Visualization Methods for Computer Vision Analysis

Mauricio Hess-Flores, Shawn Recker and Kenneth I. Joy
Institute for Data Analysis and Visualization
University of California, Davis, USA
mhessf@ucdavis.edu, strecker@ucdavis.edu, kenneth.i.joy@gmail.com

Mark A. Duchaineau
Google, Inc.
Mountain View, California, USA
duchaine@cognigraph.com

***Abstract*—We present the general idea of using common tools from the field of scientific visualization to aid in the design, implementation and testing of computer vision algorithms, as a complementary and educational component to purely mathematics-based algorithms and results. The interaction between these two broad disciplines has been basically non-existent in the literature, and through initial work we have been able to show the benefits of merging visualization techniques into vision for analyzing patterns in computed parameters. Specific examples and initial results are discussed, such as scalar field-based renderings for scene reconstruction uncertainty and sensitivity analysis as well as feature tracking summaries, followed by a discussion on proposed future work.**

Keywords*—*visualization; computer vision; uncertainty analysis; scalar fields; multi-view reconstruction;

I. Introduction

The field of computer vision has seen great advances in recent years, in fields such as object detection and tracking, and the multi-view reconstruction of scenes. Algorithms such as the Scale-Invariant Feature Transform (SIFT) [1] have allowed for very accurate feature detection and tracking, the main component behind such algorithms. For an excellent overview of many classical vision algorithms, the reader is referred to Hartley and Zisserman [2]. One drawback of current computer vision methods is that many are based on mathematical optimization of initial parameter estimates to achieve accurate results. Though such optimization is provably necessary, such as in the case of the well known *bundle adjustment* [3] in structure-from-motion, little interest has been given as to how *individual* values and patterns in parameter space affect the total cost.

The main objective of our work is to introduce visualization techniques to the vision community as a very powerful educational and algorithm design/test tool, allowing for unique visually-aided numerical exploration of the solution space in a number of applications. Visualization as a field aims to provide renderings of volumes and surfaces, including those that are time-dependent, to graphically illustrate scientific data for its understanding. Some concrete applications and results will be discussed in Section II, such as the use of scalar fields in understanding scene reconstruction uncertainty and parameter sensitivity, as well as feature tracking summaries. Conclusions and future work will be discussed in Section III.

II. Initial Applications

A. Visualization of Scene Structure Uncertainty

One specific application of visualization in computer vision involves a novel tool which we have designed, which allows for the analysis of scene structure uncertainty and its sensitivity to different multi-view scene reconstruction parameters [4]. Multi-view scene reconstruction is an important sub-field of computer vision, aiming to extract a 3D point cloud representing a scene. Detailed analysis and comparisons between methods are available in the literature [5]. Our tool creates a scalar field volume rendering, which provides insight into structural uncertainty for a given 3D point cloud position, displaying the error pattern for sampled bounding box-enclosed neighbors. A screenshot of the tool is shown in Figure 1. The combined statistical, visual, and isosurface information of the right-hand panel provides user insight into the uncertainty of the computed structure shown on the left. Uncertainty arises from error accumulation in the stages leading up to structure computation, such as frame decimation, feature tracking, and self-calibration, where larger regions of low uncertainty indicate robustness of the computed structure. Sensitivity is defined as the change in scalar field values as a specific reconstruction parameter's value changes. The scalar field is created at a user-specified size and resolution, from angular error measurements. For a reconstruction from images in Figure 2(a), a selected position in red enclosed by a green bounded region is shown in Figure 2(b) and magnified in Figure 2(c), with camera positions rendered in blue. The scalar field corresponding to the bounded region, shown in Figure 2(d), depicts lower positional uncertainties in red and higher ones in yellow and green. The red user-defined isosurface, which contains the ground-truth structure position, depicts regions of lowest uncertainty. The visible column-like shape indicates that lower uncertainty is seen in the directions along the lines of sight of the cameras. User interaction allows for uncertainty and sensitivity analysis in ways that have traditionally been achieved mathematically, without any visual aid.

B. Tracking Summaries

Another application is the creation of feature tracking summaries. A feature track is a set of pixel positions representing a scene point tracked over a set of images. A summary can be created by stacking frames vertically and observing the 'path' taken by a track. Given a 3D position computed from multi-view stereo, its reprojection error with respect to its corresponding feature track is the only valid metric to assess error, in the absence of ground truth. However, the total reprojection error value in itself does not allow the researcher to attribute a cause to the error, for example an individual bad match. We present a novel visualization technique for sequential scene reconstruction, which encodes all *individual* reprojection errors in one summary rendering. This provides insight into track degeneration patterns over time, as well as

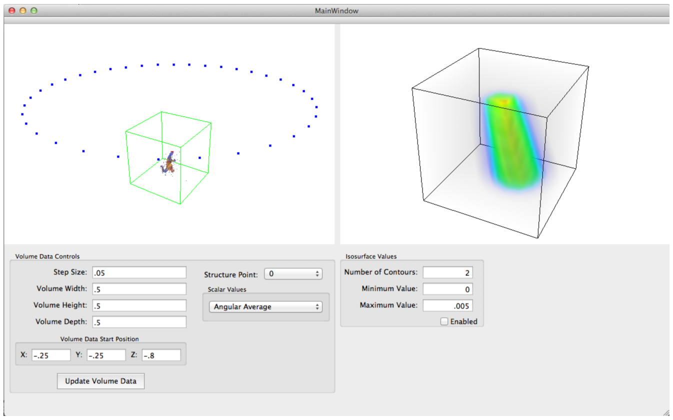
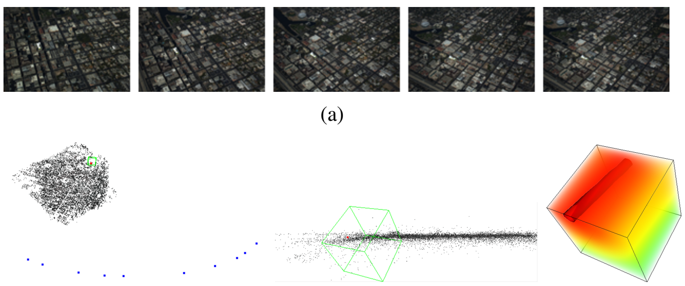


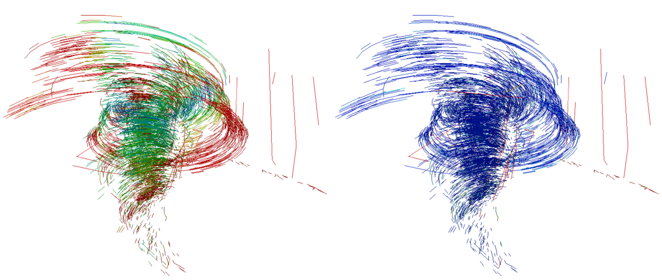
Fig. 1. Our user-interactive tool for uncertainty and sensitivity analysis in multi-view scene reconstruction, based on a scalar field analysis adopted from the visualization literature.



(a)

(b) (c) (d)

Fig. 2. Our uncertainty visualization framework. From input aerial images of a city (a), an initial point cloud reconstruction is computed in (b), with estimated camera positions as blue dots. A position on the point cloud is chosen, highlighted in red in (c) and surrounded by a bounding box for scalar-field analysis. The resulting field is shown in (d).



(a) (b)

Fig. 3. Feature track summary for the *Dinosaur* dataset [6]. Each individual track represents the evolution over time, from top to bottom, of a given feature track's pixel position at each image of a sequential image stream. In each track's rendering, low individual reprojection errors are depicted in blue, with higher ones in green (a). Overall values are much smaller after bundle adjustment optimization [3](b).

information about highly inaccurate individual track positions, all of which adversely affect subsequent camera pose estimation and structure computation. These summaries allow the user to infer the cause of a bad track, which pure total error cannot provide, allowing researchers to analyze errors in ways previously unachieved. Figure 3 shows an example.

III. Conclusion and Future Work

This paper presented the general idea of using visualization as a strong tool for computer vision research, which had not been used until now. Concrete results were shown for two applications, multi-view scene reconstruction and feature tracking, where the visualization of patterns in parameter estimates helps shed light into performance analysis for the underlying estimation algorithms. Given these initial results and the wealth of information provided by summary visualizations, we believe the future is bright with regards to incorporating visualization techniques into many other computer vision applications, which is the focus of our ongoing work.

We have identified a number of computer vision problems where applying visualization techniques could be very beneficial, besides scene strucure uncertainty [4] and feature track summaries. One such application is in 'visualizing' fields of covariance matrices, allowing for the visual analysis and comparison of multiple covariance descriptors at once. Such matrices are fundamental in computer vision, and yet the numerical data they provide is not easy to summarize and interpret. By visualizing for example the change over time of parameters related to such matrices, insight can be achieved into understanding specific patterns and trends, which may lead to a better understanding of operators such as image intensity-based descriptors. Another application is in creating scalar field or similar summaries for multi-view bundle adjustment optimization and other optimization cost functions. Yet other applications we are targeting are in video processing, such as summaries of object tracking results for performance analysis, as well as video content summaries.

References

[1] D. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal On Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[2] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, 2004.

[3] M. Lourakis and A. Argyros, "The design and implementation of a generic sparse bundle adjustment software package based on the Levenberg-Marquardt algorithm," Institute of Computer Science - FORTH, Heraklion, Crete, Greece, Tech. Rep. 340, August 2000.

[4] S. Recker, M. Hess-Flores, M. A. Duchaineau, and K. I. Joy, "Visualization of Scene Structure Uncertainty in a Multi-View Reconstruction Pipeline," in *Proceedings of the Vision, Modeling, and Visualization Workshop*, 2012, pp. 183–190.

[5] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, "A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms," in *CVPR '06: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Washington, DC, USA: IEEE Computer Society, 2006, pp. 519–528.

[6] Visual Geometry Group, University of Oxford, "Multi-view and Oxford Colleges building reconstruction," http://www.robots.ox.ac.uk/~vgg/data/data-mview.html.