RAY-TRACING CAMERA MODEL
FOR PRECISE THREE-DIMENSIONAL MEASUREMENTS
USING A PRISM-BASED ENDSOCOPIC PROBE

A.V. Gorevoy\textsuperscript{1,2}, A.S. Machikhin\textsuperscript{1,2}

\textsuperscript{1}Scientific and Technological Center of Unique Instrumentation,
117342, ul. Butlerova 17, Moscow, Russia, gorevoy.a@gmail.com

\textsuperscript{2}Moscow Power Engineering Institute,
111250, Krasnokazarmennaya ul. 14, Moscow, Russia, machikhinAS@mpei.ru

ABSTRACT

Video endoscopes are widely used for non-destructive testing of internal elements of complex objects without disassembling. In order to implement the stereoscopic method, modern video endoscopes are equipped with the attachable prism-based stereo adapters which allow to capture images from two different viewpoints on a single sensor. The obtained images are considered as captured by two virtual cameras and processed using the camera calibration, corresponding points search and 3D reconstruction methods. The key condition for achieving high measurement accuracy is the optimal choice of the mathematical model of the utilized optical system. In this research, we show that the measurement error may be significantly reduced if the conventional pinhole camera model with polynomial distortion approximation is substituted by the ray tracing model based on the vector form of Snell’s law. The effectiveness of the proposed approach is demonstrated by multiple measurements of geometric parameters for the test target and the parts of the aircraft engine. The capability to predict the measurement accuracy depending on the utilized mathematical model and the image processing algorithms is also discussed.

Key words: endoscopy, machine vision, stereoscopic imaging, prism-based stereo system.

1. Introduction

Video endoscopes are one of the main tools for non-destructive testing of internal elements of complex industrial objects (aviation and car engines, pipelines, generators, heat exchangers) without disassembling. Remote visual inspection (RVI) provides highly informative data about the quality of hard-to-reach surfaces and allows simultaneous 3D spatial measurements of the detected defects [1-3]. Endoscopic measurement technologies are based on the triangulation methods (stereo, shadow, multi-point, phase); the stereoscopic method is the most frequently used and well developed. In order to implement it, modern video endoscopes are equipped with the attachable prism-based stereo adapters which make possible to capture images from two different viewpoints on a single sensor [4,5]. The obtained images are considered as captured by
two virtual cameras and processed using the camera calibration, corresponding points search and 3D reconstruction methods as shown in Fig. 1.

![Fig. 1: The data processing pipeline of a prism-based 3D measurement endoscopy system](image)

The key condition for achieving high measurement accuracy is the optimal choice of the mathematical model of the utilized optical system. This model is used for camera calibration and 3D reconstruction algorithms, as well as for preliminary calculations at the initial design stages and for formulating the requirements for sensor and optical system [5-9]. The calculation of a merit function for the measurement system requires the technique for the estimation of 3D coordinates measurement uncertainty based on the known error in determining the image coordinates for corresponding points and the error of the calibration parameters. Furthermore, this technique is very important for displaying recommendations to operator and making decisions on the RVI results.

In this paper, we compare mathematical models for the prism-based stereoscopic imaging device and estimate them by the measurements of geometric parameters for the test target and the parts of the aircraft engine. We propose and test the technique for the measurement uncertainty estimation and discuss the influence of the assumptions on the result.

2. **Camera models**

The proposed mathematical models and calibration methods for large scale prism-based single-lens stereovision systems can be divided into two groups based on a considered model of an optical distortion induced by a prism. The first group includes pinhole models with polynomial distortion approximations [9-12]. The image acquired by the system is divided into two halves considered as captured by two virtual cameras which can be calibrated as the independent ones using common techniques [6,13]. The pinhole model assumes that all rays for the camera pass through the central point; the distortion is modeled as 2D transformation applied to the point projections in the unit plane (see Fig. 2a). In this research, we adopt the widely used model (introduced by Brown [14]) which represents distortion as the combination of radial (3rd, 5th and 7th orders) and tangential part. As a result, the model requires 10 parameters to describe each camera and 6 parameters for relative orientation of two cameras [6-8, 13].

Another group consists of the accurate geometrical optics models derived from the 3D vector form of Snell’s law [15,16]. The model combines the pinhole camera model with radial distortion and the backward ray tracing through the prism faces as shown in Fig. 2b. The model uses 9 parameters for 3 prism surfaces and one more parameter for the refractive index in addition to the internal camera parameters. In contrast to the pinhole model, the ray-tracing model can not
provide closed-form solution for the forward transformation because it requires initially unknown direction vector from 3D point (shown red in Fig. 2b). This problem is usually solved by the iterative technique called ray aiming or by look-up-table interpolation.

![Fig. 2: Pinhole (a) and ray-tracing (b) model](image)

The reconstruction of 3D point coordinates from its projections is considered as the ray intersection problem. One can choose the appropriate cost function and use the maximum likelihood estimation to solve it [7,8,17,18]. The same approach is applied to form the cost function for the optimization on the calibration stage.

The lack of techniques allowing to assess 3D measurement error and the applicability of the proposed mathematical models for different camera and prism parameters makes it problematic to apply the results for large scale prism-based systems to endoscopes. The results of computer simulation in our previous work [19] have shown that the pinhole camera model with polynomial distortion could not be equal to the ray tracing model for the typical industrial endoscope with a prism-based stereoscopic adapter.

### 3. Experimental results

We have conducted a series of experiments using industrial video endoscope Mentor Visual iQ Videoprobe made by GE Inspection Technologies (probe diameter 6.1 mm, 1/6" 440000 pix. CCD image sensor) and the forward-view stereo adapter XLG3TM616060FG with the fields of view 55°/55°. The distal end of the endoscope was fixed on a mechanical stand that allows adjustable movement along the axis of the probe. Another mechanical stand was placed in front of the endoscope to mount calibration targets. First, we captured the calibration set of images placing three plain chessboard calibration targets (with 0.5, 1 and 2 mm cell size) at the distances from 10 to 45 mm from the probe and rotating them around horizontal and vertical axes about 30° (The examples of the captured images are shown in Fig. 3a, 3b). Second, the plane of calibration target was set approximately perpendicular to the axis of the probe. We used the translation stage to shift the distal end of the endoscope and captured images with 1 mm step to obtain the test set of images in the range from 12 to 27 mm for the medium-sized target and from 22 to 40 mm for the large-sized one. All captured images were processed automatically to extract the coordinates of the chessboard corners [20]. Finally, we acquired images of the parts of the aircraft engine (see Fig. 3c, 3d for examples).

We have implemented the calibration and 3D reconstruction algorithms for two mathematical model considered in Section 2 using MATLAB iterative solver for the constrained minimization problem. The images of the calibration set were processed to estimate parameters for the
analyzed models. Next, the images of the test set were used to calculate the 3D coordinates of each chessboard corner and measure the geometric parameters, such as: 1) distances in the plane perpendicular to the axis of the probe (x- and y- segment in Fig. 3a), 2) distances from point to plane (the point is shown in red in Fig. 3a, the three other points are used to define the plane), 3) area of the square as the sum of two triangle areas.

Fig. 3: Examples of captured images and points used for measurements of geometric parameters

The calculated errors of the measurements are indicated by dots color in Fig. 4 according to the colorbar shown in the bottom of the figure. Each row represents one type of measurement, two columns corresponds to the pinhole model (Model 1, usual) and the ray-tracing model (Model 2, prism). The data is provided for the medium-sized target, the true values of segments are 1 mm, the true values of areas are 1 mm$^2$. One can see that the error value varies significantly for different parts of the observed volume and the different types of measurements. In order to perform quantitative analysis, we divided the obtained data set into zones according to the distance along the z-axis and calculated mean and standard deviation (std) of the segment length and area for every zone; the resulting graphs are placed in the right column in Fig. 4. 

The results indicate that the error for x-segments and area grows significantly on the edges of the field of view for model 1. In contrast, both models demonstrate relatively similar performance for y-segments. The standard deviation for distance-to-plane measurements is comparable for both models, but the usage of model 1 leads to unacceptable bias which increases with distance. The main conclusion based on the experimental results is that the ray-tracing model clearly outperforms the pinhole model for the distances exceeding 15 mm.
We have also compared the two considered models using the distance measurements performed by operator on the images of the aircraft engine parts (see Fig. 3c, 3d) with the same results and conclusions. The ray-tracing model provides better accuracy if the distance to the specimen or the measured distance is large (for example, the specimen shown in Fig. 3c is placed at 30 mm and the measured distance is 6.64 mm).
4. Error prediction

The optimal choice of the mathematical model for prism-based stereo-imager allows to use the developed mathematical description of calibration and 3D reconstruction problem for the prediction of measurement errors. In order to avoid uncertainties of the first order approximation we can use the unscented transformation (UT) [21, 22]. This method is not so computationally intensive as the Monte Carlo simulation and can be used for the prediction of errors during the measurement procedure. But, similar to the Monte Carlo simulation, the UT does not require analytical derivatives which make possible to use the triangulation algorithm as a black box. This fact allows us to use this technique for all considered mathematical models and measurement types despite their particular cost function.

The input data for the uncertainty estimation technique are the covariation matrices for image coordinates and calibration parameters which should be defined for all points involved in particular measurement type. The usual assumption is that the image coordinate error is independent and has the same distribution for each coordinate and each point of image [7-9], hence, the covariance matrix can be represented as the unit matrix multiplied by $\sigma$. This assumption is reasonable if the object is clearly localized in both coordinates (point, corner, cross, circle center, etc.) and the main sources of error are discretization and uniform image blurring in the presence of uncorrelated noise.

We have developed the software implementing the uncertainty estimation technique based on the UT and applied it to the 3D point coordinates and calibration target positions calculated for the test set of experimental data described in Section 3. The estimated mean errors and standard deviations for x-, y- segments and distance to plane measurements as well as real data for the ray-tracing model are shown in Fig. 5 for several values of image coordinate error $\sigma$.

![Fig. 5: Comparison of real and predicted mean and standard deviation of measured distance for different values of image coordinate errors](image)

One can see that the real values for different measurement types correspond to the predicted values for different $\sigma$. Next, we have also calculated the mean and standard deviation considering the image coordinate errors are correlated for different points and the image coordinate errors for two images are also correlated. The results indicate that the higher correlation coefficient, the lower the estimated standard deviation; this is practically equivalent to lower values of $\sigma$ for all measurement types. Hence, the correlation of image coordinate errors can not be the main reason that the values of $\sigma$ for x-segments are higher compared to other measurement types. This may be caused by the systematic calibration error or other unaccounted factors. In addition, it is noticeable that the real values of standard deviation comprise value independent of the distance, which can correspond to the calibration target error. In order to exclude the impact of the utilized uncertainty estimation technique, the calculation were repeated using the simulated data with
known covariation matrix for the same calibration parameters and target positions. The applicability of the UT-based technique was verified for all points inside the observed volume and all used covariation matrices.

5. Conclusions

We analyzed compared two mathematical models of prism-based endoscopic probes. The experimental data confirmed the results previously obtained by computer simulation: the measurement error may be significantly reduced if the conventional pinhole camera model with polynomial distortion approximation is substituted by the ray tracing model based on the vector form of Snell’s law. We have shown that the simple calibration method using a flat target allows to conduct the calibration of the endoscope with attached prism-based stereo adapter and reach the desired accuracy. The method does not require the results of other specific calibrations or the prism parameters.

The error of the measurement has a complex spatial distribution and depends strongly on the type of measurement and the orientation of the segment or plane. The experimental results indicate that the real error values do not exactly correspond to the predicted values under the assumption that the image coordinate error is independent and has the same distribution for each coordinate and each point of image. It is shown that the presence of a correlation, which is usually not taken into account, can significantly change the result. Nevertheless, the positive value of the correlation coefficient lower the error values equally for all measurement types, hence, the uncertainty estimation for zero correlation coefficient can still be used as the worst-case scenario when comparing different options and optimizing parameters at the preliminary design stage. For the same reasons, this estimation can be used to give the operator a warning about insufficient measurement accuracy. However, the absolute values of the error calculated under these assumptions should be treated with caution.

The proper analysis of other factors affecting the form of the covariance matrix of the image coordinates error for each point of the working volume is very difficult. The texture and roughness of the calibration target surface appears as random addition to ideal black-and-white pattern of marker when imaged by camera. This random addition significantly depends on the reflection properties of the surface and the position of a light source, magnification, aberrations and vignetting of the optical system. Hence, it is still a bit different for two halves of the image acquired with a prism-based endoscopic probe because of different observation angles. As a result, the assumptions about equal imaging conditions for each point can barely be fulfilled even for simple calibration target and automatic correspondence search, and the experiment to separate the impacts of different factors can not be set up. Better uncertainty estimation techniques require detailed simulation of the image registration and stereo correspondence search. This approach is not applicable for preliminary calculations, but can be used for more detailed calculations at the next design stages in conjunction with the development of image processing algorithms.

6. References