FTV - Free-Viewpoint Television System

Toshiaki Fujii

Graduate School of Engineering, Nagoya University
fujii@nuee.nagoya-u.ac.jp

Free-viewpoint TeleVision (FTV) is a next generation television which users can watch while moving their viewpoints freely. With present TVs, users can obtain only a limited view of a real 3-D world, because the view is not determined by users themselves but by the position and direction of a camera. FTV relaxes this limitation and makes it possible for users to view the 3-D world freely as if they were in the scene.

The algorithm to generate free views in our FTV system is based on the Ray-Space theory. Ray-Space consists of 4-dimensional data that represent the intensity of light rays as a function of the ray parameters. The advantage of using Ray-Space for FTV is that it can generate very high-quality, photo-realistic free-view images without a complicated rendering process. This is because Ray-Space data is "real" data and the technology used in our FTV is a signal processing method, not a computer graphics method. On the other hand, Ray-Space has problems such as cost of acquisition, transmission, and storage, because Ray-Space is constructed from an enormous amount of view images. Therefore, Ray-Space has long been thought to be only an ideal concept and not suitable for real-time applications such as FTV.

We have, however, succeeded in creating an FTV system by developing a real-time view interpolation and an efficient data compression technology. The figure below shows an overview of our FTV system. First, many video cameras capture a dynamic 3-D scene. After geometrical and intensity calibrations are applied, view interpolation is carried out and the Ray-Space is constructed. Then, data compression and coding are applied to the Ray-Space data. On the display side, an arbitrary viewpoint image is generated by extracting 2-D subspace from the Ray-Space data.

In the presentation, I first review the Ray-Space method and then explain the details of our FTV system. I then describe the data acquisition systems; the multi-camera system and one-camera with optical scan system. As for interpolation of Ray-Space, I explain adaptive filtering and table-lookup methods. Ray-Space coding is investigated using a conventional 2-D image coding method, 3-D model aided method, arbitrary-shaped DCT, vector quantization, and hierarchical subband coding. An FTV display system using head tracking is also presented. Finally, I give a demonstration of our FTV system and also introduce our MPEG (Moving Picture Experts Group) activities.

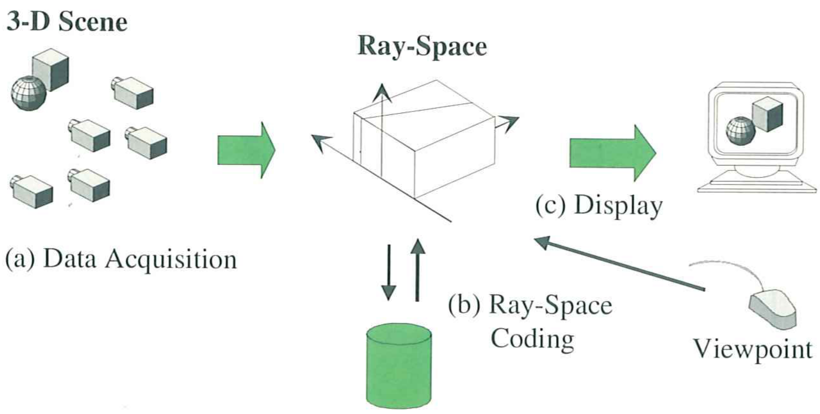


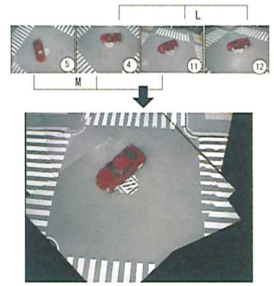
Figure: Free-viewpoint Television (FTV) system overview.

FTV –
Free-Viewpoint Television System

Toshiaki Fujii

Graduate School of Engineering
Nagoya University

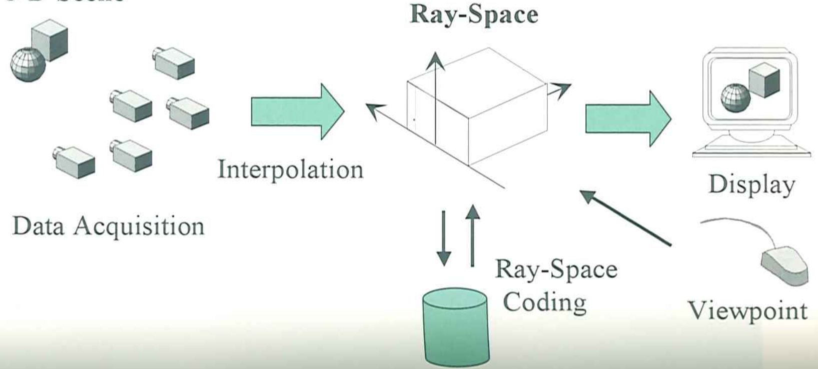
Free-viewpoint Television (FTV)



Input (multi-camera) → Processing (Interpolation) → Output (Free-view)

Free-viewpoint TV system based-on Ray-Space representation

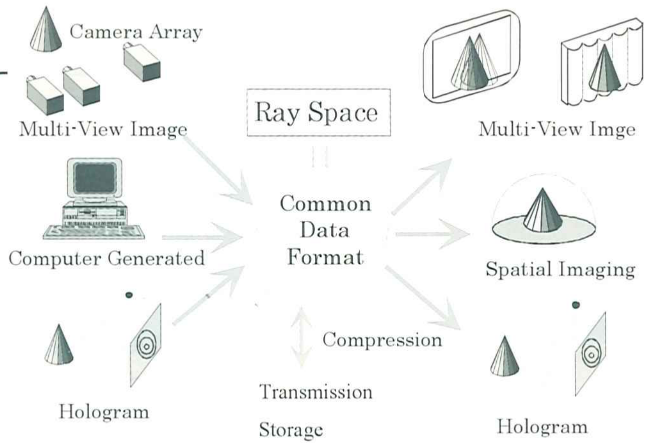
SPIE ITCom 2002, vol. 4864-22, pp. 175-189, Aug. 2002.



Outline

1. Introduction of FTV
2. Review of Ray-Space Method
3. Free-Viewpoint TV System
   (a) Data acquisition
   (b) Ray-Space coding
   (c) Display system
4. Demonstration of Free-Viewpoint TV
5. Conclusions

2. Review of Ray-Space Method

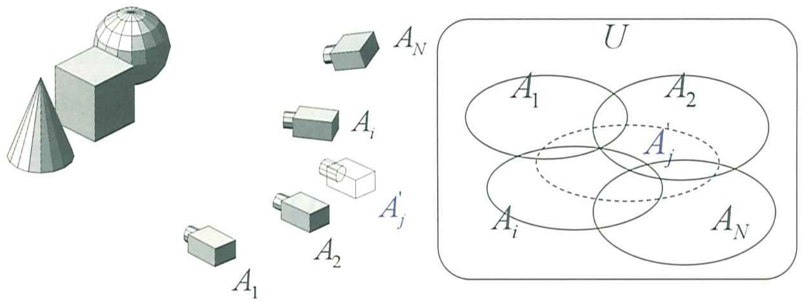


Concept of "Integrated 3-D Visual Communications" by Harashima, Fujii (1992).

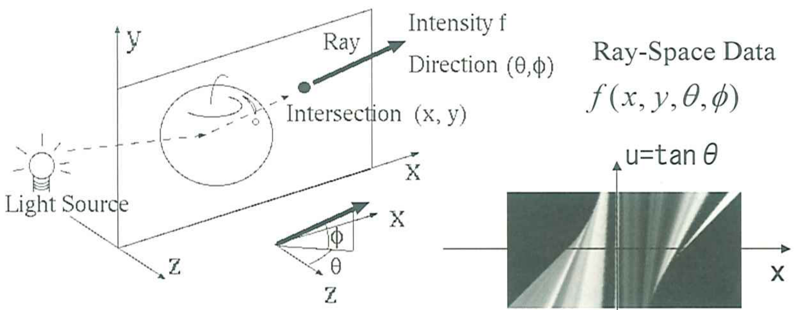
Concept of Ray-Space Representation

Multi-view image: $A_1, A_2, \cdots, A_N \subseteq U$ U : all ray data

Virtual view image: $A_j^{'} \subseteq U$



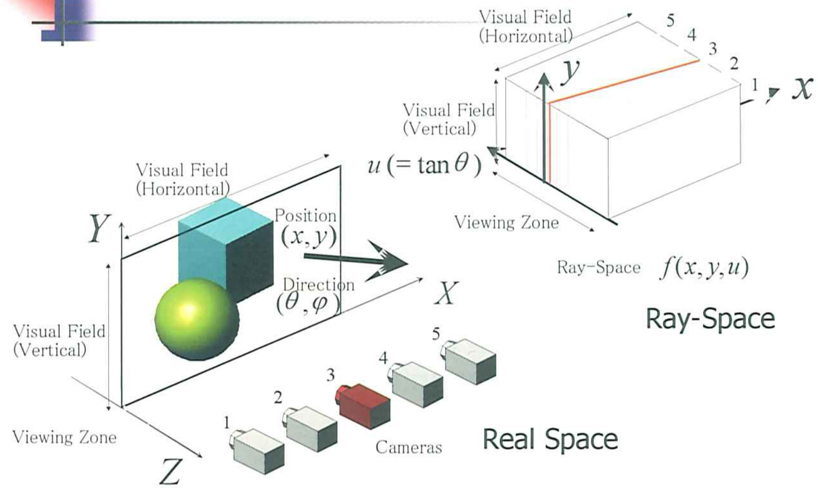
Ray-Space representation

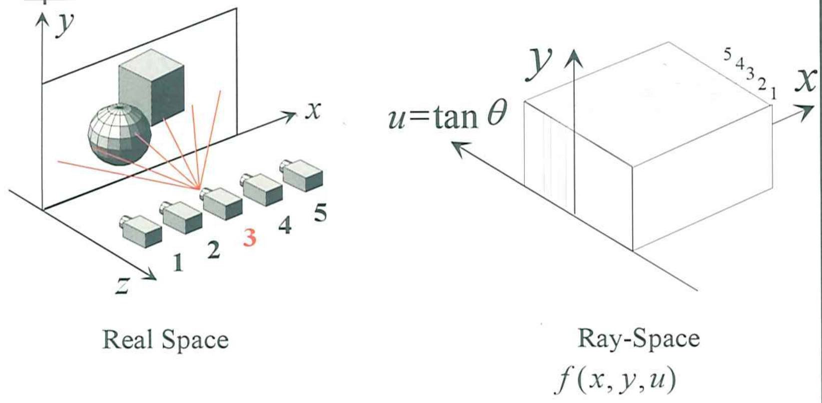


T. Fujii, "A Basic Study on the Integrated 3-D Visual Communication", Ph.D thesis of engineering, The University of Tokyo, 1994(in Japanese).

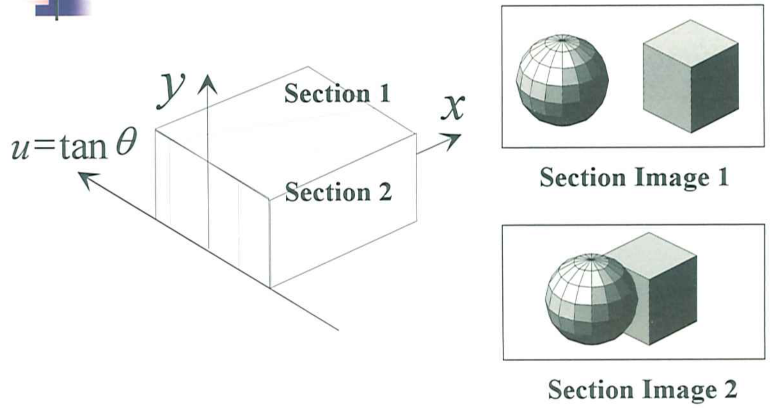
T.Fujii, T.Kimoto, M.Tanimoto, "Ray Space Coding for 3D Visual Communication", Picture Coding Symposium '96, pp. 447-451, Mar. 1996.

Acquisition of Ray-Space Data

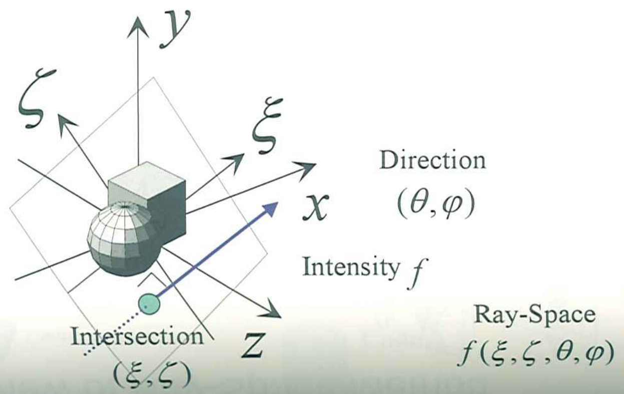




Generation of view images from the Ray-Space



Spherical Ray-Space for Convergent Views



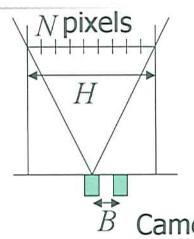
Design of Ray-Space System

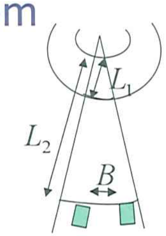
N : Number of horizontal pixels of camera

H : Width of visual field

D : Disparity of neighboring cameras

B : Baseline (camera interval)





Location Factor

$$B = \frac{H}{N} D \frac{L_2}{L_1}$$

Interpolation Factor

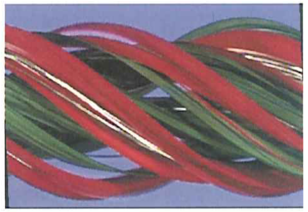
Numerical Examples:

$H = 640(mm)$, $N = 640(pixels)$

- $D = 1$, $L_2 / L_1 = 1$
  $B = 1(mm)$
- $D = 10$, $L_2 / L_1 = 1$
  $B = 10(mm)$
- $D = 10$, $L_2 / L_1 = 10$
  $B = 100(mm)$

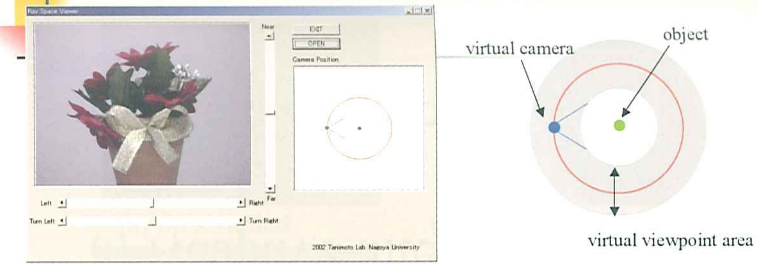
Typical examples of horizontal cross-section of the ray-space

(a) orthogonal ray-space



(b) spherical ray-space

Ray Space Viewer



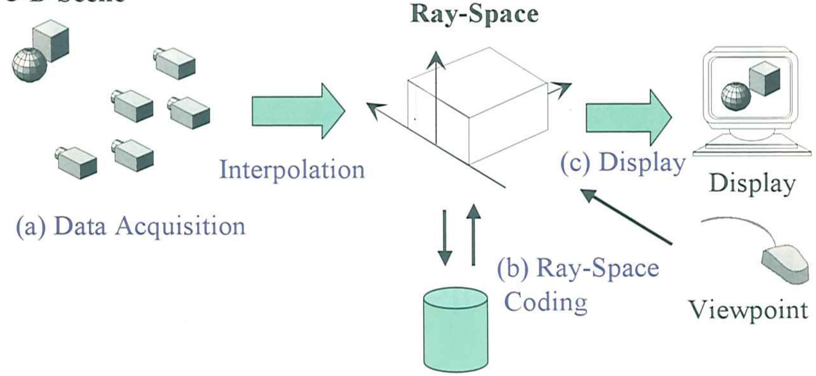
Data:

4 view images per degree , totally 1440 view images

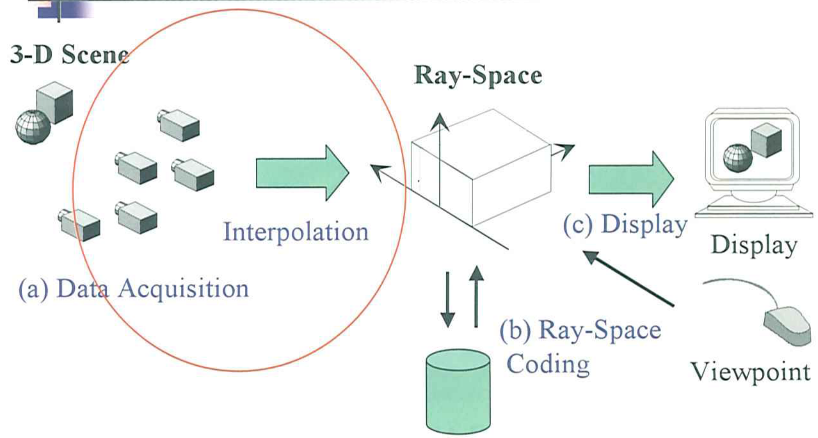
View image generation due to the motion of virtual camera:

Orbital motion
Forward and backward
Turn right and left

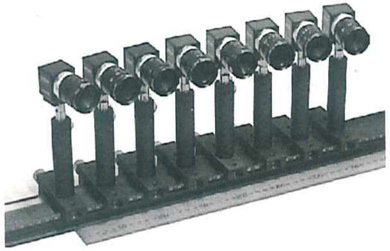
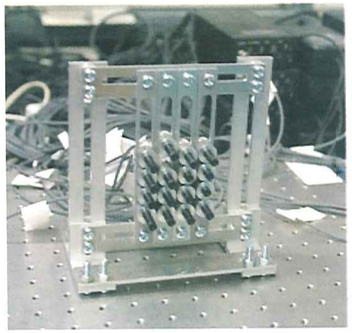
3. Free-Viewpoint TV System



3. Free-Viewpoint TV System

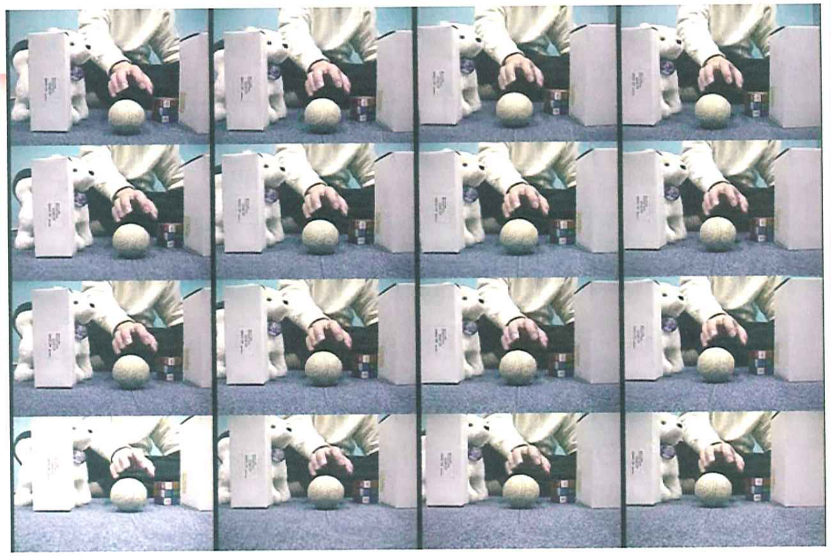


Multi-camera system

Camera platform (1D, 2D array)

Multi-camera acquisition system



Multi-camera system

- Too sparse to acquire Ray-Space
  - Interpolation is required
- Calibration is required
  - Not only geometrical calibration but also brightness, color, … etc.
- Synchronization problem

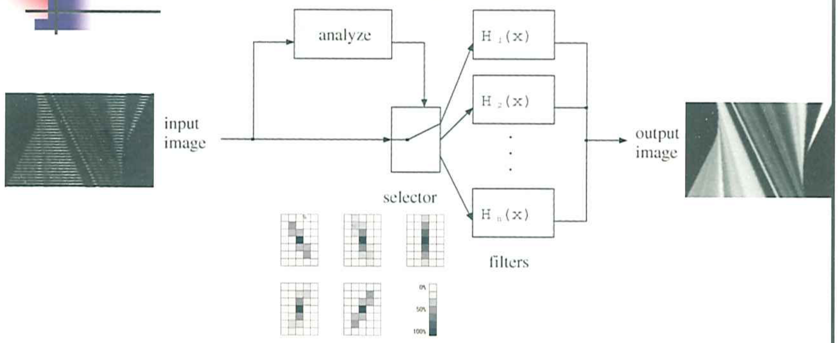
Interpolation of Ray-Space

(1) Adaptive filtering

(2) Table look-up

Ray-Space Interpolation

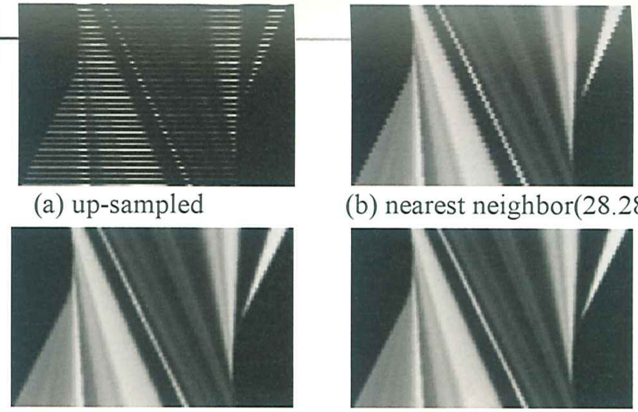
(1) Adaptive filtering



[Ref] T. Kobayashi,T. Fujii,T. Kimoto,M. Tanimoto," Interpolation of Ray-Space Data by Adaptive Filtering" , SPIE Electronic Imaging 2000, 2000.1.

Ray-Space interpolation by adaptive filtering

Comparison of interpolation methods



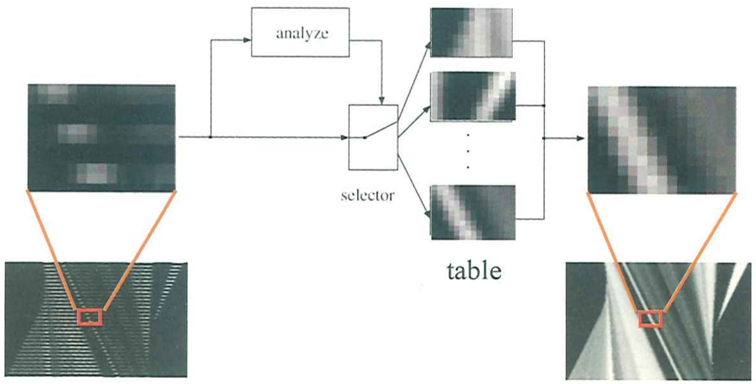
(a) up-sampled

(b) nearest neighbor(28.28dB)

(c) linear(33.47dB)

(d) proposed(35.40dB)

(2) Table look-up interpolation



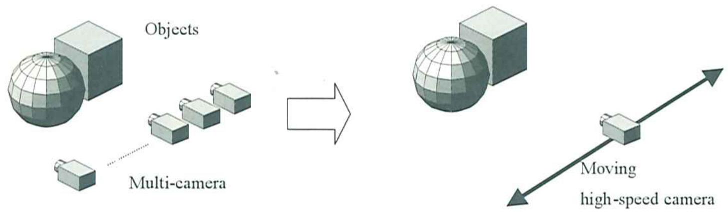
Real-time Dense Ray-Space Acquisition system

(1) Multi-camera system

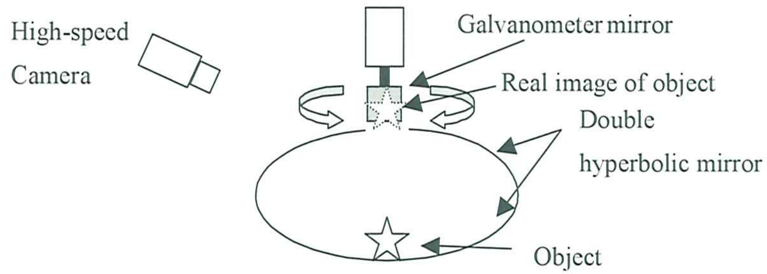
(2) High-speed camera + optical scanning system

(3) Lenslet system

(4) and more…

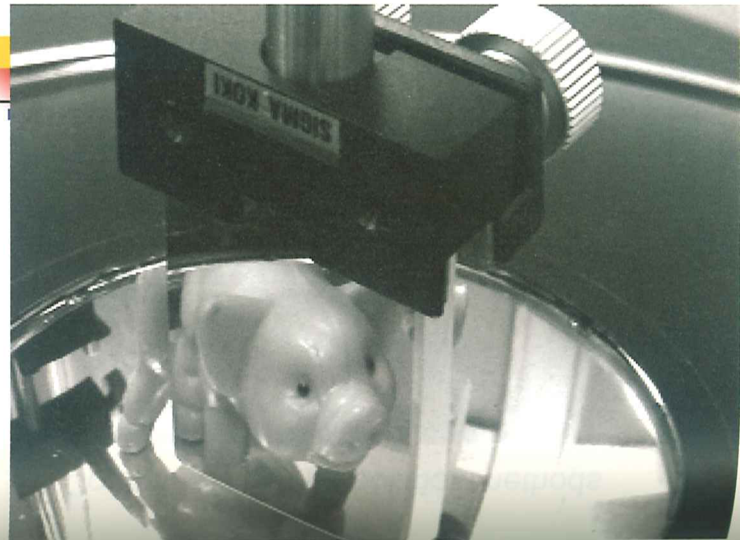


New Ray-Space acquisition system

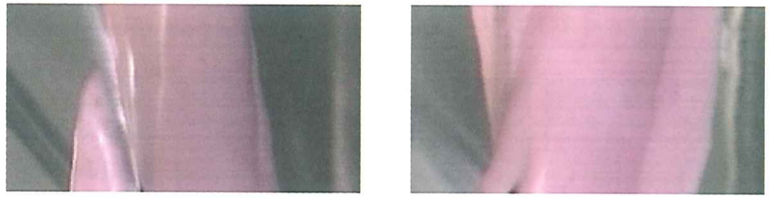


Configuration of the experimental system

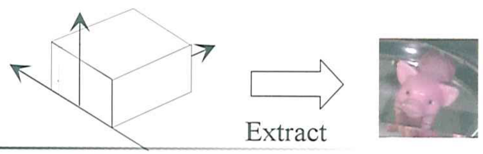
New Ray-Space acquisition system

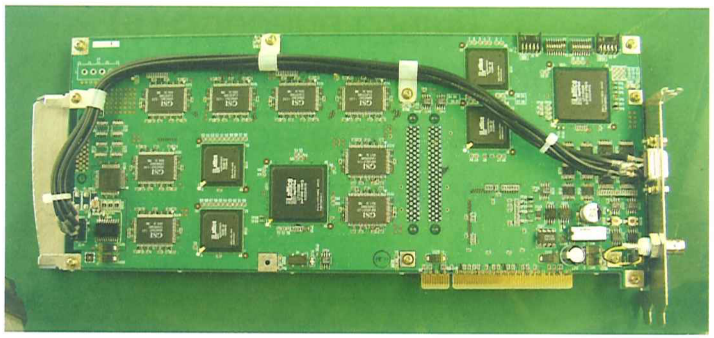


Reflection images in the galvanometer mirror



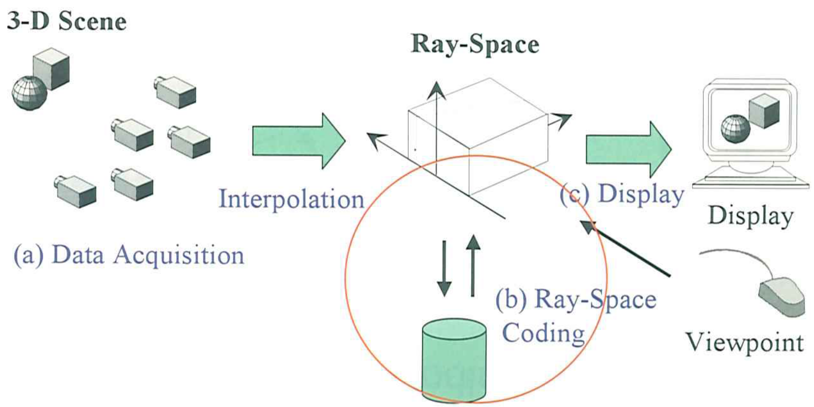
Examples of the Ray-Space obtained by the system (2-D section image)





to be presented at SPIE Electronic Imaging 2004, San Jose

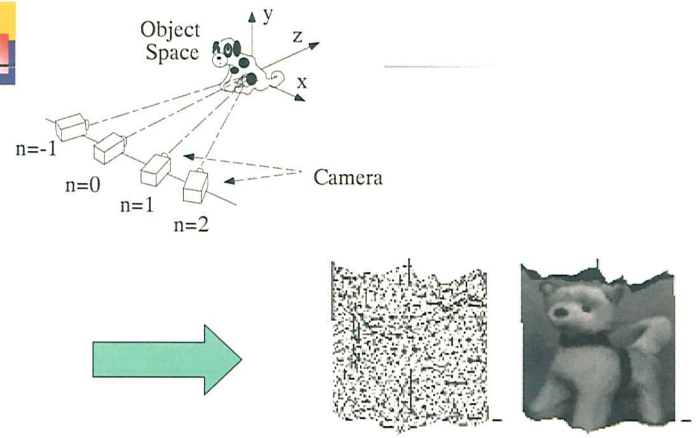
3. Free-Viewpoint TV System



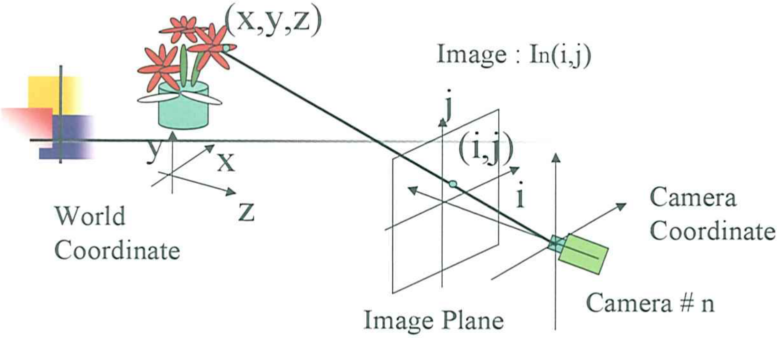
(b) Ray-Space Coding

(1) Model-based Coding
(2) Shape-adaptive DCT
(3) Vector Quantization (VQ)
(4) Subband coding

(1) Model-based Coding



[Ref] Toshiaki Fujii, Hiroshi Harashima, Data Compression and Interpolation of Multi-View Image Set IEICE Transactions on Information and Systems, Vol. E77-D, No. 9, pp. 987-995, 1994.9.



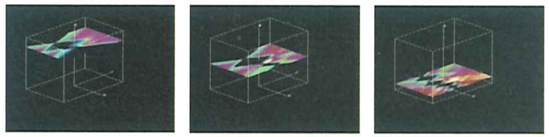
Averaged color space:

$$\text{Avg}(x,y,z) : Avg(x,y,z) = \frac{1}{N}\sum_{n=1}^{N} I_n(i,j)$$

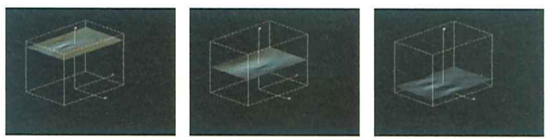
Variance space:

$$\text{Var}(x,y,z) : Var(x,y,z) = \frac{1}{N}\sum_{n=1}^{N} (I_n(i,j) - \text{Avg}(x,y,z))^2$$

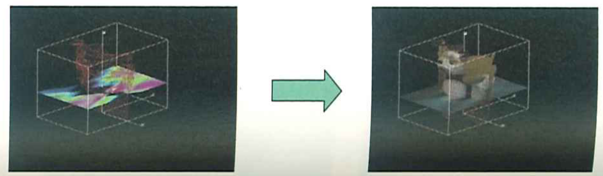
Example of Var(x,y,z)



Example of Avg(x,y,z)



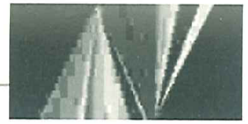
Object shape and texture estimation :



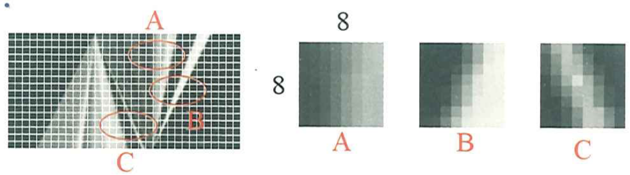
Using Var(x,y,z) **Using Avg(x,y,z)**

(2) Shape-adaptive DCT

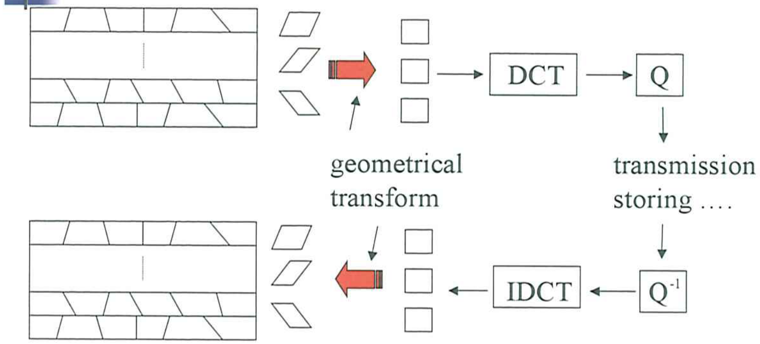
Blocking artifact arises in low bit rate.



Why?



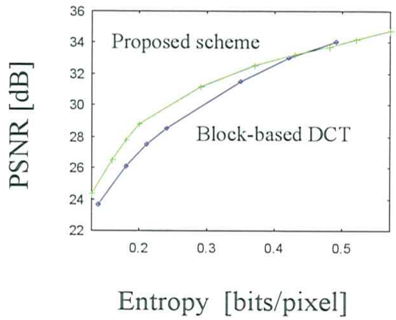
The DCT coefficients of various block patterns are quantized by the SAME Quantization Matrix.



The original image is segmented into variable shape blocks. DCT is applied to geometrically transformed blocks.

Coding experiments

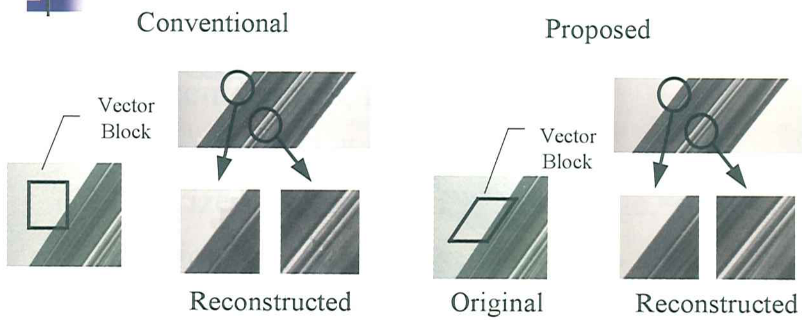
Ray-Space data:
- computer generated
- size : 256 × 128, 8 [bits/pixel]



Bits for shape :

0.047 [bits/pixel]

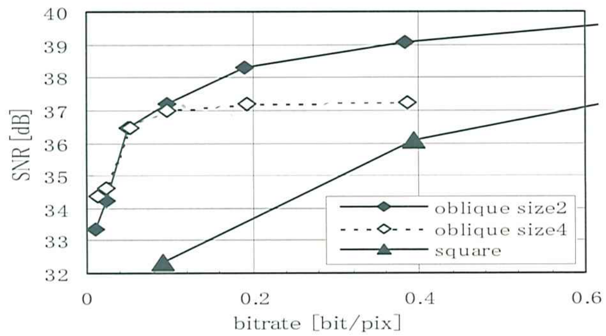
(3) Vector Quantization (VQ)



(a) Square Division

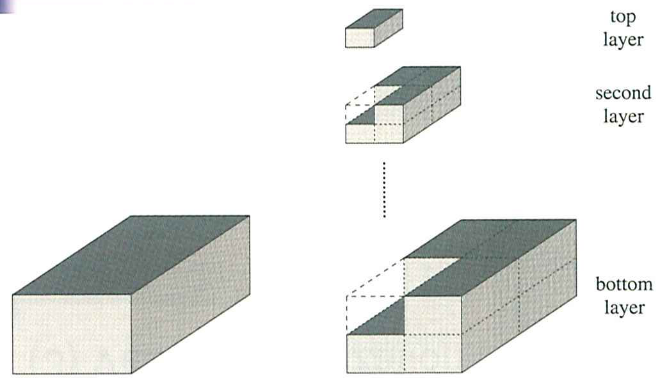
(b) Oblique Division

Coding Performance by VQ



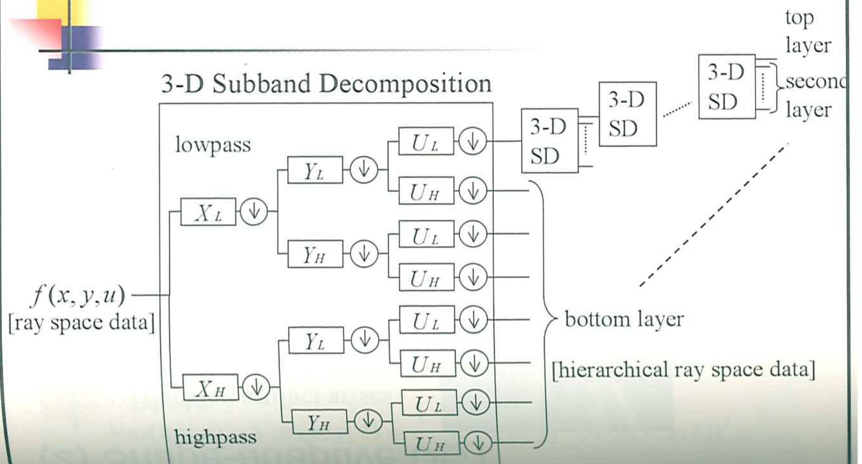
(4) Subband Coding

Hierarchical Ray-Space for Scalable Processing

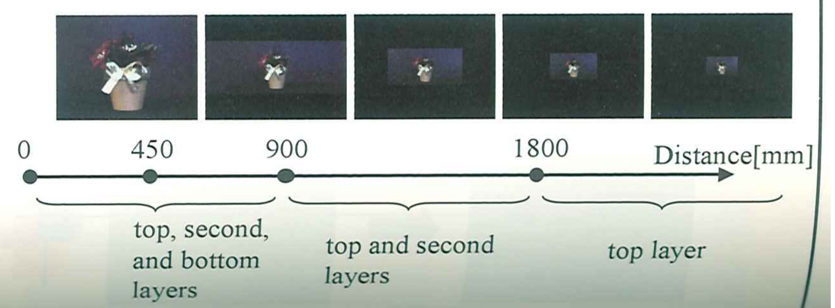


(a) Ray space (b) Hierarchical ray space

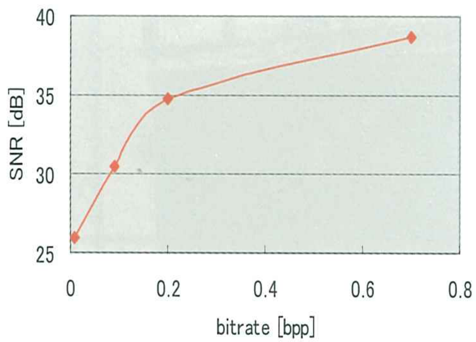
Hierarchical representation of Ray-Space



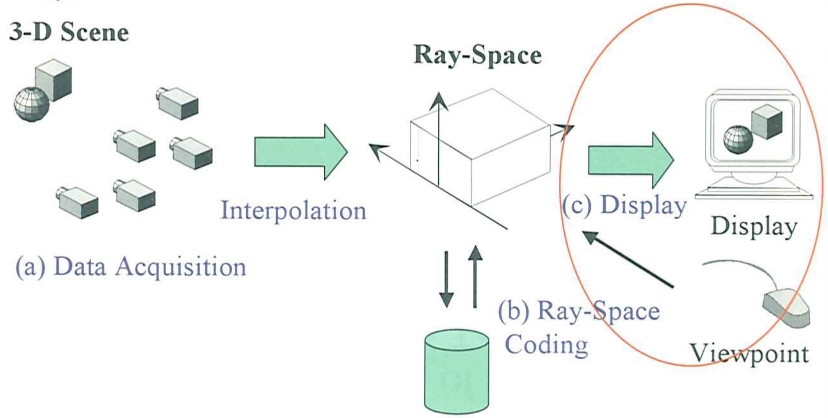
View Image Generation from the Hierarchical Ray Space



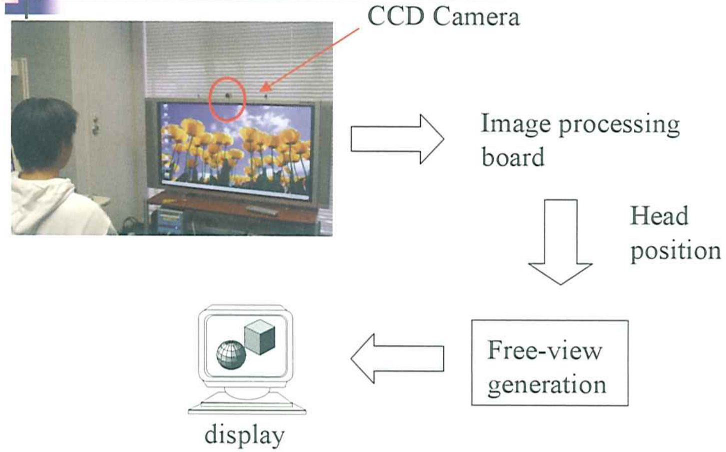
Coding Performance by SPECK



3. Free-Viewpoint TV System



Head-tracking system



Algorithm

(1) frame difference is calculated between adjacent frames, i.e.,

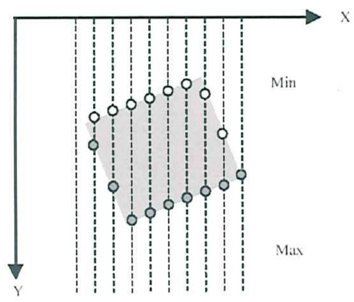
$$g(x,y,t)=|f(x,y,t)-f(x,y,t-1)|$$

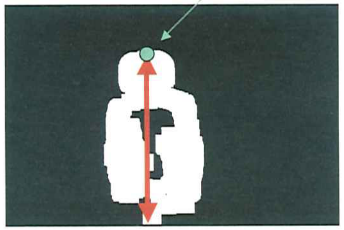
(2) binarize the frame difference image,

(3) dilation and erosion are applied to eliminate isolated points,

(4) in the object region, find the X_h that gives the maximum height Ymax-Ymin.

Head-tracking system





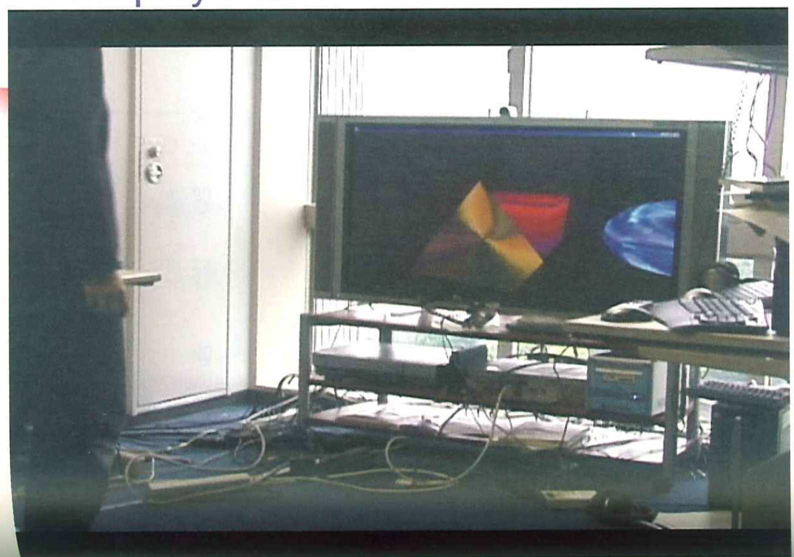
$$X_h = \arg\max_X |Y\max - Y\min|$$

Experiment

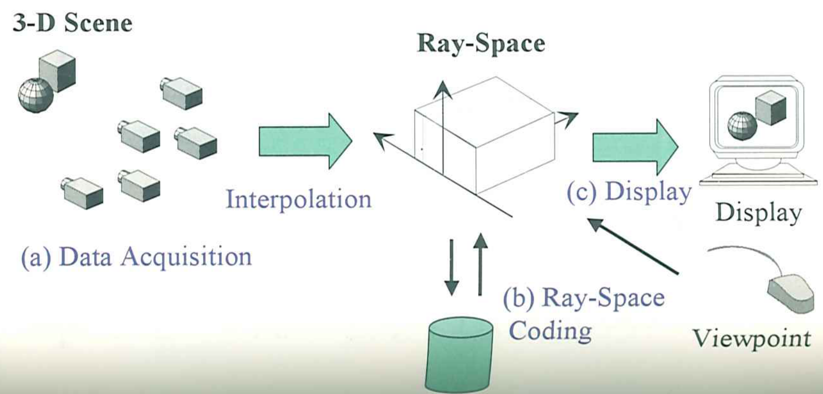
Table Specification of experimental setup

Display Monitor	Pioneer PDP-A503HD
CPU	Pentium4 2.8GHz
OS	Windows XP Professional
Compiler	C++ Builder 5
CCD camera	Toshiba IK-SM43H
Image processing board	Hicos IP7000BD

FTV Display Demo

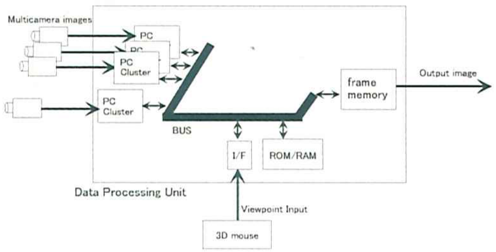


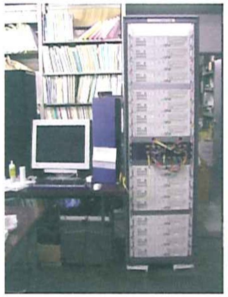
4. Demonstration of Free-Viewpoint Television (FTV)

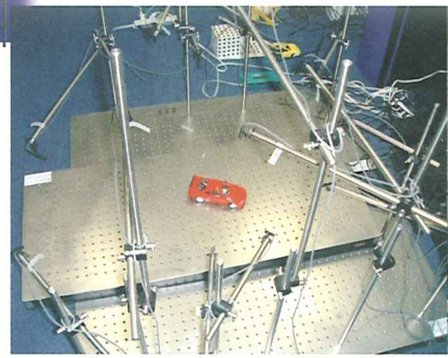


System Architecture

- PC cluster with 16 PCs
- Each PC is connected by Gigabit Ethernet
- Each PC has a image capture board

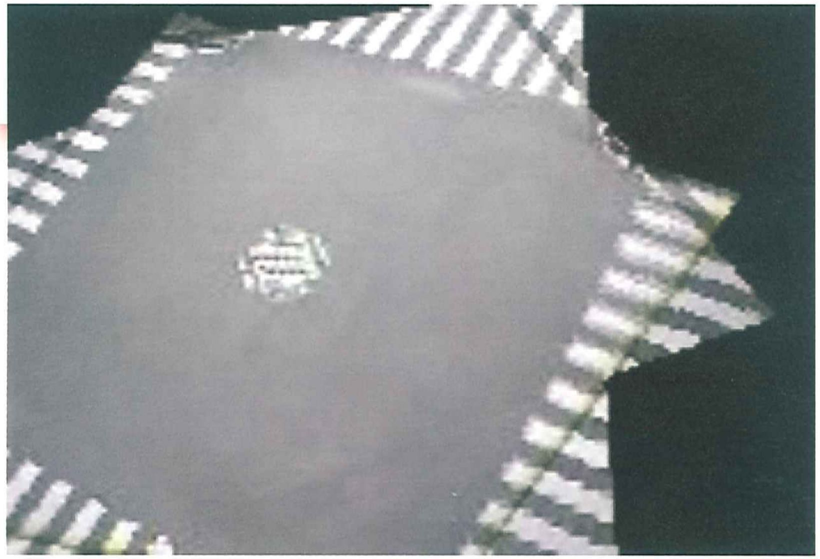






Camera setup (16 cameras)

System overview



160×120 pixels, 10 frames/sec

Specification of “aquarium” demo

- Capturing
-1. Number of Cameras : 16 cameras.
-2. Camera Arrangement : One-dimensional Arc line with 33 cm radius.
-3. Distance between each camera : 3 degree.
-4. Capturing Frame Rate : 10 fps.
-5. Captured Image Size : 320x240 pixels.
-6. Rectified Image Size : 310x230 pixels (cut 5 pixels from each edge).

-Rendering
-7. Number of Interpolated View between each camera pair: 15 View.
-8. Total View Position: 241 Views.
-9. Distance between each view : 0.1875 degree.
-10. Output Frame Rate : 10 fps.

FTV Demo - Aquarium

MPEG activity

MPEG: Moving Picture Experts Group
(3DAV Ad-hoc Group)

- **FTV — Free Viewpoint Television**,
  ISO/IEC JTC1/SC29/WG11 M8595, July 2002.
- **Comparative Evaluation of Ray-Space Representation**
  ISO/IEC JTC1/SC29/WG11 M8892, October 2002.
- **Quality Measure for Ray-Space Interpolation**
  ISO/IEC JTC1/SC29/WG11 M9238, December 2002.
- **Multi-View Video Acquisition System for FTV Experiment**
  ISO/IEC JTC1/SC29/WG11 M9472, March 2003.
- **Comparative Evaluation of Ray-Space Interpolation Methods**
  ISO/IEC JTC1/SC29/WG11 M9808, July 2003.
- **Ray-Space Coding Using Temporal and Spatial Prediction**
  ISO/IEC JTC1/SC29/WG11 M10178, October 2003.

Authors: Masayuki Tanimoto and Toshiaki Fujii

Conclusions

- Free-Viewpoint Television System based on Ray-Space method was demonstrated.
- The full real-time system covering from video capturing to display was developed.
- The following key technologies were described:
  - Data acquisition
  - Ray-Space coding
  - Display system