PANORAMIC IMMERSIVE VIDEOS
3D Production and Visualization Framework

W. J. Sarmiento and C. Quintero
Multimedia Research Group, School of Engineering, Militar Nueva Granada University, Bogota D.C., Colombia

Keywords: Panoramic videos, Immersive video, 3D production.

Abstract: Panoramic immersive video is a new technology, which allows the user to interact with a video beyond simple production line, because it enables the possibility to navigate in the scene from different points of view. Although many devices for the production of panoramic videos have been proposed, these are still expensive.In this paper a framework for production of virtual panoramic immersive videos using 3D production software is presented. The framework is composed by two stages: panoramic video production and immersive visualization. In the former stage the traditional 3D scene is taken as input and two outputs are generated, the panoramic video and path sounds to immersive audio reproduction. In the latter, a desktop CAVE assembly is proposed in order to provide an immersive display.

1 INTRODUCTION

Lately, immersive video technologies have become increasingly significant in the multimedia entertainment industry. One of them is the type in which the user has the power to control not only the time line but also his viewing position, this allows him to observe the scene from any point of view. Moreover, in higher immersion levels, the user can interact with the environment as if being present in it. These features offer the user a higher degree of immersion and interaction sensations. Panoramic video is the simplest version of immersive video. In which, videos are generated from scenes captured by a set of cameras, this allows to reconstruct the 360 degree video, but the only available points of view are those of the capturing cameras. Another important element of immersive sensation is sound, because one more of the user's senses is involved in the experience. To implement it, a multiple channel recording and a surround playing system are necessary.

The acquisition and displaying of panoramic videos present a great quantity of research challenges. At displaying level, many spatially immersive displays have been proposed (Gross et al., 2003; Katkere et al., 1997; Moezzi et al., 1996a; Neumann et al., 2000). Many of them are based on the Automatic Virtual Environment (CAVE) proposed by the Electronic Visualization Laboratory (EVL) of the University of Illinois in 1992 (Cruz-Neira et al., 1992; Cruz-Neira et al., 1993). It consists of a room in which video scenes are projected from behind the walls i.e. the user is surrounded by the projected images, giving the sensation of being immersed on the scene. At acquisition level, different applications are commercially available, like wide angle lenses or polycameras (Lin and Bajcsy, 2001; Nielseni, 2002; Nielsen, 2005), which capture different videos simultaneously from real scenes to assemble the panoramic video (Moezzi et al., 1996a). These systems are composed by a spatial configuration of cameras for each particular application. Some applications like Google Street View or immersivemedia demos are produced using these devices.

On the other hand, the implementation of 3D technology has allowed to reduce costs in audiovisual production. By using tools such as Maya, Softimage or 3D Studio, it is not necessary to built physical stages for production anymore, because they can be built and assembled in digital 3D post-production stage. Likewise, these tools allow to place virtual cameras on the scene, which have parameters that can be adjusted according to the needs of the producer and are usually in line with real cameras. Currently, these tools are low cost and accessible by even multimedia developers. Moreover, free software like Blender, Anim8or or DAZ 3D produce similar results and are available on the web.

Sarmiento W. and Quintero C. (2009).
PANORAMIC IMMERSIVE VIDEOS - 3D Production and Visualization Framework.
In *Proceedings of the International Conference on Signal Processing and Multimedia Applications*, pages 173-177
DOI: 10.5220/0002235901730177

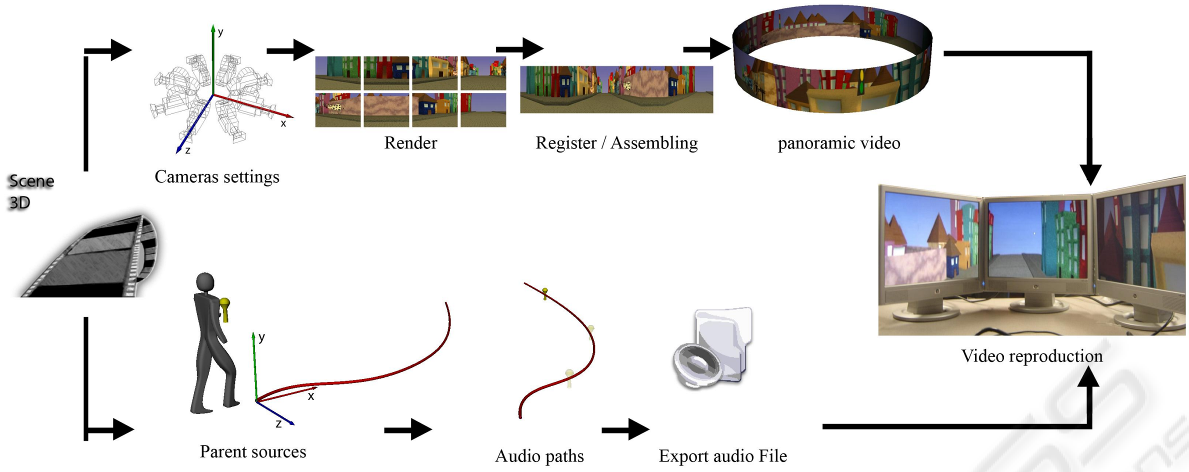


Figure 1: Schema of general architecture of proposed framework.

In order to make this applications closer to multimedia developers and final users, this paper presents a framework for production of virtual immersive panoramic videos, using only 3D production software. The second section introduces the proposed framework. In the third section details of the virtual panoramic video production are presented. Likewise, details of a low cost CAVE assembly, in which the panoramic video can be reproduced, are illustrated in section four. Finally, section five presents conclusions and future work.

2 PROPOSED FRAMEWORK

Figure 1 illustrates a scheme of the proposed framework. The initial point is a traditional 3D scene, created after pre-production stage. This 3D scene defines the background scenario, characters, objects and camera path. With all these parameters already set, the framework presents two independent branches, which may be implemented in parallel: video creation and audio path generation. The panoramic view production process starts with the building of the camera settings, followed by a rendering process which should be carried out individually for each camera and finally, a video registration process that stitches all rendered images together. On the other hand, audio path is generated parenting the sounds to the virtual scene, the 3D software creates the path where the virtual sound is placed and exported. Finally both panoramic video and audio paths are reproduced in a simplified desktop CAVE system.

3 VIDEO PRODUCTION

3.0.1 Building of Camera Setting

In order to obtain a panoramic view, this framework requires covering at least 360^o degrees in the horizontal line of vision. Therefore, N cameras are created with a different rotation value over the y axis from the main camera in the 3D scene. The number (N) of cameras depends on the horizontal aperture angle of the main camera. For example, a 35mm format camera has a horizontal angle of 49.3^o. Because of this, the implementation of 8 virtual cameras, with rotations of 45^o from the main camera, for covering the 360^o degrees, is needed. This assembly is similar to the hardware device used to capture real panoramic videos. Figure 2 shows an illustration of a camera setting corresponding to this example: a configuration of 8 cameras.

3.0.2 Rendering of Multiple Point of View Scenes

Before performing the rendering process it is required that each N_i camera is parented with the main camera, this way each of them has the camera path drawn by any traditional 3D production software. Then, N videos of the same scene are produced by rendering images generated by each camera separately. These videos correspond to the different points of view of the scene. Figure 3 shows an example of 8 images obtained by carrying out the rendering process on the camera set proposed in figure 2.

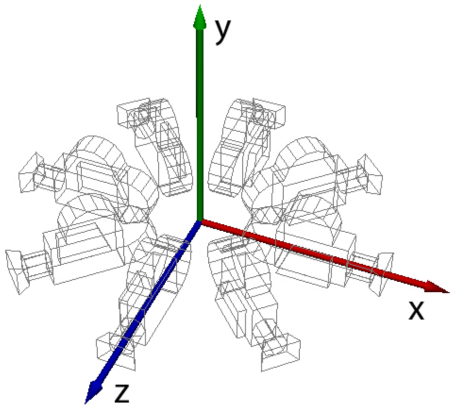


Figure 2: Camera configuration to production of panoramic 3D videos. Each camera is a rotated version of the principal camera.

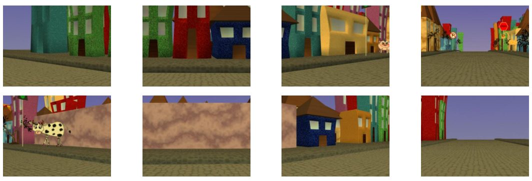


Figure 3: Scene render results obtained using the proposed set of cameras.

3.0.3 Registration of Rendered Image

Registration stage allows to stitch the scenes rendered in the previous stage in order to generate a single panoramic scene. The approach used in this stage depends on the camera features used to generate the scenes. If the rendering was performed using real camera features, resulting scenes are overlapped in the y axis. In this case a video registration approach should be applied (Shah and Kumar, 2003) . A simple registration approach which maximizes a similarity measure between adjacent scenes can be sufficient for stitching the videos in this case because the virtual cameras can be fixed on selected places in the scene. Figure 4 shows a portion of the panoramic scene generated by registering the images in figure 3. In this case a mutual information similarity measure was used, and minimization process was optimized assuming that at most a 10% overlapping exists. However, users not trained on image processing topics, can adjust the overlapping manually between images in a post-production software which can be batch propagated to all cameras. Moreover, if cameras are virtually placed on integer angle separations ($45^o, 30^o$) the rendering process is perfectly adjustable, simplifying the registration process.

Figure 4: Image obtained by registration of scenes at figure 3.

3.1 Audio Path Generation

3.1.1 Parent Audio Sources

In this framework, audio production requires that an audible atmosphere was defined in pre-production stage of the 3D scene. Likewise, a set of monophonic sounds should be defined for specific moments in the video. This can correspond to moving audible sources such as people steps or object movements. The sounds should be parented with their 3D object generator. For this, a graphical primitive (a sphere) should be created in the 3D production software which represents the sound in the scene. Any defined audio source should be parented with the corresponding object in the scene.

3.1.2 Export Sound Animation Curves

Once the audio primitives are parented with the objects on the 3D scene, these get the objects movement features i.e. they are moved according to developed animations. In this way, for any animated audio source in the scene a movement path is generated. Then, this can be exported to a text file indicating the sound position in each frame with respect to the position of the main camera.

4 IMMERSIVE PROJECTION

For panoramic video visualization an application that generates a cylindrical surface used as video projection canvas based on a texture mapping strategy is necessary. According to the video that is reproduced the texture is updated to a new video frame. Navigation is achieved by the rotation and scaling of a 3D volume view which is located in the center of the projection surface as illustrated in figure 5. This development requires tools for generating 3D graphics and audio. For this, libraries such as openGL or DirectX3D can be used.

Simultaneously immersive audio is reproduced, keeping in mind that the audio movement path for each source was built by the 3D production software

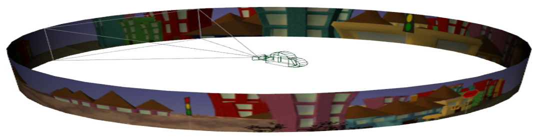
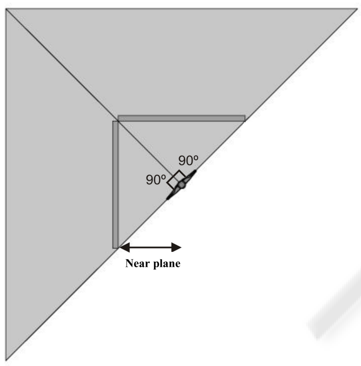
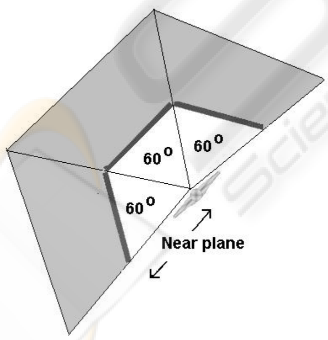


Figure 5: Visualization surface, the frame of video update the surface texture and the user may be change the camera zoom and rotation.

and exported to a file. Using the library for development of applications with 3D audio, each sound is placed in its spatial coordinate, which is changed when the video is in reproduction. Perception of surround sounds is achieved by the use of a 5.1 audio system with a sound board that codes the channels adequately. Systems with these features are commercially available and their costs are low.



(a)



(b)

Figure 6: Simplify CAVE system diagrams, (a) show a immersive volume view with 2 screen, an the (b) show a immersive volume view with 3 screen.

In order to fool the user, to provide a feeling of immersion in the virtual environment, a large display device is required that allows the user to navigate on a full vision field, as well as an enveloping audio reproduction. Immersion environment can be achieved displaying the panoramic video in any CAVE assembly (Quintero et al., 2007; McGinity et al., 2007; Moezzi et al., 1996b). However, in order to provide an affordable solution, we proposed the implementation of a desktop cave, which can be constructed using correctly placed flat panels. The only additional technical requirement is that the PC used for playing the video needs multiple independent video outputs (at least two) (Cerón et al., 2007; Jacobson et al., 2005; Quintero et al., 2007). Different immersive environments can be used according to the number of flat panels. Figure 6 shows the top view of two possible assemblies with 2 and 3 panels. The main parameters that should be set are the aperture angle and the near plane. Clearly, in a scheme all volume views have the same configuration parameters and coordinate origin, only the observation angle is changed. Based on the selected scheme a volume view is created for each designed panel using the 3D graphic library, keeping in mind that the video navigation should update the volume views simultaneously. Figure 7 shows two photographies of desktop CAVES with 2 and 3 panels displaying immersive videos generated using this framework.

5 CONCLUSIONS

In this paper a simple framework for producing and displaying virtual immersive videos has been proposed, which can be used without any specialized acquisition hardware device. On the contrary, this framework proposes the use of conventional functions on available 3D production software as 3D animation developing applications and 3D graphics and audio libraries. Video production is achieved by a sequence of simple implementation stages, which do not require advanced knowledge of mathematical issues. This fact allows this framework to be used by multimedia and art developers of any area. On the other hand, immersive videos are displayed on affordable equipments that allow immersive sensations similar to those achievable in complex visualization devices.

This framework can be used for developing and displaying immersive applications in restricted conditions of space and economical resources allowing a major access to these technologies to developing regions and research areas such as basic education and training environments.

(a)

(b)

Figure 7: Pictures to desktop CAVE system corresponding to schema with the figure 6, (a) 2 screens and (b)3 screens.

REFERENCES

Cerón, A. C., Sarmiento, W., and Sierra-Ballc, E. L. (2007). Computer graphics application for virtual robotics. In *23rd ISPE International Conference on CAD/CAM Robotics and Factories of the Future*, pages 574–577.

Cruz-Neira, C., Sandin, D. J., and DeFanti, T. A. (1993). Surround-screen projection-based virtual reality: the design and implementation of the cave. In *SIGGRAPH '93: Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pages 135–142, New York, NY, USA. ACM.

Cruz-Neira, C., Sandin, D. J., DeFanti, T. A., Kenyon, R. V., and Hart, J. C. (1992). The cave: audio visual experience automatic virtual environment. *Commun. ACM*, 35(6):64–72.

Gross, M., Würmlin, S., Naef, M., Lamboray, E., Spagno, C., Kunz, A., Koller-Meier, E., Svoboda, T., Van Gool, L., Lang, S., Strehlke, K., Moere, A. V., and Staadt, O. (2003). Blue-c: a spatially immersive display and 3d video portal for telepresence. *ACM Transactions on Graphics*, 22(3):819–827.

Jacobson, J., Kelley, M., Ellis, S., and Seethaller, L. (2005). Immersive displays for education using caveut. In *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2005*, pages 4525–4530. AACE.

Katkere, A., Moezzi1, S., Kuramura, D. Y., Kelly, P., and Jain, R. (1997). Towards video-based immersive environments. *Multimedia Systems*, 2(2):69–85.

Lin, S.-S. and Bajcsy, R. (2001). True single view point cone mirror omni-directional catadioptric system. In *Eighth IEEE International Conference on Computer Vision, ICCV 2001.*, volume 2, pages 7–14. IEEE.

McGinity, M., Shaw, J., Kuchelmeister, V., Hardjono, A., and Favero, D. D. (2007). Avie: A versatile multi-user stereo 360 interactive vr theatre. In *Proceedings or Emerging Display Technologies Conference*.

Moezzi, S., Katkere, A., Kuramura, D., and Jain, R. (1996a). Immersive video. In *Virtual Reality Annual International Symposium*, pages 17 – 24, 265. SPIE.

Moezzi, S., Katkere, A., and Kuramura, D.Y. Jain, R. (1996b). An emerging medium: interactive three-dimensional digital video. In *Proceedings of the Third IEEE International Conference on Multimedia Computing and Systems*, pages 358 – 361. IEEE.

Neumann, U., Pintaric, T., and Rizzo, A. (2000). Immersive panoramic video. In *MULTIMEDIA '00: Proceedings of the eighth ACM international conference on Multimedia*, pages 493–494, New York, NY, USA. ACM.

Nielsen, F. (2005). Surround video: a multihead camera approach. *The Visual Computer*, 21(1-2):92–103.

Nielseni, F. (2002). High resolution full spherical videos. In *International Conference on Information Technology: Coding and Computing, 2002. Proceedings.*, pages 260 – 267. IEEE.

Quintero, C. D., Sarmiento, W. J., and Sierra-Ballén, E. L. (2007). Low cost cave simplified system. In *Proceedings of Human Computer Interaction International, HCI International 2007*, pages 860 – 864.

Shah, M. and Kumar, R. (2003). *Video Registration*. Kluwer Academic Publishers.