Modelling from Multiple Photographs Using Points and Silhouettes

Tim Hawkins

University of Southern California Institute for Creative Technologies

We present a technique for recovering a 3D model from several photographs of an object. The technique makes use of a relatively small number of user-marked image correspondences as well as user-marked silhouettes. The recovered model is well-suited for rendering with view interpolation using the original photographs as textures. A view interpolation render using this technique is featured in the SIGGRAPH 2004 Electronic Theater piece *The Parthenon*.

Figure 1: Four photographs of a sculpture.

We first recover the positions of the cameras. There are a number of established techniques for doing this. In our case we made use of a known shape, a cube, in the photographs. Marking the projections of the corners of this cube in the photographs gives a fairly accurate position for each camera. We refine the camera positions using additional marked 2D points and bundle adjustment.

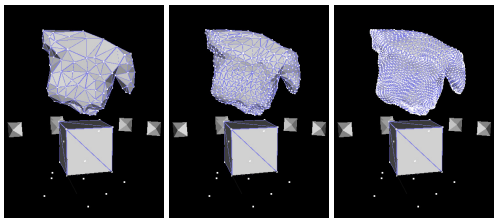

Figure 2: The original coarse mesh and two levels of subdivision.

Having recovered the camera positions, we create a very low-resolution 3D model of the sculpture by marking 2D correspondences in multiple photographs. The resulting 3D points are triangulated to form a mesh. The resulting mesh is shown on the left in Figure 2.

The coarse mesh can be used to produce view interpolated renders, but the resulting renders will have very crude silhouettes due to the coarseness of the mesh, and there will also be significant fading artifacts due to texture misalignment, particularly near the silhouette boundaries. To deal with this, we refine the coarse mesh and align it to the silhouettes of the sculpture as marked in each of the original photographs. The final high-resolution mesh is produced by subdividing the coarse mesh and aligning the subdivided mesh's projected silhouettes simultaneously to the marked silhouettes of the sculpture in all of the photographs.

We would like our refined mesh to interpolate the 3D points of the coarse mesh. We apply a simple correction to Loop subdivision to force the subdivided mesh to interpolate.

Our silhouette alignment operates on a twice subdivided mesh, shown on the right in Figure 2. We progressively deform the subdivided mesh to align it with the marked silhouettes using the following relaxation algorithm:

Repeat until close enough:

1. Initialize the total 3D step for each vertex to zero
2. For each camera
   (a) Compute the set of mesh edges which are currently on the mesh's silhouette as seen from the current camera.
   (b) For each vertex belonging to at least one silhouette edge
      i. Find the point p on the marked silhouette which is closest to the image projection of the vertex.
      ii. Compute a small 3D step which moves the vertex closer to p while maintaining the same distance from the camera.
      iii. Add the 3D step to the vertex's total 3D step.
3. For all vertices that belong to no silhouette edge, set its 3D step to be a weighted average of its neighbors, using a weighting that decreases linearly with distance. This assures that the mesh stays smooth.
4. Apply the 3D steps to all vertices.

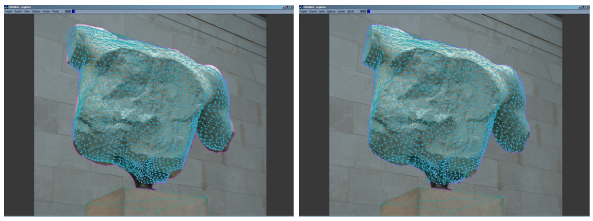

Figure 3: The mesh before and after silhouette aligment.

Finally, we render the model using a simple form of view-dependent texture mapping. As a virtual camera moves along a path which interpolates the original camera positions, we texture the model projectively by blending the photographs corresponding to the two closest cameras. We use a simple z-buffer technique to assure that surfaces which are hidden in one of the original camera views are not textured using that camera.

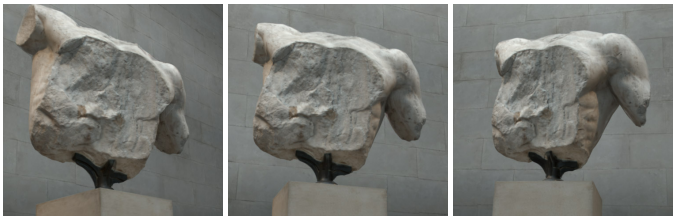

Figure 4: Three intermediate views.

References

MATUSIK, W., BUEHLER, C., RASKAR, R., GORTLER, S. J., AND MCMILLAN, L. 2000. Image-based visual hulls. In *Proc. SIGGRAPH 2000*, 369–374.

SZELISKI, R. 1993. Rapid octree construction from image sequences. *CVGIP: Image Understanding 58*, 1 (July), 23–32.

WILLIAMS, L., AND CHEN, E. 1993. View interpolation for image synthesis. In *SIGGRAPH 93*.