Check for updates

Optics Letters

Projector distortion correction in 3D shape measurement using a structured-light system by deep neural networks

SHENZHEN LV,[1,2] QIANG SUN,[1] YUYUAN ZHANG,[3] YANG JIANG,[1] JIANBAI YANG,[1,2] JIANZHUO LIU,[1] AND JIAN WANG[1,*]

[1]R&D Center of Precision Instruments and Equipment, Changchun Institute of Optics, Fine Mechanics & Physics, Chinese Academy of Sciences, No. 3888, Dongnanhu Road, Changchun, Jilin, China
[2]University of Chinese Academy of Sciences, Beijing 100049, China
[3]College of Science, Jilin Institute of Chemical Technology, Jilin 132022, China
*Corresponding author: wangj@ciomp.ac.cn

Received 15 November 2019; accepted 21 November 2019; posted 2 December 2019 (Doc. ID 383338); published 24 December 2019

In a structured-light system, lens distortion of the camera and projector is the main source of 3D measurement error. In this Letter, a new approach, to the best of our knowledge, of using deep neural networks to address this problem is proposed. The neural network consists of one input layer, five densely connected hidden layers, and one output layer. A ceramic plate with flatness less than 0.005 mm is used to acquire the training, validation, and test data sets for the network. It is shown that the measurement accuracy can be enhanced to 0.0165 mm in the RMS value by this technique, which is an improvement of 93.52%. It is also verified that the constructed neural network is with satisfactory repeatability. © 2019 Optical Society of America

https://doi.org/10.1364/OL.45.000204

Recently, three-dimensional (3D) shape measurement using structured-light system has been extensively studied [1,2] and widely applied due to its high speed, simple setup, and high accuracy [3]. To achieve high measurement accuracy, one of the crucial elements is to accurately calibrate the structured-light system. Currently, there are two popular calibration methods for the structured-light system. One is the phase-to-height conversion algorithm [4] in which the 3D shape is reconstructed by establishing the relationship between the depth and phase value. With this method, the locations of the projector and camera can be arbitrarily arranged without strict geometric constraints. The other is the stereo-vision-based model proposed by Zhang and Huang [5] in which the projector is treated as an inverse camera to reconstruct the 3D shape. This method has the advantages of simplicity, simultaneity, and high accuracy. In our previous work, we also adopted Zhang and Huang's method for the calibration of a structured-light system [6].

The distortions of lenses of the projector and camera are unavoidable owing to industrial manufacturing, and high accuracy of 3D shape reconstruction cannot be achieved without complete correction of the distortions [7]. Camera calibration has been studied extensively, and the distortion of the camera lens can be well corrected by the distortion parameters. However, the distortion correction of the projector lens remains difficult because of two main reasons. First, the projector cannot be used to directly capture images. Secondly, the correction by the distortion parameters used for the camera lens is not suitable for the projector lens due to its high optical efficiency and optical offset [8].

To circumvent the problem of projector distortions, Peng *et al.* [7] proposed an adaptive fringe projection technique for the system calibration by a phase-to-height conversion algorithm. In the process, a standard plane was randomly placed in several different locations. In the absence of lens distortions, the recovered phase distribution for each location should be in the form of a rational fraction [9]. In practice, however, there was an additional bending phase in the recovered phase due to the distortion of the projector lens. Then the additional phase distribution was fitted with Zernike polynomials. According to the fitted results, the projected fringe patterns were modified to eliminate the projector distortion before projecting. This method does not need to calibrate the projector and has a measurement accuracy of a standard plane target of 0.0213 mm in the RMS value.

For Zhang and Huang's system calibration model, Yang *et al.* [8] proposed a residual compensation method to correct the projector distortion. In this method, the projected fringe patterns were modified twice. First, pre-distortion fringe patterns were generated by modifying the fringe patterns according to the distortion coefficients of the projector lens, aiming to eliminate the main measurement error caused by the distortion. Then a residual distortion map was built based on per-pixel measurement of a planar target, and the pre-distortion fringe patterns were refined by this residual distortion map. This method has the advantage of not requiring any auxiliary equipment, and the measurement accuracy of a standard plane target is 0.0435 mm in the RMS value.

Deep learning has been successfully applied to many different fields, including computer vision and natural language processing due to the advances in hardware and software techniques, and increased data availability [10–12]. There are several examples of deep learning applications in a structured light system. Feng *et al.* [10] used the deep neural networks to perform fringe analysis from a single fringe pattern for accurately retrieving phase information. Similarly, Shi *et al.* [11] proposed a deep learning approach based on an enhanced label and patch for phase retrieval. Jeught *et al.* [12] proposed a neural network with a large simulations data set in the training process to extract the height information of the object from a single-shot fringe pattern. In this Letter, a new approach to correct the distortion of the projector lens for Zhang and Huang's system calibration model is presented, entirely based on deep neural networks. The experimental results demonstrate that the accuracy of the standard plane is 0.0165 mm (RMS) which has an improvement of 93.52%. Moreover, the neural network construction is with satisfactory repeatability, and the proposed method has the advantage of no need of changing the projecting fringe patterns.

Here Zhang and Huang's calibration method is briefly recalled as follows. The camera is a pinhole model in which the relation between a point of (x^w, y^w, z^w) on the object and its projection on the image sensor of (u^c, v^c) can be written as

$$s^c[u^c\ v^c\ 1]^T = M^c[x^w y^w\ z^w\ 1]^T,$$
$$M^c = A^c[R^c\ t^c], \tag{1}$$

where s^c is an arbitrary scale factor, A^c is the camera's intrinsic matrix, and R^c and t^c are the camera's extrinsic matrices. The projector can be regarded as an inverse camera in which a similar relation can be described as

$$s^p[u^p\ v^p\ 1]^T = M^p[x^w\ y^w\ z^w\ 1]^T,$$
$$M^p = A^p[R^p\ t^p], \tag{2}$$

where the p superscript denotes the parameters of the projector. Due to the projector not being able to capture images like a camera, Zhang and Huang presented the phase-aided method to establish the correspondence relationship between the camera and projector pixels using two sets of orthogonal sinusoidal fringe patterns. In the system calibration, the feature points on the calibrated plate are captured by the camera and the correspondent pixel coordinates on the CCD are extracted. With the established correspondence relationship, the acquired pixel coordinates on the CCD are mapped to those on the digital micromirror device of the projector. Then both the camera and projector are calibrated. When the intrinsic and extrinsic parameters of the system are acquired, 3D information on the object can be obtained using Eqs. (1) and (2).

In this Letter, the process of the distortion correction of the projector lens is shown in Fig. 1. Its detailed explanation is given below.

Step 1: Calibrate the structured-light system. The structured-light system is calibrated using Zhang and Huang's method as described above.

Step 2: Correct the camera lens distortion. The camera lens distortion can be modeled as a vector of five elements as $\mathrm{Dist}^c = [k_1, k_2, p_1, p_2, k_3]$, where k_1, k_2, and k_3 are the radial

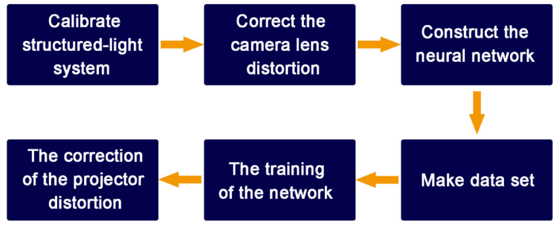

Fig. 1. Flowchart for correction of the projector distortion.

distortion coefficients, and p_1 and p_2 are the tangential distortion coefficients. Then the captured images can be undistorted based on the camera calibration results using an OpenCV camera calibration toolbox.

Step 3: Construct the neural network. The architecture of our deep neural network is shown in Fig. 2. It consists of one input layer, five densely connected hidden layers and one output layer. The layers are connected cascadedly to solve the gradient vanishing problem during training. There are 50 units in each hidden layer. Behind each densely connected hidden layer, a dropout layer with the rate setting to be 0.1 is included to reduce over-fitting. Tanh is adopted as an activation function because there are positive and negative values in the input data. The deep neural network is implemented using the Keras framework in PyCharm. It should be pointed out that the dropout layers do not appear in Fig. 2 because the network is drawn with Matlab.

Step 4: Make a data set. To make a data set, a ceramic plate with a size of 300 mm × 300 mm and flatness less than 0.005 mm is placed at 33 different locations separately, with different orientations and distances. These locations should be throughout the measurement space. Then we calculate the measurement error distribution of DZ by the least square plane fitting of the measurement result of [X, Y, Z] at each location. [X, Y, Z] and DZ are used to be the input and output of the network, respectively. During the making of the data set, we shuffle the measurement data set at each location first. Then we randomly extract 70% of the shuffled data set of each location to form a training data set. Finally, the remaining data set at each location is randomly divided into two parts, with one part used to form a validation set, and the other part is used to form a test set. In order to reduce the data loading time, we have made these data points into a Hierarchical Data Format.

Step 5: The training of the network. The network is trained with a batch size of 3000 and mean square error as the loss function. During the training, the hyper-parameters of the neural network are determined by the loss value of the validation data set. To be concrete, Adam optimization with an initial learning rate of 1×10^{-4} is used, and the learning rate is controlled by a callback function to reduce its value by 0.8 times after four epochs.

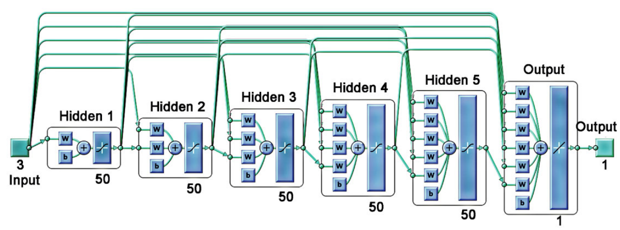

Fig. 2. Our network structure.

Step 6: The correction of the projector distortion. The error of the test set is predicted by the trained network; then the projector lens distortion is corrected by the predicted result.

The 3D shape measurement system used includes a CCD camera (DAHENG MER-131-210U3M-L) with a VTG1214-M4 lens and a projector (DLP 6500). The camera resolution is 1280×1024 pixels, and the projector resolution is 1920×1080 pixels. The multi-frequency phase-shifting algorithm with pitches of 21, 24, and 180 pixels is adopted to recover the absolute phase maps, and the windowed Fourier transform is adopted to suppress the noise in extraction of the wrapped phase [13].

We start the experiment from removing the camera lens distortion using a camera calibration toolbox in OpenCV (step 2). In step 1, the calibrated camera lens distortion is given as

$$\mathrm{Dist}^c = [-0.21741, 0.206541, 2.3 \times 10^{-4}, -2.7 \times 10^{-4}, 0]. \tag{3}$$

The calibrated target images are undistorted by the initUistortRectifyMap function in OpenCV, based on the calibrated camera internal parameters and the distortion coefficients in Eq. (3). With the undistorted target images, the camera is recalibrated and the distortion coefficients are reduced to

$$\mathrm{Dist}^c = [1.8 \times 10^{-3}, -2.557 \times 10^{-2}, \\ -1.3 \times 10^{-4}, 2 \times 10^{-5}, 0]. \tag{4}$$

It can be seen that the distortion coefficients can almost be ignored, showing the camera lens distortion is well corrected. In the following experiments, all images captured by the camera are first undistorted.

With the constructed neural network in step 3 and the training data set acquired in step 4, the neural network is trained as described in step 5. Thirty-two epochs are adopted to achieve satisfactory loss value. Figure 3 shows the loss value as function of epochs for both the training data set and validation data set. It can be seen that two curves are very close, indicating that there is no apparent over-fitting in the process of the neural network training.

With the trained neural network, the projector distortion is corrected as described in step 6. The test data set at all locations is brought into the trained neural network to evaluate the performance of the network. Then the corrected 3D shape at each location is acquired, and the error value for each input point (X, Y, Z) is predicted. Figure 4 shows the experimental results for the test data set at the first location. Figures 4(a) and 4(b) show the 3D shape and error distribution of the original data set, respectively, and 4(c) and 4(d) show the 3D shape and error distribution of the data set after correction, respectively. It can be seen that the peak-to-valley (PV) value in Fig. 4(b) is 2.0354 mm, and it is reduced to 0.1607 mm in Fig. 4(c). The experimental results for the remaining 32 locations are shown in Visualization 1, where the maximum PV value of the error after correction is 0.1644 mm, indicating that satisfactory improvement is achieved by the neural network. Moreover, the error histogram of the test data set after correction is obtained and shown in Fig. 5. From Fig. 5, we can get the RMS value of the error of the data set which is 0.0165 mm. As compared with that of the original test data set of 0.2546 mm, an improvement of 93.52% is achieved. The above experimental results prove that the constructed neural network can effectively correct the distortion error of the projector in the 3D shape measurement.

To further evaluate the performance of the neural network, another ceramic plate with a size of 300 mm $\times$ 300 mm and flatness less than 0.010 mm is placed at 11 different locations, and 3D shape measurement is conducted. At each location, the measurement data, with the error caused by the projector lens distortion, are corrected by the trained neural network, and the residual error of the measurement is listed in Table 1. The

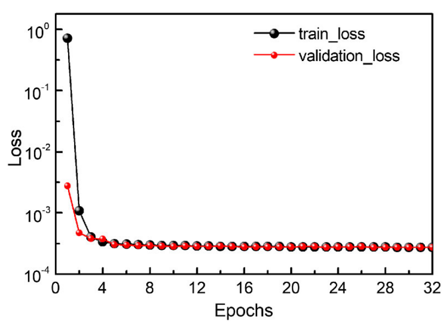

Fig. 3. Loss function.

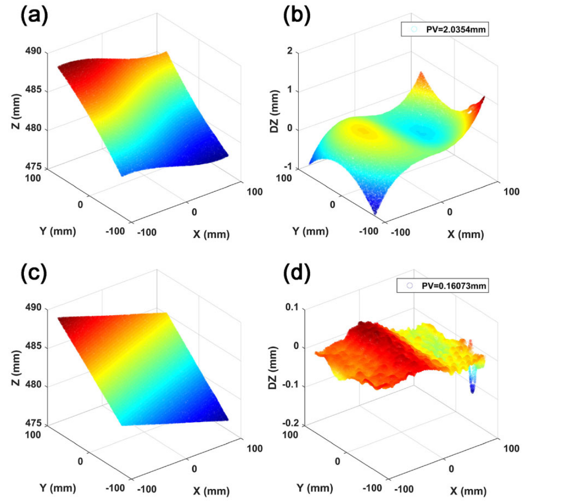

Fig. 4. Experimental results for the test data set at the first location with (a) the 3D shape of the original test data set, (b) the error distribution of the original test data set, (c) the 3D shape of the data set after correction, and (d) the error distribution of the test data set after correction.

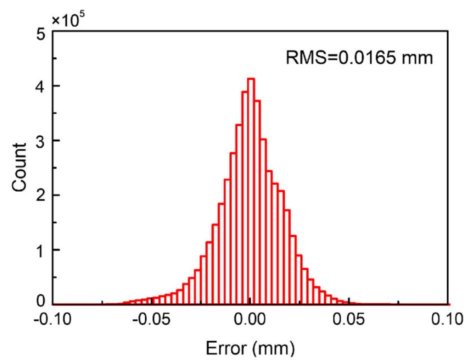

Fig. 5. Error histogram of the test data set corrected by the neural network.

Table 1. Residual Error of 3D Shape Measurement at 11 Different Locations (Unit: mm)

Location	PV	RMS	Location	PV	RMS
1	0.1252	0.0141	7	0.1308	0.0128
2	0.1113	0.0147	8	0.1380	0.0137
3	0.1313	0.0149	9	0.1121	0.0169
4	0.1260	0.0141	10	0.1095	0.0130
5	0.1549	0.0161	11	0.1229	0.0149
6	0.1247	0.0125	–	–	–

Table 2. Statistical Result of the Repeatability of the Neural Network (Unit: mm)

Time	PVmax[a]	RMS	Imp[b] (%)
1	0.1588	0.0164	93.57
2	0.1586	0.0167	93.44
3	0.1605	0.0165	93.53
4	0.1579	0.0165	93.51
5	0.1669	0.0163	93.58
6	0.1539	0.0164	93.57
7	0.1703	0.0163	93.61
8	0.1657	0.0167	93.45
9	0.1654	0.0162	93.63
10	0.1571	0.0167	93.45
mean	–	0.0165	93.53

[a]PVmax denotes maximum PV value.
[b]Imp denotes improvement of the RMS value.

second column in Table 1 lists the PV values of the residual error at 11 locations, and the third column lists the RMS values. It can be seen that the maximum and minimum RMS values of the residual error is 0.0169 and 0.0125 mm, respectively, indicating that the projector lens distortion has been satisfactorily corrected by the constructed neural network.

Finally, to assess the repeatability of the neural network construction, we repeat the experimental steps from 4 to 6 by 10 times, with the same measurement data at the 33 different locations. For each experiment, the measurement data are randomly selected again to re-form the training data set, validation data set, and test data sets. Table 2 lists the experimental results of the projector distortion correction in 3D shape measurement. The second column lists the maximum PV values of the residual errors after distortion correction, and the third column lists the RMS values. The fourth column lists the improvement of the residual error in the RMS value as compared with the correspondent error of the original test data set. It can be seen that the maximum PV value of the residual error is 0.1703 mm, the mean of the RMS values is 0.0165 mm, and the mean of the improvement is 93.53%, indicating that the neural network construction has satisfactory repeatability.

We have demonstrated that the technique of deep neural networks can be successfully applied into projector distortion correction in 3D shape measurement using structured-light system. The proposed network consists of one input layer, one output layer, and five densely connected hidden layers with a dropout layer behind each hidden layer. The experimental results show that after distortion correction by the trained neural networks the measurement accuracy of 0.0165 mm in the RMS value is achieved, which is an improvement of 93.52%. The proposed method also possesses the advantage of no need of changing the projecting fringe patterns. Moreover, the network establishment is with satisfactory repeatability, which is verified by repeating the experiment by 10 times in this Letter.

Funding. The Scientific Instrument Developing Project of the Chinese Academy of Sciences (YJKYYQ20180067).

Acknowledgment. The authors would like to express their sincere gratitude to Professor Yishi Shi and his team of the University of Chinese Academy of Sciences for their support of the network construction.

Disclosures. The authors declare no conflicts of interest.

REFERENCES

1. S. Zhang, Opt. Laser Eng. **48**, 149 (2010).
2. S. Gorthi and P. Rastogi, Opt. Laser Eng. **48**, 133 (2010).
3. S. Heist, P. Dietrich, M. Landmann, P. Kühmstedt, G. Notni, and A. Tünnermann, Light Sci. Appl. **7**, 71 (2018).
4. H. Du and Z. Wang, Opt. Lett. **32**, 2438 (2007).
5. S. Zhang and P. S. Huang, Opt. Eng. **45**, 083601 (2006).
6. S. Lv, Q. Sun, J. Yang, Y. Jiang, F. Qu, and J. Wang, Opt. Laser Eng. **122**, 65 (2019).
7. J. Peng, X. Liu, D. Deng, H. Guo, Z. Cai, and X. Peng, Opt. Express **24**, 21846 (2016).
8. S. Yang, M. Liu, J. Song, S. Yin, Y. Ren, J. Zhu, and S. Chen, Opt. Laser Eng. **114**, 104 (2019).
9. H. Guo, M. Chen, and P. Zheng, Opt. Lett. **31**, 3588 (2006).
10. S. Feng, Q. Chen, G. Gu, T. Tao, L. Zhang, Y. Hu, W. Yin, and C. Zuo, Adv. Photonics **1**, 025001 (2019).
11. J. Shi, X. Zhu, H. Wang, L. Song, and Q. Guo, Opt. Express **27**, 28929 (2019).
12. S. Jeught and J. Dirckx, Opt. Express **27**, 17091 (2019).
13. K. Qian, Appl. Opt. **43**, 2695 (2004).

159.226.165.16 {ts '2021-04-26 23:26:18'}