

Virtual Video Camera: Image-Based Viewpoint Navigation Through Space and Time

Christian Lipski, Christian Linz, Kai Berger, Marcus Magnor*
Computer Graphics Lab, TU Braunschweig

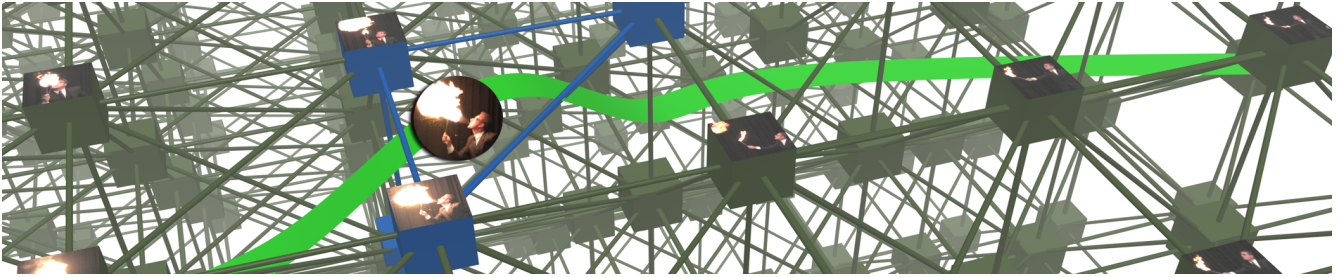


Figure 1: Viewpoint navigation space: time and camera directions span our navigation space. Each cube represents one video frame. The navigation space is partitioned into tetrahedra with video frames as vertices. Tetrahedral edges denote correspondence fields between video frames. Our virtual video camera view is interpolated by warping and compositing the four video frames of the enclosing tetrahedron (blue) in real-time.

Abstract

We present an image-based rendering system to viewpoint-navigate through space and time of complex real-world, dynamic scenes. Our approach accepts unsynchronized, uncalibrated multi-video footage as input. Inexpensive, consumer-grade camcorders suffice to acquire arbitrary scenes, e.g., in the outdoors, without elaborate recording setup procedures. Instead of scene depth estimation, layer segmentation, or 3D reconstruction, our approach is based on dense image correspondences, treating view interpolation uniformly in space and time: spatial viewpoint navigation, slow motion, and freeze-and-rotate effects can all be created in the same fashion. Acquisition simplification, generalization to difficult scenes, and space-time symmetric interpolation amount to a widely applicable Virtual Video Camera system.

Keywords: image-based rendering, video-based rendering, free-viewpoint video, viewpoint navigation

1 Motivation

The objective common to all free-viewpoint navigation systems is to render photo-realistic vistas of real-world, dynamic scenes from arbitrary perspective, given a number of simultaneously recorded video streams. Most systems exploit epipolar geometry based on either dense depth/disparity maps or complete geometry models, relying on calibrated and synchronized multi-video data. The cost, time and effort involved in recording synchronized multi-video data constitutes a major obstacle towards economically viable applications for free-viewpoint navigation. Further on, most existing systems are designed to interpolate virtual camera positions only along spatial dimensions. Temporal view interpolation requires additional scene motion information. Lacking this, viewpoint navigation systems can provide scene views only for discrete moments in time, and the scene cannot be viewed in slow-motion.

With our system, we address these limitations and propose a purely image-based approach to free-viewpoint navigation through space as well as time that accepts unsynchronized, uncalibrated multi-video footage as input.

2 Approach

Our approach is motivated by the pioneering work on view interpolation by Chen and Williams [1993]. The Virtual Video Camera systems picks up on their idea to interpolate different image acquisition attributes in a higher-dimensional space and suitably extends it to be applicable to view interpolation in the spatial as well as temporal domain. Putting the temporal dimension on a par with the spatial dimension allows for a uniform framework to continuously interpolate virtual video camera positions across space and time.

On the technical side, in our system dense image correspondences take the place of depth/disparity or 3D geometry, extending applicability to scenes whose object surfaces are highly variable in appearance, or hard to reconstruct for other reasons. Our processing pipeline makes use of known techniques and suitably adapts them to solve the problem at hand: The multi-video data is first color corrected [Snavely et al. 2006], and extrinsic camera parameters [Goesele et al. 2007] and inter-camera time offsets are estimated [Shrsth et al. 2007]. The video frames are then embedded into navigation space, adjacency of video frames is determined via constrained tetrahedralization, and dense morphing correspondences are estimated between adjacent video frames [Stich et al. 2008]. After these offline processing steps, the navigation space can be interactively explored (viewing directions and time) by real-time rendering [Mark et al. 1997].

References

- CHEN, S. E., AND WILLIAMS, L. 1993. View interpolation for image synthesis. In *Proc. of ACM SIGGRAPH'93*, ACM Press/ACM SIGGRAPH, New York, 279–288.
- GOESELE, M., SNAVELY, N., CURLESS, B., HOPPE, H., AND SEITZ, S. 2007. Multi-view stereo for community photo collections. *ICCV'07*, 1–8.
- MARK, W., McMILLAN, L., AND BISHOP, G. 1997. Post-Rendering 3D Warping. In *Proc. of Symposium on Interactive 3D Graphics*, 7–16.
- SHRSTHA, P., BARBIERI, M., AND WEDA, H. 2007. Synchronization of multi-camera video recordings based on audio. In *Proc. of MULTIMEDIA'07*, ACM Press, New York, 545–548.
- SNAVELY, N., SEITZ, S. M., AND SZELISKI, R. 2006. Photo tourism: Exploring photo collections in 3d. *ACM Trans. on Graphics* 25, 3, 835–846.
- STICH, T., LINZ, C., ALBUQUERQUE, G., AND MAGNOR, M. 2008. View and Time Interpolation in Image Space. *Computer Graphics Forum (Proc. of PG'08)* 27, 7.

*{lipski,linz,berger,magnor}@cg.cs.tu-bs.de