Imaging Based on Optic-fiber Bundles: Topological and Geometric Calibration

João Silva, Pedro Aguiar, José Gaspar
Instituto Superior Técnico / UTL, Lisbon, Portugal
joao.soares.da.silva@tecnico.ulisboa.pt, {aguiar, jag}@isr.tecnico.ulisboa.pt

ABSTRACT

Fiber Bundle cameras are imaging devices where the image is captured in one end of the fiber cable by a lens and transmitted through the cable to the other end. Commercial fiberscope cameras have their fibers arranged inside the cable to form a regular pixel grid. Breaking the fibers cable precludes the use of this camera, since due to the small diameter of each fiber and the fact that we do not know the correspondence between fibers, fixing the cable is an almost impossible task to do.

Recent research found that broken fiber cables, under certain conditions, can still be recovered. For that, one has to estimate the disorganized fiber tips positions by modeling them as a discrete camera and estimating its topology.

The lens used in front of the fibers makes rays converge. The pin-hole model is the simplest way to describe rays converging to a single point, and in this dissertation, we study if that model suits our camera. We focus on the structured light and sawtooth light patterns calibration methods, to estimate fibers topology and finding the camera parameters. Experiments are conducted to estimate topology and calibrate the camera. In the end is presented the result of topology reconstruction of an image mosaic composed by a phrase.

I. INTRODUCTION

A camera geometric imaging model describes the optical process that guide the photons into the sensitive surface of a capturing device. These models are useful to perform a camera calibration, i.e. build a map between the 3D rays in space and the pixels on the camera.

Discrete cameras are composed by a collection of pixels organized with unknown topology, different from conventional cameras which have their pixels collections organized in a rectangular grid. Neighbors cells on discrete cameras can even capture light from completely unrelated regions of the scene. The camera based on an optic fibers bundle used in this work is an example of a discrete camera. The optic fibers bundle has specific characteristics that make it different from fiberscopes. It has the fibers twisted. Therefore, the image obtained from this camera is unreadable by the human eye.

Geometric calibration involves estimating the intrinsic parameters of the camera, i.e. describe the geometric projection properties of the camera, and estimating the extrinsic parameters which allow to transform the world coordinates into camera coordinates.

In this dissertation we study an optic-fiber bundle camera, i.e. its parameters, the calibration process and even the topology of this camera in order to get a readable image with the best quality possible motivated this dissertation.

A. Related Work

Camera calibration is a subject largely covered in research and development. However new cameras with new technologies and specificities are increasingly challenging for new calibration methods or even improvements in the well-known ones.

In 1982 Posdamer and Altschuler [7] proposed the first calibration using a sequence of projected patterns based on binary encoding. A few years later, in 1995, Bergmann [3] proposed a methodology based in gray level patterns combined with phase shifting which was a great improvement since it was possible to take advantage from the best qualities of each method. In 2001, Gühring [6] improved the gray codes with phase shifting method by creating a method called line shifting which presents good results for high resolution cameras and accurate measurements. In 2004, Salvi et al. [8] grouped the methods cited before and others. The authors organized the several methods into specific types. A year after, Grossberg and Nayar [5] presented a new camera model and to validate their model they performed calibration using binary patterns.

In 2013, Bergamasco et al. [2] proposed a fully unconstrained camera model and proposed a calibration method of all individual raxels.

The main goal of geometric calibration is to estimate the camera intrinsic and extrinsic parameters. One well known methodology to estimate the camera matrix is the Direct Linear Transformation which was presented by Abdel-Aziz and Karara [1] in 1971. Tsai [11], in 1984, proposed a method where the author estimated the camera parameters. In 1999, Zhang [12] also proposed a method to estimate the intrinsic parameters by calibrating the camera with patterns at different positions and orientations.

In 2012, Fernández et al. [4] proposed different methodologies based on binary coded patterns to calibrate an optic fiber bundle camera. These methods performed well for the authors case. However, they have a fiber bundle with around 50,000 fibers, which is different from our case where our bundle has around 4000 fibers. We also explore binary coded patterns for calibrating our camera.
A. Perspective Camera vs Raxels Camera

The pin-hole model represents perspective projection. This model assumes that a lens is reduced to a small hole, the optical center, and all rays will pass through it. There is a line passing through the optical center, designated as optical axis of the camera. The intersection of that line with the image plane defines the principal point.

The camera matrix is the set of parameters that allows the transformation from the world frame to the image frame. It is divided into two types of parameters, the extrinsic and intrinsic parameters.

Extrinsic parameters describe the camera location in the world frame and which direction it is pointing at. Extrinsic parameters are represented in the form of a rigid body transformation matrix (rotation plus a translation), where the rotation is represented by a $3 \times 3$ matrix on the left side and the translation by a $3 \times 1$ column vector on the right side.

The intrinsic parameters describe the geometry of a camera. The intrinsic parameters are represented by a $3 \times 3$ matrix that allows one to transform meter coordinates into pixel coordinates. Since images are formed by pixels, this transformation is important to locate a point on the sensor surface.

The camera matrix is then of the form:

$$P = K[R|t]$$

There are cameras that are non-perspective and, as the name suggests, cannot be described by perspective projection models. These cameras do not have a single viewpoint for rays to pass through. For that, Grossberg and Nayar [5] suggested a general imaging model where the camera is defined as a set of raxels. A raxel is, in simple terms, the ray that passes through a pixel. The pixel acquires the intensity of light of that ray. One raxel can be defined as the direction of the chief ray associated to a light sensor and is characterized by a 3D position and a direction vector. The raxel model allows the geometric description of most of the cameras.

B. Perspective Camera Calibration

Camera calibration involves estimating camera parameters from images. Since seminal works, e.g. Abdel-Aziz and Karara [1] or Tsai [11], camera calibration has been done based on a 3D calibration object or a 2D pattern, both typically involving checkerboards with known dimensions. In this section we refer only the case of a 3D calibration object.

From one image of the 3D calibration object one obtains key points in image coordinates, $\{m\}$, and also the respective 3D world coordinates, $\{M\}$. The 2D points, $\{m\}$ and its 3D correspondences, $\{M\}$ can be organized as linear constraints, forming a so-called Direct Linear Transformation (DLT). The DLT allows then estimating the camera (projection) matrix $P$. In the following we present two DLT methods for estimating $P$.

The camera matrix $P$, $3 \times 4$, can be defined as an homography and one can transform it into $P = [p_1 \ p_2 \ p_3]^T$.

Having $m = PM$, where $m = (u, v)$ then

$$\begin{align*}
\lambda u &= p_1 M \\
\lambda v &= p_2 M \\
\lambda &= p_3 M
\end{align*}$$

(2)

Finally, using matrix notation for 2 one obtains:

$$\begin{bmatrix}
M^T & -uM^T \\
0 & -vM^T
\end{bmatrix}\begin{bmatrix}
p_1^T \\
p_2^T \\
p_3^T
\end{bmatrix} = \begin{bmatrix}
0 \\
0
\end{bmatrix}$$

(3)
Each pair \((m,M)\) generates two equations, since \(p = [p_1\ p_2\ p_3]^T\) is a column vector with size \(12 \times 1\), then one will have 12 unknowns. Therefore, to have a good estimation of \(P\), one should at least have 6 pairs of points in the dataset.

The linear problem in equation 3 has the form \(Ap = 0\), where \(A\) has the size \(2N \times 12\) and \(p\) is the column vector already defined. Using Singular Value Decomposition (SVD) it is possible to obtain \(p\), since it will correspond to the smallest singular value of \(A\) and will give a vectorized form of \(p\). By rescaling \(p\), it is possible to retrieve the estimated matrix \(P\) up to a scale factor.

A second method can be used to derive the calibration equations. One starts by applying a cross product to both sides of \(My = PM\), obtaining 0 in the left hand side of the equation. One has then \([mi]_xPM = 0\), where \([mi]_x\) represents a linear cross product operation as a skew-symmetric matrix of \(mi\).

The Kronecker product allows us to obtain an equation factorizing the data and variables to estimate:

\[
(M^T \otimes [mi]_x) vec(P) = 0
\]

where \(vec(P)\) is the vectorization of the matrix \(P\), stacked into a singular column vector. From each pair of points 3D-to-2D \((Mi, mi)\) is possible to obtain 3 equations in the entries of \(vec(P)\), however only two are linearly independent.

As in the previous method, to estimate \(P\), one has to have at least 6 pairs of corresponding points. From 4 one has \(Ap = 0\), where \(A = M^T \otimes [mi]_x\) and \(p = vec(P)\). This matrix \(A\) would be a stack of \(M^T \otimes [mi]_x\) having the size \(3N \times 12\). Using Singular Value Decomposition (SVD) it is possible to obtain \(p\), since it will correspond to the smallest singular value of \(A\) and will give a vectorized form of \(p\). Once again, similarly to the prior method it is possible to retrieve the estimated matrix \(P\) up to a scale factor by rescaling \(p\).

Both the calibration methodologies, as presented till here, lack the inclusion of radial distortion.

In contrast to DLT points methodology, Silva et al. [10] considers 3D lines correspondences with 2D lines in the image. It is also equipped to perform the estimation of the camera matrix \(P\) regarding radial distortion. DLT lines have the advantage of through image processing one can improve the datasets used in calibration, once it allows to fine tune the location of the lines in the image.

Trying to create a standard procedure to calibrate all types of cameras is not a new problem and there is a lot of work in that sense. Based in [5], Bergamasco et al. [2] created a fully non-conventional camera model and a novel calibration method. This calibration method deals in an efficient way with the huge amount of degrees of freedom the process. This calibration method is explored and in part applied in this dissertation.

C. Non-conventional Sensor Topology

In order to calibrate a camera there are several different methods with different types of auxiliary patterns. A representative part of them are presented and explained in [8]. Binary codes were one of the first techniques used to calibrate cameras and still are a very commune method. It was proposed by Posdamer and Altschuler [7] and consists in \(m\) patterns to encode \(2^m\) black and white stripes.

Calibration methods based on gray codes and phase shifting are an improvement relative to binary codes and also to phase shifting methods, once they have the properties of stripe patterns and also the high resolution exploited by phase shifting methods. Bergmann [3] proposed a method where he combined gray codes with phase shifting patterns to measure and calibrate surface regions. He coded sinusoidal patterns with gray levels which are then shifted along time. With the intensity information of each pixel (pixel streams), it is possible to compare phase shift for each pixel and compute the corresponding surface point in the image.

Grossberg and Nayar [5] also proposed a ray-based calibration method based in [9] binary code methodology. Camera calibration based on encoding pixel streams has a drawback since the number of calibration images is limited and it is difficult to define the pixel sizes. With the lack of knowledge about the areas of the LCD screen the photocells are capturing light, this methodology may not give the accurate description of the topology as expected.

III. Imaging Topology

In this chapter, we introduce the camera used in our work and how to estimate its topology. We start by describing the camera setup, followed by a brief discussion about camera sensors and pixel shapes, where we present our case.

A. Fiber Bundle Camera Setup

The camera setup of this work is composed by a conventional camera, figure 1(A), one cable of optic fibers mounted in front of the camera, figure 1(B), and one extra lens mounted in front of the fibers, figure 1(C). The fibers in the cable are rigidly glued to each other and randomly blended and for that reason any image appearing on the top end of the cable, figure 1(D), will be different from the bottom end, figure 1(E). The image observed at A will be a disorganized set of pixels of the image captured at C, meaning that the image coordinates of a fiber at \((u_e, v_e)\) are different from the coordinates at D \((u_d, v_d)\), i.e. \((u_e, v_e) \neq (u_d, v_d)\).

B. Conventional vs Non-conventional Sensor Topologies

There are two different types of sensors used in conventional cameras: CCD and CMOS sensors. These camera sensors capture the light that passes through the lens and then form the image. Each sensor is made of smaller photosensitive cells, i.e. pixels. In conventional cameras, the sensors usually have rectangular shapes.

Non-conventional sensors, including the case of our camera, are sensors where the topology is not trivial. In our case the sensor encompasses a fiber bundle, where each fiber works as a pixel. The tip of a fiber is circular, so this makes our pixels with circular shape, although they may be interpreted as squared shape pixels.

In the next sections we present the sensor topology representation for our case and essentially how to estimate it. We also present the implemented methodologies for finding the topology.
1) Sensor Topology Representation: Before presenting the calibration methodology, we start by defining a model for the topology of the sensor. Let

\[ f : D \in \mathbb{R}^2 \rightarrow E \in \mathbb{R}^2 \tag{5} \]

denote a sensor topology mapping, where \( D = \{(u_d, v_d)_i, i \in 1, \ldots, N\} \) and \( E = \{(u_e, v_e)_i, i \in 1, \ldots, N\} \) and \( N \) represents the number of fibers. Writing the elements of \( D \) and \( E \) in vector notation, i.e. \([u_d, v_d]^T_i\) and \([u_e, v_e]^T_i\), the \( f \) mapping of one specific element can be represented as

\[
[u_e]_i = f\left([u_d]_i, [v_d]_i\right) .
\]

We consider \( f \) to be constant. In particular we consider that the function \( f \) can be detailed as a sum of terms

\[
f\left([u_d]_i, [v_d]_i\right) = [u_d]_i + [\Delta u]_i,
\]

where \([\Delta u]_i\) denotes a constant offset for each of the image coordinates.

In the case of conventional CCD or CMOS sensors one has simply that \([\Delta u]_i, [\Delta v]_i^T = [0, 0]^T \). As introduced earlier, in the case of optic fibers \([\Delta u]_i, [\Delta v]_i^T \) has in general nonzero, unknown, real values.

2) Finding the Elements of the Sensor: In our camera the topology, i.e. the location of pixel neighbors, is unknown. Therefore, we have to estimate the sensor topology. The first step that has to be taken is to detect the elements of our sensor.

a) Fibers Detection: Since the distribution of the fibers inside the camera is irregular, we decided to define 5 \times 5 squared cells which represent the size of each fiber. Each side of a cell corresponds to the average diameter of a fiber and due to the irregular distribution correspondences between cells and fibers cannot be made, since each cell can cover more than one fiber at the same time.

To find the fiber’s coordinates one can compute the mean image over a set of frames and then apply a filter to the mean image, using non-maximum suppression operation similar to the one used on Canny edge detection. In order to assure that all fibers have received signal, the mean image is computed over a set of images in a changing scenario. Therefore the mean image will be given by:

\[
I_m = \frac{1}{N} \sum_{i=1}^{N} I_i .
\]

A 2-D order-statistic filter is applied to the mean image, i.e. a filter with approximate size of a fiber (5 \times 5 filter) is used to filter the mean image and for each position of it the maximum inside the filter domain is taken. This procedure creates a filtered image, \( I_{filt} \). Knowing that each fiber is a local maximum and having \( I_m \) and \( I_{filt} \) one is now able to find the pixels by making a logical comparison between the two images assuring that the maximum in the mean image is above a certain threshold \( T \).

As shown before in figure 1, it is possible to observe the mean image and the computed fiber’s positions showed by the blue dots.

C. Topology Find based on Structured Light

Structured light based topology find has several different methodologies already known, such the ones presented at Salvi et al. [8]. These methodologies consist in showing structured light patterns to the camera that we want to calibrate.

We use in a first step binary codes patterns and then we go to triangular waves.

Patterns are shown to the camera as illustrated in figure 4.
D. Topology Find based on Triangular Wave Light Patterns

To find the camera topology one has first to detect the fibers. The setup of Figure 4 is used for that purpose. It consists on a fiber camera and an LCD display, where the patterns for calibration are shown.

The calibration process can be divided in two main steps. The first is finding the fiber coordinates from an image captured at point D of the setup. The second step consists in determine the relative position of each fiber using an auxiliary pattern in the calibration process in order to map the fibers at the output image at D to the input image at E.

To find the topology, i.e. the fibers relative position we use triangular waves. More in detail, the pattern is a calibration gradient, in this case represented with a 2-dimensional triangular wave. Figure 3 shows the calibration images, one horizontal and another vertical, to find out the relative horizontal and vertical positions of the fibers. These images will be circularly shifted each time step.

![Fig. 3. Horizontal and vertical Calibration Patterns](image)

It is possible to obtain the relationship between two pixel streams or even between one pixel stream and the image along one axis in each time step by comparing pixel streams intensity values over time. Recurring to Fourier methods, one obtains relations between the positions of the peaks of the triangular wave, i.e. using pixel phase difference between pixel streams as measure, one obtains the relative fibers positions. For a given fiber, the captured pixel stream should correspond to a triangular wave.

Since the signal is circularly shifted, it means that choosing two different pixels there will be time delay between them. To obtain the time delay between pixel streams in the frequency domain the Discrete Fourier Transform is used to get the phase difference, using the circular time shift property:

\[
\text{DFT}(s(n)) \quad S(k) = S(n) e^{-j \frac{2\pi}{N} n_0 k}
\]

From 9, one can observe that for different pixel streams, e.g. one at position \( u_0 \) and other at position \( u_1 \), have the same amplitude spectrum, but different phases and so observing the pixel streams along time will ideally represent two triangular waves with a shift between them. Since this form of waves only has odd harmonics and the first harmonic has the biggest contribution to the signal, the first harmonic is chosen to measure the phase difference between pixel streams. Using FFT is obtained this phase difference and from these values, the relative fibers positions are computed.

IV. PROJECTION GEOMETRY

The pinhole projection model is the most common way to formalize the geometric component of the image formation process. The extrinsic and intrinsic parameters, which together form the camera projection matrix, are the main goal of the geometric calibration.

In this chapter, we detail the camera geometric calibration process, which builds on top of the acquired camera topology detailed in the previous chapter. One principal objective is proving that our camera based on an optic fibers bundle, a lens and a conventional camera, can be modeled by the pinhole camera model. Here we introduce two different methods for the geometric calibration. After testing positively for imaging centrality, we aim to fit a pin-hole model and in particular estimating the intrinsic parameters.

A. Calibration Setup

The setup to perform the geometry calibration is similar to the one used to estimate the camera topology. Figure 4 illustrates a central camera imaging a LCD screen, as a pyramid composed by optical rays starting at the screen and intersecting at the optical center of the camera.

In our real setting, the camera is placed close enough to the LCD so that the screen at the furthest distance occupies all the camera image. Consequently, the camera at closest distance images just a cropped screen.

Let us consider the camera is fixed and the LCD displaying the calibration patterns is placed at two distinct depths, e.g. \( z_1 = 10cm \) and \( z_2 = 20cm \), as shown in figure 4. Let \( M_1 \) and \( M_2 \) be 3D points defined on the LCD screen, but with respect to a single world coordinate system. More precisely \( M_1 \) closest, \( M_2 \) furthest from the camera. We label as zero the z-coordinate of \( M_1 \), hence the z-coordinate of \( M_2 \) is the value incremented by the motion of the LCD, i.e. 10cm in the numerical example of the figure. These points define the relative position to the camera for each pixel, in metric coordinates.

Having two sets of \( N \) 3D points each, one can estimate \( N \) rays that passes through both points.

Two different methods are used to calibrate the camera, one based on imaging binary patterns and the other based on imaging a triangular wave pattern, as shown in chapter III-C. Having performed the calibration, with any of the methods, one has sets of 2D points and its 3D correspondences, i.e. pairs of points \((m_1, M_1)\) and \((m_2, M_2)\) for each calibration methodology.
V. DIRECTIONS AND INCIDENCE OF OPTICAL-RAYS

In this section we use optical rays (raxels) incident to the camera and passing through known 3D points to calibrate the camera. Points on the LCD plane are named in the following as target points.

The goal of the next paragraphs is to find the camera pose, rotation \( R \), and translation \( t_c \), with respect to the world coordinate system \( \{ s \} \) as defined by the LCD location. Using this variable naming, an LCD point \([ u, v ]^T \) can be represented as \([ u, v ]^T = R^t \{ X \ Y \ Z \}^T - R^t \{ t_x \ t_y \ t_z \} \) where \([ X \ Y \ Z ]^T \) represents the point in the camera coordinate frame.

A. Calibration based on Multiple Camera Poses

In this section is briefly presented the calibration of all rays methodology proposed in [2]. Multiple (random) camera poses are considered, with the single constraint that the camera at different points where rays intersect. The two stereo cameras are considered, with the single constraint that the camera at two different locations, being each location associated with different points where rays intersect. The two stereo cameras can be solved with projective geometry.

Given \( N \) rays, one for each of the optic fibers, one may have central or non-central cameras. Contrarily to the calibration methodology proposed in [2] and briefly described in section V-A, we do not need to initialize our algorithm with a pin-hole calibration toolbox. Effectively, we compute all raxels individually, then test them for centrality (see next section) and in case of centrality we finally fit a pin-hole projection model (see later section).

C. Test of Centrality

Centrality is a property of perspective cameras. It refers to the existence of an unique optical center in the camera, which is what happens in perspective cameras.

Non central cameras do not have this property, they have different points where rays intercept. The two stereo cameras separated by a baseline is an example where non centrality can be observed.

The fiber camera used has to be tested for centrality to verify if it has an unique center of projection. This is a problem that can be solved with projective geometry.

a) Pencil of lines: Pencil of lines is the name given to a set of lines passing through the same point. Finding that point represents the problem of how to get the camera center coordinates from a set of rays. To obtain an estimation of this point, it can be used information about the rays, such as its parametric equation

\[ p_i + \alpha_i d_i = C \]

(13)

where \( C \) is the center of the camera, \( d_i, p_i \) the direction and position of the ray and \( \alpha_i \) the scale factor. Considering \( N \) optic rays and defining a vector of unknowns

\[ x = \begin{bmatrix} C & \alpha_1 & \ldots & \alpha_N \end{bmatrix}^T \]

(14)
allows defining a system of linear equations $Ax = b$, where $A$ and $b$ are:

$$A = \begin{bmatrix} -I_{31} & d_1 & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ -I_{3N} & 0 & 0 & d_N \end{bmatrix}$$

$$b = [-p_1 \ldots -p_N]^T$$  

To compute the solution of this linear equation system, one uses the Moore-Penrose pseudoinverse of $A$, which allow us to get $x = A^+b$. With $x$ the center of projection $C$ can be retrieved from the first three lines. The remaining lines of $x$ are the scale factor for each ray. The residuals of the this linear system are observed by solving $\epsilon = A \times x - b$. If the residuals are close to zero, below a certain threshold, it means that the camera is central, otherwise is a stereo camera.

b) RANSAC: To account for outliers in the datasets tested and remove them a RANSAC algorithm is used. RANSAC is an iterative algorithm that estimates parameters of a mathematical model from datasets constituted by observed data which contains outliers.

The algorithm can be divided in two main steps, the first where a small sample of the dataset is randomly acquired, a fitting model and the model parameters are computed for that data sample. The second step is where all the data is checked by the algorithm, in order to verify which of that data is consistent with the model computed in the first step. To fit the model the data can not have a deviation superior to an error threshold that defines the maximum deviation allowed from the fitting model.

D. Fit the Pin-hole Projection Model

The centrality test with a positive result is an important result. We can use methods for central cameras such as the DLT points to estimate camera matrix $P$.

Proven centrality for our camera, we are now capable describe our camera using pinhole camera model, estimating a projection matrix. For that we will use the DLT algorithm. As already mentioned in II-B, it is an algorithm that estimates the camera matrix $P$. This algorithm computes $P$ by having 2 sets of 2D and 3D points, which are correspondent to each other, i.e. 2D points obtained, with camera at a distance $z_1$ from the LCD display, and its 3D correspondences and the same for a distance $z_2$.

Once proven centrality of our camera, we move on to estimate the camera matrix $P$, which is estimated up to scale value. From $P$ is then possible to estimate extrinsic and intrinsic parameters. It is possible to estimate $K$ obtaining the camera geometry.

VI. EXPERIMENTS

This chapter describes the experiments performed to validate the methodologies presented and introduced in previous chapters, and the test to reconstruct the image topology.

Two different cameras are considered to estimate camera topology and geometry. For each camera is acquired a dataset with 2D and 3D points displayed by one LCD.

The first camera is a Logitech Quickcam Orbit AF which approximates well a pinhole camera and its purpose is to act as the control group. As a pinhole camera it has a single projection center and regular topology (CCD/CMOS sensor). The other camera is the Optical Fiber Bundle Camera for which we want to assess centrality. The bundle has approximately 4500 fibers and each fiber is imaged by approximately $5 \times 5$ pixels.

A. Imaging Sensor Topology

Two calibration processes are performed for both cameras to acquire camera topology. The first calibration method is based on binary coded patterns while the second is based on a triangular wave pattern.

1) Acquisition of Calibration Data based on Binary Patterns: The first calibration method performed is based on binary patterns placed at two different distances, one at $z_1 = 10cm$ and the second at $z_2 = 20cm$.

To check until which pattern the LCD and the camera have resolution to perceive the transition between stripes an intermediate test is performed.

This intermediate test allows one to conclude that only the first 6 bits can be used to encode the pixels positions, due to the camera resolution. The 6 bits displayed as horizontal and vertical stripes imply 12 LCD images to calibrate the camera pixel positions.

The calibration is successful and after remapping the grid it is possible to get an image perceptible to the human eye.

It is important also to understand the number of bits that is necessary to encode the fiber bundle. Since the bundle used has approximately $64 \times 64$ fibers, 6 bits are enough to encode each fiber horizontally and vertically and therefore 12 encoded images to display with the LCD.

The results are not precise enough, since several of the pixels are remapped with same coordinates due to errors.
reading the encoding shown by the LCD. Only 700 fibers get different locations. This happens due to blurry pixels which makes difficult to perceive transition between stripes. This and all other drawbacks identified in III-C implied the study of triangular waves used as calibration patterns as detailed in the next section.

2) Acquisition of Calibration Data based on Triangular Waves: The second calibration method is based on using a triangular wave pattern as described in section III-D. The calibration is performed at the same two different distances for the fiber camera. For the USB camera, new distances $z_1$ and $z_2$ had to be defined, because some pixels do not detect the triangular wave and therefore the calibration process tag those fibers as not receiving signal.

To apply topological calibration to the USB camera based on triangular wave patterns, we choose $z_1 = 30cm$ and $z_2 = 40cm$, keeping the $10cm$ difference between LCD poses.

For the USB camera, topology was estimated correctly, and one can correctly reconstruct the calibration images.

For fiber bundle camera it is a different case, once even with pixels in rearranged order with a correct topology, the low contrast of the triangle wave patterns makes hard to read the images.

For that reason, the results were not the ideal but it showed promise and with some image processing, it should be possible to recover the image captured by the camera.

B. Geometry Calibration

Given the topology, we now want to fit the geometric model to the camera. For that we perform a test of projection centrality for both cameras. A positive centrality test allows fitting a geometric projection model where one estimates the camera center and other camera parameters, based on a DLT calibration methodology.

1) Test of Projection Centrality: Having both cameras topologies allows the acquisition of two pairs of 2D and 3D points corresponding to LCD places at distances $z_1$ and $z_2$. The datasets are characterized as, $\{(m_1, M_1, m_2, M_2)\}$, where $m_1$ and $m_2$ represent the points in pixel coordinates and $M_1$ and $M_2$ represent the same points in world frame coordinates. These datasets contain outliers that interfere with the calculus of the center of the camera using pencil of lines.

a) Webcam: The results of the first test for the USB camera can be seen in figure 6 (a). This figure shows that the points obtained in the datasets, plotted as rays, intersect in one point which corresponds to the center of the camera, representing then a pinhole camera model, as was expected.

As one can observe, the distances are correct (10cm between $M_1$ and $M_2$) and the rays converge to the center, proving that this camera is central.

b) Optic Fibers Bundle Camera: The results of the first test for the Fibers camera can be seen in figure 6 (b). It shows that the rays converge and intersect in one point, the center of the camera, as in the USB camera.

This time the camera distance to the points plane is a bit far than expected. This may happen due to the distance between the tip of the lens and the beginning of the fibers cable, where our fibers sensor is present, which may conduct to measurement errors.

Although these differences, the rays still converge to the estimated center. It is possible to conclude that it is a central camera and the pin-hole camera model can be used with the fiber camera.

2) Geometric Model Fitting: After testing for centrality, now the main goal is to estimate a geometric model and if it is consistent with the estimated center.

The same datasets are used since DLT only needs 2D and its correspondent 3D points to estimate camera matrix $P$. From $P$, it is then possible to estimate intrinsic and extrinsic parameters, obtaining this way the matrix $K$.

a) Webcam: The dataset used is from calibration based on triangular waves patterns. This is the case where the calibration is performed at different distances of the LCD ($z_1 = 30cm$). The results show the USB camera modeled as a pinhole, as shown in figure 6 (c).

It is possible to observe that the camera axes are defined as expected, in particular the z-axis is orthogonal to the plane that contains the 3D points. Matrix $K$ and vector $t$ obtained are:

$$K = \begin{bmatrix} 636.97 & 18.73 & 48.10 \\ 0 & 643.34 & 40.45 \\ 0 & 0 & 1 \end{bmatrix}, \quad t = \begin{bmatrix} -2.28 \\ -1.90 \\ 30.35 \end{bmatrix}. \quad (17)$$

For this case the image principal point is at (48.10, 40.45), which corresponds the top corner of the image. This center is different from what we expected. One hypothesis to explain this, is the fact of a increased value of skew factor comparing to the previous case and the image can be suffering from some shearing effect towards the top corner. Also, the different distance from the LCD can be affecting the results of calibration performed before.

The camera center is consistent with the differences stated before. The distance from the camera to the LCD display is 30.35cm which is close to the expected 30cm.

b) Optic Fibers Bundle Camera: Finally, the camera matrix $P$ is estimated for the case of the dataset obtained from triangular waves calibration with the fibers camera. The results once again show that the camera calibration performs well when using the pinhole camera model. Figure 6 (d) shows the camera draw from the estimated matrix $P$. In this figure is also possible to corroborate the tilting of the camera observed when it was being calibrated. As can be seen the camera is tilted down.

The intrinsic parameters of the fibers camera and the translation are:

$$K = \begin{bmatrix} 448.40 & 22.94 & 173.26 \\ 0 & 426.10 & 216.27 \\ 0 & 0 & 1 \end{bmatrix}, \quad t = \begin{bmatrix} -9.79 \\ -13.51 \\ 14.80 \end{bmatrix}. \quad (18)$$

In figure 6 (d) we notice that the fibers sensor is capturing the image a bit down and left. This was also detected when performing calibration.
The camera is at 14.80 cm from the LCD display. Once again, like the case for USB camera, the difference from the expected value may be the result of non-zero residuals which introduce some errors between the rays and the camera center.

C. Mosaic Acquisition Test

To test if the topology acquired is correct an experiment was conducted. It consists in passing an image in front of the camera and try to reconstruct a mosaic of the complete image that is passed on LCD.

We used an image with a phrase taken from the dissertation title, as shown in figure 7 (a).

To perform this test, we have to have the camera pointing to the LCD at the same distance of the previous tests, for instance 20 cm. On the LCD, the image will be passed from right to the left simulating panning movement of the camera. While the image is passing the camera is taking several pictures to the LCD image, recording each frame into a video. This video contains the output of the camera at the end of the fiber bundle. It is unreadable by the human eye.

We use the estimated ground-truth fibers coordinates to compute for each frame the real position of each pixel. From this, we are now capable of determining if the topology acquired is correct. Figure 7 (b)-(c), shows the correspondences between what is seen in point E and D by the fibers.

From all the frames captured, we were able to recreate the mosaic phrase. For that, we manually selected key points from different frames and we made the correspondence to their representation in the original mosaic image.

Using a Procrustes algorithm, we were able to estimate the correspondences between the points in each frame and the original mosaic. Procrustes analysis applied to image processing allows us to find the correspondence between a set of points regarding their translation and rotation. Therefore, for each frame every pixel, i.e. every fiber, will correspond to a point in the original mosaic. Regardless some noise, we will be able to get a match and from that reconstruct the original mosaic.

We were then ready to compute a point cloud with all the points estimated and with correspondence in the mosaic. The point cloud that resulted from this process is shown in figure 7 (d).

The readability of the message confirms the applicability of the proposed calibration methodologies.

VII. CONCLUSION AND FUTURE WORK

In this work we propose methodologies for the topologic and geometric calibration of an Optical Fiber Bundle Camera. In the aspect of the topologic calibration we started by stating which methods were already available in related work and we opted to use two different calibration methods to compare the results. First, we used binary codes to estimate the camera topology but we concluded that for our camera binary codes provide modest calibration. Then we used a time multiplexing hybrid method. The method consisted in a calibration gradient pattern, namely a triangular wave repeated along lines or columns.

After having estimated the topology, we continued with geometric calibration, in particular assessing whether the pinhole camera model could be applied to our camera. For that we performed a centrality test to understand if the rays captured intersected in one point only or if we were in the presence of a non-central camera and therefore not able to use the pinhole model. We concluded that our camera, can be represented by the pinhole model and consequently we were able to estimate intrinsic parameters that represent well the camera geometry.

To test the topologic and geometric calibrations, we used the camera for imaging text, namely one phrase. The main goal was to reconstruct the phrase on a mosaic. We were able to reconstruct the mosaic image, having an image and obtained a precise enough result to be readable by a human.

In future work the aim is to automate the calibration process as well as the image reconstruction process. Also, it will be a good step forward to improve the calibration process itself, since there are still errors and fibers with wrong correspondences.

REFERENCES

[1] Y.I. Abdel-Aziz and H.M. Karara. Direct linear transformation from comparator coordinates into object space co-


