowd size, and home advantage in the English soccer Premiership. *Journal of Sports Sciences*, *26*(6), 563-568.

Knudson, D., & Kluka, D. A. (1997). The impact of vision and vision training on sport performance. *Journal of Physical Education, Recreation and Dance*, *68*(4), 17-24.

Long, G. M. (1994). Exercises for training vision and dynamic visual acuity among college students. *Perceptual and Motor Skills*, *78*(1), 1049-1050.

MacLeod, B. (1991). Effects of eyerobics visual skills training on selected performance measures of female varsity soccer players. *Perceptual and Motor Skills*, *72*, 863-866.

MacLeod, B., & Hansen, E. (1989). Effects of the eyerobics visual skills training program on static balance performance of male and female subjects. *Perceptual and Motor Skills*, *69*(2), 1123-1126.

Moore, D. S. (2006). *Basic practice of statistics* (4th ed.). New York: W.H. Freeman.

Nevill, A., Atkinson, G., & Hughes, M. (2008). Twenty-five years of sport performance research in the Journal of Sports Sciences. *Journal of Sports Sciences*, *26*(4), 413-426.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. H. (1975). *Statistical package for the social sciences* (2nd ed.). New York: McGraw-Hill.

Reilly, T., & Gregson, W. (2006). Special populations: The referee and assistant referee. *Journal of Sports Sciences*, *24*(7), 795-801.

Runyon, R. P., & Haber, A. (1988). *Fundamentals of behavioral statistics* (6th ed.). New York: Addison-Wesley.

Shapiro, A. (2003). Scheffe's method for constructing simultaneous confidence intervals subject to cone constraints. *Statistics and Probability Letters*, *64*(4), 403-406.

Tayler, M. A., Burwitz, L., & Davids, K. (1994). Coaching perceptual strategy in badminton. *Journal of Sports Sciences*, *12*, 213.

Thomas, J. R., & Nelson, J. K. (1990). *Research methods in physical activity* (2nd ed., pp. 297–320; 341–363). Champaign, Illinois: Human Kinetics.

Thomas, J. R., & Nelson, J. K. (2001). *Research methods in physical activity* (4th ed., pp. 185). Champaign, Illinois: Human Kinetics.

Vickers, J. N. (1996). Control of visual attention during the basketball free throw. *American Journal of Sports Medicine*, *24*(6), S93-S97.

Wilkinson, S. (1992). Effects of training in visual discrimination after one year: Visual analysis of volleyball skills. *Perceptual and Motor Skills*, *75*, 19-24.

Zetou, E., Tzetzis, G., Vernadakis, N., & Kioumourtzoglou, E. (2002). Modeling in learning two volleyball skills. *Perceptual and Motor Skills*, *94*(2), 1131-1142.

The Role of Databases in Sport Science: Current Practice and Future Potential

Jenny Vincent, Pro Stergiou, Larry Katz

Faculty of Kinesiology, The University of Calgary, Canada

Abstract

Databases with the tools for capture, storage, management, retrieval, integration, analysis, interpretation, reporting, and dissemination have the potential to be the single most powerful tools in sport science. Knowing how to collect, store, access, retrieve, and integrate information is critical to effective performance analysis and decision-making. Databases should form the underlying foundation of most other tools used in sport science as they provide the structure and access to the information that is the catalyst for most other applications. As the development of resources become more integrated, the value of databases increases and their role in system design becomes essential. This includes the ability to retrieve previous data for comparison with new performance and the use of data to highlight issues for deliberation. Databases are also useful as multimedia repositories of sports information. Sport involves human movement. This can be quantified with numerical data, graphics, and audio/video recordings. Multimedia resources in this context are exceptionally valuable, especially if the information can be accessed via creative interfaces that provide timely and efficient information that is tailored to each individual's unique requirements. This paper discusses the nature of databases, the role of databases in sport science including design considerations, issues of integration, examples, and future potential.

Understanding Databases

A database, by definition, is an integrated collection of logically related records or files consolidated into a common pool that provides data for one or multiple uses. At the simplest level, creation of a spreadsheet in Microsoft Excel constitutes a database. At a more advanced level, databases can be complex and customized systems that capture, store, retrieve, interpret, report, and disseminate information.

Databases form the underlying foundation of most other tools used in sport science as they provide structure and accessibility to data that is the catalyst for many other applications. This structure and access allows the user to retrieve previous data for comparison with new performance information or allows him or her to use data that highlights issues for deliberation. Sport involves human movement that can be quantified with numerical data, graphics, and audio/video recordings. Multimedia resources in this context are exceptionally valuable, especially, if the information is accessed through creative interfaces that provide timely and efficient information tailored to an individual's unique requirements. As the development of these information resources becomes more integrated, the value of databases increase.

A database is essentially a collection of interrelated data of different types. Databases for the storage and retrieval of information have been very popular in the field of Epidemiology (the study of the factors that affect health and wellness in a population). An example of such a database is the Injury Surveillance System (ISS) that is maintained by the National Collegiate Athletic Association in the United States (NCAA) (Dick, Agel, & Marshall, 2007) and used for tracking of injury data in college sports. This database collects injury data from NCAA institutions in a wide variety of sports and has been in operation since 1982. The type of information collected and stored includes data such as injuries sustained by athletes, the school the athlete attends, and the sport in which they participate. By capturing this data year after year, decision makers use this information to look at information aggregated by institution, sport, or type of injury. It has been used in tracking injury rates in new sports, on different playing surfaces, and to follow the effects of rule changes in various sports.

Database Design

Databases can be used to store all different data types including continuous or categorical data. As related to sport performance, this data can include performance data from competitions or any other type of data that a coach collects during the pre-season, competitions, or post season. Furthermore, this data can include specific biomechanical, physiological, or medical data related to each athlete. Proper planning, creation, and maintenance of a database are required to create and utilize an "active" tool that can benefit all participants. Complex, multiuser databases usually require a database expert to create and maintain.

Interpreting information from a database is only as good as the initial information that goes into it. This information must be accurate with checks in place to ensure the data is valid. This quality assurance is critical because information frequently comes from a variety of sources. Data can be input manually or automatically. Manual entry requires more time, but ultimately brings a level of confidence that the data is more accurate if the proper checks and balances are in place. Scripts (programs) can be written by a computer programmer that automatically read data, for example, from a website that contains performance data. However, using these programs still requires manual checking to ensure the accuracy of the information that has been automatically input.

A structured and thoughtful database design is essential to the success of a database. One of the key principals of good database design is avoiding unnecessary redundancy. In the Injury Surveillance System (ISS), this principal might be highlighted by the following example. All the data about an institution (for instance the name, abbreviation, and location) is not repeated for each of the 1.4 million injuries collected in the database. The institution information is stored once and then for each injury a 'referential integrity constraint' is created to link to the institution information. By avoiding the duplication of the institution information the institution information is kept consistent allowing successful retrieval and aggregation of injuries across institutions. Therefore any updates to the institution information need only be made in one location.

Another principal of database design is ensuring the uniqueness of information. For example in the ISS it is critical that every injury is recorded only once. An injury may be uniquely identified by the date and time it took place, the athlete's institution, the sport, and the body part injured. These pieces of information make up what is called a key for the injury data (note: since athlete identification information is not recorded in the ISS it can not be used in the key). The key is used in the database to enforce uniqueness. In the injury database example the key is made up of many pieces of information, a database designer will often substitute this for a single computer-generated identification number (e.g. injury id) called a surrogate key.

In any database it is very important to impose validation conditions on the data as it is entered. This is often accomplished using domain integrity, that is, by providing a list of possible choices. Domain integrity ensures consistency of the data maximizing the ability to search and aggregate the information. This type of validation condition is used heavily in the ISS and has attributed to the system's success and longevity. For example, in the collection of data on the principal body part injured a list of body parts is provided. Without this domain integrity a broken toe might be recorded under the principal body part foot, or toe, or even metatarsal making it impossible to access accurate injury rate information on this type of injury. Another type of validation condition is the designation of certain pieces of data as required (i.e. not null). Null values are a representation of missing or unknown data. A careful decision must be made for each data field that will be collected, balancing what information is needed for retrieval and aggregation with what is readily available at data entry time. A final type of validation condition is ensuring the type of data (numeric, character, date, image) and the maximum size, are appropriate for the data that will be collected in each data field. This is imperative to ensure correct sorting, grouping, and retrieval. The thoughtful consideration of all these database design elements will ensure the integrity, usability, and longevity of the database.
The principals of good database design noted above are derived from the process of normalization created by E. F. Codd, who was the first to describe the concept of relational databases (Codd, 1970). Although other database models exist (including hierarchical, network, and object) the relational database model is widely used, is very mature and the database management systems which implement it are fast and reliable. Relational databases are based on the concept of tables with columns and rows. Tables can be related to other tables through keys. Information is retrieved from tables using the Structured Query Language (SQL). In this paper all references to databases are specific to the relational database model.

Database Management Systems

A database management system (DBMS) is a set of computer programs that control the creation, maintenance, and use of a database (Wikipedia, 2009a). That is, it allows the designer to create tables with columns of a specific data type and size, and relationships between the tables; it allows for the insertion, update, and deletion of rows in the tables; and it allows for the retrieval of information from the tables using Structured Query Language (SQL). SQL is a database computer language designed for managing data. Two of the most commonly used desktop DBMS' are Microsoft Access and Apple's Filemaker Pro (cross platform). In multi-user DBMS environments, Microsoft SQL Server is a popular choice. The industrial strength Oracle DBMS is often the preference for large organizations and data warehousing applications. The freely available MySQL is another popular alternative amongst database developers.

Content Management Systems

A content management system (CMS) is a computer application used to manage workflow needed to collaboratively create, edit, review, index, search, publish and archive various kinds of digital media and electronic text (Wikipedia, 2009b). Most systems use a database to store content, metadata, and artifacts used by the system. The database will often manage security of the content allowing users to view, insert, update, or delete content as specified. The most common type of content management system is a Web CMS, which manages web site content. It will provide authoring tools to allow users with no knowledge of web site creation to maintain parts of the site. One example of a Web CMS in the sports world is a product used by hundreds of university athletic departments across North America called SideArm (Internet Consulting Services, 2009). It is used to provide information about a university's varsity sports teams allowing the team coach to easily update information as the season progresses.

Sport Databases

Databases are used in sport administration, in all of the sub-disciplines of sport science and in sport epidemiology). Examples from some of these areas are provided below together with a discussion of the key elements of good database design.

Sport Administration

An interesting possibility for databases within the sports administration stream is in the scoring of athletes competing in judged sports. A recent paper describes a gymnastics scoring system, which separates the video performance of an athlete into a set of component parts or skills and compares each skill to a database of skills and the scoring of the skill as defined in the rule book (Jeongeun Shin & Ozawa, 2008). In the rulebook for the gymnastics horizontal bar discipline, 141 skills are described and key poses for each skill displayed along with skill classification and difficulty level. In the skills database a silhouette image of a 3-D human body is augmented with the posture parameters from the key poses. Using image processing the athlete performance is then compared with the 3-D key positions from the database to provide scores for each skill and then the whole routine.

Epidemiology of Sport Injury

As noted above, epidemiology is a data intensive field of sport science. At the 2008 Summer Olympic Games in Beijing an injury surveillance system was used not only to track epidemiological information, but also to direct injury prevention programs (Junge et al., 2008). The system was shown to effectively track the large multi-sport events at the Summer Olympic Games. The researchers provided all definitions, forms, and analysis details so that the injury surveillance system could be implemented at other sports events, thus growing the overall base of standardized data. This type of work allows for comparison across various sporting competitions as well as supporting longitudinal studies nationally and internationally.

Databases in Performance Analysis

One of the most widely available types of sport performance data is competition result information. The National Performance Database is one example of a recently created database that manages performance competition data. This database was created by the Sport Technology Research Laboratory (University of Calgary) for Own The Podium 2010 (OTP) podiumcanada.ucalgary.ca:8080/OTP). OTP 2010 "… is a national sport technical initiative designed to help Canada's winter athletes win the most number of medals at the 2010 Olympic Winter Games in Vancouver, and to place in the top three nations (gold medal count) at the 2010 Paralympic Winter Games" (Own the Podium, 2009).
The National Performance Database includes competition results and performance data for all 152 sport disciplines participating in the 2010 Winter Olympic and Paralympic Games. This database includes competition results from all the World Cups, World Championships, and Olympic Winter Games since 1997 (Vincent & Childs, 2009). Using a web-based interface, administrators and the public can search for results by the various competitions or by an athlete's name. Also, the National Performance Database provides a Canadian medal count and a competition calendar. Decision-makers at OTP have a number of other interfaces that allow them to analyze Canadian athlete performance, compare Canadian results, and even view

the medal counts of the top 10 countries. Moreover, the database offers a comparison of individual athlete performances to their benchmarks, as well as, providing the OTP decision-makers with information on which countries are the strongest in various sports and disciplines. All data is live and new competitions are loaded as they occur on a weekly basis. The data is presented in both a tabular format and as well as graphs.
Figure 1 shows a sample extraction from the Own the Podium National Performance Database for short track speed skating.

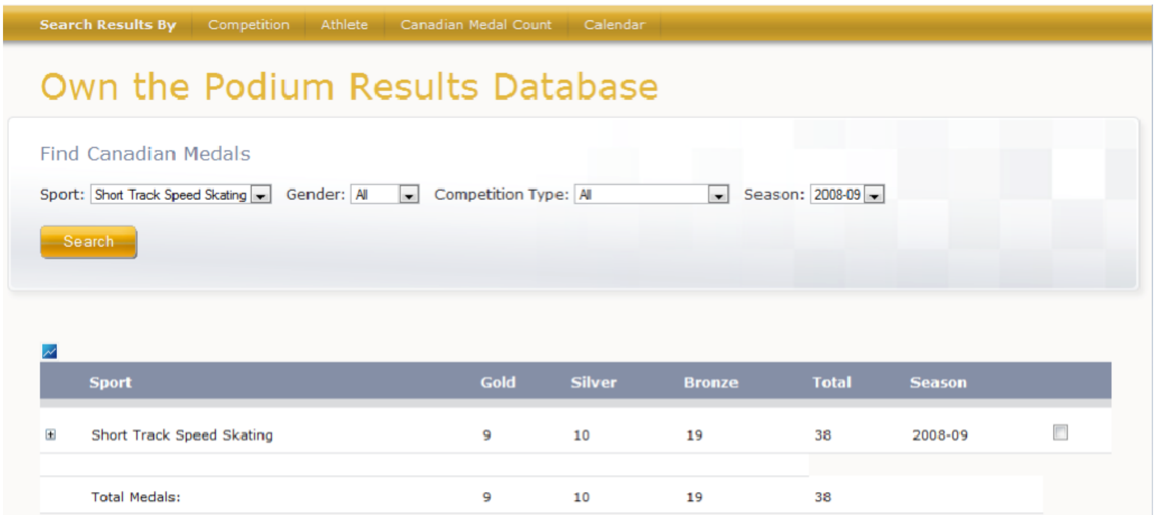

Figure 1. A query of the National Performance Database that shows information related to short track speed skating.

There were a total of 38 medals won by Canadians in 2008-2009: 9 gold, 10 silver, and 19 bronze medals. Users can drill down further by clicking on the "+" sign to the right of the word Short Track Speed Skating to see in which disciplines and genders Canadian athletes won the medals.
Another area where performance analysis databases are effective is related to physiological and/or biomechanical data. If the data is properly organized, then online database tools such as SSENTIF Sport (http://www.ssentif.co.uk). have the capabilities to collect, manage, monitor and analyze the data providing valuable tools to improve athlete performance with an interface that can be used by those without database expertise. Tables and graphs can be generated to track athlete progress and development. In the following example, biomechanical data related to how fast an athlete can throw a ball was collected including ball speed and body positions of the athlete as they performed the skill. All data was entered into the database and graphs were generated on how the athletes performed. Figures 2 below shows example graphs from two athletes, one experienced and one novice.

Figure 2. Data graphs generated from SSENTIF software. The image on top is a developmental athlete and the image on the bottom is an elite athlete.

Other programs such as Qlickview (http://www.qlikview.com) provide tools for the novice database users so that they have the opportunity to explore the data, generate visual representations of the information, and hopefully improve decision making using evidence based considerations.

Video Databases

In the analysis of sport performance, video remains a valuable element to success. Video analysis is used by both the coach and the athlete to improve skill development, reduce injuries, provide tactical information, and support notational analysis. The automatic indexing and retrieval of specific video content from full competition film is an ongoing research area. The SportsBR system is able to take in a video from the massive number of sporting competition recordings available, automatically select event-based clips from the video, and index and store them into a database for later browsing and retrieval (Hua-Yong Liu, 2005).

The system automatically selects events by looking at audio, visual and caption text information. It then indexes the clips using speech recognition technology and the extracted caption text. This integrated use of different methods to select clips from a larger video, termed multimodal analysis, is a hot spot in video analysis research. For this research they used soccer matches, selecting the events: penalty kick, free kick next to the goal box, and corner kick. Testing against a set of World Cup soccer matches, they were able to achieve an accuracy rate of 91.3% on clip selection and 97.5% on the subsequent recall of the clip. The keywords used to query the database were derived using the broadcast commentary, via speech recognition software, and from extracts of the captions appearing during the clip. This information was normalized against a sport video vocabulary database created by the researchers from observations of matches and provides some domain integrity to improve retrieval of the clips.

The research into multimodal video retrieval is not only producing databases of video clips with keywords; the research is helping to create storehouses of sports performance data. For example, the GNAVI system is focusing on the capture of the detailed information from the caption text of golf match video thus creating a collection of hole-by-hole player information (Cheolkon Jung & Joongkyu Kim, 2008). These types of systems, along with existing competition results, notational, and physiological datastores provide a mass of data available for analysis. The next step is developing the ability to analyze and extract useful information from this data.

Traditional data analysis techniques, including aggregation queries and statistical analysis, are used and researched frequently in sports. These techniques are hypothesis-driven, that is, they answer a specific question. Data mining is the process of extracting patterns from data; it is a knowledge discovery approach (Wikipedia, 2009c). Shoeman et al (2006) attempt to answer the research question of "whether proven business data mining techniques can be applied to sports games in order to discover hidden knowledge". They conducted a field study in which they data mine South African National Cricket team match information using the industry standard 'Cross-Industry Standard Process for Data Mining' to implement link analysis on the database. The researchers then examine if this type of analysis is suitable for the sports world by presenting the findings to a set of stakeholders of the South African National Cricket team. They found that business data mining techniques and applications do have the potential to discover new information, however, the transformation of match data into a usable format can be problematic. They go on to suggest that the data mining of multimedia databases, containing clips and data automatically retrieved from competition video, might be successful in getting around some of the problems.

More recently another group of researchers also mined the highly available and rapidly growing stores of cricket match data to help coaches determine game tactics and improve coaching strategies based on correlations of play patterns (UmaMaheswari & Rajaram, 2009). These researchers used an object-relational database model to store ball-shots and the varying outcomes. From the database they were able to analyze both the performance of a single player and find patterns for an entire team.

Video Databases in Coaching

Coaches and sport associations habitually collect data in various formats in a quest to improve athletes' performances; and, this data not only concentrates on their own athletes, but also on competitors as well. Managing all the data appears to be a very laborious task that can consume substantial amounts of time. Finding information to facilitate training and decision-making is especially difficult. Learning how to store meaningful information effectively is critical to efficient retrieval and analysis of data. This section discusses databases that

specialize in the storage, management, and retrieval of video data in coaching.
The Digital Scouting system offers wrestling coaches and athletes the opportunity to view and share video over the web (Digital Scouting, 2009). The application allows the user to easily organize and search for video, audio, or other digital resources using unlimited categories, events, or divisions. A permission tool then allows the user to provide a wide variety of access levels including view only, upload, or edit. Video shared through the Digital Scouting site include basic and advanced instructional material, examples of a wide variety of techniques, as well as competition video from the International, Domestic, and University scene. This application provides an example of how databases are used to manage multimedia content and security access to groups of users and various content. A sample of the Digital Scouting web interface is provided in Figure 3 and Figure 4 below.

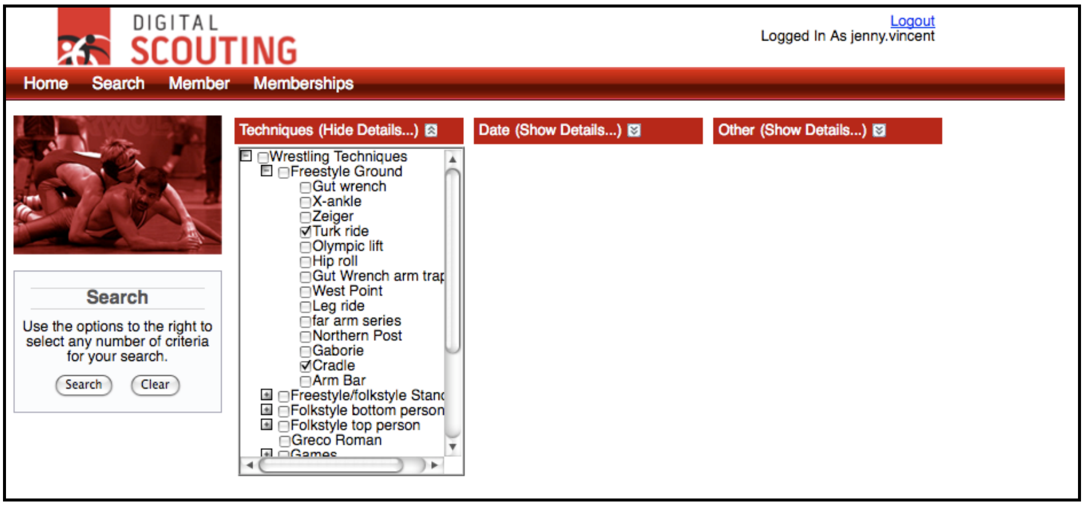

Figure 3. Wrestling video database main screen (Digital Scouting)

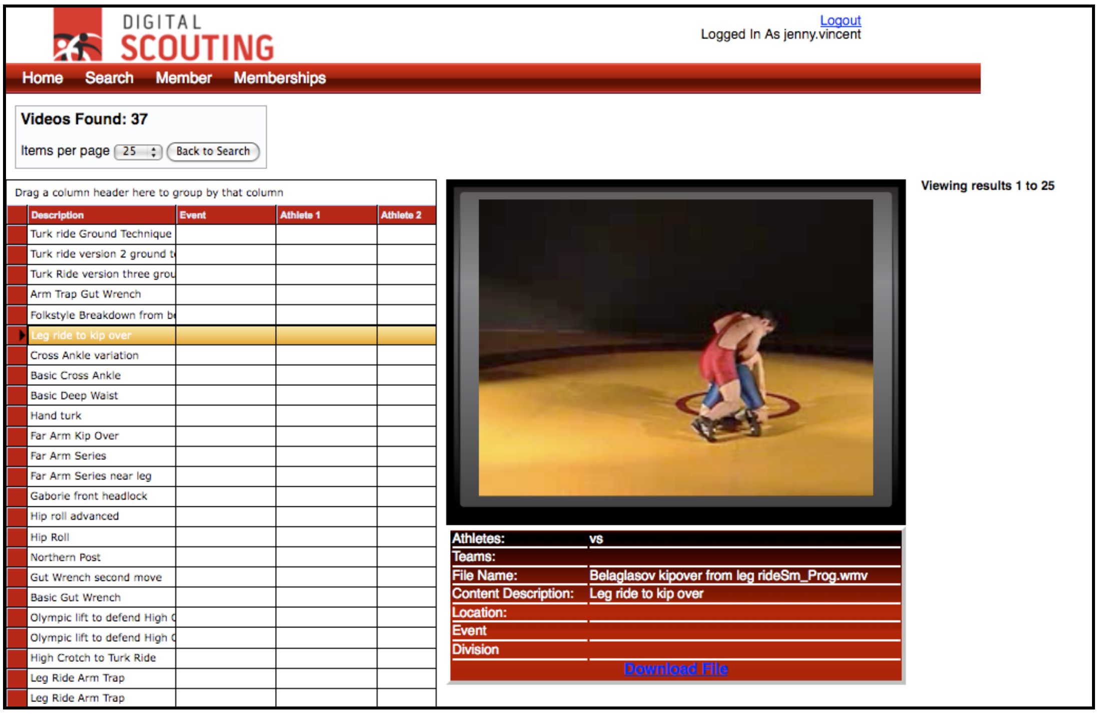

Figure 4. Wrestling video database – Video search display (Digital Scouting)

A similar web-based video sharing application called the Video Report Card has been developed by the University of Calgary's Sport Technology Research Laboratory. This system facilitates the sharing of both video and subsequent comments on the video within the database structure. This is particularly useful for athletes who may have a coach who travels frequently or who lives in another city. An example of a commented video is provided below in Figure 5.

Figure 5. Video Report Card - coach view (Sport Technology Research Laboratory)

An innovative virtual reality system has been built which allows coaches and athletes to see a virtual simulation of a complex series of gymnastic movements (Simin Li & Jinhai Sun, 2009). The coach can design or change a series of movements, looking to maximize the difficulty levels and then watch the simulation to see what the series might look like. A database is used to store the vast amount of data needed to produce the simulations. The database has 3 components: a movement storehouse, a three-dimensional model of the human body, and a movement arrangement storehouse. When realized within the three-dimensional virtual reality environment coaches and athletes can watch the series from all angles. They can then use this knowledge to return to the gym to train more effectively and safely. This example highlights the power of a database to store extremely large amounts of data and then access that data in a very efficient manner. Today's average databases can easily search and aggregate terabytes of data.

The Web is an ever-increasing dynamic tool with regards to teaching. Sports skills in this setting are no exception. Another application coming out of the University of Calgary's Sport Technology Research Lab provides such an e-learning environment for students who are learning a new skill. The Virtual Coach has the ability to compare an athlete's performance with that of a professional model, thus permitting the athlete to see the subtleties of the skill sets while simultaneously seeing both the model and themselves. Visual feedback, written evaluations and professional demonstrations of the skill sets are used to comprehensively enhance the learning process. The tool adds a common reference point between the athletes and the coaches whereby details of the sports skill can be discussed and confusion eliminated. Having sophisticated resources and training techniques available on the web on a continuous basis can have great benefits to the student and to the teacher. Sample screens of the Virtual Coach interface are shown in Figure 6 and 7 below.

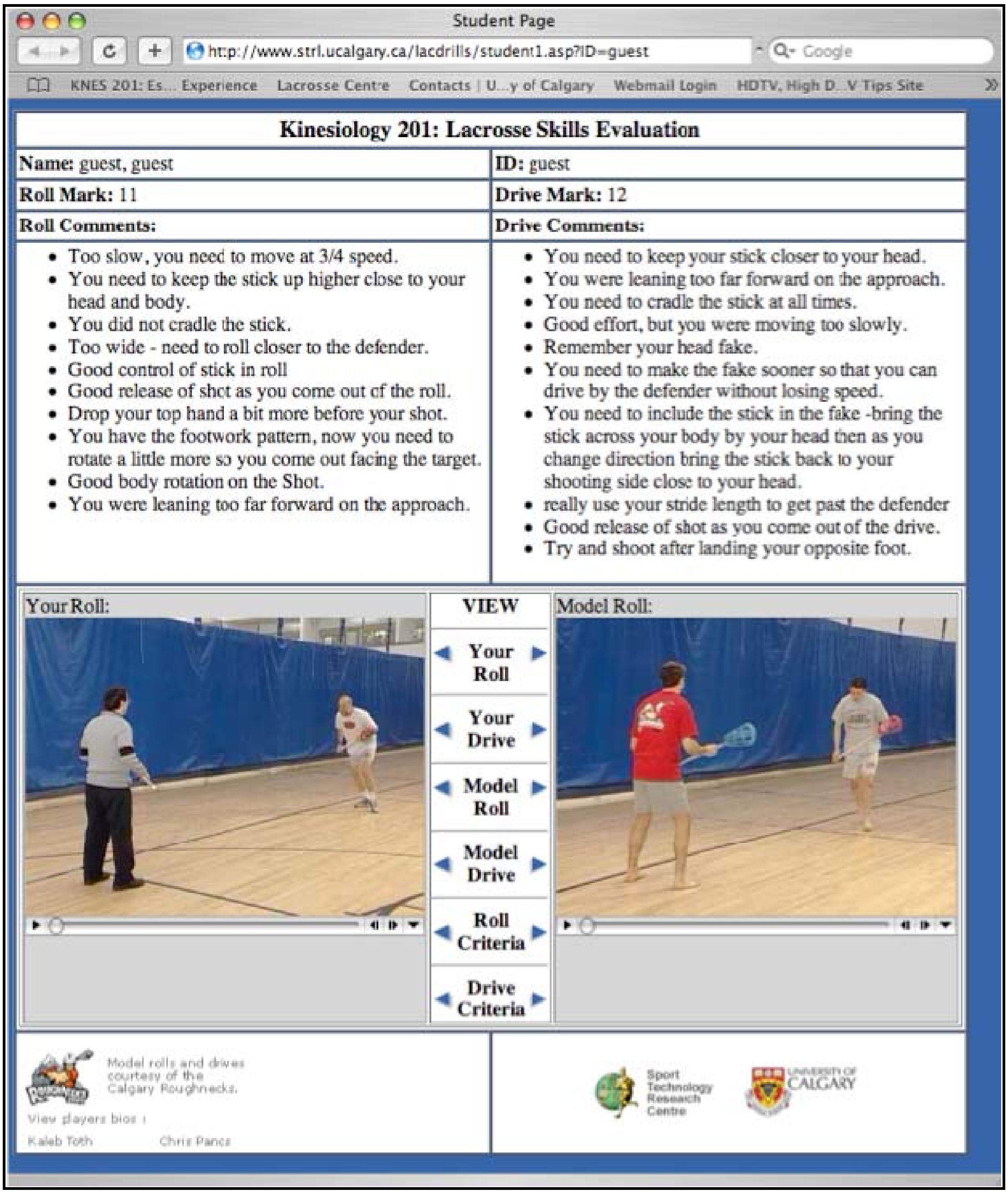

Figure 6. Virtual Coach – Student views self performance and compares to professional performance (Sport Technology Research Laboratory)

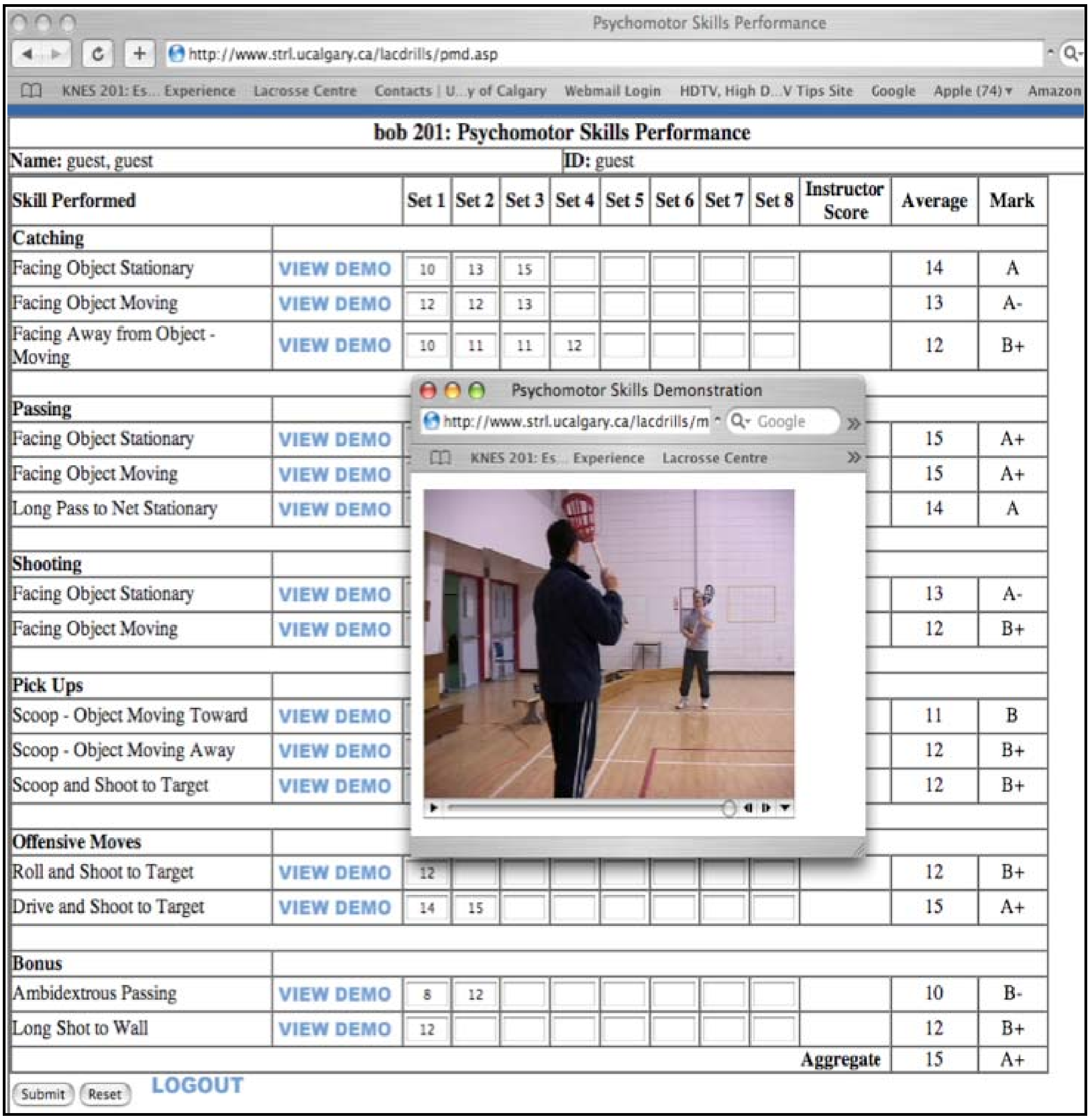

Figure 7. Virtual Coach with demonstration video of skill and online database for self recording of skill performance.

Online video databases provide individuals with the abilities to store, manage and share video content online with other people from around the world. The manager of the database can determine how to organize the videos, the content that will be available for viewing and can manage subscribers to view the content.

Dartfish, best know for its sport video analysis software, has developed Dart Fish TV an online portal which provides video database capabilities (organizes its databases in items called Channels (http://www.dartfish.tv). The image below (Figure 8) shows the front page of a channel called in this example, the Pro Stergiou Demo Channel. The content of what was created and the analysis performed on the user's desktop is up loaded in a simple manner from the Dartfish interface. An image of the process of uploading videos to the database is shown in Figure 9 below. As long as the user has a connection to the Internet, the process of uploading requires just a click of a mouse button. Additionally, content other than videos can be attached to the upload. This might include some performance data that was collected and compiled in a spreadsheet that the user would like to share with interested individuals. Any type of file can be attached to a video, similar to an attachment in an email.

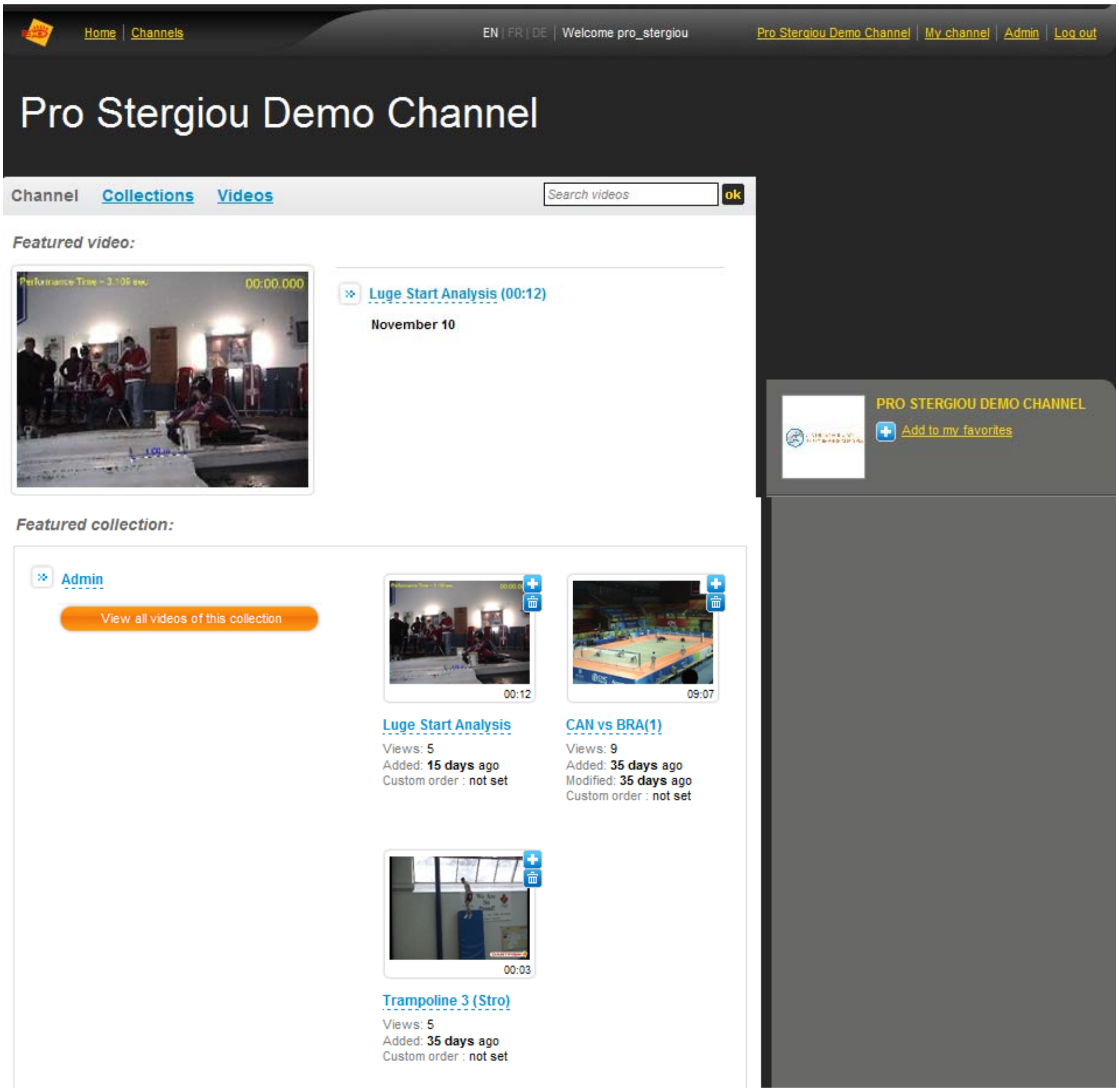

Figure 8. Image of front page of Dartfish TV website that is used to store, manage and share video content.

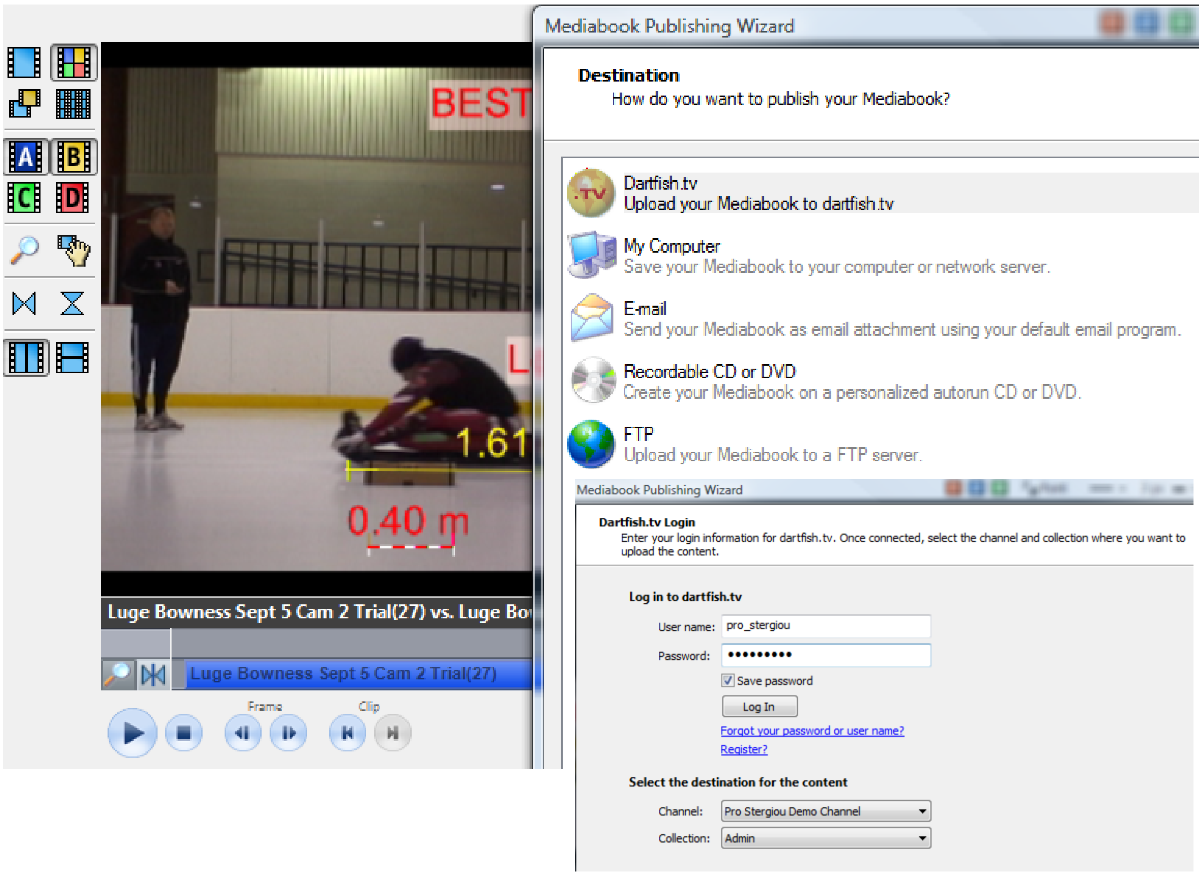

Figure 9. Process to upload video content to Dartfish TV online video database from Dartfish software interface.

Once the videos and the information are uploaded to the online video database, users can access the information from anywhere with the use of a web browser. Access to specific videos and content can be limited to levels of access and the administrator of the channel will create and dictate what content and by whom it can be seen. If analyses were performed on the video (e.g., a breakdown of a movement into key positions) then the person viewing the content would have access to the broken down movement, comments and any overlaid drawings that were placed on the video for analysis. Figure 10 below illustrates a video posted on Dartfish TV showing analysis of the start in the sport of Luge.

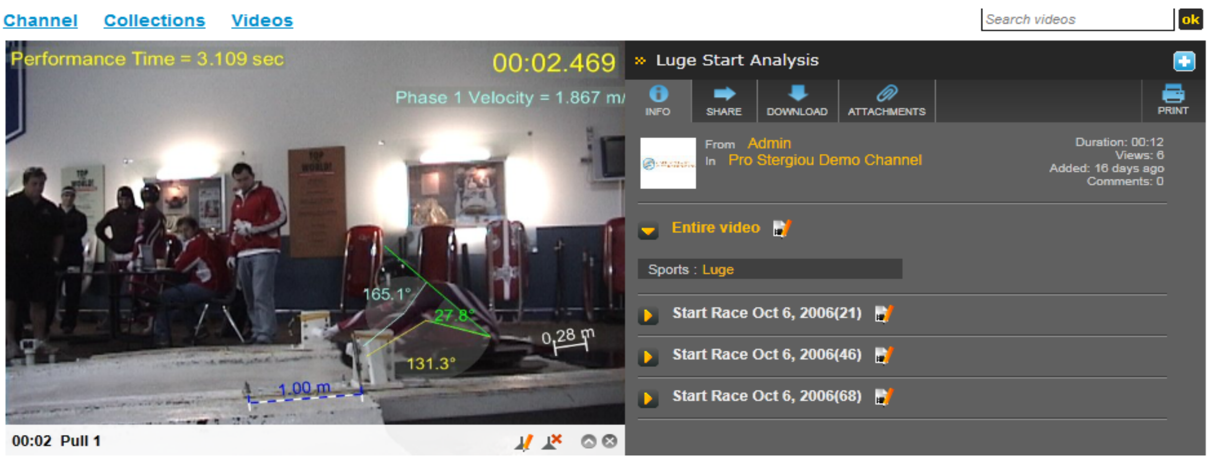

Figure 10. Video posted on Dartfish TV with analysis of the Luge start. The video is shown on the left side with the any text on top of the video. The capabilities to download this video and look at any attachments are shown on top, to the right of the video.

Videos that are related to the analysis of competition/game play can also be uploaded to Dartfish TV. The video, the events and the categorization of the events are all searchable and viewable. Figure 11 below shows a game from a team sport that has been uploaded to Dartfish TV.

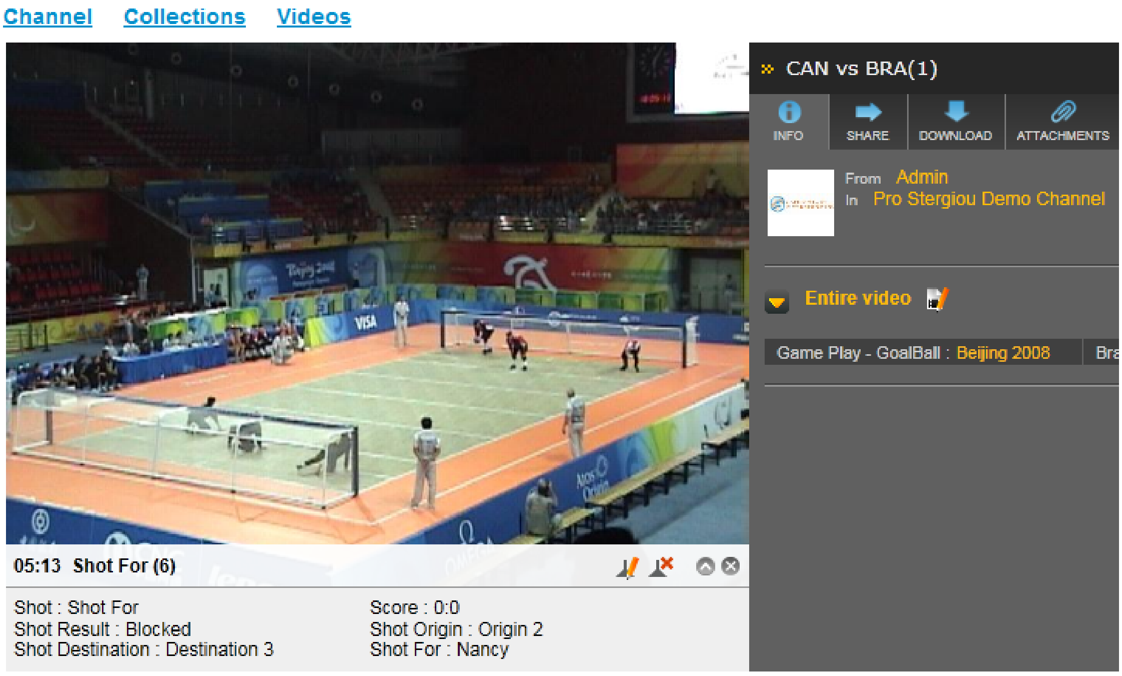

Figure 11. A screen shot of a video of game play during Beijing Paralympics 2008 of the Women's National Canadian Team. The video is playing a single shot from the game, with details located below the image categorizing the shot.

Other publicly available video databases exist for any person to post video content. "You Tube" is the best known one of these sites, however, the capabilities of these types of sites generally do not allow the user to create their own spaces, manage content nor control access. Other types of these online sites exist to store video content (Robinson, 2009; Siegchrist, 2009). Searches on the Internet will provide updated information on other available online video and content management databases using search terms like "uploading video to video sites" or "content management databases for storing sports data".

Future Considerations

The use of multimedia databases has the potential to revolutionize the way coaches, athletes, administrators and the public approach sport. It can impact the way sports are judged, the way athletes are selected to participate on teams, the way coaches approach training plans, and the public perception of sport both at the amateur and professional level.

Many skilled, elite level coaches and sport administrators make decisions based on intuition, 'gut feeling' and the influence of past experience. Effective use of performance databases can encourage sport professionals to employ an evidence based decision-making approach. This does not mean that they need to abandon their 'intuition' rather they need to integrate valuable data modeling into the decision making process. If the data does not support intuition, the data could be wrong, but at least there is a challenge to verify the decision. With the advent of online databases in sport, evidence based decision making is now practical.

For example, it is possible to collect and record experience, performance, injury, and skill data on each athlete in a given sport. In many countries team selection can be very controversial. In Canada, there is a special organization, the Sport Dispute Resolution Centre of Canada (www.crdsc-sdrcc.ca/eng/home.jsp) whose main purpose is "to help prevent and resolve disputes at the highest levels of the sport system". If sports bodies used evidence-based decision-making more often, it would reduce the number of disputes, as the decisions would be clear and less ambiguous.

However, it is incumbent on those who develop and maintain the databases to ensure that the objects (e.g. video, images, text, skill data) deposited in these repositories be properly organized and tagged so that they can be effectively and efficiently retrieved and analyzed by the appropriate decision makers.

With easy access to information, the issue of data governance is crucial. According to Wikipedia ((Wikipedia, 2009d)

> "Data governance is an emerging discipline with an evolving definition. The discipline embodies a convergence of data quality, data management, business process management, and risk management surrounding the handling of data in an organization. Through data governance, organizations are looking to exercise positive control over the processes and methods used by their data stewards to handle data."

This includes ensuring the quality of the data (completeness, accuracy, timeliness and decision usefulness of the data). Also of critical importance are the privacy policies and procedures that require special attention to storage and sharing of data that are collected on athletes (e.g. medical records). The potential of databases is great; there are many obstacles yet to be addressed; but clearly, it is worth the struggle.

References

Cheolkon Jung, & Joongkyu Kim. (2008). GNAVI: Golf navigation system based on player information, *in proceedings of the 2008 IEEE International Conference on Multimedia and Expo*, Hannover, Germany, IEEE, June, 23-26, 1413 – 1416.

Codd, E. F. (1970). A relational model of data for large shared data banks, *Communications of the ACM, 13*(6), 377-387.

Dick, R., Agel, J., & Marshall, S. W. (2007). National collegiate athletic association injury surveillance system commentaries: Introduction and methods. *Journal of Athletic Training,* 42(2), 173-182.

Digital Scouting. (2009). *Digital scouting database.* Retrieved August, 2009, from http://www.digitalscouting.com/index.html

Hua-Yong Liu, H. Z. (2005). A sports video browsing and retrieval system based on multimodal analysis: SportsBR. *Proceedings of 2005 International Conference on Machine Learning and Cybernetics,* Guangzhou, China, August 18-2, 8, 5077-5081.

Internet Consulting Services. (2009). *Why SideArm*? Retrieved August, 2009, from http://athletics.internetconsult.com/why-sidearm.aspx

Jeongeun Shin, & Ozawa, S. (2008). A study on motion analysis of an artistic gymnastics by using dynamic image processing - for a development of automatic scoring system of horizontal bar – *Institute of Electronics, Information and Communication Engineers (IEICE)* Technical Report, 108(46), 13-18.

Junge, A., Engebretsen, L., Alonso, J. M., Renström, P., Mountjoy, M., Aubry, M., et al. (2008). Injury surveillance in multi-sport events: The International Olympic committee approach. *British Journal of Sports Medicine, 42*(6), 413-421.

Own the Podium. (2009). *Own the podium - our initiative*. Retrieved August, 2009, from http://www.ownthepodium2010.com/

Robinson, R. List of Video Sharing Web Sites – Share your Videos, *The Online Video Marketing Guide – News, Tips & Trends,* Retrieved, October, 2009, from http://www.reelseo.com/list-video-sharing-websites/

Schoeman, J. H., Matthee, M. C., & van, d. M. (2006). The viability of Business Data Mining in the Sports Environment: Cricket Match Analysis as Application. *South African Journal for Research in Sport, Physical Education & Recreation, 28*(1), 109-125.

Siegchrist, G. (2009) Video Sharing Web Sites – Share Your Videos Online, *About.com:Destop Video*, Retrieved November, 2009, from http://desktopvideo.about.com/od/videohostingsites/a/vidsharing.htm

Simin Li, & Jinhai Sun. (2009). Study on simulation of gymnastics movement and the arrangement based on virtual reality technology., ICIC '09. Second International Conference on Information and Computing Science, Manchester England, May 21-22; 3, 270-273.

UmaMaheswari, P., & Rajaram, M. (2009). A novel approach for mining association rules on sports data using principal component analysis: For cricket match perspective*, in proceedings of the IEEE International Advance Computing Conference*, Patiala, India, March 6-7, 1074-1080.

Vincent, J., & Childs, T. (2009). *Database summary*. Unpublished manuscript, Sport Technology Research Laboratory, University of Calgary.

Wikipedia. (2009a). *Database management system*. Retrieved August, 2009, from http://en.wikipedia.org/wiki/Database_management_system

Wikipedia. (2009b). *Content management system*. Retrieved August, 2009, from http://en.wikipedia.org/wiki/Content_management_system

Wikipedia. (2009c). *Data mining*. Retrieved August, 2009, from http://en.wikipedia.org/wiki/Data_mining

Wikipedia. (2009d). *Data governance*. Retrieved November, 2009, from http://en.wikipedia.org/wiki/Data_governance

PART 2

APPENDED PAPERS

Sport Informatics – Historical Roots, Interdisciplinarity and Future Developments

Daniel Link & Martin Lames

Faculty of Sport Science, Munich University of Technology,Germany

Abstract

Computer science has become an important partner for sport science. Although many cases for the successful application of computers in sport exist, only a few meta-theoretical mediations on this research field can be found. To fill this gap, this paper looks at the genesis of sport informatics in Germany, discusses the interdisciplinary relation between computer science and sport science and forecasts its development up to the year 2020. Using existing models of interdisciplinarity the paper classifies research activities into four types of cooperation and emphasizes the importance of "genuine" interdisciplinary research. This discussion leads to a compact definition of sport informatics and to a graphical model, which illustrates its subject matter and the structural relations between sport science, computer science and their fields of application. The future section identifies increased computing power and network capacity, networks concepts such as pervasive or cloud computing, small one way electronics, flexible displays and positioning technologies being the most relevant technological improvements for sport. These technologies have a potential impact e.g. on the measurement and reporting of physiological and positional data, on computerized performance analysis, sport clothing and the quality of computer simulations in the fields of sports engineering, motor behavior and physiological adaptation.

KEYWORDS: SPORT INFORMATICS, THEORY OF SCIENCE, DEVELOPMENT, DEFINITION, FORECAST

Introduction

Over the past three decades, the discipline "sport informatics"[1] - also called "computer science in sports" - has become an important part in the spectrum of sport scientific research. The term covers all activities at the interface of computer science and sport science, ranging from simple tools for handling data and controlling sensors on to the modelling and simulation of complex sport-related phenomena. Whereas first applications in the 1970s used computers for information and documentation purposes only, current approaches describe virtual environments for the training of perception tasks specific to sport (Levy & Katz, 2007), the scope of computer technologies for coaching (Lames, 2008) or the automatic analysis of sport games using pattern recognition (Hoyningen-Huene & Beetz, 2009).

[1] The term sport informatics originates from a congress in Graz (Austria), organized by the International Organisation for Sports Information (IOSI) in 1975. The related proceedings were published by Recla & Timmer (1976) under the German title "Kreative Sportinformatik" (engl. "Creative Sport Informatics").

Today, computer science in sport is a well-established research field. An International Association on Computer Science in Sport (IACSS) has existed since 2002 and promotes research in this area. In many countries such as Australia, Croatia, Germany, Turkey and India national workgroups have been established, which represent sport informatics in the national scientific community and contribute new technological innovations to sport. The IACSS also maintain good relations with various other sport scientific organisations like the International Association for Sports Information (IASI), the International Council of Sport Science and Physical Education (ICSSPE) or the International Sports Engineering Association (ISEA).
Although there is no lack in research activities and scientific institutionalization, only a few meta-theoretical studies on sport informatics can be found. Most publications reports on applications of computer technology, software tools and informatic methods and paradigms in sport, but they do not deal with epistemological questions about the discipline. While some authors (Perl & Lames, 1995; Baca, 2006) suggest a rough internal structure of the discipline, others (Fischer, 1998) allude to the characteristics of interdisciplinary cooperation, but forgo a detailed discussion.
This situation prompted the authors of this paper to reflect from the point of view of sport informatics on the development of the discipline, to analyse its current situation and to speculate on its future. As the foundation of the IACSS was preceded by several national working groups, the first section of this paper will give a brief overview on developments in Germany and intends to make the German perspective available to readers outside of this country. The second section discusses the interdisciplinary relation between computer science and sport science and identifies different types of cooperation. Based on this discussion section three suggests a definition and a structural model of the discipline. The fourth issue is extrapolation: we know of some possible developments in informatics, but what will be their impact on sport mediated by sport informatics? This question is addressed in the last section, where we propose ten short theses about the state of the discipline in 2020.

Historical Roots – The Development in Germany

The discussion about interdisciplinarity in sport informatics should take into account the concepts of the individual disciplines. The first two subsections outlines the conceptual structure and conception computer science and sport science have of themselves and focuses on the epistemological discussion in the German scientific community. The third subsection gives a short overview about the institutional development of sport informatics in Germany.

Computer Science

In the sixties and the seventies of the last century, in Germany the term “Informatik” was mainly associated with questions of technology. A popular German encyclopaedia described “Informatik” as “the science of the systematic processing of information, in particular the automatic processing using digital computers” (Engesser, 1986). In terms of this definition, the discipline includes mathematical activities, which deal with algorithmic processes for the description and transformation of information and also engineering activities, concerning aspects of the development and application of computers. This technological perspective conforms to the common understanding of the discipline “computer science” in the United States or Great Britain (National Research Council, 2004).
In the beginning of the eighties, the importance of computer systems increased in almost every part of society. It became more and more clear, that the use of computer systems leads to interactions between system processes and the processes in the real world. To study these

interactions, many computer scientists adopted approaches and methods from social and behavioural sciences. These research fields were accepted as a part of the discipline "Informatik". Today, many countries use the English term "informatics" - derived from the German "Informatik" - for the science of information. Nygaard (1986) for example defines "informatics" as the "science that has as its domain information processes and related phenomena in artifacts, society and nature". This perspective separates the mathematical/logical part from the technical one and refers to the concepts of cybernetics and system theory.

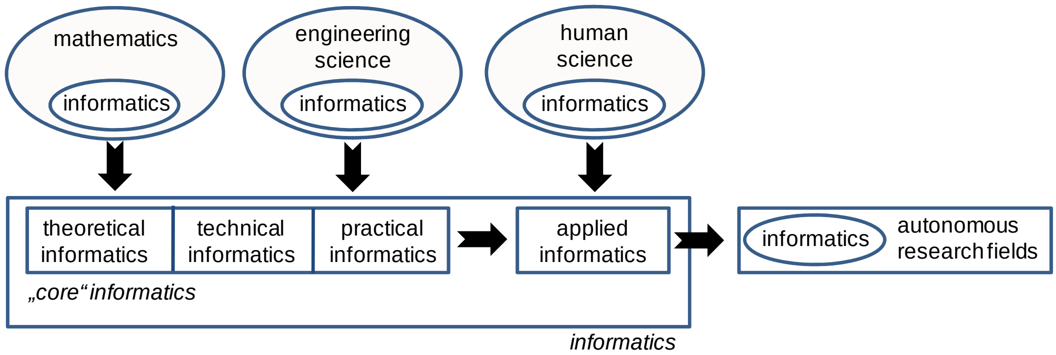

Fig. 1. Commonly used structural model of informatics. Informatics emerges by separation from mathematics and engineering science - later approaches from human sciences were integrated. The discipline is divided into the subdisciplines theoretical, technical and practical informatics, which are called "core"-informatics (Claus, 1975). The application and question related to the use of computers are studied by applied informatics. Also some autonomous research fields like bio-informatics or neuro-informatics exists, but they are not treated as a part of the discipline.

One characteristic of computer science is its ambition to support others scientific disciplines. In many cases the combination of technical expertise from computer science and specific domain knowledge leads to autonomous research fields like bio-informatics, neuro-informatics, business informatics or sport informatics. In Germany there was a debate about whether these research fields should be accepted as a part of the discipline. Luft (1992) for example, claims a strict distinction between the cooperation fields and the core area of "Informatik". Today, the discipline in Germany (also "Informatique" in France) is a kind of mixture between computer sciences and Nygaards concept of informatics (see Fig. 1), but nevertheless the question about its boundaries is still a topic of discussion.

Sport Science

German sport science has undertaken a definition process akin to that in Informatik. One common definition, published in a German Sport Encyclopaedia, describes sport science as the "collectivity of knowledge, scientific argumentation and research methods that deal with problems and phenomena related to sport" (Röthig & Prohl, 2003). While this definition is obvious, it provides no sight into the epistemological characteristics of the discipline. For example, it is quite difficult to define, what the term "sport" exactly means. One approach was the formulation of criteria, which are common to all instances of sport (e.g. motor activity, principle of organisation, non-productiveness, fair-play, performance). It is easy to see, that these criteria - however they are selected – do not apply every type of sport. To call an activity sport, it is neither necessary that all criteria are present, nor is it possible to say, which of the possible combinations can be regarded as being sport.

Another important point in the theoretical discussion is the relationship between the subdisciplines of sport science. In the late 1960s, the existence of German sport science was contentious. The proponents needed a good reason as to why a new discipline with its own structures and resources should be established in the academic landscape. Therefore the argument was put forward that the complexity of sport could not be investigated by existing research fields (Grupe, 1971). The exigency of one unified discipline, with a high degree of interdisciplinarity between its subdisciplines, was the central argument for the foundation of sport science. To support this position, Ries and Kriesi (1974) proposed a model, which shows the development of sport science in three phases, (1) detaching from base disciplines, (2) aggregation of subdisciplines within a multidisciplinary science and (3) integration of sub disciplines into a consistent and integrative science (Fig. 2).

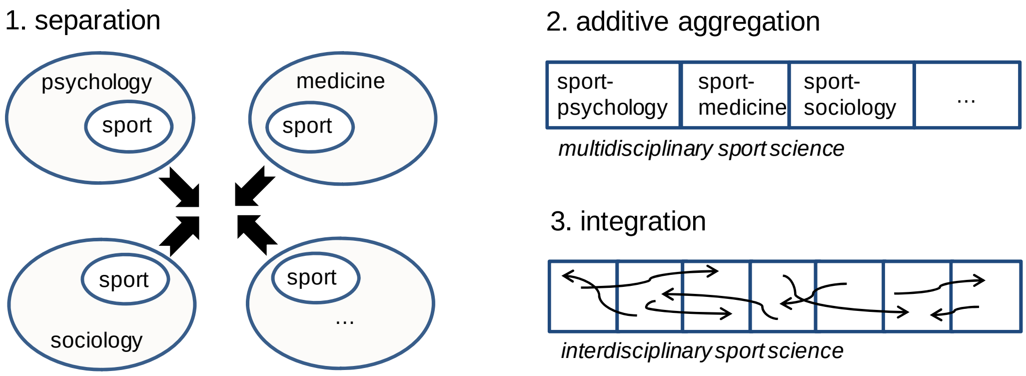

Fig. 2. Idealized model of sport science development (Ries & Kriesi, 1974). Detaching of sport related research fields from bases sciences, additive aggregation into a multidisciplinary science and integration of sub disciplines into a unified science.

The scientific reality showed, that - in contrast to this idealized model - sport was mostly studied through the eyes of each subdiscipline (e.g. sport sociology, sport psychology, exercise science). Today, German sport science has accepted that the idea of an integrative science was not a very realistic one. Sport science does not describe itself as a "unified science of sport", but as a collection of overlapping research programs, in which interdisciplinarity exists only in a sporadic and theme centred way (Höner, 2001).

Sport Informatics

The idea of an interdisciplinary science "sport informatics" in Germany was developed initially by Jürgen Perl. Together with Wolf Miethling he published the first monograph in the discipline (Miehtling & Perl, 1981) and marked the beginning of sport informatics in the Federal Republic of Germany.

Since 1985, Jürgen Perl has worked at Mainz University and founded a working group in sport informatics. Contacts with other German groups in this field of study led to workshops on sports and informatics. The first of these took place in Hochheim in April 1989. It was attended by the most important German sport science groups in this subject. This first workshop was supported by the German Institute of Sport Science indicating the strategic importance attached to this development. It was the starting point of a series of biannual conferences. The most recent workshop (the 11th German Workshop on Sport and Informatics) took place at Augsburg, Bavaria, in May 2008.

In 1994 a new strategic aim was pursued. The German Association of Sport Science (Deutsche Vereinigung für Sportwissenschaft (DVS)) represents German academic sport science with 900 members at 67 universities. It is organized in sections, representing the disciplines of sport science, and commissions giving an organizational framework for special interdisciplinary topics, like health or gender, but also for single sports like football and swimming. The strategic aim of establishing a section within DVS dedicated to research activities in the field of sport informatics possessed many attractive perspectives. First, the responsibility of caring for the research area would not be solely the task of the Mainz working group. Other groups were interested in organising the workshops. Second, being under the umbrella of DVS meant that there is much more support for example, to organise conferences, and improve access to the community of sport scientists from all other disciplines. The most important point though was establishing a section in DVS meant that sport informatics was acknowledged a discipline of sport science in Germany.

In September 1995, the general assembly of DVS supported with overwhelming majority the request to establish the Sport Informatics Section. This may be considered the formal birth date of the scientific discipline. This vote was, of course, the final act of a chain of political actions and personal discussions that were necessary to promote the idea. The scientific community was informed about the aims and methods of a scientific discipline sport informatics beforehand (Perl & Lames, 1995). The main intention of this article was to make clear to a broader community that informatics does not only provide tools for the work of sport scientists like text processing or statistics software. It offers also input at higher levels of scientific work, thus justifying the formation of a scientific discipline. The idea of interdisciplinary collaboration presented in this early article was implicitly a one-way model. Theories, methods, and tools from informatics should be applied to sports (see next section of this paper).

Soon after establishing a national association for sport informatics a new aim came into focus. The aim was to establish an international scientific association. There were good reasons to do so. As the world became a truly global village through advances in information and communications technologies, it became clear that in different parts of this village, people were addressing the same problems (Lames, 2008; Baca, 2006; Hughes, 2000; Perl, Lames & Miethling, 1997). In the area of game analysis one could for example mention the introduction of digitized boards, efforts to enter data by natural language detection software or the struggle for reliable computer-video couplings. These developments were brought forward independently for example at Mainz and Cardiff in the working groups of Jürgen Perl and Mike Hughes respectively (Lames, 2008).

After three international meetings in Cologne (1997), Vienna (1999) and Cardiff (2001) the International Association of Computer Science in Sport (IACSS) was founded at Barcelona in 2003 and Jürgen Perl became the first president. Since then, a series of biannual international conferences has been organized and members from different countries and almost all continents have joined the association. The future prospects of the association are excellent according to the present president, Arnold Baca, in his welcome message at the 11th German Workshop on Sport and Informatics at Augsburg, Bavaria, in 2008. The unique combination of sport science and informatics with the large application field of sports at any level provides these excellent perspectives. Nevertheless, the remarkable developments of the two sciences involved and in the field of sports make it necessary to reflect periodically on the levels achieved in interdisciplinary cooperation between the fields and the concept of sport informatics.

Interdisciplinarity in Sport Informatics

At first this section outlines the interests in cooperation between computer science and sport science in common projects. While the motive of sport science is quite obvious, that of computer science needs more elaborate discussion. The second part poses the question, which quality of interdisciplinarity between sport science and computer science exists today and which quality would be desirable and realistic in future? This is done by discussing existing models of interdisciplinarity and proposing a classification for research activities in sport informatics.

Common fields of interests – Why do computer science and sport science cooperate?

Here it is useful to differentiate between political, scientific and personal motivations behind cooperation. From a political perspective one must bear in mind that interdisciplinarity is considered an important research paradigm in most countries. The German Research Foundation (DFG), which is the central research funding organisation in Germany, holds the view that "scientific progress arises more and more at the borders and intersections of disciplines" (DFG, 2008). The national funding agency for sport scientific research in Germany (German Federal Institute of Sport Science) terms interdisciplinarity a "key element" of their founding policy. Announcement on funding initiatives refer to an "inter- and multidisciplinary approach […], integrated construction of theories, highly specialised choice of research methods and […] integrative presentation of results" (BISp, 2008). While the precise meaning of such catch-phrases is somewhat clouded in jargon, a scientist, whose career depends on the positive evaluation (and funding) of his research projects, is ill-advised to refuse the commitment to interdisciplinarity.

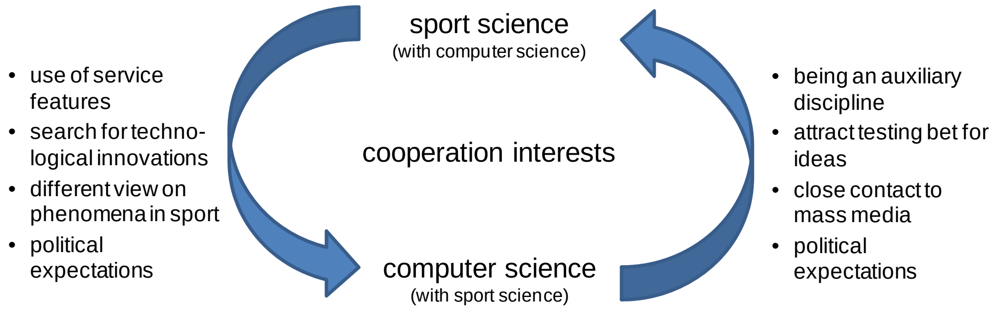

Fig. 3. Cooperation interests of computer science and sport science.

Sweeping aside this consideration, sound scientific justification for cooperation does exist (see Fig. 3). First of all - from sports science perspectives - computer science serves as an appreciated partner in those areas in which sports scientists do not excel themselves. This applies to data handling and software development, e.g. for the purpose of documenting training, controlling sensors or visualizing data. Second, information technology is an important source of innovations for training and competition. Collaborations with computer science help sport scientists to become aware of new technology and have the advantage of facilitating their availability in good time. Third, sport science expects that the approaches and perspectives of computer science should be transferable to the field of sport. For example the concept of soft computing can assist the understanding of phenomena in sport (see Section 3).

Less immediate are the benefits of cooperation with sport science from the computer science perspective. While traditionally computer science has supported other sciences, extending its own area of influence would not provide sufficient argument in favour of supporting projects in the field of sports. Of greater interests is the complexity of sport, which is well suited for testing and validation of methods and techniques of computer sciences. The existing structures in sports are neither too simple to be interesting, nor to complex to be described using mathematical models. One example of a problem is that of reconstructing human intention, e.g. as in the case of computer-aided crime detecting based on video recordings of surveillance systems (see Boghossian & Black, 2005). In fact sport science deals with sport games analyses in which similar computational requirements exists (automatic recognition of players, moves and strategies) albeit with reduced complexity (limited degrees of freedom, common rules, tactical invariants) is a similar problem. Computer science expects the development and validation of solutions for sports, to lead to knowledge, which then can be transferred to the initial problem. More generally speaking, sport could act as an attractive testing field for computer science, in which human behaviour can be observed and studied in a simplified, yet authentic field.

Additional motives for computer scientists are the societal relevance of sport and its huge role in mass media. This may give rise to the Basking in Reflected Glory (BIRG) phenomenon (Cialdini et al., 1976)). The "exotic" application field can also help them to build a reputation in the scientific community. Last but not least, many computer scientists, working with sport science, a personally involved in sport. Even if collaboration cannot be fully justified on the basis of individual involvement, political considerations and increasing publicity seem to have importance as secondary motives.

Quality of Interdisciplinarity – How do computer science and sport science work (or should work) together?

There are many ways in which the concept of "interdisciplinarity" has been classified by philosophy of science. One milestone in nomenclature was a congress in the year 1972, where the OECD proposed a classification of interactions between disciplines (OECD, 1972). In terms of this definition, multidisciplinarity is a juxtaposition of various disciplines without a connection between them. Interdisciplinarity describes any interaction between disciplines, which can range from simple communication of ideas to the integration of concepts, methodologies and epistemologies. Transdisciplinarity is the highest degree of cooperation and stands for a common set of theories and axioms for a set of disciplines (Fig. 4). On this basis, enhanced models, focussing on different aspects of interaction were developed: Heckhausen (1972) for example identifies six types interdisciplinary research, Boisot (1972) advises three categories of interdisciplinarity, Karlquist (1999) lists five modes of interdisciplinarity (an overview is provided by Chettiparamb, 2007).

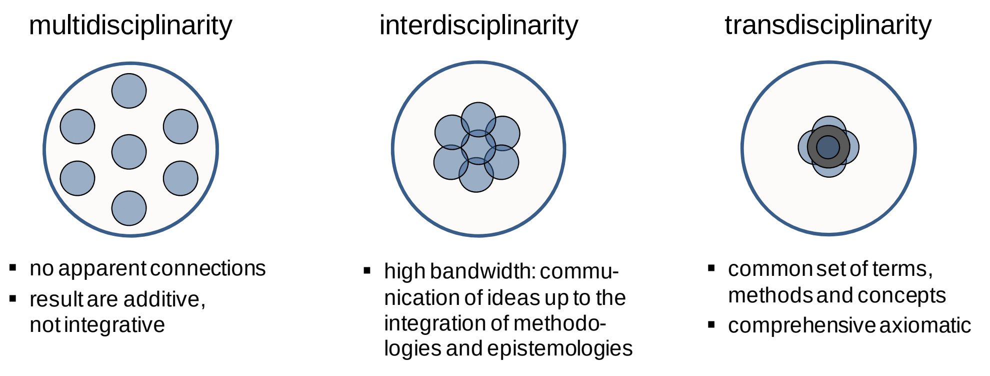

Fig. 4. Commonly used classification of types of interdisciplinarity (OECD, 1972).

When looking into the practice of sport informatics, it emerges that not any of these models is adequate to describe the existing interaction. The "borrowing" of computer scientific methods (type a, b, c) matches to Heckhausens concept of *auxiliary-interdisciplinarity*. The simple usage of pre-defined tools (type a) corresponds to the OECD-term *multidiscplinarity*. The corporate development of tools/methods (type b) can be called *pseudo-interdisciplinarity* (Heckhausen) or *restrictive interdisciplinarity* (Boisot). The use of sport scientific knowledge in computer science (type d) accords with the idea of *structural interdisciplinarity* (Boisot). In this regard, this paper proposes an own classification, using four types of cooperation (see Fig. 5):

- Type a: Sport science applies existing approaches and tools from computer science. In this case, sports science does not take part in conceptualization and development. Computer science (or - being more precise - commercial software developing companies) only act as an anonymous service provider, without contact with sport science.
- Type b: Sport science integrates knowledge from computer science. This happens, when its own area of studies needs technical solutions, which do not exist on the market. Knowledge is assimilated either by acquiring the skills necessary or by entering into partnerships with computer science e.g. by means of student or third party funded projects. One aspect of this cooperation is that computer science provides nothing but skills in software development. There is no collaboration on a scientific level.
- Type c: Computer and sport science cooperate in research programs, which are in accordance with the research interest of both disciplines. Examples are the use of artificial neuronal networks for analyzing movement patterns or application of image recognition algorithms in sport game analysis. In this cases, computer science gets new insights by validating concepts and methods which have relevance for additional - perhaps more complex - problems. Sports science benefits from an improved and faster data acquisition and by getting a different perspective on the structures of sport.
- Type d: This type is comparable with type c, with the difference that paradigms and knowledge of sport science are used in computer science. An example would be the use of kinesiological models in controlling the motion of humanoid robots.

A review of the research activities in the last 20 years reveals that many projects of type a and b, but only a very few projects of type c and d can be found. One reason, why the popularity of sport informatics in the computer scientific community is not very high (there are computer scientists, who never heard about it or do not show interest in cooperation with sport science) might be, that there is often no genuine interdisciplinary research. A deeper concentration on those fields where computer science can profit from sport scientific paradigms and knowledge (types c and d), could improve the situation. This would require better communication of sport scientific expertise and recognition of sport as a fruitful application field for computer scientists (see Fischer, 1998).

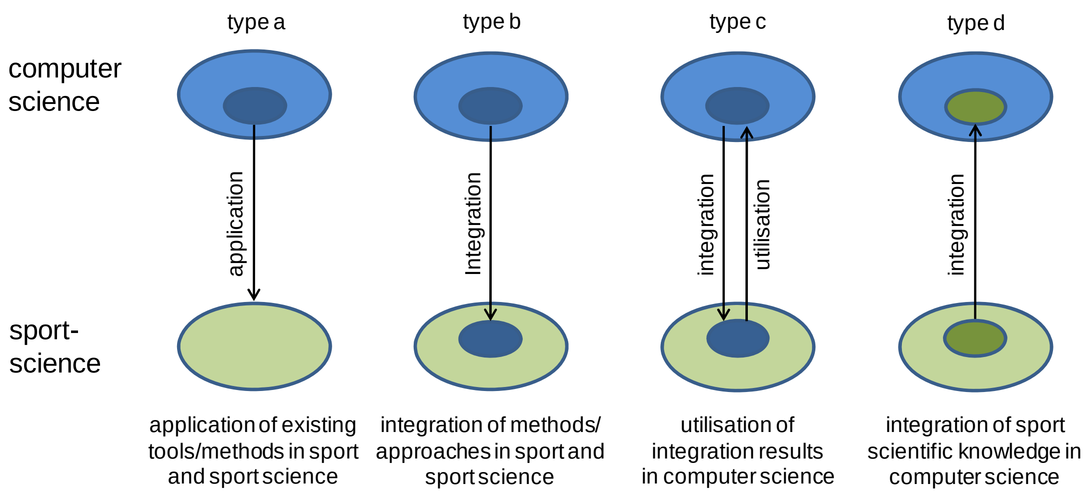

Fig. 5. Types of cooperation in sport informatics.

On the other hand, low affinity between the disciplines sports science and computer science must be taken into consideration. In contrast to other interdisciplinary linkages (like biology and chemistry, astronomy and physics or sociology and psychology), there is no common borderline, where one disciplines changes into another and no shared knowledge. Consequently sport informatics cannot be an autonomous inter-discipline like astrophysics or biochemistry. One-way-transfer is also a natural problem, because the processing of information is fundamental for all sciences, whereas the applications fields of computer sciences usually cannot provide any knowledge for the core area of computer science (apart from mathematics and electrical engineering). Even sport science and computer science have problems in creating real interdisciplinarity between their sub disciplines. They are both heterogeneous sciences without consistency in level of theoretical integration, axioms, methods and terminology (see discussion in Section 2). For this reason, while it is advisable to postulate and advance interdisciplinarity, it does well not to overcharge the idea of integration.

Mission statement – A self definition of sport informatics

Having the discussion of the last sections in mind, we suggest differentiating between sport informatics and computer science in sport. Computer science in sport stands exclusively for the use of computer technology in sport and sport science. Sport informatics also includes the application of methods and paradigms from computer/information science as well as from research programs, which try to transfer sport scientific knowledge to computer sciences. The following definition shows this enhanced self concept:

> Sport informatics is a set of multi- and interdisciplinary research programs at the interface of sport science and computer science. The materiel field is the application of tools, methods and paradigms from computer science on questions of sport science as well as the integration of sport scientific knowledge in computer science.

We can see in Figure 6 a diagram visualizing this standpoint: In both disciplines there is knowledge, that is potentially useful for the other discipline. Conversely there is a second area, which might be an application field for the findings of the other discipline. The research programs of sport informatics include parts of both disciplines and can be dedicated to one of the four types, identified in the last section.

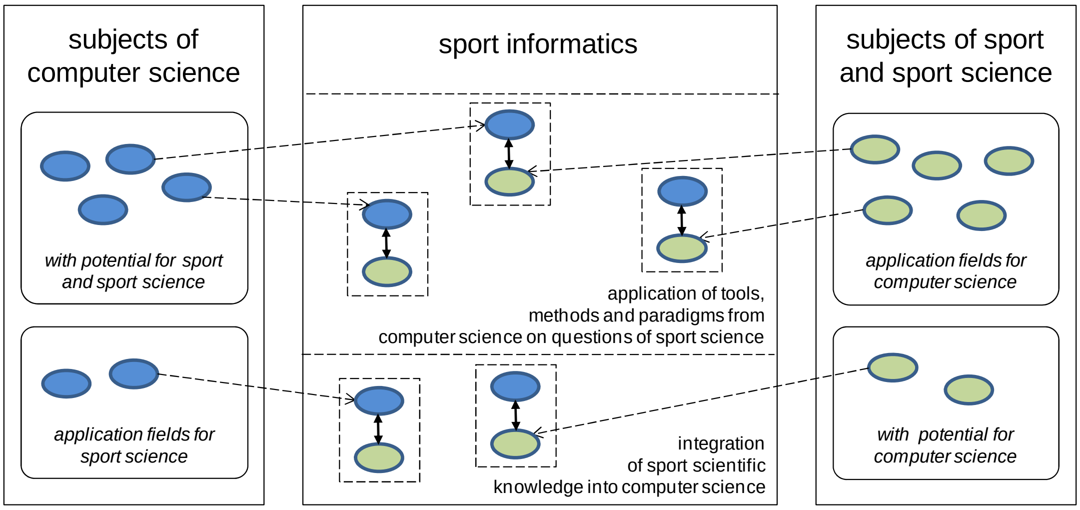

Figure 6. Basic structure of sport informatics. The discipline can be described as a set of multi- and interdisciplinary research programs. Most of these programs apply technological and methodological knowledge of computer science to study questions of sport science, but there are also some sport scientific findings, which can be useful for computer science.

Figure 7 shows a refinement of this rough structure by using a matrix with four areas. The upper both areas give examples for research fields for computer science with are useful for sport and sport science, the lower areas give examples, how computer science can profit from sport science. The next subsections discuss the two different directions of integration.

Computer Science in Sport Science

The first area in Figure 7 (top, left) shows topics in computer science, which may be useful for sport science. According to Perl and Lames (1995) we divide these area using two dimensions. The first dimension includes four main research areas of computer science, which are important to sport: (1) capturing and storing of data, (2) modelling, analysis and simulation, (3) presentation and visualization and (4) communication. The second dimension illustrates the idea, that the research field sport informatics is (or should be) more than just the simple application of tools for the recording, analysis and presentation of data. In addition to the "tool-level" (which is more information technology than computer science), there are also methods, theories and paradigms, which have a potential to support sport science (see cooperation types c and d in the last sections). The next paragraphs discuss some exemplary items in this sub matrix.

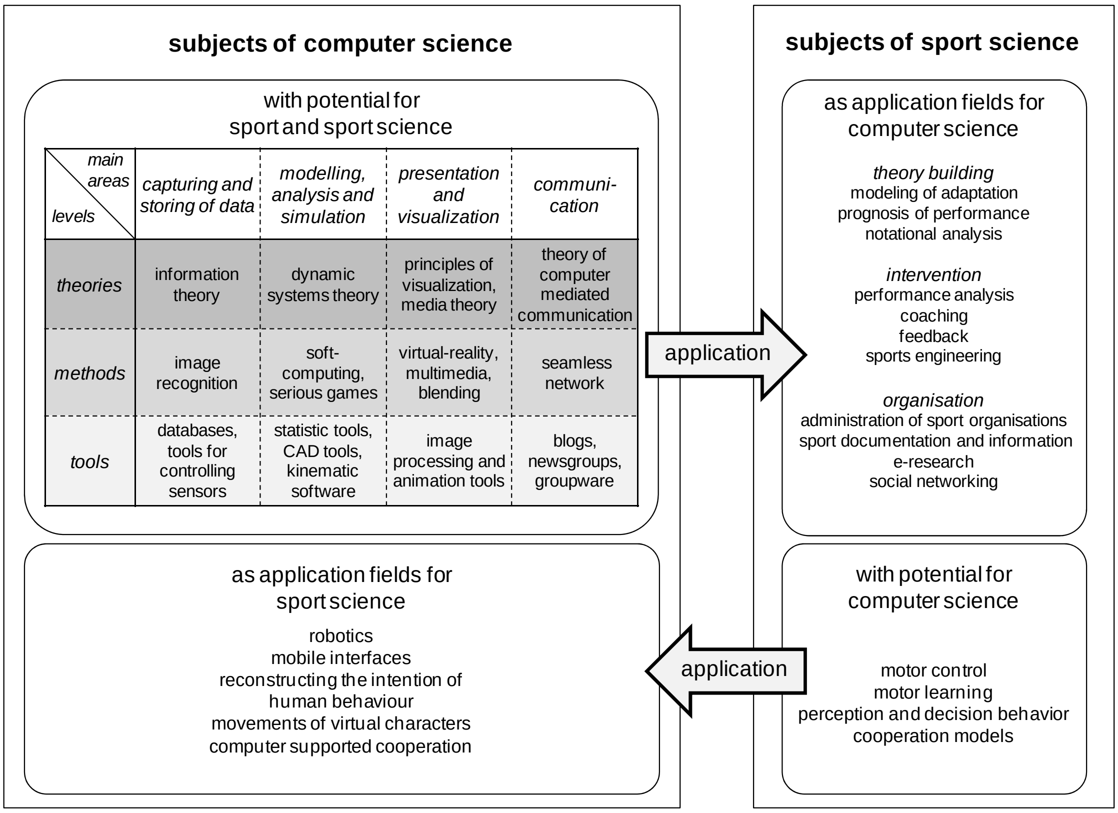

Figure 7. Subjects of sport informatics. The matrix shows examples for the research and application fields mentioned in Figure 6.

In the field of capturing and storing data, information technology provides for example database tools, which can be used for the storage of training and competition data (e.g. type, intensity and amount of training, tactical behaviour or performance) (Zeleznikow et al., 2009). This allows coaches to stay informed about the input load and the performance level of their athletes. Also many biomechanical devices like force plates, high speed cameras, GPS-devices and motion capturing systems use proprietary software tools controlling sensors and helping to manage measured data (Turner, 2009). For the modelling, analysis and simulation, sport scientists often make use of mathematical software like Matlab, Maple, SPSS or apply software for movement analysis. Sports engineers rely on tools for computer aided design (e.g. AutoCAD or Solid Edge) to develop new sports equipments.

Tools from information technology are also important for the presentation and visualization of data. In sport games for example, the linking of video sequences and databases is useful for game observation. In athletics or gymnastics, coaches use software, which allows to show additional information like force vectors, torsional movements or speed information inside a video or to superimpose pictures of two movements. Last but not least communication technology can help to organize sports: Today internet blogs or social networking websites are relevant for organizing training sessions and to find partners for team sports. Internet based groupware tools like synchronous video, video conference and whiteboards, supports athletes on international championships, by realizing communication with their coach at home (Link & Lames, 2005). In this context mobile devices for coaching become increasingly important (Link & Lames 2006; Novatchkov et al., 2009).

On the method level, computer science for example has developed techniques for image recognition, which can be helpful to capture positional and biomechanical parameters directly from video recordings (Beetz et al., 2005). Therefore computer science does not provide readymade tools, but rather general algorithms for colour, texture and shape comparison, which have to be adapted to the sport specific situation (see cooperation types b and c in the last section). In the field of analysis and simulation of sports, the methods of soft computing have become more important: artificial neuronal networks are used to simulate the relationship between training input and performance output in swimming (Edelmann-Nusser et al., 2002), to analyse tactical patterns in handball (Pfeiffer & Perl, 2006) or to identify movement patterns in basketball (Lamb, Bartlett & Robins, 2009). Generic algorithms help to find solutions in high dimensional configuration spaces e.g. to optimise the design for sport equipment (Vajna et al., 2006) or to optimise throwing movements (Bächle, 2003). Last but ot least, serious games can have a positive effect on perception, reaction and motor control and are potential useful for education and intervention in sports (overview is given by Wiemeyer, 2009).

At the theory level the theory of complex dynamic systems is an example that holds a lot of promises for sport science. Many processes in sport seem to base on non linear coupling rules, which lead to complex phenomena. The theory of complex systems helps to model and to understand for example interaction in sport games as well as processes of biological adaption due to training. Successful examples are perturbations (Hughes, Dawkins & David, 2000; Jörg & Lames, 2009), relative phases (Walter et al., 2007), chaos theory (Lames, 1999) of the paradigm of self organization (McGarry et al., 2002).

The second part of the matrix (top, right) shows examples for application fields in sport and sport science. These fields are structured which the headlines “theory building” (getting new theoretical insights into phenomena of sports), “intervention” (improving training and competition) and “organisation” (managing activities related to sport). Examples for these application fields are already mentioned in this section.

Sport Science in Computer Science

The third and fourth parts (bottom left and right in Fig. 7) shows examples for sport scientific knowledge fields, which are potentially useful for research in computer science. Here, no internal structure is needed - we just give three examples to make this direction more concrete.

The first example is software for the safe and autonomous operation of robots. Traditionally the algorithms for navigation, locomotion and objects grasping bases on two different concepts: planning and controlling (Latombe, 1991). Planning methods define the movements and the position of the joints at any time before the movement. This requires complete information about the entire environment and the objects to be manipulated. On the other hand, controlling methods rely on local status information during the movement, based on visual or force feedback. This allows reacting on unexpected events like obstacles, but without global information, it is not guaranteed, that the algorithm finds the (best) solution for a task.

Kinesiology knows that human motor control does not follow the planning or controlling paradigm, but it should better be described as self organizing process, influenced by both aspects. For example top level dart player show a substantial variability in velocity, joint angles and the sequential timing of body parts from trial to trial (Müller & Sternad, 2004). The presumably most important skill for elite players is to balance parameters during the movement and not to reproduce fixed motor programs. If sport science could understand this self organizing processes in detail, these finding could also be used in the development of new paradigms for controlling robots.

The second example is cooperation in sport games. An important factor for success in soccer is the quality of interaction between the players of a team. Interaction is needed on the level of the entire team, e.g. by shifting the team formation in dependency of the tactics and the position of the ball, as well as on the level of subgroups, for example when a striker starts running to receive a pass before the ball was played. Some of these interactions are highly trained - others arise spontaneous based on the situation. The conditions for successful or non successful interaction between players (e.g. which are components of decision making in soccer, which agreements are needed for organising the defence, which cues are used for the timing of a pass?) are of great interests for exercise science and sport psychology. If they could develop models how cooperation in sport games takes place in detail, these results would be valuable for computer scientific research fields like intelligent autonomous systems. A third example can be found in the field of mobile computing. Mobile computing - which means the use of computers during movement - is a fast growing application field for information technology. Examples are the use of handhelds in medicine (documentation of patient records), in military (geographical information for foot soldiers) or in sport (feedback about physiological parameters). One important aspect is that mobile computers places extended demands on the user's coordination and cognition (Kjeldskov & Stage, 2004). While running or walking, he has to adjust the movements of the legs with the movements of the hand-arm-system. On a cognitive level the user must pay his main attention to his forward locomotion and at the same time, he has to look on the screen to coordinate his hand movements. Up to now, our knowledge about the interrelation between walking speed, heart rate, input, reading performance and interface design has been quite vague. Experiences and research methods from kinesiology and biomechanics can help computer science to develop user friendly human-computer-interfaces, which are even usable under physiological stress.

Perspectives for 2020

The last section of this paper provides a perspective on potential developments in the field over the next ten years. Two questions are in the centre of interest:

1. Which developments in the field of information and communication technology can be expected?
2. How can these developments be used in sport und sport science?

Even when nobody can forecast future developments in detail, it is nevertheless possible to pre-estimate, which technologies are on the top of the agenda. For this purpose we refer to a study which was commissioned by the German Federal Ministry of Education and Research (Oertzen et al., 2006). The study uses a Delphi method, which relies on the phenomena "wisdom of crowds" (Surowiecki, 2004). In two rounds, a panel of independent experts (n=681) from science, economy and confederations was asked, which information and communication technologies will become important in the next 20 years. The opinions were collected, summarized and reflected back to the members. In the second round the members were asked to revise the own answer with regard to the others. As a result, the study formulated several prospects about the engineering and temporal feasibility of technologies. Some of the predicted developments, that might have an impact on sport and sport science, are:

- Miniaturization and computing power: The physical barrier is still not reached and Moore's law (number of transistors that can be placed inexpensively on an integrated circuit is doubling approximately every two years) will continue for at least ten years (Gelsinger, 2008). It is quite obvious, that the processing speed of a standard computer (price 1000$, see Kurzweil, 2005) will grow in similar dimensions.
- Network capability: Under everyday conditions, the data transfer rate will reach about 1 Gbit/s (today ca. 0,02 Gbit/s). This is true for wired data exchange via "the last

mile" and also for stationary and mobile wireless systems. Wireless networks with a small range (e.g. WLAN) will reach this capacity earlier than wireless wide area networks (e.g. UMTS).

- Networking: Miniaturization of information technology will lead to an increasing integration of processing devices into everyday life (ubiquitous computing, pervasive computing). These autonomous devices can communicate spontaneously with each other via a self-configuring wireless network. This "internet of things" (ITU, 2005) bases on common standards and can use any network infrastructure that is available in the environment (seamless network).
- Cloud computing: Many companies will offer cloud computer services like common software, data storage and computational power via the internet. The running and maintaining of an own it-infrastructure is getting more and more unnecessary (Dikaiakos et al., 2009).
- Small, inexpensive processing devices: Using synthetic material and print technologies, it is possible to produce electronic devices very cheap. This makes it cost-effective to integrate them into objects for one way use.
- Power supply: The electrical power supply of mobile devices will stay a key problem. Hardware, which small energy requirements, like sensors or RFID-devices, can use alternative energies like body heat, kinetic energy, light or sound waves.
- Display technologies: Important developments in the field of display technologies will be small and lightweight data glasses and virtual retinal display, which can draw a raster display directly onto the retina of the eye.
- Geo-spatial positioning: Additional to the American NAVSTAR-GPS and the Russian GLOSNASS-System, the Galileo-System of the European Union will be available. In particular, the combination of the systems will allow a better coverage and accuracy (<1 m, at 95 % coverage) than today. Also image tracking and delay based approaches (microwaves) will lead to an improvement of position measurement (<10 cm, at 100 % coverage) in the close-up range.
- Speech recognition: Software that coverts spoken words to machine syntax will be able to recognize most speakers without training with an accuracy of more than 90 %.
- Software components: In software engineering, the importance of predefined software components will increase. The use of 3th party services will reduce the developing time about 50 %.
- Open source: The relevance of open source software and hardware will increase. About 50% of software will be available with its source code under relaxed or non-existent copyright restrictions.

Having this in mind the question arises, in which way these advancements can be useful for sport or sport science? The example of expert systems in sport showed in the middle of the 1990s, that even well-intentioned ideas sometimes fail to succeed in the sport market. The approach in this context was to integrate knowledge of from sport practitioners and sport scientists into computer systems for training diagnosis and intervention. Even if some example for such experts systems in sport were developed (Lee & Kim, 1992; Pizzinato et al., 1998; Zeleznikow et al., 2009), a resounding success of the concept is still missing. The reasons for this are the missing quality of technological implementation (e.g. failing pattern

and speech recognition, useless statistical reporting) on the one hand, as well as missing demand from coaches and athletes on the other.
This example shows that it is quite difficult to appraise, which scientific concept can hold its ground in the dichotomy of sport market pull and technology push. Nevertheless, the following prospects P1 up to P10 try to give an optimistic outlook, how information and communication technology will be used in sport related training, competition and education in the year 2020. This "Forecast Sport Informatics 2020" considers open issues and demands from sport and sport science and the technological developments mentioned above (publications in brackets show first approaches or technological overviews):

P1 Positioning systems and lightweight sensors can be used for the capturing of total physiological and positional data (Eskofier et al., 2008). This information will be provided to trainers, athletes and mass media in real time.

P2 Based on the enhancements in the field of artificial intelligence, tactical behavior in sport games can be analyzed automatically (Atkinson & Rojasa, 2009).

P3 Virtual environments can simulate many different sports close to reality (Bideau et al., 2004). This allows a training of perception, cognition and decision in sport specific situations.

P4 New display technologies will be integrated in sport clothes (e.g. sun glasses) (Rolland & Cakmakci, 2009) and provide athletes with information during training and competition.

P5 Sport uses smart clothes including sensory (measuring and transmitting of data) and actuator (adaption of material characteristics) features for diagnosis, prevention and performance improvement (Borges et al., 2008).

P6 Biomechanical models can simulate human movements perfectly, without using motion capturing or key frames (Natural Motion, 2006). This allows the animation of virtual characters (e.g. for computer games), which look and move exactly like humans (Horswill, 2008). These models can also be used to make the motor behavior of humanoid robots more manlike.

P7 Simulation has become an important research tool for natural scientific disciplines. The simulation of training processes (e.g. performance adaption) can replace the real experiment in many cases (Ganter, Witte & Edelmann-Nusser, 2006; Churchill, Sharma & Balachandran, 2009).

P8 The increased computational power and cloud computing services allow simulating the vibration, flow and gliding of sport devices and athletes more realistic than today (Levy & Katz, 2007).

P9 The relevance of sport clubs as a local organizer of training sessions decreases. Group building via social networks on the internet gets more important for everyday sportspersons.

P10 In-class lectures at universities will lose their importance. Almost 50 % of theoretical courses and seminars in sport science will be replaced by online equivalents.

The realization and implementation of these scenarios needs both: the development and adaptation of technologies as well as the mutual exchange of sport science, computer science, sport federations and sport practitioners. Important factors for success are the added value, alternatives available, costs, usability and market acceptance of new innovations.

Conclusion

The paper has shown options for reasonable and fruitful liaisons between sport science and computer science. They hold a set of advantages for both disciplines, if projects are designed and performed with the view on genuine interdisciplinary research. As scientific progress in this area is closely connected to technological progress, sport sciences is well advised to monitor developments and to integrate partners from computer science into own research activities. Important technological improvements will be increased computing power and network capacity, networks concepts such as ubiquitous or pervasive computing, small and cheap one way electronics, thin and flexible displays as well as better speech recognition and geo-spatial positioning technologies. These technologies have a potential impact e.g. on the measurement and reporting of physiological and positional data, on computerized performance analysis, sport clothing and the quality of computer simulations in the fields of sports engineering, motor behavior and physiological adaptation. Certainly, it is the main work task for the computer-science-in-sport-community to make sure, that a check-up of the given forecast in the year 2020 - maybe on a symposium of the International Association on Computer Science in Sport - will be as successful as possible.

References

Atkinson, J. & Rojasa, D. (2009). On-the-fly generation of multi-robot team formation strategies based on game conditions. *Expert Systems with Applications*, 36(3), 6082-6090.

Baca, A. (2006). Computer science in sport: An overview of history, present fields and future applications (Part I). *International Journal of Computer Science in Sport, 5* (Special edition 2), 25-35.

Bächle, F. (2003). The optimisation of throwingmovements with evolutionary algorithms on the basis of multi-body systems. *International Journal of Computer Science in Sport (Special Edition 1)*, 6-11.

Beetz, M., Kirchlechner, B. & Lames, M. (2005). Computerized Real-Time Analysis of Football Games. *IEEE Pervasive Computing*, 4 (3), 33-39.

Bideau, B., Multon, F., Kulpa, R., Fradet, L. & Arnaldi, B. (2004). Virtual reality applied to sports: do handball goalkeepers react realistically to simulated synthetic opponents?. *Proceedings of the 2004 ACM SIGGRAPH International Conference on Virtual Reality Continuum and Its Applications in Industry* (pp. 210-216). New York.

BISp (2008). German Federal Institute of Sport Science. Press release from 2008-04-19. Access from http://www.bisp.de (welcome page) (19.05.2008).

Boghossian, B. & Black, J. (2005). The challenges of robust 24/7 video surveillance systems. *IEE International Symposium on Imaging for Crime Detection and Prevention (ICDP 2005)*, 33-38.

Boisot, M. (1972). Discipline and interdisciplinarity. In OECD (Ed.), *Interdisciplinarity: Problems of Teaching and Research in Universities* (pp. 89-97). Centre for Educational Research and Innovation (CERI). Paris: OECD Publications.

Borges, L. M., Rente, A., Velez, F.J., Salvado, L.R. Lebres, A.S., Oliveira, J.M., Araújo, P. & Ferro, J. (2008). Overview of Progress in Smart-Clothing Project for Health Monitoring and Sport Applications. *Proceeding of ISABEL 2008 - First International*

Symposium on Applied Sciences in Biomedical and Communication Technologies. Aalborg, Denmark.

Cialdini, R.B., Borden, R.J., Thorne, A., Walker, M.R., Freeman, S. & Sloan, L.R. (1976). Basking in reflected glory: Three (football) field studies. *Journal of Personality and Social Psychology, 34*, 366-375.

Chettiparamb, A. (2007). *Interdisciplinarity: a literature review*. The Interdisciplinary Teaching and Learning Group, University of Southampton. access from http://www.heacademy.ac.uk/ourwork/networks/itlg (15.05.2008).

Churchill, T., Sharma, D. & Balachandran, B.M. (2009). Correlation of novel training load and performance metrics in elite cyclists. *Proceedings IACSS09, 7th International Symposium of the International Association of Computer Science in Sport* (pp. 24-34). Canberra, Australia.

Claus, V. (1975). *Einführung in die Informatik* (Introduction to “Informatik”). Stuttgart: Teubner.

DFG (2008). German Research Foundation. Founding Policies. Access from http://www.dfg.de (welcome page) (19.05.2008).

Dikaiakos, M.D., Katsaros, D., Pallis, G., Vakali, A. & Mehra, P. (Eds.) (2009). IEEE Computing - Special Issue on "Cloud Computing", *IEEE Computing*, 12(5).

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

Engesser, H. (Ed.) (1988). *Duden „Informatik“: ein Sachlexikon für Studium und Praxis*, Stichwort Informatik (Encyclopaedia „Informatik“, Keyword „Informatik“). Mannheim: Dudenverlag.

Eskofier, B., Hartmann, E., Kühner, P., Griffin, J., Schlarb, H. Schmitt, M. & Hornegger, J. (2008). Real time surveying and monitoring of Athletes Using Mobile Phones and GPS. *International Journal of Computer Science in Sports*, 7(1), 18–27.

Fischer, G. (1998). Transcending Cultures - Creating a Shared Understanding between Computer Science and Sport. In J. Mester & J. Perl (Eds.), *Sport und Informatik V* (pp. 43-52). Köln: Strauss.

Ganter, N., Witte, K., & Edelmann-Nusser, J. (2006). Performance prediction in cycling using antagonistic models. *International Journal of Computer Science in Sport*, 5(2), 56-59.

Gelsinger, P. (2008). Opening keynote of Pat Gelsinger (Senior Vice President Intel). Intel Developer Forum (IDF) Shanghai 2008.

Grupe, O. (1971). Sportwissenschaft. Aufgaben und Probleme einer neuen Disziplin (Sport science. Tasks and Problems of a New Discipline). *Attempto*, 47/48, 20-26.

Heckhausen, H. (1972). Discipline and Interdisciplinarity. In OECD (Ed.), *Interdisciplinarity: Problems of Teaching and Research in Universities* (pp. 83-89). Centre for Educational Research and Innovation (CERI). Paris: OECD Publications.

Höner, O. (2001). Interdisziplinäre Theorienbildung als Leitorientierung für den sportwissenschaftlichen Nachwuchs(?) (Is Interdisciplinary a Profitable Research Paradigm for Young Scientists?). *Zephir*, 8(1), 16-29.

Horswill, I. (2008). Lightweight Procedural Animation with Believable Physical Interactions. *IEEE Transactions on Computational Intelligence and AI in Games*, 1(1), 39-49.

Hoyningen-Huene, N. & Beetz, M. (2009). Rao-Blackwellized Resampling Particle Filter for Real-Time Player Tracking in Sports. In A. Ranchordas & H. Araujo (Eds.), *Proceeding of Fourth International Conference on Computer Vision Theory and Applications (VISAPP)* (pp. 464-470). Lisboa, Portugal.

Hughes, M. (2000). Computers and Performance Analysis in Sport. In A. Baca (Ed.), *Computer Science in Sport* (pp. 83-102). Wien:öbv & hpt.

Hughes, M. D., Dawkins, N. & David, R. (2000). Perturbation effect in soccer. Notational Analysis of Sport III, UWIC, 1-14.

ITU (2005). *The Internet of Things*. International Telecommunication Union (ITU) Internet Reports 2005. Geneva.

Jörg, D. & Lames, M. (2009). Perturbationen im Tennis – Beobachtbarkeit und Stabilität (Perturbations in Tennis – Observability and Stability). In M. Lames, C. Augste, O. Cordes, Ch. Dreckmann, K. Görsdorf & M. Siegle (Eds.), *Gegenstand und Anwendungsfelder der Sportinformatik. 7. Symposium der dvs-Sektion Sportinformatik vom 22.-24. Mai 2008 in Augsburg* (pp. 86-96). Hamburg: Czwalina.

Karlqvist, A. (1999). Going Beyond Disciplines: The Meanings of Interdisciplinarity. *Policy Sciences*. 32 (4), 379-383.

Kjeldskov, J. & Stage, J. (2004). New techniques for usability evaluation of mobile systems. *International Journal of Human-Computer Studies*, 60(5-6), 599-620.

Kurzweil, R. (2005). *The Singularity Is Near: When Humans Transcend Biology*. New York: Viking Books.

Lamb, P., Bartlett, R. & Robins, A. (2009). Self-Organizing Maps: An Objective Method for Clustering Complex Human Movements. *Proceedings IACSS09, 7th International Symposium of the International Association of Computer Science in Sport* (pp. 100-108). Canberra, Australia.

Lames, M. (1999). Football: A Chaos Game? Applying elements of chaos theory to the genesis of goals in football. In P. Parisi, F. Pigozzi & G. Prinzi (Eds.), *Sport Science ´99 in Europe* (p. 543). Rome: Rome University.

Lames, M. (2008). Coaching and Computer Science. In P.Dabnichki & A. Baca (Eds.), *Computers in Sport* (S. 99-120). WITPress: Southhampton.

Latombe, J.C. (1991). *Robot Motion Planning* (Chapter 1). Boston: Kluwer Academic Publishers.

Lee, J.-W. & Kim, K.-J. (1992). Development of an expert system for diagnosis and training prescription of cardiorespiratory functions in athletes. In G. Tenenbaum, T. Raz-Liebermann & Z. Artzi (Eds.), *Proceeding of the international conference on computer applications in sport and physical education* (pp. 277-284). Netanya, Wingate Institute, The Zinman College.

Levy, R. & Katz, L. (2007). Virtual Reality Simulation: Bobsled and Luge. IACSS07, *6th International Symposium of the International Association of Computer Science in Sport* (pp. 241-251). Calgary: Canada.

Link, D. & Lames, M. (2005). Effects of Computer Mediated Communication on the Quality of Beach Volleyball Coaching Sessions. In F. Seifriz, J. Mester, J. Perl, O. Spaniol & J. Wiemeyer (Eds.), *1st International Working Conference IT and sport and 5th Conference dvs-Section Computer Science in sport - Book of Abstracts* (p. 172-176). Cologne: Cologne Sports University.

Link, D. & Lames, M. (2006). Handhelds in der wissenschaftlichen Trainingsberatung (Handhelds for Coaching). In K. Witte, J. Edelmann-Nusser, A. Sabo & E.F. Moritz (Hrsg.), *Sporttechnologie zwischen Theorie und Praxis IV (S. 199-208).* Aachen: Shaker Verlag.

Luft, A. L. (1992). Grundlagen einer Theorie der Informatik – „Wissen" und „Information" bei einer Sichtweise der Informatik als Wissenstechnik (Basics of a theory of "Informatik"). In W. Coy, F. Nake, J.-M. Pflüger, A. Rolf, J. Seetzen & R. Stransfeld

(Eds.), *Sichtweisen der Informatik* (pp. 49-70). Braunschweig/Wiesbaden: Vieweg-Verlag.

McGarry, T., Anderson, D. I, Wallace, S. A. , Hughes, M. D. & Franks, I. M. (2002). Sport Competition as a dynamical self-organizing system. *Journal of Sports Sciences*, 20, 771-781.

Miethling, W.-D. & Perl, J. (1981). *Computerunterstützte Sportspielanalyse (*Computer supported notational analysis*).* Ahrensburg: Czwalina.

Müller, H. & Sternad, D. (2004). Decomposition of variability in the execution of goal-oriented tasks: Three components of skill improvement. *Journal of Experimental Psychology: Human Perception and Performance*, 30, 212-233.

National Research Council (2004). *Computer science: Reflections on the Field, Reflections from the Field*. National Academic Press.

Natural Motion. (2006). *Euphoria:core motion synthesis library*. Natural Motion Inc.

Novatchkov, H., Kornfeind, Ph., Bichler, S. & Baca, A. (2009). Mobile Coaching. *Proceedings IACSS09, 7th International Symposium of the International Association of Computer Science in Sport* (p. 145). Canberra, Australia.

Nygaard, K. (1986). Programm Development as a Social Activity. In H.J. Kugler (Ed.), *Information Processing 86* (Proc. 10th IFIP World Computer Congress '86, Dublin) (pp. 189-198). Amsterdam: Elsevier Science Publishers B.V.

OECD (Ed.)(1972). Interdisciplinarity: Problems of Teaching and Research in Universities. Centre for Educational Research and Innovation (CERI). Paris: OECD Publications.

Oertzen, J., Cuhls, K. & Kimpeler, S. (2006). Wie nutzen wir Informations- und Kommunikationstechnologien im Jahr 2020? Ergebnisse einer Delphi-Befragung (How Do we Use Information and Communication Technologies in the Year 2020, Result of a Delphi Survey), *FAZIT-Schriftenreihe, Band 3*, Stuttgart, access from http://www.fazitforschung.de (15.05.2008).

Perl, J. & Lames, M. (1995). Sportinformatik: Gegenstandsbereich und Perspektiven einer sportwissenschaftlichen Teildisziplin (Sport Informatics Subject Matters and Perspectives of a Sport Scientific Sub Discipline). *Leistungssport*, 25 (3), 26-30.

Perl, J., Lames, M. & Miethling, W. (Hrsg.). (1997). *Informatik im Sport. Ein Handbuch* (Computer Science in Sport*).* Schorndorf: Hofmann.

Pfeiffer, M. & Perl, J. (2006). Analysis of Tactical Structures in Team Handball by Means of Artificial Neural Networks. *International Journal of Computer Science in Sport*, 5 (1), 4-14.

Pizzinato, A., Denis, G., Vachon, G. & Kohler, F. (1998). Knowledge based system for the simulation of decision-making of the serve return phase in tennis: ”LIFT“-system. In J. Mester & J. Perl (Hrsg.), *Sport und Informatik V (pp. 141–145)*. Köln: Strauss.

Recla, J. & Timmer, R. (Eds.)(1976). *Kreative Sportinformatik. (Creative Sport Informatics). Der Internationale Jubiläums- Kongreß 1975 in Graz*. Schorndorf: Hoffmann.

Ries, H. & Kriese, H. (1974). Scientific model for a theory of physical education and sport sciences. In U. Simri (Ed.), *Concepts of physical education and sport sciences* (pp. 175-198). Natanya.

Rolland, J. & Cakmakci, O. (2009). Head-Worn Displays: The Future Through New Eyes, Optics & Photonics News, 20, 20-27.

Röthig, P. & Prohl, R. (Eds.) (2003). *Sportwissenschaftliches Lexikon*, Stichwort Sport (Encyclopaedia of sport science, keyword sport). Schorndorf: Hofmann.

Surowiecki, J. (2004). *The wisdom of crowds. Why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations.* New York: Doubleday.

Turner, T. (2009). A WI-FI Based Embedded System for Biomechanical Data Telemetry. Proceedings IACSS09 (pp. 197-202). 7th International Symposium of the International Association of Computer Science in Sport. Canberra, Australia.

Vajna, S., Edelmann-Nusser, J., Kittel, K. & Jordan, A. (2006). Optimisation of a bow riser using the Autogenetic Design Theory. In I. Horváth & J. Duhovnik (Eds.), *Tools and Methods of Competitive Engineering Vol. 1* (pp. 593-601). Ljubljana: University of Ljubljana.

Walter, F., Lames, M. & McGarry, T. (2007). Analysis of Sports Performance as a Dynamical System by Means of the Relative Phase. *International Journal of Computer Science in Sports*, 6 (2), 35-41.

Wiemeyer, J. (2009). Digitale Spiele – (k)ein Thema für die Sportwissenschaft?! (Digital Games – Interesting or Not-interesting for Sport Science?). *Sportwissenschaft,* 39 (2), 120-128.

Zeleznikow, J., MacMahon, C. & Barnett, T. (2009). Providing automated decision support for elite athletes. *Proceedings IACSS09, 7th International Symposium of the International Association of Computer Science in Sport* (pp. 240-248). Canberra, Australia.

Enforcers, Grinders and Skilled Ice Hockey Players: Empirical Test from a Computer Simulator

Richard G. Lomax

The Ohio State University

Abstract

Ice hockey players have long been given different monikers to describe styles of play (e.g., "enforcers," "grinders," "skilled players"). The purpose of this paper was to empirically determine whether there is any basis for these categories using the APBA (American Professional Baseball Association) computer simulator for playing the 2003-2004 NHL season. Forwards and defensemen were analyzed separately, while goaltenders were not considered. Variables were initially standardized separately by position to assist interpretation of the resultant clusters. Meaningful clusters of players were generated based both on interpretability and sample size. Cluster centers were examined and descriptors were attached to each cluster depending upon relative variable importance. The descriptors were then used to assign subjective labels. There were seven forward clusters: (1) skilled attackers; (2) second-tier skilled attackers; (3) grinders; (4) enforcers; (5) least effective; (6) defensive specialists; and (7) face-off specialists. There were six defensemen clusters: (1) skilled attackers; (2) second-tier skilled attackers; (3) defensive defensemen; (4) serviceable defensemen; (5) most intimidating defensemen; and (6) least effective defensemen. From these descriptors and the specific players contained in each cluster, it was determined that there is good empirical support for the use of labels to classify hockey players from the computer simulator.

KEYWORDS: ICE HOCKEY, CLUSTER ANALYSIS, COMPUTER SIMULATION

Introduction

Over the years, ice hockey players have been given different monikers to describe different styles of play. For example, tough-nosed players often engage in hard-hitting or fighting-type behaviors, either in coming to the defense of their teammates or to try to alter the game for their team. Such players are often known by terms such as "enforcers". Other players are known as "skilled players" who are highly skilled (e.g., at passing, skating, and shooting) and play on the first line (forward) or pairing (defense) of the team. Some forwards are excellent at defending against their opponents' more skilled players; these players are known as "defensive specialists". Another type of player is known as a "grinder" whose role is to outwork the other team, particularly by being physical and hard-working.

All of these terms are in common usage by hockey coaches, general managers, scouts, the media, players, fans, and fantasy hockey participants alike. Thus when a player is described as an "enforcer", people involved with hockey in any capacity immediately know the type of player that you are talking about, and can even cite the names of some stereotypical players in that category. The purpose of this paper was to empirically determine whether there is any basis for these labels or categories using a computer simulator.

Unfortunately, until recently data sources for hockey players were restricted in terms of the number of variables collected. Basic player statistics such as goals, assists, games played, penalty minutes, goals against, saves, and the like were the only variables available to researchers. Thus, research opportunities in hockey were rather limited given the few number of variables available. Hockey research tended to be focused either on historical methods of inquiry (e.g., Wong, 2005, studied the early development of the National Hockey League <NHL>), or on summative methods using either basic statistical techniques and/or team-based variables (e.g., Lock, 1997, developed a model to predict final scores).

A plethora of variables has long been available in baseball. This has fostered much research on the game of baseball, leading to the institution of the SABR (Society for American Baseball Research) association, as well as the development of numerous SABR-metric statistics. Largely as a result of these activities, baseball research has had a long and rich quantitative history in terms of methodological sophistication.

Recently two phenomena have occurred in hockey to make more comprehensive inquiry possible. First, professional hockey began to systematically collect data on many player variables during games in real time using sophisticated software (Bruno, 1999). This allowed anyone to go online during or after a game and examine variables at the player level, such as time on ice, number of shifts played, number of shots taken, faceoffs won, giveaways, blocked shots, and hits given.

The second development was the proliferation of fantasy hockey (as well as other fantasy sports) (Lomax, 2006). Software has been developed to simulate entire hockey games, subsequently generating complete game reports and detailed statistics. In brief, these software programs require the use of player rating variables derived from actual performance. Thus players are rated on variables such as skating, passing, defensive skill, shooting ability, and checking. In other words, based on player performance in actual games, fantasy hockey commissioners across the globe get together, rate each player on many variables, and then develop a set of player ratings that all fantasy leagues can utilize. In each fantasy league, general managers assemble their own team of players, determine a particular line-up of players, and then input that information into the computer simulator which uses the player ratings to play the simulated games.

With the many variables that are now available to hockey researchers, multivariate procedures are nicely suited to take a large number of variables and reduce that information into something more manageable and meaningful. Two example multivariate studies include Williams (1994) and Williams and Williams (1998), who used principal components analysis to derive performance indices from on-ice statistics.

In this paper we consider the use of cluster analysis, whose purpose is to use a large number of variables to develop meaningful clusters of individuals who are most similar on those variables. Thus players who fall into a particular cluster will be those that are most alike with respect to the entire set of variables. More specifically, cluster analysis was used to determine the extent to which these commonly used ice hockey player labels have an empirical basis using a computer simulator. A previous pilot study with an older dataset was deemed to be promising for this more complete study (Lomax, 2001).

Methods

The database used in the present study consisted of the fantasy hockey ratings from the APBA simulator (originally the American Professional Baseball Association before becoming a company with simulators developed for many sports) for playing the 2003-2004 NHL season.

That is, the ratings were based on actual player performance from that NHL season as determined by a committee of fantasy hockey commissioners. APBA is one of several ice hockey simulators available; others include Face Off, HLS (Hockey League Simulator), and FHL (Fantasy Hockey League). Depending on the simulator, players are typically rated on each variable on scales from 0 to 5, or 0 to 9, or 0 to 99. Because every player who played during that season was rated on every variable, there was no missing data that required treatment.

Forwards (n = 350) and defensemen (n = 208) were analyzed separately because variable importance varies by position and because the monikers by which players are known are also position specific. Goaltenders were not considered for two reasons. First, goalies have a much smaller sample size than forwards and defensemen, and second because far fewer rating variables are available.

The K-means cluster analysis method of SPSS was used to analyze the data. All variables were initially standardized separately by position to assist interpretation of the resultant clusters. The penalty rating variable was normalized by a square root transformation (due to excessive skewness and kurtosis values). The following variables were included in the analysis (with possible ranges of raw scores on the rating measures): defense rating at their position (1-5), number of minutes played, shot percentage, number of shots on goal, shot on goal frequency, shot wide frequency, shot deflected frequency, shot defended frequency, skating (0-5), speed (0-5), passing (0-5), forechecking (forwards only) (0-5), physical game (0-5), throws crushing body checks (0-5), faceoff success (forwards only) (0-2), defensive clear of opponents and rebounds in front of their own goal (defensemen only) (0-2), shot blocking (defensemen only) (0-2), injury rating (0-99), assist rating (0-99), penalty rating (0-99), and overall rating (0-160). Each of the variables was found to be a significant contributor to the generation of the clusters ($p < .001$), for both the forwards analysis ($F_{6,343}$ ranged from 4.867 to 175.497 across the variables) and the defensemen analysis ($F_{5,202}$ ranged from 7.887 to 107.853 across the variables), based on comparisons of the cluster centers.

Results

Initially differing numbers of clusters were generated (i.e., 3, 4, 5, 6, and 7 cluster solutions) by position. After reviewing those initial solutions, final meaningful clusters of players were selected for each position. These final cluster solutions were based both on interpretability (i.e., wanting clusters that made sense substantively) and on sample size (i.e., not wanting too many small clusters or too few large clusters). In other words, for more clusters the sample sizes were quite small given the size of the NHL (e.g., one initial forward cluster solution had 3 players), and for fewer clusters interpretability of the clusters suffered as it was clear when two clusters were forced to be grouped together (e.g., another initial forward cluster solution had 120 players out of a total of 350 forwards, and thus it was not very homogeneous). The final cluster centers were examined (as shown in Tables 1 and 2 for the forwards and defensemen, respectively) and then compared across clusters to interpret the strengths and limitations of each cluster. That is, each individual cluster was compared to other clusters and ranked as highest (largest center of all clusters), second highest, second lowest, and lowest. Descriptors were attached to each cluster depending upon the relative importance of the variables. The descriptors were then used as the basis for assigning commonly used subjective labels. For each cluster described below, the resultant sample size and some example player names are listed. The NHL player names listed are also considered by ice hockey experts to be stereotypical names for those particular clusters, thus validating the subjectively-based player labels.

Table 1: Forward Cluster Centers

Variable / Cluster	1	2	3	4	5	6	7
Overall rating	1.73	0.23	-1.34	0.04	-.80	0.50	-0.41
Defense rating at position	0.07	-.42	-0.59	0.26	-.67	0.90	0.52
# of minutes played	1.23	0.41	-1.50	-.08	-.67	0.57	-0.20
Shot percentage	0.72	-.21	-0.85	1.10	-.38	-.16	0.13
# of shots on goal	1.83	0.00	-0.91	0.13	-.43	0.11	-0.64
Shot on goal frequency	0.76	0.73	-0.54	-.57	0.28	0.13	-1.43
Shot wide frequency	-.72	-.67	0.42	0.62	-.31	-.20	1.53
Shot deflected frequency	-.73	-.64	0.42	0.30	-.24	-.20	1.69
Shot defended frequency	-.80	-.65	0.59	0.59	-.24	-.11	1.21
Skating	1.16	0.68	-1.07	-.29	-.69	0.33	-0.34
Speed	0.91	0.70	-1.02	-.46	-.30	0.27	-0.49
Passing	1.34	0.43	-1.47	-.32	-.27	0.21	-0.22
Forechecking	0.14	-.37	-0.30	0.53	-.80	0.90	0.04
Physical game	-.20	-.74	1.17	0.47	-.56	0.59	-0.30
Throws crushing body checks	-.18	-.71	1.41	0.43	-.60	0.49	-0.35
Faceoffs	0.04	-.39	-0.60	-.09	-.25	0.59	0.69
Injury rating	-.38	-.16	0.45	0.15	0.56	-.40	-0.20
Assist rating	1.38	0.58	-1.21	-.33	-.52	0.27	-0.40
Square root penalty rating	-.05	-.35	-0.20	0.51	-.01	0.32	-0.19

There were seven meaningful clusters of forwards as follows:

(1) Skilled attackers – highest: overall rating, minutes played, shots on goal, shot on goal frequency, skating, speed, passing, assists; second highest: shooting percentage; second lowest: injury; lowest: shots wide, shots deflected, shots defended (*n* = 42; e.g., Alfredsson, Datsyuk, Forsberg, Jagr, Sakic). Thus these players are highly skilled, durable, strong shooters, heavily involved in goal scoring (goals and assists), and thus are likely to play on the first forward line and on the powerplay.

(2) Second-tier skilled attackers – second highest: shot on goal frequency, skating, speed, passing, assists; second lowest: shots wide, shots deflected, shots defended, forechecking, faceoffs; lowest: physical game, crushing body checks, penalty rating (*n* = 61; e.g., Amonte, Daze, Erat, Friesen, Smolinski). This type of player is quite skilled, a good shooter, not physical, stays out of the penalty box, very involved in goal scoring, and probably plays on the second forward line.

(3) Grinders – highest: physical game, crushing body checks; second highest: shot wide frequency, shots deflected, shots defended, injury; second lowest: defense; lowest: overall, minutes played, skating, speed, passing, shooting percentage, shots on goal, faceoffs, assists (*n* = 39; e.g., Brashear, Domi, Laraque, May, Simon). These are very physical players who are not very good shooters, injury prone, not good defensively, not very skilled, play very few minutes, not much involved in goal scoring, and thus likely to play on the third forward line.

(4) Enforcers – highest: penalty rating, shooting percentage; second highest: shots wide, shots defended, forechecking; second lowest: shot on goal frequency, passing (*n* = 40; e.g., Barnaby, Maltby, Nichol, Petrovicky, T. Stevenson). Thus these players are very

penalty prone, good forecheckers, can score but rarely put the puck on goal, not very skilled, and thus probably only play when needed to spark their team or to intimidate their opponent.

(5) Least effective – highest: injury; second lowest: overall rating, minutes played, shooting percentage, skating, physical game, crushing body checks, assists; lowest: defense, forechecking (*n* = 64; e.g., Barney, Hagman, D. Roy, Savage, Svoboda). This type of player is very injury prone, not very good defensively, plays few minutes, not very skilled, not very physical, and likely to play on the fourth forward line.

(6) Defensive specialists – highest: defense, forechecking; second highest: overall rating, minutes played, physical game, crushing body checks, faceoffs, penalty rating; lowest: injury (*n* = 63; e.g., Bates, Brind'Amour, Madden, Peca, K. Primeau). These specialists are very strong defensively, somewhat physical, good at faceoffs, take some penalties, play many minutes, not injury prone, and likely to play opposite their opponents' first or second forward line.

(7) Face-off specialists – highest: faceoffs, shots wide, shots deflected, shots defended; second highest: defense; second lowest: shots on goal, speed, penalty rating; lowest: shot on goal frequency (*n* = 41; e.g., Andreychuk, Francis, Linden, McLean, Oates). This type of player is excellent at faceoffs, not a very good shooter, good defensively, not a very fast skater, stays out of the penalty box, and probably plays at the center position and in critical game situations.

Table 2:Defensemen Cluster Centers

Variable / Cluster	1	2	3	4	5	6
Overall rating	2.34	0.46	0.19	0.17	-1.14	-.94
Defense rating at position	1.67	-0.44	0.62	-.14	-0.61	-.79
# of minutes played	1.49	0.72	0.01	0.52	-1.33	-.81
Shot percentage	0.89	0.93	-.28	0.21	-1.05	-.05
# of shots on goal	1.33	1.51	-.38	-.03	-0.64	-.39
Shot on goal frequency	0.94	1.29	-.69	0.14	-0.87	0.34
Shot wide frequency	-1.11	-1.15	0.64	-.27	1.14	-.31
Shot deflected frequency	-1.12	-1.17	0.68	-.29	1.06	-.27
Shot defended frequency	-0.68	-1.17	0.63	0.02	0.50	-.38
Skating	1.04	0.62	-.34	0.79	-1.26	-.37
Speed	0.92	0.80	-.36	0.59	-1.27	-.17
Passing	1.08	1.05	-.37	0.48	-1.38	-.17
Physical game	0.72	-0.53	0.65	-.49	0.78	-.89
Throws crushing body checks	1.11	-0.65	0.57	-.63	0.92	-.74
Defensive clear	1.20	-0.46	0.71	-.51	0.30	-.88
Shot blocking	0.09	-0.08	0.46	0.15	-0.50	-.59
Injury rating	-0.56	-0.54	-.39	-.38	0.85	1.10
Assist rating	1.18	1.41	-.51	0.41	-1.09	-.30
Square root penalty rating	0.07	-0.20	0.03	-.57	1.26	-.05

For the defensemen, there were six meaningful clusters as follows:

(1) Skilled attackers – highest: overall rating, defense, minutes played, skating, crushing body checks, defensive clears, speed, passing; second highest: shooting percentage, shots on goal, shot on goal frequency, physical game, assists; second lowest: shots wide,

shots deflected, shots defended; lowest: injury (*n* = 14; e.g., R. Blake, Foote, Lidstrom, S. Niedermayer, Pronger). This type of player is characterized by high skills, strong defensively, very durable, somewhat physical play, a strong shot, and thus likely to play on the first defensive pairing and on the powerplay.

(2) Second-tier skilled attackers – highest: shot percentage, shots on goal, shot on goal frequency, assists; second highest: overall rating, minutes played, speed, passing; second lowest: physical game, crushing body checks, injury, penalty rating; lowest: shots wide, shots deflected, shots defended (*n* = 24; e.g., Berard, Gonchar, Kubina, Mara, Stuart). Players in this category are quite skilled, good shooters, fairly durable, not very physical, and probably play on the second defensive pairing.

(3) Defensive defensemen – highest: shots defended, blocks; second highest: defense, shots wide, shots deflected, defensive clears; second lowest: shooting percentage, shots on goal, shot on goal frequency, speed, passing, assists (*n* = 62, e.g., Chelios, deVries, Sweeney, Weinrich, Wesley). Thus these players are very defensively oriented, not very skilled, not very good shooters, and likely to play on the penalty killing unit.

(4) Serviceable defensemen – second highest: skating, blocks; second lowest: crushing body checks, defensive clears; lowest: penalty rating; average on everything else (*n* = 46; e.g., Desjardins, Hedican, Malakhov, Niinimaa, Spacek). Players of this type are good skaters and shot blockers, not physical, stay out of the penalty box, not particularly skilled, and likely to play on the third defensive pairing.

(5) Most intimidating defensemen – highest: shots wide, shots deflected, physical play, penalty; second highest: shots defended, crushing body checks, injury rating; lowest or second lowest on everything else (*n* = 24, e.g., Boughner, Commodore, Exelby, Marchment, Purinton). This cluster represents players who are physical, frequently take penalties, often injured, weak shooters, not very skilled or defensive-minded, and would only see spot duty each game.

(6) Least effective defensemen – highest: injury; second lowest: overall rating, minutes played, shots on goal, skating; lowest: defense, physical play, crushing body checks, defensive clears, blocks (*n* = 38, e.g., Albelin, Delmore, Karpovtsev, Quint, Rivers). This type of player is injury prone, plays infrequently, not very good defensively, not very skilled, not physical, and likely to see limited play at the NHL level.

Discussion

In this study cluster analyses of the 2003-2004 NHL player ratings data were performed using the APBA computer simulator separately for forwards and defensemen. The resulting clusters and a comparison of the cluster centers were then used to generate a descriptor for each cluster. From these descriptors and the players contained in each cluster, it was determined that there is good empirical support for the use of labels to classify professional ice hockey players. In other words, the monikers that hockey experts attach to particular players was empirically supported through the cluster analysis of this sample. This is one indicator that the APBA ratings are related to actual NHL player performance. Obviously, based on a single study with one season’s worth of data, replication of these results would be important to achieve broader generalization. In addition, similar analyses could be conducted for other eras of hockey to determine if these or other clusters are appropriate (e.g., particularly as the game of hockey has developed over time due to changes in rules, player size, player speed, and the economics of the game). It would also be interesting to see the extent to which these results could be replicated with other hockey computer simulators.

In summary, we were able to show that the largely subjectively derived categories of ice hockey players do have an empirical basis using the APBA computer simulator. The empirically-based clusters were distinct, meaningful, and mapped nicely onto the subjective clusters commonly used among ice hockey specialists (e.g., enforcers, grinders, and skilled players).

References

Bruno, P. (1999). Hockey statistics: The new wave of numbers in the game. *The Hockey Research Journal, 4*, 59-62.

Lock, R.H. (1997). *Using a Poisson model to rate teams and predict scores in ice hockey*. Paper presented at the Joint Statistical Meetings of the American Statistical Association, Anaheim, CA.

Lomax, R.G. (2001). Using statistical methods to evaluate hockey players. *The Hockey Research Journal, 5*, 43-45.

Lomax, R.G. (2006). Fantasy sports: History, game types, and research. In A. A. Raney & J. Bryant (Eds.), *Handbook of Sports and Media* (pp. 383-392). Mahwah, NJ: Lawrence Erlbaum Associates.

Williams, W.H. (1994). Performance indices for on-ice hockey statistics. *Proceedings of the Section on Statistics in Sports*, American Statistical Association, 50-54.

Williams, W.H. & Williams, D. (1998). Performance indices for multivariate ice hockey statistics. In J. Bennett (Ed.), *Statistics in Sport* (pp. 141-155). London: Arnold.

Wong, J.C.-K. (2005). *Lords of the Rinks: The Emergence of the National Hockey League, 1875-1936*. Toronto: University of Toronto Press.