

A Study on Composition of Context-based Soccer Analysis System

Jiwon Lee, Do-Won Nam, JungSoo Lee, Sungwon Moon, Kyungill Kim, and Howon Kim

SW-Content Research Laboratory,

ETRI(Electronics and Telecommunications Research Institute), Korea

ez1005@etri.re.kr, dwnam@etri.re.kr, jslee2365@etri.re.kr, moonstarry@etri.re.kr, kki@etri.re.kr, hw_kim@etri.re.kr

Abstract— Recently, the tendency to analyze sport activities by combining sports science and ICT technology is increased. This aspect is also presented in the soccer analysis, and it is easy to find various cases of analysis according to the purpose. However, the systems presented so far do not attempt to provide a high level of content analysis. This paper discussed the limitations of the statistical analysis of the current soccer analysis system, and then presented a concrete method of constructing the system to bring it up to the content understanding base. Through the development of the proposed system, we can expect a positive change in terms of new aspects of service, industry and ecosystem.

Keywords— Event recognition, Motion recognition, Object tracking, Soccer analysis, Sports science

I. INTRODUCTION

Attempts to analyze and judge sports activities by applying sports science and ICT technology and to apply them to sports contents are increasing. For example, the Hawk-Eye [1] in the UK is widely used as an auxiliary means in tennis or volleyball for detecting line out by using high-speed cameras and sensor technology. Also, the SportVision [2] in the US provides player/ball recognition and tracking information in MLB to give fun of baseball games to the audiences.

Along with this trend, the use of IT technology and personal sensor devices as a means of measuring or increasing athlete skill is also increased. Most professional sports clubs and national athletes in the world are already using sports analysis solutions based on sports science, therefore, competition for technology development is intensifying.

These changes have had a great impact on the professional soccer market, and many products are currently being developed, but they are not enough to efficiently upgrade the current labor-intensive soccer analysis systems. In order to solve this problem, this paper first addresses the problems of the existing soccer analysis system, and then proposes a new analytical system to solve the problems.

The composition of this paper is as follows. In Section 2, we present the summary of the existing soccer game analysis systems and their limitations, and then the composition of the proposed system is presented in Section 3. Section 4 discusses the expected effects of the proposed soccer analysis system and future research directions.

II. PREVIOUS WORKS

Until now, soccer statistics/event recoding staffs watched soccer in real time and wrote a history on the recording paper whenever an event occurred for the analysis of professional soccer games. Then, team officials and content producers reworked the information from the recording paper to make the suitable contents for their purpose. This method was the simplest, but labor-intensive method. It needed a lot of cost to hire soccer recording staffs, and also recorded information only has event-based statistics that recording staffs could judge. In addition, there are many cases where the contents of the soccer game are missed during the recording process, so that many recording staffs have to analyze the game at the same time to increase the accuracy thereof.

To solve these problems and to diversify the kinds of information to be provided, many products have been launched to date. Typical products used in actual fields are shown in Table 1.

TABLE 1. MAJOR COMMERCIAL SOCCER ANALYSIS SYSTEMS

Products	Data acquisition method	special features
Sportscodel (hudl)	Manual	· Event extraction based on manual input data · Automatic video editing based on given editing parameters
TRACAB (chyronhego)	Camera	· Video-based ball / player / referee 3D real-time tracking
Viper (statsports)	Sensor	· Measurement of Athlete's speed, acceleration, maximum speed, heart rate, stress index, step balance

US Hudl's Sportscodel [3] develops and provides a system that minimizes the recording error of recording staffs by digitizing the recording paper. However, the system has a problem that it merely provides the digitized recording method to recording staffs, thus it cannot meet the needs of consumers who want more specific information.

TRACAB [4] of US Chyronhego Company has developed and provided a system to track the position of players / referees / ball in the stadium by using camera-based rocket

tracking technology and to track them in real-time and provide low-level statistical data on them. The TRACAB system can produce a richer track record by providing low-level statistical information that the recording staffs cannot provide, such as the player's position, running distance, speed, and so on. However, it does not provide high level event information such as pass, shoot, foul, etc. provided by the recording staff, and this part has a problem again to rely on the record of the recording staff.

Viper [5], Statsports Inc. in the UK, is a sensor-based device that attaches to a player's body and provides biometric information, such as heart rate and running distance. This system distinguishes itself from other systems by providing meaningful information that can calculate the biometric information such as athlete's strength, fatigue, etc., but it does not solve the need for labor-intensive recording staffs.

Recognizing the difficulty of automatically extracting high-level event information of major commercial products and attempting to solve this problem is proceeding in academia and industry. In the case of video-based event recognition systems [6][7], they proposed the system which makes patterns consisting of camera shooting styles and recognizes major events by using them in the soccer broadcast video. Here, major events mean goals, shootings, corner kicks, free kicks, penalty kicks, fouls, etc. In other words, first classifies soccer broadcasts by replay, close-up view, short view, long view, and crowd view, and then learns what patterns the corresponding views appear in for each event. Based on this, the main scene of the game can be recognized. On the other hand, in the case of voice-based event recognition system [8], a method of recognizing major events by recognizing the referee whistle sound and the reaction of the game commentator and the spectator is used.

The above-described methods for automatically recognizing existing events are not based on the behavior of an object such as a player / referee who performs a real game, but rather search for a voice /camera shooting style for recognition. Therefore, there is a limitation that the change of the shooting technique and the reaction of the spectator or the commentator cannot sufficiently reflect the difference between the recognized pattern and the predefined pattern. Furthermore, existing technologies cannot recognize each event separately because they do not accurately recognize the location and motion of the player / referee, therefore, it has a limitation that it focuses more on producing a highlight video instead of accurately classifying and analyzing the events one by one. Therefore, it is needed for a method that can accurately identify each soccer event through the location and motion information of the object directly related to the game, such as player / referee / ball, not the indirect object related to the game such as the camera movement and the reaction of spectators / commentators.

III. PROPOSED SYSTEM

As we have seen in Section 2, the soccer analysis systems released so far automatically provide low-level statistical information, but they cannot solve the problem of recording

high-level soccer events which need a lot of labor. In this paper, we propose a system to mitigate this problem by using deep learning and pattern matching techniques.

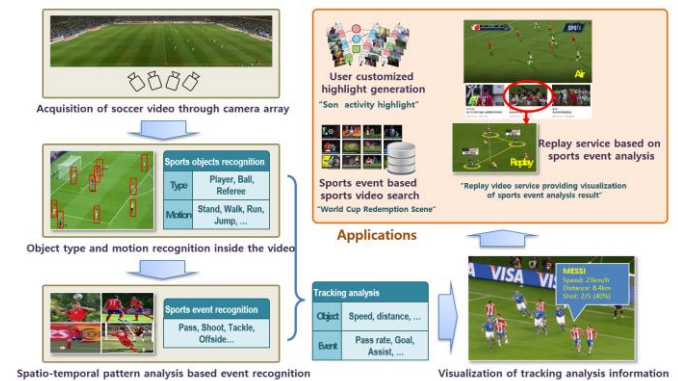


Figure 1. Overview of proposed system

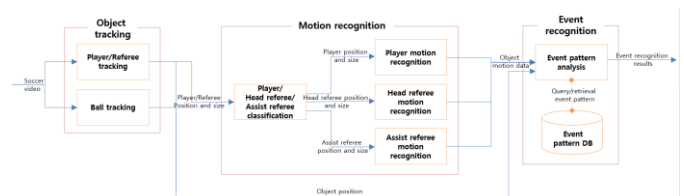


Figure 2. Top schematic of proposed system

The outline of the proposed system is shown in Figure 1 and the top schematic is shown in Figure 2. First, we distinguish each object using multi object tracking technique from given videos through multiple cameras and extract low-level statistical information such as position and speed of the object. Up to this point, it is no different than existing technologies. However, the proposed system defines a primitive motions such as running, walking, kick, jump etc. that each object can have and learns these motions through the deep learning technique to recognize learned motions from the inputted soccer video. The extracted motion information does not have great significance individually, but it is possible to recognize a sport event by patterning and analyzing a series of motions and position changes of related soccer objects. For example, as shown in Figure 3, by combining the motion information of player A with player B, the system can find out some events such as player A assists player B, player B grasps ball passed from player A and makes a goal. The extracted low-level statistical information and motion information are accumulated in the big data. Analysis of big data through pattern matching technique enables to identify tactics such as offside traps by analyzing patterns of events as well as event recognition.

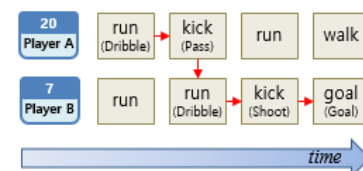


Figure 3. Motion-based event recognition example

The proposed system has a big difference from the existing service. The biggest difference is that high-level analysis of contents in the given soccer game is possible. In the case of the existing service, only statistics-oriented analysis was provided, and the content analysis was conducted only through the professional analyst, however, the proposed system can automatically understand and provide the contents of the soccer game. As the contents of the soccer game can be automatically understood, there is a big change in the search and summary of the soccer video. In the case of existing video retrieval, only tagged keyword based retrieval is allowed in the video, but content based video retrieval becomes possible through the proposed system. For example, when searching for "World Cup Park Ji-sung heading shoot", existing system only provides video with similar tag information, so it is highly likely that the desired data cannot be obtained. However, the proposed system can automatically recognize and provide the result. In this way, even for video summaries, so far only the service provider-oriented highlight video has been provided, through the proposed system, it becomes possible to provide so-called "on-demand customized sports highlight" that the user selects the desired scene, highlights are automatically created and provided.

IV. CONCLUSIONS

This paper presents a concrete method of constructing a system for raising the current soccer analysis system based on statistical analysis to the content understanding base. Through the development of the proposed system, we hope that it will contribute to the creation of ecosystem capable of reproducing various sports contents through analyzing and reusing accumulated sports data.

ACKNOWLEDGMENT

This research is supported by Ministry of Culture, Sports and Tourism(MCST) and Korea Creative Content Agency(KOCCA) in the Culture Technology(CT) Research & Development Program 2016 (R2016030044, Development of Context-Based Sport Video Analysis, Summarization, and Retrieval Technologies)

REFERENCES

- [1] Hawk-eye innovations [Online]. Available: <http://www.hawkeyeinnovations.co.uk/>
- [2] Sportvision [Online]. Available: <http://www.sportvision.com/>
- [3] Sportscode [Online]. Available: <http://www.hudl.com/elite/sportscode>
- [4] TRACAB [Online]. Available: <http://tracab.hegogroup.com/>
- [5] STATSports viper [Online]. Available: <http://statsports.com/technology/viper-pod/>
- [6] T. Y. Lee, "Event learning and detection method in soccer video," *KR patent, 10-0963744*, 7 Jun. 2010.
- [7] A. Ekin, A. M. Tekalp and R. Mehrotra, "Automatic soccer video analysis and summarization," *IEEE Trans. Image Processing*, vol. 12, no. 7, pp. 796-807, Jul. 2003.
- [8] M. Xu, N. C. Maddage, C. Xu, M. Kankanhalli and Q. Tian,

"Creating audio keywords for event detection in soccer video," in *Proc. ICME 2003*, pp. 281-284, Jul. 2003.



Jiwon Lee received the B. S. degree in Computer Engineering from Kyungpook National University, Republic of Korea, in 2008, and the Ph. D. degree in Computer Science from Korea Advanced Institute of Science and Technology (KAIST), Republic of Korea, in 2013. Since 2013 he has been a senior researcher in the Sportainment Section, SW.Content Research Laboratory, Electronics and Telecommunications Research Institute (ETRI), Republic of Korea. His research interests include multimedia security, image/video watermarking, and image/video processing.



Do-Won Nam received the B.S. degree in Computer Science from Korea Advanced Institute of Science and Technology (KAIST) in 1996 and M.S. degree in Information Technology from Pohang University of Science and Technology (POSTECH), Korea in 1998. He is working as a principal researcher in the Electronics and Telecommunications Research Institute (ETRI) since 2001. His research interests include data mining, digital rights management, digital cinema system and sports video analysis.



Jungsoo Lee received his B.S. and M.S. degrees from Jeonbuk University, Korea in 1995 and 1997, respectively and his Ph.D. degree in Electronic Engineering from Hanyang University, Seoul Korea in 2005. From 2000 to 2005, he was a senior member of MarkAny Research Institute. Currently, he is a senior member of Electronics and Telecommunications Research Institute (ETRI). His research interests are digital watermarking, fingerprinting, image processing, digital rights management, digital cinema and digital signage.



Sungwon Moon received his B.S and M.S degrees of computer science from KAIST, Korea in 2010 and 2012. Since 2012 he has been a researcher in the Sportainment Section, SW.Content Research Laboratory, Electronics and Telecommunications Research Institute (ETRI), Republic of Korea. His research interests are Digital Watermarking, Video Forensic, and Video Processing.



Kyoung Ill Kim received the Ph. D. degree in Computer Science from ChungNam National University, Daejeon, Korea, in 2014 and the M.S. degree in Computer Information System from Korea University, Chochiwon, Rep. of Korea, in 2001, respectively. Since 1983, he has worked as a principal researcher at Electronics and Telecommunications Research Institute (ETRI), Daejeon, Rep. of Korea. His research interests are cloud computing, visual communication and multimedia system.



Howon Kim received his BS degree in electronic engineering from Kyungpook National University in 1997, and his MS and PhD degree in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST) in 1999 and 2004, respectively. From Jul. 2004 to Jan. 2006, he worked at LG Electronics (CTO) as a senior researcher. Since Feb. 2006, he is currently working at Electronics and Telecommunications Research Institute (ETRI) as a principal researcher and the director of Sportainment Research Section. His current research interests include image-based 3D modeling, human motion tracking and augmented reality.