A new archiving system for TV studio sets using depth camera and global illumination

Akihiko Shirai [(1)], Kiichi Kobayashi[(1)], Masahiro Kawakita[(2)],Suguru Saito[(3)] and Masayuki Nakajima[(3)]

{shirai, kobayaki}@nes.or.jp, kawakita.m-ga@nhk.or.jp, {suguru@pi, nakajima@img.cs }.titech.ac.jp

(1) NHK Engineering Services Inc. (2) NHK Science and Technical Research Labs.

(3) Tokyo Institute of Technology

Abstract

This paper describes a new archiving system for TV broadcasting studios using the Axi-Vision depth camera and global illumination with high-dynamic-range images (HDRI).

Axi-Vision is a depth camera using near-infrared LED arrays, image intensifier, and coaxial optical systems separated by a dichroic prism. It can simultaneously generate in real-time a normal-RGB image and a grayscale image for recording depth information.

We've developed a converter to build a three-dimensional mesh from the depth image and wrap texture from the normal RGB images.

1.Introduction

Over the last few years, photo-realistic computer graphics movies and visual effects (VFX) have expanded their application fields from cinema to TV commercial films (CF). In particular, video composite technologies like lighting technologies called Image Based Lighting (IBL) using the high-dynamic-range light sources obtained from actual images, image composition using chroma key video matting and match move software have become very popular in the cinema industries.

However, these photo-realistic CG technologies still face several technical hurdles before they can be applied to TV broadcasting program production.

We've already studied basic technologies for a next-generation TV production environment for making high-quality content. In this report, we describe how to archiving images of actual TV studio sets (for making three-dimensional visual effects) by using the Axi-Vision depth camera and high dynamic range images (HDRI).

2.Requitement of the next-generation virtual studio for TV program production

Currently, most visual effects in TV programs are two-dimensional ones that are generated offline during post-production by using Adobe AfterEffects™. In particular, real-time composites that could appear in, say, live broadcasts use chroma key for making virtual sets or use Kaydara FiLMBOX™ for generating the interactions of virtual actors. Most of composites of this kind are two-dimensional. The final images also have less shadow and occlusion. The production time per frame of composites generated offline for weekly entertainment and educational programs are less than for cinema or CF (obviously this is also true for the real-time composites on news programs).

In the future, broadcasts will be predominantly HDTV (1920x1080 pixels, 30 fps, RGB 10-bit, digital), and thus the definition of "photo realistic CG" will shift to higher quality images. TV production in the future should also support a higher level of integrated interactivity. We suppose that the three keywords of CG technologies for the next generation of TV contents production are Reality, Simplicity, and Rapidness.

In this report, we focus on archiving images of actual TV sets including photo realistic VFXs. For example, consider a TV program series that broadcasts every week. Currently, the actual TV studio sets are designed by CAD, and most of the data generated during the design phase are not passed on to the people responsible for creating the VFXs. Even if the data were passed on, it wouldn't be enough to render photo-realistic graphics with commercially available CG software. Moreover, the TV sets and props that actually end up being built will often look different from the CAD design. The VFX production section should receive not only information on the TV sets but also lighting information (the lighting chosen by the lighting engineer). Of course, TV sets always get broken up after shooting for reasons of cost, lack of space, schedule, etc. As the result, in most cases, the VFX section must create a replica of the studio to render "fake photo realistic images".

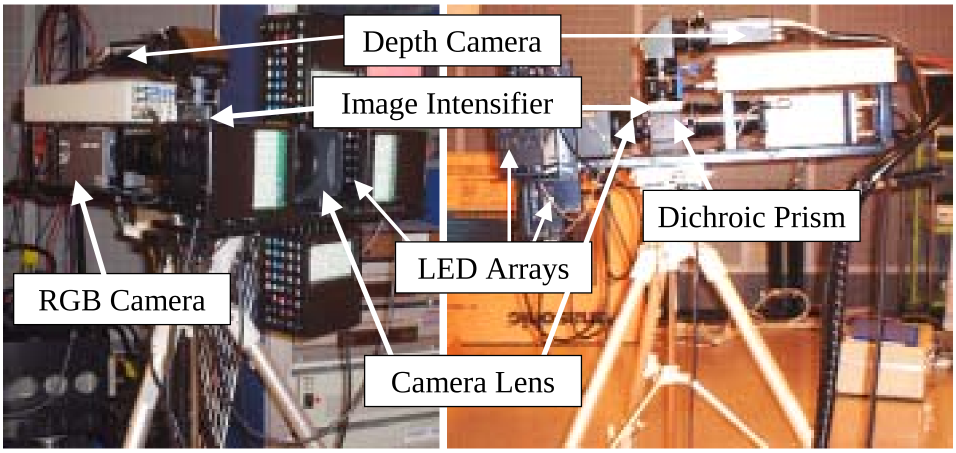

Figure 1: Photograph of Axi-Vision Camera

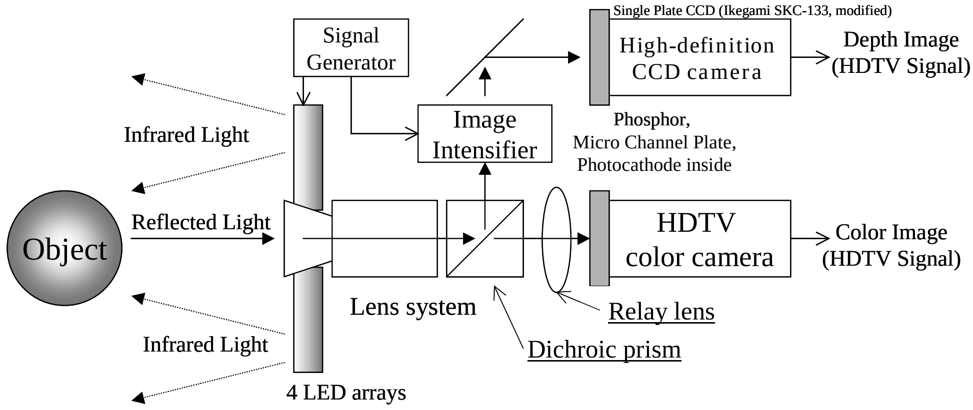

Figure 2: Configuration of the Axi-Vision Camera

Our new system overcomes these problems of retaining enough information throughout the production process. It uses HDRI techniques to archive the indoor lighting environment and a special HDTV camera, Axi-Vision, that can take depth images and RGB images at the same time. It can generate three-dimensional models and lights with normal images. It is easy to use and doesn't interfere with the main camera.

3.Axi-Vision

Axi-Vision is a special HDTV camera developed by NHK Science and Technical Research Laboratories [1]. It can simultaneously take grayscale depth images and normal HDTV images. There are several methods of getting depth information. For example, one can use the time of light flight, stereoscopic measurement with several cameras, or moiré with image processing. Most of these methods need processing time and scan at a line or point using lasers. They are difficult to use for TV VFX because of differences in resolution and image quality and the need for matching with normal (non-depth) images. A depth camera for TV VFX should be fast, and its images should be able to be automatically matched with normal images. Axi-Vision has these abilities. It can take depth grayscale images of objects in the frame and match them to RGB pixels while operating at the full rate for HDTV movies (30 fps).

The "Axi-Vision" camera system has two HDTV cameras and infrared LED arrays. These coaxial optical systems are separated by a dichroic prism. The main system is for taking normal RGB images. The other is composed of high-definition CCD camera and a specially developed image intensifier (I.I.). The I.I. acts as an ultrahigh-speed (1 nanosecond) shuttering device with high resolution (Fig.1, 2).

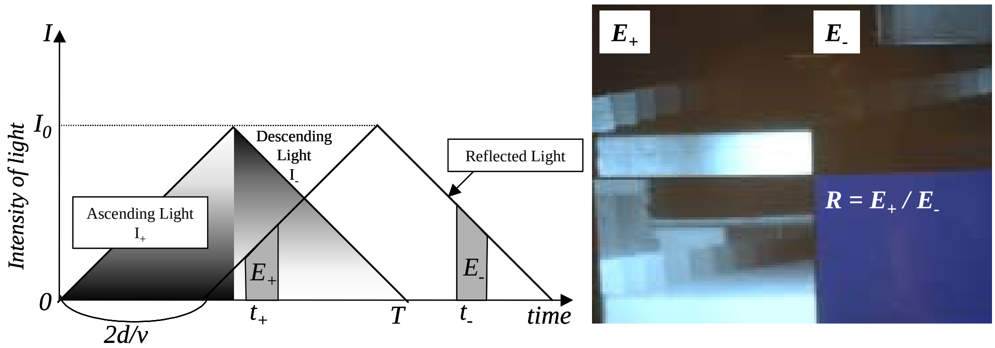

Figure 3: Principle of depth imaging

When Axi-Vision images the depth of the objects in the frame, the near-infrared LED arrays (1 W) illuminate the objects. The light intensities of LED arrays are modulated by triangular wave. Referring to Fig. 3, the ascending light intensity I_+ and descending light intensity I_-, are written as

$$I_+ = st \quad \ldots(1)$$

$$I_- = -s\left(t - \frac{1}{f}\right) \quad \ldots(2)$$

$$s = 2I_0 f, \quad \ldots(3)$$

where s is the rate of increase or decrease, I_0 is the maximum intensity of light, and f is the modulation frequency. At depth d, the intensity E_+ of light reflected from an object including reflectivity σ is

$$E_+ = \frac{\sigma s}{\left(4\pi d^2\right)^2}\left(t_+ - \frac{2d}{v}\right), \quad \ldots(4)$$

where t_+ is the instant that the shutter is opened during the ascending illumination, and v is the velocity of light. Likewise, the intensity E_- of the light reflected from the object as the illumination decreases is

$$E_- = \frac{\sigma s}{\left(4\pi d^2\right)^2}\left\{\frac{1}{2f} - \left(t_- - \frac{2d}{v}\right)\right\} \ldots(5)$$

where t_- is the instant that the shutter is opened during the descending illumination. If t_- is set as

$$t_- = t_+ + \frac{1}{2f}, \quad \ldots(6)$$

the reflectivity σ and divergence factor $1/\left(4\pi d^2\right)^2$ can be removed from the expressions by dividing E_+ by E_-. Finally, the depth d of the object is obtained as

$$d = \frac{v}{2}\left\{t_+ - \frac{1}{2f}\left(\frac{R}{1+R}\right)\right\} \ldots(7)$$

$$R = \frac{E_+}{E_-} \quad \ldots(8).$$

The depth d can thus be obtained from the ratio R of the reflected light intensity detected when the illumination intensity is

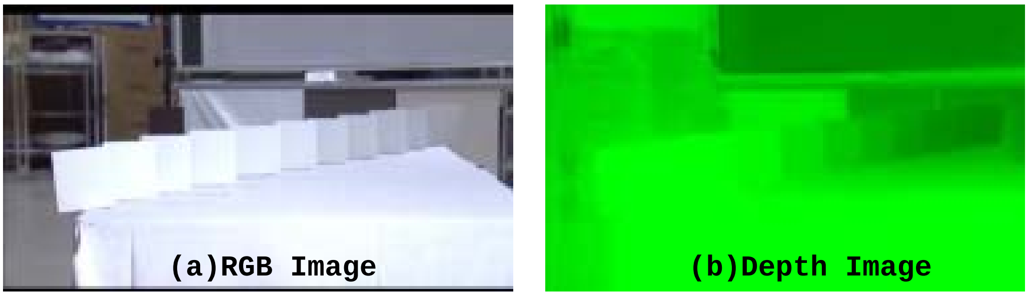

Figure 4:Experiment scene and images

ascending and descending [2].

As a result, we can take 10-bit grayscale images at each pixel depth, which can be directly related to RGB images. This camera was developed for a real-time depth key generator for virtual sets. It is tolerant to changes in surface reflectance and in optical parameters such as zooming and defocusing.

4.Modeling with depth images

Axi-Vision can get the depth image quickly and match it directly with the RGB image, unlike other methods such as image processing or laser scanning. Thus, if the Axi-Vision can precisely measure the studio set including texture and collocations, it can be expected to be a useful tool for archiving virtual sets.

We used Axi-Vision to shoot scenes that have simple paper target planes to build a three-dimensional mesh with depth image. Fig.4 (a) shows an RGB image shot with the Axi-Vision camera. There are eleven white target planes faced orthogonally to the

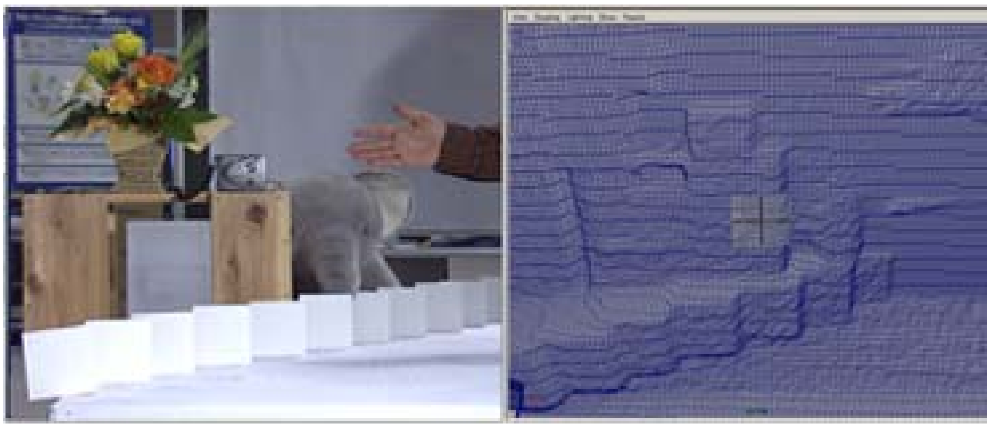

Figure 5: Original image and meshed data from depth image on Alias Maya™

camera axis. In this case, the white plane papers are located 100 mm apart 2000 to 3000 mm from camera lens center. Fig.4 (b) shows a depth image obtained by Axi-Vision. It shows clearly a difference in the density between each plane.

The images were saved as TIFF 16-bit files via Cineon 10-bit. We developed a converter that can convert TIFF 16-bit to Wavefront OBJ. To confirm the result of such a conversion, we used Alias Maya™ to load the OBJ file and rendering. Fig.5 is a part of meshed data from a depth image. It shows a clear difference in level for each plane.

5. Archiving studio lighting with HDRI

To archive the studio lighting environment, we took multi-grade intensity fisheye images using a Nikon COOLPIX 5000 with a fisheye converter CS-E5000 and neutral density filter. To convert them into HDRI, we used HDR Shop by Paul Debevec [3,4]. It is helpful when building a rough model of the set and deciding the locations of studio lights (Fig.6).

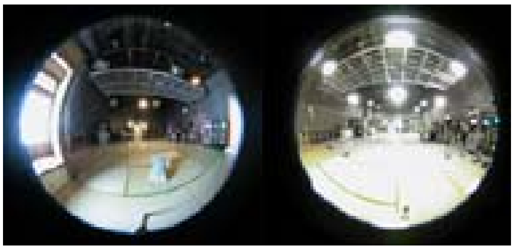

Figure 6: Fisheye images of TV sets under different light environment

6.Conclusion

We confirmed that the Axi-Vision developed by NHK Science and Technical Research Laboratories, can use to build a three-dimensional mesh model and that HDRI can archive a photorealistic lighting environment under studio light global illumination.

These results show that a depth camera, Axi-Vision and HDRI global illumination can be expected to realize a new archiving system for photorealistic TV studio sets.

Acknowledgements

This work was supported by the Telecommunications-Advancement-Organization (TAO) of Japan.

References

[1] M. Kawakita, T. Kurita, H. Hiroshi, and S. Inoue: "HDTV Axi-vision Camera," *Proc. IBC (International Broadcasting Convention) 2002*, September, Amsterdam pp.397-404 (2002).

[2] M. Kawakita, K. Iizuka, T. Aida, H. Kikuchi, H. Fujikake, J. Yonai, and K. Takizawa, Axi-vision camera (Real-Time Depth- Mapping Camera), *Applied Optics*, Vol.39, pp.3931-3939 (2000).

[3] Paul E. Debevec. Rendering Synthetic Objects into Real Scenes: Bridging Traditional and Image-Based Graphics with Global Illumination and High Dynamic Range Photography, Proc. of SIGGRAPH'98, ACM Press, pp.189–198 (1998).

[4] HDR Shop, http://www.debevec.org/HDRShop/