Journal of Intelligent and Robotic Systems **36:** 1–21, 2003.
© 2003 *Kluwer Academic Publishers. Printed in the Netherlands.*

Design and Calibration of an Opto-mechanical Appliance for 3D Non-contact Orthopedic Measurements

Part I: Mathematical Model and Laboratory Prototype

ROBERTO CARACCIOLO* and ALBERTO TREVISANI
Department of Management and Engineering (DTG), Università di Padova, 36100 Vicenza, Italy; e-mail: caracciolo@gest.unipd.it

(Received: 12 March 2001; in final form: 22 February 2002)

Abstract. An opto-mechanical system performing fast and accurate non-contact foot measurements is proposed. By processing the measured data it is possible to get three-dimensional foot models and to design custom-made shoes. This work is subdivided into two parts: in Part I, after introducing the system operating method and describing the vision technology employed, a mathematical model of the system is developed accounting for the main internal parameters of the camera. In order to reduce scanning time and the time necessary to create the 3D foot model, the measured data have been arranged in a particular structure and a suitable acquisition procedure has been developed. A laboratory prototype has been built to test the performances of the system: a detailed analysis of its geometry and of its components is provided as well as a concise description of the simple user interface developed to run and control the system.

Key words: 3D scanning, non-contact orthopedic measurements, structured light vision, image acquisition procedure.

1. Introduction

The possibility of reproducing the three-dimensional (3D) shape of an object has always been considered of the utmost importance in such fields as rapid prototyping, reverse engineering, computer animation, and artwork preservation and replication. Several technologies have been developed, tested and employed to manufacture 3D scanners, ranging from completely mechanical solutions (e.g., touching machines) to complex stereo-vision optical systems.

An interesting example of 3D scanner is the system proposed by Gonzáles et al. (1993) for digitizing wooden shoelasts in footwear industry. The machine developed consists of two CCD cameras which acquire the images of the laser lines projected onto a moving shoelast by a laser source. This system represents a typical application of the approach based on structured lightning to computer vision. A more ambitious and heterogeneous system for digitizing the shape and

* Corresponding author.

color of large fragile 3D artworks, and in particular of statues, has been presented by Levoy et al. (2000). The principal 3D scanner of the system is a laser triangulation scanner, built by Cyberware Inc., which can measure distances to within a quarter millimeter, and is mounted on a motorized gantry. The scanner head consists of a laser, a CCD sensor, a fiber-optic white light source, and a high-resolution color camera: The laser and the CCD sensor permit capturing the shape, while the light source and the color camera permit measuring the surface color of the statues scanned.

As far as the performances of 3D scanners are concerned, the time required to digitize an object, and the accuracy and resolution of the output (i.e. the 3D model of the object) vary significantly according to the technology employed, the size of the object and the desired resolution. Generally speaking, when mechanical systems are utilized, the shape and the size of the position transducer in contact with the object surface, represent physical constraints on the maximum resolution in the measurements. Moreover, mechanical systems are only suitable for reproducing the shape of objects with hard surfaces, since the load applied by the position transducer may cause a deformation of the surface which affects the accuracy of the measurements. Vision systems, both structured-light vision systems and stereo-vision systems, are usually faster than mechanical system, while their accuracy and resolution chiefly depends on the physical characteristics of the optical sensor employed and on its orientation with reference to the surface scanned.

Recently, academic and industrial researchers have been focusing on the possibility of employing technologies similar to the ones mentioned above in medical field, for applications such as manufacturing of orthoses, prostheses, and custom-made shoes for patients with severe deformities. All these applications involve measuring, and reproducing geometrically, human body parts, which can be safely, and not-invasively performed only making use of vision systems.

Yet, an obstacle to the use of vision systems for measuring purposes is the complexity of the mathematical models which are to be written for the cameras. In the paper by Gonzáles et al. (1993), the obstacle has been overcome modeling the CCD cameras employed like black boxes: the 3D real coordinates of the points sensed by the cameras are computed from the computer coordinates of the point images through empirical relations.

Several authors who have developed camera calibration techniques have also proposed interesting camera models. A complete camera model should include important internal parameters like the lens geometrical distortion, the effective focal length of the camera, the coordinates of the image center, and the horizontal scale factor. Anyway when the level of accuracy required by the application is not high, simplified models can be adopted like the one introduced by Wang and Tsai (1991), which is a pinhole camera model, neglecting both lens geometrical distortion and the horizontal scale factor. Weng et al. (1992) proposed and compared both a distortion-free camera model and a more complex model accounting for nonlinear terms such as radial and tangential distortion. A complete modelization

of the camera has been developed by Lenz and Tsai (1988): it accounts for the horizontal scale factor, the effective focal length of the camera, the radial lens distortion, and the coordinates of the image center. Batista et al. (1999) presented a very elaborate camera model also including the skew angle which is an internal parameter generally neglected by researchers. Finally, a very different approach to camera modelization has been adopted by Beyer (1992) who re-proposed a complete nonlinear mathematical model employed in photogrammetry.

In this work, which is subdivided into two parts, an opto-mechanical device is introduced. The device is designed to execute rapid, accurate, reproducible, and non-contact orthopedic measurements directly on human bodies and to create 3D models of the limbs measured.

The system operates using the structured-light vision technology and has been specifically conceived to simplify the design and manufacturing of custom-made shoes. The measurements taken by the system are employed to create the 3D model of a patient's foot; the model is then compared to a standard shoelast model, selected in a digital library, and the changes to be made to design a customized shoelast are identified. Basically, the shape of the foot measured and the one of the chosen shoelast are mixed to result in the shape of a new shoelast matching the actual shape of the patient's foot. Such a shoelast is therefore to be employed for manufacturing a custom-made shoe.

The objectives of Part I of this paper are:

- to introduce the operating method of the device presented,
- to determine an accurate kinematic and optical model of the system,
- to describe the data structure and the data acquisition procedure developed to speed up image acquisition and processing,
- to present the experimental prototype used to evaluate the system performance.

On the other hand, this paper does not deal with:

- the numerical filtering performed on the 2-D images acquired, processed and combined to get the 3D model of a foot,
- the algorithm which allows employing the measured data to modify a standard shoelast so as to design a customized shoelast,
- the geometrical smoothing algorithms implemented to improve the 3D mixed shape of a standard shoelast and a measured foot.

Part II of the paper introduces a calibration methodology for the system and summarizes the experimental results obtained.

2. System Operating Method

A schematic representation of the system is reported in Figure 1. The foot to be digitized is placed on the steady footboard while a CCD camera and two linear light sources ('lasers'), which are symmetrically positioned with respect to the camera, rotate like a rigid body around the foot. To simplify the drawing, in Figure 1 only

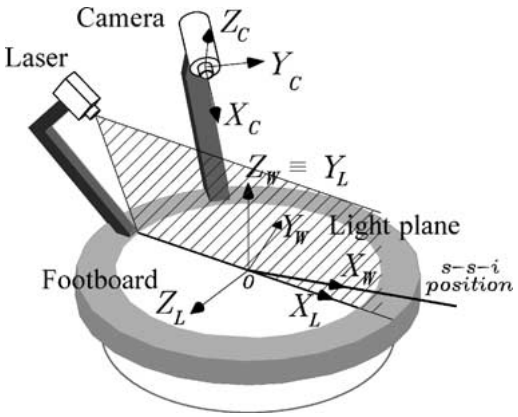


Figure 1. Schematic representation of the system showing the position and orientation of the world, light plane, and camera coordinate systems.

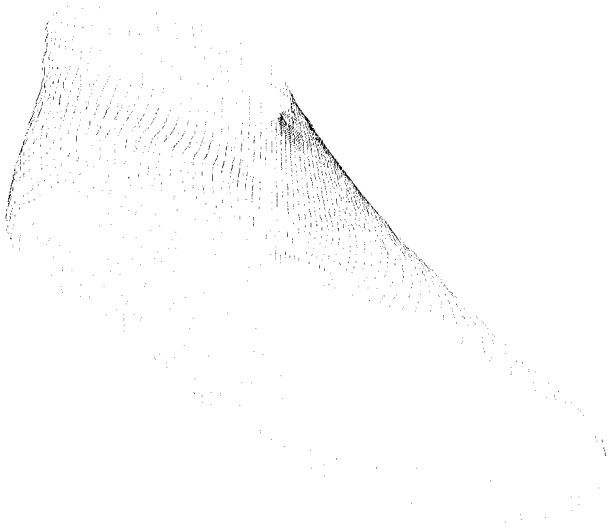


Figure 2. Aggregated image of the lines of light captured by the camera.

the laser on the right of the camera is shown. The center of the circular trajectory followed by the laser-camera system is the center of the machine steady footboard.

The camera is positioned and oriented so that it can sense the lines of light projected by the lasers on the foot surface. To prevent shadowing the two lasers are used alternately, one on the clockwise rotation of the laser-camera system about the foot, and the other on the counterclockwise rotation. While rotating, the camera observes the whole foot surface from different angles of view and, at each requested rotary position, it records the curved line segments, belonging to the light plane produced by the operating laser, and formed when the projected light strikes the foot.

As shown in Figure 2, the aggregation of all the 2-D images recorded by the camera allows reconstructing the 3D shape of the foot surface.

Subsequently, a mathematical representation of the foot surface is created, from which important geometrical information can be inferred, such as the overall length of the foot, the area of any section of it, its width and height at accurately chosen positions. Moreover, the 3D model of the foot can be employed to perform the aforementioned comparisons with the standard shoelasts in order to design customized shoelasts. *N*on-*U*niform *R*ational *B*-Splines (NURBS) have been employed to get a representation of the foot surface making use of mathematical equations. The use of the NURBS, which is not dealt with in this paper, also speeds up the comparison between a foot and a shoelast and simplifies important operations, such as the numerical filtering of the 2-D images acquired by the camera and the geometrical smoothing of the surface of the custom-made shoelast.

3. System model

The proposed vision system can be employed to perform precise 3D scans only if an accurate mathematical model of the system is developed and employed to extract information from the 2-D images of the camera. The model must take into consideration the kinematic and optical propreties of the lasers and of the camera. Figures 1 and 3, which only refers to the camera, show the position and orientation of all the cartesian coordinate systems considered in the mathematical model of the appliance. Henceforth only one of the two lasers employed is being considered: because they both operate in the same way, all the relations written with reference to one laser can be easyly modified to last also for the other one.

The world reference and light plane (laser beam) reference, marked respectively W and L, are both centered at O, which is the center of the footboard and also the center of rotation of the laser-camera system. The axis X_W lies on the footboard and is aligned to the projection of the camera optical axis at the start-stop-idle (*s*-*s*-*i*) position, which is the position where the laser-camera system starts from and stops

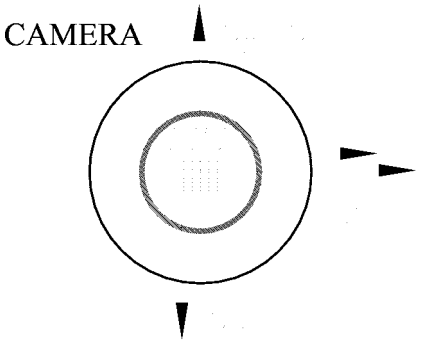


Figure 3. Schematic representation of the camera and of its CCD sensor. The camera, image plane and computer image coordinate systems are drawn.

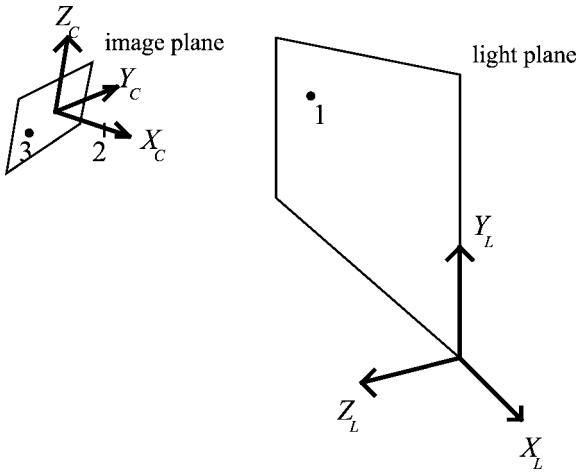


Figure 4. The three points considered lying on the same line (dashed) in the pinhole model.

at during the scans; the axis Z_W lies on the light plane. As far as the light plane reference is concerned, Z_L is perpendicular to the light plane, while X_L lies on both the footboard and the light plane, finally, Y_L coincides with Z_W. Because the light plane is orthogonal to the footboard and passes through its center, the position of the light plane in the world coordinate system can be directly determined by measuring the angle between X_W and X_L. Figure 3 shows that the camera reference and the image plane reference, which are respectively marked C and I, are both centered at the optical principal point (i.e., the intersection of the optical axis and the image plane). The axis X_C coincides with the optical axis while Y_C and Z_C are parallel respectively to the rows and columns of the sensor and coincide with Y_I and Z_I. Finally, the computer image reference, which is marked F, is centered at the top-left sensor element. The position of any point in such reference corresponds to the coordinate used to store its discrete image in the frame memory.

If a perfect pinhole camera model is adopted, initially neglecting optical distortion, a generic point on the light plane, its image on the image plane, and the lens optical center must lie on the same line. Figure 4 shows the three points aligned, whose coordinates in the camera reference are respectively (x_1, y_1, z_1), (x_3, y_3, z_3), and (x_2, y_2, z_2). In order to respect the mentioned geometric constraint, the matrix

$$\begin{bmatrix} x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{bmatrix} \tag{1}$$

must have rank 1, which can be expressed by the following system:

$$\begin{cases} (x_2 - x_1)(y_3 - y_1) - (x_3 - x_1)(y_2 - y_1) = 0, \\ (x_2 - x_1)(z_3 - z_1) - (x_3 - x_1)(z_2 - z_1) = 0, \\ (y_2 - y_1)(z_3 - z_1) - (y_3 - y_1)(z_2 - z_1) = 0. \end{cases} \tag{2}$$

Introducing the notations:

$$\mathbf{H}^* = \begin{bmatrix} y_2 - y_3 & x_3 - x_2 & 0 \\ z_2 - z_3 & 0 & x_3 - x_2 \\ 0 & z_2 - z_3 & y_3 - y_2 \end{bmatrix}, \tag{3}$$

$$\mathbf{K}^* = -\begin{bmatrix} x_2 y_3 - x_3 y_2 \\ x_2 z_3 - x_3 z_2 \\ y_2 z_3 - y_3 z_2 \end{bmatrix}, \tag{4}$$

the system in Equation (2) reduces to:

$$\mathbf{H}^* (x_1, y_1, z_1)^{\mathrm{T}} = \mathbf{K}^*. \tag{5}$$

It can be readily verified that the determinant of the matrix $\mathbf{H}^*$ is equal to zero, therefore ∞^1 solutions for (x_1, y_1, z_1) can be obtained solving Equation (5). Only if Equation (5) is solved considering the additional constraint that point 1 must also belong to the light plane, can the coordinates (x_1, y_1, z_1) be found. Assuming that the equation of the light plane in the camera coordinate system is

$$Ax_1 + By_1 + Cz_1 + D = 0 \tag{6}$$

such equation can be added to two independent equations extracted from Equation (5) to form a new system, whose matrix is called $\mathbf{H}$, and which only has one solution. In matrix notation the new system takes the form:

$$(x_1, y_1, z_1)^{\mathrm{T}} = \mathbf{H}^{-1}\mathbf{K} \tag{7}$$

where:

$$\mathbf{H} = \begin{bmatrix} -y_I & -f & 0 \\ -z_I & 0 & -f \\ A & B & C \end{bmatrix}, \tag{8}$$

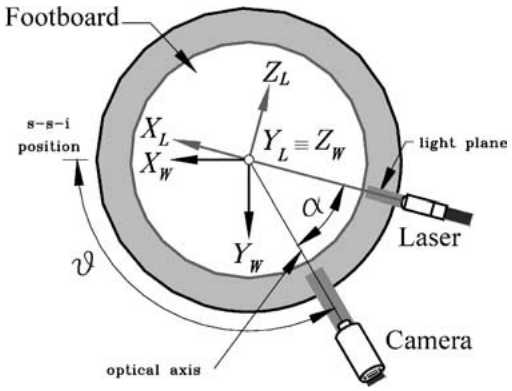
$$\mathbf{K} = \begin{bmatrix} -y_I f \\ -z_I f \\ -D \end{bmatrix}. \tag{9}$$

In accordance with a perfect pinhole model of the camera, it has been assumed that:

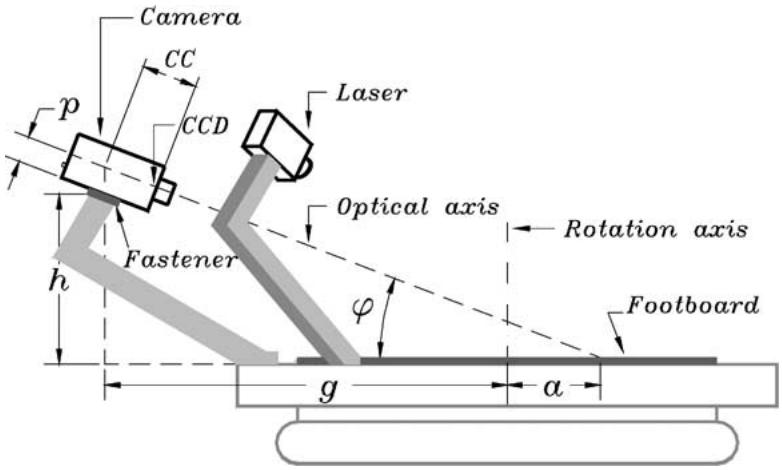
$$(x_2, y_2, z_2) = (f, 0, 0)\,, \tag{10}$$

$$(x_3, y_3, z_3) = (0, y_I, z_I) \tag{11}$$

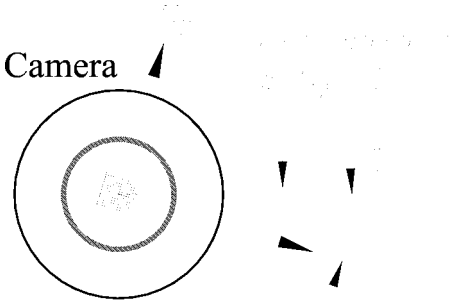
where f is the distance between the image plane and the optical center (i.e., the effective focal length), and (y_I, z_I) are the coordinates of point 3 in the image plane reference. The values of A, B, C and D depend on the geometry of the



(a)



(b)



(c)

Figure 5. Basic geometry of the system model.

system model, which is schematically illustrated in Figures 5(a), (b) and (c). The explicit expressions of A, B, C and D employed in the model and how they may be obtained are reported in the Appendix.

Equation (7) allows obtaining the position, in the camera reference, of a point in the light plane when the position of its image is known in the image plane reference. The discretization of the image caused by the CCD sensor and by the frame buffer, makes it necessary to compute the coordinates, in the 2-D image plane reference, of the image of any point starting from its 2-D computer image coordinates. The following transformation is employed to this purpose:

$$\begin{aligned} y_I &= s_y d_y (y_F - C_y), \\ z_I &= d_z (-z_F + C_z) \end{aligned} \tag{12}$$

where, with reference to Figure 3:

- (y_I, z_I) are the coordinates in the image plane reference of point 3,
- (y_F, z_F) are the computer image coordinates of the same point,
- d_y is the center to center distance between adjacent sensor elements in Y_F direction corresponding to the scan line direction,
- d_z is the center to center distance between adjacent sensor elements in Z_F direction,
- (C_y, C_z) are the computer image coordinate of the optical principal point,
- s_y is the horizontal scale factor, which is the conversion factor from the computer image pixel unit to the sensor element unit. An approximate estimate of the scale factor is obtained through Equation (13), having denoted with N_y the number of sensor elements in the scan line direction and with M_y the number of picture elements in a line as sampled by the computer:

$$s_y = \frac{N_y}{M_y}. \tag{13}$$

Lens distortion has not been considered yet. In this work only radial distortion has been introduced in the mathematical model. Tangential distortion is in fact negligible, as it usually is in most industrial machine vision applications. As a consequence of radial distortion, the true image of point 1 on the image plane does not coincide with point 3. The following transformation is employed to relate the coordinates of the distorted image of point 1 in the image plane reference to the coordinates of its undistorted image, computed in the same reference; it holds:

$$\begin{aligned} y_U &= y_D \left(1 + K_1 \rho^2 + K_2 \rho^4\right), \\ z_U &= z_D \left(1 + K_1 \rho^2 + K_2 \rho^4\right) \end{aligned} \tag{14}$$

where:

- (y_D, z_D) are the distorted image coordinates in the image plane reference,
- (y_U, z_U) are the undistorted image coordinates in the image plane reference, corresponding to the coordinates previously called (y_I, z_I),

- $\rho = \sqrt{y_D^2 + z_D^2}$ is the radial distance of the distorted image from the optical principal point,
- K_1 and K_2 are the coefficients of radial distortion.

After using Equation (7) to compute the coordinates of a generic point on the light plane in the camera reference, determining the position of the same point in the 3D world reference is straightforward. In fact, as underlined above, the light beam emitted by the laser is orthogonal to the machine steady footboard and passes through the footboard center, which is the origin of the world reference, of the light plane reference, and also the center of rotation of the laser-camera system. The position of the light plane in the world coordinate system can therefore be determined by measuring the angle between X_W and X_L which expresses the rotation of the laser-camera system about the footboard center. In the experimental prototype such an angle is measured by means of an optical encoder. The following equation expresses the rigid body transformation from the 3D camera coordinate system (x_C, y_C, z_C) to the 3D world coordinate system (x_W, y_W, z_W):

$$\begin{bmatrix} x_C \\ y_C \\ z_C \end{bmatrix} = \mathbf{R}_{WC} \begin{bmatrix} x_W \\ y_W \\ z_W \end{bmatrix} + \mathbf{T}_{WC} \tag{15}$$

where:

- $\mathbf{R}_{WC}$ is the 3×3 rotation matrix expressing the orientation of the world reference in the camera reference,
- $\mathbf{T}_{WC}$ is the 3×1 translation vector expressing the position of the world reference origin in the camera reference.

In order to determine the value of the elements of $\mathbf{R}_{WC}$, the 3D rotation about the origin necessary to match the orientation of the world reference with the one of the camera reference is expressed as a sequence of rotations ϑ, φ and σ about the axes Z_W, Y_W, and X_W. Figures 5(a), (b) and (c) shows the mentioned angles. It is possible to write:

$$\mathbf{R}_{WC} = \begin{bmatrix} -\cos\varphi\cos\vartheta & -\cos\varphi\sin\vartheta & -\sin\varphi \\ \cos\sigma\sin\vartheta + \sin\sigma\sin\varphi\cos\vartheta & -\cos\sigma\cos\vartheta + \sin\sigma\sin\varphi\sin\vartheta & -\sin\sigma\cos\varphi \\ \sin\sigma\sin\vartheta - \cos\sigma\sin\varphi\cos\vartheta & -\sin\sigma\cos\vartheta - \cos\sigma\sin\varphi\sin\vartheta & \cos\sigma\cos\varphi \end{bmatrix} \tag{16}$$

and

$$\mathbf{T}_{WC} = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} = \begin{bmatrix} \left(\dfrac{a+g}{\cos\varphi}\right) - a\cos\varphi - CC \\ a\sin\sigma\sin\varphi \\ -a\cos\sigma\sin\varphi \end{bmatrix} \tag{17}$$

where:

$$a = \left(h + \frac{p}{\cos\varphi}\right)\cot\varphi - g. \tag{18}$$

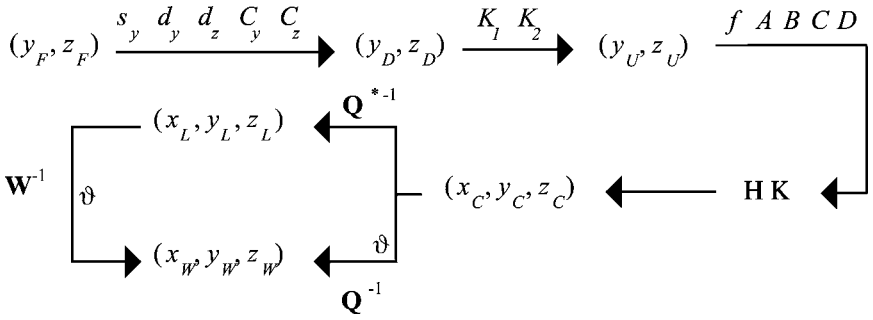


Figure 6. Transformation from 2-D computer image coordinate to 3D world coordinate.

Assembling the matrices $\mathbf{R}_{WC}$ and $\mathbf{T}_{WC}$ in the transformation matrix $\mathbf{Q}$, it holds:

$$\mathbf{Q} = \left[\begin{array}{c|c} \mathbf{R}_{WC} & \mathbf{T}_{WC} \\ \hline 0 & 1 \end{array}\right]. \tag{19}$$

There follows that the homogeneous coordinates in the world reference of a generic point on the light plane can be computed from its coordinates in the camera reference by means of the transformation:

$$(x_W, y_W, z_W, 1)^{\mathrm{T}} = \mathbf{Q}^{-1}(x_C, y_C, z_C, 1)^{\mathrm{T}}. \tag{20}$$

Similarly, it is possible to write:

$$(x_L, y_L, z_L, 1)^{\mathrm{T}} = \mathbf{Q}^{*^{-1}}(x_C, y_C, z_C, 1)^{\mathrm{T}}, \tag{21}$$

$$(x_W, y_W, z_W, 1)^{\mathrm{T}} = \mathbf{W}^{-1}(x_L, y_L, z_L, 1)^{\mathrm{T}}. \tag{22}$$

As it is reported in the Appendix, the elements of the matrices $\mathbf{Q}^*$ and $\mathbf{W}$ are computed through a procedure similar to the one described above. In the next section the role of these two marices is explained: $\mathbf{Q}^*$ becomes useful to speed up data acquisition while $\mathbf{W}$ is employed to post-process images.

To sum up, the overall transformation from the two-dimensional computer image coordinates (y_F, z_F) to the three-dimensional world coordinates (x_W, y_W, z_W) is schematically illustrated in Figure 6. The parameters that are to be known or computed at each step of the transformation are written near the arrows.

4. Data Acquisition Procedure

Two important aspects were taken into consideration when defining the data acquisition procedure:

- the knowledge of the light plane rotary position is necessary only at the very end of the overall transformation from the 2-D computer image coordinate system to the 3D world coordinate system (see Equation (20) and Figure 6);

- while scanning there are no changes in the relative position of the camera and each laser, hence a $\mathbb{R}^2$ to $\mathbb{R}^2$ biunique correspondence exists between any point of the light plane within the field of view of the camera and its projection in the camera image plane.

The overall transformation of 2-D computer image coordinates into 3D world coordinates can therefore be split into two independent transformations, the first associating the light plane coordinates of a point on the light plane with the image plane coordinates of its image, and the second associating the world reference coordinates with the light plane coordinates through the rotary position ϑ (Equation (22)).

The splitting of these two separate transformations leads to a considerable reduction in scanning time, in fact the real-time use of the formulas written in the previous section for all the points of the line of light sensed by the camera at each rotary position would imply handling trigonometric terms and computing the inverse of the matrix $\mathbf{Q}$, which slows down pixel transformation.

On the basis of the aforementioned considerations data have been arranged in an appropriate structure which reduces the calculations to be carried out while scanning. At the end of the calibration procedure (to be described in Part II of this paper), the mathematical model of the system is employed to yield two tables, called *calibration tables*, associating the computer image coordinates of all the pixels with the light plane coordinates of the corresponding points on the two laser light planes. While scanning, the acquisition program can therefore quickly convert the computer image coordinates of each pixel of the image of the line of light, into the light plane coordinates of the corresponding point in the light plane projected by the laser used. This process is repeated at each chosen position of the circular trajectory swept out by the laser-camera system. The data gathered, also including

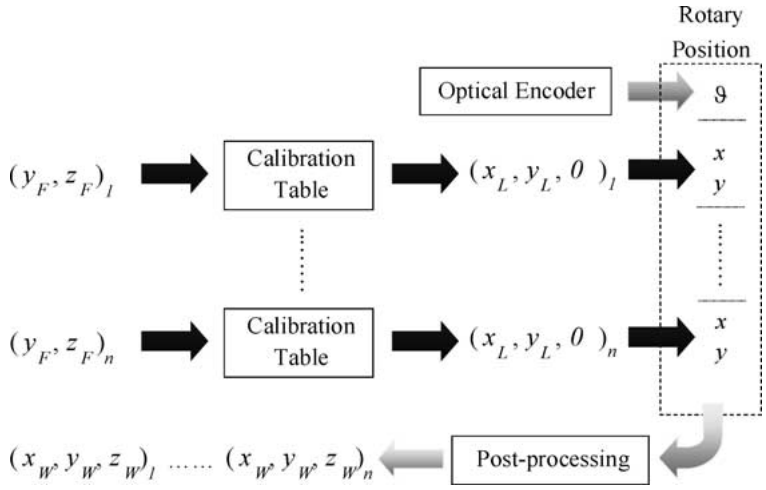


Figure 7. Scheme of the data acquisition procedure.

the rotary position ϑ, are stored in a matrix of appropriate dimensions. These data, basically providing the position of the points in cylindrical coordinates, are then post-processed using Equation (22) to compute the position, in the world reference of all the points sensed. Post-processing is carried out after scanning and hence it does not affect scanning time.

Figure 7 shows a schematic representation of the acquisition procedure. The computer image of the line of light sensed at the ϑ rotary position is supposed to be composed of n pixels. In the scheme only one rotary position is considered, therefore not a matrix but a vector of coordinates is generated.

5. Laboratory prototype

The experimental prototype of the system has been manufactured in cooperation with Label Elettronica S.r.l., an Italian company based in Padova. The prototype consists of the following components:

- one PULNIX TM 500 CCD B/W video camera with $582_h \times 500_v$ sensor elements spaced between each other 12.7 μm in the horizontal direction and 8.3 μm in the vertical direction;
- one FlashBus MV Pro frame grabber manufactured by Integral Technologies;
- one Computar C-mount lens with a focal length of 8 mm;
- two laser sources, manufactured by Lasertech S.r.l., each producing a vertical and planar laser-beam characterized by a 0.7/0.3 mrad divergence. The wavelength of the laser beam is 660 nm, corresponding to red light, and the power of the emission is 4 mW. The optics of the laser is spot-adjustable and the spot is set at 450 mm;
- one Angst & Pfister SWF 403957-W193 DC motor exerting a maximum torque of 5 Nm;
- one optical encoder with a resolution of 8000 steps/mech.rev.;
- one Advantech PCL-833 three-axis quadrature encoder card;
- one Advantech PCL-725 digital I/O card, employed to switch on and off the lasers and to control the DC motor operation;
- one Intel Pentium II 300 MHz processor PC with 64 MB RAM;
- a metallic framework including:
  - one rotating carriage,
  - one central steady footboard where the foot to be scanned is placed,
  - one elevated steady stand for the foot idle,
  - one wide handhold which patients can cling to.

The DC motor, the optical encoder, and all the other mechanical and electrical devices necessary to transmit motion and to carry electrical signals and power are fit inside the central steady footboard; Figure 8 shows a picture of the prototype. As mentioned above, two laser sources are used alternately in order to prevent shadowing, in fact it often happens that the camera cannot see the line of light projected by the laser on the foot when the line is behind a protuberant part. The relative

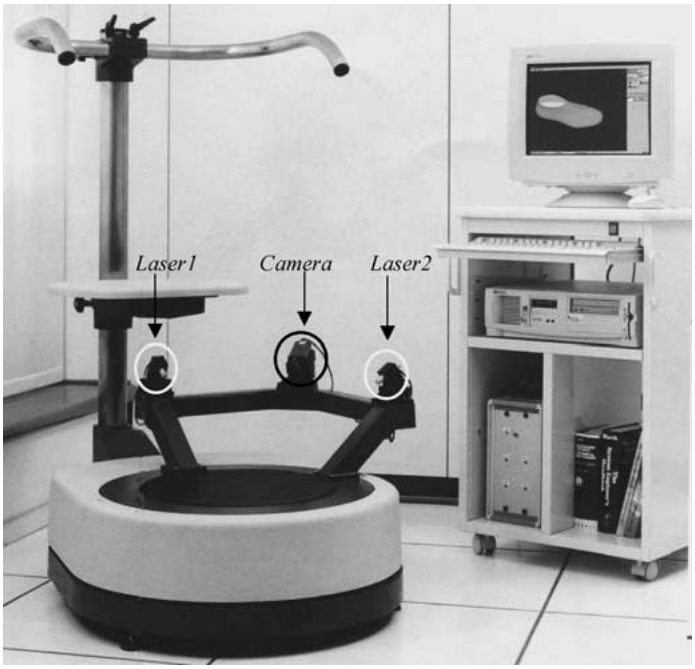


Figure 8. The experimental prototype.

position of the lasers and the camera cannot change during and after scanning, in fact, the lasers and the camera are fixed to the same carriage rotating about the center of the machine steady footboard. The two lasers forms a precise angle of $\pi/2$ rad, and they are located symmetrically with respect to the camera (they can be seen in Figure 8, one on the left, and the other on the right of the camera) so that each laser-beam forms an angle α of approximately $\pi/4$ rad with the projection of the camera optical axis on the footboard. The projection of the camera optical axis and the laser-beams pass through the center of the footboard, which coincides with the center of rotation of the laser-camera system.

So as to assure that both the projection of the camera optical axis and the light planes emitted by the lasers pass through the center of rotation of the laser-camera system, the lasers and the camera are mounted on precision guides, allowing a manual adjustment of their positioning. The relative position among the lasers and the camera can then be accurately established through the calibration procedure described in Part II.

The prototype chief geometrical characteristics are synthetically reported in Table I. The meaning of the symbol can be inferred observing Figures 5(b) and (c).

A user-friendly window interface (see Figure 9) has been developed to easily control the system operation and to modify some relevant parameters affecting the acquisition process, among which the most important is the binarization threshold. The binarization threshold is a value set to properly convert the gray scale image

Table I. Chief geometrical characteristics of the prototype

h	**g**	**p**	**CC**	φ	σ
0.210 m	0.560 m	0.0280 m	0.0669 m	23 deg.	0 deg.

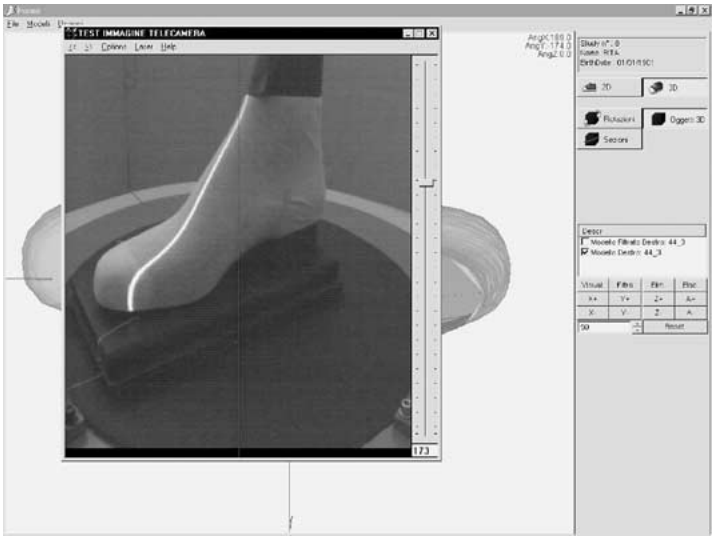


Figure 9. The user interface: a gray scale image of a shoelast and of the projected laser light is shown overlapped with a 3D image of the same shoelast. The slider used to set the value of the threshold is visible as well as the main window tool bar.

acquired by the camera into a binary image showing only the line of light projected by the laser on the foot surface. Before each scan, the choice of the most appropriate value of the threshold should be made taking into consideration the instantaneous illumination of the background. The desired value can be quickly found moving the slider shown in Figure 9 and observing the effect on the binary image on the PC display screen.

Through the same interface it is also possible to manage the patients' data, to access the shoelast database, to perform measurements on the foot model, to run the procedure computing the shape of a customized shoelast, to execute image processing, such as numerical filtering and geometrical smoothing, and to set-up and call the machine calibration procedure.

Most of the aforementioned operations are not described and analyzed here, being of limited scientific interest. Nonetheless, in this conclusive part of the paper a brief description of the operations made on the data acquired and stored is reported to prove the effectiveness of the system.

Once the 3D shape of a patient's foot is reconstructed through the aggregation of the selected images of the lines of light captured by the camera (Figure 2), a 3D

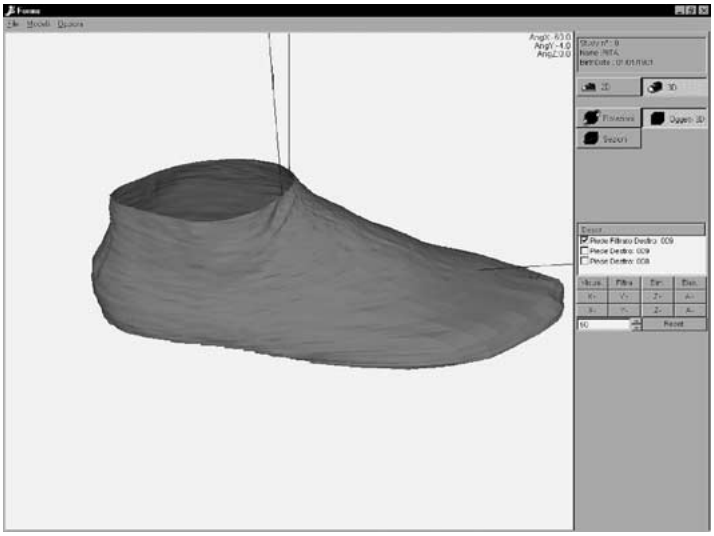


Figure 10. Reconstructed 3D image of a foot.

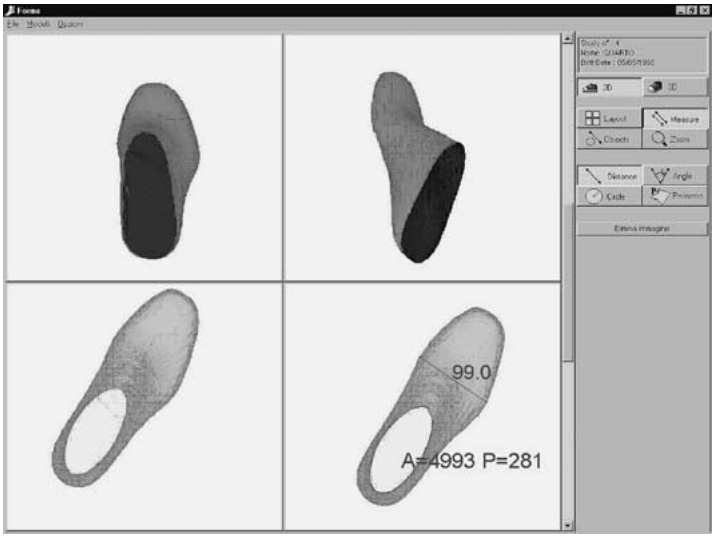


Figure 11. Examples of 2-D sections, and of geometrical features extracted by the 3D image of a foot.

mathematical representation of the foot is built by the NURBS, and a 3D image of the foot is reconstructed (Figure 10). Through the interface developed, the image can be easily rotated, enlarged, and 2-D sections of the image may be drawn on the user-defined plans. Moreover interesting geometrical features, like the area and the perimeter of a 2-D section, or the width of the foot in a selected position, can be extracted directly from the foot image (Figure 11).

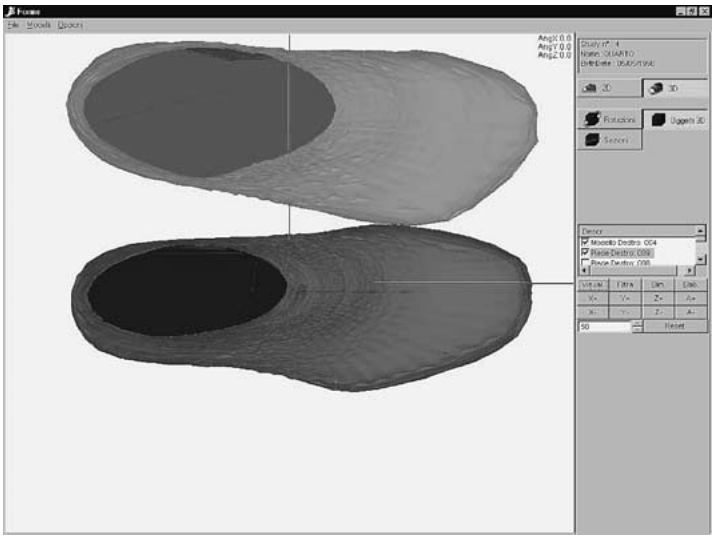


Figure 12. Plan view of a sample foot and of a standard shoelast.

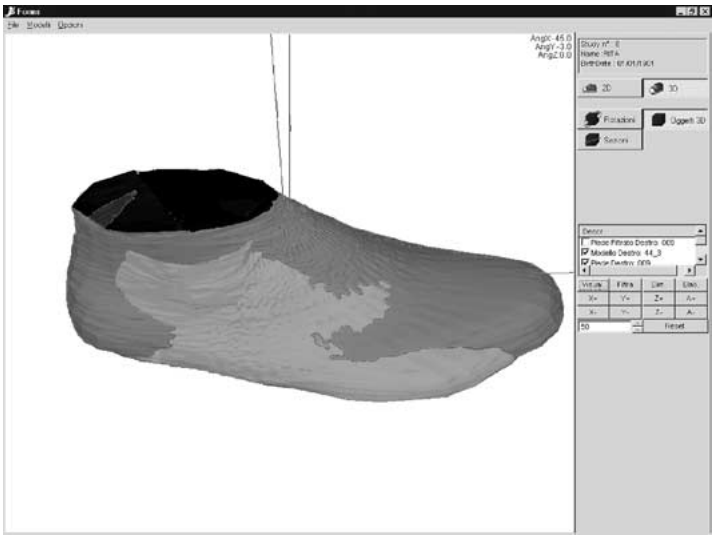


Figure 13. Customized shoelast.

Subsequently, as it has been mentioned above, the 3D model of the foot can be compared to the standard shoelasts available in a library, in order to chose the best fitting shoelast, and to modify the shoelast chosen so as to design a customized shoelast matching the actual shape of the patient's foot. Figure 12 shows the plan view of a sample foot acquired through the system (pale gray) and of a standard shoelast (dark gray). Figure 13 shows the customized shoelast: the pale-gray areas represent the areas where the standard shape of the shoelast has been modified to contain the entire volume of the foot. In the customized shoelast, geometrical

smoothing algorithms have been adopted to obtain gradual transitions between the modified and the unchanged surfaces.

6. Conclusions

The objective of executing accurate, reproducible, fast and not-invasive 3D non-contact orthopedic measurements has been achieved developing an opto-mechanical appliance which makes use of the structured-light vision technology. The system chief components are a CCD camera and two linear laser sources which rotate about the foot to be scanned. The 2-D images of the lines of light captured by the camera are aggregated and processed to create a 3D model of the foot. This model can be compared to a requested standard shoelast to design a customized shoelast.

A complete modelization of the system has been introduced in this work: it allows transforming the two-dimensional computer image coordinates of a point image into the three-dimensional world coordinates of the point. The real-time use of such a transformation leads to high scanning time because the formulas include trigonometric terms and the inverse of a matrix, which slows down the overall acquisition process. A faster acquisition procedure has therefore been implemented, which still employs the mathematical relations identified, but requires arranging data in an appropriate structure.

A laboratory prototype of the system is described, and an example of the customized shoelast obtained mixing the shape of a measured sample foot and that of a standard shoelast is shown.

An accurate calibration of the system is necessary to ensure the accuracy and reproducibility of the observations. The calibration methodology conceived and implemented is described in detail in Part II of the this paper, where also the experimental results obtained using the system prototype are summarized.

Appendix

An approach similar to the one described in Section 3 to determine the elements of matrix $\mathbf{Q}$, is employed to compute the elements of matrix $\mathbf{Q}^*$:

$$\mathbf{Q}^* = \left[\begin{array}{c|c} \mathbf{R}_{LC} & \mathbf{T}_{LC} \\ \hline 0 & 1 \end{array}\right]. \tag{A.1}$$

where

$$\mathbf{R}_{LC} = \begin{bmatrix} \cos\varphi\cos\alpha & -\sin\varphi & \cos\varphi\sin\alpha \\ \cos\sigma\sin\alpha - \sin\sigma\sin\varphi\cos\alpha & -\sin\sigma\cos\varphi & -\cos\sigma\cos\alpha - \sin\sigma\sin\varphi\sin\alpha \\ \sin\sigma\sin\alpha + \cos\sigma\sin\varphi\cos\alpha & \cos\sigma\cos\varphi & -\sin\sigma\cos\alpha + \cos\sigma\sin\varphi\sin\alpha \end{bmatrix} \tag{A.2}$$

and

$$\mathbf{T}_{LC} = \mathbf{T}_{WC} = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}. \tag{A.3}$$

The coordinates t_x, t_y and t_z express the position of the light plane origin in the camera reference, they are the same as those in Equation (17) since the origin of the light plane coordinate system coincides with the origin of the world coordinate system.

The overall rotation necessary to match the orientation of the light plane reference with that of the camera reference, is expressed as a sequence of rotations α, φ and σ about the axes Z_L, Y_L, and X_L. The meaning of the angles α, φ, and σ can be inferred from Figures 5(a), (b), and (c) where only one of the two lasers adopted is shown. In this paper, all the formulas containing the symbol α refer to the laser shown. Because the two lasers actually employed are symmetrically located with respect to the camera, if it is necessary to refer to the laser which is not shown in Figure 5, the symbol α needs to be replaced with $-\alpha$.

The matrix $\mathbf{W}$, expressing the transformation from the world coordinate system to the light plane coordinate system takes the form:

$$\mathbf{W} = \begin{bmatrix} -\cos(\vartheta+\alpha) & -\sin(\vartheta+\alpha) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sin(\vartheta+\alpha) & \cos(\vartheta+\alpha) & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{A.4}$$

In fact, a rotation of $-\pi/2$ about the axis X_L followed by a rotation $(\vartheta+\alpha)$ about the axis Z_L make the two coordinate systems coincide.

In order to compute the values of the parameters A, B, C and D of the equation of the light plane in the camera coordinate system (Equation (6)), the coordinates, in the camera reference, of three points lying on the light plane may be considered. The following three points with very simple homogeneous coordinate expressions are employed:

$$\begin{aligned} (0,0,0,1)_L^{\mathrm{T}} &\equiv (t_x, t_y, t_z, 1)_C^{\mathrm{T}}, \\ (1,0,0,1)_L^{\mathrm{T}} &\equiv \big(\cos\varphi\cos\alpha + t_x, -\cos\sigma\sin\alpha - \sin\sigma\sin\varphi\cos\alpha + t_y \\ &\qquad -\sin\sigma\sin\alpha + \cos\sigma\sin\varphi\cos\alpha + t_z, 1\big)_C^{\mathrm{T}}, \\ (0,1,0,1)_L^{\mathrm{T}} &\equiv \big(-\sin\varphi + t_x, -\sin\sigma\cos\varphi + t_y, \cos\sigma\cos\varphi + t_z, 1\big)_C^{\mathrm{T}}. \end{aligned} \tag{A.5}$$

The coordinates marked with the subscript L are expressed in the light plane coordinate system. Those marked with the subscript C are expressed in the camera coordinate system and are computed employing Equation (21). Because the three

points lie on the same plane, whose equation is assumed to be Equation (6), it holds:

$$A = \begin{vmatrix} t_y & \\ \cos\sigma\sin\alpha - \sin\sigma\sin\varphi\cos\alpha + t_y & , \\ -\sin\sigma\cos\varphi + t_y & \end{vmatrix}$$
$$\begin{vmatrix} t_z & 1 \\ \sin\sigma\sin\alpha + \cos\sigma\sin\varphi\cos\alpha + t_z & 1 \\ \cos\sigma\cos\varphi + t_z & 1 \end{vmatrix}, \tag{A.6}$$

$$B = -\begin{vmatrix} t_x & t_z & 1 \\ \cos\varphi\cos\alpha + t_x & \sin\sigma\sin\alpha + \cos\sigma\sin\varphi\cos\alpha + t_z & 1 \\ -\sin\varphi + t_x & \cos\sigma\cos\varphi + t_z & 1 \end{vmatrix}, \tag{A.7}$$

$$C = \begin{vmatrix} t_x & t_y & 1 \\ \cos\varphi\cos\alpha + t_x & \cos\sigma\sin\alpha - \sin\sigma\sin\varphi\cos\alpha + t_y & 1 \\ -\sin\varphi + t_x & -\sin\sigma\cos\varphi + t_y & 1 \end{vmatrix}, \tag{A.8}$$

$$D = -\begin{vmatrix} t_x & t_y & \\ \cos\varphi\cos\alpha + t_x & \cos\sigma\sin\alpha - \sin\sigma\sin\varphi\cos\alpha + t_y & \\ -\sin\varphi + t_x & -\sin\sigma\cos\varphi + t_y & \end{vmatrix}$$
$$\begin{vmatrix} t_z & \\ \sin\sigma\sin\alpha + \cos\sigma\sin\varphi\cos\alpha + t_z & \\ \cos\sigma\cos\varphi + t_z & \end{vmatrix}. \tag{A.9}$$

Computing the determinants of the matrices it follows:

$$\begin{aligned} A &= \cos\varphi\sin\alpha, \\ B &= -\cos\sigma\cos\alpha - \sin\sigma\sin\varphi\sin\alpha, \\ C &= -\sin\sigma\cos\alpha + \cos\sigma\sin\varphi\sin\alpha, \\ D &= -t_x\cos\varphi\sin\alpha + t_y(\cos\sigma\cos\alpha + \sin\sigma\sin\varphi\sin\alpha) \\ &\quad + t_z(\sin\sigma\cos\alpha - \cos\sigma\sin\varphi\sin\alpha). \end{aligned} \tag{A.10}$$

References

Batista, J., Araújo, H., and de Almeida, A. T.: 1999, Iterative multistep explicit camera calibration, *IEEE Trans. Robotics Automat.* **15**(5), 897–917.

Beyer, H. A.: 1992, Accurate Calibration of CCD-Cameras, in: *Proc. of 1992 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 96–101.

Gonzáles, F., Campoy, P., Aracil, R., Peñafiel, F., and Sebastián, J. M.: 1993, Three-dimensional digitizer for the footwear industry, *SPIE Comput. Vision Industry* **1989**, 332–338.

Levoy, M., Pulli, K., Curless, B., Rusinkiewicz, S., Koller, D., Pereira, L., Ginzton, M., Anderson, S., Davis, J., Ginsberg, J., Shade, J., and Fulk, D.: 2000, The Digital Michelangelo Project: 3D scanning of large statues, in: *Proc. of the Siggraph 2000 27th Internat. Conf. on Computer Graphics and Interactive Techniques*.

Lenz, R. K. and Tsai, R. Y.: 1988, Techniques for calibration of the scale factor and image center for high accuracy 3D machine vision metrology, *IEEE Trans. Pattern Anal. Mach. Intelligence* **10**(5), 713–720.

Tsai, R. Y.: 1987, A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses, *IEEE J. Robotics Automat.* **3**(4), 323–344.

Wang, L. L. and Tsai, W. H.: 1991, Camera calibration by vanishing lines for 3D computer vision, *IEEE Trans. Pattern Anal. Mach. Intelligence* **13**(4), 370–376.

Weng, J., Cohen, P., and Herniou, M.: 1992, Camera calibration with distortion models and accuracy evaluation, *IEEE Trans. Pattern Anal. Mach. Intelligence* **14**(10), 965–980.