Open Source Technologies Involved in Constructing a Web-Based Football Information System

Pedro Rodrigues[1], António Belguinha[1], Carlos Gomes[1], Pedro Cardoso[1], Tiago Vilas[1], Renato Mestre[2], and J.M.F. Rodrigues[1,3]

[1] Institute of Engineering, University of the Algarve, 8005-139 Faro, Portugal
[2] Inesting, S.A., 8005-146 Faro, Portugal
[3] Vision Laboratory, LARSyS, University of the Algarve, 8005-139 Faro, Portugal
{pcardoso,jrodrig}@ualg.pt

Abstract. The current information systems and match analysis software associated to professional football output a huge amount of statistics. Many football professionals are particularly interested in real time information about the tactical plan occurring during the match, and the relations between that information and what was prepared in the training sessions. It is fundamental to have on the bench, and on-the-fly, the most relevant information each time they have to take a decision. In this paper, we present a set of open source technologies involved in building a multi-platform web based integrated football information system, supported in three main modules: user interfaces, databases, and the tactical plan detection and classification. We show that the selected technologies are suitable for those modules, allowing field occurrences to trigger meaningful information.

Keywords: Football, open source, web technologies, information system.

1 Introduction

All historically renowned football teams have displayed a recognizable footprint in their gameplay, which has always been thought of as something observed by football experts rather than described by game statistics and scientific principles. Within the coaching process, great emphasis is placed on the coach's ability to observe and recall all the critical discrete incidents from a sport and tactical performance [1]. However, it has been shown that coaches cannot accurately observe and recall all of the detailed information that is required for a complete understanding or interpretation of the team performance [2].

Footdata is a project to build a new multi-platform product for football, based on technologies of information and communication systems, which integrates two fundamental components of the football world: i) A social network, which will have all the typical features and ii) the professional component, which features an acquisition and information system to meet all the football management needs. The latter component includes an automated platform to gather information from the teams, not only on the competition part but also on the preparation model (training

Á. Rocha et al. (Eds.): *Advances in Information Systems and Technologies*, AISC 206, pp. 715–723.
DOI: 10.1007/978-3-642-36981-0_66

sessions). This platform will be based on a prototype that will process images acquired live on-site (in matches or training sessions), or images from recorded or broadcasted matches. The processing of this information goes beyond the traditional statistical data compilations extracted from automated observation of the teams actions, adding features which will allow the analysis of the football match structure, namely in its two fundamental ways: width and depth of the team, allowing to rationalize and optimize the team's actions regarding occupation of spaces. However, the main goal is to automatically collect information which concerns the tactical plan, and on-the-fly alert the technical team for specific events. All the above should be presented in a web (browser) environment, accessible from personal computer or a mobile device (smartphone or tablet).

There are in the market commercial systems that integrate partially some of the components of this project, e.g. Kizanaro® [3], Tracab® [4], Prozone® [5], Amisco® [6], and SportVU® [7]. We can also find in the literature examples that explore some aspects of the project, like spatial and spatiotemporal analysis [8], local positioning systems in the game [9], and analysis of football strategies [10], [11] and [12].

In this paper we present a preview of what the system will be, and focus on the open source technologies to develop the professional component of Footdata(-PRO). Due to the limited size of the paper, we will not enter into full details about the reasons that led us select each specific technology. The main contribution is a compilation and discussion of the open source technologies necessary to develop a web-based football information system (IS).

The rest of the paper is organized as follows. Section 2 shows the main Footdata modules, Section 3 and 4 presents the user interface and the database technologies and in Section 5 the tactical plan detection and classification. Finally, we conclude the paper with the discussion and future work in Section 6.

2 Footdata-PRO Modules

Has mention in the Introduction, Footdata is a project to build a social network and the professional component (-PRO) in a web environment. Footdata-PRO features an acquisition and information system divided in five main modules: (a) Field, where the ball and players location in the pitch is done by computer vision (during the match or training sessions). This information is sent to a server through an internet connection, where the match tactical plans are verified, based on the information available from prepared tactical schemes, trainings, individual players, and collective information from each team. Based in the above procedure, in the coach's (user) profile and in what is occurring in real-time in the field (or in practice), selected specific information is sent to the coach's mobile device. A request can be sent to the server every time the coach requires extra information. (b) Center, where all the information and the main computation is congregated (including the processing and management of previously recorded videos). (c) Coach, where the coach's team inputs and accesses to detailed information of every topic related to the football match, scouting, medical department, etc. (d) Player, used to send specific information to the player, and the (e) Presentation, where semi-automatic presentations are prepared for the players in preparation moments.

An integrated web interface, where usability is a key factor, will combine all the above features in a single place, including video manipulation tools, tactical planning/drawing tools, semi-automatic presentation tools, etc. Figure 1 illustrates the Footdata-PRO main structure.

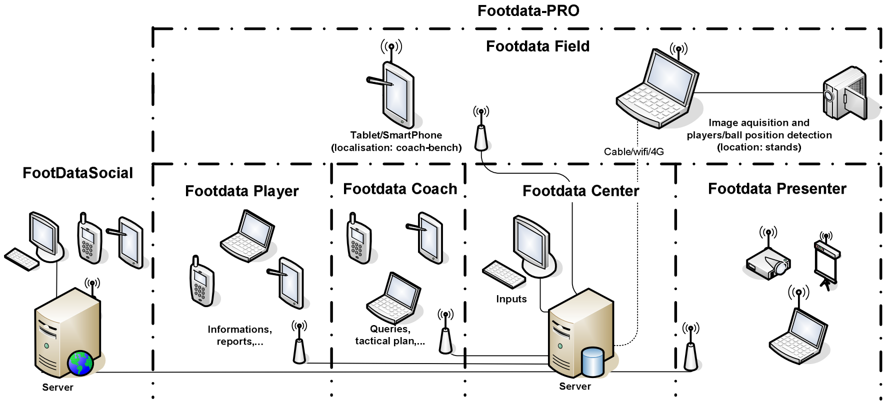

Fig. 1. Footdata-PRO main modules: Field, where the ball and players location in the pitch is done by computer vision; Center, where all the information and the main computation is congregated; Coach, where the coach inputs and accesses to detailed information; Player, used to send specific information to the players; and Presentation, where semi-automatic presentations are prepared for the players

3 User Interface Technologies

A prerequisite for the user interface was that it should be delivered as a web (browser) application. Furthermore, it should be a multi-platform system running in multiple operating systems (OS) with special emphasis to Windows, Linux, Mac, iOS, and Android. This requirement can be seen as an asset, as it is deliverable to a larger number of users, requiring only a web browser on a personal computer or on some other mobile device (phone or tablet). After the analysis of the possibilities, the decision of using HTML5 [13], [14] to its largest extension was made, maintaining the necessary fallbacks to web browser plugins or specific applications for the mobile devices. This decision was made based on two main assumptions: the majority of the HTML5 functionalities are, or soon will be, supported by the most used browsers (OS independent) and the will/need to use open source software to develop the applications (the use of open source software was a previous agreement done by all the project consortium members). Nevertheless, several difficulties are easy to find, as the embedding of video (see below).

The second phase, required the selection of a programming platform, in particular a web development framework. The choice was Django [15]. As their providers define it, Django is a high-level Python [16] web framework that encourages rapid development and clean, pragmatic design [15], which provides among others:

an object-relational mapper, elegant URL design, template system, cache system, and internationalization [17]. The option of using Django was also supported by the fact that the programming language behind it (Python) provides a very rich set functionalities on its standard library, including internet data handling, interprocess communication and networking, and concurrent execution. Furthermore, Python has also a large set of third-party extensions, in the form of modules, which extends the Python's standard library, as for example: connectors to the majority of the database systems (e.g., MySQL, PostgreSQL, Oracle, Microsoft SQL Server, MongoDB, CouchDB, etc. [18]), data analysis (e.g., Pandas, NumPy, SciPy [18]), and image processing and manipulation (e.g., OpenCV [19]).

One of the specifications made for the interface is to allow video manipulation, including the selection and displacement of a player (see Section 2). The solution planned to the video section, without using any third party plugins, will use HTML5 video support. However the current incompabilities between browsers/platforms and the video formats stand as an obstacle quite costly to solve. We have to keep in mind that legacy browsers/platforms may not support the HTML5 video element [20], this problem can addressed with the inclusion of fallbacks into plugins. By analysing the browsers supported video formats, we found out that a good format combination to work with would be MP4/WebM or MP4/Ogg. With one of those solutions, displaying compatibility is attained by the largest majority of the web browsers. Despite the format of the video being correct for a certain browser, we have to be careful with the video and audio codecs. In this case we would have the H.264 video with AAC audio codecs in a MP4 container (.mp4), VP8 video codec with Vorbis audio codec in a WebM container (.webm), and finally, Theora video codec with Vorbis audio codec in a Ogg container (.ogv) [21]. To discover if a particular browser supports a particular video format, libraries like the Modernizr Javascript Library [22] can be used. HTML5media, MediaElement.js and Video.js are also solutions to take into account in this matter.

To provide an efficient video manipulation, the native video controls that come in each browser seem to be insufficient. Fortunately, using JavaScript we can implement our own controls and design (with the help of CSS - Cascading Style Sheets) [20]. Different approaches can be found in [23], [24] where the HTML5 Video and Canvas are combined together. jQuery User Interface - jQueryUI [25] is also worth mentioning because besides being a tool for the controls design it can also solve some incompatibilities (e.g., at the time of writing, the HTML5 input range for the video seekbar, is not correctly rendered by Internet Explorer nor Firefox).

Figure 2 shows two examples of the interfaces: on the left, the team details are presented, and on the right, the video tool (some information is omitted due to the confidentiality agreement).

4 Database Technologies

The database behind this project has several levels of information to store. From the classic user profile, to the tracking data (players/ball) from matches and practice sessions. This leads to very different types of data. For example, to prepare a training session each coach has its own methodology, which goes from a detailed level, where every exercise is thought, to the ones prepared in a macro structure. Also different strategies are used to prepare a match.

The conjunction of several facts provided by the football consultants, gave us a set of uncertainties which made impossible to deliver a complete database structure at an earlier stage. As matter of fact, the process has been mainly iterative, where new ideas are constantly being proposed, which somehow led us to the adoption of an agile development strategy, attaining the principles that underlie the Agile Manifesto.

Those uncertainties made us decided to use relational databases management systems (RDBMS) [26] and NoSQL [27] databases. Comparing both we can state that NoSQL have simple access patterns, compromise on consistency for performance and Ad-hoc data format. On the other hand RDBMS allow complex queries with joins, ACID (Atomicity, Consistency, Isolation, and Durability) transactions and have well defined schemas. In particular, the project uses MySQL [28] and MongoDB [29] databases, allowing a very flexible structure where “well-structured/known” data (e.g., users data) is left on the MySQL server and data which can vary or be altered (e.g., definition of the game model or presentations from the Footdata-PRO module) is left on MongoDB.

Fig. 2. Two examples of web interfaces done in HTLM5. On the left some team information, and on the right the video tool (some information is omitted due to the confidentiality agreement).

5 Tactical Plan Detection and Classification

We can describe a football match, as the confrontation between two teams, guided by their own behaviours and rules. These rules and behaviours are called offensive and defensive principles, methods of attack and defence, and processes that help the players actions in the field, forming the team. All these concepts are part of the tactical organization, that can be grouped together forming what is known as the tactical plan (“game model”). Several systems have been developed for the above purpose and are being used by some of the well-known football clubs in Europe, e.g. [30]. Since a football match is composed with 23 moving objects (22 players and a ball), it is an application area of spatial information science, where spatial and spatiotemporal approaches can be applied to tactical analysis. However, few researches of spatial and spatiotemporal analysis have been done for sports tactical analysis [8]. Nevertheless, an approach to the analysis of time-based event records and real-time behaviour records on sport performance known as T-pattern detection is presented in [13]. Latter, a framework for analysis of soccer tactics with trajectory data is presented in [8], and in [10] the authors used tools from network theory to describe the strategy of football teams. For other approaches see [11] and [12].

For the player detection, and localization in the field we use an open source computer vision library, the OpenCV [19], [31], it has C++, Python and Java interfaces running on multiple OS. In our case C++ and the CodeBlocks IDE [32] were used. To speed up the code, the OpenMP API [33] was also used, which supports multi-platform shared-memory parallel programming.

In terms of computer vision, we prepare the system to process videos and realtime acquiring. To detect and localize the player, the algorithm can be summarized as: (a) Definition of the game zone (pitch delimitation), using the HSV colour space [34] combined with the lateral lines detection with the Canny edge detector and Hough transform [35]. We must refer that this step is only necessary on broadcasted videos. When the images are acquired by a static camera that captures the entire pitch, the key points for the contours of the field are mark by hand at the beginning of a match, using the implemented interface.

The next step (b) is to detect and assign the players to each team, for which we use a combination of median and morphological filters [35] over the binary field image, to obtain players blobs. Then, to assign them to each team, the equipment range of the Hue was used. The final step (c) consists in tracking each player by calculating the shortest Euclidean distance between each player in the present and in the previous frame. Occlusions, overlapping and players "collisions" are also solved.

In order to be able to analyse the tactical plan, it is required to obtain the correct coordinates of the players with respect to the football field. Which is attained with a perspective transformation (see e.g. [19]), from the original image (video frame) to the "model" field. In the Footdata, despite computing many of the traditional statistics, we are more interested in detecting in real-time (or in video) where the different pre-defined (by each coach; saved in the IS) tactical plan situations is accomplished or fails, labelling the different situations, and sending the information to the coach. We must refer that most game models are built from several relations: players position in the field, their position relatively to the teammates and adversaries, and the position of the ball. This is still work in progress.

Figure 3 shows a frame from a broadcasted video sequence (left), the players detection and classification, and the projection to the model field (right), where the distances are in meters, and the origin of coordinates is the interception point of the top lateral line with the midfield line.

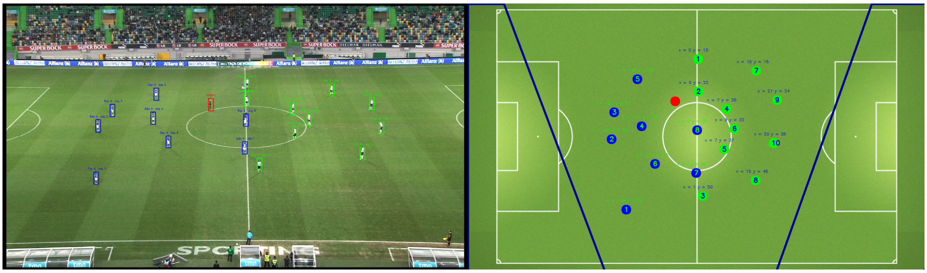

Fig. 3. The players classification (left), and the projection to the model field (right)

6 Conclusions

In this paper we present and discuss some of the main open source technologies which were involved in the building of a web integrated information system for football. The selected technologies present the most promising functionalities and compatibilities to allow us the implementation of the pretended web/cloud IS. Despite not profoundly discussed due to the lack of space, other technologies have been tested.

Other subject also not discussed in this paper is the usability of the interface. This is a fundamental key feature for the success of the product, since the final users (technical staff and players) have a very particular language (professional way to see the football), which is very different from the usual fans. Due to this reason, the interface has been discussed step-by-step with the designers, several football coaches and elements of the technical staffs. Different outputs have been (and still are) in discussion, depending on the video source to be analysed, i.e., from broadcasted videos or from a static (system) camera in the stands.

In the context of the used open source technologies, it is also worth mentioning the management and documentation of the project is done with ChiliProject [36], a web based project management system; the Git [37] is used for the distributed revision control and source code management system; and Doxygen which is used for the code documentation [38].

As future work, we plan to continue the development of the Footdata-PRO user interface. In the tactical plan detection and classification, we have already implemented the most trivial statistics, as the distance travelled by a player, the detection of passes, etc. We are now focused in the automatic detection of the offensive and defensive principles, methods of attack and defence. Also, we are still improving the players and ball tracking methods, particularly in the case of occlusions, focusing always on the fact that the system has to work in real time and has to be portable (easy to mount in any pitch or practice field by an assistant coach), even if for those reasons we have to lose a little bit of reliability.

Acknowledgments. This work was partly supported by the Portuguese Foundation for Science and Technology (FCT), project PEst-OE/EEI/LA0009/2011 and project FootData QREN I&DT, n.º 23119. We also thanks to project leader Inesting, S.A. [www.inesting.com], and the consultant football coach Domingos Paciência.

References

1. Armatas, V., Yiannakos, A., Sileloglou, P.: Relationship between time and goal scoring in soccer games: Analysis of three World Cups. Int. J. of Performance Analysis in Sport 7(2), 48–58 (2007)
2. Franks, I., Miller, G.: Eyewitness Testimony in Sport. J. of Sport Behaviour 9, 39–45 (1986)
3. Kizanaro, http://www.kizanaro.com (accessed November 2012)
4. Tracab, http://www.tracab.com (accessed November 2012)
5. Prozone, http://www.prozonesports.com (accessed November 2012)
6. Amisco, http://www.sport-universal.com (accessed November 2012)
7. Sportvu, http://www.sportvu.com (accessed November 2012)
8. Kim, H., Kwon, O., Li, K.: Spatial and Spatiotemporal Analysis of Soccer. In: Proc. 19th ACM SIGSPATIAL GIS 2011, Chicago, USA, November 1-4, pp. 385–388 (2011)
9. Leser, R., Baca, A., Ogris, G.: Local positioning systems in (game) sports. Sensors 11(10), 9778–9797 (2011)
10. Peña, J.L., Touchette, H.: A Network Theory Analysis of Football Strategies, pp. 1–6 (June 28, 2012), arXiv:1206.6904v1[math.CO]
11. Duarte, R., Araújo, D., Folgado, H., Esteves, P., Marques, P., Davids, K.: Capturing complex, non-linear team behaviours during competitive football performance. J. Syst. Sci. Complex, 1–11 (2012)
12. Moura, F., Martins, L., Anido, R., Barros, R., Cunha, S.: Quantitative analysis of Brazilian football players' organisation on the pitch. Sports Biomech. 11(1), 85–96 (2012)
13. Pilgrim, M.: HTML5 - Up and Running. O'Reilly Media, Inc. (2010)
14. HTML5, http://www.w3.org/TR/2011/WD-html5-20110525/ (accessed November, 2012)
15. Django, http://www.djangoproject.com (accessed November 2012)
16. Python, http://www.python.org/ (accessed November 2012)
17. Kaplan-Moss, J., Holovaty, A.: The Definitive Guide to Django: Web Development done Right (Expert's Voice in Web Development), 2nd edn. Apress (2009)
18. Oliphant, T.E.: Python for Scientific Computing. Computing in Science & Engineering 9(3), 10–20 (2007)
19. Bradski, G., Kaehler, A.: Learning OpenCV: Computer Vision in C++ with the OpenCV Library. O'Reilly (2012)
20. Powers, S.: HTML5 Media. O'Reilly (2011)
21. Pfeiffer, S.: The Definitive Guide to HTML5 Video. Apress (2010)
22. Casario, M., Elst, P., Brown, C., Wormser, N., Hanquez, C.: HTML5 Solutions - Essential Techniques for HTML5 Developers. Appress (2011)
23. Meyer, J.: HTML5 and JavaScript Projects. Apress (2011)
24. Schmitt, C., Simpson, K.: HTML5 Cookbook. O'Reilly (2012)
25. jQuery:jqueryui.com (accessed November 2012)
26. Balasubramaniam, R.: Data Security in Relational Database Management System. Int. J. of Computer Science and Security 6(4), 203–210 (2012)
27. Cattell, R.: Scalable SQL and NoSQL Data Stores. ACM SIGMOD Rec. 39(4), 12–27 (2011)
28. MySQL, http://www.mysql.com (accessed November 2012)
29. mongoDB, http://www.mongodb.org (accessed November 2012)

30. Borrie, A., Jonsson, G.K., Magnusson, M.S.: Temporal Pattern Analysis and its Applicability in Sport: An Explanation and Exemplar Data. J. of Sports Science 20(10), 845–852 (2002)
31. OpenCV, http://opencv.org (accessed November 2012)
32. Code: Blocks, http://www.codeblocks.org/ (accessed November 2012)
33. OpenMP, http://openmp.org/wp/ (accessed November, 2012)
34. Dearden, A., Demiris, Y., Grau, O.: Tracking Football Player Movement from a Single Moving Camera Using Particle Filters. In: Proc. 3rd Euro. Conf. on Visual Media Production CVMP 2006, London, UK, pp. 29–37 (2006)
35. Rao, D., Patil, S., Babu, N., Muthukumar, V.: Implementation and Evaluation of Image Processing Algorithms on Reconfigurable Architecture using C-based Hardware Descriptive Languages. Int. J. of Theoretical and Applied Computer Sciences 1(1), 9–34 (2006)
36. ChiliProject, http://www.chiliproject.org (accessed November 2012)
37. Git, http://git-scm.com (accessed November 2012)
38. Doxygen, http://www.stack.nl/~dimitri/doxygen/ (accessed November 2012)