Multiview Optical Tomography Datasets Overview

James Gregson

March 28, 2012

Abstract

This document is intended to provide information on the formatting and use of the multiview optical tomography datasets that were captured as a part of the Stochastic Tomography project [?].

1 License Conditions

2 Introduction

This dataset was captured using an array of up to 16 high-definition Sony camcorders (model:) in a semi-circular arrangement around a central glass capture cylinder. The cameras were synchronized and rolling-shutter effects removed using the method of Bradley et. al. [?] to obtain a set of video frames for each camera at consistent times. Geometric calibration was achieved using an extension of the two-plane method of Trifonov et. al. [?] using sparse world-to-image correspondences obtained with Atcheson et. al.'s publically available CALTag software [?]. The geometric calibration and radiometric calibration processes are described further in Sections 3.2 and 3.3 respectively.

3 Experimental Setup

The experimental setup used to capture the datasets was briefly described in [?] but more detail is provided here. An array of up to 16 Sony camcorders is arranged in a semicircle around a capture volume housed in a glass cylinder mounted concentrically on a rotation stage. The cylinder is filled with water, and additional fluids are introduced via a variety of methods, e.g. injection from the bottom, pouring, etc.

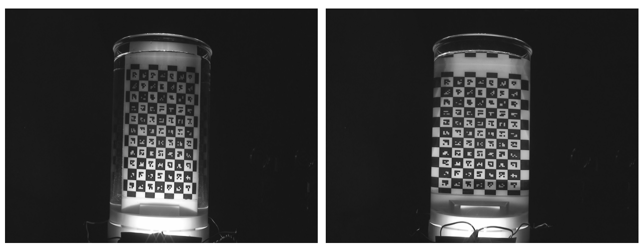
During captures, the capture volume is illuminated by LED strobe panels. These strobe panels are pulsed in order to optically synchronize the cameras using the strobe-synchronization and rolling-shutter compensation method of Bradley [?].

3.1 Calibration Procedure

A calibration target is inserted into the capture volume offset from the rotation axis by a known amount. The rotation stage then rotates the target to face each camera by known angles, each camera takes an image of the target. The stage then rotates by 180 degrees to expose a second calibration target on the reverse side of (and registered to) the first target to each camera, which again capture images. During this process, both targets are illuminated only by one of two *calibration headlights* mounted to the rotation stage. These ensure consistent illumination of the targets as viewed by all cameras as needed by subsequent radiometric calibration. This process is a slightly modified version of the procedure used by Trifonov et. al. [?].

Self-identifying tags on the two targets allow sparse world-to-image correspondences to be obtained for the two plane positions. These correspondences are then used to determine parameters for a projection model, which provides pierce points through two world-space planes for rays emanating from each image pixel. Connecting the pierce points allows the ray within the capture volume for each image pixel to be

Figure 1: Experimental setup used to capture the datasets. An array of up to 16 Sony HD camcorders is arranged in a semi-circle around a central capture volume housed in a water-filled glass cylinder. A calibration target inside the cylinder and immersed in water allows correspondences between captured images and the target plane. Mounting the cylinder on a rotation stage allows these correspondences to be obtained for all cameras, and capturing a second target at a different plane position then allows parameterization of the rays within the capture volume.



(a) Front plane position (b) Rear plane position

Figure 2: Geometric and radiometric calibration images. Two carefully registered patterns are filmed facing each camera offset from the rotation axis of the rotation stage by different amounts, referred to as the front and rear plane positions. The two diffuse patterns are lit only by a *calibration headlight* that rotates with the capture volume, ensuring that each camera observes the pattern under the same conditions. Interpolating point correspondences obtained from the CALTag self-identifying marker pattern yields per-pixel pierce points for the front and rear plane positions, which parameterizes the ray-bundle through the capture volume.

reconstructed, without needing to know precise camera extrinsics or intrinsics, or needing to account for refractive effects at the air-glass or glass-water interfaces of the capture cylinder.

3.2 Geometric Calibration and Projection Model

By interpolating point correspondences from the CALTag self-identifying marker pattern, the pierce points of rays emanating from each camera pixel on the two planes can be computed. These pierce points define the rays within the capture volume after all refractive events have occured.

We fit 2D cubic polynomial surfaces to the point correspondences for each of the u and v image coordinates. That is, if an image point $[u_i, v_i]$ is chosen, then the pierce point on the front plane $[p_{x_i}, p_{y_i}, p_{z_i} = 0]$ relative to the CALTag origin will be:

$$p_{x_i} = \beta_x^T \begin{bmatrix} u_i^3 \\ v_i^3 \\ u_i^2 v_i \\ v_i^2 u_i \\ u_i^2 \\ v_i^2 \\ u_i v_i \\ u_i \\ v_i \\ 1 \end{bmatrix} = \beta_x^T \Gamma_i \qquad p_{y_i} = \beta_y^T \begin{bmatrix} u_i^3 \\ v_i^3 \\ u_i^2 v_i \\ v_i^2 u_i \\ u_i^2 \\ v_i^2 \\ u_i v_i \\ u_i \\ v_i \\ 1 \end{bmatrix} = \beta_y^T \Gamma_i \tag{1}$$

where β_x and β_y are vectors of fit parameters, and $\Gamma_i = \left[u_i^3, v_i^3, u_i^2 v_i, v_i^2 u_i, u_i^2, v_i^2, u_i, v_i, u_i, v_i, 1\right]^T$ is the image-space cubic surface basis vector. The fit parameters are obtained for each coordinate axis and each plane position as the least-squares solution of the overconstrained problem over the N point correspondences obtained from CALTag:

$$\begin{bmatrix} \Gamma_0^T \\ \Gamma_1^T \\ \vdots \\ \Gamma_{N-1}^T \end{bmatrix} \beta_x = \begin{bmatrix} p_{x_0} \\ p_{x_1} \\ \vdots \\ p_{x_{N-1}} \end{bmatrix} \qquad \begin{bmatrix} \Gamma_0^T \\ \Gamma_1^T \\ \vdots \\ \Gamma_{N-1}^T \end{bmatrix} \beta_y = \begin{bmatrix} p_{y_0} \\ p_{y_1} \\ \vdots \\ p_{y_{N-1}} \end{bmatrix} \tag{2}$$

We have found that using these 2D cubic polynomial fits results in low reprojection errors, while maintaining a compact, smooth and differentiable camera model.

After reconstructing the pierce point for a given $[u_i, v_i]$ in pattern space, it is necessary to transform it to world-space. To do so we translate the point by the offset $[f_x, f_y, f_z]$ of the CALTag pattern origin from the known rotation axis ($x = y = 0$). This offset is known from the geometry of the calibration target and capture setup. We then rotate the point by the stage angle θ_j for the given camera (index j), which is also known from the experimental setup. Consequently the world-space pierce point for a given plane is:

$$\begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta_j & 0 & -\sin\theta_j & 0 \\ 0 & 1 & 0 & 0 \\ \sin\theta_j & 0 & \cos\theta_j & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & f_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & f_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_x^T \Gamma_i \\ \beta_y^T \Gamma_i \\ 0 \\ 1 \end{bmatrix} = \mathbf{RT} \begin{bmatrix} \beta_x^T \Gamma_i \\ \beta_y^T \Gamma_i \\ 0 \\ 1 \end{bmatrix} \tag{3}$$

Performing this process for both planes and connecting the two resulting points gives the ray segment through the capture volume for a given image-space position. Note that in the datasets, we specify the CALTag origin offset vectors $[f_x, f_y, f_z]^T$ for the front and rear calibration plane positions viewed by camera j using only θ_j, as opposed to $\theta_j + 180$.

To perform the reverse process, that is, compute an image coordinate for an known world-space coordinate as is needed by Stochastic Tomography [**?**], Equation 3 must be numerically inverted. This requires finding the pixel coordinates producing a ray that passes through the world-space point, for which we use Newton's method. We remove the rotation component of Equation 3 by multiplying the world-space point by $\mathbf{R^{-1}}$ and use as an initial guess at the centroid of the point correspondences detected by CALTag. To perform Newton's method, we define two functions $f_x(u, v)$ and $f_y(u, v)$ which evaluate the distance in the x and y directions respectively between the closest point on the ray through the current, estimated solution $[u, v]^T$ and the world-space point being projected into the image. We then iteratively solve the following equation to obtain an update to the solution $[\Delta u, \Delta v]$, evaluating the differential terms with finite differences.

$$\begin{bmatrix} \frac{\delta f_x(u,v)}{\delta u} & \frac{\delta f_x(u,v)}{\delta v} \\ \frac{\delta f_y(u,v)}{\delta u} & \frac{\delta f_y(u,v)}{\delta v} \end{bmatrix} \begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix} = \begin{bmatrix} -f_x(u,v) \\ -f_y(u,v) \end{bmatrix} \tag{4}$$

We find that this process converges rapidly, generally within 2-4 iterations. Note that this process is *not* only required for Stochastic Tomography, or when it is desirable to project from world to image space. Many typical tomography algorithms should be easily implemented without inverting 3 by simply computing rays for pixels, e.g. SART when ray-tracing is used to evalute the projection operator.

3.3 Radiometric Calibration

Since every camera captures a consistently illuminated view of each plane position, it is possible (and necessary) to radiometrically calibrate each camera to match a single master camera. The strobe-synchronization process inverts the gamma curve applied by each camera, outputting luminance images at consistent times in a linear space. To compare measurements between cameras, we compute for all cameras an effective exposure time scale factor. This is chosen as the ratio of the median gray-levels for a set of known correspondence points on the calibration target plane for every camera with respect to a single master camera. This factor is then applied to each frame of the captured data.

The correspondence points are chosen by picking a large number of points uniformly at random on the calibration target plane and inverting the camera model as described in the previous section to obtain image-space points for each of the pair of cameras. We typically use 10k points for this process.

Note that images in these datasets have not had this process performed.

4 Dataset Directory Structure and Files

Each capture consists of a root directory contain several files, plus a subdirectory per camera. In the root directory are three files:

- **capture.xml** XML formatted file listing camera and calibration parameters. The primary configuration file, with some discussion of capture conditions in XML comments.
- **caltag_pattern_scale0.46875.ps** CALTag pattern file, which matches the pattern used in the capture. Only needed if recalibration is to be performed.
- **caltag_pattern_scale0.46875.mat** CALTag matrix file, used by CALTag to find point correspondences. Only needed if recalibration is to be performed.

In the root directory there are also a number of additional subdirectories, named ***cam##***, each of which contain the data for a single camera in the capture. The contents of these directories is discussed in one of the following subsections.

4.1 capture.xml File Format

The *capture.xml* file uses an XML 1.0 file format, which should be easily parsed from most languages. The root element for the file is the <***capture***> tag, which contains a single <***target***> tag and multiple <***camera***> tags as children.

The <***target***> tag has no children and three attributes:

- ***mat_file*** The filename of the CALTag matrix file. This should match the corresponding matrix file in the top-level directory, and is only needed during recalibration.
- ***front_translation*** The translation of the CALTag pattern coordinate axis (in inches) from the world-space origin for the calibration pattern in the front plane positions. Listed as a triplet of coordinates separated by spaces.
- ***rear_translation*** Equivalent to the ***front_translation*** attribute, but for the rear plane position.

The <***camera***> tag has no children and four attributes:

- ***subdir*** The subdirectory in which this camera's data is stored relative to the root directory.
- ***stage_angle*** The angle of the rotation stage for the front plane position. The angle for the rear plane position will be the front angle plus 180 degrees.
- ***front_calib*** The filename of the CALTag point correspondence file for the front plane position, relative to the camera subdirectory, always *front_calibration.xml*.
- ***rear_calib*** The filename of the CALTag point correspondence file for the rear plane position, relative to the camera subdirectory, always *front_calibration.xml*.

4.2 Camera Subdirectory Contents

Each camera subdirectory stores a number of images and calibration files. Image files are stored as 16bit ***linear*** PNGs.

- ***front_calibration.xml*** The CALTag point correspondence file for the front plane position for this camera, in XML 1.0 format. Point correspondences between image $[u, v]$ and pattern $[x, y, z = 0]$ spaces are listed in the <***c***> tags within this file.
- ***rear_calibration.xml*** The analogous file to the ***front_calibration.xml*** file, for the rear plane position.
- ***front_plane.png*** PNG image of the front calibration plane as seen by the camera, lit only by the calibration headlight.
- ***rear_plane.png*** Same as ***front_plane.png*** but for the rear plane position.
- ***background.png*** PNG image of the capture volume without any fluorescing fluid. This is subtracted from the captured images to obtain the measurement vector.
- ***frame.XXXXXX.png*** PNG images of the capture itself, prior to background subtraction.