QUANTIFYING 3D POSITIONAL UNCERTAINTY OF RADIOLOGICAL MATERIAL FROM NUCLEAR DETONATION VIDEO

Daniel T. Schmitt, Robert Slaughter, Gilbert L. Peterson*

The Air Force Institute of Technology

Dept. of Elec. and Comp. Eng.

WPAFB, OH 45433

(937) 255-3636

daniel.schmitt@afit.edu

robert.slaughter@afit.edu

gilbert.peterson@afit.edu

Abstract

Multi-view geometry theory is applied to atmospheric nuclear tests filmed in the 1950s and 1960s, to estimate the three-dimensional locations of fixed structures and radiological material during the detonation. Results show that using bundle adjustment, points can be estimated with an average uncertainty of 0.68 meters with 0.36, 0.28, and 0.37 meters of uncertainty in x-, y-, and z-directions respectively.

Index Terms

*The views expressed in this paper are those of the authors, and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government. The authors would like to thank the Defense Threat Reduction Agency and the National Nuclear Security Administration for funding this research. This document has been approved for public release.

Atmospheric Nuclear Detonation, 3D reconstruction

I. Introduction

To determine which nuclear-capable nation provided the material in a nuclear bomb an analysis of the detonation's yield is accomplished. The detonation's yield corresponds to how enriched the nuclear material was, and further relates the material to known enrichment capabilities of nations. Accurate yield determination of a detonation helps to attribute the nuclear material's source.

Estimating the yield solely from a one-dimensional radius suffers from the ability to accurately determine the radius and has high uncertainty. Eq. 1 shows Taylor's equation [1] where R represents the radius, t is the time since detonation, p_o is the density of the medium. Any uncertainty in measuring R will compound by a factor of 5 when estimating the yield. Substituting the equation of the volume of a sphere $V = \frac{4\pi}{3}r^3$ solved for r builds Eq. 2. Presuming that a volume could be generated with similar uncertainty to that of a radius, Eq. 2 demonstrates that uncertainty in the volume of the equation lead to an uncertainty of yield by a factor of 5/3 instead of one generated from a radius with an uncertainty of 5. The error in estimating the yield of a nuclear detonation can be reduced significantly by transitioning from a radius-based model [1] to a volume-based model.

$$Yield \sim \frac{R^5 p_o}{t^2} \tag{1}$$

$$Yield \sim \frac{V^{\frac{5}{3}} p_o}{t^2} \tag{2}$$

To use a volume-based model requires a method to estimate the volume with sufficient accuracy. Volume-based models exist for nuclear detonations [2] [3], but there is no current method for verifying their accuracy because no ground truth exists for the volume of nuclear detonations. Establishing ground truth volume measurements for nuclear detonations and their interaction with the atmosphere and other structures is an important step in building a reliable volume-based model. The establishing of ground truth points in three dimensional space of atmospheric nuclear detonations is the primary contribution of this paper.

Establishing ground truth of points on a nuclear detonations in the atmosphere necessitates using data of atmospheric nuclear detonations. The United States conducted and filmed over 200

atmospheric nuclear tests during the 1950s and 1960s before it was prohibited by treaty. The films are being digitized by Lawrence Livermore National Labs (LLNL) [4] for preservation purposes. The digitization of the nuclear detonation films also enables the revalidation of fundamental nuclear explosion models using modern computer vision techniques. Applicable computer vision methods include applying multi-view geometry (MVG) theory to the films [2]. By applying MVG to the films of past nuclear detonations, a three-dimensional model could be established improving knowledge about nuclear detonation (NUDET) interactions in the atmosphere.

One popular MVG method for uncalibrated images (images where intrinsic and extrinsic parameters are unknown) uses Scale-Invariant Feature Transform (SIFT) [5] feature detection and matching [6]. For pre-calibrated cameras, 3D reconstructions can be generated from images with multi-view stereo (MVS) [7]. The NUDET films are not calibrated by today's standards, but have some information such as focal length and manually computed camera pose data. Even with what is available about the films, they present additional challenges for 3D reconstruction in three areas. First, is that the subject of the film is a non-rigid object (a radiological gas) that is feature poor, and lacks feature correspondences. Second, the films were recorded using cameras that were not synchronized in time, which makes matching frames from varied viewpoints more difficult when there is uncertainty of correct time alignment. Lastly, only a few viewpoints were recorded for each detonation and the viewpoints have wide viewpoint angles of $65-110^o$ that are outside the documented capabilities of SIFT's match repeatability [5]. Because of these issues, manually generated point correspondences are the current method for producing 3D structure for the NUDET data.

The development of an error measure for the 3D points generated by MVG requires a comparison to ground truth. Ground truth only exists on rigid bodies that are in view of the film capture. Fortunately, some of the NUDETs were detonated on the top of towers that are visible early in the detonation films.

This paper presents a method that applies MVG to films of atmospheric nuclear detonations to be used to estimate the location of 3D points of the detonation. To accomplish this, corresponding points were identified in three detonations at three different points in time on the support tower (Fig. 1) that the nuclear device was placed on. The corresponding points in multiple films and viewpoints were used to estimate the location of points in 3D-space using multi-view geometry theory. The reconstructed points of the tower were then compared to the ground

truth of the tower's location and dimensions to determine the reconstructive error otherwise known as the uncertainty of the points in space. This error was then inferred on other points in the reconstruction on the detonation where the ground truth is not known. A total of 69 reconstructions were built to collect the statistical error. The end result is the ability to estimate, with an average locational error of 0.68 meters with 0.36, 0.28, and 0.37 meters of locational error in x-, y-, and z-directions respectively, the position of radiological material during the detonation and post-detonation in the plume.

Improving the uncertainty of yield determination on nuclear detonation requires a more accurate representation of the nuclear detonation data than a statistical model of radii growth over time. A three-dimensional model and volume-base model reduces the uncertainty in estimating the yield. Creating sets of three-dimensional data on the foundational nuclear detonations provides the path to establishing three-dimensional model. This paper provides the methodology to create three-dimensional data sets on nuclear detonations by using established MVG theory enabling the three-dimensional study of nuclear detonations in the atmosphere. It provides a baseline to look at nuclear detonations in a new way by considering their three-dimensional interactions, rather than just a growth in radii over time.

II. Background

This section provides background on the NUDET films, describes the camera models and the processes of 3D reconstruction from images using triangulation and bundle adjustment.

A. Nuclear Detonation Films

From 1945-1962 the United States conducted a total of 216 atmospheric, underwater, and space nuclear detonation tests [4]. Of the atmospheric tests, some detonations were airdropped, suspended from balloons, placed on towers, ground bursts, and ship-based. Atmospheric testing was then banned following the signing of the Limited Test Ban Treaty. Much of the data on atmospheric nuclear weapons effects was produced during this 17 year period of testing leading up to the treaty.

During experimentation approximately 9,000 films were recorded of the atmospheric nuclear tests. The films were captured using different cameras by Edgerton, Germashausen & Grier, Inc.

(EG&G) using 8mm, 16mm, and 35mm film [4]. The technical films were used to capture early fireball growth, early plume rise, and late plume rise following detonation.

The type of film used most commonly was Microfile (MF) [4]. MF was developed by Kodak for the nuclear testing program and was used because of its ability to retain its chemical properties in high radiation environments and because of its high resolution. In addition, the majority of recordings were acrylate- or polyethylene-based black and white films which have an expected lifespan of 100-500 years according to ANSI Standard IT9.11-1992.

Each explosion site was carefully documented with multiple collection sites at defined locations around the explosion. Each collection site was outfitted with multiple film reels to record the explosions (Fig. 2). For each explosion, the number of collection sites ranged from 3-10. The end result was that there were about 20 videos recorded for each explosion, some from the same point of view, some with small angular differences of around 10^o and others of larger angular differences of $65 - 110^o$.

B. Homogeneous coordinates

In MVG, homogeneous coordinates are used so that points in the real world can be related to points in multiple images more easily. A point in Euclidean 3-space is represented by an ordered pair of real numbers (x, y, z). To make the coordinate homogeneous, a fourth digit is added to create the homogeneous coordinate definition that $(x, y, z) = (x, y, z, 1)$. The definition also implies that $(x, y, z) = (x, y, z, 1) = (2x, 2y, 2z, 2) = (kx, ky, kz, k)$ for any non-zero value of k. The parameter k is then sometimes referred to as a scaling factor that facilities relating images together that have different scale.

C. Camera Models

To apply multi-view geometry to the NUDET films, information recorded at the time of detonation must be interpreted using camera models. This section details the theory of the pinhole camera, which underpins the finite projective camera model. The finite projective camera is then used in Sec. III with information from the NUDET films with triangulation and bundle adjustment to create the 3D structure of the NUDET with points on the NUDET tower.

1) Pinhole Camera: The pinhole camera model is a basic model of how images are captured by a camera. Fig. 3 shows a depiction of the pinhole camera model. Letting $\mathbf{X}$ represent a point in the 3-dimensional world represented by $\mathbf{X} = (X, Y, Z)^T \in \mathbb{R}^3$ in Euclidean space. The camera center (optical center), C is located at the origin of this Euclidean space. The image plane representing the image is in the positive Z direction. The principal axis intersects the image plane in the center of the image at point $\boldsymbol{p}_0$. The image plane has an $x-$ and $y-$axis orthogonal to each other and the principal axis. Point $\mathbf{X}$ is then projected on the image plane at point $\mathbf{x}$ which is on a ray from optical center C toward point $\mathbf{X}$.

Represented in homogeneous coordinates, $\mathbf{X} = (X, Y, Z)^T = (fX/Z, fY/Z, f)^T$ where f is the focal length of the image. Another way to consider this is the captured 2D image of the 3D world is a transform from $\mathbb{R}^3 \mapsto \mathbb{R}^2$ which can also be written as a matrix multiplication in Eq. 3.

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} fX \\ fY \\ Z \end{pmatrix} = \begin{bmatrix} f & & & 0 \\ & f & & 0 \\ & & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \tag{3}$$

Now, the bracket matrix in Eq. 3 is given a name, $\mathbf{P}$, to represent the 3×4 camera projection matrix. With $\mathbf{x}$ representing the 2D image point, and $\mathbf{X}$ representing the 3D world point, the equation is then defined as in Eq. 4.

$$\mathbf{x} = \mathbf{PX} \tag{4}$$

The expression in Eq. 3 assumes that the origin of the image plane is the principal point ($\mathbf{p_0}$), which may not always be the case. As a result, it is generalized that a shift (translation) can occur to set the principal point in the middle resulting in the mapping

$$(X, Y, Z)^T \mapsto (fX/Z + p_x, fY/Z + p_y)^T$$

where (p_x, p_y) is the coordinates of the principal point. The resulting change to Eq. 3 is then

expressed in Eq. 5.

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} fX + Zp_x \\ fY + Zp_y \\ Z \end{pmatrix} = \begin{bmatrix} f & & p_x & 0 \\ & f & p_y & 0 \\ & & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \tag{5}$$

Next, the camera calibration matrix ($\mathbf{K}$) is established where

$$\mathbf{K} = \begin{bmatrix} f & & p_x \\ & f & p_y \\ & & 1 \end{bmatrix}. \tag{6}$$

Written concisely Eq. 4 can be rewritten as shown in Eq. 7 as the definition of $\mathbf{P}$ is expanded since $\mathbf{P} = \mathbf{K}[\mathbf{I}|\mathbf{0}]$. In this case, $\mathbf{I}$ is the identity matrix and $\mathbf{0}$ is the zero vector as the last column in the $\mathbf{P}$ matrix. $\mathbf{X}_{cam}$ is the same $\mathbf{X}$ in world coordinates with the emphasis that it is expressed in the camera coordinate frame. The difference between $\mathbf{X}_{cam}$ and $\mathbf{X}$ is the orientation of the camera which can be defined in terms of rotation and translation. When the $\mathbf{K}$ matrix of a camera is known, it is said to be calibrated.

$$\mathbf{x} = \mathbf{K}[\mathbf{I}|\mathbf{0}]\mathbf{X}_{cam} \tag{7}$$

2) Rotation and Translation: Fig. 4 shows the transformation from $\mathbf{X}_{cam}$ to $\mathbf{X}$ world coordinates. Letting $\tilde{\mathbf{X}}$ represent a 3-vector representing the coordinates of a point in the world coordinate frame, and $\tilde{\mathbf{X}}_{cam}$ represent that point in the camera coordinate frame, the relationship $\tilde{\mathbf{X}}_{cam} = \mathbf{R}(\tilde{\mathbf{X}} - \tilde{\mathbf{C}})$ is asserted where $\tilde{\mathbf{C}}$ is the coordinates of the camera center in the world coordinate frame. $\mathbf{R}$ is then a 3×3 rotation matrix. This relationship can be stated verbosely as in Eq. 8, or concisely as in Eq. 9.

$$\mathbf{X}_{cam} = \begin{bmatrix} \mathbf{R} & -\mathbf{R}\tilde{\mathbf{C}} \\ 0 & 1 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = \begin{bmatrix} \mathbf{R} & -\mathbf{R}\tilde{\mathbf{C}} \\ 0 & 1 \end{bmatrix} \mathbf{X} \tag{8}$$

$$\mathbf{x} = \mathbf{K}\mathbf{R}[\mathbf{I}|\tilde{\mathbf{C}}]\mathbf{X} \tag{9}$$

It is also apparent that the definition of $\mathbf{P}$ has expanded where $\mathbf{P} = \mathbf{K}\mathbf{R}[\mathbf{I}|\tilde{\mathbf{C}}]$. Looking inside of $\mathbf{P}$ we can see that it has 9 degrees of freedom: 3 for elements in $\mathbf{K}$ (f, p_x, p_y), 3 for $\mathbf{R}$,

and 3 for $\tilde{\mathbf{C}}$. $\mathbf{K}$ values are determined by the camera and are referred to as internal camera parameters or intrinsics, while $\mathbf{R}$ and $\tilde{\mathbf{C}}$ refer to the camera's orientation and position in the world coordinate system and are referred to as the external parameters or extrinsics.

3) CCD camera model: The CCD (Charged Coupled Device) camera is an extension from the pinhole camera that allows for different scales in each axial direction. The CCD camera model applies to the digitized images of the NUDET data because the images are digitized using a scanner that interprets information like a CCD camera. With image coordinates often measured in pixels, this can have the effect of unequal scale factors in each direction. As a result, it is necessary to expand the pinhole model. If the number of pixels per unit distance in image coordinates are m_x and m_y in the x and y directions, then the transformation from world coordinates to pixel coordinates is found by multiplying by a factor of diag$(m_x, m_y, 1)$. This changes the formula for the camera intrinsics ($\mathbf{K}$) as shown in Eq. 10. The new values are defined in the following manner: $\alpha_x = fm_x$, $\alpha_y = fm_y$, $x_0 = m_x p_x$, and $y_0 = m_y p_y$.

$$\mathbf{K} = \begin{bmatrix} \alpha_x & & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix} \tag{10}$$

4) Finite projective camera: Further generalizing the CCD camera, the parameter s for skew can be added to the CCD model for camera intrinsics.

$$\mathbf{K} = \begin{bmatrix} \alpha_x & s & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix} \tag{11}$$

In most cases s is zero for normal cameras. The cases where $s \neq 0$ occur when pixel elements are skewed in the CCD array. This can occur with the $x-$ and $y-$axis are not perpendicular. It can also occur when an image is taken of an image, like when a photograph is re-photographed, or a negative of an image is enlarged.

D. Triangulation

Triangulation is the process of determining a point's 3D position from a set of corresponding image locations and known camera positions. Two processes are described here, the mid-point method and the direct linear transform method. These methods, as well as other methods, are discussed and compared in [9].

1) Mid-point Method: Fig. 5 shows a diagram of the triangulation point estimation [10]. For a given image, $\mathbf{R}$ represents the rotational matrix of the image pose, $\mathbf{c}$ represents the camera center, $\mathbf{x}$ represents the 2D image location of the feature, $\hat{\mathbf{v}}$ is the direction of the ray from $\mathbf{c}$ to $\mathbf{x}$, and finally $\mathbf{q}$ is the point along $\hat{\mathbf{v}}$ that is closest to $\mathbf{p}$ and is d distance away from $\mathbf{x}$. Assuming that there are i cameras that have correspondence matches for feature $\mathbf{x}$ which gives a set of $\mathbf{x}_i$ for that feature. Based on the discussion in Sec. II-C2, there is then a set of cameras $\mathbf{P}_i = \mathbf{K}_i[\mathbf{R}_i|\mathbf{t}_i]$, where $\mathbf{t}_i = -\mathbf{R}_i\mathbf{c}_i$ and $\mathbf{c}_i$ is the i-th camera center. In Fig. 5, the rays originate at $\mathbf{c}_i$ in the direction $\hat{\mathbf{v}}_i = N(\mathbf{R}_i^{-1}\mathbf{K}_i^{-1}\mathbf{x}_i)$ where N represents the unit norm function. The nearest point to $\mathbf{p}$ on the ray is $\mathbf{q}_i$ which exists at $\|\mathbf{c}_i + d_i\hat{\mathbf{v}}_i - \mathbf{p}\|^2$ where d_i is the geometric distance and therefore has a minimum at $d_i = \hat{\mathbf{v}}_i \cdot (\mathbf{p}\text{-}\mathbf{c}_i)$. It follows that

$$\mathbf{q}_i = \mathbf{c}_i + (\hat{\mathbf{v}}_i\hat{\mathbf{v}}_i{}^T)(\mathbf{p} - \mathbf{c}_i) = \mathbf{c}_i + (\mathbf{p} - \mathbf{c}_i)_{\|},$$

and the squared distance (r_j^2) between $\mathbf{p}$ and $\mathbf{q}_i$ is

$$r_i^2 = \|(\mathbf{I} - \hat{\mathbf{v}}_i\hat{\mathbf{v}}_i^T)(\mathbf{p} - \mathbf{c}_i)\|^2 = \|(\mathbf{p} - \mathbf{c}_i)_{\perp}\|^2.$$

The optimal value for $\mathbf{p}$, is computed as a regular least squares problem by summing over all the r_i^2 and finding the optimal $\mathbf{p}$ using Eq. 12.

$$\mathbf{p} = \left[\sum_i (\mathbf{I} - \hat{\mathbf{v}}_i\hat{\mathbf{v}}_i^T)\right]^{-1} \left[\sum_i (\mathbf{I} - \hat{\mathbf{v}}_i\hat{\mathbf{v}}_i^T)\mathbf{c}_i\right] \tag{12}$$

While the mid-point method is straight-forward to comprehend, it is just one method of triangulation and is criticized for being inaccurate [9]. The direct linear transform method is considered to be more statistically optimal [10] and is discussed next.

2) Direct Linear Transform method: Consider two images $\mathbf{x}$ and $\mathbf{x}'$ and their camera matrices $\mathbf{P}$ and $\mathbf{P}'$. Written in homogeneous coordinates, $\mathbf{x} = w(x, y, 1)^T$ where (x, y) are observed point coordinates and w is an unknown scale factor. We also have the case by Eq. 4 where $\mathbf{x} = \mathbf{PX}$ and $\mathbf{x}' = \mathbf{P}'\mathbf{X}$. Calling $\mathbf{p}^{iT}$ the rows of $\mathbf{P}$, then $\mathbf{x} = \mathbf{PX}$ can be rewritten as

$$wx = \mathbf{p}^{1T}\mathbf{X}, wy = \mathbf{p}^{2T}\mathbf{X}, w = \mathbf{p}^{3T}\mathbf{X}.$$

Using the third equation, the value of w is substituted in the first two equations arriving at

$$\begin{aligned} x\mathbf{p}^{3T}\mathbf{X} &= \mathbf{p}^{1T}\mathbf{X} \\ y\mathbf{p}^{3T}\mathbf{X} &= \mathbf{p}^{2T}\mathbf{X}. \end{aligned} \tag{13}$$

Repeating this for $\mathbf{x}' = \mathbf{P}'\mathbf{X}$ similar equations are arrived at for $\mathbf{x}'$ and $\mathbf{P}'$

$$\begin{aligned} x'\mathbf{p}'^{3T}\mathbf{X} &= \mathbf{p}'^{1T}\mathbf{X} \\ y'\mathbf{p}'^{3T}\mathbf{X} &= \mathbf{p}'^{2T}\mathbf{X}. \end{aligned} \tag{14}$$

Eq. 13 and 14 are then solved for 0 arriving at

$$x(\mathbf{p}^{3T}\mathbf{X}) - (\mathbf{p}^{1T}\mathbf{X}) = 0$$

$$y(\mathbf{p}^{3T}\mathbf{X}) - (\mathbf{p}^{2T}\mathbf{X}) = 0$$

$$x'(\mathbf{p}'^{3T}\mathbf{X}) - (\mathbf{p}'^{1T}\mathbf{X}) = 0$$

$$y'(\mathbf{p}'^{3T}\mathbf{X}) - (\mathbf{p}'^{2T}\mathbf{X}) = 0.$$

These equations can be combined into a form $\mathbf{A}\mathbf{X} = \mathbf{0}$ by factoring out $\mathbf{X}$ with $\mathbf{A}$ shown in Eq. 15 [9].

$$\mathbf{A} = \begin{bmatrix} x\mathbf{p}^{3T} - \mathbf{p}^{1T} \\ y\mathbf{p}^{3T} - \mathbf{p}^{2T} \\ x'\mathbf{p}'^{3T} - \mathbf{p}'^{1T} \\ y'\mathbf{p}'^{3T} - \mathbf{p}'^{2T} \end{bmatrix} \tag{15}$$

This solution of equations can be solved using Singular Value Decomposition that relates to the smallest singular value of $\mathbf{A}$. The solution to $\mathbf{A}$ with known $\mathbf{P}$ then gives the 3D points in $\mathbf{X}$.

E. Bundle Adjustment

Bundle adjustment is an accurate way to recover structure and motion simultaneously for a three-dimensional reconstruction [10]. Bundle adjustment achieves these improvements by simultaneously calculating structure and camera poses. It can also scale for large data sets of tens of thousands of images [6].

Bundle adjustment performs the 3D triangulation from a set of three-dimensional points $\mathbf{X}_j$ which is in view by a set of cameras with matrices $\mathbf{P}^i$. Denoting $\mathbf{x}_j^i$ as the coordinates of the j-th point seen by the i-th camera. By Eq. 4, $\mathbf{P}^i\mathbf{X}_j = \mathbf{x}_j^i$ is stated. If measurements in the image are noisy then it is not likely that $\mathbf{P}^i\mathbf{X}_j = \mathbf{x}_j^i$ has an exact solution.

A Maximum Likelihood solution is then sought after assuming the noise is Gaussian with the goal of estimating projection matrices $\hat{\mathbf{P}}^i$ and points $\hat{\mathbf{X}}_j$ which project exactly to the image

points $\hat{\mathbf{x}}_j^i$ which also minimizes the distance between the reprojected point and measured image points $\mathbf{x}_j^i$ for every view in which the point appears. The reprojection used in Eq.16 [8] where $d(\mathbf{x}, \mathbf{y})$ is the geometric image distance between homogeneous points $\mathbf{x}$ and $\mathbf{y}$.

$$\min_{\hat{\mathbf{P}}^i, \hat{\mathbf{X}}_j} \sum_{ij} d(\hat{\mathbf{P}}^i \hat{\mathbf{X}}_j, \hat{\mathbf{x}}_j^i)^2 \tag{16}$$

Minimizing the reprojection error is not without uncertainty. Past work [11] [12] has narrowed the uncertainty to be associated with the 3D depth of a scene and the camera motion. While bundle adjustment can be used without knowing the locations of cameras, the uncertainty is minimized when the initial estimate of camera positions is as correct as possible. Using the recorded camera positions in the NUDET films as initial camera pose positions, minimizes the negative effect of improper initialization which is the largest contributor to uncertainty in reconstructions. From the initial pose projection, the structure and pose parameters are then improved iteratively with bundle adjustment using the multiple point correspondences. To accomplish bundle adjustment, existing software packages are available [13].

III. Methodology

The determination of the uncertainty of points in 3D space generated from triangulation and bundle adjustment using the NUDET films provides a baseline by which other three-dimensional models of the detonation can be measured. To accomplish this, two reconstructions were generated for each combination of two points of view. One reconstruction used the triangulation method that solved for 3D points using singular value decomposition as described in Sec. II-D2. The second reconstruction was run through a bundle adjustment algorithm to refine the camera intrinsics and extrinsics and create a bundle adjusted reconstruction.

To quantify the error, 3D reconstructions were accomplished on tower detonations. Points were manually selected to be on different floors of the tower. The locations and size of the towers was known, so the reconstruction was compared to ground truth. The error was then the difference between the reconstructed point and the ground truth.

Ground truth was established by the survey sheets of the detonation which depicted ground zero within a foot's accuracy. Total tower height was also known for each detonation (Table I), and by counting the number of floors, the height per floor was determined for each detonation.

TABLE I

HEIGHT OF TOWER DETONATIONS [4].

Detonation	Tower Height (ft)	Floors
Tesla	308	25
Turk	508	35
MET	408	30

The width of the tower was documented as well. Once the reconstruction was complete, the reconstruction was rotated to match that of the ground truth. Each reconstruction yielded ten points on the tower that were used to calculate the error as the differences between the point and the ground truth. In addition to the total distance, the distance in x,y, and z dimensions was calculated. The mean and variance was then calculated from the collection of points for all the reconstructions. The process was repeated three times for three different points in time on each detonation.

A. Workload

The requirements for the workload for this test were that the films had to be of a tower detonation, with multiple camera angles, and have already been digitized. At the time of testing for this project, six detonations met this criteria, three from Operation Teapot (Tesla, Turk, and MET), and three from Operation Plumbbob (Boltzmann, Galileo, Kepler). The number of films and potential reconstructions of the films are shown in Table II. To select the source data for this experiment, the digitized films and viewpoints were analyzed. Reliable matching tower correspondences could not be found in Boltzmann because of the 110^o viewpoint change between points of view. Detonations Tesla, Turk, and MET were chosen because of the larger number of films and combinations of reconstructions that could be made. By selecting these three detonations, 77% of the available reconstructions were tested on 65% of the available films. With the workload chosen as three detonations (Tesla, Turk, and MET), at three different points in time, all available films were tested pair-wise across two locations for a total of 69 $(Tesla(9) + Turk(10) + MET(4)) \times 3\ timestamps)$ tests accomplished for both triangulation and bundle adjustment.

TABLE II

AVAILABLE DIGITIZED TOWER DETONATIONS [4].

Operation	Detonation	Films Loc. 1	Films Loc. 2	Num. Reconstr.
Teapot	Tesla	3	3	9
Teapot	Turk	2	5	10
Teapot	MET	2	2	4
Plumbbob	Boltzmann	2	2	4
Plumbbob	Galileo	1	2	2
Plumbbob	Kepler	1	1	1

B. Estimating Camera Intrinsics ($\mathbf{K}$)

To apply triangulation and bundle adjustment, the films must be calibrated using the information that was documented when they were recorded. To calculate the camera intrinsic values ($\mathbf{K}$) for the NUDET data, each of the parameters of $\mathbf{K}$ must be determined. Remembering the portions of $\mathbf{K}$ from Sec. II-C3 as

$$\mathbf{K} = \begin{bmatrix} \alpha_x & s & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix}$$

and that $\alpha_x = fm_x$, $\alpha_y = fm_y$, $x_0 = m_x p_x$, and $y_0 = m_y p_y$ where f is the focal length, m_x and m_y are the pixels per unit distance in image coordinates in the $x-$ and $y-$directions, and p_x and p_y are the coordinates of the principal point, and s is the skew.

For now, the assumption will be made that the skew is zero ($s = 0$). The focal length is also recorded in the original documentation [4]. This leaves four parameters that need to be determined (m_x, m_y, x_0, and y_0).

1) Estimating m_x and m_y: The parameters m_x and m_y, are determined by relating the pixel dimensions of the scanned film to the original 16mm film. Fig. 6 shows the dimensions of the film that was used in the Tesla filming [4]. All of the images used were perforated two edges with Eastman Kodak short perforation (2R-2994). The dimensions between perforations width-wise (distance F in Fig. 6) were used to determine the x dimensions. The dimensions between each perforation vertically (distance B in Fig. 6) were used to determine the y dimensions. Samples

of each video were taken, and the distance measured in the images in pixel length and related to the pixel lengths in the $x-$ and $y-$direction (10.49 mm and 7.605 mm respectively). As a result, an equation can be set for m_x and m_y.

$$m_x = \frac{horizontal\ pixel\ length\ between\ perforations}{10.49mm}$$
$$m_y = \frac{vertical\ pixel\ length\ between\ perforations}{7.605mm}$$

Measurements were taken in the $x-$ and $y-$direction and the pixel difference was averaged for each film. To determine the vertical distance, subsequent frames were stacked vertically to recreate the continuous feel of the original film. When this was accomplished it was determined that there was some overlap between subsequent images of 10 pixels which could be visually determined by repetition in the pixel pattern. After this 10 pixel overlap was removed from the $y-$distance the results were accurately computed and are shown in Table III.

TABLE III

RESULTS OF ESTIMATING m_x AND m_y. TOLERANCE FOR m_x WAS 0.33 AND TOLERANCE FOR m_y WAS 0.19.

Film	m_x ($\frac{pixels}{mm}$)			m_y ($\frac{pixels}{mm}$)		
	$\bar{m_x}$	2σ	Max Err.	$\bar{m_y}$	2σ	Max Err.
02	114.59	0	0.328	114.67	0.15	0.326
03	114.68	0	0.328	113.93	0.15	0.345
04	114.49	0	0.327	113.87	0	0.194
06	114.49	0	0.327	113.80	0.15	0.345
07	114.68	0	0.328	113.80	0.15	0.345
08	114.49	0	0.327	114.26	0	0.194
09	114.49	0	0.327	113.80	0.15	0.345
10	114.49	0	0.327	114.26	0	0.194
11	114.49	0	0.327	114.26	0	0.194
12	114.49	0	0.327	113.83	0.15	0.345
16	114.49	0	0.327	114.26	0	0.194
17	114.49	0	0.327	113.80	0.15	0.345
18	114.82	0.14	0.462	113.80	0	0.194

Table III shows that there is very little variation in the measurement of $x-$distance. The maximum error was calculated based on the worse case scenario that both the measurement

error (σ) and the tolerance of the perforation of the film were additive when in fact the error is likely that it is the larger of these two errors. With this in consideration, all the variance in measurement is within tolerance of the production of the film perforations.

On the other hand, there appears to be evidence of skew. The 95% confidence interval of m_x is (114.347, 114.772) while the 95% confidence interval of m_y is (113.46, 114.59). A two-sample t-test with unequal variance gives the probability that these two groups are from the same population with a probability of 0.00001. It's possible that either the film shrunk in the vertical direction, or something in the digitization process of these image is shortening them vertically, or lengthening them horizontally. Regardless, values for m_x and m_y are now known.

2) Estimating x_0 and y_0: Estimating x_0 and y_0 is preferable above estimating p_x and p_y because x_0 and y_0 units are in terms of pixels and can be determined from the scanned image, while p_x and p_y units are in terms of mm, and would require analyzing the films exactly, or first estimating x_0 and y_0 and extracting p_x and p_y from their m_x and m_y relationships.

The parameters x_0 and y_0 are intended to represent the center of the image plane. The film has borders on the left and right for the perforation of the film. By removing the perforated areas, the actual area that contains images can be isolated. This can remove any error that would result in horizontal misalignment of the film when they were scanned. Estimating x_0 and y_0 can be accomplished by choosing the center of the viewable portion of the image.

Prior to determining a center, a standard image crop was determined that removed the perforated edges off the left and right side. To accomplish this, all the pixel values of all the images of several detonations were summed column-wise. This method was chosen because the interior of the perforated edges had relatively dark vertical lines and summing the columns of images would provide a complete summary of where these areas were on the images. Then the local minimums on either side of the plot was used as the cropping in the x-direction. This plot is shown in Fig. 7. The end result is that cropping the image at 352, 1490 would be the best standard cropping dimensions.

3) Estimating $\mathbf{K}$: All the parameters are in place to estimate $\mathbf{K}$, the camera intrinsic matrix. The unit values of $\mathbf{K}$ are in terms of pixels. Referring to Table III, the extimates of x_0 and y_0 from the center of the viewable image, and the focal lengths in the load sheets [4], $\mathbf{K}$ can now be estimated. Eq. 17 shows $\mathbf{K}$ for Tesla camera 2 [4] using the center of the viewable area as

the principal point.

$$\begin{aligned}\mathbf{K} &= \begin{bmatrix} \alpha_x & s & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix} = \begin{bmatrix} f*m_x & s & x_0 \\ & f*m_y & y_0 \\ & & 1 \end{bmatrix} \\ &= \begin{bmatrix} 102.1*114.585 & 0 & 561 \\ & 102.1*114.436 & 441.5 \\ & & 1 \end{bmatrix} \\ &= \begin{bmatrix} 11699 & 0 & 561 \\ & 11683 & 441.5 \\ & & 1 \end{bmatrix}\end{aligned} \tag{17}$$

C. Estimating $\mathbf{P}$

With $\mathbf{K}$ estimated, $\mathbf{P}$ can be estimated. Recalling that $\mathbf{P} = \mathbf{KR}[\mathbf{I}|\tilde{\mathbf{C}}]$ (Sec. II-C2), $\mathbf{K}$ is known, $\mathbf{I}$ is the identity matrix, only $\mathbf{R}$ and $\tilde{\mathbf{C}}$ need to be identified.

$\mathbf{R}$ is the rotation of the camera that can be defined from the bearing (azimuth) from the ground zero location and the camera locations. $\tilde{\mathbf{C}}$ is the camera locations relative to ground zero in world coordinate space.

Both $\mathbf{R}$ and $\tilde{\mathbf{C}}$ are determined from data recorded at the time that the NUDET tests were accomplished in the survey data. The relative coordinates $(\Delta N, \Delta E, \Delta Z)$ from ground zero of the survey can be used for $\tilde{\mathbf{C}}$. $\mathbf{R}$ can be determined by using the angles (bearing and tilt). With θ representing bearing, ϕ representing tilt, and $\psi = 0$ the rotation matrix ($\mathbf{R}$) can be represented using a directional cosine matrix as show in Eq. 18.

$$\mathbf{R} = \begin{bmatrix} \cos\theta\cos\psi & a & b \\ -\cos\theta\sin\psi & c & d \\ \sin\theta & -\sin\phi\cos\theta & \cos\phi\cos\theta \end{bmatrix}$$

$$\begin{aligned} a &= \cos\phi\sin\psi + \sin\phi\sin\theta\cos\psi \\ b &= \sin\phi\sin\psi - \cos\phi\sin\theta\cos\psi \\ c &= \cos\phi\cos\psi - \sin\phi\sin\theta\sin\psi \\ d &= \sin\phi\cos\psi + \cos\phi\sin\theta\sin\psi \end{aligned} \tag{18}$$

IV. Results

Fig. 8 shows one example of manual point correspondences between two images of Tesla which has a 72^o angle between viewpoints. This can be compared to the reconstruction shown in Fig. 9. The number of points on the tower were similar to the number of points on the detonation. This was deliberate so the reconstruction would be equally weighted by both points on the tower and on the detonation. The points that were chosen on the detonation were chosen for their ability to be accurately matched from alternate points of view, and provided a variety of features including bright spots, folds in the detonation, and endpoints of cone protrusions.

Ten points were reconstructed of the tower for each triangulation and bundle adjusted reconstruction. Each point was then compared to the ground truth location of where the point should be. The error of each point in the Euclidean distance between the ground truth location of the points of the tower and their estimated position in the reconstruction. The error was then analyzed by detonation and isolating the x-, y-, and z-dimensions.

Table IV shows the average error using triangulation and bundle adjustment with N representing the number of points. The average error using triangulation was 0.83 meters and the average error using bundle adjustment was 0.68 meters. The expectation was that bundle adjustment would improve (decrease) the error of reconstruction and that was demonstrated in each case, but without statistical significance. A 95% confidence interval on triangulation has the true mean in the range of (0 ... 2.25) meters. A 95% confidence interval on bundle adjustment has the true mean in the range of (0 ... 1.7) meters.

TABLE IV

AVERAGE ERROR OF POINT CLOUD RECONSTRUCTION.

Detonation	Triangulation (m)		Bundle Adj. (m)		N
	AVG	σ	AVG	σ	
Tesla	0.56	0.65	0.54	0.59	270
Turk	0.76	0.35	0.68	0.29	300
Met	1.62	0.93	0.98	0.59	120
Total	0.83	0.71	0.68	0.51	690

TABLE V

DIMENSION OF ERROR OF POINT CLOUD RECONSTRUCTION.

Dimension	Triangulation (m)		Bundle Adj. (m)	
	AVG	σ	AVG	σ
X	0.46	0.65	0.36	0.38
Y	0.27	0.24	0.28	0.25
Z	0.44	0.49	0.37	0.42
Total	0.83	0.71	0.68	0.51

Table V shows the error with regard to $x-$, $y-$, and $z-$dimension. The dimensionality of the error is pretty evenly spread between dimensions. This is expected because the viewpoint angles of the detonations are close to 90^o (Tesla 72^o, Turk 63^o, MET 68^o). If the angles were smaller than 45^o, the expectation would be that the $z-$dimension would dominate the error. On the bundle adjustment results, a 95% confidence interval is $(0 \ldots 1.12)$ meters in the x-direction, $(0 \ldots 0.78)$ meters in the y-direction, and $(0 \ldots 1.21)$ in the z-direction.

Ground truth of the locations of points on the detonations are not known. To estimate the locations of points on detonations the techniques of triangulation and bundle adjustment were applied to points on a detonation and points on the detonation's tower where ground truth was known. Since the process of choosing points and the technique applied in reconstruction are consistent for the reconstruction, the error associated with the points with known ground truth can be inferred on the points where ground truth is not known. The locational error of the points of the tower then become the locational error of the points on the detonation.

V. Discussion

The results shown in Table V provides the information necessary to estimate the spacial accuracy of points in an atmospheric nuclear detonation. With this new information, 3D point correspondences can be generated over a period or region of interest in a detonation enabling the study of regional areas of a detonation whose behavior is different than existing models. These point correspondences can be used to test the accuracy of other models by comparing the models to the point correspondence during that period in time.

One example of using this information for further research would be to track how points on the detonation change over time. The change of location of the same points over time can be used to model how the detonation propagates in complex areas of the detonation, where the detonation interacts with objects or changes in atmospheric density. Using data assimilation [14] to fuse established mathematical models with the nuclear images and how they change over time would be a logical way to improve the accuracy of estimating the positional accuracy.

Another example would be to verify an estimate of volume generated from a single image. An assumption could be made that since the edges of a detonation appear spherical, then the entire detonation is spherical. The results of a bundle adjusted reconstruction can generate points on the detonation recreating the depth component of the detonation at those points. The points can then be compared to a spherical rendering of the detonation to determine how accurate the spherical assumption is. If the edges of the sphere are within the error of reconstruction, the assumption would be confirmed as acceptable. Further assumptions could be made that the detonation is not spherical, but rather an ellipsoid with symmetric horizontal and depth components, but a differing vertical component. The differences between 3D point reconstructions and sphere could be compared to that of a 3D point reconstructions and an ellipsoid leading to conclusions that one model is more accurate than another. These comparisons are not possible without the understanding of the accuracy of the point correspondence grounded in truth locations that this paper establishes.

VI. Conclusions and Future Work

This paper presents a method to quantify the positional error of points in a 3D point cloud generated from triangulation and bundle adjusted reconstructions of nuclear material in a nuclear detonation. To quantify the positional error, 138 reconstructions were accomplished on 50% of

applicable digitized detonations to estimate the positions of the support tower used in detonations. Ten points in each reconstruction were then compared to location information recorded at the time of detonation to estimate the positional error of a point. Reconstruction via triangulation was compared to reconstruction using bundle adjustment. Using bundle adjustment over triangulation improves the error by 0.15 meters on average, without statistical significance. The error of a given point with bundle adjustment can be estimated with an average positional error of 0.68 meters with 0.36, 0.28, and 0.37 meters of uncertainty in x-, y-, and z-directions respectively.

The three dimensional data sets created and the quantified error associated with 3D point reconstructions provides a method to verify volume-based models of nuclear detonations improving the accuracy of yield estimates and improving the ability to attribute the nuclear material used in a detonation to a specific nation. Future work includes generating volume estimates of detonations and evaluating how accurate they are using the results of this paper. These volume estimates might be based on single points of view, or integrating multiple points of view.

REFERENCES

[1] Geoffrey Taylor, "The formation of a blast wave by a very intense explosion," in *Proceedings of the Royal Society of London*. 1950, vol. 201, pp. 159–174, Series A, Mathematical and Physical Sciences.

[2] Daniel Schmitt, "Position and volume estimation of atmospheric nuclear detonations from video reconstruction," Nov 2014, Air Force Institute of Technology.

[3] Robert Slaughter, "Multidimensional analysis of atmospheric nuclear detonations," Air Force Institute of Technology, Jan 2014.

[4] Spriggs Gregory D., "Film Scanning Project," Lawrence Livermore National Laboratory, Aug 2011.

[5] David G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, Jan. 2004.

[6] Noah Snavely, Steven M. Seitz, and Richard Szeliski, "Photo tourism: Exploring photo collections in 3d," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 835–846, July 2006.

[7] Curless B. Diebel J. Scharstein D. Szeliski R. Seitz, S. M., "A comparison and evaluation of multi-view stereo reconstruction algorithms," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 519–526, 2006.

[8] Richard Hartley and Andrew Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, New York, NY, USA, 2 edition, 2003.

[9] Richard I Hartley and Peter Sturm, "Triangulation," *Computer vision and image understanding*, vol. 68, no. 2, pp. 146–157, 1997.

[10] Richard Szeliski, *Computer Vision: Algorithms and Applications*, Springer-Verlag, London, 2011.

[11] John Oliensis, "The least-squares error for structure from infinitesimal motion.," *International Journal of Computer Vision*, vol. 61, no. 3, pp. 259–299, 2005.

[12] Richard Szeliski and Sing Bing Kang, "Shape ambiguities in structure from motion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 5, pp. 506–512, May 1997.

[13] Andrea Vedaldi and Brian Fulkerson, "Vlfeat: An open and portable library of computer vision algorithms," in *Proceedings of the International Conference on Multimedia*, New York, NY, USA, 2010, MM '10, pp. 1469–1472, ACM.

[14] F.-X. Le Dimet, I. Souopgui, O. Titaud, V. Shutyaev, and M. Y. Hussaini, "Toward the assimilation of images," *Nonlinear Processes in Geophysics*, vol. 22, no. 1, pp. 15–32, 2015.

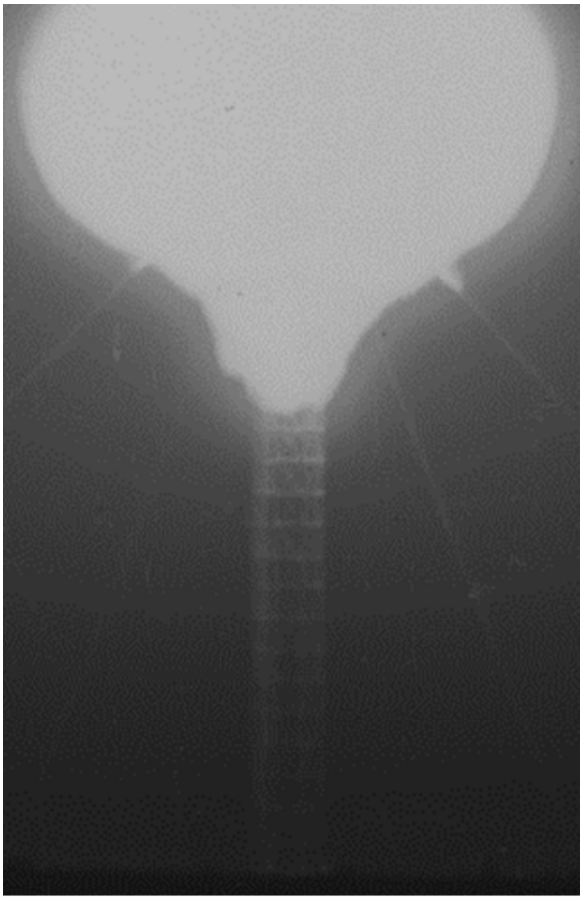


Fig. 1. An example NUDET image of Tesla detonation with tower visible below the detonation [4].

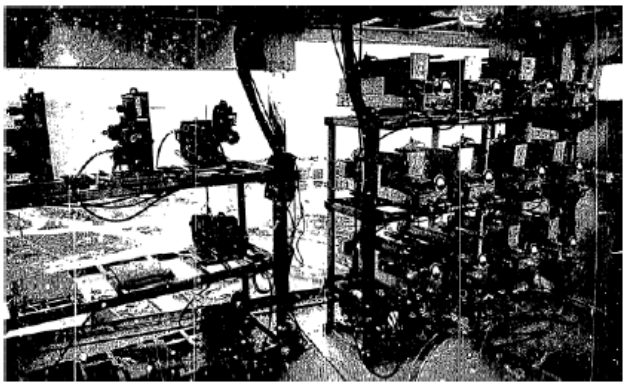


Fig. 2. A typical arrangement of cameras within a collection site [4].

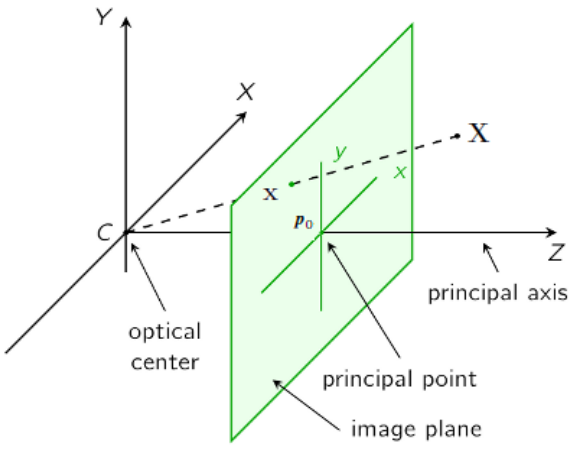


Fig. 3. The pinhole camera model.

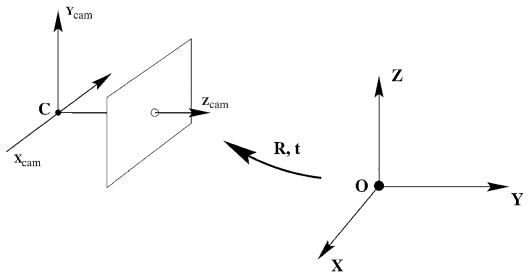


Fig. 4. The Euclidean transformation between world and camera coordinate frames is comprised of rotation ($\mathbf{R}$) and translation ($\mathbf{t}$) [8].

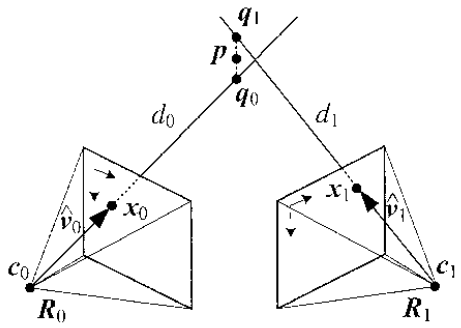
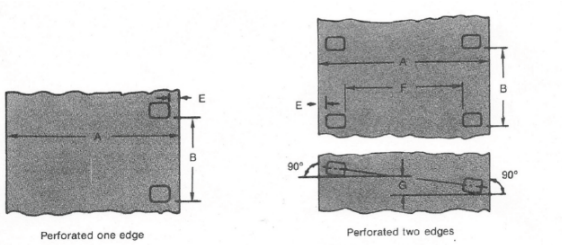


Fig. 5. 3D point triangulation [10]. Point $\mathbf{p}$ is estimated using rays traced from the centers of camera at $\mathbf{R}_1\mathbf{c}_1$ and $\mathbf{R}_2\mathbf{c}_2$.



Perforation Type and ANSI Number										
Dimension	1R-2994 (PH22.109)		1R-3000 (PH22.12)		2R-2994 (PH22.110)		2R-3000 (PH22.5)		Tolerances ±	
	Inches	mm	Inches	mm	Inches	mm	Inches	mm	Inches	mm
A*	0.628	15.95	0.628	15.95	0.628	15.95	0.628	15.95	0.001	0.03
B	0.2994	7.605	0.3000	7.620	0.2994	7.605	0.3000	7.620	0.0005	0.013
E	0.0355	0.902	0.0355	0.902	0.0355	0.902	0.0355	0.902	0.0020	0.051
F					0.413	10.49	0.413	10.49	0.001	0.03
G (max)					0.001	0.03	0.001	0.03	—	—
L†	29.94	760.5	30.00	762.0	29.94	760.5	30.00	762.0	0.03	0.8

*This dimension also represents the unperforated width.
†This dimension represents the length of any 100 consecutive perforation intervals.

Fig. 6. Dimensions of the 16mm film [4].

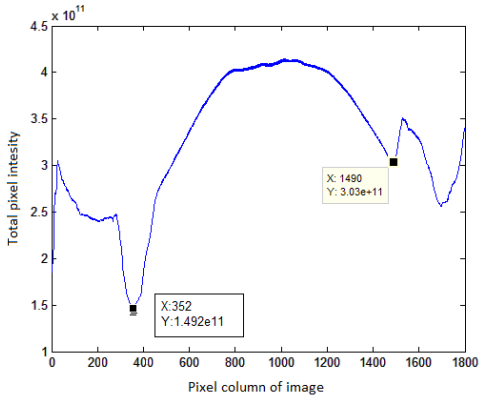


Fig. 7. A summary of pixel intensities by column. The low points on either side represent where the best area to crop the images would be.

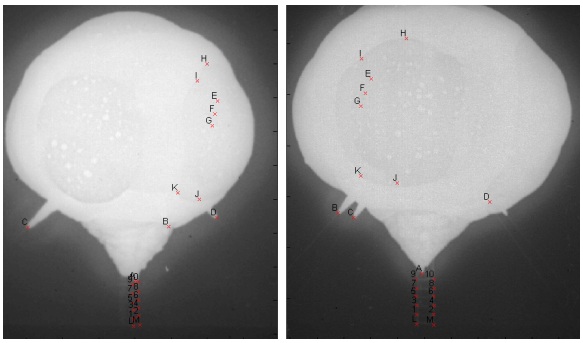


Fig. 8. Matched correspondences of Film 04 with Film 10 of Tesla.

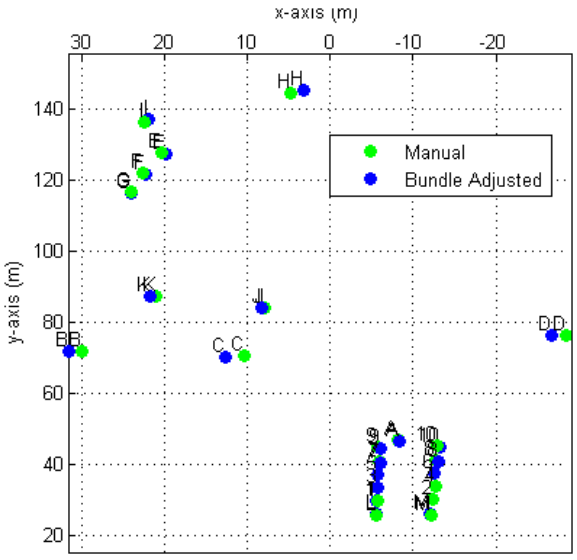


Fig. 9. 3D Reconstruction of Film 04 with Film 10 of Tesla.