ICARUS: Interactive Reconstruction from Uncalibrated Image Sequences

Simon Gibson, Jon Cook, Toby Howard and Roger Hubbold*
Advanced Interfaces Group,
Department of Computer Science,
University of Manchester, UK

1 Introduction

The ICARUS system is a suite of software packages, developed at the University of Manchester, that allows geometric models to be quickly and easily reconstructed from image and video sequences captured with uncalibrated digital cameras. The system combines automatic and semi-automatic camera calibration algorithms with an easy-to-use interactive model-building phase (Figure 1). Surface textures are automatically extracted from images and mapped onto the reconstructed models.

Previous algorithms for model reconstruction from images have taken two different approaches. Fully automatic algorithms (e.g. [Pollefeys et al. 1999]) use a large number of closely-spaced images to reconstruct a dense point-cloud and highly detailed triangular mesh. Approaches such as this have the benefit of ease of use, but the models they produce lack structure and contain missing regions caused by object occlusion. By way of contrast, semi-automatic approaches (e.g. [Debevec and Malik 1996; Hakim 2000]) rely on user assistance to calibrate the cameras and build scene structure, but have the benefit of being able to incrementally construct an object hierarchy. They also take advantage of user knowledge when reconstructing the environment, thereby overcoming problems caused by object occlusion. The main disadvantage of semi-automatic systems is the amount of user interaction required to identify enough common features between the images so that calibration may be accurately performed.

2 The ICARUS System

The approach taken by the ICARUS system falls between these automatic and semi-automatic extremes, and its robustness and ease of use comes from taking advantage of the strengths of both types of algorithm. The main source of input data for the ICARUS system is video sequences captured with a digital video camera. We combine automated feature selection and tracking with robust hierarchical structure-from-motion and novel self-calibration algorithms to automatically estimate the camera parameters for each frame of the sequence. This calibration data is then used to assist the user in interactively reconstructing a model of the environment. The new calibration algorithm has been found to be robust for a wide variety of different camera motions and sequence lengths, and has also been used for augmented video-production. Reconstruction is achieved by manipulating the position, orientation, and size of primitive objects such as polygons, boxes, cylinders etc. so that their projections into the sequence match the projections of real objects. Manipulation is aided by a variety of user-specified constraints, ranging from strictly-enforced hierarchical parent-child relationships to image-based constraints on the location of object vertices projected into each frame. Image-based constraints are evaluated using the information obtained during the calibration stage. We employ a non-linear optimization algorithm that is capable of manipulating the parameters of these objects in real-time, so that the constraints specified by the user are best satisfied.

*e-mail: {sg|cookj|toby|roger}@cs.man.ac.uk

This novel approach of combining automatic calibration of video sequences with interactive geometry reconstruction allows the user to spend more time modeling important features of the scene, rather than preparing the system for calibration. The system is not limited to video sequences, however, since we also take advantage of alternative calibration techniques for video sequences captured with tripod mounted cameras, as well as for small collections of photographs, and even single images. This gives the user maximum flexibility in terms of how the input image data can be captured, without the need to employ different reconstruction algorithms for each type of input. Example reconstructions from a variety of sources are shown in Figure 1. Typical reconstruction times for each model were less than 30 minutes. Further examples, and an evaluation copy of the software, may be downloaded from *http://aig.cs.man.ac.uk/icarus*

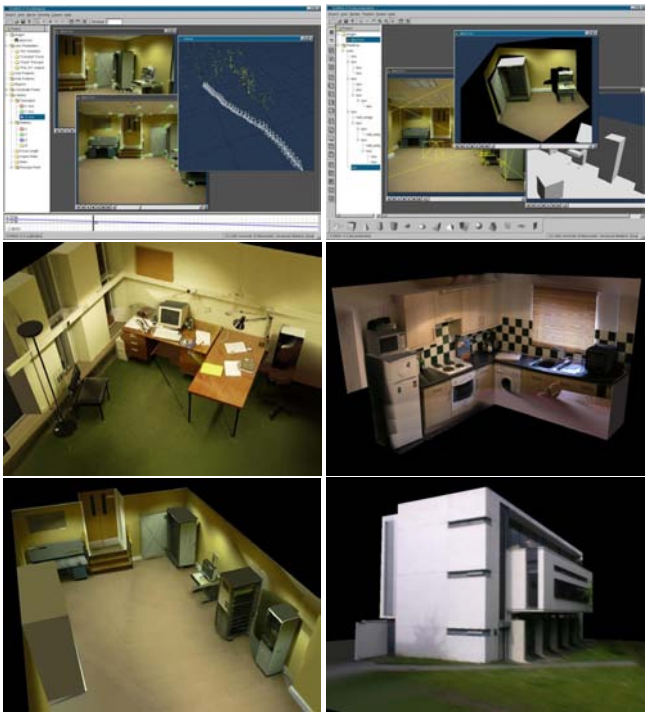

Figure 1: The ICARUS calibration and reconstruction user-interfaces (top row). Example reconstructions from a single image (middle-left), a video camera mounted on a tripod (middle-right), and from hand-held video camera footage (bottom row).

References

DEBEVEC, P., AND MALIK, J. 1996. Modeling and rendering architecture from photographs: A hybrid geometry and image-based approach. In *Proceedings of SIGGRAPH 1996*, 11–20.

HAKIM, S. E. 2000. A practical approach to creating precise and detailed 3d models from single and multiple views. *Int. Archives of Photogrammetry and Remote Sensing 33* (July), 122–129.

POLLEFEYS, M., KOCH, R., AND GOOL, L. V. 1999. Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters. *Int. Journal of Computer Vision 32*, 1, 7–25.