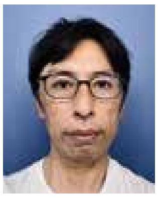
Broadcasting Professional Baseball by using Volumetric-video Technology

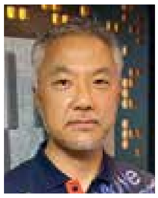
Takayuki Shinoda
Nippon Television Network Corporation

Satoshi Sasaki
Nippon Television Network Corporation

Hajime Sato
Canon Inc.



Yuji Watanabe
Nippon Television Network Corporation



Taiji Kamiya
Canon Inc.

Atsushi Date
Canon Inc.

1. Introduction

Seventy years have passed since Nippon Television (NTV) became the first commercial broadcaster to broadcast professional baseball on August 29, 1953. What started as a live broadcast with two cameras has evolved to the point that approximately 100 specialized cameras can instantly generate “volumetric video” (or “free-viewpoint video image”) that provide a special viewing experience (Figure 1). Introduced for baseball broadcasts as a world’s first in April 2022, free-viewpoint video image was implemented for five games in the 2022 season and for all 68 home games (at Tokyo Dome) of the Yomiuri Giants in the 2023 season.

This implementation of free-viewpoint video image has made it possible to generate video replays of any scene, including batting, running between bases, and brilliant defensive coordination, with 360-degree free camerawork. It creates a series of video images from viewpoints that cannot be captured with normal cameras. The series may include video images from the perspective of players that make the viewer feel as if they were on the field or highlights that stop at the moment of an outstanding play and move around freely up, down, left, and right.

■ Figure 1: Free-viewpoint video images in professional baseball broadcast

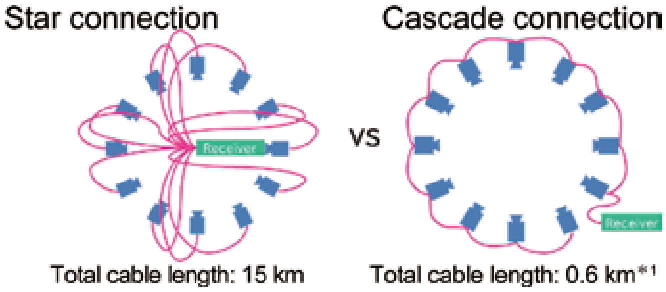
2. Overview of volumetric-video technology

Volumetric-video technology is a means of capturing images of a subject with a number of cameras and generating 3D spatial data from the captured images. Multiple cameras surrounding the subject capture all temporal and spatial changes and generate 3D models. The texture at the time of the captured pictures is attached to the generated 3D models, and free-viewpoint videos images are generated by providing viewpoints from any position or angle in the space.

3. Three features of the implemented volumetric-video system

The three main features of the implemented volumetric-video system are described as follows. The first feature is the method of connecting the cameras. To configure the system by installing multiple cameras in a stadium or arena, it is necessary to wire each camera by cable to a dedicated outside-broadcasting van housing a video-generation server. In that case, if a “star-type connection” is used to connect the cameras and server on a one-to-one basis, the number of cables and distances required would be enormous. If the total circumference of the stadium is estimated as 600 m, the total cable length would be 15 km. Given that fact, we configured the implemented system as a "cascade connection"; that is, the cameras are connected to each other and data is transmitted in a "bucket relay" fashion. This configuration makes it possible to reduce the total cable length and significantly shorten the installation period of the system (Figure 2).

■ **Figure 2: Star- and cascade-type cable connections**

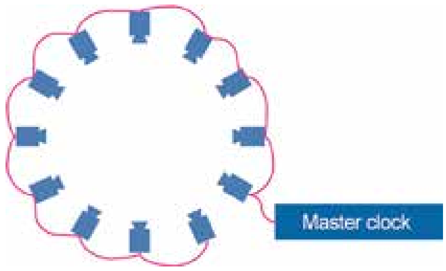


The second feature is a system for highly accurate synchronization of the cameras. Generating free-viewpoint videos images for sports events necessitates accurately capturing fast-moving objects such as bats and balls. To meet that need, it is necessary to synchronize a large number of installed cameras with high precision and capture images accurately at any given moment and without any misalignment of the cameras at any installation point. In response to that necessity, Canon has developed a high-precision video-recording technology that uses a master clock on the network to control a large number of cascaded cameras and synchronize the recording timing of each camera to the order of microseconds (Figure 3). This technology has made it possible to capture the high-speed pitches, bat swings, and struck balls of professional baseball players accurately and generate free-viewpoint TV videos images.

The third feature is the wide image-capturing range.

While working on professional baseball broadcasts, we came to realize that a valuable highlight scene is not only the moment of the batter hitting the ball but also the entire play, including the action leading up to the outs during a double play. In other words, since incredible plays can occur anywhere on the field, homogeneous images must be generated across the entire field. Expanding the image-capturing area to the entire field by using the viewpoint method would require a huge number of cameras, which is unrealistic. Accordingly, through detailed simulations of camera placement, we developed a "full-field method" that covers the entire field with the optimal number of cameras. This new method makes it possible to generate high-quality volumetric-video images from anywhere in the image-capturing area.

■ **Figure 3: High-precision synchronization of cameras**

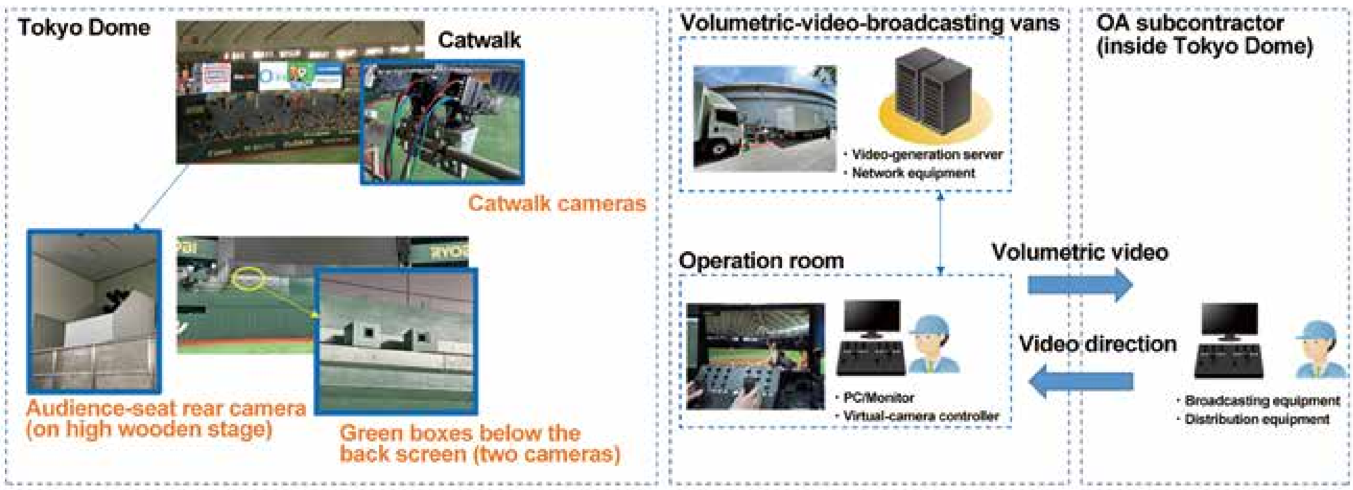


4. Configuration of system

The configuration of the system used for capturing volumetric-video images during a live baseball game at Tokyo Dome is shown schematically in Figure 4. The area from which volumetric video images were captured was the infield portion of the Dome's baseball field, which was covered by a determined number and placement of cameras. The cameras and their control boxes were clamped to the railing of the catwalk at the top of Tokyo Dome. Below the giant screens, which are inaccessible from the catwalk, a temporary wooden stage above and behind the audience seating area was constructed, and cameras were installed on it. With prior permission from the relevant authority, cameras were also installed in green boxes under the back screen in a way that would not obstruct the view of the players.

For rehearsals and the actual game, an outside-broadcast van was brought in from Canon and parked outside Tokyo Dome, and the free-viewpoint video images were generated from the van. The van was equipped with servers for generating free-viewpoint video images and connected to approximately 100 sets of cameras and their control boxes attached to the catwalk by optical cable via a network switch between them. A system-control unit was also installed in an operation room in the Dome, where Canon's technicians took charge of system operation. For

■ **Figure 4: Optimal placement of cameras**



generating the volumetric-video images, it is necessary to operate a virtual camera; accordingly, NTV's camera operators were trained to generate two streams of free-viewpoint video images under the instructions of a director so as to generate 1080i video and 48-kHz/24-bit audio data. One of these streams continuously assigns camera-motion paths from the start to the end of the game and continues to generate free-viewpoint video every about three seconds, and the other stream assigns camera-motion paths to the immediately preceding play and generates replay video. The audio data collected by NTV is imported into the volumetric-video-generation servers and synchronized by the servers to generate the free-viewpoint video.

5. Graphics

The volumetric-video images of the players are backed by a view of the Tokyo Dome, which is also computer-generated. The key to making a live baseball game look realistic is the realism of the 3D graphics. The space for the volumetric-video images is constructed from 3D data and textures scanned inside Tokyo Dome in advance. To minimize the difference between the volumetric-video images and images captured by regular broadcast cameras, we paid close attention to details like grass and soil, and we individually adjusted each object, repeatedly checking and correcting them, to create a space that looked like the inside of the real Tokyo Dome (Figure 5).

■ Figure 5: Created CG images of realistic field and bases

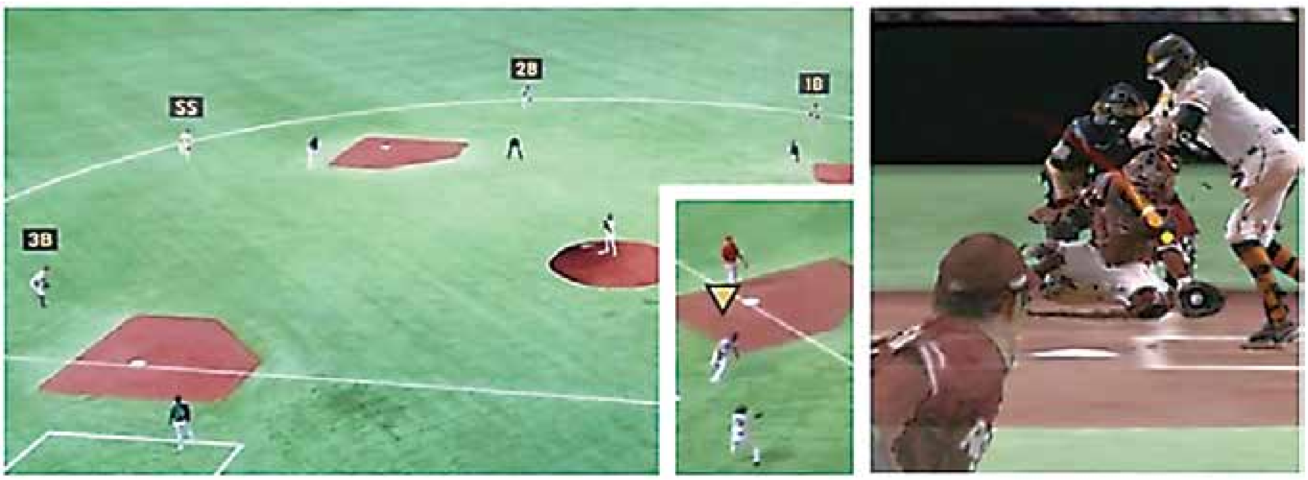
From the opening-season game in 2023, a dedicated camera was installed for capturing real-time texture with respect to the giant-screen content above the back screen and spectator seating, and the captured textures were texture-mapped to a background model to reproduce the giant screens, ribbon display, and spectator seating. The player-introduction video and direction of cheering are reflected realistically, and the excitement of the ballpark is conveyed as realistically as possible by the realistic background (Figures 6 and 7).

■ Figure 6: Camera for capturing background texture (blue frames)

■ Figure 7: Realistic background and placement of imaginary objects

Objects that do not actually exist inside the stadium, such as airships and banners, were placed in the volumetric-video space, and they were also used for program production and advertisements. AI recognition of objects in the space is combined with advanced CG expressions such as drawing the trajectory of pitches, tracking of players in wide-angle images to make it easier to see defensive formations, and tracking of markers in accordance with the movements of runners of interest (Figure 8).

■ Figure 8: AI representation of pitch trajectory and player tracking



6. Operation

Although a lot of know-how about live broadcasting using volumetric video is available, camera operation in virtual space is focused on hereafter. To manipulate the camera in virtual space, eight parameters are used: “x-axis,” “y-axis,” “z-axis,” “pan,” “tilt,” “roll,” “forwards/backwards,” and “zoom,” and one person must control them simultaneously using their fingers only (Figure 9). In the case of recording with a normal broadcast camera, although multiple operations may be performed simultaneously, the physical limitations on the camera’s range of motion make it easy to get an idea of the camera’s movement. However, in the case of recording with the virtual-space broadcast camera, the free viewpoint allows the camera to move infinitely within the virtual space; as a result, it is very difficult to get a sense of where the camera is positioned with respect to its central axis, and without sufficient training, the production is not viable. It therefore took some time to get used to the camera while adjusting the speed and sensitivity of the movement for each parameter.

Figure 9: Operation status

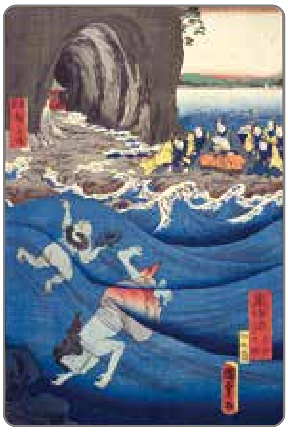
Moreover, the production director had too much freedom in using words to convey their own impressions of the volumetric-video images, and it sometimes took time for communication among the people in charge to get used to the system. More than a regular program, a live volumetric-video broadcast is not possible without each person having a common purpose and visual impression. Our know-how and experience inherited from 70 years of baseball broadcasting and the great efforts of numerous personnel with a common goal made it possible to take on new creative challenges and provide a new viewing experience.

7. Conclusion

The number of cameras set up around the catwalk giant display in the Tokyo Dome started with 87 and exceeded 100 as we improved the volumetric-video technology. The 3D models created instantly from the TV images captured by these cameras produced a series of images from angles and perspectives never seen before in a manner that lead to the creation of a new viewing experience. Being not limited to baseball, this volumetric-video technology could also be applied to other sports. We will continue to be constantly aware of “new perspectives” and strive to produce programs that convey the fun of sports even more than before.

Finally, we thank all those involved for their cooperation in the above-described broadcasting through the introduction of volumetric-video (free-viewpoint video image) technology.

Cover Art



Enoshima, from Famous Views of Tokaido Road

Utagawa Kunisada (1786-1864)

Source: National Diet Library, NDL Image Bank (https://rnavi.ndl.go.jp/imagebank/)