Interactive Virtual Simulation for Multiple Camera Placement

Jeff Williams
jwilliam@site.uottawa.ca
University of Ottawa

Won-Sook Lee
wslee@uottawa.ca
University of Ottawa

1 Keywords

Object modelling, virtual environments, simulation, 3D reconstruction, multiple cameras, mutual viewable volume.

Abstract

Camera placement plays an integral role in image-based 3D object reconstruction. Poor camera placement can lead to poor reconstruction if important silhouette or texture details are not captured in one view, if a camera is too close and the object silhouette is clipped, or if a camera is too far to provide a high-resolution capture. Some of these constraints are in contention. Manually adjusting real world cameras to satisfy these criteria is difficult and tedious as cameras may have to be unmounted and moved, their output checked for fitness of view, then readjusted. An interactive virtual camera positioning utility is presented that helps the reconstructor quickly and easily choose intuitively good positions, with emphasis on preventing clipping and resolution loss by visualizing the camera-set's mutual viewable volume.

2 Introduction

The quality of image-based 3D object reconstruction depends on camera placement. Many commonly used image-based 3D reconstruction techniques require an apparatus involving multiple cameras fixed about an area in which the object or person to be reconstructed is located. For volume carving and volume intersection techniques, an accurate construction of a visual hull from a set of images depends not only on the quantity of cameras, but of the fitness of the camera positioning.

There is a number of factors that contribute to a good camera position and orientation. Since 3D reconstruction techniques often build shape using silhouettes, the most shape information is revealed when viewing from a direction *perpendicular* to the normal of the points in the object's silhouette perimeter. However, the most texture information is measured when the viewing direction is *parallel* to the object's surface normal, as this limits aliasing due to foreshortening.

Further, camera position affects reconstruction through resolution and clipping. Poor resolution occurs when the object is *too far* from a viewing digital camera and varying texture details become blended into the same pixel of the finite resolution camera. On the other hand, clipping can occur when the object is *too close* to the camera, and the silhouette of the object to be reconstructed is not fully captured in the image. Since many image-based reconstruction methods proceed by finding the intersection of the frustum produced by projecting the 2D object silhouette in a photograph into 3D, if a silhouette is clipped in even one photograph, the reconstructed object is similarly truncated.

To satisfy these constraints one requires the use of virtual camera positioning. By allowing the user to position the cameras in a virtual environment quickly, gracefully, and efficiently, much monotonous work can be avoided. Our contribution is twofold. First, we provide a virtual camera simulation that allows the user to easily position cameras to intuitively capture the most shape and texture detail from an object. Second, our system computes a polyhedron which visualizes the *mutual viewable volume* of a set of cameras, which is defined as the intersection of all camera's view frustums. Depicting this volume allows the user to clearly detect if the object being reconstructed will be truncated. This allows the user to position the cameras so that the correct bounding volume for the object to be reconstructed is chosen. For static scenes, the cameras can be positioned for the highest resolution possible without clipping the object silhouette, and arranged to gather the most important texture and shape detail. For dynamic scenes such as motion capture, the cameras can be positioned to allow a larger volume that will capture the entire range of motion.

The benefits of our system can be realized by many existing applications where multiple cameras are used for reconstruction. The setup time of outdoor reconstruction systems in [3] and [2] can be significantly sped up, and telepresence systems such as Blue-C, 3D Live, or users of Zaxel systems would obtain better results [4][5][6].

The system is currently implemented to operate using calibration data generated by the Virtual Viewpoint (VV) 3D reconstruction system offered by Zaxel Systems, Inc. The VV system consists of a variable number of cameras, a processing workstation for every three cameras, and proprietary software to perform 3D reconstruction. The system we used for experiments consists of twelve Firefly CCD digital video cameras manufactured by Point Grey Research. The VV system implements a derivative of the image-based visual hull 3D reconstruction technique, and as such is able to produce 3D video from a new viewpoint using video feeds from multiple cameras. The separate video feeds are first captured and stored uncompressed on the workstations connected to the camera, then the feeds are processed offline to generate the reconstruction. For our experiments, the VV cameras are mounted at various heights on the walls of a small room having four walls. The room's walls and floor are painted a constant green colour so that the VV system can use chroma keying to aid in the background subtraction processing.

On the topic of positioning virtual camera modelling, some previous results exist. A number of papers deal with cinematic navigation path calculation, in order to find the optimal camera placement for scene observation at a given time. In [7] the authors develop a method to optimize placement so that resolution of camera paths is highest, with resolution constrained in terms of effect of distance and foreshortening. In [8] a semi-automated method for controlling camera placement based on user input of cinematography idioms relating to camera position is presented. Virtual Director is a system for intuitively choosing camera positions over time in a virtual reality environment [9]. However, these do not deal with the problems of camera positioning for scene reconstruction.

There is much research into virtual camera control. In [10], Ware and Osborne explore various ways for a user to control and position a virtual camera, and attempt to determine in which situations each control method is appropriate. In [11] the authors present a method for allowing a user to control multiple cameras simultaneously. In [12], the authors aim to improve the realism by adding classical physics constraints to camera motion.

There is much work related to camera placement to limit occlusion and resolution loss. In [13] the authors implement a system to model virtual cameras with varying calibration parameters for outputting scenes. In [14] the authors describe algorithmic approaches to determine camera positions that limit occlusion and image degradation. In [15] and [16], the authors optimize the camera position based on field of view, visibility, depth of field, and minimizing occlusion. However, none deal with reconstruction truncation.

In [17], State et al. have developed an interactive system for the placement of virtual cameras, dubbed *Pandora*. The system was created to intuitively optimize the positions of cameras used to obtain 3D video recordings of surgical procedures. Pandora allows the user to position the virtual cameras and observe the effect of their choices. To show the effect of each camera's placement, the camera's rectangular viewing volume is projected into the scene like a movie projector emitting light, and is drawn onto objects in the scene. The effect of different cameras is distinguished by using different colours for each camera's projection. However, they don't explicitly deal with object truncation as our system does.

3 Proposed Approach

Our virtual camera modelling system provides an interactive virtual environment which models the real world VV apparatus. Each camera is modelled in the virtual environment with the position, orientation, and focal length of the real camera. The walls of the VV room are drawn in wireframe for reference as shown in Fig. 2. The user can import a model of an object into the scene. The tool also generates and displays the current mutual viewable volume for the cameras in realtime. Using the mouse, the scene can be rotated in two degrees of freedom about the centroid of the cameras, the view of the scene can be tilted and panned, translated, and zoomed.

By clicking a camera in the scene and dragging, the user can reposition cameras with respect to the wall they are attached to, for the purpose of examining the effect of placing a camera at a different location on the wall. In the real world, camera lenses are not flush with the wall that a camera is mounted on. The camera body and mounting system push the camera some distance away. In our virtual system, this is modelled by determining which wall is closest to the virtual camera and the distance from the camera to the wall. When the user drags a camera it moves parallel to the wall and maintains its distance from the wall. The user may also arbitrarily position the camera by switching to its viewpoint and moving or rotating the scene; the camera will translate and rotate so that the observed viewpoint is the camera's new viewpoint.

The system provides two key methods for qualitatively determining the fitness of a particular camera's placement. First, the user may switch to any camera's viewpoint to observe through its lens. Second, the cameras are each coloured differently, and facets of the mutual viewable volume polyhedron which are bounded by a particular camera's frustum are coloured the same as the camera. This allows a global view of which cameras are limiting the mutual viewable volume. Cameras shown to be placed poorly can be dragged into better position while the mutual viewable volume polyhedron is updated in real-time.

Since our system uses the Zaxel VV system for the experiment, our virtual system contains the same number of cameras as the real system, and processes views the same way.

The system can calculate a transformation that takes real physical coordinates to virtual coordinates given four special real-world coordinates.

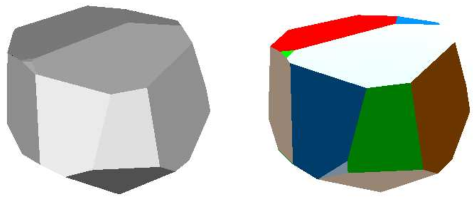
The virtual camera modelling application is built using OpenSceneGraph [18] and uses calibration data calculated by the VV system.

3.1 Results

The system allows the user to visualize the mutual viewable volume (MVV) of all the cameras. The MVV defines a volume which is the total space available in which to place objects to reconstruct. The purpose of this volume is to allow the user to quickly decide if an object or space is too large to be reconstructed using the current camera configuration, or if the volume is overly large and so the cameras can be moved closer for a higher fidelity reconstruction.

The faces of a polyhedron representing the mutual viewable volume are calculated by intersecting the viewing frustums of all the cameras. Since each camera produces a rectangular image we can consider the frustum of each camera to be a rectangular pyramid. If we limit the height of this pyramid, then the boundary of each frustum is composed of four triangular polygonal segments of planes (faces), each with a normal pointing outwards from the cameras viewable volume. By intersecting each of the faces from a frustum with the faces of other frustums, and only keeping the portions of the faces being cut that are towards the inside of the cutting faces, we can produce a closed set of faces that are the boundary of the mutual viewable volume.

To prove that the computed MVV is correct, we generated a model of the MVV using the Virtual



(a) (b)

Figure 1: The MVV generated experimentally is shown in (a) while the MVV calculated by our system is shown in (b). The facets on the boundary of (b) may appear smaller since different viewing focal lengths were used between VV and the virtual system.

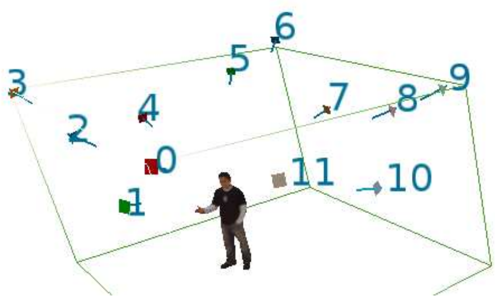


Figure 2: View of camera arrangement. Virtual cameras are coloured pyramids with lines representing viewing direction, labelled with the VV camera number. The walls of the VV room are modelled as a wireframe.

Viewpoint system itself. The VV system includes a tool that can export a full VRML 3D model from one frame of the captured video. By calibrating the system with the lights on, and then capturing frames of video in the dark, each camera is fooled into believing that it's entire frustum is foreground. When reconstruction is performed after this procedure, it generates an effective MVV. The results are shown in Fig. 1 (a). The calculated MVV in the virtual environment shown in Fig. 1(b) is very close to the experimentally generated MVV in Fig 2(a). The area of some faces in (b) are over or under estimated, and the virtual MVV has a few small extra faces.

The presented figures demonstrate the virtual camera application. In Fig. 2 the user is viewing the scene from no virtual camera. The 12 cameras are each labelled with an index, with cameras labelled 0-2 on one wall, 3-5 on an adjacent wall, 6-8 on another,

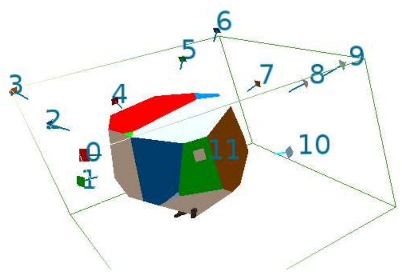


Figure 3: The mutual viewable volume of all cameras calculated using the above method is displayed. The facets are coloured the same as the camera that limits the MVV to create that facet.

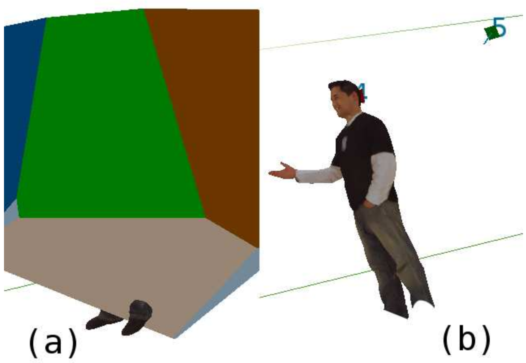


Figure 4: On the left (a) the MVV demonstrates that the camera positions will cause reconstruction to be truncated. On the right (b) is the maximum portion of the figure that can be reconstructed.

and 9-11 on the last. A model is loaded into the scene for determining fitness of each camera position. Figure 3 presents a similar view, except that the mutual viewable volume is shown as a multicoloured polyhedron, and depicts the volume within which objects must be contained to avoid clipping objects and creating an incorrect reconstruction. By placing a model of an appropriate size into the scene, it can be easily seen if the volume is large enough for reconstruction.

The MVV polyhedron construction takes less than a second to calculate and render on a 600MHz Athlon system.

Figures 4, 5, and 6 demonstrates the utility of the mutual viewable volume visualization. We see in Fig. 4 that two cameras are positioned poorly such that 3D reconstruction will not reconstruct the bottoms of the figure's legs, causing the reconstructed object to be truncated. We see that the cameras causing the problem are the two with colours corresponding to

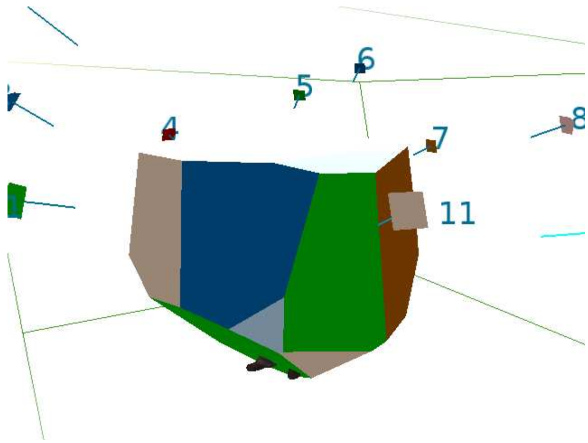


Figure 5: By carefully moving camera 11 lower parallel to the wall it's mounted on, we can partially fix the problem.

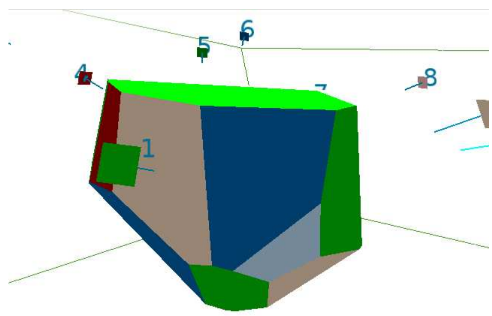


Figure 6: After moving camera 1 lower, the problem is solved.

the lower polyhedron faces of the MVV. By moving those cameras lower, as done in Fig. 5 and Fig. 6 this problem is solved. Since each camera was moved along a wall in the virtual system, it is easy to mimic this movement in the real world by moving the corresponding physical camera along the same wall.

4 Discussion

This paper focusses on virtual camera modelling, and mutual viewable volume calculation and visualization.

An automatic optimization method can be added that adjusts the camera positions so that an object is viewed better from individual cameras. That is, if the object is not centered in the camera, pan or tilt the camera so it is, or if the object is very small in the camera, move the camera forward, etc. This optimization can be kept basic, or made robust by optimizing the camera position based on the virtual 3D reconstruction of a model. The goal would be to minimize the error between the reconstructed model and the original model.

Although outputting the real-world position of camera lenses is possible using the change of basis from real to virtual coordinates described above, a method for adjusting the real camera pan and tilt is not implemented. This can be achieved by having the virtual system output a convenient point, such as on the floor or another wall, that should be in the center of the camera's field of view, and rotating the camera accordingly.

Simple positioning convenience features can be implemented to increase ease of positioning, such as automatically rotating all cameras to view the same user-defined point, and allowing the user to move the point so that the mutual viewable volume is positioned at the right location for reconstruction.

The MVV, while used mainly for truncation detection, might also be useful to measure how well the object shape information is being reconstructed. A face on the MVV that has a large area can mean that only the one camera causing the face is generating shape information for the region tangent to that face, and so other cameras should be moved to break the large area face into smaller areas. The cameras to be moved should be those that generate only small faces in the MVV. The ideal camera configuration that results in the best global shape reconstruction given a limited number of cameras could be that where the area of the faces of the MVV are most uniform. Additional constraints such as limited choice for camera position make this problem interesting.

5 Conclusion

In this paper, we have presented an interactive virtual system that can be used to calibrate cameras in the real world. Real-world calibration information from the Zaxel VV system is imported accurately to create a multiple camera virtual environment. Virtual cameras can be moved and rotated, including constraining their movement to be parallel to a wall's surface, and their viewpoints observed from with the correct focal length and principle point. Object models can be used to estimate fitness of view, and the system calculates and visualizes the mutual viewable volume which marks the boundary of the usable volume for shape-from-silhouette methods.

We have shown that camera positioning and orientation affects the quality of silhouette based 3D reconstruction. Multiple conflicting constraints make this problem difficult to solve; there is no simple configuration that is ideal in all cases. A method of determining the best camera positions is necessary. Manual camera placement is not an answer as it is tedious and error prone. By using a virtual environment to position cameras intuitively with extra cues not visible in the real world, such as the mutual viewable volume, this burden is lifted. This system can be used to save time and effort, and to produce higher quality output in reconstructed objects.

6 Acknowledgements

Thanks to François Malric for his help with configuration, calibration, and various other tasks regarding the Zaxel VV equipment.

References

[1] Mundermann, L., Corazza, S., Chaudhari, A.M., Alexander, E.J., Andriacchi, T.P.: Most favorable camera configuration for a shape-from-silhouette markerless motion capture system for biomechanical analysis. Volume 5665., San Jose, CA, USA, SPIE (2005) 278–287

[2] Debevec, P., Taylor, C., Malik, J.: Modeling and rendering architecture from photographs: A hybrid geometry and image-based approach. In: Proceedings of SIGGRAPH 1996. (1996)

[3] El-Hakim, S., Beraldin, J.: Detailed 3d reconstruction of monuments using multiple techniques. In: Proceedings of the International Workshop on Scanning for Cultural Heritage Recording - Complementing or Replacing Photogrammetry. (2002) 13–18 NRC 44915.

[4] Gross, M., Wurmlin, S., Naef, M., Lamboray, E., Spagno, C., Kunz, A., Koller-Meier, E., Svoboda, T., Gool, L.V., Lang, S., Strehlke, K., Moere, A.V., Staadt, O.: blue-c: A spatially immersive display and 3d video portal for telepresence. In: Proceedings of SIGGRAPH 2003. (2003)

[5] Prince, S., Cheok, A.D., Farbiz, F., Williamson, T., Johnson, N., Billinghurst, M., Kato, H.: 3d live: Real time captured content for mixed reality. In: Proceedings of 2002 IEEE / ACM International Symposium on Mixed and Augmented Reality, IEEE Computer Society (2002)

[6] Inc., Z.S.: Us patent no. 6864903 (2003) Internet system for virtual telepresence.

[7] Bodor, R., Schrater, P.R., Papanikolopoulos, N.: Multi-camera positioning to optimize task observability (2005) Computer Science and Engineering Technical Report 05-009.

[8] Amerson, D., Kime, S.: Real-time cinematic camera control for interactive narratives (2000)

[9] : (Virtual director) http://virdir.ncsa.uiuc.edu/virdir/.

[10] Ware, C., Osborne, S.: Exploration and virtual camera control in virtual three dimensional environments. In: SI3D '90: Proceedings of the 1990 symposium on Interactive 3D graphics, New York, NY, USA, ACM Press (1990) 175–183

[11] Fukatsu, S., Kitamura, Y., Kishino, F.: Manipulation of viewpoints in 3d environment using interlocked motion of coordinate pairs. In: INTERACT. (2003)

[12] Turner, R., Balaguer, J.F., Gobbetti, E., Thalmann, D.: Physically-based interactive camera motion control using 3D input devices. In Patrikalakis, N.M., ed.: Scientific Visualization of Physical Phenomena: Proceedings of CG International Tokyo, Conference held in Tokyo, Japan, Springer-Verlag Inc. (1991) 135–145

[13] Trujillo-Romero, F., Ayala-Ramirez, V., Sanchez-Yanez, R., Ibarra-Manzano, O.: A virtual active camera for scene modelling in robotic vision tasks (2004)

[14] Xing, C., Davis, J.: Camera placement considering occlusion for robust motion capture (2000)

[15] Tarabanis, K.A., Tsai, R.Y., Kaul, A.: Computing occlusion-free viewpoints. In: IEEE Trans. Pat. Anal and Mach. Intel. Volume 18(3). (1996) 279–292

[16] Sujan, V.: Task directed imaging in unstructured environments by cooperating robots. In: submitted to the 2002 Third Indian Conference on Computer Vision, Graphics and Image Processing, Ahmedabad, India (2002)

[17] State, A., Welch, G., Ilie, A.: An interactive camera placement and visibility simulator for image-based vr applications. In: Proc. of the Engineering Reality of Virtual Reality 2006. (2006)

[18] Burns, D., Osfield, R.: Open scene graph a: Introduction, b: Examples and applications. In: VR, IEEE Computer Society (2004) 265

[19] Charlton, E.: An Octree Solution to Conservation-laws over Arbitrary Regions (OSCAR) with Applications to Aircraft Aerodynamics. PhD thesis, University of Michigan (1997) Section 'Computing Cut-Faces'.