

Creation of Immersive UX Services

Ryuji Kubozono, Akihito Akutsu, Norihiko Matsuura, Kenichi Minami, and Akira Ono

Abstract

The NTT Group is researching and developing communication technology and media processing technology amid growing expectations for the creation of services that provide users with a highly realistic experience of events such as sports matches and live entertainment shows. We are also working with various corporate partners to conduct feasibility studies and service trials aimed at creating new value. This article describes the direction of our research and development aimed at evolving the key technologies needed to implement immersive user experience (UX) services, and presents an overview of our efforts to create entirely new immersive UX services.

Keywords: high reality, public viewing, live entertainment

1. Current state of media services

In the television (TV) corners of large consumer electronics retailers nowadays, most of the TV sets are 4K-compatible. With the ongoing advances in the resolution of consumer video cameras and digital cameras, we are now starting to see 4K TV broadcast services, and even trials of 8K broadcasting. The move to higher resolution is also driving development in the audio market, and it is expected that this trend will continue in the future as the world's media-related services and products migrate towards higher definition and higher resolution.

In addition to this market trend, 2016 has also been called the first year of the virtual reality (VR) era due to the emergence of various products and services using VR and/or augmented reality (AR), especially in the game/amusement sector. In addition to specialist VR/AR equipment, the market is now awash with inexpensive head-mounted displays (HMDs) and smartphone-based VR/AR applications, and we are starting to see a wide variety of content being aimed at these terminals. For example, Pokémon GO was a major success around the world and became the first such application to really capture the public's imagination.

Meanwhile, amusement parks and movie theaters aim to provide a more realistic viewing environment,

and in addition to making advances in three-dimensional (3D) video and higher-quality audio with more channels, systems such as MX4D™ [1] and 4DX® [2] are also incorporating other elements besides audio and video in order to stimulate the other senses with, for example, vibrations, water sprays, mists, or aromas. When combined with audio and video being displayed in front of the audience, these services provide an even greater sense of realism. Services of this sort have been a popular feature of amusement parks for many years, and efforts are being made to advance this technology into movie theaters so that a trip to the movies will become more of an experience than simply watching a film.

There are also facilities that achieve a heightened sense of realism with audio and video by projecting video productions onto 360° screens or dome-shaped screens, or by covering the audience's field of view with three screens (in front and on both sides) [3], and we are expecting to see more movie theaters and other entertainment facilities introducing immersive environments in the future.

2. Current state of the entertainment sector

In the field of sports, attending games and competitions has been a popular past time for many years. However, not everyone is able to attend such events at

the venues where they are occurring. Consequently, public viewings are becoming more common in Japan for games and competitions that attract a lot of interest, especially when Japanese teams are competing at international events. Various forms of public viewings are held in places such as sports stadiums and public halls so that the action can be enjoyed by fans who were unable to travel to the actual venue.

Meanwhile, movie theaters are showing non-movie content (ODS: other digital source) during intervals in movie screenings and are promoting measures aimed at attracting new customers. It is predicted that by 2020 this will have grown into a ¥63.3 billion market (including ¥31.8 billion for live broadcasts) [4]. Sports-related public viewings are often held free of charge, but we are starting to see viewings that charge an entrance fee. As progress is made in the resolution of issues relating to content rights, there will probably be a growing number of monetized cases.

In the area of music entertainment, the music industry has known for many years that sales of packaged music such as compact discs are set to decline and is shifting towards holding live events where music fans can get a once-in-a-lifetime experience even if the artist is playing the same set list.

Compared with western countries, Japan still has relatively strong sales of packaged music, but the number of live performances is rising, and it is no exaggeration to say that this is part of a major worldwide trend [5]. Although the live music market is continuing to grow, many concerts offer new experiences through the use of elaborate stage productions and gimmicks to ensure that the appetite of audiences for repeated viewings is undiminished. The latest technologies and production methods are being introduced for this purpose, and we can expect this trend to continue into the future.

Digital technology has also recently been used in various ways for productions other than music concerts. Many different types of works based on comics, films, and role-playing games are showcased on stage with the latest digital technology, and it is possible to reproduce the world view of the original creator to a high degree and to enhance the audience experience. The latest digital technology has also been incorporated into traditional stage performances. In a production of *The Tempest* at the Shakespeare Theatre in London [6], motion capture and projection mapping technologies were used to create mysterious special effects even within a traditional theatrical production. In Japan, performances that combine traditional

kabuki theater with the latest projection mapping techniques have attracted considerable interest.

3. NTT's vision of immersive UX services

At NTT, in light of the abovementioned market trends, we aim to create immersive user experience (UX) services to provide audiences with new experiences. To realize this goal, we aim to develop various applications, including *ultra-immersive public viewing*, where people can experience the strength and speed of the world's top sports players even when they are far away from the actual sporting venue, *ultra-immersive live viewing*, where people at a remote location can enjoy traditional performing arts in the same way as people in the theater or hall where the performance is actually taking place, *enhanced live performances*, where people can experience something newer than an ordinary live performance, and *stadium solutions*, where people can experience seeing a sports match in person, but with a great deal of added fun.

3.1 Ultra-immersive public viewing and live viewing

The aim of ultra-immersive public viewing and live viewing (**Fig. 1**) is to implement a world where sports events or live performances can be transmitted in their entirety to remote locations, where people can experience the events as if they were there themselves. During large international sporting events, we already have public viewings using video footage displayed on large screens in stadiums throughout Japan. However, this is not a highly immersive experience. Instead, we aim to use a combination of display control technology that enables people to view the event as if they were in the same venue as the sports players or performers, presentation control technology that enables the audience to feel as if they had entered an actual sports venue, and audio technology that enables them to experience new sounds. Furthermore, we aim to provide a new viewing experience by detecting and tracking the subjects appearing in the video images and linking them to various kinds of metadata, resulting in a service that would be impossible to implement with TV broadcasting.

3.2 Enhanced live performances

We are looking at ways of providing greater emotional impact in traditional performing arts (**Fig. 2**). Since 2016, NTT has been providing technology to the Cho-Kabuki series of events in partnership with

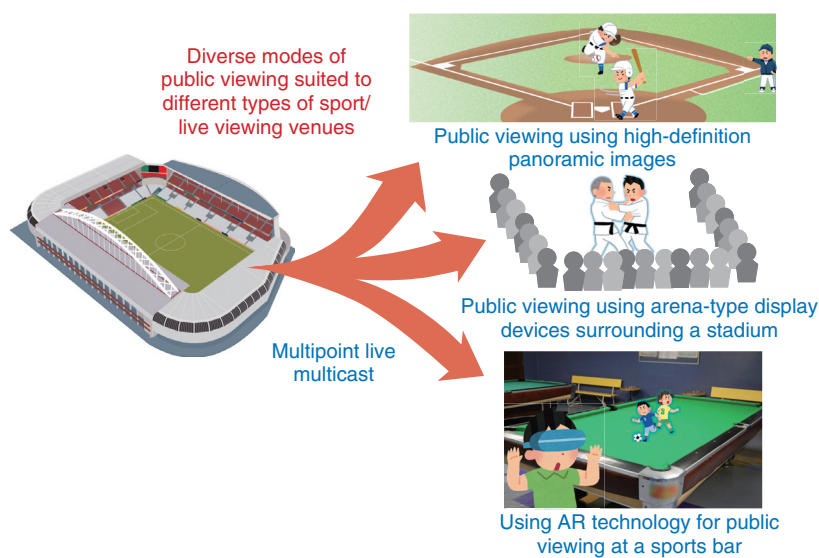


Fig. 1. Ultra-immersive public viewing concept.

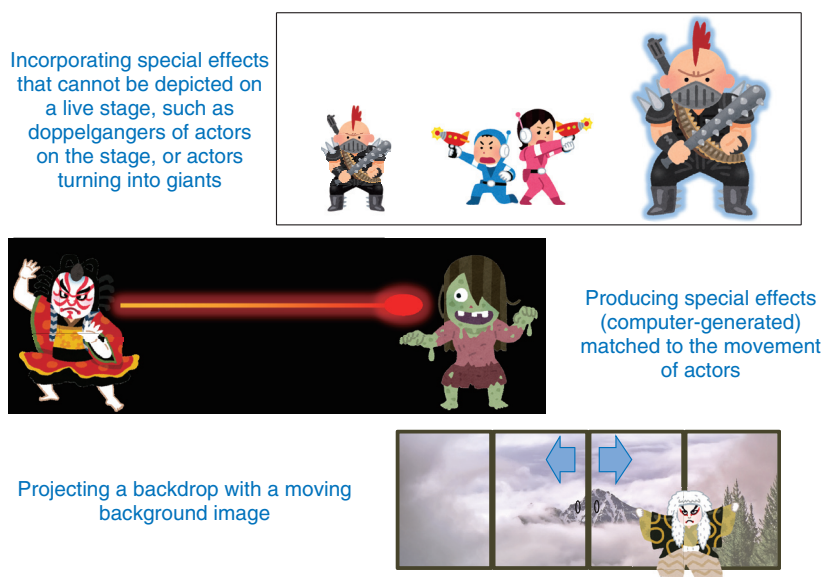


Fig. 2. Enhanced live performances.

Dwango Co., Ltd. and Shochiku Co., Ltd. in order to provide new performance events by combining the traditional art of kabuki with modern technology. We have also been working at identifying potential applications for this technology and extracting new technical issues.

3.3 Stadium solutions

A stadium solution (**Fig. 3**) is a way of entertaining spectators at a sports event before and after the game, and increasing their enjoyment of the game itself so that people will want to visit the stadium even on days when no games are taking place. In this way, we aim to revitalize stadiums and their surroundings. For example, NTT is working on the creation of a showcase

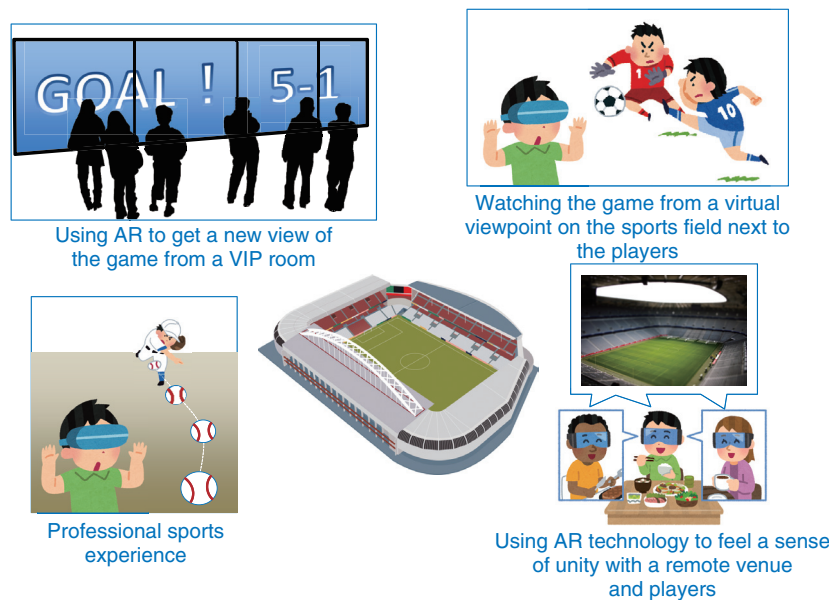


Fig. 3. Stadium solution concept.

where people can use VR technology to experience what it would be like to face balls thrown or hit towards them by professional baseball and tennis players, and is evaluating its feasibility as a VR experience.

4. Main technological components of an immersive UX service

We can provide new ways to enjoy sports by allowing people to see and hear things that they would not normally be able to experience. This opens the door to a completely new way of looking at sporting events. One technology that can be used to implement such a system is called free-viewpoint video synthesis technology [7]. This enables people wearing HMDs to watch sports from the viewpoint of a restricted part of the stadium that they cannot enter. The use of high-speed cameras makes it possible to show viewers the exact course traveled by the ball and players, and high frame rate video encoding technology can be used to compress the video pictures with high definition and a high frame rate to show subtle movements and differences in dynamism [8].

In addition, the use of surround video stitching and synchronous transmission technology makes it possible to realize a super high definition panoramic video picture covering the entire field of view instead of simply presenting the pictures on a large TV screen

as in previous public viewings [9]. The use of wave field synthesis technology can position a sound image at the location of the subjects shown in the video, or at a position away from the front of the display screen, enabling multiple audience members to experience relayed sound without having to provide separate headphones or other audio devices for each individual [10].

Furthermore, the use of goggle-free 3D video screen technology that enables the viewing of stereoscopic images without having to wear 3D goggles will make it possible to provide a completely new kind of highly immersive experience [11]. To convert images into 3D, it is necessary to extract the subject of interest from video pictures captured from many different angles by means of arbitrary background real-time object extraction technology [12]. For sports and entertainment events, we must consider how accurately this can be done in real time against backgrounds that contain moving objects.

In addition, by displaying various additional information together with the audiovisual content, we can enhance people's enjoyment of sports and entertainment events and help them to learn about the players and performers. This increases the appeal of relayed broadcast events. An important component of such systems would be moving object detection technology that can detect the positions and postures of players on a sports field and of performers on a stage [13].

We expect that this will be the subject of active research and development in the future.

5. Future prospects

In addition to the technologies introduced here, the NTT Group is pursuing open innovation in various fields to create new kinds of immersive UX services based on diverse media processing and communication technologies. As a result, although we are currently pursuing many technologies with a view to business development, there are still many others that remain at the level of feasibility studies. Going forward, we will continue to pursue open innovation with various players in order to establish media processing and communication technologies and create services that can be used to address social issues and stimulate local growth.

References

- [1] Website of Toho Cinemas' MX4D (in Japanese), <https://www.tohotheater.jp/service/mx4d/>
- [2] Website of United Cinemas' 4DX (in Japanese), <http://www.unitedcinemas.jp/4dx/>
- [3] Website of United Cinemas' ScreenX (in Japanese), <http://www.unitedcinemas.jp/screenx/>
- [4] Press release issued by GEM Standard on February 3, 2016 (in Japanese). <https://gem-standard.com/information/100>
- [5] PIA Research Institute, "2016 White Paper on Live Entertainment," 2016 (in Japanese).
- [6] Website of d3 Technologies, "The Tempest," <http://www.d3technologies.com/en/showcases/theatre/the-tempest/>
- [7] K. Okami, K. Takeuchi, M. Isogai, and H. Kimata, "Free-viewpoint Video Synthesis Technology for a New Video Viewing Experience," NTT Technical Review, Vol. 15, No. 12, 2017. <https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201712fa2.html>
- [8] Y. Omori, T. Onishi, H. Iwasaki, and A. Shimizu, "A 120 fps High Frame Rate Real-time Video Encoder," NTT Technical Review, Vol. 15, No. 12, 2017. <https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201712fa3.html>
- [9] T. Sato, K. Namba, M. Ono, Y. Kikuchi, T. Yamaguchi, and A. Ono, "Surround Video Stitching and Synchronous Transmission Technology for Immersive Live Broadcasting of Entire Sports Venues," NTT Technical Review, Vol. 15, No. 12, 2017. <https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201712fa4.html>
- [10] K. Tsutsumi and H. Takada, "Powerful Sound Effects at Audience Seats by Wave Field Synthesis," NTT Technical Review, Vol. 15, No. 12, 2017. <https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201712fa5.html>
- [11] M. Makiguchi and H. Takada, "Smooth Motion Parallax Glassless 3D Screen System that Uses Few Projectors," NTT Technical Review, Vol. 15, No. 12, 2017. <https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201712fa6.html>
- [12] H. Nagata, H. Miyashita, H. Kakinuma, and M. Yamaguchi, "Real-time Extraction of Objects with Arbitrary Backgrounds," NTT Technical Review, Vol. 15, No. 12, 2017. <https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201712fa7.html>
- [13] T. Tokunaga, Y. Tonomura, J. Shimamura, F. Hamamura, and H. Kakimoto, "Real-time Moving Object Detection Technology and Trial of Stone Location Delivery at a Curling Venue," NTT Technical Review, Vol. 15, No. 12, 2017. <https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201712fa8.html>

Trademark notes

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


Ryuji Kubozono

Vice President, NTT Service Evolution Laboratories.

He received a B.S. and M.S. in physics from Kagoshima University in 1987 and 1989. He joined NTT Human Interface Laboratories in 1989. He was with NTT WEST from 1999 to 2003 and from 2006 to 2008, and with NTT Smartconnect from 2012 to 2016.


Akihito Akutsu

Executive Research Engineer, Proactive Navigation Project, NTT Service Evolution Laboratories.

He received an M.E. in engineering from Chiba University in 1990 and a Ph.D. in natural science and technology from Kanazawa University in 2001. Since joining NTT in 1990, he has been engaged in research and development (R&D) of video indexing technology based on image/video processing, and man-machine interface architecture design. From 2003 to 2006, he was with NTT EAST, where he was involved in managing a joint venture between NTT EAST and Japanese broadcasters. In 2008, he was appointed Director of NTT Cyber Solutions Laboratories (now NTT Service Evolution Laboratories), where he worked on an R&D project focused on broadband and broadcast services. In October 2013, he was appointed Executive Producer of 4K/8K HEVC (High Efficiency Video Coding) at NTT Media Intelligence Laboratories. He received the Young Engineer Award and Best Paper Award from the Institute of Electronics, Information and Communication Engineers (IEICE) in 1993 and 2000, respectively. He is a member of IEICE.


Norihiko Matsuura

Executive Research Engineer, NTT Media Intelligence Laboratories.

He joined NTT in 1994 and engaged in R&D of a 3D shared space system for communication via avatars in a virtual environment. He is currently in charge of the Visual Media Project, where he is conducting R&D in areas such as video encoding/decoding and artificial intelligence technologies for visual media.


Kenichi Minami

Executive Research Engineer, Natural Communication Project, NTT Service Evolution Laboratories.

He received a B.E. in electronic engineering and an M.S. in biomedical engineering from Keio University, Kanagawa in 1991 and 1993. He received an MBA from Thunderbird, Global School of Management, Arizona, USA, in 2002. He has been engaged in R&D management in the development of "Kirari!" immersive telepresence technology since 2016. During 2012–2014, he was responsible for the development of mobile application services at NTT DOCOMO. His research interests include image and audio processing, user interfaces, and telepresence technologies. He is a member of IEICE.


Akira Ono

Senior Research Engineer, Supervisor, Natural Communication Project, NTT Service Evolution Laboratories.

He received an M.E. in computer engineering from Waseda University, Tokyo in 1992. He joined NTT in 1992 and was involved in R&D of video communication systems. From 1999 to 2010, he worked at NTT Communications, focusing on network engineering and creating consumer services. He moved to NTT Cyber Solution Laboratories (now, NTT Service Evolution Laboratories) in 2010. Since 2015, he has been studying the immersive telepresence technology called "Kirari!".