4D View Synthesis: Navigating through Time and Space (sap_0551)

Mingxuan Sun* Georgia Inst. of Tech.
Grant Schindler† Georgia Inst. of Tech.
Sing Bing Kang‡ Microsoft Research
Frank Dellaert§ Georgia Inst. of Tech.

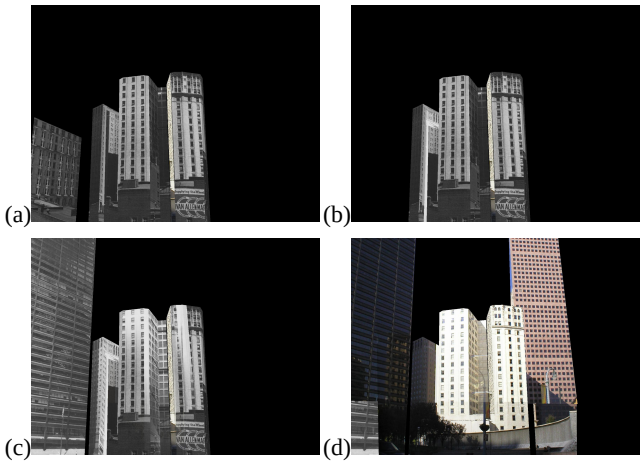

Figure 1: Building disappearing and appearing: time-lapse at a synthesized view based on time-dependent geometry.

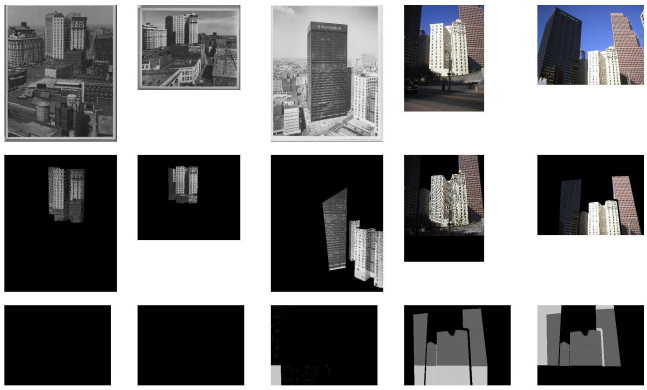

Figure 2: First row: input ordered images. Second row: visibility test through time and space. Third row: weighted mask on temporal distance warped to final view (brighter pixels have higher weights). The result of blending all textures weighted based on time is show in Figure 1(d).

In this sketch, we present a 4D view synthesis technique for rendering large-scale 3D structures evolving in time, given a sparse sample of historical images. We built a system to visualize urban structure that is a function of the time selected, thereby allowing virtual navigation in space and time. While there is a rich literature on image-based rendering of static 3D environments, e.g., the Facade system [Debevec et al. 1996] and Photo Tourism [Snavely et al. 2006], little has been done to address the temporal aspect (e.g., occlusion due to temporal change). We construct *time-dependent geometry* to handle the sparse sampling. To render, we use time-and-view-dependent texture mapping and reason about visibility both in time and space. Figure 1 shows the result of view synthesis based on time-dependent geometry.

We represent time-dependent geometry as a set of independent 3D models, each of which is associated with a time tag indicating its existence period. The 3D models can be reconstructed using a structure from motion approach and the time tag can be inferred from images or with the method of [Schindler et al. 2007]. Each image is associated with a virtual camera and a time tag.

With reconstructed time-dependent geometry, we obtain the set of 3D structures which exist at a given time. Time-and-view-dependent texture mapping requires the extraction of texture from appropriate parts of source images for synthesis. This is achieved by projecting 3D structure and testing for occlusions. Note that geometry is visible in an image if and only if it exists at the time which the image is taken and is not occluded by any other model at that time. A weight mask is computed for each given image to determine how much each texture contributes to the synthesized view. Figure 2 shows the process to synthesize an image at a given time and viewpoint. The submitted video (part of the supplementary material) shows a representative virtual navigation session.

Future research will focus on fully automating the geometry recovery process and obtaining more accurate geometry.

*e-mail: cynthia@cc.gatech.edu
†e-mail: schindler@cc.gatech.edu
‡e-mail: SingBing.Kang@microsoft.com
§e-mail: dellaert@cc.gatech.edu

References

DEBEVEC, P. E., TAYLOR, C. J., AND MALIK, J. 1996. Modeling and rendering architecture from photographs: A hybrid geometry- and image-based approach. In *Proceedings of SIGGRAPH 2000*, ACM Press / ACM SIGGRAPH, vol. 30, ACM, 11–20.

SCHINDLER, G., DELLAERT, F., AND KANG, S. B. 2007. Inferring temporal order of images from 3D structure. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, IEEE.

SNAVELY, N., SEITZ, S., AND SZELISKI, R. 2006. Photo tourism: Exploring photo collections in 3D. In *Proceedings of SIGGRAPH 2000*, ACM Press / ACM SIGGRAPH, ACM.