A Variational Approach to Problems in Calibration of Multiple Cameras

Gozde Unal, *Member*, *IEEE*, Anthony Yezzi, *Member*, *IEEE*,
Stefano Soatto, *Member*, *IEEE*, and Greg Slabaugh, *Member*, *IEEE*

Abstract—This paper addresses the problem of calibrating camera parameters using variational methods. One problem addressed is the severe lens distortion in low-cost cameras. For many computer vision algorithms aiming at reconstructing reliable representations of 3D scenes, the camera distortion effects will lead to inaccurate 3D reconstructions and geometrical measurements if not accounted for. A second problem is the color calibration problem caused by variations in camera responses that result in different color measurements and affects the algorithms that depend on these measurements. We also address the extrinsic camera calibration that estimates relative poses and orientations of multiple cameras in the system and the intrinsic camera calibration that estimates focal lengths and the skew parameters of the cameras. To address these calibration problems, we present multiview stereo techniques based on variational methods that utilize partial and ordinary differential equations. Our approach can also be considered as a coordinated refinement of camera calibration parameters. To reduce computational complexity of such algorithms, we utilize prior knowledge on the calibration object, making a piecewise smooth surface assumption, and evolve the pose, orientation, and scale parameters of such a 3D model object without requiring a 2D feature extraction from camera views. We derive the evolution equations for the distortion coefficients, the color calibration parameters, the extrinsic and intrinsic parameters of the cameras, and present experimental results.

Index Terms—Calibration, variational methods, color calibration, lens distortion calibration, camera parameters refinement.

1 Introduction

The problem of recovering a 3D representation of a scene from multiple 2D images has been one of the main research interests in computer vision. Many of the existing stereo techniques involve preprocessing the camera images to extract 2D features such as corners, lines, and contours of objects in the scene. These features are then used to find correspondences between camera views. In practice, searching for features and establishing correspondences is not an easy task due to noise and local extrema. Early variational approaches to the 3D reconstruction problem were pioneered by Faugeras and Keriven [1] who also relied on local feature matching. A more recent variational approach by Yezzi and Soatto [2], [3] proposed a joint region-based image segmentation and simultaneous 3D stereo reconstruction technique. This paper addresses camera calibration techniques built on this latter stereo reconstruction framework that avoids searches for local correspondences and is versatile enough to accommodate the new applications to be shown. A tradeoff is achieved by making a piecewise smooth object assumption and a constant background assumption; however, extraction of 2D features from given camera views are not required.

Camera calibration refers to the problem of finding the mapping between the 3D world and the camera or image plane. For most computer vision algorithms aimed at reconstructing reliable digital representations of 3D scenes, accurate camera calibrations are essential. There has been a great deal of research on the camera calibration problem as early as the 1970s [4]. In most of the previous techniques, some sets of features were extracted from images of a known calibration pattern and intrinsic camera parameters as well as camera pose and orientation (extrinsic camera parameters) are estimated by a minimization of an overall cost functional [5], [6], [7], [8], [9], [10], [11], [12], [13]. Many calibration techniques use both nonlinear minimization and closed-form solutions as in [14].

In this paper, we develop a coordinated refinement technique for the extrinsic and intrinsic camera parameters: lens distortion, focal lengths, skew, and estimation of camera color calibration parameters in a coupled way within a multiple camera system.[1] For geometrical measurements, an intrinsic camera parameter, the camera lens distortion, is an important issue and will result in inaccurate 3D reconstructions if not taken into account. Another common problem in multiview stereo techniques is caused by color miscalibrations between cameras due to different sensor characteristics. Extrinsic parameters of the cameras, on the other hand, determine the relative poses and orientations of cameras and their correct estimation is one of the first phases of a camera calibration system.

- *G. Unal and G. Slabaugh are with Siemens Corporate Research, 755 College Road East, Princeton, NJ 08540. E-mail: {gozde.unal, greg.slabaugh}@siemens.com.*
- *A. Yezzi is with the School of Electrical Engineering, Georgia Institute of Technology, 777 Atlantic Dr. NW, Atlanta, GA 30332-0250. E-mail: ayezzi@ece.gatech.edu.*
- *S. Soatto is with the Computer Science Department, University of California, Los Angeles, Boelter Hall 3531D, 405 Hilgard Ave., Los Angeles, CA 90095-1596. E-mail: soatto@ucla.edu.*

Manuscript received 18 Apr. 2005; revised 3 Apr. 2006; accepted 15 Sept. 2006; published online 18 Jan. 2007.
Recommended for acceptance by L. Quan.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number TPAMI-0202-0405.
Digital Object Identifier no. 10.1109/TPAMI.2007.1035.

1. An initial version of this work that addresses lens distortion and color calibration can be found in [15] and [16], and an initial work addressing extrinsic camera calibration appears in [3].

Fig. 1. Three out of five camera views of the real calibration object shown on the left.

Fig. 2. Photograph of the color calibration object.

1.1 Relation to Previous Work and Contributions

1.1.1 Lens Distortion

The ideal pinhole camera model leads to imaging of world lines as lines on the image plane and simplifies many computations and considerations [6]. However, for most real cameras with wide-angle or inexpensive lenses, this assumption does not hold and nonlinearities introduced by a well-known phenomenon referred to as a lens distortion should be taken into account. The corresponding distortion parameters should be estimated for each camera.

In many existing calibration techniques, good estimates for extrinsic and intrinsic camera parameters are first obtained by a pinhole camera model neglecting lens distortion. Then, distortion calibration is performed while holding the other parameters fixed [17], [18], [19]. This is possible because the mapping from 3D world coordinates to the 2D image plane can be decomposed into a perspective projection and a mapping that models the deviations from the ideal pinhole camera.

A popular group of lens distortion calibration methods in the literature, mainly under the category known as plumb line methods, rely on the first step of extracting edges from the images. Either a user manually selects the image curves or there must be a way to reliably estimate image edges that correspond to linear 3D segments in the world. An optimization problem is set up by defining a measure of how much each detected segment is distorted. The curved lines in the image that do not really correspond to 3D line segments will constitute outliers in this optimization procedure [17], [20], [21], [22], [23]. Other techniques such as [24] rely on point correspondences. Given a set of 3D points, the associated epipolar and trilinear constraints are arranged into a tensor, which is computed with estimated distortion parameters at each step to minimize a reprojection error in an iterative manner. In another group of methods as in [25], [26], [27], a direct solution strategy is employed to find camera calibration parameters by incorporating lens distortion as well.

Our contribution is a new distortion calibration technique that does not rely on extraction of edges and search for point correspondences. The former may not be an easy task due to noise and local extrema. Instead, we devise an integrated calibration technique in which the distortion parameters of cameras are computed in a tightly coupled framework. The desired coupling of multiple camera views comes from estimating a common 3D object (in this case, the calibration object). In other words, we minimize the cost between the reprojection of the 3D calibration object and the image measurements by evolving the distortion parameters of the cameras. In our distortion calibration algorithm, we use a white bar object made from a foam core, as shown in Fig. 1 on the left. We capture its views before a dark background with the multiview stereo rig system, a desktop multicamera system designed for remote multimedia collaboration, developed by HP Labs [28]. The images of the calibration object captured from three of the five cameras in the rig are given in Fig. 1. Many desktop multicamera systems use wide-angle and inexpensive cameras that produce severe distortion effects as can be observed in the given images.

As we will show, with this technique, we can also incorporate other parameters of calibration into the same variational framework and get their locally optimal estimates as well.

1.1.2 Color Calibration

Another common problem in multiview stereo techniques is caused by color miscalibrations between cameras, resulting from variations in camera responses due to different sensor characteristics and ambient conditions like temperature, manufacturing differences, and so on. These yield different color measurements between cameras, and affect the algorithms that depend on these measurements. Camera color calibration refers to the problem of estimating the color calibration parameters of cameras to overcome these unwanted effects. A common approach taken toward this problem is to calibrate each camera independently through comparisons with known colors on a color calibration object/environment [28], [29].

The color calibration object we use, shown in Fig. 2, is a color cube with patches of known colors whose images are captured from each camera. Demosaicing coefficients are calculated independently for each camera based upon the absolute colors of the calibration object and the measured color responses of each camera. Slight errors and differences that arise from this independent calibration procedure sometimes lead to noticeable seams or discontinuities in the texture mapping process during the transition of the texture map between neighboring cameras. Our goal is to help even out these discrepancies by devising a *relative intercamera* color calibration technique in which the demosaicing parameters of cameras are calculated jointly in a tightly coupled framework rather than just one camera at a time.

Similar to our approach to lens distortion calibration, the desired coupling of the multiple camera views comes from estimating a common 3D shape and, in addition, a common radiance function for the calibration object (in this case, the color cube). We take advantage of the fact that the object shape is known up to location and scale to simplify the problem. Hence, we estimate the pose parameters of the cube, the radiance function on the cube and the color calibration coefficients for each camera.

1.1.3 Extrinsic and Intrinsic Calibration

Following the same philosophy as mentioned in the other two calibration problems above, extrinsic and intrinsic calibration parameters can be estimated in a variational framework using the general stereoscopic framework of Yezzi-Soatto.

It should be noted that due to the differential nature of the estimation equations derived, the extrinsic and intrinsic update equations require rough initial values. This is a well-known feature of almost all of the recent state-of-the-art energy functionals used in segmentation (for example, Mumford-Shah energy, geodesic energy, ...), that is, the

solutions are locally optimal, hence, starting far away from the real solution may lead to solutions that get stuck at local extrema far from the desired solution. Nevertheless, the usefulness of a refinement stage in extrinsic and intrinsic camera parameters will be demonstrated via the improvement in the final 3D reconstructions. A nice feature of the methodology presented in this paper is that it can integrate several different problems in geometric and color calibration into an overall unified system based on the joint segmentation framework to evolve pose, color, distortion, extrinsics, and intrinsics.

The organization of this paper is presented as follows: We first present a variant of the Yezzi-Soatto algorithm in which a 3D object is allowed to move with a semiaffine motion model in Section 2. We developed this scheme for our applications in calibration, where the 3D object shape is roughly known (up to three scales and rigidity) to obtain more efficient and faster algorithms. We then present a novel technique for lens distortion calibration in Section 3 and a novel technique for relative intercamera color calibration in Section 4. We apply the same calibration ideas for intrinsic camera calibration in Section 5 and for extrinsic camera calibration problem in Section 6. Conclusions and discussions are given in Section 7.

2 Evolution Equations of 3D Object Motion Parameters

The Yezzi-Soatto 3D stereo reconstruction model builds a cost on the discrepancy between the reprojection of a model surface with a radiance $f : \mathbb{R}^3 \longrightarrow \mathbb{R}$, the background (infinitely far away) with radiance $b : \mathbb{R}^3 \longrightarrow \mathbb{R}$, and the actual measurements from multiple camera views. Let g_i denote the transformation from world coordinates to camera coordinates: $g_i : \mathbf{X} \longrightarrow \mathbf{X}_i = (X_i, Y_i, Z_i)^T$ and let π denote the perspective transformation from camera frame to the image plane: $\pi : \mathbf{X}_i \longrightarrow \mathbf{x}_i = (x_i = \frac{X_i}{Z_i}, y_i = \frac{Y_i}{Z_i})^T$.

On the image plane, the cost functional for the Yezzi-Soatto model can be written as a joint segmentation problem over regions of n camera images I_i with domain $\Omega_i = R_i \cup R_i^c$ (R_i denotes the foreground region) and with three color channels $k \in (R, G, B)$:

$$E = \sum_{k=R,G,B} \sum_{i=1}^{n} \int_{R_i} \left[f^k\Big((\pi \circ g_i)^{-1}(\mathbf{x}_i) \Big) - I_i^k(\mathbf{x}_i) \right]^2 d\Omega_i + \sum_{k=R,G,B} \sum_{i=1}^{n} \int_{R_i^c} \left[b^k - I_i^k \right]^2 d\Omega_i. \tag{1}$$

This energy can be lifted back onto surface S:

$$E(S) = \sum_{k=R,G,B} \sum_{i=1}^{n} \int_S \Big[\Big(f^k(\mathbf{X}) - I_i^k(\pi \circ g_i(\mathbf{X})) \Big)^2 - (b^k - I_i^k)^2 \Big] \mathcal{X}_i(\mathbf{X}) \sigma(\mathbf{X}_i) dA, \tag{2}$$

where σ is the Jacobian of the change of coordinates from the image plane to the surface, $\mathcal{X}_i$ is the visibility function of a voxel on the surface, and dA is the area measure of surface S. The deformation of the surface S with respect to this energy or data fidelity measure is then obtained by finding the partial differential equation (PDE) that is the gradient descent flow of the energy E. A popular class of numerical techniques, known as Level Sets Methods [30], is utilized to evolve the surface S via the evolution of a 3D function $\Psi : \mathbb{R}^3 \longrightarrow \mathbb{R}$. Nevertheless, an update of the level set function is required after each iteration of the associated PDE, and even with more efficient narrowband schemes [31], there is a considerable amount of computation involved. For our intended applications, in which there is a calibration object whose shape can be roughly known a priori, rather than deforming the surface of the 3D object, we will evolve its pose and scale parameters instead. Next, using the energy E in (2), we will derive the ordinary differential equations (ODEs) to update the parameters of the surface motion modeled by a semiaffine transformation, which is more general than a similarity but less general than a fully affine transformation.

Let the original rigid surface be denoted by S_o, then $S = g^s(S_o)$, or $\mathbf{X} = g^s(\mathbf{X}_o) = \mathbf{R}_s \mathbf{X}_o + \mathbf{T}_s$, and let λ denote parameters of the rigid motion g^s of the surface S_o with rotation $\mathbf{R}_s$ and translation $\mathbf{T}_s$. Then, the gradient of the energy E with respect to λ is given by

$$\begin{aligned} \frac{\partial E(\lambda)}{\partial \lambda} &= \sum_{k=R,G,B} \int_S \sum_i F_i^k(\mathbf{X}) \left\langle \frac{\partial \mathbf{X}}{\partial \lambda}, \mathbf{N} \right\rangle dA + \frac{\partial F_i^k(\mathbf{X})}{\partial \lambda} dA \\ &= \sum_{k=R,G,B} \int_{S_o} \sum_i F_i^k(g^s(\mathbf{X}_o)) \left\langle \frac{\partial (g^s \mathbf{X}_o)}{\partial \lambda}, \mathbf{R}_s \mathbf{N}_o \right\rangle dA_o \\ &\quad + 2(f^k(g\mathbf{X}_o) - I_i^k) \left\langle \frac{\partial (g^s \mathbf{X}_o)}{\partial \lambda}, \nabla f_S^k \right\rangle dA_o, \end{aligned} \tag{3}$$

where $F_i^k = [(f^k(\mathbf{X}) - I_i^k(\pi \circ g_i(\mathbf{X})))^2 - (b^k - I_i^k)^2] \mathcal{X}_i(\mathbf{X}) \sigma(\mathbf{X}_i)$ is the Mumford-Shah term from (2) (also in [2]). The derivation follows from shape optimization tools [32] that provide the shape derivatives in curve and surface evolution framework. $\mathbf{N}$ denotes the surface normal vector. Note that the visibility function $\mathcal{X}_i(g^s(\mathbf{X}_o))$, included in the data term $F_i^k(\cdot)$, is computed using the original visibility function but compensated by $\mathbf{R}_s^T(\mathbf{C}_i - \mathbf{T}_s)$), where $\mathbf{C}_i$ is a camera center. The second term in (3) is the region term corresponding to the foreground object, whereas the first one is the boundary term. In our applications, the background is modeled by a piecewise constant radiance; therefore, we omit the background region term in the equation.

For translation parameters, we have

$$\left\langle \frac{\partial (g^s \mathbf{X}_o)}{\partial \lambda}, \mathbf{R}_s \mathbf{N}_o \right\rangle = \mathbf{R}_s \mathbf{N}_o.$$

For rotation parameters, we have

$$\left\langle \frac{\partial (g^s \mathbf{X}_o)}{\partial \lambda}, \mathbf{R}_s \mathbf{N}_o \right\rangle = \left\langle \mathbf{R}_s \underbrace{\begin{bmatrix} 0 & Z_o & -Y_o \\ -Z_o & 0 & X_o \\ Y_o & -Xo & 0 \end{bmatrix}}_{\hat{\mathbf{X}}_o}, \mathbf{R}_s \mathbf{N}_o \right\rangle = \left\langle -\mathbf{R}_s \hat{\mathbf{X}}_o, \mathbf{R}_s \mathbf{N}_o \right\rangle, \tag{4}$$

where we utilize exponential coordinates (see [33] for details on this representation) for the global rotation parameters of the surface. We note that a matrix in an inner-product expression, when operated on a vector, will incorporate each of its row vectors in the inner product to result in a vector: $\langle \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n, \mathbf{y} \rangle = (\langle \mathbf{x}_1, \mathbf{y} \rangle, \langle \mathbf{x}_2, \mathbf{y} \rangle, \ldots, \langle \mathbf{x}_n, \mathbf{y} \rangle)$.

For further flexibility in initializing a model surface, we add three scaling parameters along the X, Y, and Z axes. Then, the semiaffine transformation for a point $\mathbf{X}_o$ on the surface becomes: $\mathbf{X} = g^s(\mathbf{X}_o) = \mathbf{RSX}_o + \mathbf{T}$, where

$$\mathbf{S} = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & s_z \end{bmatrix}.$$

The gradient of the energy with regard to the scaling parameters $\lambda = s_j$ is derived similarly to the above:

$$\sum_{k=R,G,B} \int_{S_o} \sum_i F_i^k(g^s(\mathbf{X}_o)) \left\langle \frac{\partial(g^s\mathbf{X}_o)}{\partial\lambda}, \mathbf{R}_s\mathbf{N}_o \right\rangle dA_o,$$

where

$$\begin{aligned} \left\langle \frac{\partial(g^s\mathbf{X}_o)}{\partial\lambda}, \mathbf{R}_s\mathbf{N}_o \right\rangle &= \left\langle \mathbf{R}_s \frac{\partial S}{\partial\lambda}\mathbf{X}_o, \mathbf{R}_s\mathbf{N}_o \right\rangle \quad \text{with, e.g.,} \\ \frac{\partial(g^s\mathbf{X}_o)}{\partial s_x} &= \mathbf{R}_s \frac{\partial S}{\partial s_x}\mathbf{X}_o = \mathbf{R}_s \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{X}_o \\ &= \mathbf{R}_s \begin{bmatrix} \mathbf{X}_o.x \\ 0 \\ 0 \end{bmatrix}, (X_o = \mathbf{X_o}.\mathbf{x}), \text{ then} \\ \left\langle \frac{\partial(g^s\mathbf{X}_o)}{\partial\lambda}, \mathbf{R}_s\mathbf{N}_o \right\rangle &= \Bigg\langle \underbrace{\mathbf{R}_s \left[\begin{bmatrix} X_o \\ 0 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ Y_o \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 0 \\ Z_o \end{bmatrix} \right]}_{\mathbf{R}_s^X}, \\ &\qquad \mathbf{R}_s\mathbf{N}_o \Bigg\rangle. \end{aligned} \tag{5}$$

The evolutions for the rigid motion parameters λ are then given by the following gradient descent equations:

$$\frac{\partial\lambda}{\partial t} = -\frac{\partial E}{\partial\lambda} = -\sum_{k=R,G,B} \int_{S_o} \underbrace{\sum_i F_i^k(g^s(\mathbf{X}_o))}_{F^k} \ \mathbf{R}_s\mathbf{N}_o \, dA_o, \tag{6}$$

(for translation),

$$\frac{\partial\lambda}{\partial t} = -\frac{\partial E}{\partial\lambda} = -\sum_{k=R,G,B} \int_{S_o} F^k \big\langle -\mathbf{R}_s\hat{\mathbf{X}}_o, \mathbf{R}_s\mathbf{N}_o \big\rangle \, dA_o, \tag{7}$$

(for rotation),

$$\frac{\partial\lambda}{\partial t} = -\frac{\partial E}{\partial\lambda} = -\sum_{k=R,G,B} \int_{S_o} F^k \big\langle \mathbf{R}_s^X, \mathbf{R}_s\mathbf{N}_o \big\rangle \, dA_o, \tag{8}$$

(for scaling).

Here, note that the visibility function $\mathcal{X}_i(g^s(\mathbf{X}_o))$ is computed using the original visibility function but compensated by the $\mathbf{S}^{-1}\mathbf{R}_s^T(\mathbf{C}_i - \mathbf{T}_s))$, where

$$\mathbf{S}^{-1} = \begin{bmatrix} 1/s_x & 0 & 0 \\ 0 & 1/s_y & 0 \\ 0 & 0 & 1/s_z \end{bmatrix}.$$

Note that we can generalize this idea in a straightforward fashion by considering $\mathbf{S}$ to be more general than a simple diagonal matrix in order to accommodate a fully affine motion of the surface.

We will use (6), (7), and (8) in updating the pose of the surface S to estimate its correct placement in the 3D space for the calibration applications presented in Sections 3 and 4.

3 Lens Distortion Calibration

The lens distortion is usually modeled by a function defined from the ideal image plane to the distorted image plane. One approach is to decompose it into two terms: radial and tangential distortion [17]. The radial distortion is a deformation along the radial direction from a center of distortion point to an image point and the tangential distortion is a deformation in a direction perpendicular to the radial direction and is negligible for many cameras. To model the radial distortion effects, a commonly used distortion function $D(r)$ is given by $(1 + k_1r^2 + k_2r^4 + \ldots)$, where r is the radius from the center of distortion to a point on the ideal image plane. The principal point (u_0, v_0) is often used as the center for radial distortion [6], which we will also adopt. Below, $\hat{\mathbf{x}}_i$ is the distorted image coordinates and D is the distort function:

$$\hat{\mathbf{x}}_i = D\mathbf{x}_i = (1 + k_1^i r^2 + k_2^i r^4 + \ldots)\mathbf{x}_i, \tag{9}$$

where $r^2 = (x_i^2 + y_i^2)$, and k_j^i is the jth distortion coefficient for camera i. In (9), we assume that $k_0 = 1$, which can be changed to an arbitrary k_0 value.

3.1 Calibration of the Lens Distortion Parameters

Notation. Our notation from world coordinates $\mathbf{X}$ to image coordinates (u, v) is:

$$\mathbf{X} \underbrace{\longrightarrow}_{g_i} \mathbf{X}_i \underbrace{\longrightarrow}_{\pi} \mathbf{x}_i = \begin{pmatrix} \frac{X_i}{Z_i} = x_i \\ \frac{Y_i}{Z_i} = y_i \\ 1 \end{pmatrix} \underbrace{\longrightarrow}_{\begin{bmatrix} L_u & 0 & u_0 \\ 0 & L_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}} (u, v),$$

where D is the distort function in (9) and L_u and L_v are the focal lengths. The gradient of the energy (1), assuming a single image channel over the distorted image plane, with regard to distortion parameters k_j^i is given by

$$\frac{\partial k_j^i}{\partial t} = -\frac{\partial E}{\partial k_j^i} = -\int_{\hat{c}_i} F_i\Big((D \circ \pi \circ g_i)^{-1}\hat{\mathbf{x}}_i\Big) \left\langle \frac{\partial\hat{\mathbf{x}}_i}{\partial k_j^i}, \hat{\mathbf{n}}_i \right\rangle d\hat{s}, \tag{10}$$

where $F_i = (f - I_i)^2 - (b - I_i)^2$, subscript i corresponds to each camera view, and $\hat{\mathbf{n}}_i$ denotes the normal vector to the occluding boundary $\hat{c}_i$ of region R_i on the distorted image plane. We only consider the boundary term ($\hat{s}$ is the arclength of the contour $\hat{c}_i$ on the image plane: The distorted or actual image coordinates), as we assume that the foreground and background have constant radiance. We design the lens distortion calibration object to satisfy this assumption.

We want to lift this integral back onto occluding boundary C_i of the surface. Note that $\frac{\partial\hat{\mathbf{x}}_i}{\partial k_j}$ are given by

$$\frac{\partial\hat{\mathbf{x}}_i}{\partial k_1^i} = r^2\mathbf{x}_i, \quad \frac{\partial\hat{\mathbf{x}}_i}{\partial k_2^i} = r^4\mathbf{x}_i, \quad \ldots \frac{\partial\hat{\mathbf{x}}_i}{\partial k_j^i} = r^{2j}\mathbf{x}_i,$$

hence

$$\begin{aligned}\left\langle \frac{\partial \hat{\mathbf{x}}_i}{\partial k_j^i}, \hat{\mathbf{n}}_i \right\rangle d\hat{s} &= \left\langle r^{2j}\pi(\mathbf{X}_i), J\frac{\partial}{\partial s}(D\circ\pi)\mathbf{X}_i \right\rangle ds \\ &= \left\langle r^{2j}\pi(\mathbf{X}_i), JD'\circ\pi'\frac{\partial}{\partial s}\mathbf{X}_i \right\rangle ds,\end{aligned} \tag{11}$$

where J denotes the 2×2 90 degree rotation matrix, $D' = (1 + k_1 r^2 + k_2 r^4 + \ldots)$, and

$$\pi' = \frac{1}{Z_i^2}\begin{bmatrix} Z_i & 0 & -X_i \\ 0 & Z_i & -Y_i \end{bmatrix} \tag{12}$$

is the Jacobian of the perspective projection π. We can continue to simplify

$$\begin{aligned}&\left\langle \frac{\partial \hat{\mathbf{x}}_i}{\partial k_j^i}, \hat{\mathbf{n}}_i \right\rangle d\hat{s} = r^{2j}D'\Bigg\langle [\pi(\mathbf{X}_i)]_{2\times 1}, \\ &\frac{1}{Z_i^2}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} Z_i & 0 & -X_i \\ 0 & Z_i & -Y_i \end{bmatrix}\left[\frac{\partial \mathbf{X}_i}{\partial s}\right]_{3\times 1}\Bigg\rangle ds \\ &= \frac{r^{2j}D'}{Z_i^2}\left\langle \begin{pmatrix} X_i \\ Y_i \end{pmatrix}\frac{1}{Z_i}, \begin{bmatrix} 0 & Z_i & -Y_i \\ -Z_i & 0 & X_i \end{bmatrix}\left[\frac{\partial \mathbf{X}_i}{\partial s}\right]_{3\times 1}\right\rangle ds \\ &= \frac{r^{2j}D'}{Z_i^3}\left\langle \begin{bmatrix} 0 & -Z_i \\ Z_i & 0 \\ -Y_i & X_i \end{bmatrix}\begin{pmatrix} X_i \\ Y_i \end{pmatrix}, \frac{\partial \mathbf{X}_i}{\partial s}\right\rangle ds \\ &= \frac{r^{2j}D'}{Z_i^3}\left\langle \begin{pmatrix} -Z_iY_i \\ Z_iX_i \\ 0 \end{pmatrix}, \frac{\partial \mathbf{X}_i}{\partial s}\right\rangle ds.\end{aligned}$$

Noting that

$$\begin{pmatrix} -Z_iY_i \\ Z_iX_i \\ 0 \end{pmatrix} = \mathbf{X}_i \times \begin{pmatrix} X_i \\ Y_i \\ 0 \end{pmatrix},$$

we have

$$\left\langle \frac{\partial \hat{\mathbf{x}}_i}{\partial k_j}, \hat{\mathbf{n}}_i \right\rangle d\hat{s} = \frac{r^{2j}D'}{Z_i^3}\left\langle -\begin{pmatrix} X_i \\ Y_i \\ 0 \end{pmatrix}\times \mathbf{X}_i, \frac{\partial \mathbf{X}_i}{\partial s}\right\rangle ds,$$

and

$$\begin{aligned}\left\langle \frac{\partial \hat{\mathbf{x}}_i}{\partial k_j^i}, \hat{\mathbf{n}}_i \right\rangle d\hat{s} &= -\frac{r^{2j}D'}{Z_i^3}\left\langle \mathbf{X}_i \times \frac{\partial \mathbf{X}_i}{\partial s}, \begin{pmatrix} X_i \\ Y_i \\ 0 \end{pmatrix}\right\rangle ds \\ &= -\frac{r^{2j}D'}{Z_i^3}\left\langle \|\mathbf{X}_i\|\mathbf{N}_i, \begin{pmatrix} X_i \\ Y_i \\ 0 \end{pmatrix}\right\rangle ds.\end{aligned} \tag{13}$$

Substituting (13) into (10), we get the calibration equation

$$\frac{\partial k_j^i}{\partial t} = \int_{C_i} F_i \frac{r^{2j}D'\|\mathbf{X}_i\|}{Z_i^3}\left\langle \mathbf{N}_i, \begin{pmatrix} X_i \\ Y_i \\ 0 \end{pmatrix}\right\rangle ds \tag{14}$$

for the lens distortion parameters k_j^i. Note that the distortion calibration method we propose can handle different models of distortion by changing the D function and related derivatives in (14).

Fig. 3. Initialized surface model shown from three different vantage points.

3.2 Using Several Poses of the Object

When camera views from multiple poses of the object are available, we can take advantage of the existence of variously distorted views in calibrating the lens distortion. In the first phase, we estimate both pose and distortion coefficients from separate experiments. To simplify the explanation, let us assume that we want to solve for only one distortion coefficient k_1^i for each camera i. Once we obtain rough estimates for the object pose and distortion coefficients k_1^i, we can fuse a "common distortion" $\tilde{k}_1^i$ from these separate experiments for each camera i and then jointly evolve $\tilde{k}_1^i$s as follows:

$$\frac{\partial \tilde{k}_1^i}{\partial t} = \sum_{m=1}^{M\text{poses}} \int_{C_{i,m}} F_{i,m}\frac{r^{2j}D'\|\mathbf{X}_{i,m}\|}{Z_{i,m}^3}\left\langle \mathbf{N}_{i,m}, \begin{pmatrix} X_{i,m} \\ Y_{i,m} \\ 0 \end{pmatrix}\right\rangle ds. \tag{15}$$

At the same time, we evolve the pose parameters of separate poses of the object as described in Section 2, with the only difference being the incorporation of the new "common distortion" in the equations. For instance, we evolve any of them for a given pose as follows:

$$\frac{\partial \lambda}{\partial t} = -\sum_i \int_{S_o} F_i(g^s(\mathbf{X}_o))\left\langle \frac{\partial (g^s\mathbf{X}_o)}{\partial \lambda}, \mathbf{R}_s\mathbf{N}_o\right\rangle dA_o, \tag{16}$$

where F_i includes computation of $I_i(D\cdot\pi\cdot g_i(g^s(\mathbf{X}_o)))$ with the new common distortion coefficients $\tilde{k}_1^i$ in the multiplying distortion factor D.

3.3 Experimental Results

For our calibration algorithm, we initialize a surface model of the real calibration object, which is shown from several vantage points in Fig. 3. After initializing the surface, the first phase of our algorithm is to evolve its pose parameters to position the 3D object model roughly in the correct location in 3D space. For the experiments presented here, captured via a HP Labs' stereo rig system, we used three different poses of the calibration object, but we can increase the number of poses used in the process. Example evolutions of the pose parameters are shown in Fig. 4 for three different pose captures of the calibration object in each column (showing only one camera view for each pose). The distortion coefficients are also evolved at a slower pace. That is, the time step used in the associated ODE is small in the first phase. In the experiments, the distortion function D in (9), with one distortion coefficient k_1 for each camera, is used.

After the separate evolutions for each of the poses have converged, common initial distortion coefficients are computed as the average of the results from phase 1. In the second phase of the algorithm, we evolve the distortion coefficients for each camera, again, separately but summed over different poses. We show sample views of poses 1, 2, and 3 in row 1 of Figs. 5, 6, and 7. As the distortion coefficients converge to true

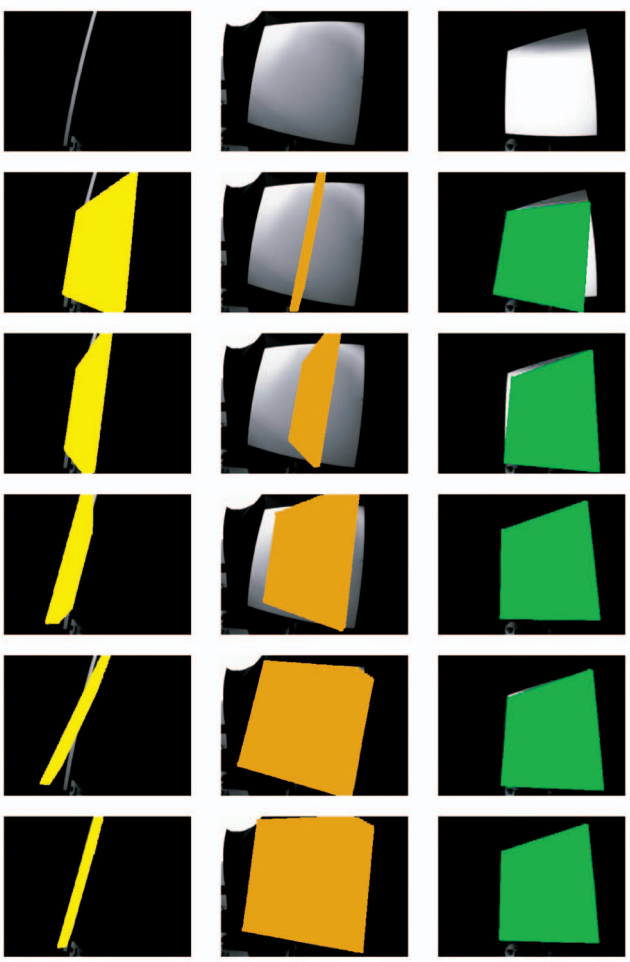

Fig. 4. Column 1: pose 1. Row 1: one camera image shown; row 2: with projection of initialized surface (orange mask); rows 3-5: during evolution of the pose parameters of the surface; and row 6: with converged pose parameters. Columns 2 and 3: same as column 1 for poses 2 and 3, respectively.

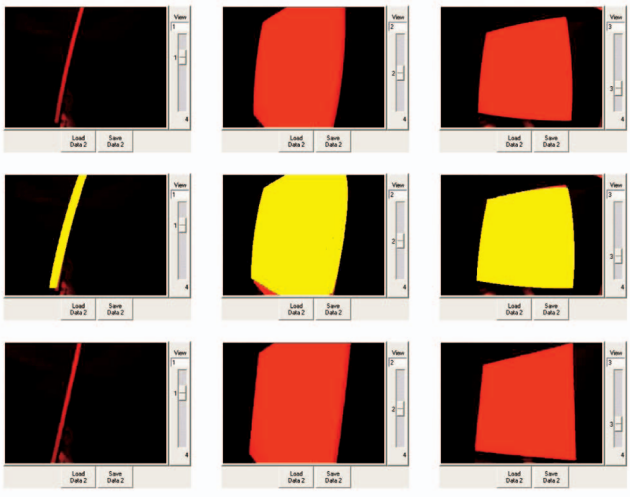

Fig. 5. Pose 1. Row 1: three out of five captured views. Row 2: projected surface after distortion parameters have converged. Row 3: undistorted with the obtained distortion coefficients.

values, the reprojection of the surfaces onto the distorted views results in a better match to the image data and continues to minimize the overall energy. Such images with

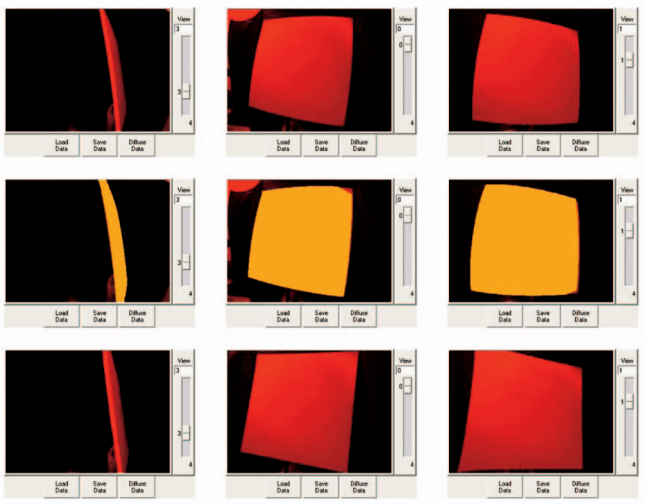

Fig. 6. Pose 2. Row 1: three out of five captured views. Row 2: projected surface after distortion parameters have converged. Row 3: undistorted with the obtained distortion coefficients.

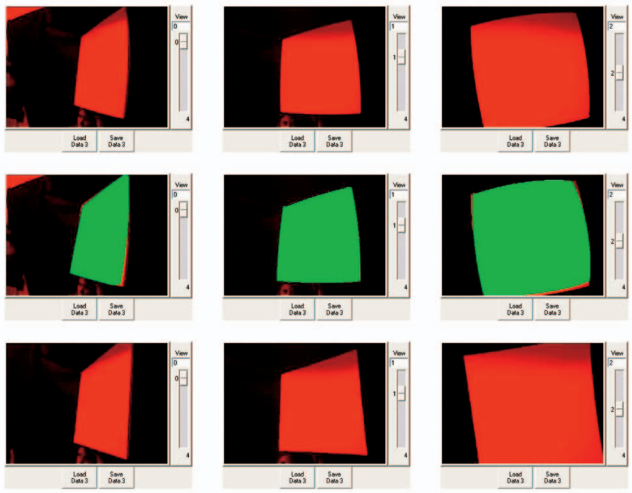

Fig. 7. Pose 3. Row 1: three out of five captured views. Row 2: projected surface after distortion parameters have converged. Row 3: undistorted with the obtained distortion coefficients.

reprojections are shown on the second row of Figs. 5, 6, and 7. The undistorted views are shown as well on the third row. The straightening effect of this operation on the curved lines can be clearly observed in these images.

4 Color Calibration

For color calibration, the differences in absolute colors measured in the response of each camera are modeled by a simple multiplicative factor in each of color red, green, blue (RGB) channel measurements and an additive offset parameter.

The first variation of our energy functional E using this model leads to gradient descent flows

$$\begin{aligned}\frac{\partial E}{\partial \alpha_{i,k}} = &- \int_{R_i} \left[f^k - \left(\alpha_{i,k} I_i^k + \beta_{i,k} \right) \right] I_i^k d\Omega_i \\ &- \int_{R_i^c} \left[b^k - \left(\alpha_{i,k} I_i^k + \beta_{i,k} \right) \right] I_i^k d\Omega_i \end{aligned} \tag{17}$$

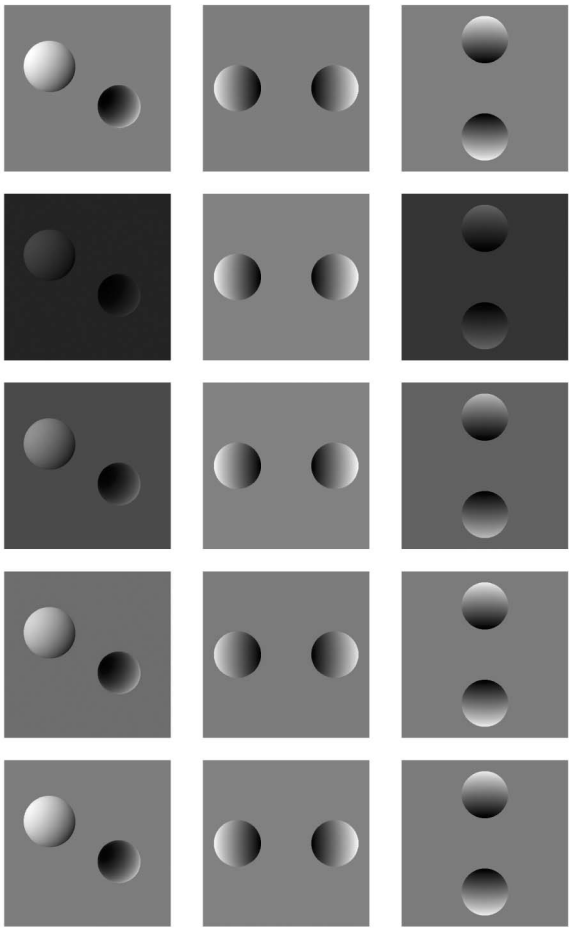

Fig. 8. Row 1: three original views (cameras 1, 7, and 9). Row 2: the same three views after deliberate simulated miscalibration of the grayscales. The same three views while evolving the calibration parameters: rows 3 and 4, intermediate stages; row 5, the views after evolution of the calibration parameters was completed.

$$\frac{\partial E}{\partial \beta_{i,k}} = -\int_{R_i} \left[f^k - \left(\alpha_{i,k} I_i^k + \beta_{i,k} \right) \right] d\Omega_i - \int_{R_i^c} \left[b^k - \left(\alpha_{i,k} I_i^k + \beta_{i,k} \right) \right] d\Omega_i \tag{18}$$

for the color calibration parameters $\alpha_{i,k}$ and $\beta_{i,k}$ for each camera i, and $k \in \{R, G, B\}$, where I_i^k, f^k, and b^k are from one of the three color channels $\{R, G, B\}$. Note that one can extend this framework to RGGB images in a straightforward fashion.

In our test calibration experiments, we utilized white noise additive offsets and multiplicative scaling coefficients to perturb the measured images, thereby exaggerating the effect of color miscalibrations. On a synthetically created example in Fig. 8, where the correct geometry and radiance function are known, we show such miscalibration effects on the original views, and views during the evolution of α_is and β_is in (17) and (18) and views after these parameters have converged. In addition, in Fig. 9, the curves depict the true α and β values for all nine camera views and the convergence of the estimated parameters toward the real values.

Similarly, in Fig. 10, the color cubes with original colors are shown from some camera views first, then shown after their color calibration parameters are perturbed. Finally, the convergence of the color parameters results in a corrected set of colors as shown in the views. Also shown in Fig. 11 are the evolutions of the color calibration parameters for the shown views. We have to note here, again, that due to relative calibration framework among cameras, the updated parameters may not always result in absolute values but still provide useful outputs for the multicamera systems.

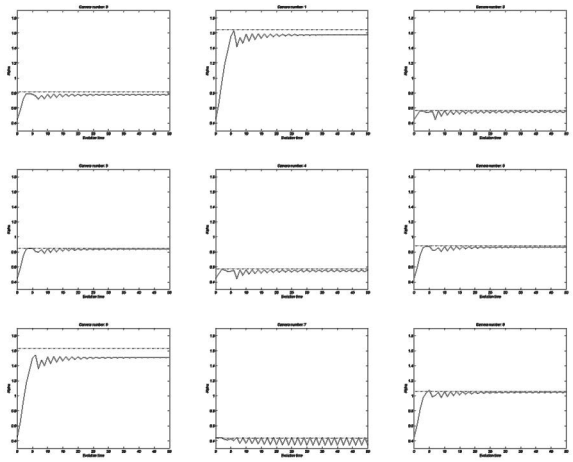

Fig. 9. Evolution of the parameter α for different camera views. True α value is shown as a dotted line.

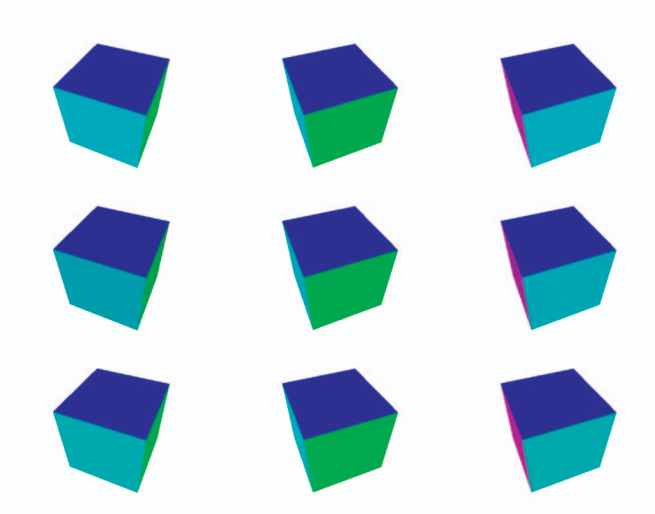

Fig. 10. Some camera views shown during the evolution of the color calibration. Top: original views. Middle: perturbed views. Bottom: final views after convergence. Note the color similarity in the top and bottom rows.

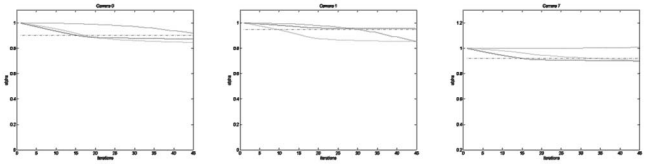

Fig. 11. Evolution of the parameter α for different views for RGB channels of the synthetic color cube. True α value is shown as a dotted line.

5 Intrinsic Parameter Calibration

We show the evolution of three of the main intrinsic camera parameters: focal lengths, denoted by L_u and L_v for each of the coordinates on the image plane, and the skew parameter a. Inclusion of the skew parameter between the two plane coordinates leads to an intrinsics matrix of the form

$$\pi = \begin{pmatrix} L_u & a & u_0 \\ 0 & L_v & v_0 \\ 0 & 0 & 1 \end{pmatrix},$$

then the Jacobian of the perspective transformation becomes (compare to (12))

$$\pi' = \begin{pmatrix} L_u/Z_i & a/Z_i & -L_uX_i/Z_i^2 - aY_i/Z_i^2 \\ 0 & L_v/Z_i & -L_vY_i/Z_i^2 \end{pmatrix} = \frac{1}{Z_i^2}\begin{pmatrix} L_uZ_i & aZ_i & -L_uX_i - aY_i \\ 0 & L_vZ_i & -L_vY_i \end{pmatrix}.$$

The derivatives of the image coordinates with respect to each of the intrinsic parameters are computed from the overall energy functional as before (similar to our derivations of the lens distortion calibration parameters in Section 3):

$$\frac{\partial E}{\partial L_j} = \sum_{k=R,G,B} \int_{C_i} \underbrace{\left[(f^k - I_i^k)^2 - (b^k - I_i^k)^2\right]}_{F^k} \left\langle \frac{\partial \mathbf{x}_i}{\partial L_j}, \mathbf{n}_i \right\rangle ds.$$

For the focal length parameter L_u, we have

$$\begin{aligned} &\left\langle \frac{\partial \mathbf{x}_i}{\partial L_u}, \mathbf{n}_i \right\rangle ds \\ &= \left\langle \frac{\partial \pi \mathbf{C}_i}{\partial L_u}, \frac{\partial}{\partial s} J\pi' C_i \right\rangle ds \\ &= \left\langle \frac{\partial}{\partial L_u}\begin{pmatrix} L_uX_i/Z_i + aY_i/Z_i \\ 0 \end{pmatrix}, \right. \\ &\quad \left. \frac{1}{Z_i{}^2}\begin{bmatrix} 0 & L_vZ_i & -L_vY_i \\ -L_uZ_i & -aZ_i & L_uX_i + aY_i \end{bmatrix}\begin{bmatrix} \partial \mathbf{X}_i \\ \partial s \end{bmatrix} \right\rangle ds \qquad (19) \\ &= \frac{1}{Z_i^3}\left\langle \begin{bmatrix} 0 & -L_uZ_i \\ L_vZ_i & -aZ_i \\ -L_vY_i & L_uX_i + aY_i \end{bmatrix}\begin{pmatrix} X_i \\ 0 \end{pmatrix}, \frac{\partial \mathbf{X}_i}{\partial s} \right\rangle ds \\ &= \frac{1}{Z_i^3}\left\langle L_v \begin{pmatrix} 0 \\ Z_iX_i \\ -Y_iX_i \end{pmatrix}, \frac{\partial \mathbf{X}_i}{\partial s} \right\rangle ds. \end{aligned}$$

Noting that

$$\begin{pmatrix} 0 \\ Z_iX_i \\ -Y_iX_i \end{pmatrix} = \mathbf{X}_i \times \begin{pmatrix} X_i \\ 0 \\ 0 \end{pmatrix},$$

then for the focal length parameter L_u, we obtain

$$\begin{aligned} \left\langle \frac{\partial \mathbf{x}_i}{\partial L_u}, \mathbf{n}_i \right\rangle ds &= \frac{1}{Z_i^3}\left\langle -L_v \begin{pmatrix} X_i \\ 0 \\ 0 \end{pmatrix} \times \mathbf{X}_i, \frac{\partial \mathbf{X}_i}{\partial s} \right\rangle ds \\ &= -\frac{L_v}{Z_i^3}\left\langle \mathbf{X}_i \times \frac{\partial \mathbf{X}_i}{\partial s}, \begin{pmatrix} X_i \\ 0 \\ 0 \end{pmatrix} \right\rangle ds \qquad (20) \\ &= -\frac{L_v\|\mathbf{X}_i\|}{Z_i^3}\left\langle \mathbf{N}_i, \begin{pmatrix} X_i \\ 0 \\ 0 \end{pmatrix} \right\rangle ds. \end{aligned}$$

Due to the skew parameter, the equations for the second focal length parameter L_v will be slightly different. This time, incorporating the derivative with respect to L_v in (19), we derive

$$\begin{aligned} \left\langle \frac{\partial \mathbf{x}_i}{\partial L_v}, \mathbf{n}_i \right\rangle ds &= \frac{1}{Z_i^3}\left\langle \begin{bmatrix} 0 & -L_uZ_i \\ L_vZ_i & -aZ_i \\ -L_vY_i & L_uX_i + aY_i \end{bmatrix}\begin{pmatrix} 0 \\ Y_i \end{pmatrix}, \frac{\partial \mathbf{X}_i}{\partial s} \right\rangle ds \\ &= \frac{1}{Z_i^3}\left\langle \begin{pmatrix} -L_uZ_iY_i \\ -aZ_iY_i \\ L_uX_iY_i + aY_i^2 \end{pmatrix}, \frac{\partial \mathbf{X}_i}{\partial s} \right\rangle ds. \end{aligned}$$

Again, noting that

$$\begin{pmatrix} -L_uY_iZ_i \\ -aZ_iY_i \\ L_uX_iY_i + aY_i^2 \end{pmatrix} = \mathbf{X}_i \times \begin{pmatrix} -aY_i \\ L_uY_i \\ 0 \end{pmatrix},$$

then for the focal length parameter L_v, we have

$$\left\langle \frac{\partial \mathbf{x}_i}{\partial L_v}, \mathbf{n}_i \right\rangle ds = -\frac{\|\mathbf{X}_i\|}{Z_i^3}\left\langle \mathbf{N}_i, \begin{pmatrix} -aY_i \\ L_uY_i \\ 0 \end{pmatrix} \right\rangle ds. \qquad (21)$$

Note that when the skew parameter a is 0, which is a widely used convention, the above equation reduces to a symmetric form of (20) derived for L_u.

Finally, we derive similarly the update equations for the skew parameter a:

$$\begin{aligned} \left\langle \frac{\partial \mathbf{x}_i}{\partial a}, \mathbf{n}_i \right\rangle ds &= \frac{1}{Z_i^3}\left\langle \begin{bmatrix} 0 & -L_uZ_i \\ L_vZ_i & -aZ_i \\ -L_vY_i & L_uX_i + aY_i \end{bmatrix}\begin{pmatrix} Y_i \\ 0 \end{pmatrix}, \frac{\partial \mathbf{X}_i}{\partial s} \right\rangle \\ &= \frac{1}{Z_i^3}\left\langle L_v \begin{pmatrix} 0 \\ Z_iY_i \\ -Y_i^2 \end{pmatrix}, \frac{\partial \mathbf{X}_i}{\partial s} \right\rangle ds. \end{aligned}$$

This time, noting that

$$\begin{pmatrix} 0 \\ Z_iY_i \\ -Y_i^2 \end{pmatrix} = \mathbf{X}_i \times \begin{pmatrix} Y_i \\ 0 \\ 0 \end{pmatrix},$$

then for the skew parameter a, we have

$$\left\langle \frac{\partial \mathbf{x}_i}{\partial a}, \vec{n_i} \right\rangle ds = -\frac{L_v\|\mathbf{X}_i\|}{Z_i^3}\left\langle \mathbf{N}_i, \begin{pmatrix} Y_i \\ 0 \\ 0 \end{pmatrix} \right\rangle ds. \qquad (22)$$

The final evolution equations for the three intrinsic parameters for each camera i are then given by

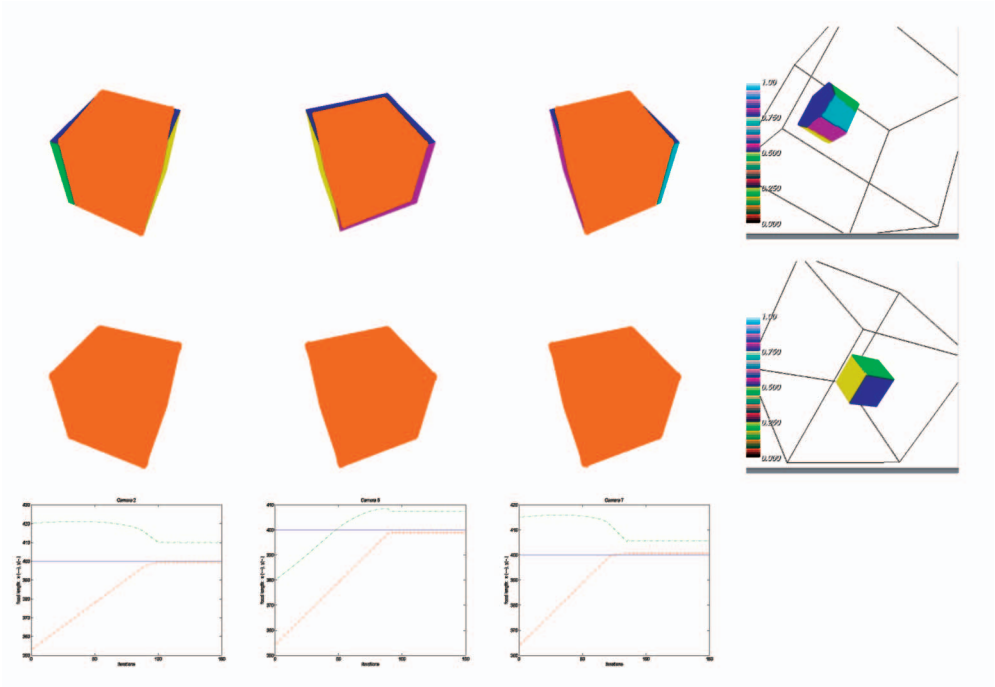

Fig. 12. Top: three camera views shown during the evolution of the intrinsic parameters of an initial cube with projections from the initial surface. Middle: final views after convergence of the intrinsic parameters of the surface. Also shown at the bottom are the evolution of the two focal length parameters for each shown camera view (red and green curves) along with the true (blue curve) focal lengths.

$$\frac{\partial L_{u,i}}{\partial t} = -\sum_{k=R,G,B} \int_{C_i} F^k \frac{L_v}{Z_i^3} \left\langle \|\mathbf{X}_i\| \mathbf{N}_i, \begin{pmatrix} X_i \\ 0 \\ 0 \end{pmatrix} \right\rangle ds, \tag{23}$$

$$\frac{\partial L_{v,i}}{\partial t} = -\sum_{k=R,G,B} \int_{C_i} F^k \frac{1}{Z_i^3} \left\langle \|\mathbf{X}_i\| \mathbf{N}_i, \begin{pmatrix} -aY_i \\ L_u Y_i \\ 0 \end{pmatrix} \right\rangle ds, \tag{24}$$

$$\frac{\partial a_i}{\partial t} = -\sum_{k=R,G,B} \int_{C_i} F^k \frac{L_v}{Z_i^3} \left\langle \|\mathbf{X}_i\| \mathbf{N}_i, \begin{pmatrix} Y_i \\ 0 \\ 0 \end{pmatrix} \right\rangle ds. \tag{25}$$

In Fig. 12, a synthetic color cube example is shown. The intrinsic parameters, focal lengths L_u and L_v, are initialized to perturbed values and when the intrinsic calibration update equations have converged, both the projections of the cube surface onto the images and the evolution of the focal lengths are shown.

6 Extrinsic Camera Calibration

We now consider the same energy functional as a function of the extrinsic calibration parameters $\Lambda_i = (\lambda_{i1}, \ldots, \lambda_{i6})$ for each camera image I_i. Notice that the only term in our energy functional E, which depends upon Λ_i, is the corresponding fidelity term in E_{data} (due to the dependence of π_i^{-1}), assuming a constant background radiance in the scene:

$$\begin{aligned} E_{data,i}(S, f, b, \Lambda_i) = &\sum_{k=R,G,B} \int_{R_i} \left(f^k(\pi_i^{-1}(\hat{\mathbf{x}})) - I_i^k(\hat{\mathbf{x}})\right)^2 d\Omega_i \\ &+ \int_{R_i^c} \left(b^k - I_i^k(\hat{\mathbf{x}})\right)^2 d\Omega_i, \end{aligned} \tag{26}$$

where $\hat{\mathbf{x}}$ denotes image coordinates as before (for simplicity of discussion, distortion $D = 1$).

6.1 Initial Expression of Gradient

If we let $\hat{c}_i = \partial R_i$ denote the boundary of R_i, then we may express the partial derivative of E with respect to one of the calibration parameters λ_{ij} as follows:

$$\begin{aligned} \frac{\partial E}{\partial \lambda_{ij}} &= \text{boundary term} + \text{foreground term} \\ &= \sum_{k=R,G,B} \int_{\hat{c}_i} \left(\left(f^k(\pi_i^{-1}(\hat{\mathbf{x}})) - I_i^k(\hat{\mathbf{x}})\right)^2 - \left(b^k - I_i^k(\hat{\mathbf{x}})\right)^2 \right) \\ &\quad \left\langle \frac{\partial \hat{c}_i}{\partial \lambda_{ij}}, \hat{\mathbf{n}}_i \right\rangle d\hat{s} \\ &\quad + \sum_{k=R,G,B} 2 \int_{R_i} \left(f^k(\pi_i^{-1}(\hat{\mathbf{x}})) - I_i^k(\hat{\mathbf{x}})\right) \\ &\quad \left\langle \nabla_S f^k\left(\pi_i^{-1}(\hat{\mathbf{x}})\right), \frac{\partial}{\partial \lambda_{ij}} \pi_i^{-1}(\hat{\mathbf{x}}) \right\rangle d\Omega_i. \end{aligned} \tag{27}$$

In the boundary term, $d\hat{s}$ denotes the arclength measure of $\hat{c}_i$, and $\hat{\mathbf{n}}_i$ denotes its outward unit normal. In the foreground term, ∇_S denotes the natural gradient operator on the surface S.

6.2 Rewriting the Boundary Term

Ultimately, we will compute all quantities by integrating along the current estimate of the surface since that is the actual object represented by our data structures. Thus, it is more convenient to express the contour integral around $\hat{c}_i(\hat{s})$ in the image plane as a contour integral around $C_i(s)$ on the surface S instead (where $\pi_i(C_i) = \hat{c}_i$ and where s is the arclength parameter of C_i). They may be related as follows:

$$\left\langle \frac{\partial\,\hat{c}_i}{\partial\lambda_{ij}}, \hat{\mathbf{n}}_i \right\rangle d\hat{s} = \left\langle \frac{\partial}{\partial\lambda_{ij}}\pi_i(C_i), \frac{\partial}{\partial s} J\pi_i(C_i) \right\rangle ds,$$

$$\text{where } J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

$$= \left\langle \frac{1}{Z_i^2}\begin{bmatrix} Z_i & 0 & -X_i \\ 0 & Z_i & -Y_i \end{bmatrix} \frac{\partial\,\mathbf{x}_i}{\partial\lambda_{ij}}, \frac{1}{Z_i^2}\begin{bmatrix} 0 & Z_i & -Y_i \\ -Z_i & 0 & X_i \end{bmatrix} \frac{\partial \mathbf{X}_i}{\partial s} \right\rangle ds$$

$$= \frac{1}{Z_i^3}\left\langle \frac{\partial \mathbf{X}_i}{\partial s}, \begin{bmatrix} 0 & -Z_i & Y_i \\ Z_i & 0 & -X_i \\ -Y_i & X_i & 0 \end{bmatrix} \frac{\partial \mathbf{X}_i}{\partial \lambda_{ij}} \right\rangle ds$$

$$= \frac{1}{Z_i^3}\left\langle \frac{\partial \mathbf{X}_i}{\partial s}, \frac{\partial \mathbf{X}_i}{\partial \lambda_{ij}} \times \mathbf{X}_i \right\rangle ds = \frac{1}{Z_i^3}\left\langle \frac{\partial \mathbf{X}_i}{\partial \lambda_{ij}}, \mathbf{X}_i \times \frac{\partial \mathbf{X}_i}{\partial s} \right\rangle ds$$

$$= \frac{\|\mathbf{X}_i\|}{Z_i^3}\left\langle \frac{\partial \mathbf{X}_i}{\partial \lambda_{ij}}, \mathbf{N}_i \right\rangle ds$$

(since $\mathbf{X}_i$ and $\dfrac{\partial \mathbf{X}_i}{\partial s}$ are perpendicular tangent vectors to S).

Thus, the boundary term written as an integral on the surface S (along the occluding contour C_i) has the following form:

$$\frac{\partial \lambda_{ij}}{\partial t} = \sum_{k=R,G,B} \int_{C_i} \left(\left(f^k - I_i^k\right)^2 - \left(b^k - I_i^k\right)^2 \right) \frac{\|\mathbf{X}_i\|}{Z_i^3} \left\langle \frac{\partial g_i}{\partial \lambda_{ij}}, \mathbf{N}_i \right\rangle ds, \tag{28}$$

which is also the update equations for the extrinsic parameter j for camera i with a piecewise constant assumption on the foreground and the background radiance.

6.3 Rewriting the Foreground Term

The first step in rewriting the foreground/background integrals is to reexpress the derivative of the back-projected 3D point $\mathbf{X} = \pi_i^{-1}(\hat{\mathbf{x}}, \Lambda_i)$ with respect to the calibration parameter λ_{ij} in terms of the derivative of the forward projection $\pi_i(\mathbf{x}, \Lambda_i) = \pi(g_i(\mathbf{X}, \Lambda_i))$, since π_i has an analytic form while π_i^{-1} does not. We begin by fixing a 2D image point $\hat{\mathbf{x}}$ and note that

$$\hat{\mathbf{x}} = \pi_i\left(\mathbf{X}(\hat{\mathbf{x}}, \Lambda_i), \Lambda_i\right)$$
$$\left(\text{where } \mathbf{X}(\hat{\mathbf{x}}, \Lambda_i) = \pi_i^{-1}(\hat{\mathbf{x}}, \Lambda_i) = g_i^{-1}\left(\pi^{-1}(\hat{\mathbf{x}}), \Lambda_i\right)\right)$$

and, thus, differentiation with respect to λ_{ij} yields

$$0 = \frac{\partial}{\partial \lambda_{ij}}\pi_i(\mathbf{X}, \Lambda_i) = \frac{\partial \pi_i}{\partial \mathbf{X}}\frac{\partial \mathbf{X}}{\partial \lambda_{ij}} + \frac{\partial \pi_i}{\partial \lambda_{ij}}$$

$$= \frac{1}{Z_i^2}\begin{bmatrix} Z_i & 0 & -X_i \\ 0 & Z_i & -Y_i \end{bmatrix} \frac{\partial g_i}{\partial \mathbf{X}}\frac{\partial \mathbf{X}}{\partial \lambda_{ij}} + \frac{1}{Z_i^2}\begin{bmatrix} Z_i & 0 & -X_i \\ 0 & Z_i & -Y_i \end{bmatrix} \frac{\partial g_i}{\partial \lambda_{ij}}$$

$$\begin{bmatrix} Z_i & 0 & -Z_i \\ 0 & Z_i & -Y_i \end{bmatrix} \frac{\partial g_i}{\partial \mathbf{X}}\frac{\partial \mathbf{X}}{\partial \lambda_{ij}} = -\begin{bmatrix} Z_i & 0 & -X_i \\ 0 & Z_i & -Y_i \end{bmatrix} \frac{\partial g_i}{\partial \lambda_{ij}}.$$

(29)

Notice, though, that (29) does not uniquely specify $\partial\mathbf{X}/\partial\lambda_{ij}$ but merely gives a necessary condition. We must supplement (29) with the additional constraint that $\partial\mathbf{X}/\partial\lambda_{ij}$ must be orthogonal to the unit normal $\mathbf{N}$ of S at the point $\mathbf{X}$ in order to obtain a unique solution:

$$\frac{\partial \mathbf{X}}{\partial \lambda_{ij}} \cdot \mathbf{N} = 0 \qquad \left(\text{or, equivalently, } \frac{\partial g_i}{\partial \mathbf{X}}\frac{\partial \mathbf{X}}{\partial \lambda_{ij}} \cdot \mathbf{N}_i = 0\right). \tag{30}$$

Now, combining (29) and (30), we have

$$\begin{bmatrix} Z_i & 0 & -X_i \\ 0 & Z_i & -Y_i \\ N_{ix} & N_{iy} & N_{iz} \end{bmatrix} \frac{\partial g_i}{\partial \mathbf{X}}\frac{\partial \mathbf{X}}{\partial \lambda_{ij}} = -\begin{bmatrix} Z_i & 0 & -X_i \\ 0 & Z_i & -Y_i \\ 0 & 0 & 0 \end{bmatrix} \frac{\partial g_i}{\partial \lambda_{ij}}$$

$$\frac{\partial \mathbf{X}}{\partial \lambda_{ij}} = -\left(\frac{\partial g_i}{\partial \mathbf{X}}\right)^{-1} \begin{bmatrix} Z_i & 0 & -X_i \\ 0 & Z_i & -Y_i \\ N_{ix} & N_{iy} & N_{iz} \end{bmatrix}^{-1} \begin{bmatrix} Z_i & 0 & -X_i \\ 0 & Z_i & -Y_i \\ 0 & 0 & 0 \end{bmatrix} \frac{\partial g_i}{\partial \lambda_{ij}}$$

$$\frac{\partial \mathbf{X}}{\partial \lambda_{ij}} = \frac{-\left(\frac{\partial g_i}{\partial \mathbf{X}}\right)^{-1}}{Z_i(\mathbf{X}_i \cdot \mathbf{N}_i)} \begin{bmatrix} Z_i N_{iz} + Y_i N_{iy} & -X_i N_{iy} & X_i z_i \\ -Y_i N_{ix} & Z_i N_{iz} + X_i N_{ix} & Y_i z_i \\ -Z_i N_{ix} & -Z_i N_{iy} & Z_i z_i \end{bmatrix}$$

$$\begin{bmatrix} Z_i & 0 & -X_i \\ 0 & Z_i & -Y_i \\ 0 & 0 & 0 \end{bmatrix} \frac{\partial g_i}{\partial \lambda_{ij}}\frac{\partial \mathbf{X}}{\partial \lambda_{ij}} = \frac{-\left(\frac{\partial g_i}{\partial \mathbf{X}}\right)^{-1}}{\mathbf{X}_i \cdot \mathbf{N}_i}$$

$$\begin{bmatrix} \mathbf{X}_i \cdot N_i - X_i N_{ix} & -X_i N_{iy} & -X_i N_{iz} \\ -Y_i N_{ix} & \mathbf{X}_i \cdot N_i - Y_i N_{iy} & -Y_i N_{iz} \\ -Z_i N_{ix} & -Z_i N_{iy} & \mathbf{X}_i \cdot N_i - Z_i N_{iz} \end{bmatrix} \frac{\partial g_i}{\partial \lambda_{ij}}$$

$$\frac{\partial \mathbf{X}}{\partial \lambda_{ij}} = -\left(\frac{\partial g_i}{\partial \mathbf{X}}\right)^{-1} \left(\mathbf{I} - \frac{\mathbf{X}_i \otimes \mathbf{N}_i}{\mathbf{X}_i \cdot \mathbf{N}_i}\right) \frac{\partial g_i}{\partial \lambda_{ij}},$$

(31)

where $\otimes$ is the Kronecker product and $\mathbf{I}$ is the 3×3 identity matrix.

The second step proceeds in the same manner as outlined earlier in rewriting the data fidelity terms in E_{data} by noting that the measure in the image domain $d\Omega_i$ and the area measure on the surface dA are related by $d\Omega_i = \sigma(\mathbf{x}_i, \mathbf{N}_i)dA$, where $\sigma(\mathbf{X}_i, \mathbf{N}_i) = (\mathbf{X}_i \cdot \mathbf{N}_i)/Z_i^3$. Then, the foreground term in (27) is given by

$$\sum_{k=R,G,B} 2\int_{R_i} \left(f^k - I_i^k\right) \left\langle \nabla_s f^k\left(\pi_i^{-1}(\hat{\mathbf{x}})\right), \frac{\partial}{\partial \lambda_{ij}}\pi_i^{-1}(\hat{\mathbf{x}}) \right\rangle d\Omega_i$$
$$= \sum_{k=R,G,B} 2\int_{\pi_i^{-1}(R_i)} \left(f^k - I_i^k\right) \left\langle \nabla_s f^k(\mathbf{X}), \frac{\partial \mathbf{X}}{\partial \lambda_{ij}} \right\rangle \frac{\mathbf{X}_i \cdot \mathbf{N}_i}{Z_i^3}\, dA.$$

(32)

Therefore, the following foreground term will be added to the update equation of the extrinsic parameter in (28):

$$\frac{\partial \lambda_{ij}}{\partial t} = \sum_{k=R,G,B} -2\int_{\pi_i^{-1}(R_i)} \frac{\left(f^k - I_i^k\right)}{Z_i^3} \left\langle \nabla_s f^k(\mathbf{x}), \left(\frac{\partial g_i}{\partial \mathbf{X}}\right)^{-1} \left((\mathbf{X}_i \cdot \mathbf{N}_i)\frac{\partial g_i}{\partial \lambda_{ij}} - \left(\frac{\partial g_i}{\partial \lambda_{ij}} \cdot \mathbf{N}_i\right)\mathbf{X}_i \right) \right\rangle dA.$$

(33)

In Fig. 13, several photos from a set of 32 images of a toy skater doll are shown. When the initial extrinsic parameters are off, as observed in the projections of the foreground object onto the images (shown by an orange mask), a visual hull created using the uncorrected extrinsic camera parameters is significantly away from the real doll surface. After the extrinsic calibration equations (28) plus (33) are evolved to convergence, visual hull created using the updated extrinsic parameters demonstrates the correction and true refinement provided by the derived equations. In Fig. 14, we depict the extrinsic refinement stability by showing the

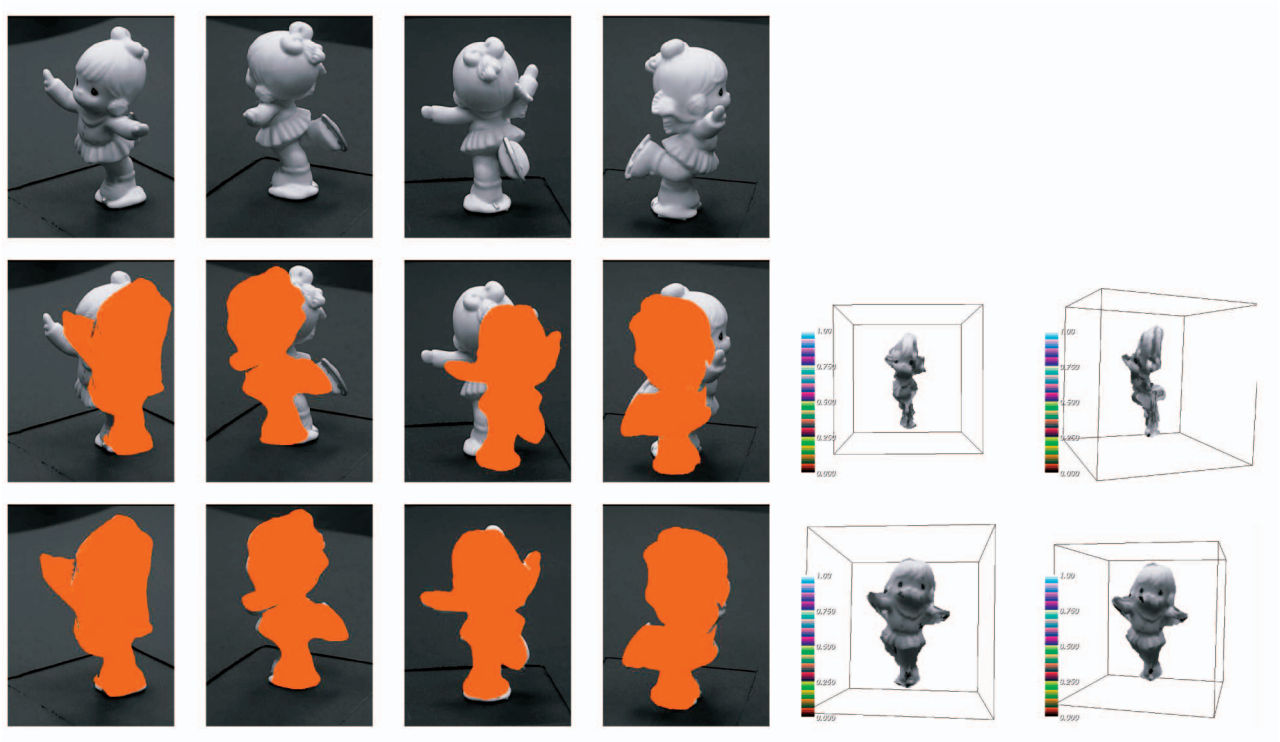

Fig. 13. Four camera views shown (top) during the evolution of the extrinsic parameters of an initial surface of a toy skater. Row 2: views shown with projections from the initial surface. Row 3: final views after convergence of the extrinsic camera parameters. Visual hull generated using the miscalibrated initial extrinsic parameters (row 2, right); visual hull generated using the converged extrinsic parameters (row 3, right).

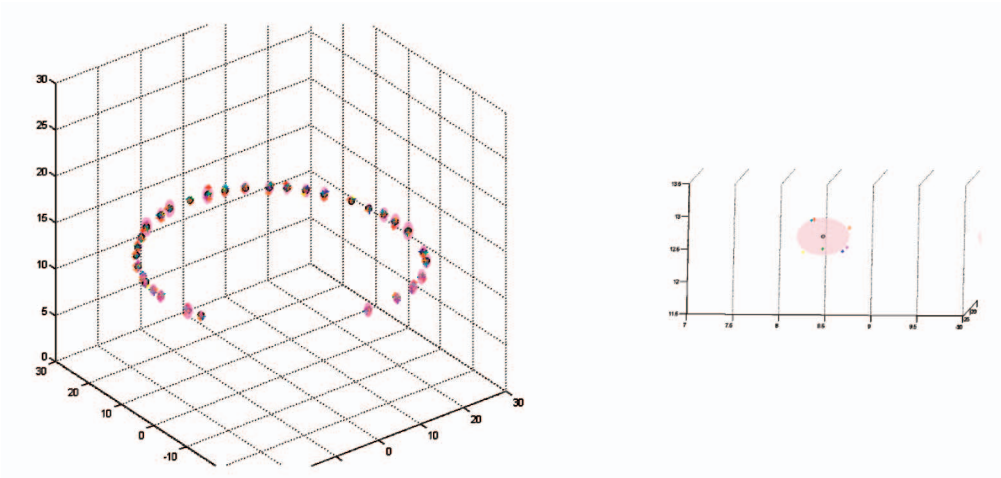

Fig. 14. Uncertainty ellipsoids drawn around each camera center for the toy skater data show the extrinsic refinement stability (right: zoomed into one camera's perturbations).

uncertainty ellipsoids drawn around each camera center. Parameters were perturbed in the x, y, and z directions randomly several times and converged properly for variations up to 8 percent.

7 Results and Conclusions

The toy skater example shown in Fig. 15 demonstrates the simultaneous evolution of the extrinsic and intrinsic parameters for the 32 cameras, along with the projections of the foreground surface. The visual hulls created with, again, the initial set of camera parameters and the evolved set of camera parameters display a correct refinement of the camera parameters.

For most of the experiments, we utilized a 128^3 volume and a $140 \times 150 \times 360$ volume for the Bust data set. With a volumetric signed distance representation in our C++ implementation without any code optimization on a Pentium 2.40-GHz processor, each single iteration to compute all calibration gradients takes on the order of 10 seconds, depending on the number of camera views as well, and convergence takes about 50-400 iterations, depending on the initialization, hence, a computation time of about 8-60 minutes. However, a mesh representation on the object may be

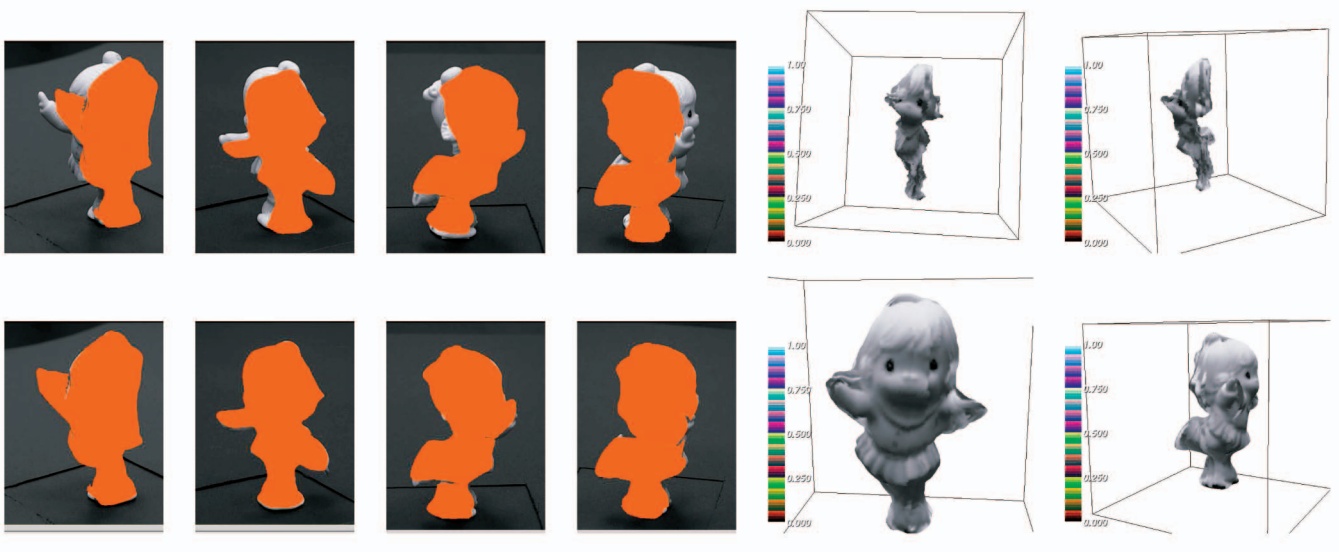

Fig. 15. Row 1: four camera views during the evolution of the extrinsic plus intrinsic parameters of a toy skater with projections of the initial surface. Row 2: final views after convergence of the camera parameters. Visual hull generated using the miscalibrated initial parameters (row 1, right); visual hull generated using the converged parameters (row 2, right).

easier to work with since the parameter update equations we derived are ODEs.

A common issue for any calibration procedure is that when there are shape symmetries or constant radiance on the object, camera pose parameter estimation is not stable; however, these do not affect the 3D reconstruction (for example, multiple views on a sphere do not allow estimating camera pose, but they still allow estimating the shape of the sphere). Regarding the radiance assumptions, because our algorithm integrates information globally on the entire collection of images, it is far less sensitive to this accident than algorithms based on local statistics such as point feature correspondences. Therefore, symmetries are not an obstacle since our goal is not to obtain the absolute calibration parameters (ground truth) but to help refine 3D reconstruction. From this perspective, the only criterion of concern is the reprojection error. We experimented with a full-turn head sequence using Intel's Van Gogh Bust data for testing the issue of shape complexity. We utilized only 16 camera views from the available 330 camera images for ease and speed of computations. We computed reprojection errors: a Type II error (error of omission) and Type I error (error of commission) by counts of voxels for several camera views used during our experiments both after perturbation of the camera parameters and after evolution of the parameters as shown in Table 1. After the refinement stage, the Type I error dropped by 95 percent, and the Type II error dropped by 40 percent. As remarked above, our goal is not to obtain absolute camera parameters but to help 3D reconstruction algorithm to obtain objects correctly, which is achieved.

The Bust data comprise numerous views and this facilitated the following experiment to show the practicality of our calibration correction. For the three camera views, out of the 16 views, we deliberately used wrong camera calibration parameters, which belong to that of the neighbor views in the sequence in Fig. 16. This represents a possible perturbation in a real-life scenario, that is, the cameras are accidentally moved a little bit after the calibration and the views that are captured afterward are a little bit off. The 3D reconstruction of the Bust object on the top right shows the erroneous surfaces obtained in this case. With our coordinated refinement of the extrinsic parameters using (28) and (33), the improvements in the reprojection errors and the 3D reconstruction are observed in Fig. 16.

A real color calibration experiment is carried out using the HP Labs stereo rig system. We captured images, shown in Fig. 17, of the color calibration object from five cameras. Notice that the first picture is somewhat darker than the others, second and third pictures appear lighter, and there

TABLE 1
Type I and Type II Errors in Counts of Voxels for Several Camera Views for Bust Data (Fig. 16) after Perturbation of Camera Parameters (Initial) and after Evolution of Parameters (Final)

	Type II error		Type I error	
Camera	Initial	Final	Initial	Final
Cam 1	3584	47	5593	4476
Cam 2	13831	77	17762	4807
Cam 55	2841	191	5574	4618
Cam 77	8894	46	14013	14700
Cam 105	8339	1344	10032	6724
Cam 166	1005	170	4021	4001
Cam 200	4901	467	7971	6414
Cam 207	7339	97	9764	6783
Cam 239	9251	213	10615	6492
Cam 244	12467	263	16956	3110
Cam 321	1733	65	6501	7365

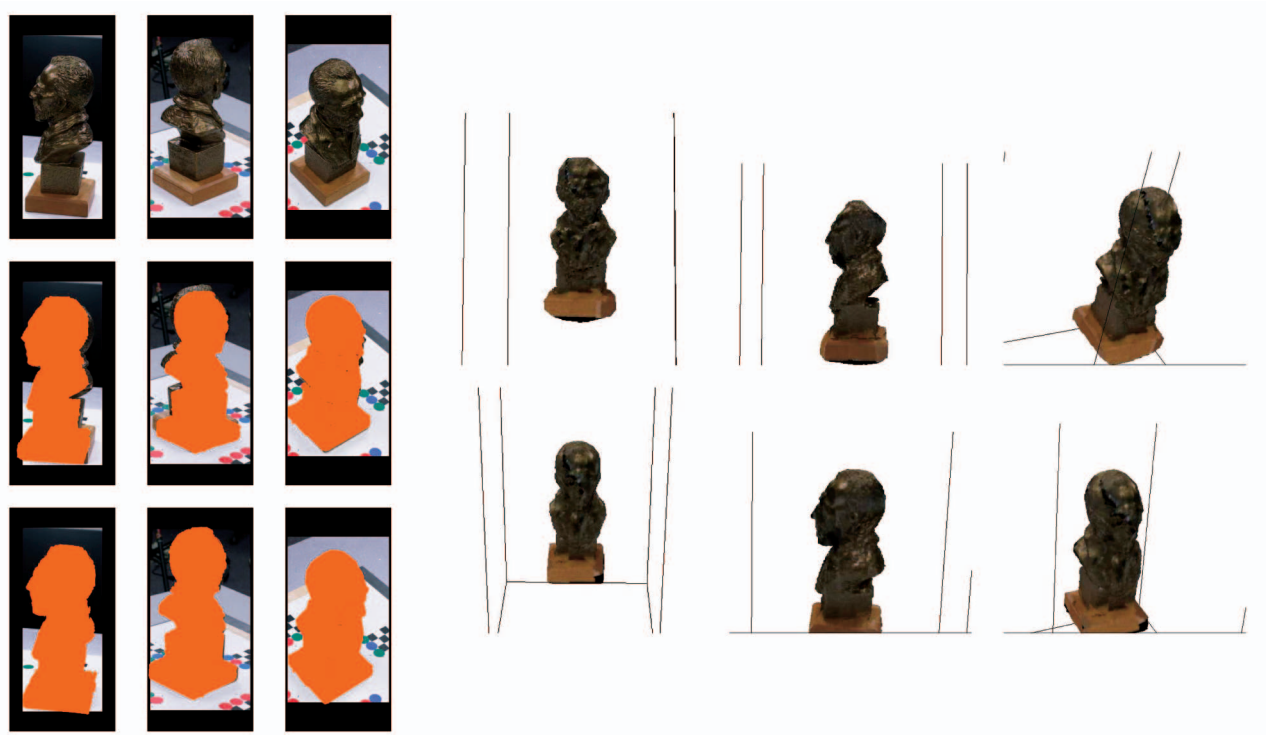

Fig. 16. Camera views 78, 167, and 240 in the top row are used deliberately with camera calibration parameters of camera views 77, 166, and 239 of the Van Gogh Bust data set. Top: three camera views shown with projections from the initial surface in row 2; here, note the resulting initial mismatch in projected silhouettes. Row 3: final views after convergence of the camera parameters. Visual hull surfaces obtained by using wrong calibration parameters for views 78, 167, and 240 on the right (top row) and surfaces with corrected calibration parameters in the bottom row.

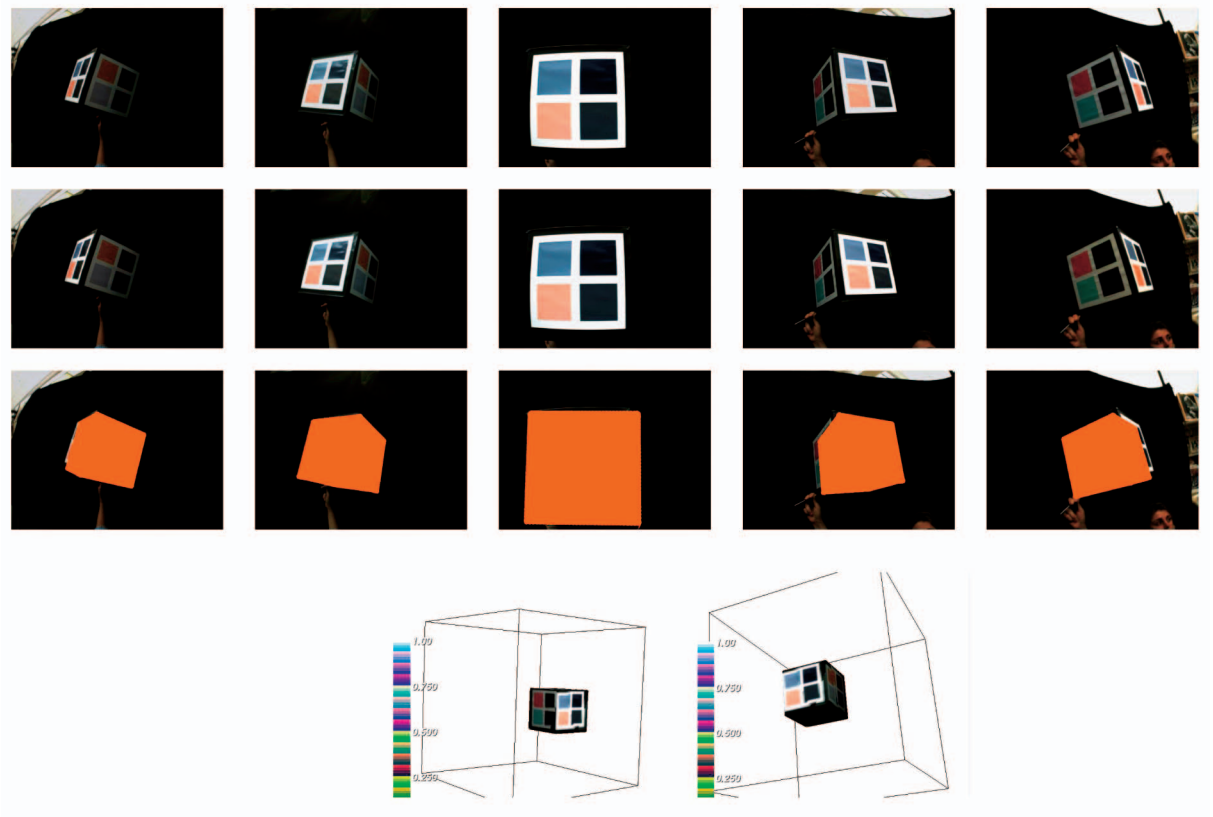

Fig. 17. Some camera views shown during the evolution of the color calibration parameters of the HP color calibration object surface. Top: five camera views. Row 2: final views after convergence of the extrinsic camera parameters. Row 3: same shown with projections of the converged cube. Bottom: color calibration cube with reconstructed radiance on the surface from two different vantage points.

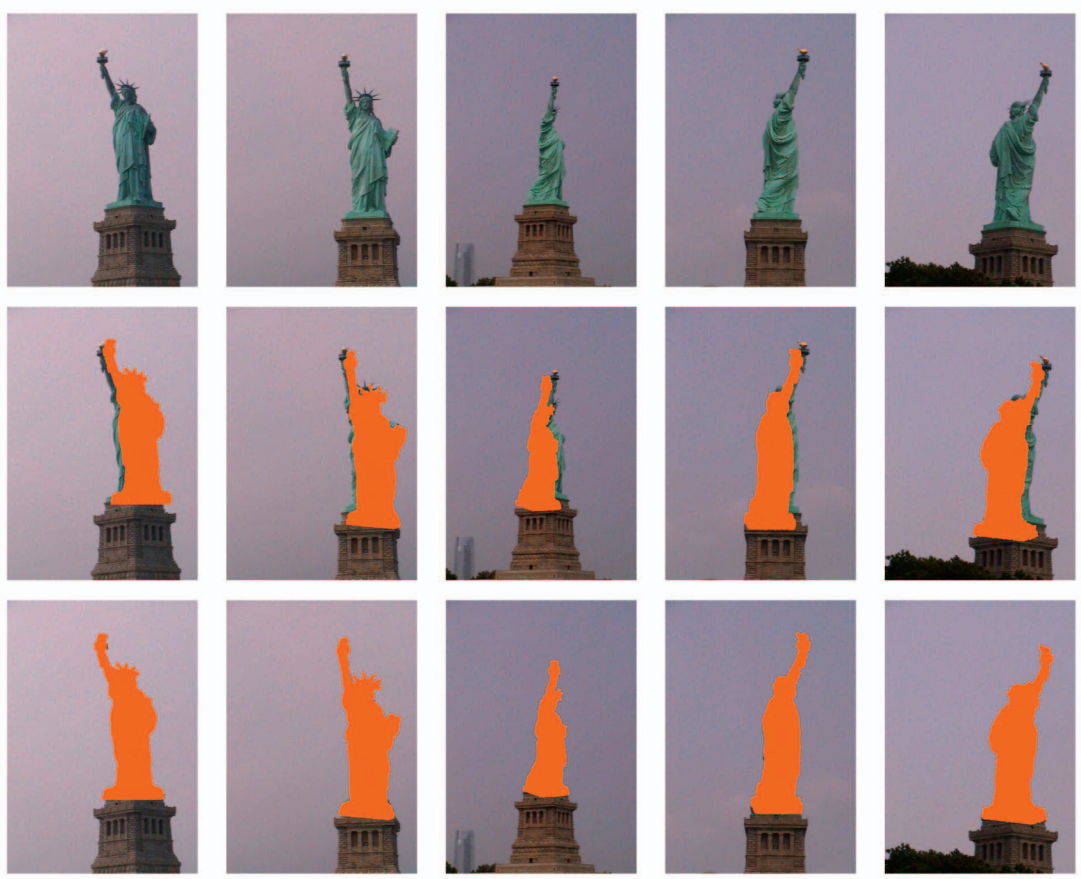

Fig. 18. Some camera views shown during the evolution of the camera calibration parameters of the Statue of Liberty surface. Top: five camera views shown with projections from the initial surface in row 2. Row 3: final views after convergence of the camera parameters.

is a color mismatch. A cube surface is rigidly registered with the scene; also, the radiance function on the cube is estimated as shown in the bottom row of Fig. 17. The second row shows views after the evolution of color calibration coefficients are completed. The third row shows the projections of the model surface onto the views. It can be visually assessed that color responses of the cameras have achieved a balancing effect and helped to obtain a better texture mapping as well.

Next, we demonstrate a calibration experiment using pictures from a handheld camera with no camera calibration information available. In this scenario, the variational calibration techniques we presented require some rough initial values that we obtained through a self-calibration software currently under development. We have a set of 13 pictures taken around the Statue of Liberty, covering about $220^\circ/360^\circ$ of a circle around the statue, with a few of the views shown in Fig. 18.[2] We obtained initial camera parameters: extrinsics and intrinsics, including the skew parameter. A rough calibration results in the projections shown in Fig. 18. After evolution of the camera parameters—extrinsics, intrinsics, including the skew parameter, and color parameters—the comparison is done with the visual hulls of before and after evolution camera parameters in Fig. 19. One can observe the correction in the Statue of Liberty surface with a better set of camera parameters obtained with the derived update equations throughout the paper. We also show blow-up regions in Fig. 20 from some of the camera views before and after the evolution of the color camera parameters and the colors are modified toward achieving some relative agreement among the cameras, which can, however, only be subjectively judged.

2. We thank our colleague Irwin Sobel at HP Labs for providing these pictures.

7.1 Discussions

One may argue that the requirement of some rough initial extrinsic and intrinsic camera parameters limits the usability of this technique. However, the refinement or correction of camera parameters from a perturbed state of a previous calibration is a real-world problem that constantly presents obstacles to the usage of multiple camera systems. After a very good initial calibration, the cameras over time may see small changes in their parameters. For instance, extrinsic parameters will often be changed particularly due to unwanted accidental motion. Similarly, the intrinsics and color parameters of the cameras may go through small variations due to ambient conditions and wear-off. Therefore, the presented camera calibration framework proves to be a useful tool for multicamera systems.

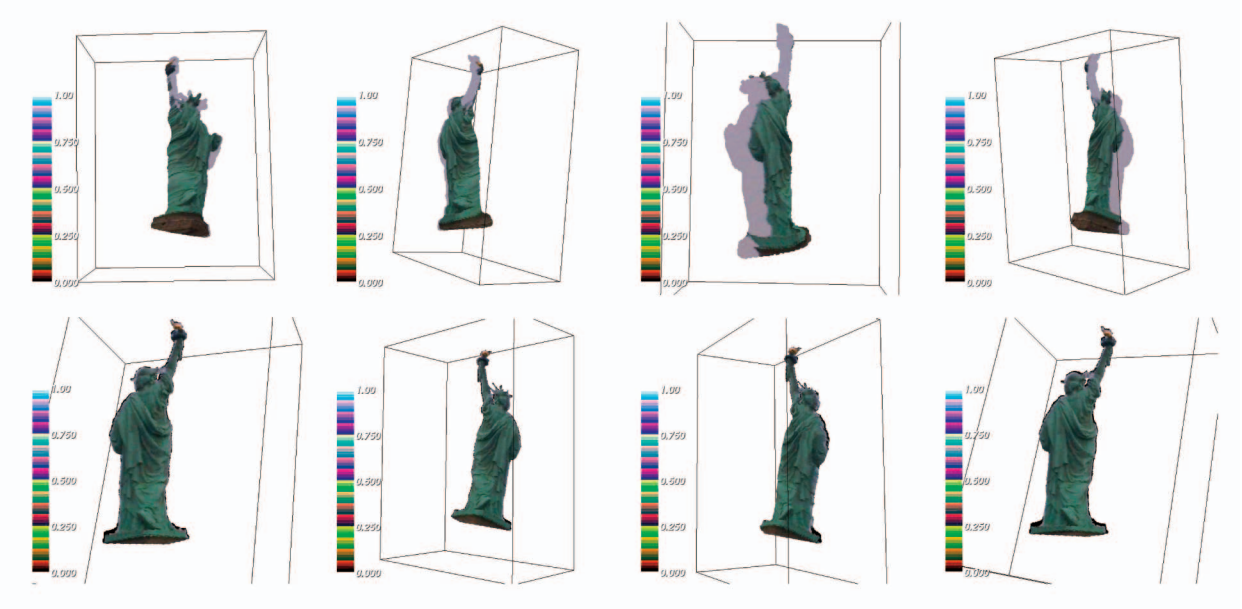

Fig. 19. Visual hull surfaces with initial rough calibration parameters (top), and with refined calibration parameters (bottom), also with radiance texture mapped onto the surfaces.

7.2 Conclusions

In this paper, we employed the 3D stereo techniques based on variational ideas to various camera calibration refinement problems. We have presented new multiview stereo techniques to

- evolve pose parameters of a 3D model object to take advantage of the known shape of calibration object, and to reduce computational complexity,
- evolve distortion parameters of cameras given a 3D model shape,
- evolve color calibration parameters of cameras given a 3D model shape,
- evolve intrinsic parameters of cameras, and
- evolve extrinsic parameters of cameras.

Pros and cons of this technique are discussed as follows:

- A nice feature of the methodology presented in this paper is that it can integrate several small and different problems such as distortion calibration and color calibration into an overall unified system based on the joint segmentation framework and simultaneously evolve pose, color, distortion, extrinsic, and other parameters as well.

Fig. 20. Some camera views before and after the color calibration for the Statue of Liberty.

- We make piecewise smooth object assumption and a constant background assumption, which may be a limitation if the background is to be modeled as well. However, a background model may be added to this framework if needed.
- The presented methods eliminate the need for search of image edges, point correspondences from images, which can be very sensitive to pixel-level noise, whereas our approach, being based on image regions for comparisons, is not as sensitive to noise.
- Another advantage of our framework is that it easily accommodates additional data. In the more classical approaches to stereo, bringing in more data or adding more images to the algorithm might not help all the time; that is, if something goes wrong in the independent segmentation phase of even one image, it destroys the whole process of reconstructions and geometry. On the other hand, adding more data to this joint segmentation framework will only improve robustness, providing more tolerance toward errors.
- For the distortion calibration method, more improvements may be obtained with utilizing more poses, hence, many more camera images of the calibration object and more than one distortion coefficient in the model selected. One can also utilize more general/complicated distortion models than the simple polynomial D function.
- Currently, we have an implicit representation of the calibration objects, that is, the cube or the rectangular bar. Computing surface normals, visibility functions for the surface occluding boundary from this implicit representation is not perfectly exact, and the quantities are slightly smeared. A future direction toward more efficient algorithms, is to use an explicit representation of the calibration object to more accurately describe the occluding boundaries.

With this approach, 3D grids are not needed for the data structures, resulting in increased accuracy and speed and decreased memory requirements.

- Camera calibration is particularly suited to our framework, since it does not have to be done in real time, and the environmental conditions may be allowed to vary to a degree (for example, our choice of a constant colored foreground object before a dark background).

ACKNOWLEDGMENTS

The authors acknowledge HP Labs, Palo Alto, California, for their support to G. Unal and A. Yezzi through grants to Georgia Tech for funding of this work. The authors would like to thank their colleagues at HP Labs: Bruce Culbertson, Harlyn Baker, Irwin Sobel, Tom Malzbender, and Donald Tanguay for fruitful discussions and their support. The authors also thank Hailin Jin for providing them with the Intel's Bust data set. A. Yezzi was partially supported by US National Science Foundation grants CCR-0133736 and IIS-0208197. S. Soatto was supported by US National Science Foundation grant IIS-02080197.

REFERENCES

[1] O. Faugeras and R. Keriven, "Variational Principles, Surface Evolution PDES, Level Set Methods and the Stereo Problem," technical report, INRIA, 1996.
[2] A. Yezzi and S. Soatto, "Stereoscopic Segmentation," *Int'l J. Computer Vision,* vol. 53, pp. 31-43, 2003.
[3] A. Yezzi and S. Soatto, "Structure from Motion for Scenes without Features," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 525-532, 2003.
[4] I. Sobel, "On Calibrating Computer Controlled Cameras for Perceiving 3D Scenes," *Artifical Intellligence,* vol. 5, pp. 185-198, 1974.
[5] R.Y. Tsai, "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses," *J. Robotics and Automation,* vol. 3, pp. 323-344, 1987.
[6] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision.* Cambridge Univ. Press, 2000.
[7] B. Triggs, "Camera Pose and Calibration from 4 or 5 Known 3D Points," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 278-284, 1999.
[8] P. Gurdjos and P. Sturm, "Methods and Geometry for Plane-Based Self-Calibration," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 491-496, 2003.
[9] O.D. Faugeras, Q.T. Luong, and S.J. Maybank, "Camera Self-Calibration: Theory and Experiments," *Proc. European Conf. Computer Vision,* pp. 321-334, 1992.
[10] Y. Seo, A. Heyden, and R. Cipolla, "A Linear Iterative Method for Auto-Calibration Using the DAC Equation," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 880-885, 2001.
[11] J. Oliensis, "Fast and Accurate Self Calibration," *Proc. Int'l Conf. Computer Vision,* pp. 745-752, 1999.
[12] M. Pollefeys, R. Koch, and L. Van Gool, "Self-Calibration and Metric Reconstruction in Spite of Varying and Unknown Intrinsic Camera Parameters," *Int'l J. Computer Vision,* vol. 32, no. 1, pp. 7-25, 1999.
[13] T. Svoboda, D. Martinec, and T. Pajdla, "A Convenient Multi-Camera Self-Calibration for Virtual Environments," *PRESENCE: Teleoperators and Virtual Environments,* vol. 14, no. 4, 2005.
[14] J. Heikkila and O. Silven, "A Four-Step Camera Calibration Procedure with Implicit Image Correction," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 1106-1112, 1997.
[15] G. Unal and A. Yezzi, "A Variational Approach to Problems in Calibration of Multiple Cameras," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 172-178, 2004.
[16] G. Unal, A. Yezzi, H. Baker, and B. Culbertson, "A Variational Approach to Calibration of Multiple Cameras," Technical Report HPL-TR-2004-219, HP Labs, Palo Alto, Calif., 2004.
[17] F. Devernay and O. Faugeras, "Straight Lines Have to Be Straight," *Machine Vision and Applications,* vol. 13, pp. 14-24, 2001.
[18] J. Weng, P. Cohen, and M. Herniou, "Camera Calibration with Distortion Models and Accuracy Evaluation," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 14, pp. 965-980, 1992.
[19] Z. Zhang, "A Flexible New Technique for Camera Calibration," Technical Report MSR-TR-98-71, Microsoft Research, 1998.
[20] D.C. Brown, "Decentering Distortion of Lenses," *Photogrammetric Eng.,* vol. 32, 1966.
[21] M.T. El-Melegy and A.A. Farag, "Nonmetric Lens Distortion Calibration: Closed-Form Solutions, Robust Estimation and Model Selection," *Proc. Int'l Conf. Computer Vision,* pp. 554-559, 2003.
[22] S.B. Kang, "Semiautomatic Methods for Recovering Radial Distortion Parameters from a Single Image," Technical Report CRL 97/3, CRL, 1997.
[23] R. Swaminathan and S.K. Nayar, "Nonmetric Calibration of Wide-Angle Lenses and Polycameras," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 2413-2418, 1999.
[24] G.P. Stein, "Lens Distortion Calibration Using Point Correspondences," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 602-608, 1997.
[25] A.W. Fitzgibbon, "Simultaneous Linear Estimation of Multiple View Geometry and Lens Distortion," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* 2001.
[26] F. Du and M. Brady, "Self-Calibration of the Intrinsic Parameters of Cameras for Active Vision Systems," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 477-482, 1993.
[27] H.S. Sawhney and R. Kumar, "True Multi-Image Alignment and Its Application to Mosaicing and Lens Distortion Correction," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 450-456, 1997.
[28] H.H. Baker, D. Tanguay, I. Sobel, D. Gelb, M.E. Goss, B.W. Culbertson, and T. Malzbender, "The Coliseum Immersive Teleconferencing System," Technical Report HPL-2002-351, HP Labs, 2002.
[29] E. Marszalec, "On-Line Color Camera Calibration," *Proc. Int'l Conf. Pattern Recognition,* pp. 232-237, 1994.
[30] S. Osher and J.A. Sethian, "Fronts Propagating with Curvature Dependent Speed: Algorithms Based on the Hamilton-Jacobi Formulation," *J. Computational Physics,* vol. 49, pp. 12-49, 1988.
[31] D. Adalsteinsson and J.A. Sethian, "A Fast Level Set Method for Propagating Interfaces," *J. Computational Physics,* vol. 118, pp. 269-277, 1995.
[32] J. Sokolowski and J.P. Zolesio, *Introduction to Shape Optimization, Shape Sensitivity Analysis.* Springer, 1992.
[33] R.M. Murray, Z. Li, and S. Sastry, *A Mathematical Introduction to Robotic Manipulation.* CRC Press, 1994.

Gozde Unal received the PhD degree in electrical engineering from North Carolina State University, Raleigh, in August 2002. She worked as a postdoctoral fellow at the Georgia Institute of Technology from the fall 2002 to the spring 2003 and visited HP Labs, Palo Alto, California, as a postdoctoral researcher during the summer of 2003. Since the fall of 2003, she has been a research scientist in the Intelligent Vision and Reasoning Department, Siemens Corporate Research, Princeton, New Jersey. Her research interests include variational techniques with connection to information theory and probability theory, applications to various computer vision, and image processing problems such as stereoscopic camera calibration, 2D/3D image segmentation and registration, filtering and enhancement, and stochastic particle systems. Her current research is focused on medical image analysis, segmentation, and registration techniques with applications to clinically relevant problems in MR, US, CT, and intravascular images. She is a member of the IEEE.

Anthony Yezzi received the PhD degree from the University of Minnesota, Minneapolis, in 1997. After completing a postdoctoral research position in LIDS, MIT, Cambridge, he began his faculty position at the Georgia Institute of Technology as an assistant professor in 1999. Currently, he is an associate professor in the School of Electrical Engineering, Georgia Tech. He has consulted for a number of medical imaging companies including GE, Picker, and VTI. His research lies primarily within the fields of image processing and computer vision. His interest within these fields includes anisotropic diffusion for image smoothing, active contours, segmentation, multiframe shape from shading, stereoscopic reconstruction, and shape analysis. His work in anisotropic smoothing and segmentation has been largely motivated and directed toward problems in medical imaging applied to MRI, ultrasound, CT, and OCT modalities. Two central theme of his research, in general, are curve/surface evolution theory from differential geometry and partial differential equations. He is a member of the IEEE.

Stefano Soatto received the PhD degree in control and dynamical systems from the California Institute of Technology in 1996 and the DIng degree (highest honors) from the University of Padova, Italy, in 1992. He joined the University of California, Los Angeles in 2000 after being an assistant and then associate professor of electrical and biomedical engineering at Washington University and research associate in applied sciences at Harvard University. His general research interests are in computer vision and nonlinear estimation and control theory. In particular, he is interested in ways for computers to use sensory information (for example, vision, sound, and touch) to interact with humans and the environment. He is the recipient of the David Marr Prize (with Y. Ma, J. Kosecka, and S. Sastry of the University of California, Berkeley) for work on Euclidean reconstruction and reprojection up to subgroups. He also received the Siemens Prize with the Outstanding Paper Award from the IEEE Computer Society for his work on optimal structure from motion (with R. Brockett of Harvard University). He is a member of the IEEE.

Greg Slabaugh received the BS degree in engineering physics from the University of Michigan and the MS and PhD degrees in electrical engineering from the Georgia Institute of Technology under the supervision of Dr. Ronald W. Schafer. After graduating from the University of Michigan, he developed 3D graphics for video games at a startup called Friendly Software. He is a member of technical staff at Siemens Corporate Research, Princeton, New Jersey His main research interests are computer vision, image processing, and computer graphics. More specifically, he is interested in medical image processing, 3D reconstruction from multiple images, new view synthesis, image and surface registration, tracking, segmentation, geometric modeling, adaptive filtering, PDEs, radial basis functions, deformable surfaces, and level set methods. He received a President's Fellowship and a Schlumberger Foundation Fellowship, as well as Outstanding Research and Outstanding Service Awards from the Center for Signal and Image Processing in the School of Electrical and Computer Engineering, while he was a graduate student. He is a member of the IEEE.