ENABLING AUTOMATIC SOCCER ANALYSIS THROW WI-FI REAL TIME TRACKING

Pedro Abreu¹, Vasco Vinhas¹, Pedro Mendes¹, Luis Reis¹, Julio Garganta²

¹Faculdade de Engenharia Universidade do Porto, DEI/LIACC, Rua Dr. Roberto Frias, s/n 4200-465, Porto, Portugal, pha@fe.up.pt, vvm@fe.up.pt, pmcm@fe.up.pt, lpreis@fe.up.pt
²Faculdade de Desporto da Universidade do Porto, DF, Rua Dr. Plácido Costa, 91, 4200-450, Porto, Portugal, jgargant@fade.up.pt

Abstract. Nowadays soccer is more than a sport, it is a business industry. Providing financial support to a senior team is a constant challenge for club managers. Having this in mind the most relevant factors for a club's life are the game results. They can ensure, amongst other aspects, the continuity of a soccer coach. As a consequence of that, systems supporting decisions represent fundamental tools not only for match analysis but also to support training sessions. In the literature most academic advances are camera-based solutions and still face some technical issues like occlusion problems and high computational demands. This research project presents new portable and inexpensive soccer analysis software using as base a Wi-Fi real-time tracking system able to identify typical player movements all over a penalty box. The data gathered over a training session constitutes good results especially in what concerns accuracy, thus confirming its robustness and applicability as a tool for scientific soccer analysis and as a valuable support for practical soccer training session. In the future the usage of high-end wireless equipment will increase the system's accuracy and also the coverable area.

Keywords: Wi-Fi, Real-Time Tracking, Business Intelligence, Soccer Analysis.

1 Introduction

Collective sports games (CSG) appeared in the 7th century b.c in Colombia when the Inca civilization played a game called Pok ta Pok which is believed to have many similarities with basketball. In the Pok ta Pok game the field is constituted with two baskets located at three or four meters of the ground. The main goal was to pass a ball throw those targets without using the feet or the hands. Since then many other games emerged and in the 19th century the CSG hit its apogee with many participants playing them all over the world [3].

Nowadays there are more than two hundred CSG and in some cases one game might have the same name but different rules in distinct world regions. In order to classify these games a systematization process has been conducted in the recent years [7][22]. It consists in grouping the games having into account different sets of characteristics such as: ball shape, type of the field (indoor or outdoor) and its regions, game rules, definition of opponents and teammates etc.

The football or Soccer (designation used in North America) is a CSG that appears, as we know today, in the 19th century in England [31]. One of the game's characteristics is that the players have the possibility to make real time choices, having the restrictions imposed by the rules, defined in a training session or in a professional competition, always in mind. This point is an element of freedom and consequently creativity. This reality should be managed by another individual called Coach whose purpose is to train the team for matches. In the literature many are the authors that tried to define Coaching. For Hughes [10] coaching a soccer team is mainly a task of enhancing performance by providing feedback about the performance of the athletes and team.

This performance may be measured having in mind different perspectives like: Physical, Technical, Tactical and Psychological. Although these points of view are extremely relevant for the game, tactical perspective is one that has occupied the academic society over the several years. The game's tactical analysis allows the coach to identify the relationships amongst play events, considering not only the individual but also the collective factors. These performance indicators reflect the importance of teamwork, pace, fitness and movement, and therefore to target strengths and weakness [11].

Human observation and memory are not reliable enough to provide accurate and objective information from athletes in high-performance environments. Consequently, measurement tools constitute an important instrument by providing information about movement, statistical compilation of game events.

More then offline tools, the use of real time soccer systems provide the coach one opportunity to change some tactical or even technical aspects in a training session or in a real game situation.

The majority of tools that exist in the market are very expensive and in some cases still present some technical issues like occlusion problems. This research work presents a platform that could be used for a coach in a training session or even a match. It is capable to automatically calculate the full path of a player in the field in a specific interval of time, the most populated zone in field all over a game allowing the coach to review the performance of his team. Using a Wi-Fi network and a positioning engine on top of it, this system provides a visualization tool for such data on a real time basis. This information includes fully scalable concentration grids, a vision inference assuming that the tracked entities are associated with soccer players. In addition to what was
exposed the system also works as a statistical collector meaning that it is possible to use data mining techniques predicting and categorizing typical player paths and also detecting their behavior patterns all over the game.

This paper is organized as follow. Section 2 describes the current state of the art regarding the most relevant approaches concerning CSG specially related with soccer, section 3 presents an overview of the system’s global architecture and describes some of its most relevant modules and details; section 4 exposes the results obtained from one system's instantiation; and finally, in the last section, conclusions are presented and future work trends are discussed.

2 State of the Art

Nowadays the key factor in a soccer club’s life is the game results. They determine-represent the success of the club and in many cases the coach’s future. Because of that club coaches need to have maximum technico-tactics information about the game events and the way that it was played by the players [1].

Currently many are the computer systems that support coaching decisions before and after the game [16]. These types of systems may be divided in three distinct groups: performance evaluation, strategy development and real time competent assistant system. These last are the most complex ones because they involve some particular features like real time objects tracking, identification and classification of player movements and game events detection. In order to build an indispensable tool for a coach, these systems should be automatically capable of recognizing intentional activities in a multiagent environment with continually acting agents.

In the next subsections a group of generic of-the-shelf and academic tracking systems shall be presented.

2.1 Generic Tracking Systems

In the literature there are many generic tracking systems that emerged over the past few years. These solutions are divided in two distinct groups: image based and non-image based.

2.1.1 Non-Image Based Systems

The Global Positioning System (GPS) is a satellite based solution that began being used by the U.S military forces for the planning of their operations particularly in arid and mountainous terrains. Since the 80’s this technology became available for general public use and today it is normally used to do real time tracking analysis of different types of vehicles and as a base to analyze their motion [30][19].

The Radio Frequency Identification (RFID) is an automatically wireless identification method that is capable of tracking objects and even people using radio waves. In terms of the required hardware, this technology uses a receiver and a set of tags which could be classified in: passive tags that are only detectable within a range lower than 13 meters of the receiver active tags that can be found within 40 meters of the receiver but need to have their own internal source power. Although the use of this technology could be an interesting solution for some areas, the high cost of the receiver and the active tag’s average unit price is still an issue [6].

Wi-Fi is the name given for a popular wireless networking technology that uses radio waves to provide wireless high-speed Internet and network connections. Having this technology as a base it is easily to create a wireless data network that could be used in historical urban environments, academic campus etc. This technology could be used also for designing a tracking system. By reusing the wireless data network and it is possible to create a tracking system on top of this infrastructure. Another advantage of this approach is the possibility of tracking an object using only a single access point though in this particular situation the precision will diminish due to the lack of signal triangulation.

By comparing this last with other technologies the risks of occlusion and signal loss in this kind of approach can be considered very low mainly in environments that present low levels of metal concentration [17].

Bluetooth is a wireless protocol present in almost mobile phones in the market. Although this protocol could be used in a tracking system, the high battery consumptions [12], the short coverable area and the non-transparent connection establishment process transform this approach inadequate for an efficient tracking system.

2.1.2 Image Based Systems

Thermal Signature is one of the most expensive tracking technologies existing in the market. It consists in detecting thermal signature of the objects tracked. The main purpose of these solutions is the reconnaissance and processing of thermal images. Although these systems present good results in some environments like oceans (within the objective of tracking living entities) [21] the high cost of the equipment and in some cases the inexistency of a detectable thermal signature restricts the use of these kind of systems to a very controlled set of situations.
Multi-Camera video surveillance is the most popular tracking technology. In terms of hardware required this technology utilize a set of cameras spread in a trackable environment and a particular network. Despite being used in distinct scenarios some important issues still remain. The need to have high resolution equipment, dedicated network and the computational demands are still major problems that researchers have trying to optimize by using overlapping camera views [13][14].

In terms of generic tracking solutions and as a conclusion, all approaches have their strengths and weakness. Having in mind the characteristics of the CSG the best alternative seems to be a Wi-Fi based system. The competitive tag cost (that could be put for instance in a player shirt) and the high level of accuracy (in average less than 3 meters) shall constitute an important advantage. Another huge advantage is the fact that this technology is almost immune to the majority of occlusion problems that affect other approaches.

Another technology that could constitute a good solution is the RFID based one. In spite of the fact that this technology requires standardization and consequently the cost of equipment will fall out (specially the receiver), the use of active tags could allow the increase of coverable area and accuracy levels have already reached good values. In spite of this, occlusion issues, related with liquids and metal still persist.

Multi camera surveillance systems are also quite common. In this kind of system some problems still remain like the camera’s cost, the computational demands and the occlusion problems that constitute a very important factor that ought to be optimized. The other approaches like Bluetooth, GPS and thermal signature are not applicable in the CSG universe.

2.2 Sports Video Analysis

One of the major research areas in the CSG is the sports video analysis. In football/soccer domain researchers had focus their work in problems like shot classification [9], scene reconstruction [29], structure analysis [26][28], event extraction [2][18] and rule-based semantic classification [20]. These approaches used the image transmitted by the television and recorded them for posterior processing (after the match ended).

These kinds of systems are categorized by Ekin [8] in two main groups: cinematic and object-based ones. The object based uses algorithms to detect objects in a video while the cinematic uses features from video composition and produce rules.

2.2.1 Cinematic Approaches

Xu et. al [28] present a cinematic approach using for it the feature’ dominant color ratio’ to segment video. They defend that video reports should focus on play yard to extract game situations.

Xie et. al [26] used a Hidden Markov Models approach to detect two restricts events: play and break, in a video game. The complexity of this process is higher than in other sports like tennis or volley because for instance in soccer it is hard to determine if the game is stopped by a decision of the referee or by other highlights of the game-goal, corner, kick, shot, etc.

Other works like [23] tried to expand Xie's work and detect more game events like focus and replay in order to define new features/structures that they called Attack.

2.2.2 Object Base Approaches

The object base approach demands more computational resources but it allows more high-level domain analysis. In order to detect a large number of game events the work developed by Gong et. al [9] analyzes the ball’s trajectory and the relationship between the players’ moves over the match. In the literature there are also many works that tried new approaches like merging audio and video information [2]. Although this kind of approach could constitute more high-level domain analysis one big issue is the asynchronies between audio and video queues.

2.2.3 Real Time Tracking Systems

Over the past few years new approaches appeared that use a multi camera tracking system to track players which promote new kinds of features like a near real time analysis. By comparing with classic approaches analyzed in the previous subsections, this system uses a fix number of stationary video cameras placed in a traceable environment. This type of approach increases the overall field of view reducing the dynamic occlusion problems and improves the accuracy and robustness of the information collected.

Cai and Aggarwal [5] and Khan et. al [15] track the object using the best view camera and if the trackable object leads the field of view they change it to a neighbored camera. Other authors like Stein [24] and Black [4] assume that all trackable objects are in the same plane and compute the homography transformation between the coordinates of two overlapping images captured with uncalibrated cameras.

In the Xu et. al [27] work eight cameras were used and were calibrated in a ground plane coordinated system using Tsai’s algorithm [25]. Unfortunately this work presents some technical difficulties like problems
with sparse landmarks in the coverable area that decrease the accurate calibration and data association and situations involving more than two players grouped in the same game region.

Summing, in CSG and more specifically in soccer the unique tracking systems that already exists in real environments are camera based. As demonstrated previously these systems still have to optimize some features like occlusion problems, computational demands, material cost and lack of portability.

3 Project Description

In this section, the undertaken project is described in detail in what regards its several components and analysis perspectives.

Having this in mind, the electrical infrastructure is detailed and after that the system’s global architecture is depicted, in order to have an overall glimpse. The database model is further explained and the final two subsections are dedicated to the tools’ individual description.

3.1 Electrical Infrastructure

Most of professional soccer coaches state that the training session should have the same length as a conventional soccer match-ninety minutes. Consequently any training support system should stay active for all of this period. To fulfill this goal an electronic system was designed. In this approach a conventional 45A car battery is used directly connected to a 600w UPS. The UPS battery is also connected to the car battery in order to increase the autonomy of the system. This electrical infrastructure (Figure 1) is capable to provide power for more than 120 minutes.

![Figure 1. Electrical Infrastructure and WI-FI Network](image)

In order to increase the WI-FI network's density, a star topology approach is used. A router is connected directly to the battery's electrical extension and it is placed behind the goal. The access points (APs) were placed in specific points of the penalty box as shown in Figure 1.

3.2 Global Architecture

In this subsection the system's global architecture will be illustrated as well as of its modules and how they interact and therefore extract not only system components dependencies but also information flow analysis. All of the system's components and their relationships are exposed in Figure 2.
Having the above mentioned in mind, and paying a closer attention to the numbers in figure, one is able to identify the system’s modules as follows: Offline map editor; Wi-Fi enabled localization tag; Position Engine; Database for data integration and storage; Real-time monitoring application and Web enabled real-time and historical business intelligence.

Although most of these elements are object of further explanation in the next subsections, one ought to undertake a brief description of those whose nature is not obvious and, in order to, clarify their interaction.

The first action, that ought to be conducted, in offline mode, consists in conducting a complete map creation-edition. The user shall specify, amongst other details, depicted in subsection Map Editor, the image file representing the soccer field layout and the used scale. This information is compiled in a XML file for both the position engine and real-time monitoring tool and submitted to the mentioned database for the historical BI application.

The Wi-Fi tag consists in an active 802.11 a/b/g board with a couple of power batteries attached. These are configured to connect to a specific Wi-Fi network – security, DHCP but another network configurations are also possible – in order to directly communicate with the position engine. By using this kind of wireless technology, it is possible to reuse partially or totally the spot’s arena network infrastructure, having only, for special requirement, a high density of access points as the accuracy naturally increases with this factor.

The used position engine periodically collects data from the tags and updates their position against a pre-loaded localization model. This model requires a previous offline site survey for measuring Wi-Fi signal strength and for network items – routers and access points – precise localization. The engine is also web-enabled and supports a HTTP/XML API so that third-party applications can interact with it, therefore accessing localization and status information regarding each individual registered tag.

Using this communication protocol, the developed real-time monitoring server is responsible for gathering, at a specific periodicity that typically equals to the position engine frequency – every tag’s valid location data. With this information, this module is directly responsible for updating the database and caching the session’s data for the real-time monitoring application.

Having the continuous up-to-date database as a solid information reference, it was possible to enable both real-time and historical business intelligence applications. For real-time knowledge extraction, it was only used data referring to active sessions, for historical analysis, and delegating all the process effort to the database engine, specific and dynamic time windows were used to filter data. Despite the additional explanations that are given in subsection Real-Time and Historical BI Application, the versatility of such application must be referred as it congregates both web-enabled features and zero data process as it is all delegated to the database engine and allocated in a dedicated server – enabling its usage in a wide range of devices, including PDAs and mobile phones, alongside with traditional notebooks and desktop computers.

As a synopsis, one might refer the system’s architecture as fairly distributed, where offline information regarding soccer field layout and wireless network definitions team up with a real-time web-enabled position engine, which enables third-party applications to collect and store data, so that diverse specific end-users can
access both real-time and historical knowledge in a wide range of equipments, therefore enhancing coaching efficiency levels.

3.3 Database Model

In this subsection, the designed database model is revealed and justified.

Having into consideration the specific reported system’s application in the soccer domain – usually characterized for multiple player movements all over the field combined with the project’s idiosyncrasies – specifically in what concerns to localization tracking frequency – the database model paradigm followed consists in a hybrid form of a data warehouse star architecture with a slight normalization flavor, as illustrated in Figure 3.

Regarding the strong star model, it is supported for the high data production levels, and perhaps most important, the fact that all data insertions are machine responsibility, as depicted in the above subsection, therefore preventing human error. It is also vindicated by historical analysis need that may cover hundreds of thousands and even millions of records.

On the other hand, some database normalization was introduced in order to cope with real-time requirements that would not be compatible with computing hundreds of records out of a table with millions of records, in a continuously systematic way. Another argument in favor of database normalization resides in the soccer's field layout.

Referring to specific database model items, one shall point the central relevance of rtls log as central table responsible for storing all localization data. For each pair of tag/session identification, a particular position is recorded in a given layout with a specific timestamp. The concept of session may be different in each training session according to coach's decision. A new session could be related to three distinct situations: a player substitution (when a player is substituted by a colleague), a player out of the field (for instance to receiving medical assistance) or other situation when the player is out of the limits of the region that was defined by the coach for a specific situation in training session.

In order to achieve real-time requirements, some redundancy has been introduced in what concerns active session identification, so that active players identification could be easily, and most important, efficiently retrieved.

3.4 Map Editor

Map Editor is a traditional, network enabled, desktop application responsible for complete soccer field layout definition. The soccer coach shall open an image file and provide the interface with the drawing scale – in order to convert pixels to meters and vice-versa. Afterwards, the tool offers the possibility to pinpoint and draw, over the original layout, spawn areas – concept that will allow the detection of new sections.

Once the layout is completely defined, the coach is able to save map characterization in a XML file in any available location and/or commit it to a specified database – with the previously described database model implemented.

The XML file will be an input for both the position engine and the real-time monitoring server, and, on the other hand, the committed database information is ground for historical computation and analysis.

Summarizing, the Map Editor constitutes itself as an auxiliary tool, vital for system’s setup and dynamic enough to cover the entire analyzed soccer field. Its dual output enables a flexible usage for several system components and, simultaneously, due to XML openness, enables third-party development and integration.

3.5 Real-Time and Historical BI Application

In order to extract significant business intelligence knowledge based on historical data and not only real-time information, the authors decided to make an immediate use of the raw position data stored in the database.
Taking advantage of using Oracle as equally laboratory and production database, Oracle’s Application Express was used to generate a web application responsible for processing data and, most important, aggregate information in an understandable way.

As depicted in Figure 4, the Apex’s engine is directly embedded in the database, thus directly dealing with coach’s web requests. With this architecture, several systems can easily access BI application as all heavy processing is the database server’s responsibility, leaving the coach device with only chart rendering computation.

![Figure 4. The Apex engine](image)

### 4 Results

The results exposed in this section concern to the data gathered over a training exercise conducted with four human players in a real soccer field’s penalty box with its dimensions as well as the goal’s as recommended by the Federation International de Football Association (FIFA). The exercise’s purpose was to train a player’s shot accuracy after receiving a pass from a winger. For that matter a goalkeeper, two wingers and a striker participated in this experience having each of them a Wi-Fi tag attached to their shirts.

The penalty box was also divided in a 10*4 grid for calibration purposes and also to guide the site surveying process. The following picture (Figure 5) exposes how the exercise was conducted.

![Figure 5. Soccer Exercise](image)

To clarify the Wi-Fi network’s density one ought to first specify the access points’ positioning. A router was placed behind the goal as well as the batteries and the entire electrical infrastructure described in the previous section. The remaining three access points were also used and positioned over the center of the remaining lines that define the penalty box (excluding the one which contains the goal line). To maximize the signal’s strength all the Wi-Fi devices emitting a signal were put on top of a structure that allowed them to gain 1.20 meters of height. They were also put twenty centimeters away from the real lines so that the players’ moves were not affected by their presence. Figure 6 shows the signal’s strength and noise levels on this particular scenario:

![Figure 6. Signal Strength and Noise for WI-FI network](image)

Since this is an outdoor environment the authors believe that the gathered noise values are the main cause for the error on the player detection because they are not being compensated by refraction and reflection phenomena which are typical in indoor environments. One ought to point out that this test was conducted with high-end devices and so there is a high probability of diminish the noise’s impact just by changing the hardware to high-end artifacts, as their value mostly differs on the applied power on signal emission.
Even so, the next figures clearly demonstrate that the system was able to track the players during this exercise which lasted about thirty minutes. For instance, on figure 7, showing the box’s density over the entire exercise with the scale depicted at the bottom of the picture, one can observe a red cell on the goal area which undoubtedly corresponds to the goalkeepers’ presence waiting for the striker’s shots. The neighbor cells are also highlighted as the goal keeper moved a bit during the exercise in order to better cover the striker’s shots on goal. The other highlighted cells demonstrate how the other three players moved during this training session.

Figure 7. Box density over an exercise

Figure 8 shows a real time screenshot of the player density where one can observe the wingers’ position after having one of them passes the ball.

Figure 8. Player density in the game field

And finally on Figure 9, one can observe the left winger’s and striker’s position during a pass. On this particular figure the player’s are represented as blue dots over the field. In this case the error between the obtained position and the real one did not exceed two meters for each player, which also justifies the fading green cells on the box’s corner (shown on Figure 7) as the wingers could decide from where they wanted to perform the pass as long as their distance to the box’s limits did not overcome three meters.

Figure 9. Striker Position during a pass

Overall the system remained stable during the whole training session thus confirming its robustness and applicability as a tool for scientific soccer analysis.

5 Conclusion and Future Work

In this section, the project’s main conclusions shall be presented as well as major future work areas and potential collateral applications are deputed.

In this paper a new tracking system tested in a penalty box area in the universe of soccer is presented. As mentioned in section 3, with the construction of a portable and inexpensive system that includes basic wireless network, a car battery and UPS it was possible to track players on a real time basis all over a penalty box in a soccer field. This project shows that a Wi-Fi based technology could constitute an excellent solution for
satisfactory- in average less than 3 meters- using low brand equipments- router and access points. With this inexpensive tracking solution any team’s coach has detailed reports about the performance of a specific player or the all team in a training session or even in a soccer mach. The possibility of having real time player positions in a specific situation and historical player paths constitutes an important tactical indicator for any soccer coach. In this particular item the Oracle’s Apex Technology proved to be a solid solution. It allows multiple simultaneous accesses and, consequently, dramatically enhance analysis empowerment, while, at the same time, eliminates heavy data computation from end-users terminals. These characteristics allow accesses from unconventional systems such as PDAs, smartphones and not only notebooks and desktop computers through their web-based interface. This particular feature has a great importance for, technical staff that for instance is spread through the soccer stadium in a match.

Regarding future work areas there are many possibilities that could wide potential the robustness and reliability of this system and new applications.

Concerning to the positioning engine; the algorithm of tracking could be improved. This particularly point merged with the use of better wireless equipment will increase the accuracy of the system.

Secondly, another item that could be always improved is the tag dimension. After this optimization it will be possible to incorporate this type of tags in a player’s shirt without any risks for physical integrity.

Another feature that could be interesting to explore as future work is the transformation of the actual system in a complete support decision framework for soccer coaches. For that purpose it is necessary to build a hybrid tracking system made by two synchronous modules. One module will be responsible for tracking the players and for this the actual system could be a solution and the other one should be responsible for tracking the ball. In this last problem one of two solutions could be adopted: a camera based classic solution with the advantage of only needing to track a specific object (with particular dimensions and color) decreasing so, the occlusion problems or adopt a new type of approach using, for instance, a chip inside the ball.

The second step for this new system will be the construction of soccer ontology. This point has particular importance because it helps to define concepts relationship with events of the game like: a pass, a shot, a corner etc. After that it is possible to construct a tracking system that will be capable to automatically detect game events, calculate historical player paths and in an advance face automatically detect player behavior relationship not only with their positionning but also with ball’s. This system will definitely fill a gap in the market.

Taking into account the current project’s features and also the identified future enhancements, there have been identified several application domains that go beyond the soccer or even CSG.

Amongst these, one shall mention the possible system’s adoption by large warehouse management where traffic jams are not unusual. The proposed system would permit live vehicle tracking that in conjunction with a planning module would enable efficient traffic control, therefore avoiding bottlenecks, without compromising warehouse storage capacity.

Another possible application would reside in health care institutions where it would be useful for medical staff tracking around the facilities, in order to efficiently contact them in case of emergency. Also within this domain, especially in mental institutions, patient tracking could be a great advantage.

Security applications are also easy to imagine, not only to track assets in a closed environment but also potential human targets such as children in public areas – such as malls hotels or conventional centers.

As a summary, it is fair to state that the project’s initial ambitions were fully met and that the cooperation with an important university in the sports area was extremely important for better measuring the system’s positive impact. The technology’s transparency, allied with the future work areas, is believed to greatly improve potential applications, thus significantly widening the project’s initial horizons.

6 Acknowledgements

The authors would like to thank Professor Eugénio Oliveira for excellent laboratory conditions as well as financial support for equipment acquisition. A special mention is also due to FADEUP for its will to be an active partner of this project as well as to provided the installations of the football field.

References


