

Virtual Reality-based Sports Training System and Its Application to Baseball

Dan Mikami, Kosuke Takahashi, Naoki Saijo, Mariko Isogawa, Toshitaka Kimura, and Hideaki Kimata

Abstract

This article introduces a system we are developing that provides a first-person-view experience to users. It is aimed at preparing athletes to engage in sports competitions by enabling them to experience game situations from their own point of view before the game. The use of video has rapidly become popular in the sports field in various ways such as for scouting the opposing team before the game. Such videos are usually captured from a position different from that of players while they are playing the game. Thus, they cannot simulate the game very well. Our system provides a virtual reality-based first-person-view experience by using information captured or measured from locations that will not disturb the game. We describe in this article how our system worked in a trial with a professional baseball team.

Keywords: VR, sports training, baseball

1. Introduction

The use of information and communication technology has rapidly permeated into sports in recent years. One of the most popular uses is statistical analysis and its visualization because of the easy correspondence to big-data mining and broadcasting. In baseball, statistical analysis verifies factors that correlate to victory. One such type of analysis is known as sabermetrics and is mainly used for evaluation and recruiting in team operations.

In contrast, our research mainly focuses on supporting individual athletes to help them perform at their best. For sports in which players directly confront each other, for example, baseball, tennis, and soccer, it is important to not only improve the player's own skill but also to help players adjust to their opponents. We believe that this will be especially important for professional athletes because of the high skill level of their competitors.

For athletes to perform at their best in a game or sport, it is important for them to know their opponent.

Video-based scouting has been popularized for this purpose in recent years. It enables intuitive understanding of the opposition, which words or statistics cannot easily provide.

However, there are limitations to what can be achieved with the current video-based scouting. One such limitation is viewpoint. Because the scouting video is captured from outside of the baseball field, the viewpoints of the actual game situation and of the scouting video are different. Because of this difference, the scouting video does not fully support athletes in their pre-game or pre-sport preparation.

To address this, we tackled the problem of how to generate and provide a first-person-view experience before the game or sport was actually played. We have mainly focused on the sport of baseball in this article because it is one of the most appropriate ways in which the technology can be applied.

2. Background

In this section, we briefly explain some key features

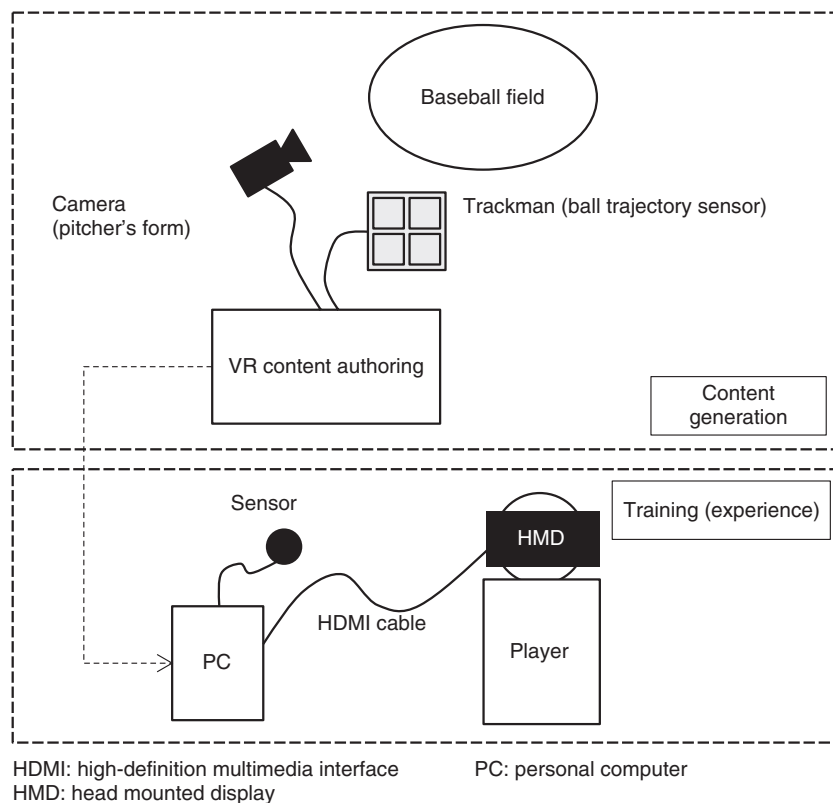


Fig. 1. Structure of VR-based imagery training system.

of batting in baseball and then discuss the required elements of the technology we are developing.

2.1 Features of baseball batting

A baseball game always starts when the pitcher throws the ball to the first batter. During a game, starting pitchers usually throw about 100 pitches, and batters typically have about three opportunities at bat. Therefore, the batters have to adjust to the way the pitcher throws the ball during the game. Radio broadcasters commenting on the game are likely to report on how batters adjust to the way the pitcher throws the ball the second or third time they face the pitcher. Batters can therefore use what they have learned from virtual reality (VR)-based technology to improve their batting performance.

2.2 Definition of requirements

We held discussions with professional players and coaches to determine the necessary capabilities of the system. We extracted the following four requirements:

- (1) The system can provide the ball trajectory

with depth information.

- (2) The system can reproduce the correct ball trajectory at arbitrary positions within the batter's box.
- (3) The system can be taken to any place and set up easily.
- (4) The system can be updated with the latest information about the pitcher.

In the next section, we describe the system we propose that fulfills these requirements.

3. VR-based imagery training system

The system structure is depicted in **Fig. 1**. The proposed system consists of two phases: a generation phase and an experience phase. The generation phase generates an experience in a three-dimensional (3D) virtual space based on a pitcher's motion and on ball tracking. We use captured videos to represent a pitcher's motions, while the ball is depicted by computer graphics (CG) based on the measured 3D ball position.

The experience phase provides an experience to

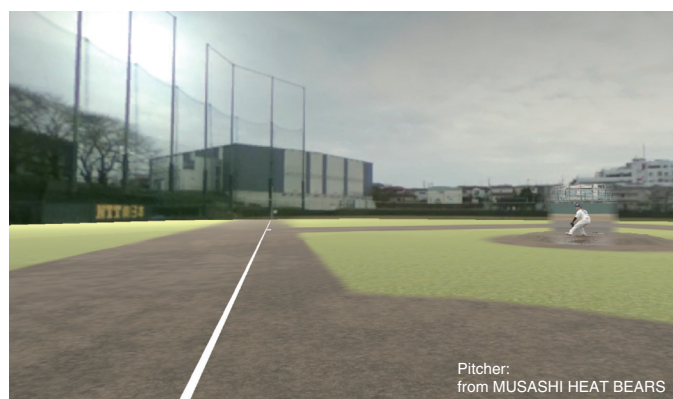


Fig. 2. Example snapshot of batter's view.

users through a head mounted display (HMD). The position and orientation of the HMD are measured so that the view that is rendered to the user changes depending on the user's posture. We used Oculus Rift as the HMD. Another well-known HMD with position tracking is VIVE, but we selected the Oculus Rift model because it is lighter. An example snapshot of a batter's view wearing the HMD is shown in **Fig. 2**. In actual use cases, the HMD displays two different images to each eye to provide parallax.

The system that we briefly introduce here effectively provides the correct ball trajectory from arbitrary viewpoints, and it can be used in arbitrary places because the HMD we adopted is highly portable. We describe the generation phase in the next section. This is the main contribution of this article from a technical aspect and mainly focuses on how easily updates can be made by using the latest information.

4. Generation phase

The generation phase updates VR content with the latest information. This makes it easy even for non-experts to update operations while maintaining stable quality. We adopted a hybrid approach that combines a billboard representation of a pitcher and a CG-drawn ball at a measured 3D position.

4.1 Billboard-based pitcher representation

There are several possible ways to represent a pitcher in 3D virtual space. In one 3D CG example, the pitcher's posture and texture are measured. However, it is difficult to stably update VR content using this solution because pitchers will almost never per-

mit their posture to be captured with a motion capture system. This is to prevent the opposing team from getting hold of such content and watching them before the game.

A well-known method called billboard-based video representation [1] does not rely on the pitcher's posture. Roughly speaking, one advantage of this method is that it mainly makes use of captured videos. Another advantage is that it can easily represent subtle nuances in pitchers such as changes in their facial expressions, which are very hard to measure and represent through the use of CG. We used these advantages to provide billboard-based representations of pitchers. First, we captured a video of the pitcher's motions from outside the playing field. We then placed a virtual panel (called a billboard) at the pitcher's location in a virtual stadium depicted by 3D CG. Finally, we played the captured video on the billboard. In baseball, the positions of the pitcher and batter do not significantly change. Thus, the billboard-based representation suits this task.

4.2 Representation of ball trajectory by CG

Unlike a pitcher's motion, the thrown ball is a significantly difficult target for the billboard approach to represent. This is because it includes large and abrupt position changes; a thrown ball moves about 18 m in about 0.5 s. Therefore, we do not use captured video directly to represent the ball. Instead, we first measure the ball trajectory and then use CG to render the ball. This solution enables us to provide an accurate ball trajectory at arbitrary positions from the batter's box area.*

However, a severe disadvantage occurs in combining billboard-based pitcher representation and



(a) Duplication of actual filmed ball and CG rendered ball



(b) Filmed ball has been removed

Fig. 3. Duplication of the filmed ball and the CG rendered ball, and the same image after removing the filmed ball.

CG-based ball rendering. That is, there is a duplication between the actually filmed ball and the CG rendered ball (**Fig. 3(a)**). To avoid this duplication, the filmed ball should be removed by image processing. We used the inpainting method proposed by Isogawa et al. [2] to remove the ball in the filmed video. We verified that it significantly improves the quality of experience, as shown in **Fig. 3(b)**.

5. Experiments and future direction

We verified the validity of our system with the cooperation of a team in the Nippon Professional Baseball (NPB) organization from two aspects. The first was the degree to which NPB players and coaches would accept it. The second was the system's operability, especially in updating VR content.

5.1 Acceptability

There are various indexes to evaluate batters; one of the most popular is batting average. However, we obtained users' (i.e., batters) subjective evaluations in a system trial conducted throughout a baseball season. This made it hard to make comparisons under a controlled setting. Furthermore, the task of successfully hitting a baseball is a quite uncertain one; a batter who gets a hit only three times in ten attempts is considered to be a very good hitter. This makes it very hard to evaluate the performance of our system on the basis of an objective score such as batting average.

(1) Depth sensing

Almost all professional baseball players have used videos to check out their opponents. However, watching videos on a 2D display does not give sufficient

depth information since it provides neither binocular parallax nor motion parallax. In contrast, the system we propose reproduces 3D ball positions in virtual space and renders videos depending on the position and orientation of the HMD. It also renders them for both the right eye and left eye positions. Thus, it provides an HMD experience with depth cues, that is, binocular parallax and motion parallax. Many athletes who used our system said that with it, they really felt that the ball was coming from the pitcher.

(2) Simulation training for adjusting to pitcher's motion

Many batters say that it is very helpful for them to be able to see simulations of the opposing pitcher before the game. This enables them, for example, to check on how they will adjust to the pitchers' motions based on the changes the pitchers make in the speeds at which they throw the ball. They feel that this will allow them to make full use of their batting chances, even in their first time at bat.

(3) The system as a communication device

We received a comment from a staff member who said that using this system made it possible to expect better communication with players since both coaches and players could get the same ball trajectory.

However, some points arose that need to be addressed, including the low resolution provided due to the use of the HMD and the insufficient field of view for batting. One of the tasks we will tackle in the future is to carry out a more detailed analysis of the

* Ball tracking systems have been rapidly popularized in baseball stadiums. We believe it will be possible to accurately reproduce ball trajectories within a few years.

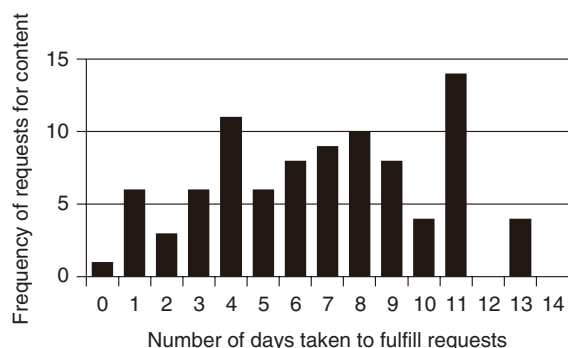


Fig. 4. Elapsed time between receiving a request for content from a team and the provision of the content.

system. The most important but most difficult thing we need to achieve is how to objectively measure the perceived correctness of the batter's experience. We will need to tackle this point to further develop the system.

5.2 Operation ability throughout a season

We conducted a trial of the system throughout the 2016 baseball season. The season comprised 143 games, but some of them were lacking in data to create VR content. Nevertheless, we created VR content for 96 pitchers, who threw an average of 15.6 pitches per game.

A histogram of the elapsed time between a request from a team for content and the time we returned the completed VR content to the team is shown in **Fig. 4**. Note that we could have returned the content sooner than we did if it had been necessary to meet the

team's schedule. Sometimes, due to schedule changes or other factors, we were requested to provide content within a short time, even within a single day. The relatively high per-day submission frequency shown in the figure indicates that with our system we were able to handle such requests appropriately.

6. Summary

This article described a VR-based imagery training system we are developing and experiments we conducted on it in cooperation with a professional baseball team. We believe that our use of VR will provide people with advance experience in improving their actions in the real world. This will help them to make decisions under severe temporal constraints and thus will be helpful for various VR developers.

Future tasks will include clearly detailing factors such as HMD resolution, field of view angles, and system delay. This will make it essential to conduct experiments using athletes as subjects. We will attempt to carry out such experiments in a way that will mutually benefit the athletes and ourselves.

References

- [1] Y. Ohta, I. Kitahara, Y. Kameda, H. Ishikawa, and T. Koyama, "Live 3D Video in Soccer Stadium," *International Journal of Computer Vision (IJCV)*, Vol. 75, No. 1, pp. 173–187, 2007.
- [2] M. Isogawa, D. Mikami, K. Takahashi, and A. Kojima, "Image and Video Completion via Feature Reduction and Compensation," *Multi-media Tools and Applications*, Vol. 76, No. 7, pp. 9443–9462, 2016.

Trademark notes

All brand names, product names, and company names that appear in this article are trademarks or registered trademarks of their respective owners.



Dan Mikami

Senior Research Scientist, Visual Media Project, NTT Media Intelligence Laboratories.

He received a B.E. and M.E. from Keio University, Kanagawa, in 2000 and 2002, and a Ph.D. from the University of Tsukuba in 2012. He joined NTT in 2002. His current research activities are mainly focused on computer vision and virtual reality. He was awarded the Meeting on Image Recognition and Understanding 2009 Excellent Paper Award 2009, the IEICE (Institute of Electronics, Information and Communication Engineers) Best Paper Award 2010, the IEICE KIYASU Zen'iti Award 2010, the IPSJ SIG-CDS (Information Processing Society of Japan, Special Interest Group on Consumer Devices & Systems) Excellent Paper Award 2013, and the IEICE Human Communication Award 2015.



Kosuke Takahashi

Research Scientist, Visual Media Project, NTT Media Intelligence Laboratories.

He received a B.Sc. in engineering and an M.Sc. in informatics from Kyoto University in 2010 and 2012. He joined NTT Media Intelligence Laboratories in 2012. His research interests include computer vision. He received the Best Open Source Code award Second Prize at the 2012 IEEE Conference on Computer Vision and Pattern Recognition.



Naoki Saijo

Senior Research Scientist, Sports Brain Science Project, NTT Communication Science Laboratories.

He received a B.E. and M.E. in mechanical engineering from Waseda University, Tokyo, in 1999 and 2001, and a Ph.D. from Tokyo Institute of Technology in 2010. He joined NTT in 2001 and has been engaged in research on human information processing. His research interests include human sensorimotor mechanisms, especially visuomotor control properties. He is a member of the Society for Neuroscience (SfN), the Japan Neuroscience Society (JNS), the Society of Instrument and Control Engineers, and IEICE.



Mariko Isogawa

Researcher, NTT Media Intelligence Laboratories.

She received her B.S. and M.S. from Osaka University in 2011 and 2013. She is currently a researcher at NTT Media Intelligence Laboratories and has been a Ph.D. candidate at Osaka University since 2016. Her research interests include computer vision topics such as image completion and quality assessment.



Toshitaka Kimura

Senior Research Scientist, Human Information Science Laboratory, NTT Communication Science Laboratories.

He received a Ph.D. in human neuroscience from the Graduate School of Arts and Sciences, the University of Tokyo, in 2003. He joined NTT Communication Science Laboratories in 2003. He has been engaged in research on sensorimotor control and action-perception interaction in humans. He has recently been focusing on brain functions in athletes and the use of information and communication technology in sports. He is a member of SfN, JNS, and IEICE.



Hideaki Kimata

Senior Research Engineer, Supervisor, High-Reality Visual Communication Group Leader, NTT Media Intelligence Laboratories.

He received a B.E. and M.E. in applied physics in 1993 and 1995 and a Ph.D. in electrical engineering in 2006 from Nagoya University, Aichi. He joined NTT in 1995 and has been researching and developing a video coding algorithm, visual communication systems, and machine learning. He is a Chief Examiner of IPSJ Special Interest Group on Audio Visual and Multimedia Information Processing.