Optical Engineering

OpticalEngineering.SPIEDigitalLibrary.org

Modification of calibration and image processing procedures for precise 3-D measurements in arbitrary spectral bands by means of a stereoscopic prism-based imager

Alexander S. Machikhin
Alexey V. Gorevoy
Demid D. Khokhlov
Arseny O. Kuznetsov

Alexander S. Machikhin, Alexey V. Gorevoy, Demid D. Khokhlov, Arseny O. Kuznetsov, “Modification of calibration and image processing procedures for precise 3-D measurements in arbitrary spectral bands by means of a stereoscopic prism-based imager,” *Opt. Eng.* **58**(3), 033104 (2019), doi: 10.1117/1.OE.58.3.033104.

Optical Engineering 58(3), 033104 (March 2019)

Modification of calibration and image processing procedures for precise 3-D measurements in arbitrary spectral bands by means of a stereoscopic prism-based imager

Alexander S. Machikhin,[a,b,*] Alexey V. Gorevoy,[a,b] Demid D. Khokhlov,[a,b] and Arseny O. Kuznetsov[b]
[a]National Research University "Moscow Power Engineering University," Institute of Automatics and Computer Engineering, Department of Electrical Engineering and Introscopy, Moscow, Russia
[b]Russian Academy of Sciences, Laboratory of Acousto-Optic Spectroscopy, Scientific and Technological Center of Unique Instrumentation, Moscow, Russia

Abstract. Nowadays, stereoscopic devices are widely used for precise three-dimensional (3-D) measurements of hard-to-reach elements in industrial and biomedical applications. The most common approach for their implementation is the utilization of prism-based optical systems, which allows simultaneous acquisition of two images from different viewpoints on a single sensor. Normally, they are equipped with a wideband (WB) white-light source, but contrast visualization of the inspected object and, therefore, accurate quantitative characterization of its parameters, often requires narrow-band (NB) spectral imaging. We show that the standard geometrical calibration may lead to significant measurement errors when obtained using WB illumination and applied to NB images. As the criterion to evaluate the calibration error, we have chosen the difference between reference and measured lengths of the 1 mm segment along transverse x, y, and longitudinal z axes. If WB calibration data are applied to NB images, the measurement bias increases from the middle of the working spectral range to its edges and can reach significant values: up to 0.1 mm along x axis and 0.15 mm along z axis in 10 to 25 mm distance range. To overcome this, we propose the calibration and image processing procedures based on a proper choice of a few spectral bands for calibration and interpolation of the calculated calibration parameters. Results of multiple experiments using stereo video endoscope confirmed that the proposed technique allows a decrease in the measurement bias by three times in comparison to conventional WB calibration for all wavelengths of the visible range, which essentially improves the measurement accuracy. The impact of WB calibration on random errors of measurements and the quality of image rectification was also analyzed and shown to be insignificant. [DOI: 10.1117/1.OE.58.3.033104]

Keywords: prism-based stereoscopic optical system; three-dimensional measurements; video endoscope; multispectral imaging; geometrical calibration; image rectification.

Paper 181170 received Aug. 11, 2018; accepted for publication Feb. 25, 2019; published online Mar. 20, 2019.

1 Introduction

Stereoscopic video endoscopes have become one of the main tools for nondestructive testing (NDT) in aircraft, automobile, energy, and other industries. These tools are widely used for quantitative characterization of hard-to-reach elements and defects inside complex objects without their disassembly. Remote visual inspection (RVI) using stereoscopic video endoscopes allows noncontact three-dimensional (3-D) geometrical measurements of the structural elements accessible only by small-diameter (4 to 8 mm) probes.[1,2]

The mostly used approach to the implementation of this technique is the utilization of a miniature prism-based optics for acquisition of two images from two different points on a single sensor (Fig. 1). In this case, an interchangeable stereo tip adapter has to be attached to the distal end of the video probe in front of the integrated lens.[2,3] Illumination from the white-band (WB) light source is delivered to the inspected object via the fiber optic light guide located inside the probe.

To compute 3-D coordinates of the object points, conventional stereoscopic technique may be implemented.[4–6] It employs a calibration procedure, image rectification, and other processing algorithms for pixel matching prior to the start of the measurement process resulting in a full 3-D surface map of the inspected object. An advanced ray tracing camera model allows to take into account a specific distortion introduced by the prism.[7–10]

In some applications, contrast visualization of the inspected surface using WB illumination and, therefore, defects localization and quantitative characterization is not effective or even impossible. In these cases, fluorescence, Raman, and other narrow-band (NB) spectral imaging techniques may be helpful to increase the performance of the RVI procedure.[11,12] For example, fluorescent penetrant inspection requires selection of ultraviolet light to ensure good contrast between the glow emitted by the penetrant in the defected areas and the unlit surface of the material.[13] For the material identification, the reflectance spectra in the wide range are effective.[14] In this case, it is especially important to analyze both local spectral properties of the object and spatial distribution of these properties. Therefore, spectral imaging feature integrated into the stereoscopic video endoscopes makes

---

*Address all correspondence to Alexander S. Machikhin, E-mail: aalexanderr@mail.ru

0091-3286/2019/$25.00 © 2019 SPIE

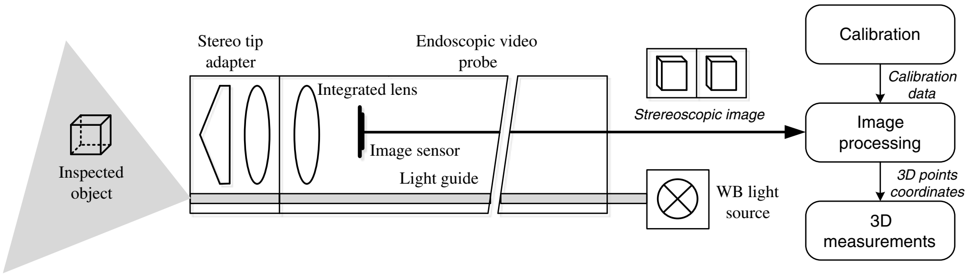


Fig. 1 Conventional concept of a video endoscope for 3-D stereoscopic measurements.

these tools more informative and flexible for real-time NDT applications. This feature may be implemented by means of the spectral filtration of the illumination irradiated by the WB light source, for example, using the tunable filter or a filter wheel conjugated with this source. Figure 2 shows examples of the effectiveness of spectral imaging for contrast visualization of the defect.

Conventional geometrical calibration procedure is implemented for white-light illumination limited only by spectral width of the light source and sensitivity of the sensor. Such calibration is not optimal for NB illumination and may limit the achievable measurement accuracy due to inevitable chromatic aberrations of the prism-based optical system.[15] Dispersion of light caused by the prism leads to the spectral dependence of the focal distance, distortion coefficients, and other parameters of the whole optical system. This must be taken into account at all stages of stereoscopic image processing: calibration, rectification, etc. Therefore, for proper 3-D measurements under NB illumination, it is necessary to modify conventional algorithms.

In this paper, we experimentally demonstrate this limitation and describe the calibration procedure proposed in our previous work[16] and based on a proper choice of a few spectral bands to decrease the measurement errors provided by stereo video endoscopes in NB light. We analyze a ray tracing model and spectral dependence of its parameters, show that for operation in any selected spectral band within visible range, the system may be calibrated in a few NBs with further interpolation of calibration parameters, compare the effectiveness of the proposed approaches to calibration and image rectification procedures with the conventional ones.

2 Mathematical Models and Data Processing

Processing of stereoscopic images includes stereo matching, calculation of 3-D point coordinates, geometrical measurements and/or reconstruction of 3-D surface. In order to increase the accuracy and speed of image matching, some preprocessing techniques, such as rectification, distortion correction, and noise reduction may be implemented. For precise rectification as well as for calculation of 3-D point coordinates, one needs a mathematical model, which converts 2-D image coordinates into 3-D ray coordinates in the object space. Preliminary calibration of the system allows obtaining the parameters of this model. Since the ray tracing camera model provides better measurement accuracy for prism-based stereoscopic systems,[10] all theoretical considerations below are based on this model.

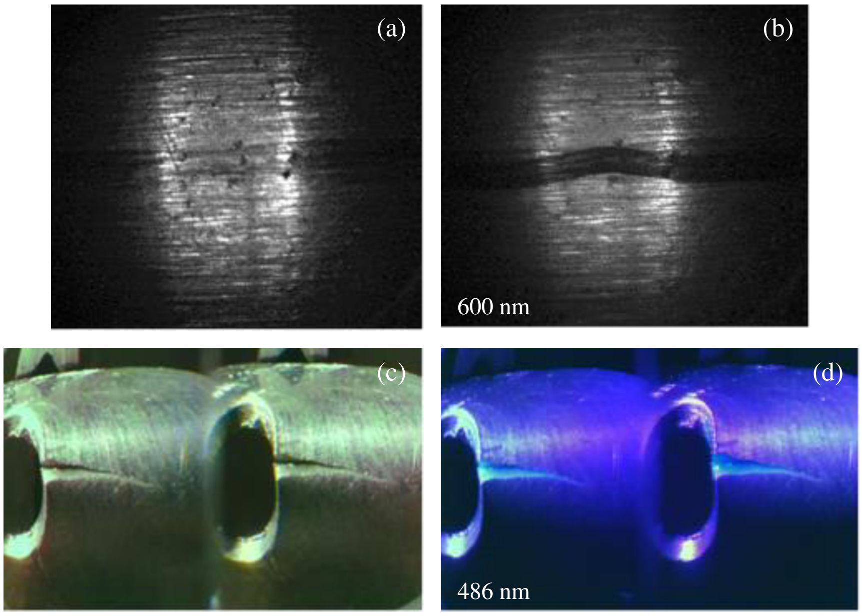


Fig. 2 (a, b) WB and NB images of the oil on metal surface obtained using conventional video endoscope and (c, d) a crack using fluorescent penetrant inspection via stereoscopic video endoscope.

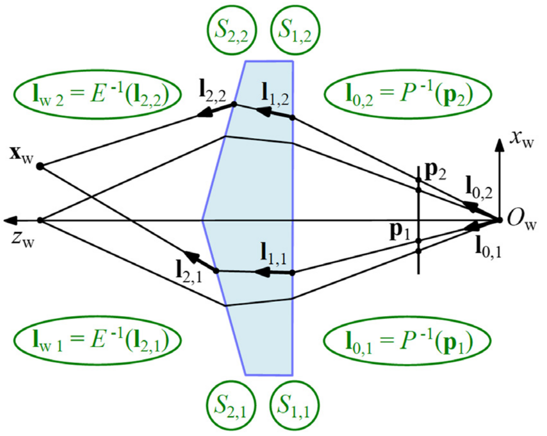


Fig. 3 Ray tracing camera model for prism-based stereoscopic system.

2.1 *Ray Tracing Camera Model*

The image acquired by the prism-based stereoscopic system consists of two parts. Each of them is formed by the rays passing through one part of the biprism. To make the mathematical description applicable to multiocular prisms,[7] we consider that the image consists of N parts obtained via N parts of the prism and further refer to them as "i'th image part" and "i'th prism part," $i = 1,2\dots N$. We formulate the model using a ray tracing from the image plane to the object space.[8,10] We define the vector of ray coordinates $\mathbf{l} = (\mathbf{c}^T, \mathbf{v}^T)^T$ as the concatenation of 3-D coordinates of the origin point $\mathbf{c}$ and the direction vector $\mathbf{v}$ and use $\mathbf{l}_{m,i} = S_{m,i}(\mathbf{l}_{m-1,i})$ notation to describe the refraction of each ray $\mathbf{l}_{m,i}$ on the m'th surface for i'th image part. Each transformation $S_{m,i}$ is derived according to the vector form of Snell's law.[10] To find the ray coordinates $\mathbf{l}_{2,i}$ in the object space for each image point $\mathbf{p}_i$, we consider the pinhole model with radial distortion for the main lens, apply it to define ray coordinates $\mathbf{l}_{0,i}$, and further use $S_{1,i}$ and $S_{2,i}$ to trace this ray through the prism (see Fig. 3).

The forward pinhole model is usually formulated as the combination of basic transformations:

$$\mathbf{p} = P \circ E(\mathbf{x}_\mathrm{w}) = A \circ D \circ F \circ E(\mathbf{x}_\mathrm{w}), \tag{1}$$

where E corresponds to Euclidean mapping $\mathbf{x} = E(\mathbf{x}_\mathrm{w}) = \mathbf{R}\mathbf{x}_\mathrm{w} + \mathbf{t}$, translating 3-D point $\mathbf{x}_\mathrm{w}$ from the world coordinate system (WCS) to the coordinate system (CS) of the camera ($\mathbf{R}$ is the rotation matrix and $\mathbf{t}$ is the translation vector), F is the central projection onto the unit plane $\mathbf{x}' = F(\mathbf{x})$, D is the 2-D distortion transformation $\mathbf{x}'' = D(\mathbf{x}')$ and A is the 2-D affine transformation to the image coordinates in pixels $\mathbf{p} = A(\mathbf{x}'')$.[4,17] Here, we need the inverse pinhole model in the form:

$$\mathbf{l}_{0,i} = P(\mathbf{p}_i) = F^{-1} \circ D^{-1} \circ A^{-1}(\mathbf{p}_i), \tag{2}$$

with the following basic transformations for $\mathbf{p}_i = (u, v)^T$:

$$\mathbf{x}'' = A^{-1}(\mathbf{p}_i)\colon \begin{pmatrix} x'' \\ y'' \end{pmatrix} = \left(\frac{u-u_0}{f_\mathrm{u}}, \frac{v-v_0}{a_\mathrm{f} f_\mathrm{u}}\right)^T, \tag{3}$$

$$\mathbf{x}' = D^{-1}(\mathbf{x}'')\colon \begin{pmatrix} x' \\ y' \end{pmatrix} = (1 + \rho_1 r^2 + \rho_2 r^4 + \rho_3 r^6) \begin{pmatrix} x'' \\ y'' \end{pmatrix},$$
$$r^2 = x''^2 + y''^2, \tag{4}$$

$$\mathbf{l}_{0,i} = F^{-1}(\mathbf{x}')\colon \mathbf{l}_{0,i} = ((0{,}0{,}0), (x', y', 1)/|(x', y', 1)|)^T. \tag{5}$$

The set of intrinsic camera parameters f_u, k_f, u_0, v_0 and the distortion parameters ρ_1, ρ_2, ρ_3 are used to describe the inverse pinhole model. Since the inverse transformation P^{-1} depends only on the parameters of the image sensor and the main lens, it is the same for all image parts. As a result, the transformation for the ray tracing model can be represented in the form:

$$\mathbf{l}_{\mathrm{w}\,i} = E(\mathbf{l}_{2,i}) = E^{-1} \circ S_{2,i} \circ S_{1,i} \circ F^{-1} \circ D^{-1} \circ A^{-1}(\mathbf{p}_i), \tag{6}$$

where Euclidean mapping E is the same for all image parts due to our choice of the same CS for all parts of the prism. We further consider the camera CS and the WCS to be coincident, so E is omitted. Hence, the direction of the z_w axis is along the optical axis of the main lens.

In contrast to the pinhole model, the ray tracing model cannot provide closed-form solution for the forward transformation $\mathbf{p}_i = P_i \circ E_i(\mathbf{x}_\mathrm{w})$ because it requires initially unknown direction vector of the line $\mathbf{l}_{\mathrm{w}\,i}$ from 3-D point $\mathbf{x}_\mathrm{w}$. This problem is usually solved by the iterative technique called ray aiming or by look-up-table interpolation. We can formally use the notation for the forward transformation if we consider that this notation stands for the iterative solver or the interpolation technique.

We assume that the prism is located in the air and use the same description as in the paper.[10] Hence, we need nine parameters for three faces of the prism and one more parameter for the refractive index n. The complete vector of parameters $\mathbf{k}$ includes 4 intrinsic camera parameters, 3 distortion parameters, and 10 prism parameters. From the theory of optics, we can conclude that the focal distance f_u and the refractive index n should significantly vary for different wavelengths, the coordinates of principal point u_0, v_0, and the distortion parameters ρ_1, ρ_2, and ρ_3 should slightly vary (actually, this depends on peculiarities of the chromatic aberration correction for particular optical system design), but the ratio of focal lengths a_f and 9 parameters for the location of prism faces should not change. Consequently, vector $\mathbf{k}$ should be divided into two parts: the first one is common for all spectral bands and the second one is different. The parameters of the second part should either be calibrated separately for every spectral band or represented as a function of wavelength.

2.2 *Calculation of 3-D Point Coordinates*

After point matching has been finished, the reconstruction of 3-D point coordinates $\mathbf{x}_\mathrm{w}$ from N corresponding image points $\mathbf{p}_i$ may be considered as the optimization problem:

$$\hat{\mathbf{x}}_\mathrm{w} = \arg\min_{\mathbf{x}_\mathrm{w}}(C(\mathbf{x}_\mathrm{w}, \mathbf{p}, \mathbf{k})), \tag{7}$$

where $\mathbf{p} = (\mathbf{p}_1^T, \mathbf{p}_2^T, \dots, \mathbf{p}_N^T)^T$ is the composite vector of image coordinates. The selection of the appropriate cost function is not trivial in the general case, because it depends on a priori data about the position of target point $\mathbf{x}_\mathrm{w}$, its projections $\mathbf{p}_i$, the statistics of coordinate measurement errors, and the properties of camera transformations P_i.

For simplicity, we assume that a priori information is not available and assign equal weights to all points and coordinates. Since the ray tracing model provides closed-form solution for the inverse projection, we use the cost function based on the sum of distances d from the point $\mathbf{x}_\mathrm{w}$ to rays $\mathbf{l}_{\mathrm{w}\,i}$:

$$C(\mathbf{x}_\mathrm{w}, \mathbf{p}, \mathbf{k}) = \sum_{i=1}^{N} d^2(\mathbf{x}_\mathrm{w}, \mathbf{l}_{\mathrm{w}\,i}) = \sum_{i=1}^{N} \mathbf{b}_i{}^T \mathbf{b}_i,$$

$$\mathbf{b}_i = (\mathbf{Id}_{3\times 3} - \mathbf{v}_{\mathrm{w}\,i}\mathbf{v}_{\mathrm{w}\,i}{}^T)(\mathbf{c}_{\mathrm{w}\,i} - \mathbf{x}_\mathrm{w}), \tag{8}$$

where $\mathbf{Id}_{3\times 3}$ is 3×3 identity matrix.

In this case, the solution of the minimization problem formulated in Eq. (7) can be found by the direct least-squares method.[4] The method for $N = 2$ is reduced to determining the midpoint of the common perpendicular and does not require iterations.

In most applications of 3-D measurement endoscopic systems, the calculated 3-D point coordinates are necessary to measure the geometric parameters, such as segment length, point-to-line distance, point-to-plane distance, surface area, etc.[1] Details on equations used for geometric measurements and uncertainty estimation may be found in the book.[18]

2.3 *Image Rectification*

Conventional rectification methods for pinhole camera models use the assumption that each camera should be rotated around the center of projection. Calculation of the rectification transformation parameters for i'th camera can be done by determining the rotation transformation E_i and the projective transformation P_i for the new (virtual) camera.[4,19,20] Unlike the pinhole model, the ray tracing one has no single center of projection,[10] which makes pinhole rectification techniques inapplicable. The proposed method for the ray tracing model includes the following stages.

I. Applying the inverse ray tracing [Eq. (6)] for point grid on both halves of the image. The maximal convergence points of rays $\mathbf{l}_{\mathrm{w}\,i}$ are considered to be the projection centers O_1 and O_2 of virtual cameras (Fig. 4).

II. Rotation of the x_i axis of each virtual camera until it becomes parallel to the line O_1O_2. The direction of the y_i axis has to be defined from the condition of perpendicularity to the axes x_i and z_w. The direction of the z_i axis is assigned perpendicular to the axes x_i and y_i. After the algorithm calculates the orientation and the position of virtual cameras' CS, it is possible to determine the parameters of the Euclidean transformation E_i. These transformations translate 3-D point $\mathbf{x}_\mathrm{w}$ from the world CS to the CS of i'th virtual camera: $\mathbf{x}_i = E_i(\mathbf{x}_\mathrm{w}) = \mathbf{R}_i\mathbf{x}_\mathrm{w} + \mathbf{t}_i$.

III. Calculation of the parameters of the projective transformations P_1 and P_2 corresponding to the virtual stereopair, which define the coordinates of rectified points. In the first approximation, these transformations are considered to be equal to the transformation P of the real camera, but without the polynomial distortion approximation. P_1 and P_2 need to be adjusted in order to maximize the useful area on rectified images and to make the scale of these images close to the initial one.

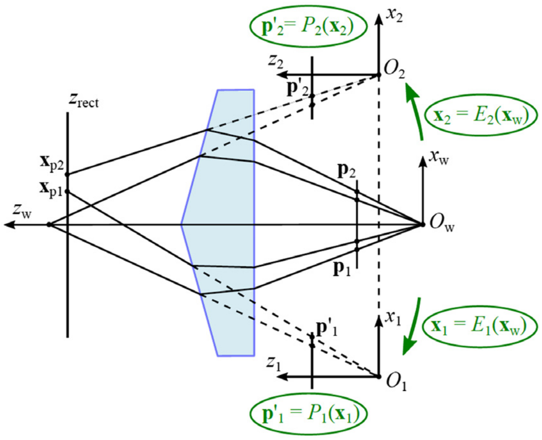


Fig. 4 Determination of virtual cameras for image rectification using the ray tracing camera model.

Since rays $\mathbf{l}_{\mathrm{w}\,i}$ have no single intersection point, we should specify a position of the rectification plane in order to define points $\mathbf{x}_{\mathrm{p}1}$ and $\mathbf{x}_{\mathrm{p}2}$ used for the projection. This plane is assumed to be parallel to the $x_i\ O_i\ y_i$ planes of both virtual cameras. After the rectification transformation is defined, one can calculate the coordinates of the point $\mathbf{p}_i'$ on the rectified image for every point $\mathbf{p}_i$ on i'th half of the initial image. First, the inverse ray tracing [Eq. (6)] is applied for every point $\mathbf{p}_i$. Next, the point $\mathbf{x}_{\mathrm{p}i}$ is calculated by intersecting the refracted ray $\mathbf{l}_{\mathrm{w}\,i}$ with the rectification plane. Finally, this point is projected to the point $\mathbf{p}_i'$ on the virtual camera's image plane by applying transformations E_i and P_i. Similarly, we can find the inverse rectification transformation from $\mathbf{p}_i'$ to $\mathbf{p}_i$ and use it to calculate look-up-tables for image interpolation as in conventional pinhole rectification technique.

In order to achieve higher accuracy of rectification, the plane must be placed as close to the real object position as possible. The impact of the distance to the rectification plane needs additional analysis and is outside of the scope of this paper. Based on multiple experiments, we can conclude that one rectification distance is usually enough to cover the working range typical for prism-based stereoscopic systems.

We should notice that the proposed image rectification technique uses the vector of parameters $\mathbf{k}$, which may vary for each spectral band. Hence, the impact of this factor on the accuracy of image rectification should also be estimated as a part of our analysis of WB and NB calibration applicability in this paper.

3 Calibration

3.1 *Standard Calibration Procedure*

The aim of the calibration procedure is to determine the parameter vector $\mathbf{k}$ using captured images of a calibration target. These targets may be manufactured as flat, step-like, or cube objects with many contrast markers, with centers that can be easily determined on the image (chessboard pattern, circles, line grid, or other geometric patterns). We consider that the calibration target has M points and 3-D coordinates of the points $\mathbf{x}_{\mathrm{t}j}$, $j = 1,2\ldots M$ are known with a certain accuracy in the CS of the calibration target. During the image capturing, the target is placed at R different positions to acquire images. Next, the image coordinates $\mathbf{p}_{i,j,k}$, $k = 1,2\ldots R$ are calculated for each point and each position using image processing. We should notice that for i'th camera (i.e., for i'th image part) and k'th position, only a part of M points will be visible, so the following description assumes that the point may be detected in the image and its image coordinates may be calculated. The complete projection transformation for these points can be written as $\mathbf{p}_{i,j,k} = P_i \circ E_i \circ E'_k(\mathbf{x}_{\mathrm{t}j})$, where E'_k stands for the Euclidean mapping from the CS of the calibration target to the WCS. In addition to the parameter vector $\mathbf{k}$ for transformations P_i and E_i, we also need to find vector $\mathbf{k}_{\mathrm{t}}$ describing transformations E'_k for each k. If we introduce the composite vector $\mathbf{x}_{\mathrm{t}}$ of all 3-D points and the composite vector $\mathbf{p}$ of all projections, we can write the optimization problem for calibration as follows:

$$\hat{\mathbf{k}}, \hat{\mathbf{k}}_{\mathrm{t}} = \arg \min_{\mathbf{k},\mathbf{k}_{\mathrm{t}}} (C(\mathbf{x}_{\mathrm{t}}, \mathbf{p}, \mathbf{k}, \mathbf{k}_{\mathrm{t}})). \tag{9}$$

The cost function for the ray tracing model should use the sum of squared distances from 3-D points to backprojected rays: $d^2(\mathbf{x}_{\mathrm{t}j}, \hat{\mathbf{l}}_{\mathrm{t}i,j,k})$, where $\hat{\mathbf{l}}_{\mathrm{t}i,j,k} = \hat{E}_k'^{-1}(\hat{\mathbf{l}}_{\mathrm{w}\,i,j,k})$, $\hat{\mathbf{l}}_{\mathrm{w}\,i,j,k}$ is calculated from $\mathbf{p}_{i,j,k}$ using Eq. (6) with current estimation of $\hat{\mathbf{k}}$ and $\hat{E}_k'^{-1}$ is described by current estimation of $\hat{\mathbf{k}}_{\mathrm{t}}$. If we use a flat calibration target, we can replace rays $\hat{\mathbf{l}}_{\mathrm{t}\,i,j,k}$ by points $\hat{\mathbf{x}}_{\mathrm{t}\,i,j,k}$ of ray intersections with XOY plane in the CS of the target and write the cost function as follows:

$$C(\mathbf{x}_{\mathrm{t}}, \mathbf{p}, \hat{\mathbf{k}}, \hat{\mathbf{k}}_{\mathrm{t}}) = \sum_{i,j,k} (\mathbf{x}_{\mathrm{t}j} - \hat{\mathbf{x}}_{\mathrm{t}\,i,j,k})^T \mathbf{\Sigma}_{\mathrm{x}\,i,j,k}^{-1} (\mathbf{x}_{\mathrm{t}\,j} - \hat{\mathbf{x}}_{\mathrm{t}\,i,j,k}), \tag{10}$$

where $\mathbf{\Sigma}_{\mathrm{x}\,i,j,k}^{-1}$ is the inverse covariance matrix of calibration target coordinate measurement error for $\mathbf{x}_{\mathrm{t}\,i,j,k}$. The sum over all i, j, k actually means the sum over all points, which image coordinates $\mathbf{p}_{i,j,k}$ were calculated. Due to the lack of a priori information about the distribution of image coordinate measurement error, the covariance matrix $\mathbf{\Sigma}_{\mathrm{x}\,i,j,k}$ is usually either eliminated from Eq. (10) (considered to be the identity matrix) or replaced by the weighting coefficient $1/z^2_{\mathrm{w}\,i,j,k}$ for simplicity. In this paper, we use the Levenberg–Marquardt method[4] to solve the minimization problem Eq. (9). In order to improve stability and avoid local minimum, one may apply global search optimization techniques, such as particle swarm optimization and genetic algorithms.[21,22]

3.2 *Modification for Multiple Spectral Bands*

Here, we consider that the calibration problem is solved for multiple spectral bands simultaneously. Hence, we acquire images of the calibration target in L spectral bands and calculate the image coordinates $\mathbf{p}_{i,j,k,l}$ for each band $l = 1,2\ldots L$. If we assume that the vector of calibration parameters $\mathbf{k}_l$ is unique for each band, the composite vector $\mathbf{k}$ should be defined as $\mathbf{k} = (\mathbf{k}_1^T, \mathbf{k}_2^T, \ldots, \mathbf{k}_L^T)^T$. According to our analysis above, vector $\mathbf{k}_l$ should be divided into two parts: $\mathbf{k}_{\mathrm{c}}$ is common for all spectral bands and $\mathbf{k}_{\mathrm{v}\,l}$ is different. Then, we define the composite vector $\mathbf{k}$ as $\mathbf{k} = (\mathbf{k}_{\mathrm{c}}^T, \mathbf{k}_{\mathrm{v}1}^T, \mathbf{k}_{\mathrm{v}2}^T, \ldots, \mathbf{k}_{\mathrm{v}\,L}^T)^T$. Depending on image capturing procedure, number of positions R can also be different for each band. If we ensure that the calibration target does not move when switching spectral bands, we can use single vector $\mathbf{k}_{\mathrm{t}}$ to describe transformations E'_k. Otherwise, we need vector $\mathbf{k}_{\mathrm{t}\,l}$ to describe R_L transformations $E'_{k,l}$ for l'th spectral band and the composite vector $\mathbf{k}_{\mathrm{t}} = (\mathbf{k}_{\mathrm{t}\,1}^T, \mathbf{k}_{\mathrm{t}\,2}^T, \ldots, \mathbf{k}_{\mathrm{t}\,L}^T)^T$. Using these notations, we can write the optimization problem as in Eq. (9). The cost function can also be formulated similar to Eq. (10) replacing indices i, j, k by i, j, k, l.

In this work, we assume that $\mathbf{k}_{\mathrm{v}l}$ includes two variables: the focal distance f_{u} and the refractive index n. All other parameters of the ray tracing model are considered to be the same for all spectral bands, so they are included in $\mathbf{k}_{\mathrm{c}}$. Our equipment allows us to capture images in all spectral bands for each position of calibration target, which makes it possible to use single vector $\mathbf{k}_{\mathrm{t}}$ for all bands. If the number of spectral bands is high, it seems impractical to capture images in all these bands for calibration. In this case, we should select small number of bands for calibration and interpolate parameters $\mathbf{k}_{\mathrm{v}l}$ by wavelength for other bands. The choice of this number and the type of interpolation is determined by the assumed dependence of the parameter on the wavelength. Since the focal distance f_{u} and the refractive index n are usually smooth continuous functions of wavelength, we suppose that three properly chosen bands will be sufficient for calibration in the visible range.

4 Experimental Verification

4.1 *Experimental Setup*

To implement a multispectral calibration, we have assembled a setup shown in Fig. 5. We used a self-made video endoscopic probe with prism-based stereoscopic system, which has field of view 40 deg × 45 deg in each channel and range of working distances 5 to 40 mm. Video camera at the distal end of the probe has 1/6" color image sensor with resolution 1920 × 1080 pixels. The probe was fixed on the stage. For illumination of the calibration target, we used a white-light LED. For spectral selection, we inserted a filter wheel between the target and the LED. The wheel contained 6 NB filters with center wavelengths 460 nm, 503 nm, 550 nm, 600 nm, 650 nm, 704 nm, and a FWHM 10 nm and an empty aperture for WB image acquisition. Figure 6 shows the spectral dependency of LED relative power distribution and NB filters transmission curves.

To acquire images for calibration and tests, we utilized a plane calibration target with black and white chessboard pattern. Due to sufficient range of distances and magnifications, three samples of calibration target with 0.5 mm (small-sized), 1 mm (middle-sized), and 2 mm (large-sized) chessboard square size have been manufactured. Each target had a grid of 25 × 25 squares. A chessboard pattern was produced by chrome etching on glass with inaccuracy about 1 μm.

Without filters, i.e., in WB mode, and with each NB filter, we have captured two image series: calibration sequence and test sequence. The calibration sequence included images

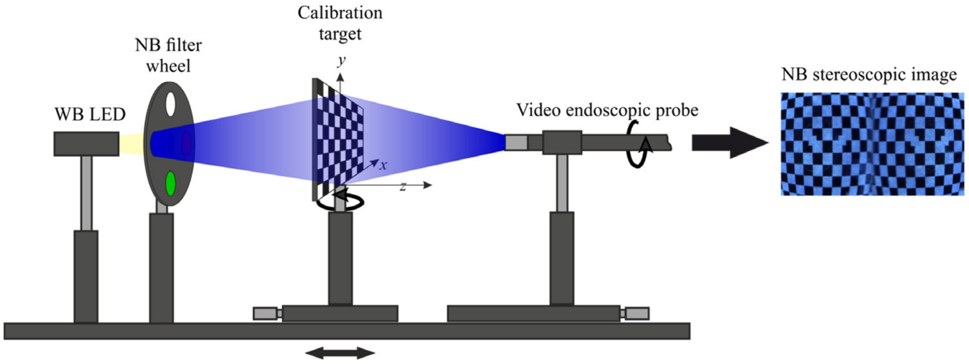


Fig. 5 Experimental setup for WB and NB calibration of the stereoscopic endoscopic probe.

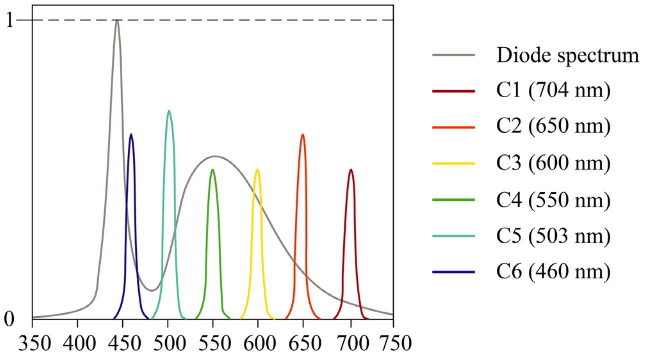


Fig. 6 LED spectral power distribution and transmission curves of utilized NB filters.

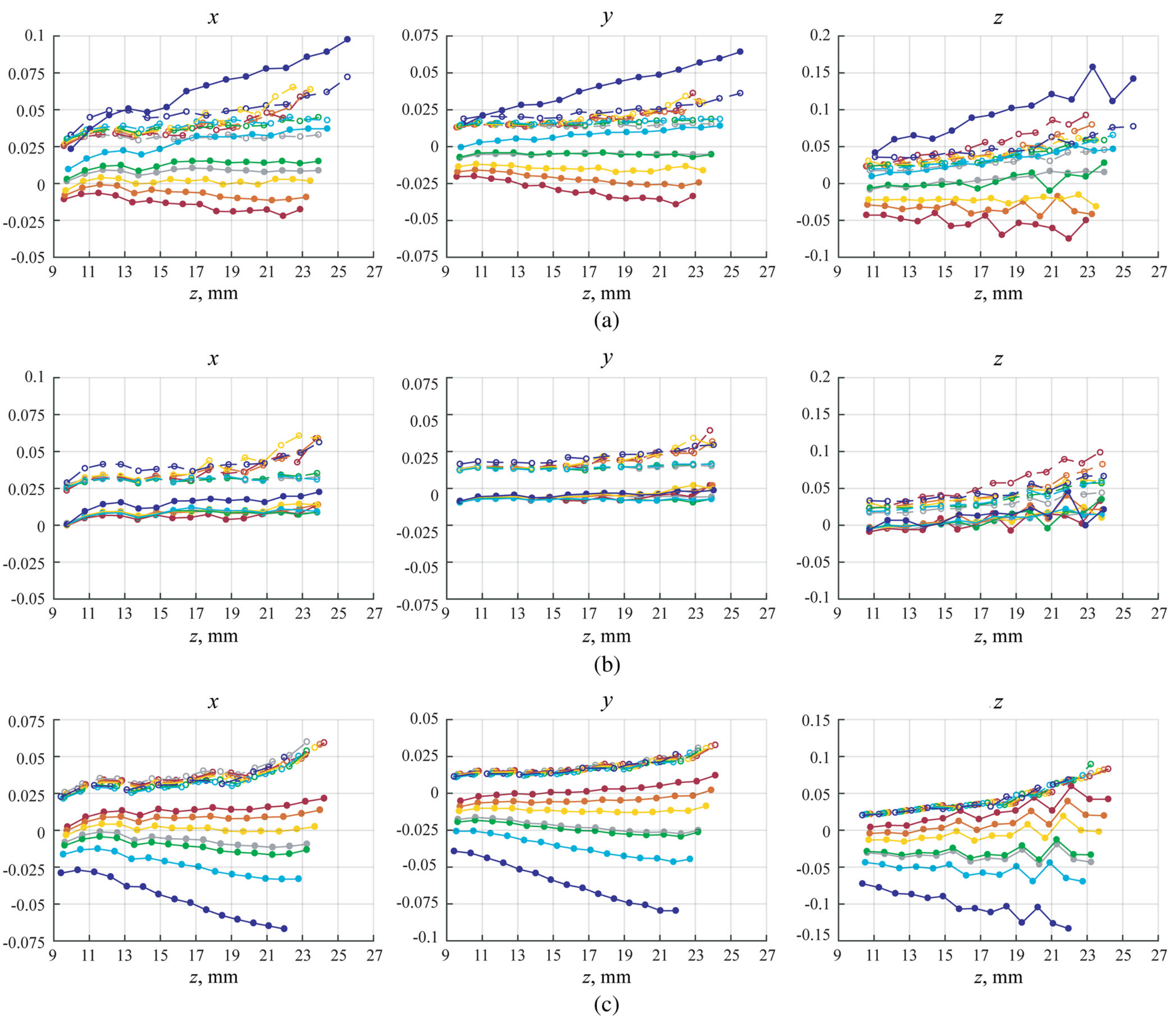
captured at different positions of small-sized, medium-sized, and large-sized calibration targets covering the total range of distances from 8 to 32 mm. Each of these targets has been placed approximately perpendicular to z-axis of the stand at the distance, which allows to see at least eight markers in a horizontal row on both image halves. Then, the target was consequently rotated by 30 deg and 45 deg around x and y axes and by 45 deg around z axis to capture images at oblique angles. Finally, we have positioned the target perpendicular to z-axis at approximately 1.5× longer working distance than the initial position. As a result, we have acquired seven images (corresponding to WB light and 6 NB filters) for each of 32 positions of calibration targets.

Test sequences were collected for small-sized and medium-sized calibration targets in the following way. Again, the calibration target has been placed approximately perpendicular to z-axis of the stand at a reasonably close distance and oriented so that horizontal rows of markers were approximately horizontal on the image. We used translation stage to shift the calibration target along z-axis and captured images with 1 mm step. Thus, we have obtained series containing 6 images for small-size target and 15 images for medium-sized one in the range from 10 to 24 mm for WB light and each NB filter. The range of distances was limited by the depth of field of the optical system and the condition of minimal chessboard square scale suitable for robust image processing.

4.2 *3-D Geometrical Measurements*

After obtaining WB and all NB calibration images, we processed them to calculate image coordinates for each chessboard node. The developed software for image processing allows a user to set the initial correspondence between image points and points on calibration target using two specific chessboard nodes (see the image of the calibration target in the right part of Fig. 5) and further to refine image coordinates using the technique described in the paper.[23] All automatically processed images of calibration and test sequences have been checked manually afterward to delete points in low-illuminated areas, around glare, etc. Since the matching between image points and points on calibration target has been done for both image parts, the stereo correspondence for left and right half-images was obtained automatically. Using these coordinates and the calibration technique for the ray tracing camera model,[10] we obtained the set of calibration parameters for WB image series as well as for each of NB series independently. The images of the test sequences have been used to calculate the 3-D coordinates for each node applying the ray tracing model with different calibration sets. To characterize and compare the measurement accuracy achieved using different camera models and calibration procedures, a few criteria may be used. The straightforward approach is to carry out multiple measurements in each point and calculate mean value and standard deviation (STD). In this case, mean value characterizes the systematic measurement error caused by nonideal calibration and some random factors that vary from point to point, but do not change when several measurements are made at the same point, for example, the errors of test-object manufacturing. STD characterizes only random error caused by noise in the image and varies from point to point due to illumination nonuniformity, vignetting and aberrations. It is time-consuming and challenging to show and analyze these 3-D distributions within the whole field of view and for several spectral bands. Therefore, in the field of endoscopic measurements and other imaging applications, accuracy is normally characterized by the mean value and STD calculated over the field of view at the distance z (i.e., over all points of calibration target at each position). These values depend only on the distance z to calibration target and are easy to illustrate.

We chose orientation of the calibration targets so that the grid lines were approximately parallel to x- and y-axis of the stand. Thus, we used the distance between chessboard nodes as x- and y-segments. Points to measure the segment along z axis were taken either from the previous or from the next image of the series, i.e., when the test chart is shifted by 1 mm. We should note that the measured segments are not exactly aligned along x, y, and z axes in any CS, but



(a)

(b)

(c)

Fig. 7 Dependence of mean value (solid, in mm) and STD (dotted, in mm) of difference between reference and measured lengths of the 1-mm segment along *x* (left), *y* (center), and *z* (right) axes on the distance to the target: (a) if WB calibration is applied to NB measurements, (b) if spectral ranges of calibration and measurements coincide, and (c) if WB and NB calibrations are applied to the measurement at 650 nm. Colors correspond to the same spectral ranges, as shown in Fig. 6.

we consider this data to be sufficient to analyze measurement uncertainty for different orientation of segments.

Figure 7 shows the mean value and STD of difference between reference and measured lengths of the 1 mm segment along *x*, *y*, and *z* axes. If WB calibration data are applied to NB images from the test sequence, the measurement bias increases from the middle of the working spectral range (550 nm) to its edges (460 and 704 nm) and can reach significant values: up to 0.1 mm along *x* axis and 0.15 mm along *z* axis [Fig. 7(a)], which is more than three times larger than STD for the WB calibration applied to WB measurements (about 0.03 mm along *x* axis and up to 0.05 mm along *z* axis). Applying the WB calibration data to NB images in the green part of the spectrum, where the sensor sensitivity and LED intensity are maximal, always provides almost the same mean value and STD as for WB images. As expected, the best results are achieved, when the spectral ranges of calibration data and test images are the same [Fig. 7(b)]. In this case, mean value error does not exceed 0.025 mm along *x* axis, 0.01 mm along *y* axis, and 0.05 mm along *z* axis, so it is about three times smaller than corresponding STD. One more interesting example is applying all the obtained calibration packs to a particular NB image series. We chose NB filter C2 (650 nm) for illustration [see Fig. 7(c)]. One can see that the greater the distance between the spectral range used for calibration and the band used for the measurement, the larger the measurement bias. At the same time, the values of STD are almost independent of the choice of spectral band for calibration.

Actually, the mean values shown in Fig. 7 correspond to systematic error of 3-D coordinates mainly caused by uniform or nonuniform scaling. The effect of scaling is clearly seen from Figs. 7(a) and 7(c) as the movement of data points along the z axis of the plot. In contrast, STD values are affected by both random errors originated from errors of image coordinates and systematic errors caused by more complex distortion of 3-D coordinates than scaling (such as adding curvature to originally flat objects). According to Fig. 7(c), applying different calibration parameters to one NB image set does not lead to significant variation of STD values, hence, the impact of systematic error on STD is rather small. Slight variations of STD values are mainly the result of the overall scaling. Hence, the STD values actually represent random errors. The increase with distance is typical for stereoscopic measurements.[4]

As can be seen from Fig. 7(b), the STD values for independent NB calibrations are quite different. The main reason is the Bayer color mask, which actually leads to two times higher image resolution for C4 and C5 filters (using G image channel) than for C1, C2, C3 filters (using R) and C6 (using B). Additionally, the combination of white LED spectrum, NB filter transmission curve, and camera spectral sensitivity provides lower signal-to-noise ratios for C1 and C6. As a result, the STD of random errors for C4 and C5 is smaller in comparison to other bands.

Even performing independent calibrations for all NB channels does not allow to exclude bias completely [Fig. 7(b)]. Small residual values are probably induced by the limitation of the applied camera model, which does not separately consider the pupil aberration of the main lens. This is the subject of a future research.

We have applied the calibration technique for all six NBs and WB jointly, as proposed in Sec. 3.2 (denoted further as *All*). Since this approach requires too many images, it is barely applicable in practice, but it allows to estimate the spectral dependencies of the calculated values f_u and n and find the optimal type of interpolation (Fig. 8). Possible candidates are linear interpolation using values for C1 and C6 *C1C6* or C2 and C5 *C2C5* and Conrady's equation[24] for interpolation using three values: C1, C4, and C6 *C1C4C6*. As shown in Fig. 8, the latter one provides a very close approximation.

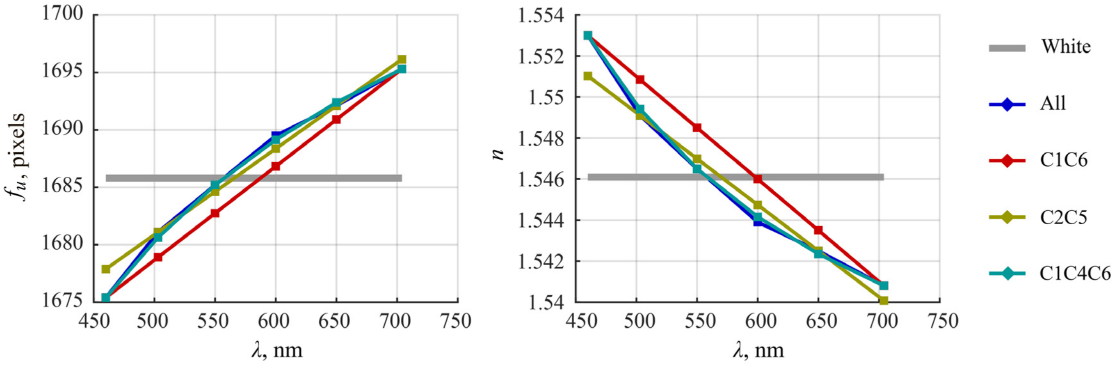
For practical testing, we utilized two series for C1 and C6 filters for joint calibration and calculated f_u and n for other filters using linear interpolation *C1C6*. Next, we repeated this method for C2 and C5 filters *C2C5* and three series for C1, C4, and C6 filters using Conrady's equation *C1C4C6*.

To estimate the effectiveness of these approaches, we compared them with the calibration in WB *White* and the case of independent calibration for all spectral ranges *Own*. Figure 9 illustrates the results for the NB spectral range C2 (650 nm). One can see that the measurement bias may be radically decreased by simultaneous calibration. The results for *C1C4C6* and *All* almost coincide with the independent calibration *Own*. Since the parameters for this spectral band (C2) were not interpolated for *C2C5*, the results for *C2C5* are also the same as for *All*. The STD values are mainly caused by random errors and are not significantly affected by the choice of calibration technique.

In order to choose the optimal calibration technique, we summed the absolute values of the mean values across all spectral bands and considered the resulting values as the overall "score" for accuracy assessment focused on measurement bias (Fig. 10). Since the score for *C1C4C6* matches the score for the joint calibration using all series *All* and the independent calibration *Own*, we can conclude that three NBs located in the middle and on the edges of the working spectral range are enough to provide the desired measurement accuracy. The linear interpolation using two nonedge NBs *C2C5* is slightly worse but it still significantly increases the accuracy compared to WB calibration.

4.3 *Image Rectification and 3-D Shape Reconstruction*

Since the proposed image rectification technique uses the vector of calibration parameters, it may also be affected by the choice of the calibration technique. In order to evaluate the rectification accuracy, we assessed the vertical mismatch Δy for each pair of corresponding points in rectified images. We introduced the criterion $R = A/B$, where A is the number of pixels with Δy less than $\Delta y_{\max}$ pixels, B is the total number of pixels having the corresponding points.



(a) (b)

Fig. 8 Spectral dependencies of the calculated values (a) f_u and (b) n. Curves for *All* and *C1C4C6* almost coincide.

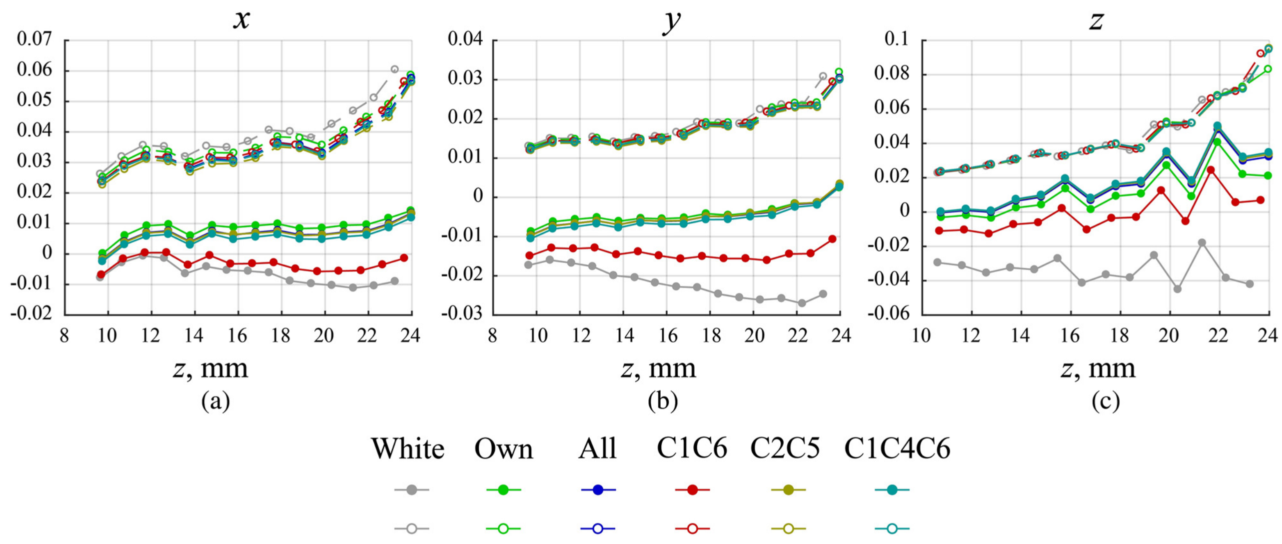


Fig. 9 Dependence of mean value (solid) and STD (dotted) of difference between reference and measured lengths of the 1-mm segment along (a) x, (b) y, and (c) z axes on the distance to the target, if various types of calibration data interpolation are applied.

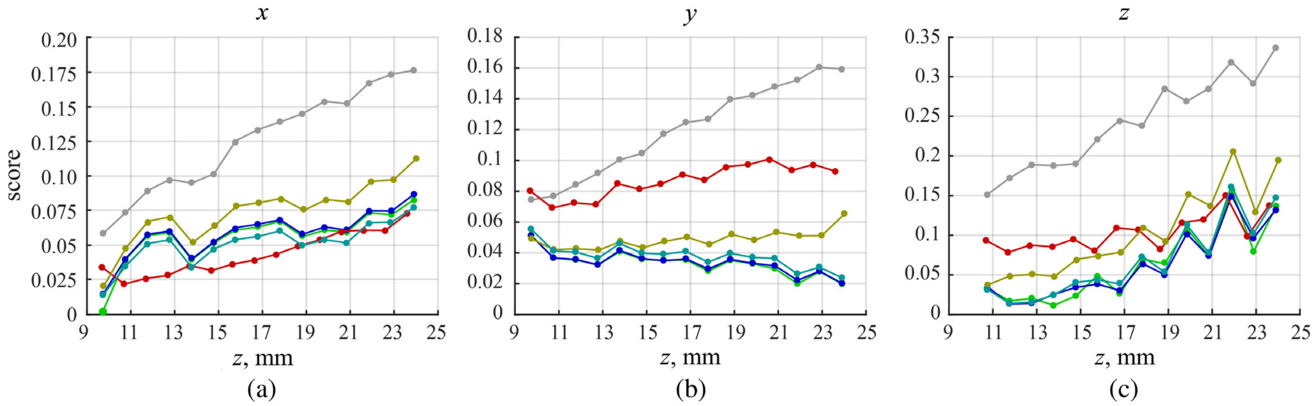


Fig. 10 Sum of absolute values of mean values across all spectral bands. Colors correspond to the same calibration sets, as shown in Fig. 9. Types of measurements (a, b, and c) are the same as in Fig. 9.

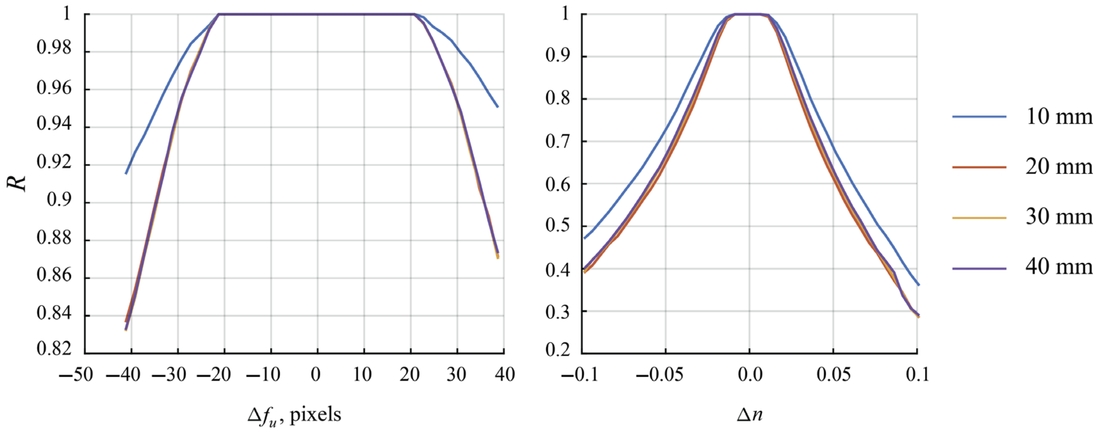
Proportional to the radius of the point spread function, the value of $\Delta y_{\max}$ was chosen equal to 4 pixels. We assume that this value is the maximum tolerable one for point matching based on the experimental results with images of various real objects.

To analyze the sensitivity of the rectification to the deviation of f_u and n, we have obtained the dependencies of R on these parameters using the computer simulation. For this purpose, we have used the ray tracing model with *White* set of parameters to acquire the synthetic image of the point grid located at various distances $Z_{\text{obj}} = Z_{\text{rect}}$ (10, 20, 30, and 40 mm). Next, we acquired synthetic images with deviated values of f_u and n, rectified these images using the original *White* set of parameters and estimated Δy and R (see Fig. 11). We have chosen *White* set because it corresponds approximately to the center of the working spectral band. As shown in Fig. 8, f_u varies in the range ±10 pixels and n varies in the range ±0.006 across the wavelength band 450 to 750 nm. Figure 11 demonstrates that under these variations, R is stable and equal to 1. Hence, we can predict that the usage of WB calibration should not significantly lower image rectification accuracy.

To compare the effectiveness of WB and NB calibration for the rectification of real images, we have used the test image sequence of the calibration targets and calculated the vertical deviation (Δy) maps of the chessboard nodes (Fig. 12). One can see that the difference between WB calibration *White* and independent calibrations for each NB *Own* does not exceed 1 pixel.

Finally, we have calculated the 3-D structure of the metal test specimen at 460 nm, 503 nm and 704 nm using WB and NB calibration to show the influence of rectification on the quality of the reconstructed 3-D surfaces (Fig. 13). We have used normalized cross-correlation[5,8] for stereo matching on rectified images. Distance to the object was 12.5 mm.

Figure 13 indicates that WB calibration almost does not lead to the rectification errors and, therefore, does not influence the accuracy of point matching. However, it causes



(a) (b)

Fig. 11 Calculated dependence of parameter R on (a) f_u and (b) n.

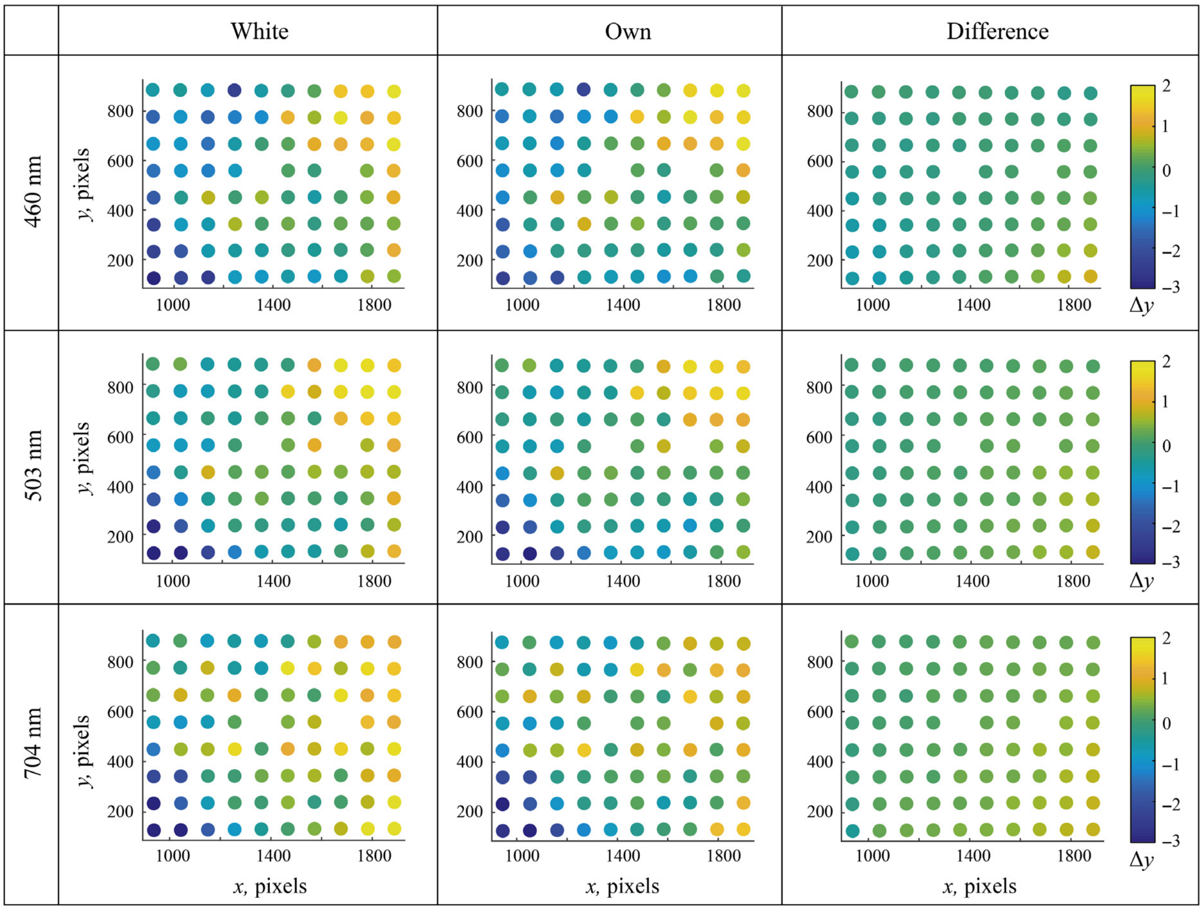


Fig. 12 The values of vertical deviation Δy (in pixels) versus pixel coordinates on the right half-image. Dot maps represent experimental data for the cases of different wavelengths (460, 503, and 704 nm). Δy values, obtained using WB calibration (*White*, left), were subtracted from the ones, obtained with NB ones (*Own*, center), resulting the difference distribution map (right).

significant geometrical distortion of the reconstructed 3-D shape, which corresponds well to the results of Sec. 4.2. The bottom images in Fig. 13 show that under WB calibration, 3-D models obtained at different wavelengths have different depth and scale. The results for different variants of joint calibration (*All*, *C1C6*, *C2C5*, *C1C4C6*) confirm that the impact of the calibration technique on the quality of rectification is insignificant.

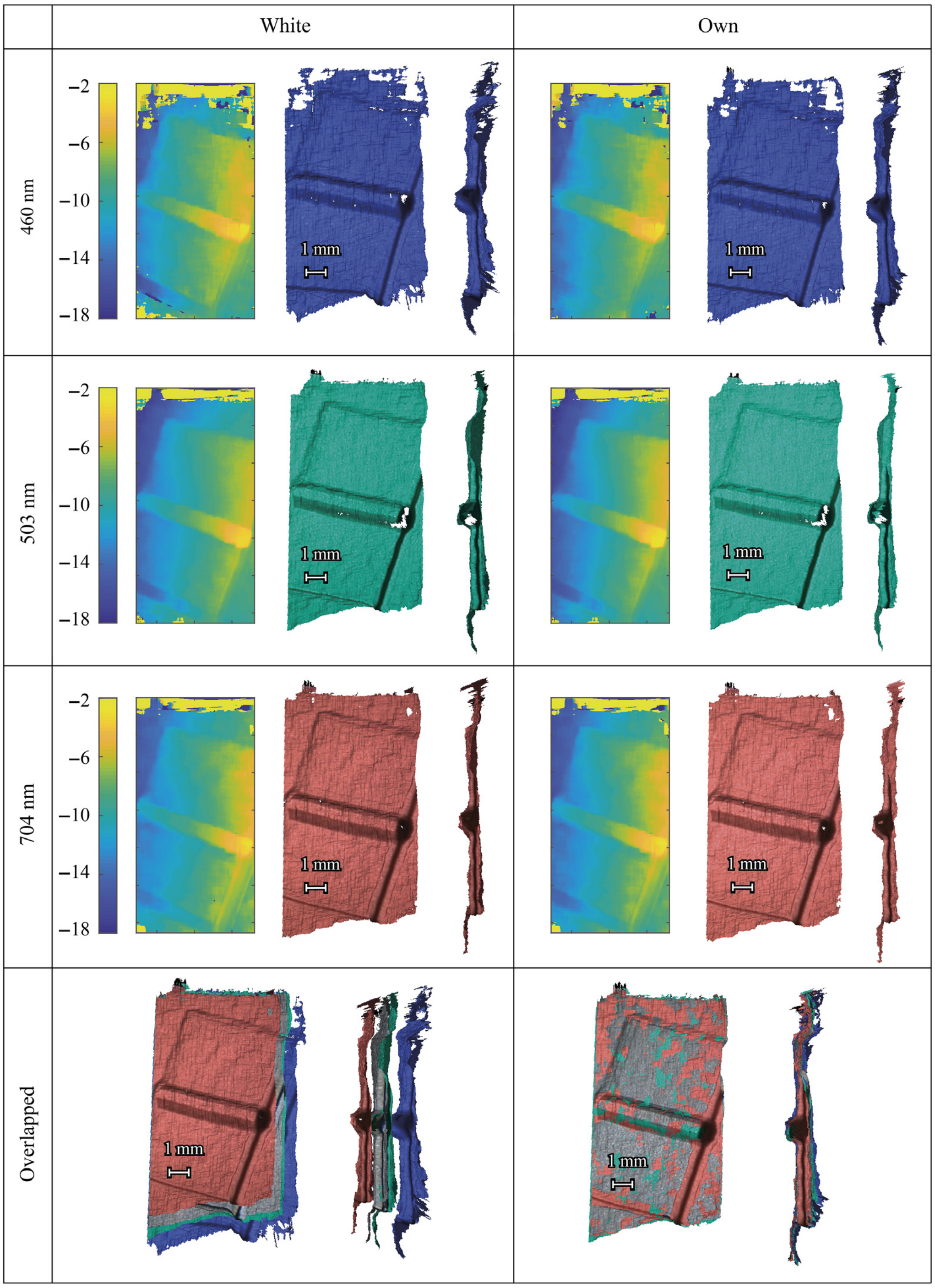


Fig. 13 Disparity maps and reconstructed 3-D models for the metal measurement target obtained using WB (*White*, left) and NB (*Own*, right) calibrations.

5 Conclusion

NB imaging provides a reduction of chromatic image aberrations introduced by a prism-based optical system and, therefore, better accuracy of stereo matching. The experimental results have confirmed the theoretical assumption that applying WB calibration parameters is not optimal if NB images are used for the measurements. Using WB calibration for NB images leads to significant bias and has almost no effect on STD across field of view. For accurate quantitative characterization of the inspected objects and 3-D shape reconstruction, it is necessary to calibrate the system either in a certain NB, where contrast visualization is achieved, or in a few NBs with further interpolation of calibration parameters for operation in any selected spectral band. We have shown that for measurements in any NB within the whole visible range with the accuracy comparable to the accuracy of independent NB calibrations, it is enough to calibrate the system in three NBs corresponding to its center (in the green part of the spectrum) and edges (in the blue and red parts of the spectrum).

Applying WB calibration to NB images does not lead to significant rectification errors and, therefore, does not lower the accuracy of point matching in comparison with the accuracy obtained using NB calibration.

The results of our experiments may be useful for the development of RVI tools for spectral detection and quantitative characterization of low-contrast objects with increased sensitivity and measurement accuracy.

Acknowledgments

The Russian Science Foundation (project #17-19-01355) financially supported the work. In the part of rectification algorithms analysis, the research is supported by Russian Fund for Basic Research (project # 17-29-03469).

References

1. F. Lafleur, "Videoscope trends: improvements and new developments," https://www.qualitymag.com/articles/93455-videoscope-trends-improvements-and-new-developments.
2. E. Hubben and D. Jervis, "Advances in three dimensional measurement in remote visual inspection," in *Proc. 18th World Conf. Nondestr. Test.* (2012).
3. V. Batshev, A. Machikhin, and Y. Kachurin, "Stereoscopic tip for a video endoscope: problems in design," *Proc. SPIE* **10466**, 104664D (2017).
4. R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed., Cambridge University Press, Cambridge (2004).
5. D. A. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*, 2nd ed., Pearson, New Jersey (2012).
6. K. B. Lim and Y. Xiao, "Virtual stereovision system: new understanding on single-lens stereovision using a biprism," *J. Electron. Imaging* **14**, 043020 (2005).
7. X. Cui et al., "Accurate geometrical optics model for single-lens stereovision system using a prism," *J. Opt. Soc. Am. A* **29**, 1828–1837 (2012).
8. L. Wu, J. Zhu, and H. Xie, "Single-lens 3D digital image correlation system based on a bilateral telecentric lens and a biprism: validation and application," *Appl. Opt.* **54**, 7842–7850 (2015).
9. K. B. Lim and Q. Qian, "Biprism distortion modeling and calibration for a single-lens stereovision system," *J. Opt. Soc. Am. A* **33**, 2213–2224 (2016).
10. A. Gorevoy and A. Machikhin, "Optimal calibration of a prism-based videoendoscopic system for precise 3D measurements," *Comput. Opt.* **41**(4), 535–544 (2017).
11. H. F. Grahn and P. Geladi, *Techniques and Applications of Hyperspectral Image Analysis*, Wiley, Chichester (2007).
12. C. Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*, Springer Science & Business Media, New York (2003).
13. L. Cartz, *Nondestructive Testing: Radiography, Ultrasonics, Liquid Penetrant, Magnetic Particle, Eddy Current*, ASM International, Ohio (1995).
14. M. Sardela, *Practical Materials Characterization*, Springer, New York (2014).
15. J. Sasian, *Introduction to Aberrations in Optical Imaging Systems*, Cambridge University Press, New York (2013).
16. A. Machikhin et al., "Calibration of a stereoscopic video endoscope for precise three-dimensional geometrical measurements in arbitrary spectral bands," *Proc. SPIE* **106770**, 106770F (2018).
17. J. Kannala and S. S. Brandt, "A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses," *IEEE Trans. Pattern Anal. Mach. Intell.* **28**, 1335–1340 (2006).
18. K. Kanatani, *Statistical Optimization for Geometric Computation: Theory and Practice*, Dover Publications, New York (2005).
19. A. Fusiello, E. Trucco, and A. Verri, "A compact algorithm for rectification of stereo pairs," *Mach. Vision Appl.* **12**(1), 16–22 (2000).
20. J. Mallon and P. F. Whelan, "Projective rectification from the fundamental matrix," *Image Vision Comput.* **23**, 643–650 (2005).
21. L. Deng et al., "A novel camera calibration technique based on differential evolution particle swarm optimization algorithm," *Neurocomputing* **174A**, 456–465 (2016).
22. J. A. M. Rodríguez and F. C. M. Alanís, "Binocular self-calibration performed via adaptive genetic algorithm based on laser line imaging," *J. Mod. Opt.* **63**(13), 1219–1232 (2016).
23. J. Mallon and P. F. Whelan, "Which pattern? Biasing aspects of planar calibration patterns and detection methods," *Pattern Recognit. Lett.* **28**, 921–930 (2007).
24. A. E. Conrady, *Applied Optics and Optical Design*, Dover Publications, New York (2011).

Alexander S. Machikhin is an assistant professor in the National Research University "Moscow Power Engineering University" and a leading researcher in the Scientific and Technological Center of Unique Instrumentation of Russian Academy of Sciences. He received his BEng and MEng degrees from Bauman Moscow State Technical University in 2005 and 2007, respectively, and his PhD degree in optics from the Scientific and Technological Center of Unique Instrumentation of Russian Academy of Sciences in 2010. His current research interests include spectral imaging, acousto-optics, holography, and optoelectronic systems. He is a member of SPIE.

Alexey V. Gorevoy is an engineer in the National Research University "Moscow Power Engineering University" and a researcher in the Scientific and Technological Center of Unique Instrumentation of Russian Academy of Sciences. He was graduated from Bauman Moscow State Technical University in 2010. His current research interests include optical design, digital image processing, 3-D imaging, and computer graphics.

Demid D. Khokhlov is an engineer in the National Research University "Moscow Power Engineering University" and a junior researcher in the Scientific and Technological Center of Unique Instrumentation of Russian Academy of Sciences. He received his BEng and MEng degrees from Bauman Moscow State Technical University in 2015 and 2017, respectively.

Arseniy O. Kuznetsov is an engineer in the Scientific and Technological Center of Unique Instrumentation of Russian Academy of Sciences. He received his BEng from Bauman Moscow State Technical University in 2018.