Vision, Modeling, and Visualization (2012)
M. Goesele, T. Grosch, B. Preim, H. Theisel, and K. Toennies (Eds.)

Visualization of Scene Structure Uncertainty in a Multi-View Reconstruction Pipeline

S. Recker[1], M. Hess-Flores[1], M. A. Duchaineau[2], and K. I. Joy[1]

[1]University of California, Davis
[2]Google, Inc.

Abstract

This paper presents a novel, interactive visualization tool that allows for the analysis of scene structure uncertainty and its sensitivity to parameters in different multi-view scene reconstruction stages. Given a set of input cameras and feature tracks, the volume rendering-based approach first creates a scalar field from angular error measurements. The obtained statistical, visual, and isosurface information provides insight into the sensitivity of scene structure at the stages leading up to structure computation, such as frame decimation, feature tracking, and self-calibration. Furthermore, user interaction allows for such an analysis in ways that have traditionally been achieved mathematically, without any visual aid. Results are shown for different types of camera configurations, where it is discussed for example how over-decimation can be detected using the proposed technique, and how feature tracking inaccuracies have a stronger impact on scene structure than the camera's intrinsic parameters.

Categories and Subject Descriptors (according to ACM CCS): I.2.10 [Image Processing and Computer Vision]: Vision and Scene Understanding—3D/Stereo Scene Analysis

1. Introduction

Recently, there has been a lot of work in the field of multi-view scene reconstruction. In the reconstruction process, three-dimensional (3D) objects and scenes can be computed from a collection of images taken from different camera viewpoints. Most common reconstruction algorithms produce a point cloud representing the scene's structure. In the literature, such a reconstruction typically involves a number of stages, such as feature tracking, frame decimation, self-calibration, camera pose estimation, structure computation, and parameter optimization. State-of-the-art algorithms [GSC*07] provide very accurate final scene reconstructions. These are based on sparse feature detection and matching, such as SIFT [Low04] and SURF [BETVG08].

The accuracy of a multi-view reconstruction relies fundamentally on accurate feature tracking, as this affects the subsequent camera intrinsic and extrinsic calibrations, as well as the computation of scene structure. Even when using robust estimation procedures and outlier detection, such as RANSAC [HZ04], lighting conditions, occlusions, and repetitive patterns limit feature tracking efficacy and skew subsequent calibration and structure estimation. These stages are prone to additional sources of error and numerical instability. Furthermore, the absence of ground-truth camera and structure parameters forces multi-view algorithms to resort to non-linear optimization of parameters to reduce reprojection error in order to obtain accurate point clouds. However, high numbers of scene points and cameras can make such *bundle adjustment* an expensive element in a reconstruction pipeline, and must be used judiciously, despite efficient sparse implementations [LA00].

The main contribution of this paper is to present a novel interactive tool, which allows for the analysis of scene structure uncertainty and its sensitivity to different multi-view scene reconstruction parameters. The tool allows for both a visual and numerical analysis, by use of a simple angular error metric to create a scalar field, at a user-specified resolution. The scalar field itself provides insight into structural uncertainty, measuring the error in a given 3D grid position with repsect to the correct structure position. In this context, sensitivity is defined as the change in scalar field values as a specific parameter's value changes. The main objective of our work is to introduce visualization techniques to the scene reconstruction community, allowing for unique visually-aided numerical exploration of the solution space in

DOI: 10.2312/PE/VMV/VMV12/183-190

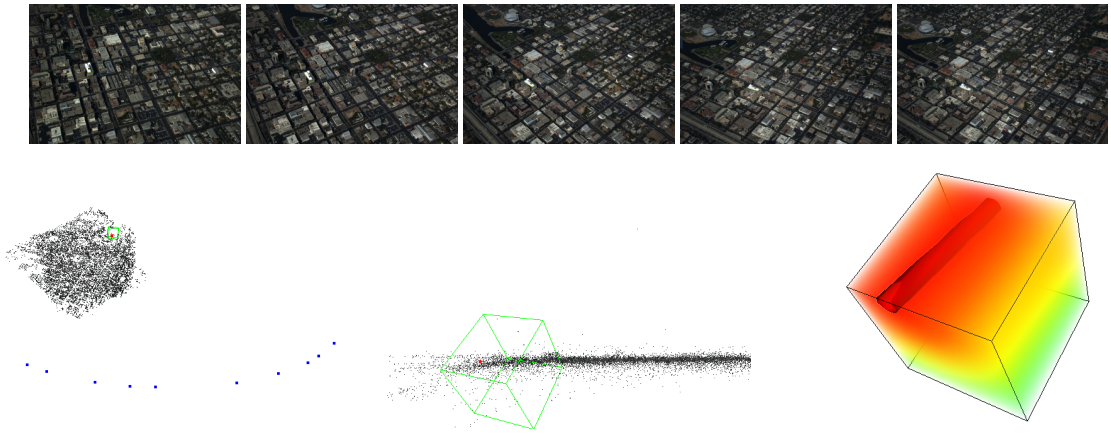


Figure 1: *Images (top) used for the reconstruction view (left and middle) and structure uncertainty view (right) in our tool. Camera positions are shown in blue. The bounded region in green (left and magnified in the middle image) corresponds to the visualized scalar field (right). The scene point under consideration is highlighted in red (left and middle). The scalar field (right) depicts lower uncertainties enclosed in a red isosurface (also containing the scene point) and higher ones in yellow and green, showing a column-like shape where greater uncertainty is seen in the directions along the plane of the scene.*

structure computation. A screenshot of our tool is shown in Fig. 1.

2. Related Work

As mentioned, point cloud scene reconstruction obtains a 3D representation of the underlying scene from a collection of images. The following sequential stages are necessary for performing multi-view reconstruction, keeping in mind that there are many different algorithms for this purpose and that these are the most common steps. Corresponding pixels, known generally as *feature tracks*, can be computed using dense or sparse algorithms. This is the most important process in scene reconstruction, as errors in this stage will affect all subsequent stages [HZ04]. Frame decimation [Nis01] should be applied at this point to filter out frames that lead to very small or very large baselines. Numerical instability occurs with small baselines, while large baselines lead to feature tracking inaccuracies. Next, camera intrinsic calibration is performed by a process known as self-calibration, which aims to recover the cameras' intrinsic parameters, for example focal length [HZ04]. Also, the 'epipolar geometry' can be estimated from matches between consecutive image pairs or triplets [HZ04]. The epipolar geometry mathematically encapsulates the intrinsic projective geometry between groups of views, and is directly related to pose estimation, or the recovery of the cameras' extrinsic parameters of translation and rotation [HZ04]. Between pairs and triplets of views, only relative extrinsic parameters can be computed, but with a previously-computed scene structure, the Direct Linear Transformation [HZ04] can be used to estimate absolute poses. Once the camera parameters are estimated, computation of the scene's 3D structure can be achieved by methods such as 'linear triangulation' [HZ04]. In the absence of ground-truth information, bundle adjustment is the only valid geometrical evaluation of accuracy and is performed to optimize all or a number of the different camera and structure parameters [LA00]. Typically, the Levenberg-Marquardt algorithm is used to minimize the 'reprojection error' of all computed structure points across all cameras with respect to the fixed feature tracks.

There are numerous algorithms in the computer vision literature based on the described pipeline stages. For example, Akbarzadeh et al. [AFM*06] introduced a method for dense reconstruction of urban areas from a video stream. Pollefeys et al. [PNF*08] used a similar approach for real-time urban reconstruction. Goesele et al. [GSC*07] presented a reconstruction pipeline for large, unstructured collections of online photographs of a scene, based on an adaptive view selection technique that robustly computes depth maps along with 3D models of the scene.

There has been some work on uncertainty analysis specifically for scene structure computation [HZ04, BCGvdH01, CFA98, ZN96], but it has been mainly a mathematical analysis which has not been enhanced by visualization techniques. For example, Rodehorst et al. [RHH08] introduced a ground-truth based approach to evaluate camera pose estimation, while Knoblauch et al. [KHFDK09] introduced a geometric error extraction of both feature matches and camera pose errors. This method does not rely on ground-truth data or any other assumptions about the scene. Our presented framework sheds light into structural uncertainty and sensitivity across the different stages of a pipeline, such as feature tracking, frame decimation and self-calibration.

3. Procedure

As input to the visualization tool, it is assumed that the projection matrices for N cameras are known, or alternatively their individual intrinsic and extrinsic parameters. Additionally, a set of feature tracks across the images and the resulting scene structure are required. However, it is important to note that for the simulations discussed in the Results section, ground-truth feature tracks, cameras, and structure were used in order to validate the output of the proposed visualization tool.

In the first step, one of the computed scene points is chosen by the user. Ideally, it should be a point that is seen by all or the greatest number of cameras, as this allows for a more accurate uncertainty analysis. A scalar field

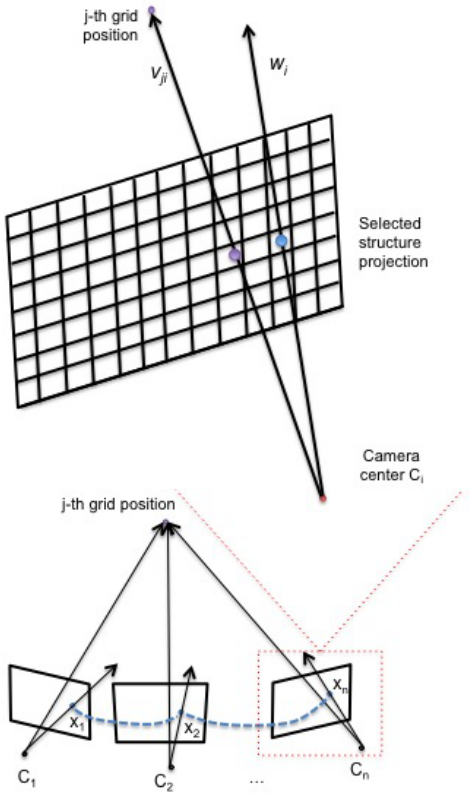


Figure 2: *Scalar field ray calculations. A scalar field is created by computing a sum of angular errors between the rays viewing the selected point and all grid points.*

over 3D space is then rendered for the chosen point in the visualization tool. To calculate each value in the scalar field, as shown in Fig. 2, for each camera center C_i a unit vector v_{ji} is computed between all 3D positions that lie on a regularly-spaced grid with M samples, denoted with subindex j. The resolution M of this grid can be specified by the user. The user can also specify the spatial location of this grid, along with its dimensions along the X, Y and Z world axes.

A second unit vector from each camera center, w_i, is obtained by casting a ray from each center C_i through the visible projection of the chosen structure point on each image plane (blue image plane dot in Fig. 2). This projection generally does not coincide with the projection of a given grid point with vector v_{ji} (purple image plane dot in Fig. 2), and hence there is typically a non-zero angular difference between each possible v_{ji} and w_i. Then we compute, at each grid position for all cameras, the radian angles formed by vectors v_{ji} and w_i, as shown in Eq. 1 for a single camera.

The visualization tool supports both *average* and *range* scalar field types. To obtain the *average* scalar field $S_{j,ave}$, at every j^{th} grid position the previously-computed N angles are added and averaged, as shown in Eq. 2. To obtain the *range* scalar field $S_{j,range}$, at every j^{th} grid position the range between maximum and minimum angles is obtained, as shown in Eq. 3. As will be discussed in the Results section, the power of the visualization tool lies in interpreting the joint information provided by these two scalar field types, along with their associated statistics.

$$S_{ji} = cos^{-1}(v_{ji} \cdot w_i) \tag{1}$$

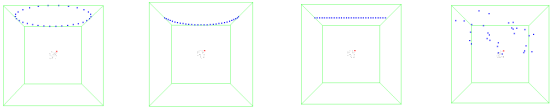
$$S_{j,ave} = \frac{\sum_{i=1}^{N} S_i}{N} \tag{2}$$

$$S_{j,range} = S_{ji,max} - S_{ji,min} \tag{3}$$

Upon providing the necessary information to the tool, the cameras' positions, computed structure, scalar field dimensions, and chosen structure (shown in red) are displayed on the left. The right-hand panel displays the resulting scalar field. The scalar field visualization was implemented in VTK [Inc12], which utilizes a ray casting technique for volume rendering. Samples are color-coded such that red indicates low structural uncertainty in a particular region, whereas blue indicates high structural uncertainty. The opacity for a region can be adjusted by the user. In addition, a VTK marching cubes implementation [Inc12] is used to generate an isosurface, which encloses sub-volumes (shown in dark red) of the best possible structure locations.

3.1. Analyzed Test Cases

Several tests were conducted to analyze the sensitivity of a reconstructed point to parameters across different stages of a reconstruction pipeline. To this end, tests were performed on four different types of camera configurations, in synthetic scenes with ground-truth information available, as displayed in Fig. 3. The first configuration represents a set of cameras positioned in a circle above the scene. The second configuration is similar to the first, except using only a semi-circle of cameras. The third configuration involves a set of cameras in a line above the scene. The fourth configuration involves a set of cameras that were placed randomly, representing an



(a) Circle (b) Semi-circle (c) Line (d) Random

Figure 3: *Synthetic camera configurations. The examined scene structure is shown in red and the volume analyzed is displayed in green. The cameras are shown in blue.*

unstructured collection of images. Each configuration consists of 30 cameras, each looking towards the origin, $(0,0,0)$, of the scene. It was assumed throughout all tests that the ground-truth position of the analyzed structure point was located at $(0.1,0.1,0.1)$ in world space and the same physical camera was used to acquire every view.

The first test was designed to examine the effect of frame decimation [Nis01] on structure computation. The second test measured the sensitivity of feature tracking errors on multi-view reconstruction. Finally, the last tests evaluated the effect of inaccuracy in self-calibration, through respective experiments on varying focal length and principal point parameters. For the simulation, the principal point for each camera was translated in the same random direction from the $(0,0)$ image plane origin, by amounts corresponding respectively to 1%, 2%, 5%, 10%, and 20% of the image plane size. For every test, both the *average* and *range* scalar fields were computed.

3.1.1. Frame Decimation Simulation

The goal of this simulation was to study the effect of 'frame decimation' on a multi-view reconstruction, from the point of view of scalar field analysis. To this end, for the four tested camera configurations, cameras were evenly decimated from the original 30 down to 15, 10, 8, 4, and finally 2 cameras, such that the baseline between consecutive cameras increased each time, with equal spacing between each.

3.1.2. Feature Tracking Simulation

The objective of this simulation was to simulate inaccuracy in feature tracking, and then observe the effect on the obtained scalar fields. To simulate feature matching error, the correct projected position of the analyzed structure point at $(0.1,0.1,0.1)$ was moved in a random direction on each camera's image plane, by the same fixed amount. The tested amounts were 1%, 2%, 5%, 10%, and 20% of the image plane size.

3.1.3. Self-Calibration Simulation

Principal point variation. The goal of this simulation was to investigate the effect of varying each camera's principal point to positions other than the $(0,0)$ center of the image plane. This test, along with a similar one for focal length, were designed to study the effect of inaccuracy in the self-calibration process.

Focal length variation. For the last test, focal length was varied with respect to its initial ground-truth value, similarly to the principal point simulation. Focal length was decreased by 1%, 2%, 5%, 10%, and 20% of its original value.

4. Results

All tests were conducted on a MacBook Pro machine with an *Intel Core i7* processor at 2.66 GHz with 4 GB of RAM, running Mac OS X Lion 10.7.3. Analysis was performed on the two types of scalar fields defined in Section 3: *average* and *range*. The power of our proposed tool comes from the joint analysis of these two fields and their associated statistics, which produce both visual and numerical results for structure uncertainty and multi-view reconstruction stage sensitivity analysis. The key lies in correctly *interpreting* the information it provides, to use if most efficiently for studying the effects of parameter variation on multi-view structure computation. For the two types of scalar fields, the following four statistics were computed across the entire field: average μ, standard deviation σ, volume V of lowest uncertainty enclosed by a given isosurface value, and ratio R of the longest to shortest sides of the bounding box that encloses the isosurface, in order to analyze its shape.

4.1. Simulation Results

Frame decimation simulation results. Trend charts for this simulation are shown in Fig. 4. The *range* scalar fields for each of the six tests performed on the *circle* configuration are shown in Fig. 8. From Fig. 4, it can be seen that the average μ of the *average* scalar field remains fairly constant as the number of cameras is reduced, but there is a more pronounced increase in the average μ (and standard deviation σ) of the *range* scalar field below 10 cameras. This indicates that proper frame decimation can filter out a great number of frames without structure uncertainty being affected much. With over-decimation, however, uncertainty increases due to the lack of rays for triangulation, as evidenced by the range values for fewer cameras. Isosurface volume, V, for the *range* field initially increases due to the maintained good conditioning, but falls apart very rapidly with over-decimation, as seen for the four and two-camera cases. The visual effect of this is clear in Fig. 8. For 30 cameras down to around 8, the isosurfaces show a column-like structure near the middle of the camera positions, with a bowl-like shape at a certain depth from the cameras. This is directly related to the concept of *baseline-to-depth ratio*, an important factor in scene reconstruction. The bowl-like structure essentially provides a range of 'good' baseline-to-depth ratios *and* wherever it appears it indicates that a proper frame decimation is being performed. With over-decimation, however, enough information is missing such that the solid bowl shape disappears and the lowest uncertainty values appear only in

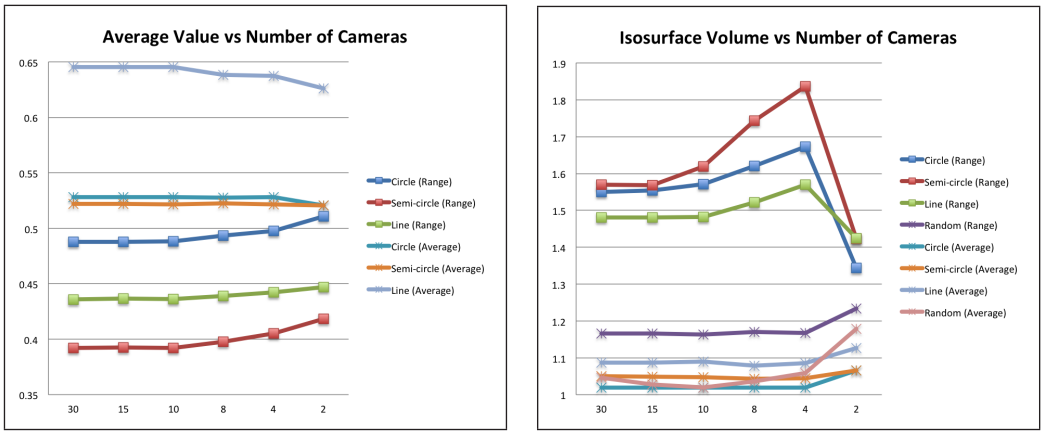


Figure 4: *Frame decimation trend charts. The average* μ *for the* average *and* range *scalar fields vs. number of cameras (left) and the isosurface volume V vs. number of cameras (right) are shown for each configuration. An isovalue of* 0.05 *was used in both fields for* circle, semi-circle *and* line. *A value of* 0.07 *was used in* random *for the* average *field, and* 0.4 *for* range. *Results show that frame decimation maintains structural stability until around four frames, where over-decimation begins to manifest.*

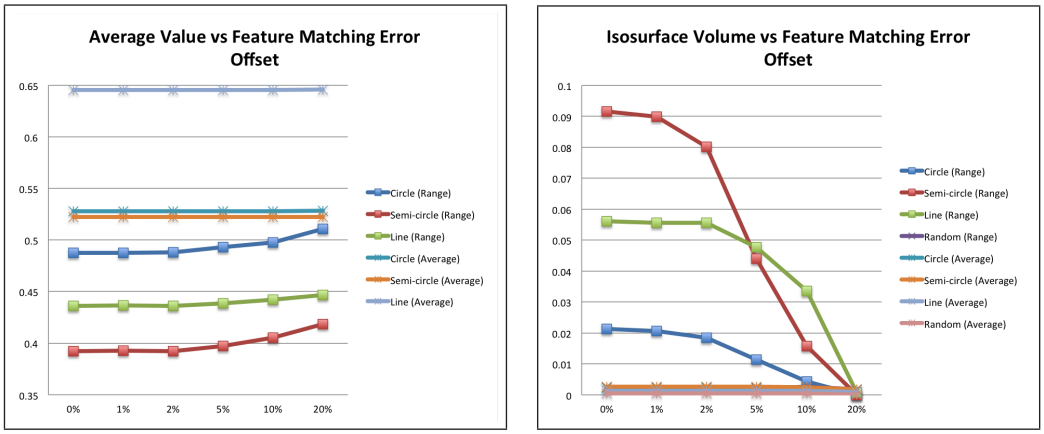


Figure 5: *Feature matching trend charts. The average* μ *for the* average *and* range *scalar fields vs. feature matching offset error (left) and the isosurface volume V vs. feature matching offset error (right) are shown for each configuration. An isovalue of* 0.05 *was used in both fields for* circle, semi-circle *and* line. *A value of* 0.07 *was used in* random *for the* average *field, and* 0.4 *for* range. *Results confirm that scene structure is very sensitive to feature tracking errors.*

the line of direction of each camera, as can be seen in the interesting shapes for the four and two-camera top views in Fig. 8. Furthermore, with over-decimation, feature tracking suffers from inaccuracy due to perspective, illumination and occlusion changes in the viewed scene.

Feature tracking simulation results. Trend charts for this simulation are shown in Fig. 5. The *range* scalar fields for each of the six tests performed on the *semi-circle* configuration are shown in Fig. 9. From Fig. 5, it can be seen that the average μ of the *average* scalar field remains almost exactly constant, while the average μ (and standard deviation σ) of the *range* field increases slightly with feature matching error, mainly after 2%. This coincides with a decrease in isosurface volume V for the *range* scalar field. Changes are more pronounced for the *circle* and *semi-circle* configurations. It can be seen in Fig. 9 how the size of the isosurface-enclosed region diminishes with increasing error, indicating that it is unlikely to obtain an accurate scene structure as feature tracking becomes inaccurate, confirming its known sensitivity to tracking errors from the literature. Notice how the isosurface for the 'no error' case differs in shape from the bowl-like structure seen in Fig. 8 for the *circle* configuration, such that good baseline-to-depth-ratios have a different range for this setup. Also, it was noticed that interpreting the visual results for the *average* scalar field provides more information than analyzing the numerical results, as isosurface

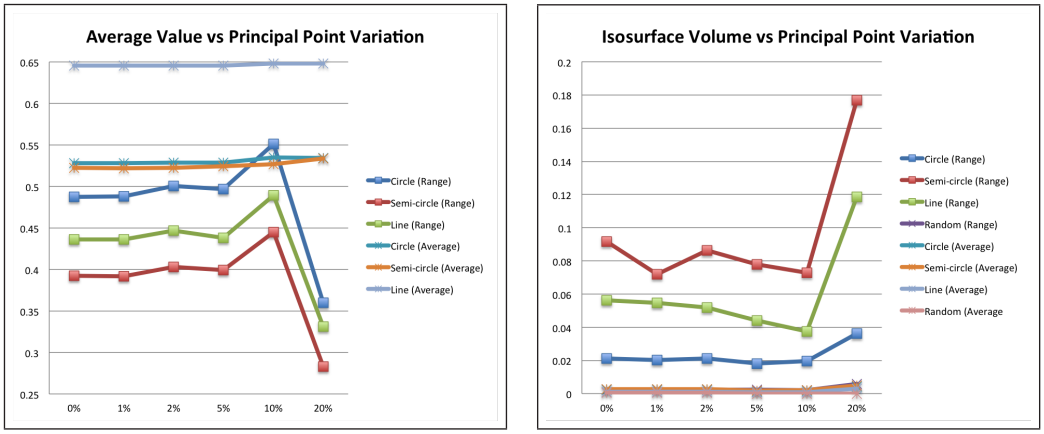


Figure 6: *Principal point trend charts. The average* μ *for the* average *and* range *scalar fields vs. principal point offset error (left) and the isosurface volume V vs. principal point offset error (right) are shown for each configuration. An isovalue of* 0.05 *was used in both fields for* circle, semi-circle *and* line. *A value of* 0.07 *was used in* random *for the* average *field, and* 0.4 *for* range. *Results show that scene structure is not very sensitive to principal point variation, mainly affecting its position.*

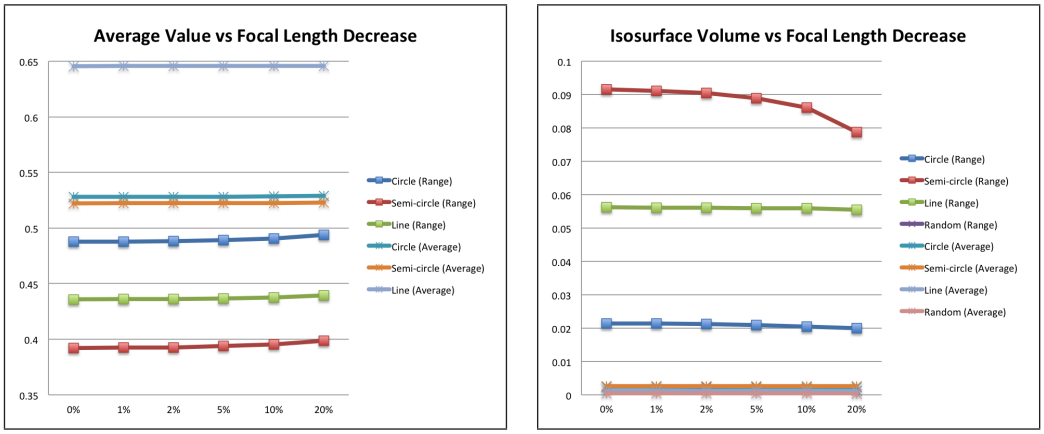


Figure 7: *Focal length trend charts. The average* μ *for the* average *and* range *scalar fields vs. focal length decrease (left) and isosurface volume V vs. focal length decrease (right) are shown for each configuration. An isovalue of* 0.05 *was used in both fields for* circle, semi-circle *and* line. *A value of* 0.07 *was used in* random *for the* average *field, and* 0.4 *for* range. *Results show that scene structure is not very sensitive to focal length variation, mainly affecting its scale.*

shape provides visual evidence related to the main direction of uncertainty, whereas isosurface shape, measured with the 'ratio' parameter R, remains fairly constant.

Self-calibration simulation results. Trend charts for the principal point simulation are shown in Fig. 6. The *range* scalar fields for each of the six tests performed on the *random* configuration are shown in Fig. 10. From Fig. 6, it is interesting to observe that with increasing principal point deviation, the average μ (and standard deviation σ) for the *range* scalar field remains fairly constant but eventually actually decreases for large deviations, and isosurface volume V increases similarly. Average μ and isosurface volume V of the *average* scalar field tend to increase very slightly, as seen in Fig. 10. This interesting effect seems to indicate that the final computed scene structure is not very sensitive to small principal point deviations, unlike with other parameters such as feature tracks unless very inaccurate. It affects mainly the *position* of the final structure in 3D due to the shift in image plane ray intersections.

As for the focal length simulation, trend charts are shown in Fig. 7, with *average* scalar fields for each of the six tests performed on the *line* configuration displayed in Fig. 11. In general, average μ (and standard deviation σ) of the *range* scalar field increase insignificantly as focal length decreases,

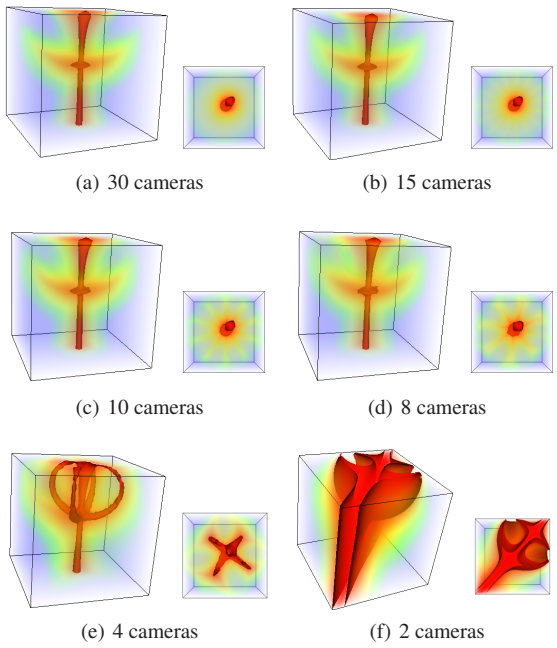


Figure 8: Range *scalar fields and isosurfaces for the frame decimation simulation applied on the* circle *configuration. In all images an isovalue of* 0.05 *was used, with side views on the left and top views (parallel to the plane of the cameras) on the right for each case.*

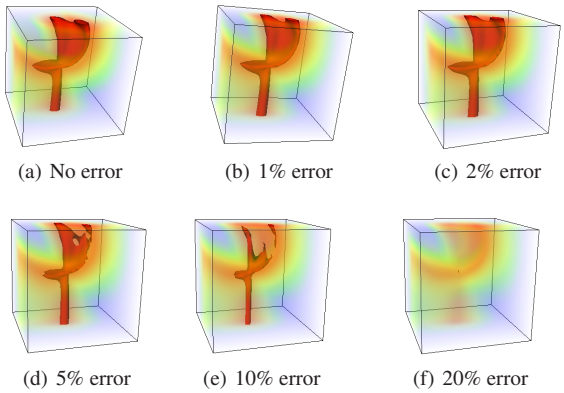


Figure 9: Range *scalar fields and isosurfaces for the feature matching simulation applied on the* semi-circle *configuration. In all images an isovalue of* 0.05 *was used.*

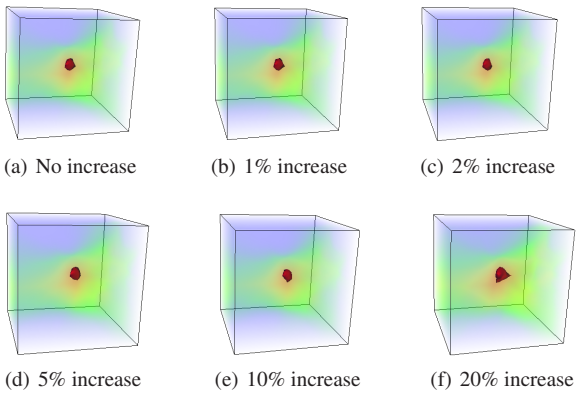
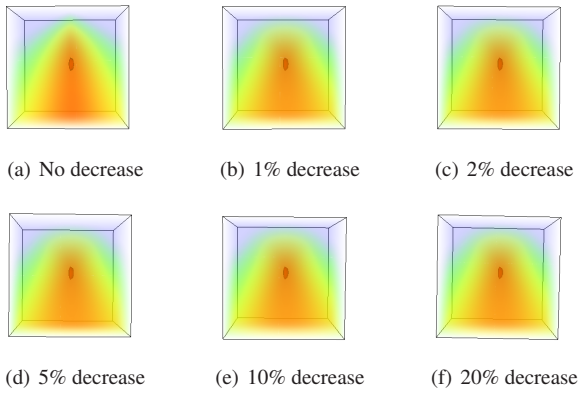


Figure 10: Range *scalar fields and isosurfaces for the principal point simulation applied on the* random *configuration. In all images an isovalue of* 0.4 *was used.*

across all camera configurations. The average μ and standard deviation σ for the *average* field also remain basically constant. The volume V of the generated isosurfaces decreases very slightly as the focal length decreases, except for the *line* configuration, where it remains constant. The isosurface 'ratio' values R remain fairly unchanged. This analysis indicates that scene reconstruction is not distorted or sensitive to large changes in focal length, which mainly affects its scale but not its stability. This has been verified in multi-view reconstruction tests, where a wide range of input focal length values produced very similar final reconstructions.

4.2. Discussion

The performed tests focused on analyzing the effect of frame decimation, feature matching inaccuracy, and self-calibration on structure computation. For frame decimation, removing cameras up to a certain point does not cause drastic visual or statistical changes. On the other hand, our results confirm the effect of over-decimation, where critical frames are discarded such that information is lost and a higher structural uncertainty can be expected [Nis01]. The results for feature matching show that scene reconstruction is very sensitive to feature tracking inaccuracies. The isosurface volume, especially for the *range* scalar field, decreases dramatically across all configurations as error increases. In the experimentation of principal point inaccuracy, it was surprising to notice how *range* scalar field average eventually decreases and isosurface volume increases with principal point offset, indicating that scene structure is not very sensitive to small variations. Modifying the camera focal length in general resulted in very small variations, affecting mainly the scale of the final reconstruction but not distorting it nor affecting its accuracy much, even over a large range of values. Another important observation is that the different camera configurations produced differently-shaped isosurfaces. For example, the *circle* and *random* configurations produce more spherical *average* scalar field isosurfaces, while the *semi-circle* and *line* configurations produce more elliptical regions, where the axis with most spread indicates the direction of higher uncertainty, which appears to lie orthogonal to the actual



(a) No decrease (b) 1% decrease (c) 2% decrease

(d) 5% decrease (e) 10% decrease (f) 20% decrease

Figure 11: Average *scalar fields and isosurfaces for the focal length simulation applied on the* line *configuration. In all images an isovalue of* 0.05 *was used.*

camera configuration. For all of our tests only a single parameter was varied each time, leading to a single scalar field from which sensitivity to the particular parameter could be analyzed. However, analysis in the case of varying multiple parameters, such as in pose estimation where there are 5n parameters for n cameras, is more complicated. This leads to an exponential number of scalar field evaluations which is beyond the scope of this paper.

5. Conclusions

In this paper, a novel user-interactive visualization and statistical tool is presented, which provides insight into structure uncertainty and its sensitivity to parameters in multi-view scene reconstruction. Given a set of input camera parameters, feature tracks and scene structure, the user is able to generate a scalar field visualization, based on an angular error metric, along with corresponding statistical data, which enables sensititivy analysis in reconstruction stages such as frame decimation, feature tracking and self-calibration. This includes the ability to modify opacity and render isosurfaces. To validate the proposed tool, a number of synthetic tests were performed using four typical camera configurations, and also applied to real datasets. Results show that the joint analysis of two scalar field types, along with corresponding isosurfaces and statistical data, allows the user to infer structural uncertainty and sensitivity to the underlying parameters involved in multi-view reconstruction.

6. Acknowledgements

This work was supported in part by Lawrence Livermore National Laboratory and the National Nuclear Security Agency through Contract No. DE-FG52-09NA29355. We thank our colleagues in the Institute for Data Analysis and Visualization (IDAV) at UC Davis for their support.

References

[AFM*06] AKBARZADEH A., FRAHM J.-M., MORDOHAI P., CLIPP B., ENGELS C., GALLUP D., MERRELL P., PHELPS M., SINHA S., TALTON B., WANG L., YANG Q., STEWENIUS H., YANG R., WELCH G., TOWLES H., NISTER D., POLLEFEYS M.: Towards urban 3d reconstruction from video. In *3D Data Processing, Visualization, and Transmission, Third International Symposium on* (june 2006), pp. 1 –8. 2

[BCGvdH01] BROOKS M. J., CHOJNACKI W., GAWLEY D., VAN DEN HENGEL A.: What value covariance information in estimating vision parameters? In *ICCV'01* (2001), pp. 302–308. 2

[BETVG08] BAY H., ESS A., TUYTELAARS T., VAN GOOL L.: Speeded-up robust features (SURF). *Comput. Vis. Image Underst. 110* (June 2008), 346–359. 1

[CFA98] CHEONG L., FERMÜLLER C., ALOIMONOS Y.: Effects of Errors in the Viewing Geometry on Shape Estimation. *Comput. Vis. Image Underst. 71*, 3 (1998), 356–372. 2

[GSC*07] GOESELE M., SNAVELY N., CURLESS C., HOPPE H., SEITZ S. M.: Multi-view stereo for community photo collections. In *Proceedings of ICCV 2007* (2007). 1, 2

[HZ04] HARTLEY R. I., ZISSERMAN A.: *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, 2004. 1, 2

[Inc12] INC. K.: Vtk: Visualization toolkit, 2012. 3

[KHFDK09] KNOBLAUCH D., HESS-FLORES M., DUCHAINEAU M. A., KUESTER F.: Factorization of correspondence and camera error for unconstrained dense correspondence applications. In *5th International Symposium on Visual Computing* (2009), pp. 720–729. 2

[LA00] LOURAKIS M., ARGYROS A.: *The Design and Implementation of a Generic Sparse Bundle Adjustment Software Package Based on the Levenberg-Marquardt Algorithm*. Tech. Rep. 340, Institute of Computer Science - FORTH, Heraklion, Crete, Greece, August 2000. 1, 2

[Low04] LOWE D.: Distinctive image features from scale-invariant keypoints. *International Journal On Computer Vision 60*, 2 (2004), 91–110. 1

[Nis01] NISTÉR D.: Frame decimation for structure and motion. In *SMILE '00: Revised Papers from Second European Workshop on 3D Structure from Multiple Images of Large-Scale Environments* (London, UK, 2001), Springer-Verlag, pp. 17–34. 2, 4, 7

[PNF*08] POLLEFEYS M., NISTÉR D., FRAHM J.-M., AKBARZADEH A., MORDOHAI P., CLIPP B., ENGELS C., GALLUP D., KIM S.-J., MERRELL P., SALMI C., SINHA S., TALTON B., WANG L., YANG Q., STEWÉNIUS H., YANG R., WELCH G., TOWLES H.: Detailed real-time urban 3d reconstruction from video. *International Journal of Computer Vision 78* (2008), 143–167. 10.1007/s11263-007-0086-4. 2

[RHH08] RODEHORST V., HEINRICHS M., HELLWICH O.: Evaluation of relative pose estimation methods for multi-camera setups. In *International Archives of Photogrammetry and Remote Sensing (ISPRS '08)* (Beijing, China, 2008), pp. 135–140. 2

[ZN96] ZHAO W., NANDHAKUMAR N.: Effects of Camera Alignment Errors on Stereoscopic Depth Estimates. *Pattern Recognition 29*, 12 (December 1996), 2115–2126. 2