Vol.2012-AVM-76 No.5
2012/2/23

# 自由視点画像生成ためのデータ表現

パナヒプル テヘラニ メヒルダド[†]
藤井 俊彰[†]、谷本 正幸[†]

近年、自由視点画像生成を目的とするシステムがいくつか提案されているが、そのシステムが用いるデータ表現はそれぞれ異なっている。これらのシステムの入力データは、主にカメラアレイによって取得される。その取得したデータを表現する際、システムの要求条件に基づいてテクスチャおよびジオメトリ情報を使用する・しないなど、さまざまな異なる形式で表現されている。本稿では、それらのシステムとそのデータの表現方式を検討する。まず自由視点画像生成をするさまざまなシステムを紹介し、それらのデータ表現を分類・整理して各方式の得失を論じる。

# Data Representation for Arbitrary View Synthesis

MEHRDAD PANAHPOUR TEHRANI[†]
TOSHIAKI FUJII[†], MASAYUKI TANIMOTO[†]

Several systems targeting arbitrary view synthesis have been proposed, where each has different data representation. The input data for these systems are mainly captured by camera array. However, the captured data are represented by texture and geometry information in different format based on the system requirement. This paper aims to review those systems and their data representations scheme. Firstly we introduce different systems representations and systems for arbitrary view synthesis, and finally discuss and categorize them based on their data representations.

## 1. Introduction

Current TV provides us only a single view of a 3D world and users cannot control their viewpoint, which is different from what we experience in our life. 3DTV with multiview displays (stereoscopic and autostereoscopic) has more views, and gives us depth feeling, while the viewing zone of these 3DTV is very limited. Multiview system sets many cameras in a large space and displays captured views by switching them, which can realize motion parallax with a large viewing zone, while the number of viewpoints is fixed. In these systems, users cannot set their viewpoints freely. If we have the model of 3D scene, we can generate free-viewpoint images easily in a virtual world. However, free-viewpoint image generation in a real world is a challenging task.

## 2. Components of 3D Systems

In general, it is possible to separate the 3D representation of free viewpoint into three main parts: acquisition, transmission and display parts as demonstrated in Figure 1. In this paper [1], we will focus on data representation issue within the three components.


**Figure 1**: Basic block diagram of 3D Systems

**Acquisition:** In order to capture dynamic 3D contents, the data acquisition system has to be extended from a single camera system to stereo camera or multi-camera systems. The multi-camera system is overweight the stereo one as it can acquire larger amount of 3D scene information. The number of cameras varies depend on the system size. This component can consist of capturing, camera calibration, color/geometry correction, and data representation.
**Compression:** The transmission of 3D content is generally similar to the transmission of 2D content in the video-based system. It means that the 3D data can be transmitted through a common communication channel, such as the internet or wireless channel. However, the 3D information of a scene is usually a lot larger than the 2D information of the same scene. Thus a compression is compulsory. However, in a scenario with limited processing and communication abilities of nodes, a distributed source coding scheme is suitable.

[†]名古屋大学
Nagoya University

**Display:** There are many various ways to display the captured 3D information based on the display devices. Those display devices can be separated into two groups. The first group is the display devices that use in 2D visual system such as normal 2D display or television. The other group includes a special display devices designed for 3D visual system such as binocular glasses or 3D television. In order to provide freedom on viewpoint selection, rendering or free viewpoint synthesis has to be done within this component.

## 3. 3D Systems and Representation Methods

There are five main representation methods that have been widely used to represent 3D information: (1) Image Domain, (2) Integral Photography, (3) Image-based rendering –IBR, (4) Model-Based Rendering – MBR, and (5) Combination of IBR and MBR representations.

**(1) Image Domain:** In image domain representation, original multi-view images are directly obtained from a multiple camera system.

**Acquisition:** The camera setup can be the parallel setup, convergent setup, or divergent setup. To generate a new view image by e.g. image warping technique, the corresponding points between original images have to be specified. Automatically finding correspondences requires computation costs, such as feature extraction, matching, and correlation calculation. In addition to that, if we want to recover the depth value of the corresponding points, camera intrinsic and extrinsic parameters have to be measured through camera calibration.

**Compression:** Original multi-view images can be compressed as moving pictures. Each view image can be compressed by using e.g. JPEG and multi-view image set can be compressed by using multiview image coding techniques. More sophisticated approach is to use shape or depth of the scene. In [2], 21 view images are compressed to the amount of 1-2 view images.

**Display:** In image domain representation, no special rendering process is required. It may be warping of original multi-view images or rendering from 3D object to 2D image, and these processes can be done in real-time.

**(2) Integral Photography:** Similar to Image domain representation, original multi-view images are directly obtained from a multiple camera system.

**Acquisition:** In this system, an array of CCD camera densely captures the view. The camera setup is precisely aligned. In this method, the captured views are optically converted.

**Compression:** The IP representation is similar to image domain representation, so that original multi-view images can be compressed as moving pictures.

**Display:** In IP representation, no special rendering process is required. However, the same view is back projected through multiple microlense arrays. Therefore, human eye can render. Integral Photography represents the 3D information in image domain; however the amount of captured information is large, which makes it similar to the ray-space representation concept.

**(3) Ray-Space, Light Field, and Lumigraph (IBR):**Ray-Space representation is obtained by converting the original multi-view images to "ray" parameters.

**Acquisition:** Ray-Space[3][4] is defined in the world coordinate; the conversion requires camera intrinsic and extrinsic parameters. Therefore, camera calibration has to be done. In real experimental systems, it is difficult to obtain dense Ray-Space data. To solve this problem, it is a good solution to generate intermediate ray data, i.e., to "interpolate" Ray-Space data from sparsely sampled Ray-Space data. Light field or Lumigraph are similar to Ray-Space.

**Compression:** For the Ray-Space compression, the following redundancy of the Ray-Space data can be utilized. One is the same redundancy as that the 2D still images have. A cross section of the Ray-Space corresponds to a 2D still image. Therefore, the Ray-Space has the same statistical properties as 2D still images. The other redundancy comes from the fact that the intensity of the surface point is similar regardless of the viewing direction. The extreme case is that the objects are covered with the Lambertian surfaces. In this case, the intensities of all the rays that come from the same point are constant regardless of the viewing direction. This redundancy is so high that it should be exploited for the efficient Ray-Space coding. In [5][6][7],ray-space coding such as subband coding, vector quantization are briefly described.

**Display:** Since a view image is a 2D subspace of 4D space, the renderer only resamples a 2D slice of line from the 4D Ray-Space or Light field. This process requires only a memory access procedure and no computation costs.

**(4) Model Based (MBR):** Model-based representation is obtained by converting the original multi-view images to a 3D-model, and texture information.

**Acquisition:** Model-based also needs camera calibration and registration as the Ray-Space and Light field. Model-based rendering usually recovers the geometry of the real scene and then renders it from desired virtual view points. Methods for the automatic construction of 3D models have found applications in many field, including Mobile Robotics, Virtual Reality and Entertainment. These methods generally fall into two categories, active [8][9] using laser and passive [10]using multiple 2D photographs of a scene methods.

**Compression:** To robustly reconstruct a 3D model of a scene, the Multi-Hypothesis Volumetric Reconstruction (MHVR) algorithm is used [11]. In contrast to methods that rely on feature-matching [12], no corresponding points needs to be found in MHVR. A voxel model is built directly from calibrated images that depict a static scene from multiple viewpoints, and no image segmentation is necessary. The algorithm is based on discretizing the space containing the object into volume elements (voxels).

**Display:** Model-based, the required viewpoint is constructed simply by having the texture map and 3D model information. The texture is mapped simply using the 3D information.

**(5) Surface Light Field Mapping (IBR & MBR):** Surface light field mapping is a representation method that combines the IBR and MBR approaches.

**Acquisition:** Surface light field also needs camera calibration and registration as the Ray-Space and Light field. The difference lies on that in a surface light field; the first pair of parameters of ray-space describes the surface location. This means that we need the shape of a 3D object. In the case of a synthetic scene, since the computer has a model of an object, surface light field can be easily obtained. However, in a real scene, it is a troublesome procedure to measure the 3-D object and construct triangular mesh. Another process of surface light field is data resampling. It includes normalization of texture sizes and resampling of viewing directions.An "interpolation" process of light rays, which is only 0-order interpolation is sufficient, since views are very dense.

**Compression:** In surface light field, their approach is to approximate light field data using the dimensionality reduction over small surface patches. Precisely, they approximate the discrete 4-dimensional surface light field function as a sum of a small number of products of lower-dimensional functions. The compression ratio in this stage depends on the truncation of the sum obtained from the decomposition. In [13], they report that 100:1 compression ratio is obtained by the approximation, 10:1 by vector quantization, 6:1 hardware texture.

**Display:** Surface light field requires the most complicated rendering process. First, the algorithm texture maps the surface primitive using the surface map. Then, it texture maps the surface primitive using the fragment of the view map determined by the view-dependent texture coordinates. Finally, it performs pixel-by-pixel multiplication of the results of the two texture mapping operations. The algorithm is tuned to use hardware engine.

**(6) Multiple local ray-space (IBR & MBR):**

**Acquisition:** Multiple local ray-space also needs camera calibration and registration as the Ray-Space. Local ray-space (LRS) method for walk-through view generation is proposed in [14][15]. In figure 2(a) a single ray space is applied for representation of the whole 3-D space that needs dense camera array. In this method the captured information is not sufficient especially for the walk-through view experience, and makes this task very expensive to be realized. This approach can be used for sparse camera density, and divides the space into several local regions, called a local ray-space. Ray information of each local region is recorded using cameras' parameters, and the estimated 3-D model of objects that also used to generate enhanced dense LRSs, using stereo matching algorithms [16].

**Compression:** Similar approaches for ray-space can be applied to each local ray-space.

**Display:** Occlusion area can be rendered and walk-through view can be generated by reading the corresponding local regions in the range of requested viewpoint.

**(7) Multi-view and multi-depth (IBR & MBR):**

**Acquisition:** It needs camera calibration. Multiview video plus depth was proposed as a data format for free viewpoint systems. To realize walk-through view among objects Multilayer Multiview Video plus Depth (MMVD) [17] can be used.. MMVD is generated in several layers among objects by using the local ray-space method..

**Compression:** In multiview video data, there is correlation among views and depth. Thus, efficient multiview coding requires algorithms that exploit temporal and interview/depth redundancy. Moreover, there is common characteristic between same viewpoint image and depth that worth considering. These redundancies can be removed by interview predication, and extracting the similarities between view and depth. This is based on the spatiotemporal analysis on multiview data set. An extension of MVC would be a solution.

**Display:** Using MMVD representation, realtime walk-through view can be synthesized with low complexity, using 3D warping and interpolation scheme [18]. MMVD approach follows the standardized data format, i.e. multiview video plus depth for free viewpoint video generation and transmission.

## 4. Overview of 3D Data Representations

Table 1 summarizes the comparative evaluation of each representation mentioned.

| Representation Method | Calibration& Registration | Data Conversion | Object Dependency | Rendering /Display Method | Free View Quality |
|---|---|---|---|---|---|
| Image Domain | No | No | No | Warping /Projection | Not-Precise |
| Integral Photography | No | Optically | No | No (Visually) | Precise |
| Ray-Space Light field/Lumigraph | Required | Coordinate Transform | No | Memory access/ Interpolation | Precise |
| Model-Based | Required | 3D model Texture | High | Texture mapping | Not-Precise |
| Surface Light field | Required | Decomposition Approximation | High | Texture mapping /pixel-by-pixel multiplication | Not-Precise |
| Multiple Local Ray-space | Required | Coordinate Transform | High | Memory access/ Interpolation | Not-Precise |
| Multiview and Multidepth | Required | No | Medium | Warping /Interpolation | Precise |

**Table 1:** Comparison of free viewpoint representation methods

## 5. 3D Systems for Arbitrary View Synthesis

### 5.1 Image Domain

**3D-TV:** The system was proposed by Mitsubishi Electric Research Laboratories (MERL) [19]. A prototype system was implemented with real-time acquisition, transmission and 3D display of dynamic scenes. A distributed, scalable architecture to manage the high computation and bandwidth demands is used. It consists of an array of cameras, clusters of network-connected PCs, and a multi-projector 3D display. Multiple video streams are individually encoded and sent over a broadband network to the display. The 3D display shows high resolution (1024x768) stereoscopic color images for multiple viewpoints without special glasses. The systems are implemented with rear-projection and front-projection lenticular screens.

**Eye Vision:** The system was proposed by the Carnegie Mellon University (CMU) [20]. It involves shooting multiple video images of a dynamic event, such as a football game, from multiple cameras placed at different angles (more than 30 cameras). The video streams from these cameras are combined by computer and the resulting images reach viewers in a format that will make them feel as if they are flying through the scenes they see.

### 5.2 Integral Photography

**Integral Photography (IP) System:** The system was proposed by NHK [21]. It is a well-known technique of autostereoscopic photography proposed by Lippmann [34]. No polarized glasses are needed in order to experience the three-dimensional effect. One distinctive feature is that the appearance of the displayed 3D image changes depending on where the viewer is located, just the way a real object appears to the viewer. Unlike other 3D display methods that can only present this 3D effect horizontally, IP offers 3D images vertically as well, realizing a so-called “full parallax”. IP can be considered in the same group as holography for its ability to reproduce spatial images. In these spatial image reproduction systems, the images constructed from light rays appear real to the eye of an observer.

**1D-II 3D Display System:** The system was proposed by Toshiba [22], and called “1D Integral Imaging (1D-II 3D Display System)”. It is based on the integral photography (IP) method proposed by M. G. Lippmann [23]. Two prototypes of 3D PC-based display system are developed. Horizontal resolution of 300 with 32 parallaxes in 20.8-inch type and 18 parallaxes in 15.4-inch type were developed due to adoption of mosaic pixel arrangement of the display panel. These prototypes have wide viewing area and a viewer can observe 3D images with continuous motion parallax. By using lenticular sheets, these prototypes can also display the same bright images as an ordinal display in spite of having many parallaxes. Plug-ins for creating 3D image for CG software are also developed.

### 5.3 IBR

**FTV:** The first system was developed in Nagoya University [24][25], and called "Free viewpoint TV (FTV)". In this system, one user can freely control the viewpoint position of any dynamic real-world scene. This system can cover a limited space. It uses image based rendering method based on ray-space concept. The system includes 16 clients node with a server node computer. Each node is a PC cluster powered by Intel Pentium III 800MHz as CPU with 256Mbyte RAM. Each PC is a general-purpose PC that has an image capturing board mounted in a PCI bus. A Gigabit Ethernet connects sensor node with a central node.

**Light Field Video Camera System:** The basic concept of this system is, similar to the basic concept of FTV. However, system designing approach and display rendering algorithm are carried in the very different idea. The system was proposed by Stanford University [26], and called "Light Field Video Camera System".In this system, the special cameras called the “Light Field Video Camera” (LFVC) are used to construct a camera array.

### 5.4 MBR

**3D Room:** The 3D room [27] system aims for capturing and modeling a real time-varying event inside the viewing zoon. The room is 20 feet (L) x 20 feet (W) x 9 feet (H). 49 cameras are currently distributed inside the room. A PC cluster system (17 PCs) is used to digitize all the video signals from the cameras simultaneously as uncompressed and lossless color images. The input image sequences provide depth image sequences by applying multiple baseline stereo frame by frame. The depth images of all cameras are merged into a sequence of 3D shape models, using a volumetric merging algorithm.

**3D Video:** The system was proposed by Kyoto University [28]. 3D video is a real 3D movie recording the object’s full 3D shape, motion, and precise surface texture. This system uses a parallel pipeline processing method for reconstructing a dynamic 3D object shape from multiview video images, by which a temporal series of full 3-D voxel representations of the object behavior can be obtained in real time. To realize the realtime processing, first a plane-based volume intersection algorithm is introduced, which represents an observable 3D space by a group of parallel plane slices, then back-projects observed multiview object silhouettes onto each slice, and finally applies two-dimensional silhouette intersection on each slice. Then, a method to parallelize this algorithm is performed using a PC cluster, where five-stage pipeline processing are employed in each PC as well as slice-by-slice parallel silhouette intersection. The method is improved by introducing IBR techniques.

### 5.5 Multi-texturing

**Multi-texturing:** The system was proposed by Heinrich Hertz Institute (HHI). Different types of model-based approaches have been proposed. Some of these approaches require some kind of coloring or texturing of the models. It is known from computer graphics that simple static textured models do not offer convincing realism when viewed from different positions, due to reflectance properties of the materials and lighting conditions of the scene. This is the so-called “painted shoe box effect”. Imagine a piece of velvet that changes appearance extremely depending on the viewpoint. An approach that requires few cameras and works well with video is based on multi-texture surfaces as presented in [29].

**3D Mixed-Reality Video:** The system was proposed by Tsukuba University [30]. This system realizes a 3D video display system that can capture video from multiple cameras, reconstruct 3D models and transmit 3D video data in real time. A target object with a simplified 3D model consisting of a single plane and a 2D texture extracted from multiple cameras is represented. This 3D model is simple enough to be transmitted via a network. A 3D video of a typical soccer scene that includes a dozen players was processed at 26 frames per second. The system consists of a capturing unit, a server PC and a rendering PC. The capturing unit consists of a video recording unit that includes multiple cameras and a scene analysis PC. The multiple video recording unit records multiple videos, which are completely synchronized with each other, to the memories or the hard disks of the capturing PCs.

### 5.6 IBR & MBR

**Bird’s-Eye View System:** The system was proposed by Nagoya University [31]. It assists the driver and generates an arbitrary viewpoint bird's-eye view and shows to a driver in real-time by using 16 CCD cameras mounted on a dome structure over a cross section. The 16 CCD cameras are set at each vertex of triangular patches, which compose a hemisphere dome with the radius of 50cm over the one twenty-fourth size of diorama. These cameras are arranged at about 20cm interval in the shape of an equilateral triangle. One twenty-fourth size of miniature car is used as an object. Triangular patch is considered for view synthesis.

**Surface Light Field System:** The system was proposed by Intel [32]. A light field parameterized on the surface offers a natural and intuitive description of the view-dependent appearance of scenes with complex reflectance properties. To enable the use of surface light fields in real-time rendering a compact representation suitable for an accelerated graphics pipeline is needed. The approximated the light field data is done by partitioning it over elementary surface primitives and factorizing each part into a small set of lower-dimensional functions. After acquisition, it can be further compressed using standard image compression techniques leading to extremely compact data sets that are up to four orders of magnitude smaller than the input data. Finally, an image-based rendering method and light field mapping are used that can visualize surface light fields directly from this compact representation at interactive frame rates on a personal computer.

**Walk-through view system:** The system was proposed by KDDI R&D LABS [13][14][16]. The first system was captured with 30 HD cameras, with circular arrangement (r=6m). 3D model of the abject is generated and used to generate multiple local ray-spaces. Given the required walk-through view, corresponding LRSs are read and the arbitrary view is synthesized. By converting LSRs to MMVD, walk-through can be realized in realtime.

## 6. Discussion

Various data representations can be discussed in the view-geometry domain [33] as shown in Figure 2, where horizontal/vertical axis denote the amount of texture/geometry information.

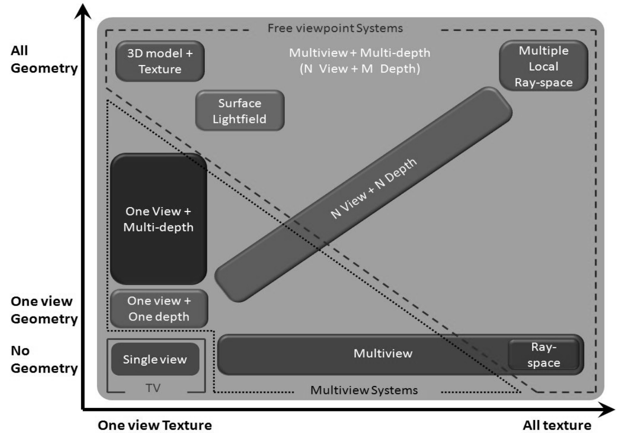


Figure 2: texture-geometry representation

Conventional TV system is realized by a single view image. Representation with more view than one view is multiview representation. If number of views is very large, multiview representation is equivalent to ray-space. Using one view plus one depth representation we can achieve a free-viewpoint image generation in a narrow range. By adding more geometrical information or having multi-depth, we can represent the free-viewpoint system as one view plus multi-depth, such as LDI (Layered-Depth Images) [34]. LDI contains more geometrical information, so that it can provide wider range of free-viewpoint. With all geometrical information or 3D model of the scene, we have 3D model plus texture representation that can provide a wide range of free-viewpoint. Surface lightfield contains less amount of geometrical information in comparison with 3D model while several views are required for this representation. This representation can provide wide range of free-viewpoint images. (N View + N Depth) is considered by MPEG. It covers a large area of the view-geometry representation. The region bounded with dashed line is feasible for implementation of free viewpoint systems. Dotted line is the area where the multiview systems are feasible. Note that free viewpoint systems can be realized using multiview representation when the views are sufficiently dense.

## 7. References

1 M. P. Tehrani, "3D Image Processing and Communication in Camera Sensor Network – Free Viewpoint Television Networking -". PhD Thesis, Nagoya University, Oct 2004
2 T. Fujii and H. Harashima, "Data compression and interpolation of multi-view image set" IEICE Transactions on Information and Systems, Vol. E77-D No.9, pp. 987-995, Sept. 1994.
3 T. Fujii, T. Kimoto, and M. Tanimoto, "Ray space coding for 3D visual communication", Proc. Picture Coding Symp.'96, vol.2, pp. 447-451, March 1996.
4 T. Fujii, T. Kimoto, and M. Tanimoto, "A New Flexible Acquisition System of Ray-Space Data for Arbitrary Objects", IEEE Tran. on Circuit and Systems, vol.10, pp. 218-224, March 2000.
5 T. Fujii, T. Kimoto, M. Tanimoto, "Ray-Space Coding Based on Arbitrary-Shaped DCT", SPIE Electronic Imaging '98, Jan. 1998.
6 M. Kawaura, T. Ishigami, T. Fujii, T. Kimoto, M. Tanimoto, "Efficient Vector Quantization of Epipolar Plane Images of Ray Space By Dividing into Oblique Blocks", Picture Coding Symposium 2001, pp.203-206, Apr. 2001.
7 M. Magnor and B.Girod, "Data Compression for Light-Field Rendering", IEEE Trans. on Circuits and Systems for Video Technology, vol. 10, no. 3, pp. 338-343, Apr.2000.
8 U.Castellani, S. Livatino, and R.B. Fisher, "Improving environment modelling by edge occlusion surface completion", 1st International Symposium on 3D Data Processing Visualization and Transmission (3DPVT)*,* Padova, Italy, June 2002.
9 J.Davis, S.M. Marschner, M.Garr, and M.Levoy, "Filling holes in complex surfaces using volumetric diffusion", 1st International Symposium on 3D Data Processing Visualization and Transmission (3DPVT), Padova, Italy, June 2002.
10 T.Kanade, P.Narayanan, and P.Rander, "Virtualized reality: Concepts and early results", IEEE Workshop on Representation of Visual Scenes, pp. 69-76, June 1995.
11 P. Eisert, E. Steinbach, and B. Girod, "Multi-hypothesis volumetric reconstruction of 3-D objects from multiple calibrated camera views", Proc. International Conference on coustics, Speech, and Signal Processing ICASSP'99 Phoenix, USA, pp. 3509-3512, March 1999.
12 M. Pollefeys, R. Koch, M. Vergauwen, and L. van Gool. Metric 3D surface reconstruction from uncalibrated image sequences. Lecture Notes in Computer Science, 1506, pp. 139-154, 1998.
13 W-C. Chen, R. Grzeszczuk, J-Y. Bouguet, "Light Field Mapping: Hardware-Accelerated Visualization of Surface Light Fields", SIGGRAPH 2001 Course Notes for Course #46.
14 A. Ishikawa, M. Panahpour Tehrani, S. Naito, S. Sakazawa, and A. Koike, "Free viewpoint video generation for walk-through experience using image-based rendering", Proc. ACM Multimedia 2008, Vancouver, Canada, Oct.-Nov. 2008.
15 M. P. Tehrani, A. Ishikawa, S. Sakazawa, A. Koike, "Enhanced Multiple Local Ray-spaces Method for Walk-through View Synthesis", Proc. of IEEE Computer Society, International Symposium Universal Communication, ISUC 2008, Osaka, Japan, Dec. 2008.
16 D. Scharstein and R. Szeliski, "Ataxnomy and Evaluation of dense two-frame stereo correspondence algorithm", International Journal of Computer Vision, 47(1/2/3):7-42, April-June 2002.
17 M. P. Tehrani, A. Ishikawa, S. Sakazawa, A. Koike, "Enhanced Multiple Local Ray-spaces Method for Walk-through View Synthesis", Proc. of PCSJ/IMPS2008, I-2.19, 59-60, Japan, Oct. 2008.
18 L. Yang, M. P. Tehrani,, T. Fujii and M. Tanimoto, "High-quality virtual view synthesis in 3DTV and FTV," 3D Research, vol. 2, no. 4, pp. 1-13, Dec. 2011.
19 W. Matusik, H. Pfister, "3D TV a Scalable System for Real-Time Acquistion, Transmission and Autostereoscopic Display of Dynamic Scenes", ACM SIGGRAPH, August 2004.
20 http://www.ri.cmu.edu/events/sb35/tksuperbowl.html.
21 M. Okui, F. Okano, "Real-time 3-D Imaging Method based on Integral Photography", Report (for AHG on 3DAV), ISO/IEC JTC 1/SC 29/WG 11 m10088, Brisbane, Australia, Oct. 2003.
22 T. Saishu, K. Taira, R. Fukushima and Y. Hirayama, "Distortion Control in a One-Dimensional Integral Imaging Autostereoscopic Display System with Parallel Optical Beam Groups", Proc. of SID, 53-3, Seattle, USA, May 2004.
23 G. Lippmann, Comptes-Rendus Academie des Sciences, 146, 446-451, 1908.
24 P. Na Bangchang, T. Fujii, M. Tanimoto, "Experimental System of Free viewpoint TeleVision", Proc. IST/ SPIE Symposium on Electronic Imaging, Santa Clara, CA, USA, vol. 5006-66, pp. 554-563, Jan 2003.
25 M. P. Tehrani, P. Na Bangchang, T. Fujii, M. Tanimoto, "The Optimization of Distributed Processing for Arbitrary View Generation in Camera Sensor Networks", IEICE Transaction on Fundamentals of Electronics, Communication and Computer Sciences, E87-A(8), 1863-1870, Aug. 2004.
26 B. Goldlucke, M. Magnor, and B. Wilburn, "Hardware accelerated Dynamic Light Field Rendering", VMV 2002, Nov. 2002.
27 H. Saito, S. Baba, M. Kimura, S Vedula, T. Kanada, "Appearance-Based Virtual View Generation of Temporally-Varying Events from Multi-Camera Images in the 3D Room", Proceed on 3-D Imaging and Modeling, Oct. 1999.
28 T. Matsuyama, X. Wu; T. Takai, T. Wada,*"* Real-time dynamic 3-D object shape reconstruction and high-fidelity texture mapping for 3-D video*", IEEE Trans. on* Circuits and Systems for Video Technology, Vol. 14, No. 3, March 2004.
29 K. Müller, A. Smolic, M. Droese, P. Voigt, and T. Wiegand, "Multi-Texture Modelling of 3D Traffic Scenes", Proc. ICME 2003, IEEE International Conference on Multimedia & Expo, Baltimore, MD, USA, July 6.-9. 2003.
30 T. Koyama, I. Kitahara, Y. Ohta, "Live Mixed-Reality 3D Video in Soccer Stadium", Proc. The 2nd IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR2003), pp.178-187, 2003
31 M.Sekitoh, T.Fujii, T.Kimoto, and M. Tanimoto, "Bird's Eye View System for ITS", IEEE, Intelligent Vehicle Symposium, pp. 119-123, 2001.
32 W. Chao Chen, J. Y. Bouguet, M. H. Chu, R. Grzeszczuk, "Light Field Mapping: Efficient Representation and Hardware Rendering of Surface Light Fields". ACM Transactions on Graphics. 21 (3), pp. 447-456, 2002.
33 M. Tanimoto, M. P. Tehrani, T. Fujii, T. Yendo, "Free-Viewpoint TV", IEEE Signal Processing Magazine, Vol. 28, No. 1, pp. 67-76, Jan. 2011.
34 J. Shade, S. J. Gortler, L. He, R. Szeliski, Layered Depth Images, Proc. SIGGRAPH 98, pp. 231-242, Jul. 1998.