6 Image Transmission and Contents Technology

6-1 Production of 3D Standard Test Contents

KIMURA Kazuhiro, ARAKAWA Yoshiki, OI Ryutaro, and YAMAMOTO Kenji

NICT have produced 3D standard test contents, and developed 2D/3D conversion software. The contents can be used for performance evaluation of 3D equipment, such as 3D monitors and projectors, R&D of 3D image processing technology, and assistance of 3D movie production. 2D/3D conversion software is also useful for development and learning 2D/3D conversion technologies.

Keywords
3D image, Auto-stereo, Standard contents, Test chart, 2D/3D conversion

1 Introduction

History has repeated itself in regard to three dimensional (3D) images, with a number of rises and falls in the past, and since around 2008, 3D images have entered into a third boom[1]. Currently, 3D images are rapidly transitioning into a practical stage due to the numerous 3D movies being shown around the world, the sale of stereoscopic 3D televisions by all major electrical appliance manufacturers and other factors. This is a stark contrast to when 3D images were mainly shown at science museums.

During the previous two booms, special projectors were required to display 3D images and televisions with sufficient resolution were not capable of being introduced into households. However, the recent explosive increase in the popularity of 3D technology is considered to be a result of easier access to digitalized 3D projectors. In terms of televisions for domestic use, regular televisions have become capable of displaying 3D images through the addition of a few additional functions due to the popularity of high-resolution flat panel displays and digitalization.

Conversely, 3D contents creation technology and stagecraft, etc., is still in a development and immature stage. Although binocular test charts are currently being distributed by the Institute of Image Information and Television Engineers[2] in regard to comparative performance evaluations of all types of 3D image display devices (including glasses-free systems) and standard test contents to become to standard for the development of image processing technology such as 3D image compression and conversion, such technology cannot respond to all types of 3D image systems (including glasses-free systems) nor be freely utilized for the various research objectives and demonstrations, etc., conducted by researchers. In this recent boom, the number of 3D image creators is increasing, but there are many obstacles in the way of copyright in the reuse of images and it is extremely difficult for persons other than the relevant creators to use such images for research and development and exhibitions, etc. Although manufactures of 3D monitors, etc., create images to evaluate their products, companies independently create their own images since third parties are not permitted to the images of other companies.

NICT has performed the research and development known as “3D Image Support Technology” for the purpose of supporting research and development and personnel training for 3D images. This research and development created standard test contents that can be used for the future development of glasses-free system and the current mainstream binocular 3D image technology. In addition, NICT has developed 2D/3D conversion software for the purpose of broadening the range of content creators and supporting the research and development and human resource development in the area of 3D image content. This content and software has already been distributed to approximately one hundred users to be freely used for research and development, the evaluation of equipment, etc., standardization reviews and non-profit exhibitions.

2 Circumstances of standard content production

For some time, many people have called for the creation of 3D image standard content, and even within the Ultra-Realistic Communications Forum (URCF), it has remained an important issue for promoting 3D image technology. However, as the creation of this kind of content involves great cost, it was unable to be achieved by the cooperation of the URCF’s members alone. Even under these circumstances, the URCF established the Standard Test Content Consideration Working Group to find issues relating to the creation of necessary standard content for the purpose of making preparations to enable the creation of such content should the opportunity arise.

In April 2009, the Ministry of Internal Affairs and Communications got a supplementary budget to perform the public offering for outsourcing of “Research and Development of Glasses-Free 3 Dimensional Image Technology – 3 Dimensional Image Support Technology–”. Then, following the NICT’s successful submission of a research proposal based on the examination results of the URCF, the following content was created by securing the required capital.

3 The formulation of the standard test content basic specifications

As it is important to reflect the opinions of people such as researchers who will most likely use the content, the “Standard Test Contents Creation Planning Project” was established in the URCF by related interested persons and the results of the deliberations were used when the creation specifications were being determined.

The Project drafted a content creation request list which was used by NICT to determine the creation specifications. The following five types of content were specified in this request list.

(1) Multi-view 3D image content
(2) Ultra-high definition stereo 3D image content
(3) 3D image content created by a rangefinder camera
(4) 3D CG content
(5) 3D image content created by a scanner-type camera

The requested specifications relating to the above (1) are shown in Table 1, the requested specifications common to (2) ~ (4) are shown in Table 2, the individual requested specifications related to (2), (3) and (4) are each shown in Table 3, Table 4 and Table 5 respectively and the requested specifications relating to (5) are shown in Table 6.

The above specifications were reflected in the creation of the standard test content to the greatest extent possible. There was the basic idea that, live-action images and their simulated computer graphics (CG) images of indoor-outdoor objects, as opposed to test patterns required for setup and adjustment of 3D image display devices, should be created so that they could also be utilized for the research and development of image processing technology such as compression, coding etc. and for demonstration exhibitions of developed image display devices. The image data was recorded in a file for each still image frame (as opposed to a general video file) on the assumption that it be used for various research and development purposes. The data was uncompressed and frames that filmed a checker board (angled in frontal, vertically and horizontally tilted directions) and a Macbeth chart were attached for calibration purposes. In addition, data such as camera parameters at the time of filming were attached.

Since it became apparent that it would be difficult to create all the requested content due to budget restrictions, the real-time rendering CG content of 3D CG content and the multi-view 3D image content were not created.

4 Details of the created standard content

Although the content created this time basically focused on the display of glasses-free 3D images, the content was also created to be used with binocular 3D images. The created content

Table 1 *Requested specifications relating to multi-view 3D image content*

Type of Content	A) The creation of filming material with a camera interval of 1 m or more for the main purpose of generating interpolating images B) The creation of filming material with a camera interval of approximately 6.5 cm for the purpose of displays on glasses-free monitors
Camera Arrangement	- 10 cameras are equally positioned in a straight line for both A) and B) - Circular positioning would also be ideal - Ensure cameras have a positional relationship that enables correlative image processing after filming using some kind of method - Attach information relating to the position and accuracy of cameras
Adjustment for Field Angles and Directions	- Effort on matters such as light axis when filming for A) and B) to the maximum extent possible - Attach information relating to image revision (geometry strain, etc.) after filming for A) and B) - Attach information regarding field angle and camera direction settings
Camera Type	- Ensure the camera specifications are attached when filming - Ensure that all ten cameras are the same and there is a high level of image quality uniformity and reliability
Lens Type	- Ensure that the lens specifications are attached - Ensure that the lens parameters are described at the time of filming - The lens performance should be of a program broadcast filming level quality
Inter-camera Synchronization	- Ensure that the synchronizing accuracy is of a GenLock level for broadcasting
Resolution	- Ensure that the resolution is 1920 × 1080 and the field frequency is 60 (59.94) or more
Color Adjustment	- Ensure that the color for all ten cameras is matched before filming - Ensure that filming is performed using a color calibration pattern and information that enables color correction is attached
Compression	- Ensure that recordings are uncompressed - Ensure that the signal route is displayed in detail from filming until recording (ensure that there are no compressed components in the signal route)
Lighting Conditions	- When filming in a studio i) lighting is applied as evenly as possible or ii) lighting suits the relevant scene - When filming outdoors i) Avoid conditions where colors cannot be sufficiently shown such as rainy and cloudy conditions, unless when intentional for staging purposes
Sound	- Ensure sound is recorded using LR
Copyright/Image Rights	- Ensure that images are processed without problems as "Standard Content for Distribution"
Detailed Content Specifications	1) Scenes should be able to be used for research and evaluations by anyone 2) Ensure that 1 sequence is one minute or more 3) Ensure that objects are 2 for A) and 4 for B) types or more
Camera Calibration	1) Specific camera filming methods such past performance records should be attached if available 2) Likewise Method for image adjustment when filming and image revisions after filming should be proposed if available

Table 2 *Common requested specifications for ultra-high definition stereo 3D image content, 3D images by a rangefinder camera and 3D CG content*

1) Create plans on the assumption of the following two purposes i) Standard content for creators Distribution for production agencies, creators and students, etc. - Utilize standard content as "Guidelines" or an "Introduction Guide" in order to create 3DCG. - Show the most appropriate stereoscopic parameters when being viewed on representative display devices. - Use standard content as material for examining 3D image creation and processing methods and conducting research and development. ii) Standard content for subjective evaluation purpose Use standard content to enable evaluations for not only image quality evaluations and physiological evaluations but also high-level staging effects. Utilize content common to i) Prioritize the purpose of i) and examine response within in a possible scope even for ii) - Construct scenes with the same depth of expressivity as the practical use content. - Appropriately implement plans and issue directions in view of the purpose of use specified in this paper by staff with experience in creating practical use content.
2) The scenarios and stories are not absolutely necessary. However, content that differs from practical use content is not effective as guidelines for content creation. - Also consider variations such as spatial structure, composition, brightness, color, frequency distribution and movement.
3) Ensure that the same scenes are presented with differing stereoscopic parameters and the visual effects are comparatively examined. - In the case of live-action stereo, include captured images containing the same objects with differing stereoscopic parameters.
4) Ensure that several types of image effects are comparatively confirmed (in the case of live-action stereo images) and pre-rendering CG) - In addition to the methods common to 2D such as camera work, movement, scene changes, lighting and texture, ensure that the effects of expression methods suitable for 3D images are incorporated and comparatively examined with such matters (overlap expression utilizing depth and brightness expression utilizing visual rivalry, etc)
5) Ensure that the recognizable size of objects such as people and familiar items (mobile phones, buildings, etc) are used and the reproduction of the size of such items can be examined. - Add detailed filming data and creation data in order to ensure that filming conditions and creation conditions can be analyzed. - Add calibration data (geometry, sensitivity characteristics, color, etc) according to the each format. - Provision data should be non-calibrated and calibration data should be attached using separate methods.

Table 3 *Requested specifications regarding ultra-high definition stereo 3D image content*

1) Ensure that the impact of changes to the convergence distance (visibility, feeling of fatigue, visual effects, etc) through scene changes and movements can be compared and confirmed. Guidelines for appropriate creation conditions should be provided for the above
2) The effect of changes to the convergence distance (when images of the entire scene protrude or retreat) can be compared and confirmed. Guidelines for appropriate creation conditions should be provided for the above
3) Ensure that unnecessary and unnatural effects (cardboard effects and puppet theater effects, etc) can be confirmed and their relationship with camera parameters can be confirmed. Guidelines for appropriate creation conditions should be provided for the above. Assume use for the purpose of examining the possibility that the impact level of these effects will change according to the resolution (the ability to repress by improved resolution) on a test basis.
4) Assume use with a 56 inch LCD (mainly for creation purposes) to a 200 inch projector (mainly for public exhibition purposes)
5) Other - Image data is 4K uncompressed data. - Create indoor and outdoor scenes with the addition of the following conditions. i) Indoor - Ensure that the reproducibility of object substances (metal, cloth, transparency, etc) can be examined (relationship with resolution and relationship with 3D information, etc). - Ensure that the changes of lighting effects caused by the stereoscopic view and viewing environment can be examined. Ensure that different lighting conditions can be compared and examined. - Assume the examination of the possibility that 3D technologies can represent the sense of depth in spite of weak contrast, etc. ii) Outdoor - Ensure that the expression effect of the sense of perspective can be examined, including scenes that distribute objects in a deep direction. - Ensure objects have a sufficient sense of detail which can confirm the impact of the resolution. - Ensure that changes of image effects caused by filming angles can be compared. - Assume images to be used for examining the relationship with screen size, viewing distance, filming angles, and image effects.

Table 4 *Requested specifications regarding 3D image content by the rangefinder camera*

1) Simultaneously conduct filming by positioning two HD cameras on both sides of the Axi-vision camera.
2) Image data shall be HD uncompressed and distant images.
3) Limit the object distance to a scope of approximately 2 m.
4) In regard to occlusion, use multiple types of objects with different characteristics within the same image (shape, etc) (Assume the usage of generating disparity images using distant images)
5) Use objects which move indoors.
6) Create three or more scenes with variations.

Table 5 *Requested specifications regarding 3D CG content*

1) Pre-rendering CG - Create high definition CG model with abundant expression and movie data with increased expressivity after post-processing such as synchronizing. - The CG model shall be created in a commercial modeling software data format capable of creating high quality CG such as 3DsMAX and Maya. - The high definition CG model shall be capable of being displayed on various 3D displays such as binocular and multi-view 3D by format conversion software. - A CG model with background and CG model without background shall be created. - The CG model shall provide scene data (including effects such as camera, animation and lighting) and be capable of mechanically creating video data by rendering the provided data. The camera setting shall be binocular stereo at the time data is provided. - The movie data to be provided shall be provided in a binocular stereo format. In regard to other 3D viewing formats, assume the user will perform separate rendering from the CG model. - The movie data shall be HD uncompressed on the assumption that the display has approximately 30 inch to 200 inch viewing environment. - Create a light-weight CG model. However, create such model in consideration of increasing the expressivity such as lighting effects and shading as much as possible.
2) Real-time CG - Provide 3D images while changing the stereoscopic viewing parameters in real time. - Create a simulation function that enables creators to easily comprehend the impact of the stereoscopic viewing parameters. - Enable stereoscopic viewing parameters (convergence distance, camera distance, field angle) to be changed in real-time and the relationship between the position of objects in CG scenes and stereoscopic viewing parameters to be schematically shown to users in an easy to understand manner. - Normally compatible with representative binocular stereo system (polarized 3D LCDs and projectors) - The CG data created by users can be used with modeling software such as Maya and 3DsMAX. - Data exchange and conversion software in standard formats should be provided.

Table 6 *Requested specifications regarding 3D image content using a scanner-type camera*

1) Camera - Type: 3 chip type color camera - Field Angle: 30 degrees or higher - Resolution: VGA or higher - Movement Range: horizontal 300 mm, vertical 200 mm (NICT facility specifications) - Camera Interval: 2 mm
2) Content (limited to still objects) - Objects that are difficult to estimate depth (objects with many hidden aspects, transparent objects, specular objects): plants, gem stones, glass objects and metal objects, etc. - Diorama (items with which scenery filmed outdoors can be assumed)
3) Necessary equipment and software, etc., when filming - Lighting equipment (equipment that is stable for long periods of time) - Internal parameter estimation and lens distortion correction software - Checker board for internal parameter estimation - HD (for saving data) - DVD-ROM (for data distribution) * Use the XZ stage and control PC of NICT
4) Distribution data - Format: bmp - Adjunct Data: Internal parameters - Filming conditions (camera specifications, viewpoint interval and distance to the object) - Images: Completion of lens distortion correction However, raw data and checker board information to be distributed if desired
5) Other - Data shall be distributed by sending several DVD-ROMs.

can be classified into the following four types.

- Ultra-high definition stereo 3D image content
- 3D image content produced by a rangefinder camera
- 3D image content produced by a scanner-type camera
- 3D CG content

4.1 Ultra-high definition stereo 3D image content

This content is binocular content that is filmed using two 4K cameras. This is ultra-high definition content that also assumes the performance of processes such as multi-view type interpolation processing. By down-converting this content to high-definition image quality, this content can also be displayed on binocular displays. The data is 4K resolution and uncompressed 30p contained in a left and right image sequential number TIFF file for each frame. Filming conditions (camera parameters) are included with all cuts and filming was conducted indoors and outdoors.

4.1.1 Purpose of creation

In recent years, 3D image technology has been steadily progressing towards commercial use. In particular, the quality of 3D images has been improved for commercial purposes. In order to produce 3D image technology for commercial purposes, it is essential that all display, filming, generation, transmission, recording and playback technology is evenly established at a high level. Furthermore, research and development of both devices and 3D image content creation technology is a vital remaining issue for the commercial use of 3D image technology.

In light of the above, standard ultra-high definition stereo 3D image content was created to contribute to the development of ultra-high definition 3D image technology and the technology to create its content.

4.1.2 Content overview

The creation of ultra-high definition stereo 3D image content focused on stereo 3D image content consisting of two 4K ultra-high definition images (horizontal 3,840×vertical 2,160 pixels) which were live-action video images filmed using 4k ultra-high definition stereo 3D cameras (two 4K cameras).

This content is used as standard content for the examination of 3D image creation methods and their processing formats and the research and development of 3D image technology. Consequently, the creation of this content was based on knowledge previously acquired, especially by the creation of 3D image content (3D high-definition), and the aim of clarifying the deference between HD images and 4K images.

4.1.3 Content details

The matters common to both indoor and outdoor content creation are detailed as follows.

(1) This content was created on the assumption that it is to be distributed to content production agencies, content creators and researchers and students researching 3D images and related fields.
(2) This content will be used as content material for the examination of 3D image creation methods, the examination of processing formats and research and development of 3D image technology.
(3) Scenes were created with the same amount of expressivity at the content used for practical and commercial purposes.
(4) Content was created with scenarios and stories.
(5) Variations such as spatial components, structures, brightness, color, frequency distribution and movements were considered.
(6) The visual effects of the same scenes with different stereoscopic parameters were compared and examined.
(7) Images of the same objects filmed with different stereoscopic parameters were included.
(8) Several types of image effects were compared and confirmed.
(9) Methods such as camera work, movement, scene changes, lighting and texture common to 2D were incorporated.
(10) Expression method effects sufficient for 3D images were incorporated.
(11) Recognizable size of objects such as peo-

ple and familiar items (mobile phones, buildings, etc) were used and the reproduction of the size of such items was examined.

An overview of the creation of indoor image content is detailed as follows.

(1) The impact of changes to the convergence distance (visibility, feeling of fatigue, visual effects, etc) through scene changes and movements was compared and confirmed.
(2) Content that provides guidelines for appropriate creation conditions regarding (1) was created.
(3) The effect of changes to the convergence distance (when images of the entire scene protrude or retreat) was compared and confirmed.
(4) Content that provides guidelines for appropriate creation conditions regarding (3) was created.
(5) Unnecessary and unnatural effects (cardboard effects and puppet theater effects, etc) were confirmed and content to confirm their relationship with camera parameters was created.
(6) Content that provides guidelines for appropriate creation conditions regarding (5) was created.
(7) Content was created on the assumption that it will be used for the purpose of examining the possibility that the impact level of these effects will change according to the resolution (the ability to repress by improved resolution) on a test basis.
(8) Content that enables the examination of the reproduction of objects (relationship with resolution and relationship with 3D information, etc) was created.
(9) Changes of lighting effects caused by the stereoscopic view and viewing environment were examined.
(10) Different lighting conditions were compared and examined.
(11) The examination of the possibility that 3D technologies can represent the sense of depth in spite of weak contrast, etc. was assumed.

An overview of the creation of outdoor image content is detailed as follows.

(1) The impact of changes to the convergence distance (visibility, feeling of fatigue, visual effects, etc) through scene changes and movements was compared and confirmed.
(2) Content that provides guidelines for appropriate creation conditions regarding (1) was created.
(3) The effect of changes to the convergence distance (when images of the entire scene protrude or retreat) was compared and confirmed.
(4) Content that provides guidelines for appropriate creation conditions regarding (3) was created.
(5) Unnecessary and unnatural effects (cardboard effects and puppet theater effects, etc) were confirmed and content to confirm their relationship with camera parameters was created.
(6) Content that provides guidelines for appropriate creation conditions regarding (5) was created.
(7) Content was created on the assumption that it will be used for the purpose of examining the possibility that the impact level of these effects will change according to the resolution (the ability to repress by improved resolution) on a test basis.
(8) Content that enables the expression effect of the sense of perspective to be examined (including scenes that distribute objects in a deep direction) was created.
(9) Objects with a sufficient sense of detail which can confirm the impact of the resolution level were filmed.
(10) The changes of image effects caused by filming angles were compared.
(11) Content was created on the assumption that it be used for examining the relationship with screen size, viewing distance, filming angles, and image effects.

4.1.4 Filming camera

Two RED ONE cameras were used on a half mirror pedestal (camera interval adjustment approximately 0 ~ 100mm) and when broadening effects such as puppet theater effects, a regular 3D pedestal (130mm ~) was used. The two RED ONE cameras attached to the half mirror pedestal are shown in Fig. 1 and the camera main specifications are specified in Table 7.

4.1.5 Indoor content scenario

A house studio with a modern image indoor set was used. This shows a bright scene where a young female washes her face, puts on makeup, happily chooses which dress to wear out and puts on a one person fashion show.

Ultra-high definition images were shown through skin texture, metals, the texture of cosmetics items such as sponges and various eye makeup. Furthermore, scenes settings that enabled 3D effects within natural movements such as water splashing when face washing and throwing rouge in a jovial manner at the camera when putting on makeup were established and typical 3D scenes settings that were full of convergence angle changes such as the inevitable back and forth motion of objects in the indoor one person fashion show were also established. The explanation points and content of this scenario is shown in Table 8.

4.1.6 Outdoor content scenario

Objects interweaved among trees were filmed. People walking and scattered objects appeared with a forest in the background. Easily comparable scenes were created with and without 3D standard objects in the foreground, middle ground and background. The explanation points and content of this scenario is shown in Table 9.

In particular, filming was performed in consideration of being able to sense the real differences between HD and 4K 3D effects. The anticipated 3D effects by 4K images are detailed below.

(1) Increased depth information of tree branches and leaves according to the level of resolution (effects of overlap recognition, etc)
(2) Increased sensation of emersion resulting from shortening the viewing distance through high definition (drawing closer to the display or screen)

Fig.1 *Ultra-high definition stereo 3D camera*

Table 7 *RED ONE camera main specifications*

Sensor	12 Megapixel Mysterium™
Lens Size	24.4mm × 13.7mm (Super 35 mm)
Effective Pixel Array	4520(h)×2540(v)
Full Pixel Array	4900(h)×2580(v)
Dynamic Range	> 66dB
Object Depth of Field	Equivalent to a 35 mm Cine Lens (S 16 mm window sensor)
Acquired System	4K 16:9
Delivery Formats	4K RGB
Project Frame Rates	23. 98, 24, 25, 29. 97, 30 fps 4K
Video Output	Single/Dual Link HD-SDI
REDCODE Codec	12 bit RAW 4K, 1-30 fps
Audio	4 Channel uncompressed, 16/24 bit, 48kHz

4.1.7 Example of filming and images

The actual filming is shown in Fig. 2 and an example of the image of a scene is shown in Fig. 3.

4.2 3D image content created by a rangefinder camera

4.2.1 Content overview

A content creation method that combined the image with depth data that could be obtained by a rangefinder camera at unprecedented high resolution with the regular 2D image filmed from the same viewpoint was developed, by which content creation was performed.

More specifically, four programs comprising of images/sounds with approximately one minute uncompressed color HD images and depth data were created using a rangefinder

Table 8 *Indoor content scenario*

Explanation Points	Content
(1) Change of convergence distance by scene changes and movements	- When the convergence distance is greatly changed by scene conversion from LS to UP (more significant than from UP to LS), a scene that makes it difficult to recognize 3D instantly occurs. - Measures: Reduce the convergence distance changes even with the same field angle
(2) Effects of changing convergence distance	- Compare the 3D effects when images protrude or retreat as a screen or screen surface standard. Most of the 3D range can be used and dynamic 3D effects can be achieved by appropriately using the foreground and background, so long as images do not protrude excessively. - However, when the convergence angle greatly changes, it is necessary to examine the editing connection method during editing.
(3) Unnatural Effects and causes	- Cardboard effect: Images where the depth of objects is unnaturally compressed. Images become to resemble thin paper cutout plays. - Cause: Caused by less disparity in left and right images such as narrow camera distance and telescopic filming. Suitable for miniature filming. - Puppet theater effect: Images that seem to have an unnatural object ratio. Also referred to as “miniature effect”, people appear small in comparison to the surroundings. - Cause and Measures: Caused by a large camera distance and images in the foreground and background being over emphasized. This is also an effective method for distant view and airborne filming.
(4) Comparisons with the containment (emphasis) of unnatural effects by 4K	- Examine whether the cardboard effect and puppet theater effect images are contained (emphasized) by 4K high definition images. Compare and examine two screens by 4K/HD.
(5) Examination of lighting conditions and 3D effects	- Examine the 3D effect by changing the lighting conditions. - What is the effect of protruding images in bright backgrounds? What is sense of depth? - What is the 3D effect when the background is bright and the foreground is dark? - Compare factors such as the impact of spotlights.
(6) Comparison of 2D creation methods and 3D	- Compare 2D and 3D by camera work (PAN, cranes, etc) and editing effects (dissolve, wipe, etc). - Compare cuts of 2D views and 3D views using the same materials.

Table 9 *Outdoor content scenario*

Explanation Points	Content
(1) Change of convergence distance by scene changes and movements	- When performing 3D filming outdoors, the depth may be more than expected and 3D images sometimes become more difficult to view if there are no obstructions such as walls. - When the convergence distance greatly changes due to the movement of objects, it is necessary to adjust 3D images within the scope of the fusion of the background and objects.
(2) Effects of changing convergence distance	- Compare the 3D effects when images protrude or retreat as a screen or screen surface standard. Most of the 3D range can be used and dynamic 3D effects can be achieved by appropriately using the foreground and background, so long as images do not protrude excessively.
(3) Unnatural Effects and causes	- The cardboard effect is likely to occur outdoors when filming from a distance. When filming in 3D from a distance, it becomes necessary to adopt measures such as intentionally broadening the camera distance - However, appropriate camera distance adjustment is required since the puppet theater effect can also occur by broadening the camera distance. - Examine the parameters of such adjusted camera distances.
(4) Comparisons with the containment (emphasis) of unnatural effects by 4K	- Examine whether the cardboard effect and puppet theater effect images are contained (emphasized) by 4K high definition images in outdoor filming conditions. Compare and examine two screens by 4K/HD.
(5) 3D expression effect by images of objects being distributed in a deep direction	- When target objects are placed in a deep direction and people move, examine the change of the 3D effect caused by the change in the line of view of the viewers. - Also examine using 4K since this active guidance of the line of view affects the 3D effect when creating 3D images.
(6) Change of image effects due to different filming angles	- Although the 3D effect is emphasized for the closer objects by filming with a wide angle, the depth and balance breaks down. - Adjust the balance during actual filming in consideration of the final screen size and viewing distance.

camera with uncompressed HD resolution possessed by NICT, and they were directed at the same level as an actual program. The four programs related to a living room, an office, a dart bar and dancing were produced.

In order to avoid the usual problem of occlusion (shadows) when processing 3D content using depth maps, this technology newly performed filming in a horizontal direction with a 3 lens structure and only the middle camera was attached with depth data. Furthermore, the upper, lower, left, right and center views of a checkered design (chess board) positioned on a flat surface was filmed for each cut as an indicator to regulate the relative position between each camera when filming and one shoot of the whole chess board as an indicator that detects the distortion rate of each camera lens. Then, in addition to filming a Macbeth chart and gray-scale chart for each scene and enabling camera sensitivity characteristics to be requested after filming, the distance between each object of the object scenes and the center camera lens was measured at the time of filming. The electronic data of the distance and scene structure chart was attached.

The resulting content consisted of four 1,920×1,080/59.94 i system images (horizontal 3 lenses + depth data) and two 48 kHz/24 bit system sounds (LR), totaling approximately four minutes of data. The data volume was approximately 500 GB.

Fig.2 *Filiming of ultra-high definition stereo 3D content*

Fig.3 *Scenes for ultra-high definition 3D content*

4.2.2 Creation background

Currently, systematic methods of filming, transmission and displaying 3D image content do not exists and a number of display formats are currently being examined. Research and development of standard content creation technology for these types of 3D image content is being conducted and as a result, the creation of content that enables the evaluation of display devices is required in this field.

The image with depth data in itself not only constitutes the structural model of object space when conducting an observation from a certain viewpoint, but also entails the following adaptations. 1) In multi-view images, by using depth data, more accurate interpolation is possible when synthesizing an image located in the position actually without viewpoints by means of interpolation processing. 2) It can be used as a benchmark providing the anteroposterior relation of objects when outputting the image synthesized with other 3D contents. 3) It is used as the pinpoint extraction processing of an only object, for instance, on a particular surface with a certain depth (referred to as depth key processing or Z key, etc.) when displaying the original 3D data on a existing 2D display. 4) The method for converting into a higher grade 3D image format such as holography is known[3][4]. Consequently, it can be expected that images with depth data will be used for conversion to various types of display formats and 3D image research.

Looking at domestic and international trends, as is seen in MPEG's proposal[5], the importance of images with depth data has been rising. However, in article[5], the depth data is measured by a low-resolution sensor with horizontal 174×vertical 144 pixels, which is lower than the SD, and the depth filming is operated with the light axis position being different that of color images.

On the other hand, the rangefinder camera possessed by NICT can radiate in the object space the modulated photocurrent of the near infrared rays that is harmless to people and derive the distance of object from the obtained reflected light per pixel in the real time. This is the full-spec HD resolution camera that can obtain the accurate distance data (correct value) of object space. It is possible to obtain distributable data (16bit-TIFF format individual files and Wav files) by using an auxiliary input device fitting to camera output interface.

Since there has never been a precedent that the image with depth/the left and right image/sound/detailed imaging parameters is provided as standard data in the HD video sequence at a resolution of 207 million pixels, development for 3D image content creation technologies using rangefinder camera is of great significance.

4.2.3 Technical specifications of imaging devices

As a rangefinder camera, Axi-Vision made by NHK Engineering Services, Inc., the highest resolution camera that is currently available on a commercial basis was used. The overview of this camera is shown in Fig. 4.

The method for distance detection with the Axi-Vision system is shown in Fig. 5. When radiating modulated light whose output light intensity increases with time (increased modulated light) to an object from the light source mounted on the camera, the reflected light from the object returns to the camera side with delay corresponding to the distance between objects. When filming that reflected light by a high-speed shutter camera for a short time, for obtained images, luminance difference corresponding to the distance to the object occurs, by which the distance to each point of object can be evaluated

When radiating increased modulated light to the object from light source and filming it with a high-speed shutter camera for a short length film, the reflected lights from objects O_1 and O_2 generate to its round-trip time a difference corresponding to distance difference. For this reason, in image A exposed for a short time, the distance information between O_1 and O_2 emerges as the difference in image contrast. The luminance value of O_1 and O_2 in this image at this time is pegged as I_1, I_2 respectively. Since the contrast[brightness] of image A is affected by such factors as the reflection rate of object, the spatial unevenness of irradiating light quantity and attenuation effect by diffusion light, these are corrected by radiating the light whose light intensity decreases with time (decreased modulated light) and filming image B.

Both image A and image B equally contains coefficients indicating the attenuation rate of light quantity due to the reflection rate of each object. Thus, if inter-image luminance ratio is calculated, these coefficients are cancelled and the range image D that represents the distance of camera to object as image contrast is derived.

In the rangefinder camera, distance d is derived from formula (1). Here, r stands for the luminance proportion between image A for increased modulation and image B for decreased modulation, T for modulating frequency, v for light speed and t_s for shutter timing.

$$d = \frac{1}{2} v \left\{ t_s - \frac{T}{2} \left(\frac{r}{1+r} \right) \right\} \tag{1}$$

Fig.4 Rangefinder camera (Axi-Vision camera)

In the Axi-Vision system, as is shown Fig. 6, by radiating the intensity-modulated light with 4 different phases to the object for 1 video frame duration, the variation of the reflected light from the object is detected and the distance is calculated. In this case, distance value is drawn form the formula (2).

$$d = k * \arctan \left\{ \frac{h(V_1 - V_3) - (1-h)(V_2 - V_4)}{(1-h)(V_1 - V_3) + h(V_2 - V_4)} \right\} \tag{2}$$

Here, d stands for distance, $V_n (n = 1, 2, 3, 4)$ for nth image signal, k, h for invariable, and this h for the coefficient that varies every horizontal scanning line of image. In regard to the measurement range of range image, the effective range of distance detection is adjustable according to object by making the frequency and phase of intensity-modulated light and the constant h of formula (2) variable.

The distance detection device (Fig. 4) composed of the distance detection principle discussed above has the light source part that radiates intensity-modulated light to object and the camera lens that captures images. The distance detection part has the image intensifier as the

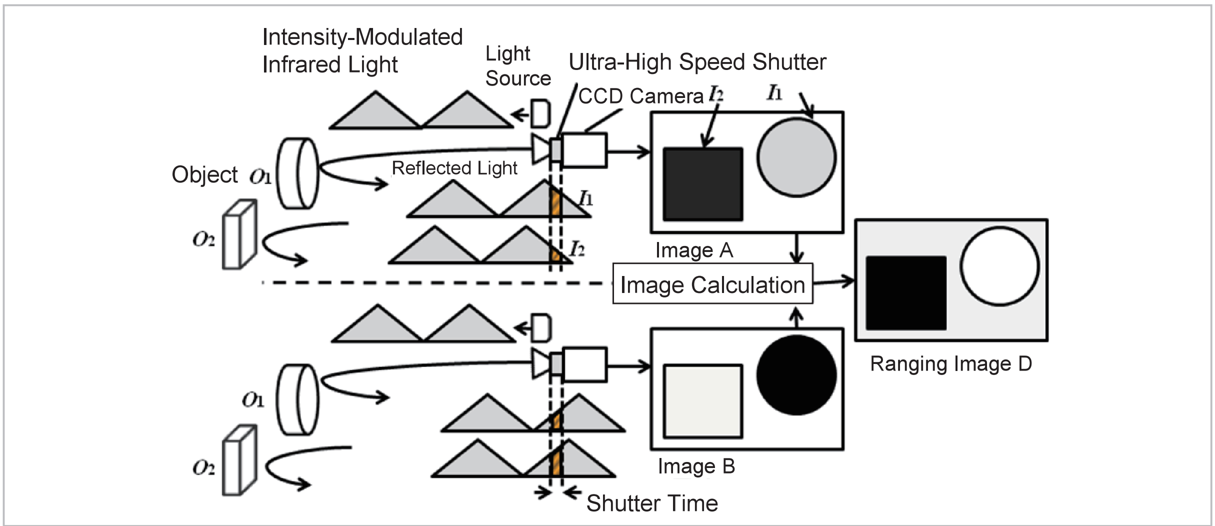


Fig.5 Basic principle for distance detection

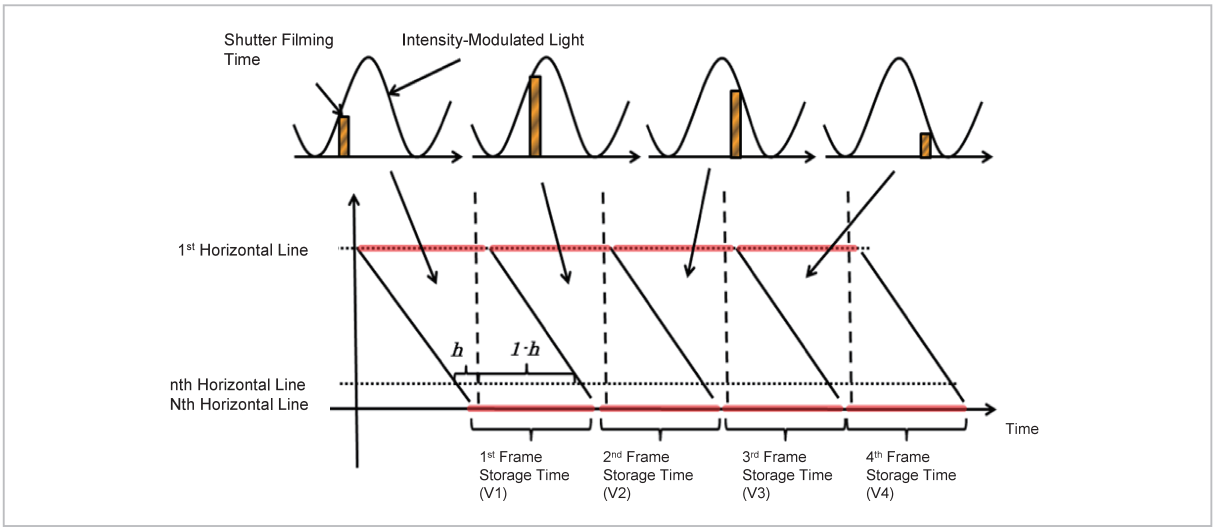


Fig.6 *Method for distance detection with Axi-Vision camera*

imagining part bearing a high-speed shutter, and consists of the CMOS sensor that stores and forwards the output per horizontal scanning line at a high speed. The intensity modulation of output light and the shutter motion of the image intensifier are activated synchronically by the signal from signal generator. In addition, based on the synchronization signal, the phase switching is controlled per image frame.

In this content creation, in order to avoid the occlusion issue that has always become controversial in the processing of 3D content using depth map, filming was performed by 3 color cameras aligned horizontally, among which only the center camera was provided with depth data. Specifically, as shown in Fig. 7, on the both sides of the Axi-Vision camera, 2 high-definition cameras of the same performance were horizontally arrayed to conduct filming simultaneously. Color images were filmed in the 3 positions of left (L), center (C) and right (R), depth image in one position of center (D). Thus, synchronized images worth 4 K×2 K were filmed in total.

4.2.4 Calibration data

A checkered design (a black and white chess board consisting of a 10 cm square grid) positioned on a flat surface was filmed in advance as a calibration chart for the geometry before and after filming. There were two patterns on the chess board.

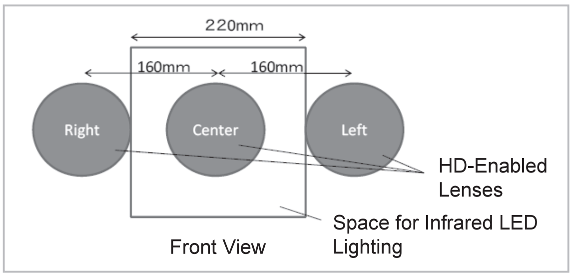


Fig.7 *Camera alignment*

(1) Full screen chess board

It became apparent during testing that the geometric distortion pattern changed according to the zoom value and focus value of the video lens (FUJINON HA13×4.5 BRD) used for actual filming. In order to enable the chess board to correct lens distortion on each of the lenses of the trinocular cameras, the chess board pattern that covered the entire field angle of each video camera in the object focus position was filmed for all cuts (4 – 7 cuts) consisting of four scenes. In addition, the entire cut of the image of the filmed full screen chess board was collected and attached at the end of each scene of the video.

(2) Center screen chess board

Although the left, center and right cameras were used for the actual filming after the field angle, optical axis and vertical positioning was physically adjusted, it became apparent during testing that the cameras were not completely aligned as a result of the camera weight and slight zoom lens movements. Consequently, the front, upper, lower, left and right patterns were filmed while changing the angle of the chess board in the center portion of screen in order to use such images as an indicator for determining the three camera positions after filming. The images of center screen chess board were also filmed for all cuts (4 – 7 cuts for each scene) consisting of four scenes. In addition, the entire cut of the image of the filmed center screen chess board was collected and attached at the end of each scene of the video.

Even though HD cameras/lenses for broadcast use were used, there are variations of sensitivity characteristics and response characteristics regarding color. As a result, a Macbeth chart and grayscale chart was filmed for each of the four scenes in order to enable the camera sensitivity characteristics to be estimated later. Furthermore, the iris value and object brightness (brightness measured by regular filming) for each camera when filming was saved as electronic data.

In order to determine how many meters the distant images captured with the Axi-Vision camera responded in the physical space, the distance between representative objects and the center Axi-Vision camera in each cut of the four scenes was measured and the values saved.

(3) Macbeth chart and Greyscale

A color checker chart (Macbeth) and a greyscale chart (for sensitivity characteristics use) were filmed one time or more for each scene when the post-filming cuts were linked and the four living, office, dart bar and dance scenes were created. Furthermore, the filmed Macbeth and greyscale chart images were attached to the end of each scene of the video.

(4) Distance data from the front lens of the center camera to each object

The distance from the front lens of the center camera (C: color HD Image, D: Axi-Vision depth data image) to the main objects contained in the scenes was measured by laser measuring equipment for all cuts in the four scenes. The measurement data was entered with the drawing of cut in electronic data form separately to the video. More specifically, the data was entered into a MS Word file and saved in a folder containing each of the living, office, dart bar and dance scenes.

(5) Camera iris values, focus values, object brightness and Axi-Vision setting values

The left camera (L), center camera (C), right camera (R) lens iris values, zoom values, focus values filming surface height, optical axis field, object brightness in the object central surroundings and the color temperature were measured for all cuts in the four scenes. In addition, the repetitive frequency, I.I. gain values, I.I. gate time, ch5 phase angle data which constitute the setting parameters of Axi-Vision were also measured for each cut in the same manner. The measurement data was entered with the data specified in (4) in electronic data form separately to the video. More specifically, the data was entered into a MS Word file and saved in a folder containing each of the living, office, dart bar and dance scenes.

4.2.5 Examples of created images

Examples of the created four program images are shown in Fig. 8 through Fig. 11.

4.3 3D image content produced by the scanner-type camera

4.3.1 Content overview

A creation method for filming image series comprising of equivalently highly dense horizontal and vertical camera array and possessing both horizontal and vertical narrow pitch and an extremely large number of viewpoints was developed using a scanner device controlled by computer, and content was created using this method.

More specifically, using a 3 band color camera (uncompressed IEEE1394 digital camera) with an even space resolution in each RGB wavelength and props and objects actually used for broadcasted programs, two scenes were created with a 2 mm horizontal and 2 mm vertical lens pitch, approximately 1.2 million pixels for each 1 viewpoint and still image content each produced by 15,000 points (total 30,000 points). The two created scenes were about crystal and diorama.

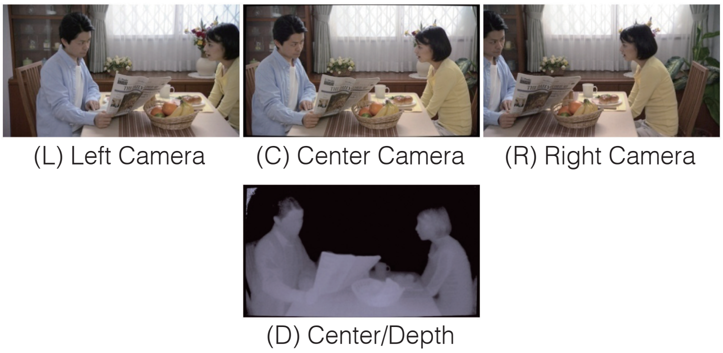


Fig.8 *3D image content produced by the rangefinder camera (Living room)*

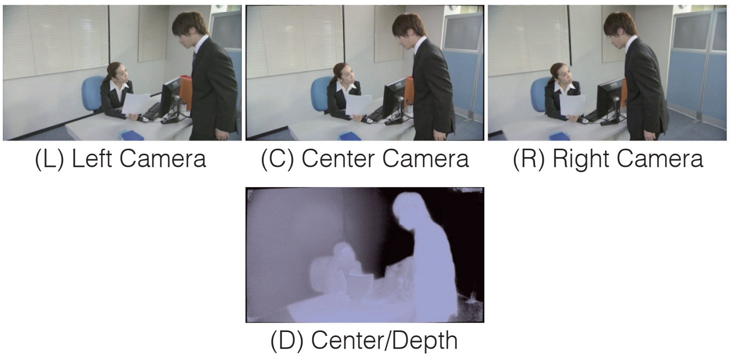


Fig.9 *3D image content produced by the rangefinder camera (Office)*

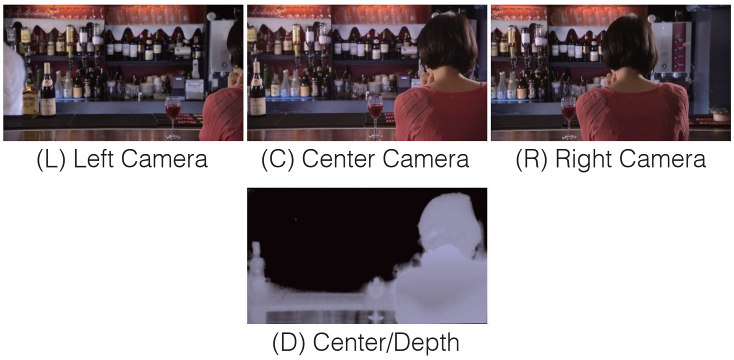


Fig.10 *3D image content produced by the rangefinder camera (Dart bar)*

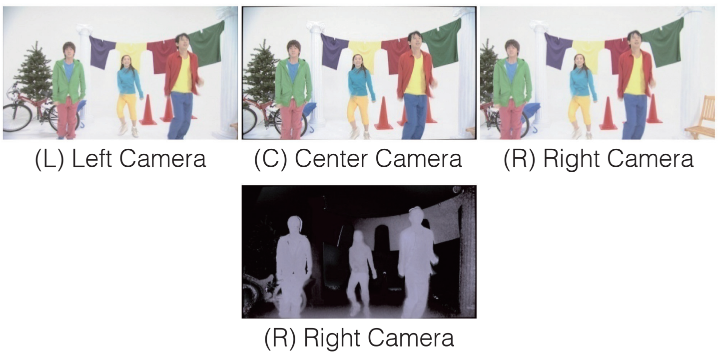


Fig.11 *3D image content produced by the rangefinder camera (Dance)*

The upper, lower, left, right and center sections of a checkered design (chess board) positioned on a flat surface were filmed for each scene as an indicator for specifying the camera parameters of the center position camera when filming with this technology. Then, a Macbeth chart and grayscale chart for each scene was filmed and the camera sensitivity characteristics were enabled to be estimated after filming. When filming, the advantages of directly controlling the cameras by using the IIDC protocol in IEEE1394 were especially utilized, and each parameter such as shutter speed, gain, gamma characteristics, pedestals, white balance and frame rates were set to the conditions regarded to be generally suitable as standard data.

Filming was operated with the shutter speed fixed at 1/15 sec., amplifier gain in 0dB, gamma correction OFF (linear response characteristics), the pedestal at default value, auto-adjust white balance (AWB/One Push) by filming a solid white picture per scene, the fixed white balance in the scene, and the frame rate at 7.5 frames/sec. Furthermore, the iris was focused to the highest extent possible, filming was performed under pan focus conditions and the impact of miss focusing was avoided.

The resulting content was a two scene approximately 30,000 viewpoint (150 horizontal viewpoint × 100 vertical viewpoint) image data consisting of a 15,000 viewpoint 1,280 × 960 pixel/RGB24 bit map (BMP) file and calibration data (5 viewpoint chess board + 1 viewpoint Macbeth chart). The data volume was approximately 120 GB.

4.3.2 Creation background

Currently, systematic methods of filming, transmission and displaying 3D image content do not exists and a number of display formats are currently being examined. Research and development of standard content creation technology for these types of 3D image content is being conducted and as a result, the creation of content that enables the evaluation of display devices is required in this field.

Since the 1990's, multi-viewpoint image data has been raised in the much researched topics of 3D image technology relating to the Ray-Space theory and Light Field Rendering. (1) Parallel cameras aligned in a high-density horizontal and vertical position are often used as a method of imputing real images into the light field or positioning virtual cameras in CG. As shown in Fig. 12, since the camera alignment responds to the s-t plane of the light field and the camera internal pixel directly responds to the u-v plane by the use of this data set, post-possessing becomes comparatively easy. Furthermore, it has become apparent that the use of multi-viewpoint image data enables the achievement of comparatively good image rendering results that retain the occlusion relation-

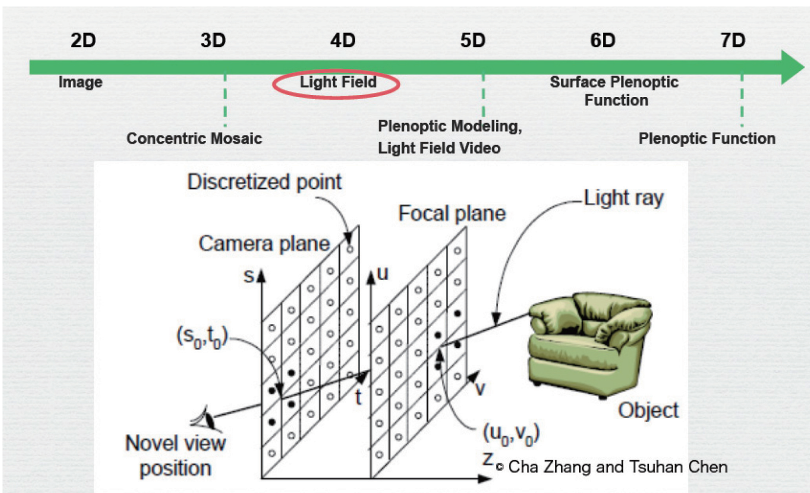


Fig.12 Light field (source: Carnegie Mellon University)

ship in real 3D images. Moreover, in the hologram field, (2) an extremely large number of 2D image sequences are used in the method known as holographic stereogram[6] for some time. (3) Although the establishment of the focus point is slightly different, multi-viewpoint image data is also used in the recent research of digitalized holograms[7]. While obvious, (4) the structure of object scenes can also be estimated through reverse projection from multi-viewpoint images and numerous applications currently exist. Consequently, in addition to the conversion to various types of display formats, it is expected that multi-viewpoint images will be widely used in 3D image research.

Even looking at the movements in both Japan and overseas and as seen in the No. 91 MPEG[8], the importance of multi-viewpoint image data is increasing. In Japan, since 1994, standard images have existed in the "Tsukuba University Multi-Viewpoint Image Data Base" created by a group lead by Professor Ohta Yuichi of Tsukuba University. These images consist of 9×9 viewpoint VGA image (640×480 pixels) filmed with an interval of several centimeters and such images have been often used by research groups both in Japan and overseas.

In order to acquire higher-density and multi-viewpoint standard image data, NICT has introduced new filming equipment. For this creation, an object positioned 705 mm ~ 1000 mm from the front lens of the camera was filmed with 150×100 viewpoint in a 2mm vertical and horizontal pitch and a higher density due to the recent advancements in computer equipment.

As the multi-view image data often goes through the image processing and rendering processing after its use rather than given chance only to be used as it is, the use of a 3-chip color camera whose sampling points for each R/G/B completely coincided was considered appropriate to be used when filming it, which assumed an aspect of raising the threshold of creation. Although there are some examples[9] of ultra-high density filming of the ray space or the light field, there very few examples of such filming methods and filming data sets being provided as standard content and the development of 3D image content creation technology by a scanner-type camera has become very significant due to the research being conducted on such filming methods and data sets.

4.3.3 Filming device and object specifications

An overview of the scanner-type camera (multi-viewpoint camera) specifications required when determining the size and position of objects are shown in Table 10 and the object distance derived from the sampling conditions are specified as follows.

- Object - Camera Distance: 705 ~ 1,000 mm

The filming field of view for the front surface of the object derived from the above conditions, including the field angle and object distance, is specified as follows.

- Width: approximately 668 mm
- Height: approximately 477 mm

Consequently, the object was aligned sufficiently fine and dense within a cubic shape of 30 cm for width, depth and height. The position of the object from the scanner-type camera as seen from above is shown in Fig. 13.

As multi-viewpoint filming was conducted with a camera with the maximum resolution equivalent to SXGA, the object was created with a sufficient accuracy (texture fineness) and attention was paid to thoroughly immerse the object with the possible scope of the scan-

Table 10 Scanner-type camera overview

Filming System	3 Chip Color Camera
Field Angel	Horizontal 30° or more
Format	1/2 Optical System ~ 1/3 Optical System
Resolution	VGA or higher
Resolution when filming	SXGA ~ XGA
Camera Movement Range (Horizontal)	300 mm
Camera Movement Range (Vertical)	200 mm
Camera Movement Amount (Depth)	0 mm
Camera Movement Amount (Rotation)	0 degrees
Camera Movement Interval	2 mm

ner-type camera. Since the standard content of a 3D image, two types of objects with different characteristics in the same screen were created in relation to shadowing (occlusion). One type of object was a “crystal” object that is difficult to estimate the depth, including numerous hidden matter, transparent matter and strong specular matter and other type of object was a “diorama” that assumes outdoor filming of a scene containing objects such as streets, buildings and trees. Each object is shown in Fig. 14 and Fig. 15 respectively.

4.3.4 Examples of created images

Examples of created images are shown in Fig. 16 and Fig. 17. These Figures are an extract of only the images filmed from nine representative points and they are lined up 3×3. Of the nine images shown in each Figure, the upper left images are the images filmed from the most upper left position and the upper right images are the images filmed from the most upper right position. The center images are images

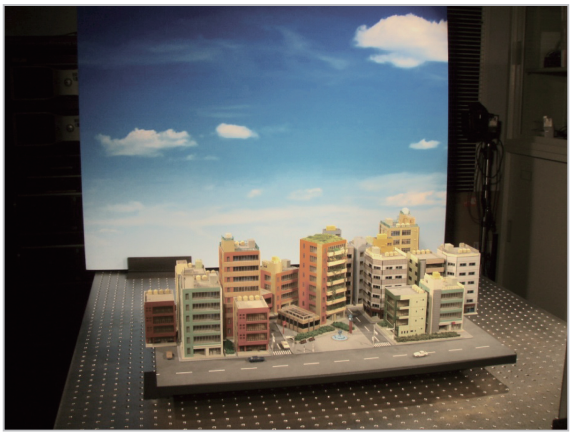


Fig.15 Object (Diorama)

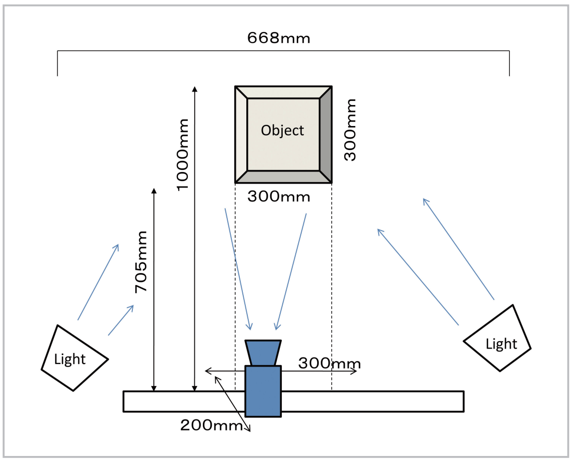


Fig.13 Object spatial position to the scanner-type camera (2D diagram)

Fig.16 Example of scanner-type camera content image (Crystal)

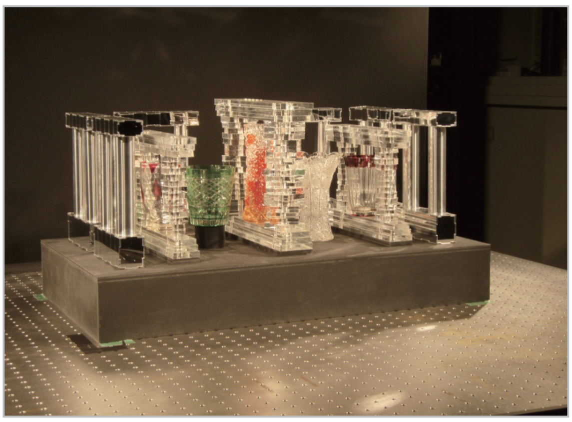


Fig.14 Object (Crystal)

Fig.17 Example of scanner-type camera content image (Diorama)

filmed from the center of the scan.

4.4 3D CG content

4.4.1 3D CG content overview

3D CG content consists of a high-definition CG model with rich expressivity and movie data with increased expressivity produced by post-processing such as synthesis.

The high definition CG model is not merely primitive simple modeling data but a model consisting of six scenes based on appropriately complex object and scene designs. The scenes were created in consideration of variations of factors such as spatial structure, structural outline, brightness, frequency distribution and movement. This model constitutes scene data completed only with 1 pass off-line rendering. The format consists of a AutodeskMaya 2010 native format and rendering was made possible through the use of Mentalray for MAYA. Although the initial setting is the intersection method using stereo (2 lens) cameras, it is capable of being used with the multiple 3D display format by setting three or more cameras. Furthermore, the model can respond to various viewing conditions by changing the parameters.

Each scene is 15 to 30 seconds in 30p and image expressions with obstructions in the binocular format have also been included. In addition, all types of video effects such as camera work, movement, scene changes, lighting and texture can be compared and confirmed for the 3D images.

The movie data is a rendered 122 second high-definition video in 30p contained in a TIFF file (left and right eye sets). The content is a future image of ultra-realistic communication technology containing variations such as eye distance, convergence angles, field angles, camera motion, object movement and color.

4.4.2 High-definition CG model scene overview

(1) Mystical Ocean Floor Temple

This scene shows a gentle light band filtering through the ocean surface, rocks covered with moss, swaying seaweed and floating marine snow. The remains of the decayed temple stand quietly, showing a mystical scene. Bubbles occasionally can be seen rising and sharks form a group while swimming around leisurely. Suddenly, one shark breaks from the pack, closes into the forefront, then turns around to disappear into the deep ocean after swimming through the temple. Representative images of this 15 second scene are shown in Fig. 18.

(2) Micro World – Bees and Flower Field

This scene shows a flower field covered with a profusion of various types of beautiful flowers that unfolds before the viewer's eyes. Flower pedals sway in the wind and then form a group and float in the air. If the viewer constantly looks at the image they will see bees flying. The camera dynamically crosses the field to one flower from the perspective of a bee. When moving through the micro world,

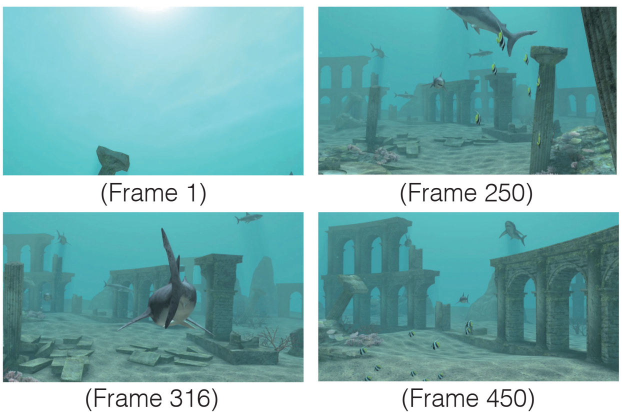


Fig.18 *Example of a CG scene "Mystical Ocean Floor Temple"*

the viewer can see one bee aiming for nectar at the center of a flower. If the viewer carefully looks at the ultra-close up of the center of the flower, they will see the nectar shine under the sunlight. Representative images of this 15 second scene are shown in Fig. 19.

(3) Space Station Floating in a Small Planet

This scene shows a space station built on a small planet in the enormous universe. Construction work is being performed by workers wearing spacesuits and a giant robot. A small-scale rocket stands-by the inside gate of the space station waiting to travel to other small planet. The space station can be seen from an open glass door inside the gate. The small-scale rocket shoots past the robot after take-off. Representative images of this 30 second scene are shown in Fig. 20.

(4) Strange World – Noble Metal Dissolution

This scene shows a dream-like strange world reminiscent of the works of the artist “Salvador Dali”. Beautiful precious metals such as sparkling gorgeous gem stones and an impressive shiny pocket watch are positioned in many places around the scene. Gradually, these items defy gravity and float in space. Then they melt and change shape. Viewers can see these beautiful precious metals in this strange world under various light effects such as regular lighting, candle light and a naked flame. Representative images of this 15 second scene are shown in Fig. 21.

(5) Virtual Light Museum

This scene shows an imaginary museum referred to as “Virtual Light Museum”. It displays “light art” that displays various color

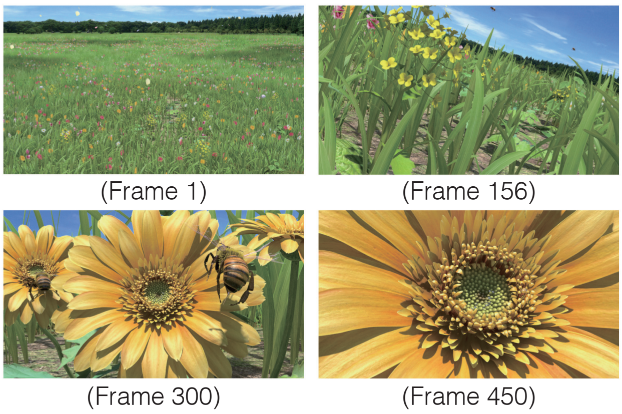


Fig.19 *Example of a CG scene image “Micro World – Bees and Flower Field”*

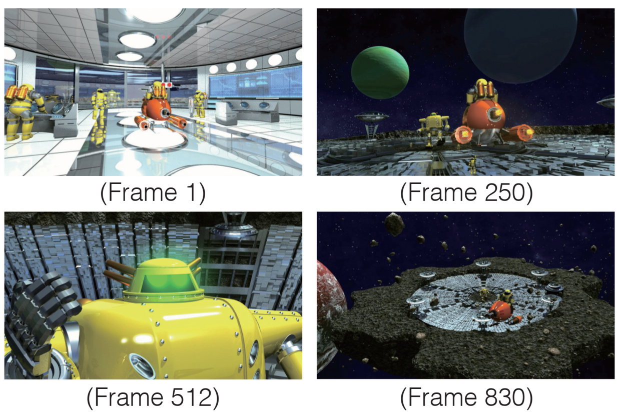


Fig.20 *Example of a CG scene image “Space Station Floating in a Small Planet”*

tones rich with a "sense of volume". This "light art" consists of an accumulation of countless shining lights and thin lines which constantly change slowly. The camera moves smoothly among the "light art" works on a crane and shows various expressions from overhead and close-up viewpoints of the constantly changing "light art". Representative images of this 15 second scene are shown in Fig. 22.

(6) Park Scenery

This scene shows a sunny regular park scene where people are walking, children are playing with balls and couples are talking on a bench. The camera moves at a calm regular walking pace. Representative images of this 15 second scene are shown in Fig. 23.

4.4.3 3D CG movie scene overview

Geometric 2 dimensional diagrams with rich colors commence floating movements in 3D. Following this, the diagrams slowly scatters in a random manner and begins to float in space. Gradually the scene focuses on the diagram in the huge area of space with a sensation of depth. The geometric diagrams formed in this large space then becomes a huge stage and this is where various research theme images regarding ultra-realistic communications is developed. Each research theme image scene such as ultra-realistic broadcasts, communications, remote medical care, sensory communications and technology transmissions is displayed. Representative images of this 122 second movie are shown in Fig. 24.

5 The development of 2D/3D conversion software

Along with the creation and provision of the standard test content discussed in the previous chapter, the free experience of 3D image creation by creators of future 3D images and human resource development in the field of research and development of 3D image support technology, 2D/3D conversion software has been developed in the aim of supporting 2D/3D conversion technology creators.

2D/3D conversion refers to technology that estimates the depth of objects taken from images filmed with one camera using experimental rules and then creates depth images and left and right eye images.

5.1 Developed 2D/3D conversion software overview

Recently various types of 2D/3D conversion technology have been developed and commercialized. This software differs from the full-fledged professional use software since it enables beginners to freely create 3D images. The incorporated algorithms have been kept comparatively simple in order not to infringe any patents, estimate depth by applying a filter to inputted images in order and generate depth images and left and right eye images. Although the accuracy of automatic conversion is not sufficient, this software can be used as a development support tool as it contains an environ-

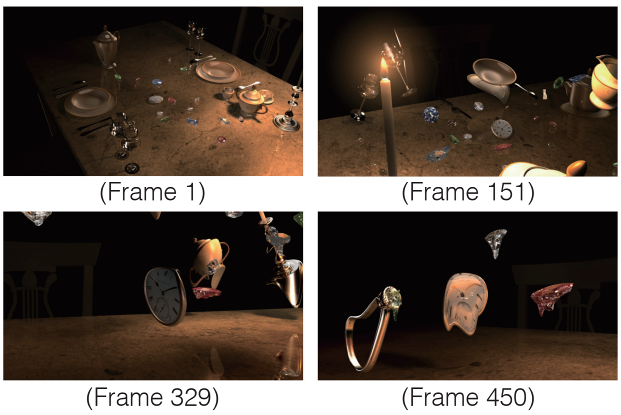


Fig.21 *Example of a CG scene image "Strange World – Noble Metal Dissolution –"*

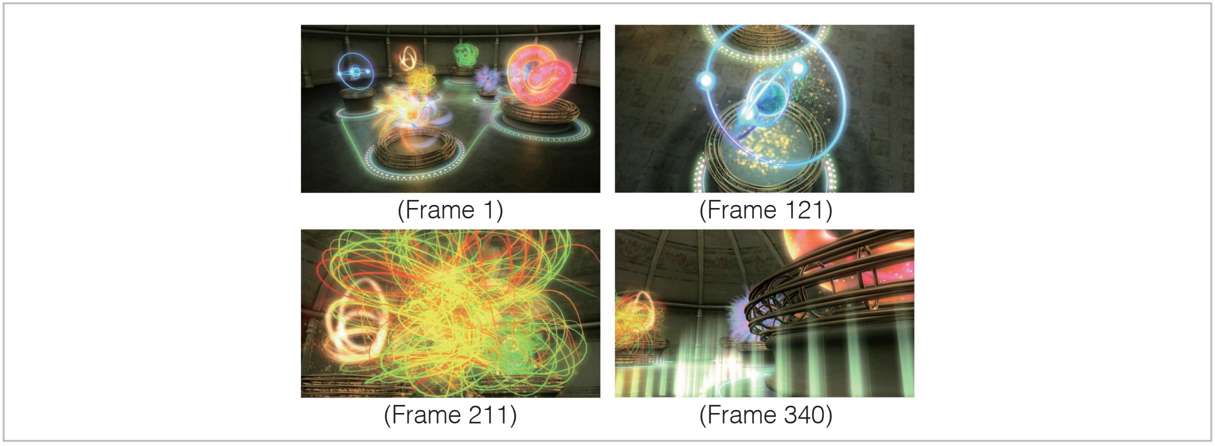


Fig.22 *Example of a CG scene image "Virtual Light Museum"*

Fig.23 *Example of a CG scene image "Park Scenery"*

Fig.24 *Example of CG movie image "Ultra-Realistic World"*

ment that enables trial and error by incorporating the manual correction function and conversion filter developed by the user. In addition, this software can also be used as an image processing educational tool that enables users to try the effects of various image processing filters.

This software operates with PCs installed with a Windows environment. When a one viewpoint image is inputted, a depth image or left and right eye image is outputted with the relevant viewpoint. Filter combinations can be easily changed by the method of estimating depth while applying various filters. The OpenCv function that may be able to be used for estimation purposes has also been implemented as a filter. Users can also create and easily control their own filters using the C# language. Some estimations can be performed automatically and then modified in detail manually.

5.2 Example of 2D/3D conversion

An example of conversion using this software is shown in Fig. 25. Although errors can be seen in the depth estimations of the sky and clouds as a result of the automatic conversion shown in Fig. 25 (A), mostly correct conversion results can be achieved by performing simple manual corrections in approximately ten minutes as shown in Fig. 25 (B).

6 Content and software distribution

Developed standard content and 2D/3D conversion software are being distributed to researchers, universities and companies with addresses and places of business in Japan free of charge (applicants are responsible for hard disks containing recorded content and return delivery expenses). Applications for delivery can be made through the following website.

http://3d-contents.nict.go.jp/

Delivery for content can be requested by entering into the content delivery server from the website, applying for an ID and completing registration.

Content and software may be used freely after accepting the following purposes of use.

(1) Use as an evaluation management video for research and development of 3D related equipment
(2) Use for industry standard examination purposes or as an quality training video
(3) Use as an evaluation management video in creation, editing and screening processes
(4) Use as an exhibition video for industrial samples and exhibitions
(5) Use for other research and development purposes (including research and development such as image processing, conversion technology, compression transmission technology) and non-profit purposes (including

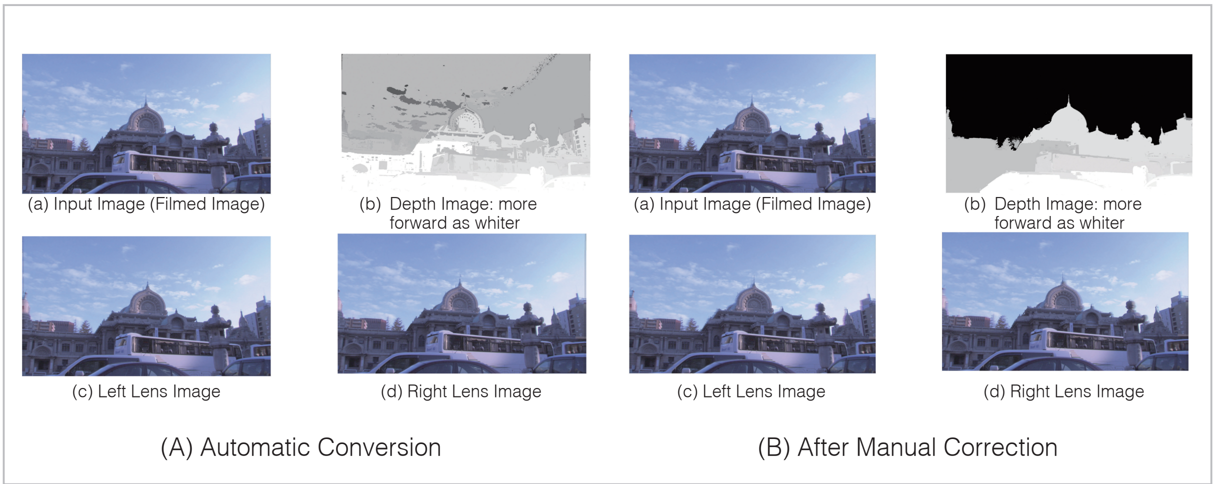


Fig.25 *Conversion example by 2D/3D conversion software*

education, training and information campaigns)

As a general rule, the following acts are not permitted unless approved by a separate license.

- Acts relating to the redistribution or duplication by persons other than the license holder (not including redistribution for use in the same company)
- Acts relating to screening, exhibition, broadcasting or transmission for profitable purposes
- Acts relating to the re-editing of the video (alterations or fraudulent use, etc)

Image processing for research and development purposes such as down conversion, frame rate modifications, compression, coding, format conversion and interpolation processing for the purpose of conforming the video to the user's display device is not considered to be an alteration. However, in the event images are modified in such a manner and screened or announced in essays, etc., such modifications must be noted.

When test images are to be provided to international standardization organizations (MPEG, ISO, IEC, ITU, etc), NICT does not directly distribute such image but supplies the content to the persons in charge of such standardization and grants a license to distribute such images to the relevant persons of working groups, etc. (including persons overseas).

Content is distributed in an uncompressed sequential number still image file on the assumption that such content will be used for the purpose of wide-ranging research and development. However, this content will be provided in a converted format such as the BlueRay format which can be used without conversion processing on the user side when numerous participants are used to perform impact evaluation experiments on persons in 3D images or performance evaluations, etc., of commercial 3D television.

7 Conclusions

We have developed and commenced the free distribution of standard test content and 2D/3D conversion software that enables the use by anyone for the purpose of performing comparisons and evaluations of 3 dimensional image equipment and research and development of image processing technology.

A number of issues came to light as a result of performing the creation of the standard test content. Ultra-high definition stereo 3D image content was created at a 30 p frame rate due to delays in obtaining cameras. Since we received strong requests for standard test content to be created at a 60 p frame rate, content needs to be recreated using the acquired 60p cameras.

This time, as some content was not created, the content should be enhanced in the future.

Acknowledgements

This report is a result of "Research and Development of Glasses-Free 3 Dimensional Image Technology – 3 Dimensional Image Support Technology" research outsourced by the Ministry of Internal Affairs and Communications.

We would like to thank everyone in the Standard Test Content Creation Planning Project of the Ultra-Realistic Communications Forum (URCF) for all their advice regarding the detailed specification settings for the standard test content.

References

1 "3D Display, The 3rd honestry.," Nikkei Electronics, No. 2008.9.22, Sep. 2008. (in Japanese)

2 "3D Image Standard Chart," The Institute of Image Information and Television Engineering (ITE) Web site, http://www.ite.or.jp/shuppan/testchart_index.html (in Japanese)

3 R. Oi and M. Okui, "A Method for Displaying a Real-scene Fresnel Hologram," ITE Technical Report, Vol.

30, No. 43, pp. 3–6, 2006. (in Japanese)

4 R. Oi, K. Yamamoto, and M. Okui," Electronic generation of holograms by using depth maps of real scenes," Proceedings of SPIE, Vol. 6912, No. 21, pp. 1–11, 2008.

5 GIST (Gwangju Institute of Science and Technology), "3D Video Test Sequence and Camera Parameters," ISO/IEC JTC1/SC29/WG11, MPEG2009 / M16949, 2009.

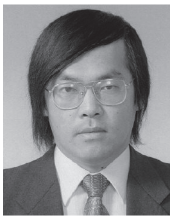
6 J.T. McCrickerd and Nicholas George, "Holographic stereogram from sequential component photographs," Applied Physics Letters, Vol. 12, No. 10–2, 1968.

7 R. Oi, T. Mishina, M. Okui, Y. Nojiri, and F. Okano, "A Fast Hologram Calculation Method for Real Objects," The Journal of the Institute of Image Information and Television Engineers, Vol. 61, No. 2, pp. 198–203, 2007. (in Japanese)

8 Fraunhofer HHI, "HHI Test Material for 3D Video," ISO/IECJTC1/SC29/WG1, MPEG2008/M15413, 2008.

9 Marc Levoy and Pat Hanrahan, "Light Field Rendering," Proceedings of ACM SIGGRAPH, pp. 31–42, 1996.

(Accepted Sept. 9, 2010)



KIMURA Kazuhiro

Director, Project Promotion Office, Universal Media Research Center

3D Image, Satellite Communication

ARAKAWA Yoshiki, *Dr. Eng.*

Senior Researcher, Project Promotion Office, Universal Media Research Center

Super High Definition Image, 3D Image, Image Transmission

OI Ryutaro, *Dr. Sci.*

Senior Researcher, 3D Spatial Image and Sound Group, Universal Media Research Center

Optical Wave Propagation Analysis, Holography, 3D Imaging Technology, Image Sensor

YAMAMOTO Kenji, *Dr. Eng.*

Senior Researcher, 3D Spatial Image and Sound Group, Universal Media Research Center

Electronic Holography, 3D Image Technology