

A Comparative Study on Preprocessing Methods for Object Tracking in Sports Events

Sungwon MOON*, Jiwon LEE*, Dowon NAM*, Wonyoung YOO*, Wonjun KIM**

*ETRI(Electronics and Telecommunications Research Institute), Korea

**KONKUK University, Korea

moonstarry@etri.re.kr, ez1005@etri.re.kr, dwnam@etri.re.kr, zero2@etri.re.kr, wonjkim@konkuk.ac.kr

Abstract— Due to the development of machine learning, interest in automatic recognition of objects in videos is increasing. In the sports field, there is a growing demand for a technology for automatically recognizing and tracking various objects in a sports game video. However, it is very difficult to recognize and track an object without proper preprocessing due to its small size and large sports field. In this paper, we compare background subtraction and edge detection, which are typical preprocessing methods for automatic recognition and tracking of objects in sports images, and analyze advantages and disadvantages of each method.

Keywords— Multi object tracking, Visual tracking, Sports events, Video processing, Preprocessing

I. INTRODUCTION

With the rapid development of machine learning technology represented by deep learning, there is an increasing demand to automatically recognize and track objects in various videos. In the sports industry as well, there is a demand to support the work of the operator who is proceeding by the manual operation with the automated system. However, due to the nature of the sports video, it is difficult to automate the tracking of the objects in the video. The objects in the sports video are very small in size compared to the entire stadium, and unlike ordinary video objects, they are not only irregular and fast moving but also wear similar clothes as shown in Figure 1, making it difficult to apply the existing object tracking algorithm.

Applying the appropriate preprocessing algorithm can be a good solution to solve the above problem. Traditional preprocessing algorithms, background subtraction and edge map extraction, can effectively isolate the foreground region from the stadium area in the sports video. Attempts to recognize and track objects in binary video generated through preprocessing, such as background separation or edge map detection, are less difficult than object recognition in the original video. Therefore, the application of proper preprocessing algorithm for real - time sports video object tracking leads to application of object tracking algorithm with lower complexity.

In this paper, we apply the traditional preprocessing algorithm, Mixture of Gaussian(MOG) background subtraction[1] and Canny edge detector[2], to sports video

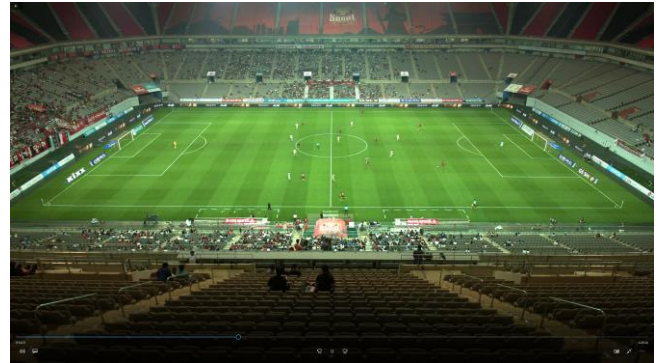


Figure 1. Example of soccer video

and compare the results. In addition, we apply the two methods to confirm the improvement of object tracking performance in real sports video.

This paper is organized as follows. We explain the previous works in section II. Experimental results are given in section III and concluding remarks are given in section IV.

II. RELATED WORKS

MOG2[3] is a traditional effective background subtraction algorithm. The algorithm is robust to a variety of disturbing factors such as slow lighting changes, shadows, and reflections. This algorithm is suitable for overcoming obstacles such as change of weather and brightness in the stadium and extracting the foreground region. Also, since the shape of the sports arena does not change much during the competition, background separation can be considered as a preprocessing method for object tracking. However, due to the nature of sports video, which is a very small and long-time object, disturbance caused by shaking of the camera temporarily fails to extract the foreground region, which is also a fatal drawback to continuous object tracking.

The Canny Edge Detector, an older edge extractor, offers excellent performance in edge extraction. Featuring low error rate, edge detection accuracy, and single edge point response, the Canny Edge Detector delivers excellent performance in edge extraction when applied to sports video. The extracted edges can be used to detect objects in sports video through connected component labeling. However, most sports stadium



Figure 2. Example of ETRI K-League database video

have a lot of lines, which is an obstacle that makes it difficult to detect only the desired objects, even after passing through the edge extractor.

Jung et al.[4] conducted a quantitative comparative evaluation study of background separation algorithms for soccer video analysis in 2017. This study applied FGD[5], GMG[6], KNN[7], MOG[1], and MOG2[3] to soccer video and compared the results. As a result, KNN is proposed as the best background separation algorithm because of similarity between foreground and ground truth. In this paper, we propose MOG2 as the most suitable algorithm for robustness and shadow region extraction in various situations.

III. EXPERIMENTAL RESULTS

The video used in the experiment is ETRI K-League database video in Figure 2, 22 players and 3 referees are shown in the video. The CPU used in the experiment is Intel i7-6500U CPU@2.50GHz and also implemented OpenCV 3.1.0 in Visual Studio 2015 environment for implementation. We used MOG2 as the background subtraction algorithm, Canny edge detection as the edge map extraction algorithm, and detected the objects using C Grana 's Connected component labelling method[8] for each result.

The preprocessing using MOG2 has good overall results and very high accuracy for object detection especially when the motion of objects is active. Instead, the method does not detect that object if the movement of a particular object, such as a goalkeeper, sharply decreases. Also, even after applying the mathematical morphology, the shape of the object is not clear and it is difficult to specify the exact position. In addition, when the camera shakes while taking a picture, it is difficult to trace a stable object because a failure occurs in the entire screen and a simultaneous detection of each object fails.

TABLE 1. COMPARISON BETWEEN THE PREPROCESSING METHODS

Method	Avg. false positive	Avg. false negative
MOG2	15.4%	11.4%
Canny	30.1%	31.8%
MOG2+Canny	13.6%	12.2%

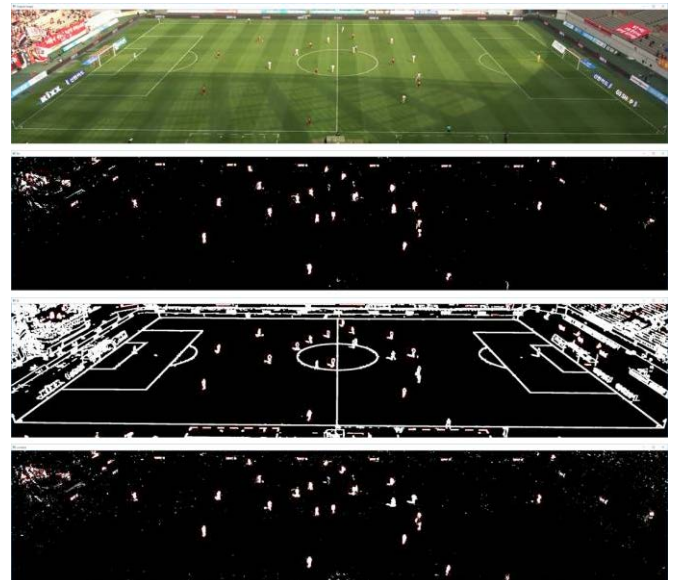


Figure 3. Example of original image, background subtraction image, edge map extraction image from the top in order

The method of using Canny edge detector has the advantage of knowing the exact position of the detected object because the shape of the object close to the ground truth can be confirmed. Apart from these advantages, it is difficult to detect the object when the object is near the line in the stadium due to the nature of detecting all edges in the video. Also, if the shape of the object in the stadium near the stadium is similar to the shape of the object to be detected, it may be unnecessarily detected.

The combined use of background separation and edge extraction can be a good preprocessing method for object tracking in sporting events. In the case of applying the background separation to the binary image, it is not much different from the result of applying the background separation only in the general case. However, when the camera is shaken, noise is not generated in the field, so that existing object tracking can be maintained. In addition, if an existing background separation image is additionally used for the corresponding image, the object in the form close to the edge extraction can be traced while the lines in the field are removed as shown in Figure 3.

If the resources are sufficient, it is possible to increase the success rate of object tracking by performing both preprocessing methods simultaneously and selectively using edge extraction and background separation results. Because the lines in the stadium are fixed, background detection results are used near those areas, and edge detection results can be used in other areas to improve both object detection and morphing accuracy.

IV. CONCLUSIONS

There are many preprocessing methods for object tracking in video, and it is important to select a suitable method because the results differ depending on the characteristics of video. In this paper, we compare the preprocessing method and the results of applying each preprocessing method which

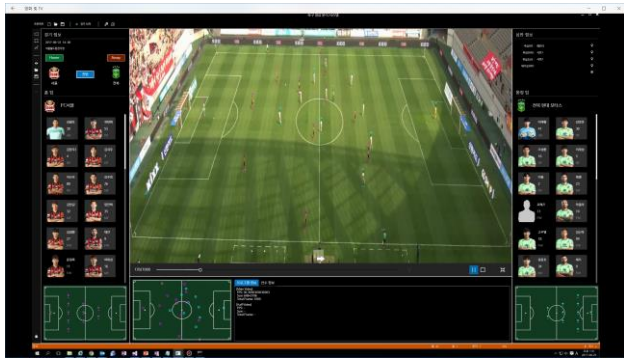


Figure 4. Tracking system with improved performance through preprocessing

can help the object tracking technology in sports image with wide stadium. As a result, it is confirmed that it is possible to enhance object tracking performance by simultaneously using background separation, object shape confirmation, and edge map extraction, which are advantageous for detecting moving objects and objects near lines, and advantageous for detecting moving objects. As shown in the Figure 4, the developed object tracking system showed about 15% improvement in object tracking performance when the preprocessing method was not applied.

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] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," in Proc. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), vol. 2, pp. 246-252, June 1999.
- [2] J. Canny, "A computational approach to edge detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 8, no. 6, pp. 679-698, June 1986.
- [3] Z. Zivkovic, "Improved adaptive gaussian mixture model for background subtraction," in Proc. IEEE Int. Conf. Pattern Recognition, Aug. 2004
- [4] Chanhoo Jung. "Objective Evaluation of Background Subtraction Algorithms for Soccer Video Analysis: An Experimental Comparative Study." The Journal of Korean Institute of Communications and Information Sciences, 42.1 (2017.1): 42-45.
- [5] L. Li, W. Huang, I. Gu, and Q. Tian, "Foreground object detection from videos containing complex background," in Proc. ACM Multimedia, pp. 2-10, 2003.
- [6] A. B. Godbehere, A. Matsukawa, and K. Goldberg, "Visual tracking of human visitors under variable-lighting conditions for a responsive audio art installation," in Proc. Am. Control Conf., Jun. 2012.
- [7] Z. Zivkovic and F. van der Heijden, "Efficient adaptive density estimation per image pixel for the task of background subtraction," Pattern Recognition Lett., vol. 27, no. 7, pp.773-780, 2006.



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 Reserch Laboratory, Electronics and Telecommunications Research Institute (ETRI), Republic of Korea. His research interests are Digital Watermarking, Video Forensic, and Video Processing.



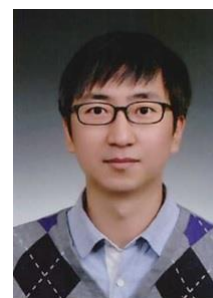
Jiwon Lee received the B. S. degree in Computer Engineering from Kyungpook National University, Republic of Korea, in 2008, and the Ph. D. degree in Computer Science from Korea Advanced institute of Science and Technology (KAIST), Republic of Korea, in 2013. Since 2013 he has been a senior researcher in the Sportainment Section, SW.Content Reserch 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.



Wonyoung Yoo received his Bachelor's, Master's and Ph. D. degrees in Electronic Engineering from Jeonbuk National University, Republic of Korea, in 1996, 1998 and 2003. since 2001 He is working as a Managing Director of Electronics and Telecommunication Research Institute(ETRI). His research interests include image/video watermarking, fingerprinting and video analysis



Wonjun Kim(Korea, 1980. 09. 06.) received the B.S. degree in electrical engineering from Sogang University, Seoul, Korea, the M.S. degree from the Department of Information and Communications, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, and the Ph.D. degree from the Department of Electrical Engineering, KAIST, in 2006, 2008, and 2012, respectively. From Sept. 2012 to Feb. 2016, he was a Research Staff Member in Samsung Advanced Institute of Technology (SAIT). Since Mar. 2016, he has been with the Department of Electronics Engineering, Konkuk University, Seoul, Korea, where he is currently an Assistant Professor. His research interests include image and video understanding, computer vision, pattern recognition, and biometrics, with an emphasis on saliency detection, face and action recognition, multiple object tracking, etc.