VR@Home: A Personal VR Studio Platform

Wonwoo Lee and Woontack Woo

***Abstract*—Recently, people emerge as a contents provider with increasing usage of Internet in our daily life. People produces large portion of multimedia contents served on the Web. With this trend the personal broadcasting is being focused as a new application area of the Internet services. In this paper, we proposed VR@Home, a personal VR studio platform for personal broadcasting at home. VR@Home platform creates 3D scene by merging the real world contents and the virtual contents. 2.5D background and 3D objects generated from images of a real scene provide more immersion to users. The real-time soft shadows created by the estimated light sources keep the consistency between the real world and the virtual world. Through the personal virtual studio platform, a user can make his/her own contents and distribute or share them with the others on the Internet. Our platform is applicable to education in remote site, entertainment, and internet TV applications.**

***Index Terms*— Virtual Reality, Image processing, Machine vision, Multimedia systems**

I. Introduction

TRADITIONALLY, people has accepted the contents produced by contents providers, such as broadcasting company and Potals on the web. They read newspapers, listen to radio programs, or watch TV everyday. However, people do not remain as contents consumers anymore nowadays. Instead of being a passive consumer, each of them becomes an active producer of multimedia contents. People not only create multimedia contents, such as pictures, movie clips, and flash animations, but also share theirs with others on the Internet (e.g. Blogs, Google video service). It becomes popular to upload multimedia contents on the web and to share them with other people. Recently, there are several services which enables a person to broadcast personally at their home (e.g. PodCast [1], Internet radios). This phenomenon is accelerated by recent high-performance PCs. With this trend, personal broadcasting solution market grows faster and the requirement of them increases.

The broadcasting has been usually managed by companies. However, personal broadcasting is done in a home environment where several restrictions exist. Since conventional broadcasting systems require expensive AV devices and large studio, it is not possible to set up all those large and expensive devices at home. There emerge several challenges, such as 3D contents creation, user segmentation, and realistic synthesis of the real and virtual scene, when we bring the conventional VR studio to our home. From this viewpoint, augmented reality (AR) and virtual reality (VR) technologies can provide good solutions to overcome the restrictions of a home environment.

AR and VR technologies are already adopted by many broadcasting companies and applied to advertisements, sports, and documentary programs successfully. BBC produced 3D contents by modeling actors and a real indoor scene in 3D [2]. In MixTV project, BBC took experiments of applying AR technologies on different environments such as schools, homes, and studios [3]. NHK tried to remove spatial restriction of the virtual studio with blue screen by modeling the real studio and merging it with the real scene [4]. Recently, AR technologies become to be used in personal applications. Virtual objects such as a eyeglass, and virtual masks are augmented on a user's head in real-time in [5]. Since conventional broadcasting systems require expensive AV devices and large studio, it is not possible to set up all those large and expensive devices at home. There emerge several challenges, such as 3D contents creation, user segmentation, and realistic synthesis of the real and virtual scene, when we bring the conventional VR studio to our home.

In this paper, we propose VR@Home, a personal VR studio platform for personal broadcasting at home. The objective of our platform is to let users create and broadcast 3D virtual contents at their home through internet services. VR@Home platform creates 3D scene with 2.5D background model, 3D object model, the segmented user, and other CG contents. 2.5D background and 3D objects generated from images of a real scene provide more immersion to users. The lighting condition of the real world is estimated and applied to the virtual world. It preserved the consistency of the real world and the virtual world. From the estimated light source, soft shadows are rendered in real-time. It makes the virtual objects are integrated to the virtual scene more realistically. In addition, VR@Home enables users to interact with virtual contents through intuitive user interface. VR@Home consists of the following five components which have the same functionalities as the conventional VR studio system does.

- **Camera calibration module**: Camera calibration is necessary before we exploit images captured from it for modeling real scene. We provide camera calibration module which uses a specially designed pattern for image-based modeling.
- **Image-based modeling module**: We provide modules for 2.5D background modeling and 3D object modeling. From image sequences our modeling module generate 3D mesh model from the image sequences automatically.

This research is funded by ETRI OCR.

The authors are with GIST U-VR Lab., Gwangju 500-712, S. Korea (e-mail: {wlee, wwoo}@gist.ac.kr).

- **Realistic composition module**: To generate more realistic 3D scene, the lighting condition of the real world is estimated from images. When rendering 3D scene, the shadows are generated in real-time from the estimated light information.
- **Real-time user segmentation module**: Without blue screen, we segment a user from a complex static background in real-time. Our segmentation module learns the background of a user for segmentation and update background for robustness.
- **Tangible AR interface module**: We adopt a tangible AR interface for user interaction. The user can move and rotate the virtual object freely on his hand through the tangible AR interface.

Through the personal virtual studio platform, a user can make his/her own contents and distribute or share them with the others on the Internet.

This paper is organized as follows. In chapter 2, we explain five modules of VR@Home platform. Implementation of VR@Home is described in chapter 3. We present conclusion and future works in chapter 4.

II. PERSONAL VR STUDIO PLATFORM

A. Camera calibration module

Camera calibration is estimation of the intrinsic parameters (focal length, principal point) and the extrinsic parameters (position and orientation) of a camera from images captured by it. In VR studio system, camera calibration should be preceded to integrate the virtual contents with real scene. Another application of camera calibration is image-based 3D modeling and rendering.

Our camera calibration is designed for 3D object modeling. When a user captures an image sequence of an object, our camera calibration module estimates the camera parameters. We use a specially designed pattern for robustness of camera calibration. The pattern contains square markers each of which has a unique ID.

(a) (b) (c)

Fig 1. Camera calibration with vanishing points (a) Input image (b) Binary image for marker detection (c) Estimated lines from marker's corners

An image is converted to binary image and the marker's corner converted to binary image and the marker's corner is detected and the lines passing through the detected corner points are estimated. Finally, Camera parameters are estimated from the vanishing lines and vanishing points.

B. Image-based modeling module

To enable users to make personalized contents, we provide image-based modeling modules. Since image-based modeling can reduce intervention of user for 3D modeling which occurs frequently in modeling with 3D modeling tools such as Maya or 3DS MAX, it can be used by the users who have not much knowledge of 3D modeling. Our modeling module consists of 2.5D background modeling part and 3D object modeling part.

2.5D background of an indoor scene is reconstructed from images and corresponding 3D point data provided by a multi-view camera. The images of a scene are captured at several positions, and we obtain corresponding 3D point data of entire scene. After data acquisition, 3D points are integrated to a single point cloud [6]. Finally, a mesh model of the scene is generated from the point cloud. The images are used as textures of the mesh model [7]. Textures captured from the real scene improve the realism of the reconstructed model.

Fig 2. The reconstructed 2.5D background model from images

For 3D object modeling, we used a single camera instead of a multi-view camera. We assume we know the intrinsic and extrinsic camera parameters accurately by calibration. The captured images used in camera calibration and the estimated camera parameters are used as input. The object is segmented from the pattern by comparing colors in normalized RGB space. We exploit silhouette and texture information for reconstruction. Using silhouette a visual hull of the object is retrieved as an initial 3D model. The initial model is refined using silhouette, smoothness, and texture constraints. After iterative refinement, the final model of the object is created. Fig 3 shows 3D reconstruction of a teapot.

(a) (b) (c)

Fig 3. 3D object reconstruction (a) Selected silhouettes of a teapot (b) Visual hull reconstruction as initial model (c) Refined model after deformation

C. Realistic composition module

When we merge the real scene and the virtual scene there exists inharmony between them. One way to improve the realism of the scene is adding shadows to inserted real scene. To make the shadows consistent with the real space lighting condition we have to know it.

Light source estimation is the calculation of the lighting information of the real scene from images. We place two or more spherical objects which reflect the light well in the scene. After the scene is captured by a camera, point light sources are estimated from the reflection of the light on the objects. In the image, we find pixels which have strong intensities in object area. The pixels are matched among the object images, and the corresponding 3D position is calculated by back-projection of the ray reflected on the specular points S_1 and S_2 from a camera center C in spherical coordinates. Fig 4 shows how the point light source can be found.

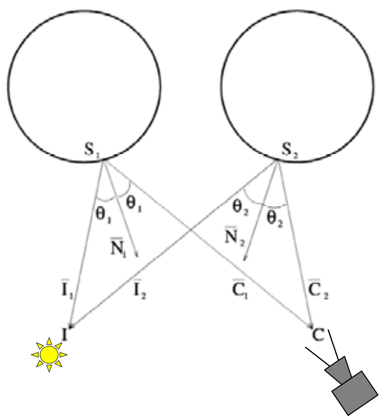
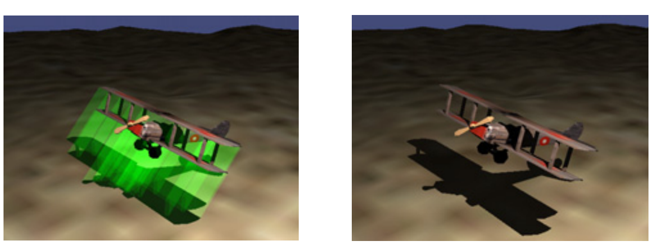


Fig 4. Point light source estimation

Shadows are one of effective components which improve the realism of the rendered scene. We can recognize the relative position and size among different objects and the shape of a surface where shadows are drawn [8]. In addition, shadow generation can be implemented in real-time.

With the estimated light sources, we generate soft shadows in real-time. The soft shadows generation is based on the algorithm proposed in [9]. Multiple shadow volumes are calculated and simple shadows are generated. The simple shadows are now converted to soft shadows by applying dialation operation. For better performance, we exploit a graphics hardware in computation of shadow volumes.



(a) (b)

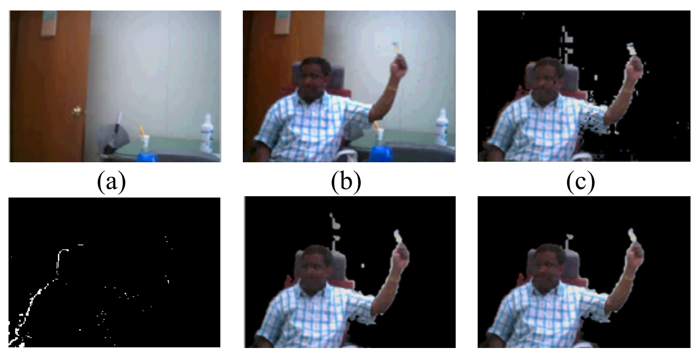
Fig 5, Shadow generation from shadow volume (a) shadow volume of the object (b) Generated shadow

D. Real-time user segmentation module

The most VR studio use blue screen for actor segmentation. However, the blue screen is not convenient and requires large space. It is not convenient to be used in home environment practically. Thus, we need to segment a user in general complex scene in real-time.

Our user segmentation module learns a static complex indoor scene without moving objects. The mean and standard deviation of each pixel of the image is calculated. When a user enters to the scene after learning, we examine how the color values of pixels change among consecutive frames. If the difference of the current color and the color accumulated during previous frames, the pixel belongs to a moving object. If not, it is classified as background.

Fig 6 shows the user segmentation result. Spot noises are removed by median filtering and connected component analysis. Shadows of a user also cause wrong classification result. For robust segmentation, we detect shadows by using HSV color scheme and remove it as shown in Fig 6(d) and (e). The mean and standard deviation of the background pixels are updated as time passes. This update makes the segmentation more robust under a small variation of lighting condition. Fig 6(e) shows the final segmentation result.



(a) (b) (c)

(d) (e) (f)

Fig 6. User segmentation result (a) Background (b) Input frame (c) Initial segmentation (d) Detected shadow (e) Shadow elimination (f) Final segmentation result after noise removal

E. Tangible AR interface module

In conventional VR studio system, the actors have to move following to a pre-defined scenario or make actions assuming that there exist virtual objects to interact with. This interaction approach causes unnatural action of an actor. In personal virtual studio system, the problem may be worse. Tracking devices can be a solution, but they are expensive to use personally.

In personal VR studio, a computer vision-based interface can be a good solution by exploiting users' USB cams. Among many types of vision-based interface, tangible AR interface provides intuitive interaction between users and virtual objects. A user can move and rotate the virtual object freely on his hand by moving and rotating a tangible object. Thus, the user does not need to pretend that there is virtual object in the real space.

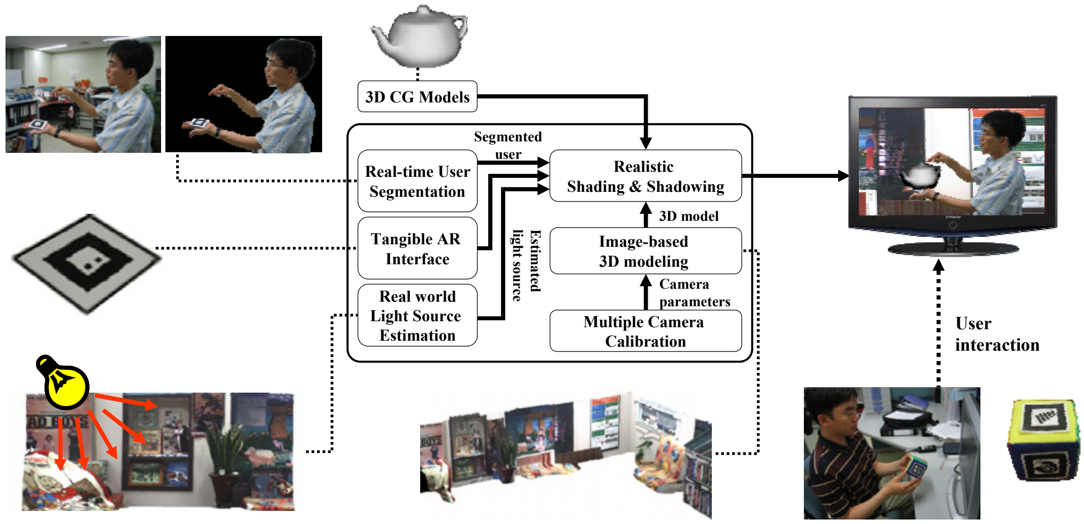


Fig 7. Flow diagram of VR@Home

III. Implementation

By integrating all modules explained in previous chapter to a single framework, we implemented VR@Home, our first prototype of personal VR studio platform. Fig 7 shows the flow diagram of VR@Home. 2.5D background and 3D object is modeled in off-line and inserted to the virtual space. Light source of the background is estimated from background images and applied to rendering. User segmentation and rendering with shadows are done in real-time. 2.5D background, 3D models, and the segmented user are integrated by rendering the virtual space. Tangible user interface is implemented with ARToolkit which support marker detection and camera pose estimation.

VR@Home works on the workstation with Dual Xeon 3.0Ghz, 4GB RAM, and QuadroFX3400. We use a USB camera to capture images for 3D object modeling and user segmentation. For 2.5D background modeling Digiclops [11] manufactured by Point Grey Research is adopted.

IV. Conclusion

In this paper, we proposed VR@Home, a personal VR studio platform for personal broadcasting at home. VR@Home platform creates 3D scene with 2.5D background model, 3D object model, the segmented user, and other CG contents. 2.5D background and 3D objects generated from images of a real scene provide more immersion to users. The real-time soft shadows created by the estimated light sources keep the consistency between the real world and the virtual world. Through the personal virtual studio platform, a user can make his/her own contents and distribute or share them with the others on the Internet. Our platform is applicable to education in remote site, entertainment, and internet TV applications.

References

[1] http://www.podcast.net.

[2] O. Grau, T. Pullen and G.A. Thomas, “A combined studio production system for 3D capturing of live action and immersive actor feedback,” IEEE Transactions on Circuits and Systems for Video Technology. Volume 14, No. 3, March 2003.

[3] V. Lalioti and A. Woolard, “Mixed Reality Productions of the Future,” Proc. International Broadcasting Convention, Amsterdam, Netherlands, September 2003.

[4] Yuko Yamanouchi, Hideki Mitsumine, Takashi Fukaya, Masahiro Kawakita, Nobuyuki Yagi, Seiki Inoue, “Real space-based virtual studio seamless synthesis of a real set image with a virtual set image,” Proc. of VRST2002, pp. 194-200, 2002.

[5] Logitech Video Effects, http://www.logitech.com

[6] S. Kim and W. Woo, "Indoor Scene Reconstruction using a Projection-based Registration Technique of Multi-view Depth Images," 6th Pacific-Rim Conference on Multimedia, LNCS Vol. 3768, pp. 759-771, 2005.

[7] W. Lee and W. Woo, "Panoramic Mesh Model Generation from Multiple Range Data for Indoor Scene Reconstruction," 6th Pacific-Rim Conference on Multimedia, LNCS Vol. 3768, pp. 1004-1014, 2005.

[8] J.M. Hasenfratz, M. Lapierre, N. Holzschuch and F.X. Sillion, "A survey of Real-Time Soft Shadows Algorithms," Eurographics2003 State-of-the-Art Report, pp.753-774, 2003.

[9] S. Brabec, H.P. Seidel, "Shadow volumes on programmable graphics hardware," Computer Graphics Forum (Eurographics 2003), 25(3), 2003.

[10] ARToolkit, http://www.hitl.washington.edu/artoolkit/

[11] Digiclops, http://ptgrey.com/products/digiclops/index.asp