DOI: 10.20472/AHC.2016.001.004

TEPPEI ERIGUCHI
Demand Co.,Ltd., Japan

FREE VIEWPOINT REAL VIDEO STREAMING SYSTEM FOR 360° THREE-DIMENSIONAL VIEWING

Abstract:

The goal of this study was to develop a unique free viewpoint real video technology for 360° three-dimensional (3-D) viewing of photographic subjects and to utilize it in online video transmissions. In existing multi-angle video systems, the simultaneous synchronization and playback of large volumes of video images significantly burdens the central processing unit (CPU). Because repeated play and pause processing of multiple videos must be performed, a time lag occurs with every change in viewpoint. In this context, smooth transitions between videos of multiple viewpoints are challenged.

In response to this situation, the author developed a method to create a single composite image of a subject photographed in 360° using multiple cameras and display the image during transmission by partially trimming it and sliding the display positioning under the direction of the user.

This method resulted in a system that enables not only smooth viewpoint switching and 360° 3-D viewing of photographic subjects; it does so at a quality and size possible for online streaming. Therefore, in 2013, a Japanese patent was granted for this technology, it was used in an advertising promotion for Sharp’s 4K television, and it was used in the television Asahi web program “Danceta!” Additional patents were granted in 2015 in Korea, China, and the United States. The European patent is pending.

Keywords:

virtual reality, free viewpoint real video, three-dimensional viewing, patent

JEL Classification: C60, M37

 ISBN 978-80-87927-24-3, IISES

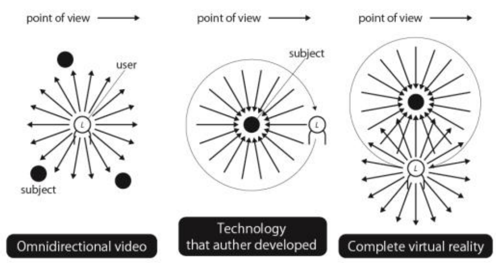
Introduction

This study involved the development of video streaming technology and the analysis of the results of streaming a Stock Video that enables 360° three-dimensional viewing that was realized in June 2013. In particular, this study discusses the problems that were associated with multi-angle video technology that existed in 2012 and earlier, the solutions that address these problems, and the unique imaging technology that was born from there. The author acquired a patent for this technology in Japan in 2013. In 2015, patents were granted in the US, South Korea, and China, and a European patent is pending[1].

In 2015, YouTube started offering the delivery and posts of omnidirectional videos that can be viewed at an angle of 360 degrees with the user at the center, and this was followed in 2016 by Facebook2. Other media have since been enabled for presenting omnidirectional video, known as VR（Virtual Reality）content. Many reports about the expansion and expectations of the virtual reality video market have been published3. However, the implementation of virtual reality in the true sense has some essential requirements. An omnidirectional video is a live-action video rather than being CG. As shown in Figure 1, the video can be viewed 360° around the user. The technique can be understood by imagining the ability to view any object from the surrounding 360°. VR is the fusion of the two technologies (omnidirectionality and virtual reality).

The purpose of this study is to analyze the potential of this technology, which allows any object to be viewed from the surrounding 360°, to realize true virtual reality.

Figure 1. True virtual reality and its enabling video technology



Challenges of existing multi-angle video technology

The realization of a video that allows objects to be viewed at 360° requires some conditions to be met. First, as shown in Figure 2, there are a certain number of viewpoints, each of which needs to maintain the correct angle of view as well as the angle and the distance to the subject. In addition, as shown in Figure 3, it is necessary to perform continuous changeover of the viewpoint at a frame rate that is less than the speed of the video. It is only necessary to align these two conditions to enable users to obtain a smooth video experience, such as wrapping around the periphery of the subject. As a result, the user experiences an object in three dimensions.

Figure 2 Point of view to maintain the exact angle of view and the angle and the distance to the subject

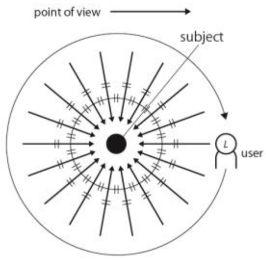
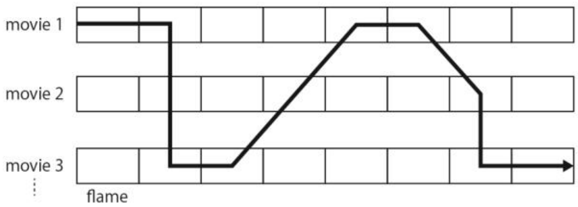


Figure 3 Continuous viewpoint switching at a frame rate less than the speed of the video



On the other hand, in the years since 2012, as multiple Web sites and DVD packages contained content that allowed multi-angle images to be viewed from multiple angles. Content such as this is created even now. However, as shown in Figure 4, only those videos, in which a subject is simply viewed from multiple angles, were not used to view the object stereoscopically. Thus, there is no causal relationship between the different viewpoints; instead, this relationship simply provided the difference between the distance and angle of the object. Moreover, aggressive technology development has not been performed, because the speed of response to viewer's instructions when switching images from one perspective to another viewpoint is not important. As shown in Figure 5, these existing technologies have been created using either of the following two methods. The first way is to display a large number of images simultaneously, and to switch the image displayed on the front in accordance with the viewer’s instruction. The second method requires a decision as to when to switch the point of view, repeat the play, and stop the video. In this case, one of the programs is loaded and sent to the CPU of the device for reproducing the video. The complexity of processing results in a time lag occurring in the viewpoint switching.

Figure 4 Point of view of existing multi-angle video

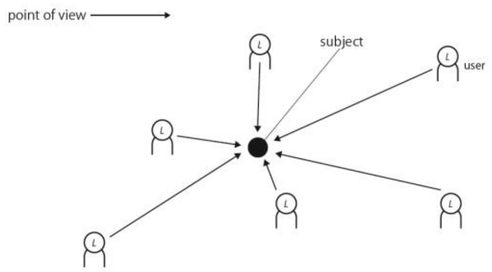
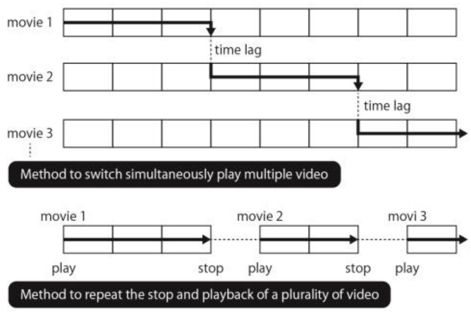


Figure 5 Procedure for playing existing multi-angle video



In other words, applying existing multi-angle video technology to produce video technology that can be viewed in a three-dimensional manner from 360° around the subject presents two major problems. The first is maintaining a constant distance to the subject to obtain an image of the same field angle from multiple locations. The second is to eliminate the time lag when switching the angle over multiple angles. The author has solved the challenges of existing multi-angle video technology by developing unique imaging method and data creation program.

Precise adjustment of the angle of view

At first, we developed a method for obtaining an image of the same field angle from multiple locations by maintaining a constant distance relative to the object. Achieving this necessitates the material that is captured by the plurality of cameras to be matched in a number of data points. These points are: the distance relative to the camera of the subject, the angle of the camera to the subject, length of the camera with each other, time synchronization between the camera, camera settings such as aperture and sensitivity, etc. Especially the first four points are important, because these are not only simple camera settings, they control the camera in dense measurement, which means it is necessary to install and set based on a certain rule. This was achieved by applying a method that adjusts the angle of view, which is also utilized in a time slice image.

We start from that match and the distance between the subject and the camera, and the distance between the different cameras. First, we install a guide shape as shown in Figure 6 in a position to reach the center of the object or multiple objects, after which we draw a circle around its guide. In addition, radially from the guide, we obtained a straight line at a constant interval, to install a camera in the position where the straight line and circle intersect. Thereby, the position of the subject and the camera is determined, as shown in Figure 7, where the distance between the subject and the camera, and the distance between the different cameras becomes constant.

Figure 6 Guide to be used at the time of shooting

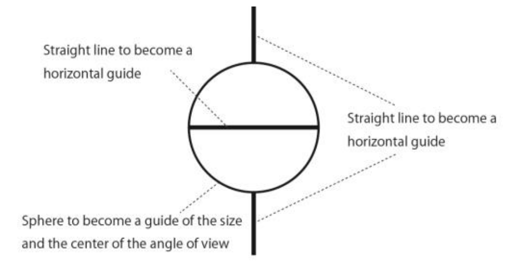
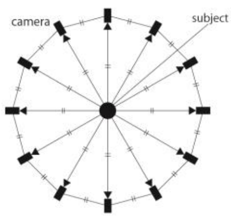
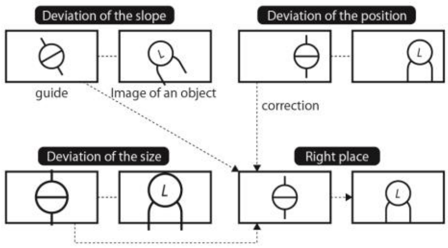


Figure 7 Distance between the camera and subject, and distance between different cameras



Secondly, we had to adjust the angle and the angle of view of the camera. While checking the camera image which serves as a reference, we determined the position and size of the guide on the viewfinder. While assuming the size of the subject, we adjusted the camera tilt and zoom to determine the angle of view of the reference. The inclination of the camera is perpendicular to the angle of view. Furthermore, as shown in Figure 8, while looking at the image on each of the viewfinders of the other cameras we adjusted the camera tilt and zoom, tailored to the camera as a reference. Upon completion, these values become the fixed inclination and the angle of each of the cameras fixed. Although the adjustment of the camera angle of view is basically completed, obtaining the effect of a smoother perspective switching requires adjustment at the time of editing, because it is difficult to make adjustments for fine displacement at the shooting stage.

Figure 8 Adjustment of the angle of view using the guide



Finally, we examined the time synchronization between the cameras. First, it is necessary to make sure that you can identify a particular frame by always using a flash of light for each take. More specifically, we captured movement to turn on or off the penlight on all cameras. When editing, looking for a timing that has disappeared or the timing at which the light was on, we adjusted the timing. Thus, the time required for synchronization between the cameras was recorded and we obtained the material necessary for a smooth angle switching.

Switching angles without a time lag

The second challenge was the development of a mechanism to ensure smoothness with no delay switching between multiple angles. Although this problem could be resolved by applying the theory, the reproduction of the streaming video on a PC required the amount of data to be reduced as much as possible. In turn, this necessitated a mechanism to achieve a simple data structure and programs. Therefore, the author created a composite video image from the material captured by a plurality of cameras to devise a method for displaying trimmed video on a PC. We verified the proposed mechanism by using material taken by 16 cameras of a scene including ballet dancers at an angle of 180°.

The overall flow of the editing process is as shown in Figure 9. First, higher resolution images with a size of about 4K that were taken from 16 angles were placed in a 4 × 4 arrangement starting from the upper left corner and ending at the lower right corner, as shown in Figure 10. As a result, the plurality of video material in a large video was used to create a synthetic image lined in a grid pattern. In addition, we recorded the coordinates of the center of the angles to obtain the resolution of each of the videos.

Figure 9 Flow diagram illustrating the shooting, editing, and programming

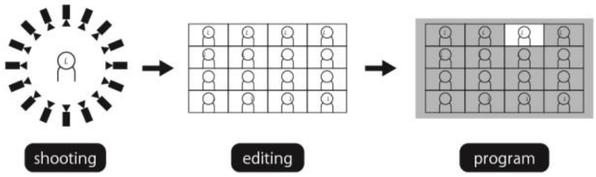
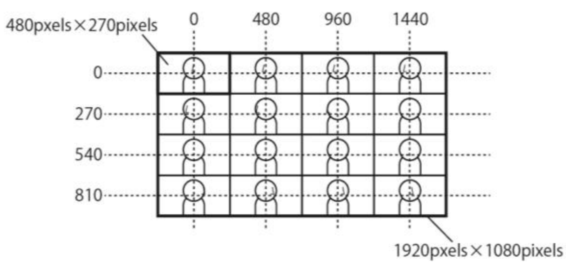
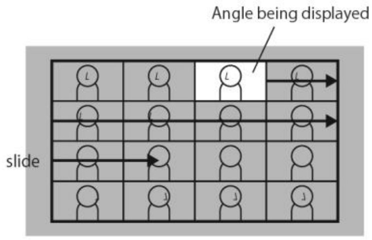


Figure 10 Structure of the data to be edited



Next, this high-resolution image is saved on the local hard disk of the PC, to construct a mechanism to display the material on the PC through a special program. First, the point to be played on a PC browser while reading the movie data is the same as for general video playback. However, in this test, as shown in Figure 11, while the video is being played, a mechanism is adopted to display the video on the screen by partial trimming. Specifically, of the 16 images arranged in the grid pattern, only one is displayed as determined by the program. Then, only the image acquired at angle 1 is displayed when the video playback is started. In addition, the user is able to continuously switch between viewing angles to switch the trimming position of the original material, in accordance with the resolution and the coordinates of each angle on the material obtained in the synthesis of the angle shown in Figure 10. This mechanism achieves a smooth changeover between the different points of view. At this time, the user who is viewing the image is not aware that the image acquired at 16 angles has been synthesized. In fact, it is simply perceived as though a video of more than one angle is switched on continuously. This enables the user to enjoy an experience similar to watching wraps around the subject from around 360°. The result is that a subject is experienced stereoscopically.

FIGURE 11. Works to be displayed by trimming one angle only



Streaming playback of a free viewpoint video, and the acquisition of a patent

The author addressed the two challenges by devising a solution based on a mechanism such as the above. Then, the method was tested in a stand-alone environment by placing the data on the local hard disk of the PC, to confirm proper operation. Subsequently, the data was actually placed on a streaming video server, to carry out tests while streaming the video over the Internet. This test was performed by using a program to identify problems such as the loading time while streaming the 4K video. The loading time was reduced by applying compression that allows playback to start while the data is being read. Further, a method that considers the length of the material and the user of the communication environment of the material was developed that provides a choice that has also been found effective for material acquired from a number of angles. The results of their verification confirmed that if there is a certain amount of loading time, then the streaming playback of a free viewpoint video over the Internet would work properly. The advantages of this method are that the number of images that are suppressed in one aspect of the image to be reproduced is that the processing change can be performed by simply changing the trimming position. Thus, the CPU is not burdened with playing a large number of images simultaneously as is the case with conventional multi-angle video techniques. Thus, there is no need to repeat the play and stop the video every time view switching occurs, which leads to the realization of smooth switching with no time lag angle.

On the basis of the verification result, the author submitted a patent application of this technology to the Japanese Patent Office, where an examination was carried out. After review, the Patent Office produced examples of other techniques that have already been filed and the author received a refusal notification from the Patent Office. However, these examples are not in the range of existing multi-angle technology described above; therefore, the differences between existing technologies and the newly developed technology were described in detail and were submitted as a written opinion to the notification. As a result, a patent was awarded on May 9, 2013. The patent document clearly describes the novelty and the inventive step against existing technology. After that, this technology was actually used for the first time on the Sharp 4K TV promotion site "AQUOS REAL LIVE".

AQUOS REAL LIVE and Danceta!

"AQUOS REAL LIVE" is the first application that utilizes the newly developed free viewpoint video4.

This Web site, as part of the promotion of SHARP's first 4K TV AQUOS UD1, was published on June 27, 2013. The realism of projecting the image to the 4K TV with normal full HD television has demanded a resolution that is four times the resolution of a technique for representing on the PC. This was achieved by adopting technology that enables the viewer to feel a sense of realism by watching free from the surrounding 360°.

In addition, as a motif to express the sense of realism, WORLD ORDER has been appointed, the world-famous dance performance group. Their performance was captured by 32 cameras shooting from around 360 °. Further, one camera was placed overhead to acquire images. The content was constructed based on material taken by a total of 33 cameras. The user is able to view theWORLD ORDER performance over approximately 1 minute 40 seconds around 360° by switching the viewpoint freely. This site has, owing to its innovative video technique, been featured in a number of media5, it has won several advertising awards6. This is considered a result of the new video technology that made possible the streaming playback of a free viewpoint video on the Web site that could not be realized until then.

Subsequently, a content that is published by the TV Asahi web video on June 1, 2015 "Danceta!" is another program which is constructed using the same technique7. In this content, Beat Buddy Boi, ALMA, Hilty & Bosch, etc., it is possible to watch a free dance performance of the top dance groups in Japan from around 360° on the PC. Furthermore, within this site, the user can also view another image using this technique. This video consists of another dancer adding commentary about the technique while showing a dancer's performance in the video from 360° by using a PC. This is different from the method utilized for the original technique and it can be said that the potential of this technique presents a further extension of the utilization method.

Conclusion

As mentioned at the beginning of this paper, the realization of virtual reality by live-action video in the real sense of the term, requires the video to be watched from around 360° around the subject as an essentiality. The author has developed proprietary video technology, to achieve a streaming playback on the PC of material such as "AQUOS REAL LIVE" and "Danceta!" Currently, users cannot select the central axis of the object. In addition, the construction of virtual reality space by live-action is problematic; for example, the fusion of this technology with the celestial omnidirectional video. However, there is no doubt it has become an integral step towards its realization.

The results that were achieved with the dance performance content have sufficiently shown the effectiveness of the free viewpoint video. Considering the extent of the effect of the video technology, in addition to dance content in the future, live music, sports, and educational content is expected to take advantage of a variety of situations.

References

Parisi, T 2015, *Learning Virtual Reality: Developing Immersive Experiences and Applications for Desktop, Web, and Mobile*, Sebastopol, CA : O'Reilly Media

Nite, S 2014, *Virtual Reality Insider: Guidebook for the VR Industry*, OPAL LOCKA, FL : New Dimension Entertainment

Miller, JB 2014, *Internet Technologies and Information Services*, Santa Barbara, CA : libraries unlimited

Allen, A 2016, *The Inevitable: Understanding the 12 Technological Forces That Will Shape Our Future*, New York, NY : Viking

Carr, N 2011, *The Shallows: What the Internet Is Doing to Our Brains*, New York, NY : W. W. Norton & Company

---

[1] Patent No. 40966285, International Application No. PCT / JP2012 / 074896.

[2] Page of 360 ° video content on YouTube.
https://www.youtube.com/channel/UCzuqhhs6NWbgTzMuM09WKDQ
Timeline of 360 ° videos on Facebook.
https://www.facebook.com/Facebook360/

[3] Forbes article to report that Mark Zuckerberg has been speaking about the VR and the SNS.（February 24, 2016）
http://www.forbes.com/sites/kathleenchaykowski/2016/02/24/mark-zuckerberg-has-a-plan-to-make-virtual-reality-social/#1045a6577959PLAY
Chinese People's Daily Online article to report that Lei Jun, CEO of Xiaomi, has been speaking about entry into the VR market.（January 19, 2016）
http://j.people.com.cn/n3/2016/0119/c94476-9005893.html
Nihon Keizai Shimbun article to report about the announcement of the VR headset "PS VR" to connect to the STATION4.（March 16, 2016）http://www.nikkei.com/article/DGXLASFK16H1I_W6A310C1000000/

[4] Content page of “AQUOS REAL LIVE” on the archive.
http://awards.1-10.com/2014/aquos-real-live/archives/

[5] Asahi Shimbun articles that cover a story about the publication of the “AQUOS REAL LIVE”.（2013 年 6 月 27 日）
http://www.asahi.com/tech_science/cnet/CNT201306270103.html

[6] AQOUS REAL LIVE page of on the Webby Awards website.
http://webbyawards.com/winners/2014/web/general-website/consumer-electronics/aquos-real-live/

[7] Content page of "Danceta!" on TV-Asahi web video。
http://www.tv-asahi.co.jp/douga_mv/danceta/