Lara Hoffmann*, Ines Fortmeier, and Clemens Elster

Deep learning for tilted-wave interferometry

Deep Learning für die Tilted-Wave Interferometrie

<https://doi.org/10.1515/teme-2021-0103>

Received September 27, 2021; accepted November 4, 2021

Abstract: The tilted-wave interferometer is an interferometrical measurement system for the accurate optical form measurement of optical aspheres and freeform surfaces. Its evaluation procedure comprises a high-dimensional inverse problem to reconstruct the form of the surface under test from measured data. Recent work has used a deep learning hybrid approach to solve the inverse problem successfully in a simulation environment. A quantification of the model uncertainty was incorporated using ensemble techniques. In this paper, we expand the application of the deep learning approach from simulations to measured data and show that it produces results similar to those of a state-of-the-art method in a real-world environment.

Keywords: Deep learning, ensemble learning, uncertainty quantification, interferometry, optical form measurement, inverse problem, aspheres, metrology.

Zusammenfassung: Das Tilted-Wave Interferometer ist ein interferometrisches Messsystem für die hochgenaue optische Formmessung von optischen Asphären und Freiformflächen. Dabei muss für die Formrekonstruktion einer Prüflingsoberfläche ein hochdimensionales inverses Problem gelöst werden. Kürzlich wurde ein hybrider Ansatz, der auf der Anwendung von tiefen neuronalen Netzwerken basiert, zur Lösung des inversen Problems vorgestellt. Dieser wurde erfolgreich auf simulierten Daten angewandt. Mit Hilfe von Ensembletechniken kann außerdem eine Modellunsicherheit für die rekonstruierte Prüflingsoberfläche angegeben werden. In dieser Arbeit erweitern wir die Anwendung von Simulationen auf gemessene Daten und zeigen an einem Beispiel, dass der hybride Deep-Learning-Ansatz ähnliche Ergebnisse liefert wie eine moderne Messmethode, die dem aktuellen Stand der Technik entspricht.

*Corresponding author: Lara Hoffmann, Physikalisch-Technische Bundesanstalt (PTB), Abbestr. 2-12, 10587 Berlin, Germany, e-mail: lara.hoffmann@ptb.de

Ines Fortmeier, Physikalisch-Technische Bundesanstalt (PTB), Bundesallee 100, 38116 Braunschweig, Germany

Clemens Elster, Physikalisch-Technische Bundesanstalt (PTB), Abbestr. 2-12, 10587 Berlin, Germany

Schlagwörter: Maschinelles Lernen, Unsicherheitsquantifizierung, Interferometrie, optische Formmessung, Asphären, Metrologie.

1 Introduction

The need for accurate measurement techniques increases with ongoing technological advancements. A manufactured good can be fabricated only as accurately as it can be measured. Non-spherical optics, like aspheres or freeform surfaces, are indispensable for modern optical systems [1]. In order to track the state-of-the-art in the form measurement of such optical surfaces, interlaboratory comparison studies are performed to analyze the differences between the various high accuracy measuring systems, including tactile and optical instruments [2, 3, 4]. The tilted-wave interferometer (TWI) [5, 6] is one of the state-of-the-art interferometrical measurement techniques for optical form measurements of optical aspheres and freeform surfaces.

The evaluation procedure of the TWI takes account of both measured data and simulated data, with the deviation of the specimen from its known design topography derived from the differences between the measured and simulated data, thus yielding a high-dimensional inverse problem. Conventional TWI methods consist of three major steps for reconstructing the surface form: the prior correction of the interferometer model (“calibration”) [5, 7, 8]; the reconstruction of the form deviation represented by Zernike polynomial functions [9]; and the reconstruction of the remaining high-frequency form deviations [10]. Measurement uncertainties are as yet not available for tilted-wave interferometry, but previous work addresses a number of relevant investigations [9, 11, 12, 13]. Recent work [14, 15] suggests the use of a deep learning hybrid approach to solve the inverse topography reconstruction problem. In this hybrid approach, the surface form is reconstructed by a deep learning model, while the prior calibration is performed by the conventional method using the simulation tool box SimOptDevice [16, 12, 17] developed at the Physikalisch-Technische Bundesanstalt (PTB). The recent approach incorporates a quantification of the model uncertainty and was shown to achieve accurate results and reliable uncertainty estimates in a simulation-based environment.

Deep learning methods have become very popular due to their ability to learn complex relationships from given data. They have been successfully applied to various applications ranging from object detection [18] and computational imaging [19] to natural language processing [20] and autonomous driving [21]. Deep learning applications in optics include single-frame interferogram demodulation [22], adaptive optics that deblur retinal images [23], correcting freeform surface misalignments in non-null interferometers [24], and wavefront sensing [25] where Zernike coefficients of some aberrated wavefronts are estimated from a single intensity image. The main drawback of using deep learning models is their black box character. There are two main approaches to overcoming this problem. One is to incorporate an uncertainty quantification in the network prediction [26]. The other applies explainability methods to gain a better understanding of how the network output relates to the input signal [27].

Ensembles [28, 29, 30] and dropout [31, 32] are currently the most popular uncertainty methods in deep learning. The total uncertainty of a new prediction can be divided into an aleatoric and an epistemic part [33]. While the aleatoric uncertainty refers to the irreducible part of the uncertainty, the epistemic uncertainty relates to the model uncertainty of the trained network and can be reduced, for instance, by using more training data. The epistemic part of the uncertainty is of great interest when the goal is to infer a regression function, as is the case for the application of computational optical form measurement under consideration here. By contrast, the aleatoric uncertainty is crucial when sampling new data points from the posterior predictive distribution. A reliable deep learning model should generalize well to out-of-distribution data. This means that the uncertainty quantification should relate to the gap between the model prediction and its ground truth not only with respect to the data originating from the training data domain but also for data generated from outside.

The novelty of this paper is to show that the recently developed deep learning hybrid approach for computational optical form measurement [15] can be applied to measurement data it has never before seen, even if it was trained on purely simulated data. The results obtained include an estimate of the model uncertainty and are shown to be reliable by comparing the deep learning approach to a state-of-the-art method for optical form measurements, namely the tilted-wave interferometer distributed by Mahr GmbH (MarOpto TWI 60) [34].

This paper is structured as follows. The basic principle of the TWI is introduced in Section 2. This is followed in Section 3 by a brief recapitulation of the deep learning

hybrid approach [15], a description of how it is applied to measured data, and a presentation of the form reconstruction result for a typical apshere. In Section 4, the reconstruction result of the hybrid deep learning approach is compared to the result from a state-of-the-art form measurement system. Finally, conclusions are given in Section 5.

2 Tilted-wave interferometry

The experimental setup of the tilted-wave interferometer (TWI) [5, 6] is shown in Figure 1. Incoming light is reflected at the specimen, then interferes with coherent light from the reference arm at the beam splitter, and the resulting interferograms are collected at the camera. The two main characteristic features of the TWI setup are the 2D microlens array, which acts like a 2D point light source array, and a beam stop in the Fourier plane of the imaging optics. With this setup, depending on the local slope of the specimen, many subinterferograms with resolvable fringe densities are generated at the camera. To ensure that there is no interference between neighboring light sources at the camera, four disjoint masks are used in the focal plane of the microlens array. For each mask, the optical path length differences (OPDs), i. e., the differences between the optical path length of the reference arm and that of the measurement arm, are calculated.

Any specimen can be represented by a known design topography and an additional unknown difference topography, i. e., the deviation of the specimen from its design. The TWI form reconstruction is based on the difference between the OPDs measured for the specimen and the OPDs calculated for the design topography. State-of-the-art TWI methods can be summarized into the following three steps:

1. Prior calibration to correct the interferometer model such that it more closely approximates the real measurement system [7]
2. Iterative reconstruction of the form deviations represented by Zernike polynomial functions [35, 9]
3. Reconstruction of the remaining high-frequency form deviations [10]

3 Deep learning hybrid approach

This section will first recapitulate the recently introduced deep learning hybrid approach for computational optical form measurements [15] and then describe how it can

Figure 1: a) The experimental setup of the TWI is illustrated here. Note that only the center point source is activated. b) An example of simulated interferograms that result when different masks are used on the 2D microlens array.

Figure 2: a) An interferometer model simulates optical path length differences (OPDs) for the given topography T . b) The deep neural network is illustrated as a black box. Its input is the difference between the OPDs from the specimen and the OPDs from the design topography. The output is the difference topography, i. e., the deviation of the specimen from its known design topography. c) Having independently trained multiple networks, the ensemble prediction ($y_\mu = \frac{1}{K} \sum_{i=1}^K y_i$) and model uncertainty ($u^2(y_\mu) = \frac{1}{K} \sum_{i=1}^K (y_\mu - y_i)^2$) are the mean and uncorrected sample variance of the individual member predictions. The operations are performed pixel-wise.

be applied to measurement data to reconstruct the surface of a real specimen. All simulations are realized using the simulation toolbox SimOptDevice developed at PTB [17].

3.1 Method

The deep learning hybrid approach [15] is based on the TWI evaluation procedure. In contrast to the conventional method (Section 2), the surface is here reconstructed by applying a deep learning model. However, the calibration step (the first step that corrects the interferometer model) is still performed in the conventional manner.

For a fixed design topography, the deep learning model learns to predict a difference topography from the corresponding differences between the measured OPDs and the simulated OPDs (cf. Fig. 2b). The model is trained on purely simulated data, which means that all data for the design topography and for the specimen are simulated (cf. Fig. 2a). Moreover, no positioning error is included for the topographies. Various specimens can be generated for a fixed design by randomly drawing Zernike coefficients modeling the difference topography. Zernike polynomials are a set of orthogonal polynomials on the unit disk [36]. The difference topographies in the training data are generated by sampling the first 136 coefficients with the exception of the offset and tilt, i. e., the first three coefficients

always equal zero. The deep learning model is trained on perfectly calibrated data, which means that the same interferometer model is used for the simulated data of both the design topography and the specimens.

The deep learning approach is based on deep ensembles [29], which represent a popular method for incorporating an uncertainty quantification into a deep neural network prediction and for making the prediction itself more robust. Multiple networks are independently trained instead of a single one. As sketched in Figure 2c, the ensemble prediction y_μ is the mean value of the individual network predictions and the (squared) model uncertainty $u^2(y_\mu)$ corresponds to the uncorrected sample variance. In [29], the uncorrected sample variance was introduced as a heuristic. It can, however, be justified from a Bayesian perspective [30]. The approach chosen here uses an ensemble of eight deep neural networks, all having the same U-net architecture [37]. The high-dimensional inverse problem of reconstructing the optical surface is treated as an image-to-image regression problem, and U-nets have previously been successfully applied to various imaging tasks [19]. The training took about 45 minutes per network on a single GPU (Tesla V100-NVLink on an Intel Xeon Gold 6226) [15].

The input and output images of the deep neural networks have a resolution of only 64×64 pixels. Less memory is needed to store the data and train the neural networks for low resolution images. Furthermore, generating the training data is much faster because less rays need to be traced through the optical system. This low resolution is sufficient for the purposes of this paper, but it can be adapted as required in future work. Note that the input has four channels due to the disjoint masks on the 2D microlens array (Fig. 1). The deep learning hybrid approach has been validated on purely simulated data. Nonetheless, it was shown to produce reliable results on a disjoint test data set that included input images that were clean, noisy, or simulated with a corrupted interferometer model. Further information on the training of the deep neural networks and more detailed results are presented in [15] (p. 9, Table 1).

An additional post process is added to the deep learning hybrid approach when applied to measured data. It modifies the individual ensemble member predictions before calculating the mean and uncorrected sample variance (cf. Fig. 2c). The steps are summarized as follows:

1. Add the predicted difference topographies to the design topography to reconstruct the specimen.
2. Rotate and shift the point clouds of the reconstructed specimens into the design to align the coordinate sys-

tems of the design function and the reconstructed topography [3].

3. Subtract the design topography to deduce the resulting form deviation of the specimen from its design.
4. Calculate the average and uncorrected sample variance over the ensemble members.

3.2 Reconstruction results on measured data

The deep learning hybrid approach will now be applied to measured data from a real specimen using the trained deep learning model from [15]. The underlying design topography is an asphere.¹ The measured interferograms have a resolution of 2048×2048 pixels. Their OPDs are calculated by using a five-step phase shifting algorithm [39] and the Goldstein unwrapper [40], while the OPDs corresponding to the design topography are simulated by the calibrated interferometer model. The differences between the measured and the simulated OPDs are shown in Figure 3a. Note that the images are sampled down by selecting every 32nd pixel because the input dimension of the deep neural networks is 64×64 for each channel.

The trained deep neural network ensemble from [15] then reconstructs the form deviation of the specimen from its design taking the differences of the OPDs (Fig. 3a) as an input. In addition to the post processing step described above, Zernike polynomial functions are fitted to the resulting difference topography in order to compare the results to the results of a state-of-the-art measurement system (with different resolution) as presented in Section 4. The reconstructed difference topography is plotted in Figure 4a. Figure 4b shows the horizontal and vertical cross sections of the predicted difference topography together with the 95 % symmetric credible intervals of the model uncertainty, i. e. $y_\mu \pm 1.96 \sqrt{u^2(y_\mu)}$ (cf. Fig. 2c).

Positioning errors directly influence the measured OPDs and hence the topography reconstruction. However, no misalignment was taken into account during network training. The difference topography was therefore estimated a second time including a prior misalignment correction. The estimation of the specimen positioning error is realized by minimizing the differences between the measured and simulated OPDs in a least-squares sense considering the parameters of the degrees of freedom of an as-

¹ The conic constant and the paraxial surface radius characterising the asphere [38] are $\kappa = -1$, and $R = 0.0202$ m. The aspherical coefficients are $A_4 = 5.4145 \cdot 10^3 \text{ m}^{-3}$, $A_6 = -8.0413 \cdot 10^5 \text{ m}^{-5}$, $A_8 = -2.9871 \cdot 10^9 \text{ m}^{-7}$, $A_{10} = -1.4918 \cdot 10^{12} \text{ m}^{-9}$, $A_{12} = 1.3777 \cdot 10^{15} \text{ m}^{-11}$, $A_{14} = 4.4258 \cdot 10^{18} \text{ m}^{-13}$, and $A_{16} = -3.4928 \cdot 10^{21} \text{ m}^{-15}$.

Figure 3: a) The down-sampled differences between the measured and simulated OPDs represent the four channels of input to the deep neural networks. b) The simulated OPDs differ from a), because the positioning error between the specimen and the design topography is corrected prior to the simulation.

Figure 4: a) This is the prediction of the deep learning hybrid approach for the measured data of the asphere specimen. b) The horizontal and vertical cross sections of the profile in a) are plotted together with the 95% symmetric credible interval of the model uncertainty. c) Analogous to a) but with the misalignment of the specimen corrected prior to data simulation. d) Analogous to b) but considering the profile from c).

where, i. e., the rotation angle parameters and the position in space, but excluding the position along the optical axis, which is measured with a distance measuring interferometer. The input to the deep neural networks then comprises the difference between the measured OPDs, which are the same as before, and the simulated OPDs computed after correcting the misalignment of the topography. The input data is now more centered, as shown in Figure 3b. Figures 4c and 4d show the resulting reconstructed difference topography. The root-mean deviation of the variance corresponding to the model uncertainty drops from 113 nm without prior misalignment correction to 77 nm when the prior correction of the topography position is included.

4 Comparison

The reconstruction results from the deep learning hybrid approach are compared to the results produced by a state-of-the-art method for optical form measurements, namely the tilted-wave interferometer distributed by Mahr GmbH

(MarOpto TWI 60) [34]. The comparison procedure is similar to the one used in recent interlaboratory comparison studies [3, 4]. Again, all processing steps are carried out using the simulation toolbox SimOptDevice [17].

4.1 Comparison procedure

In a first step, the point clouds of the reconstructed specimens need to be transformed into a format that is comparable. To this end, the data are restricted to the largest common aperture radius, and the point clouds are aligned in the Cartesian coordinate system by fitting the reconstructed topographies into the design topography in a least-squares sense. This then allows the estimated difference topographies to be compared after subtracting the design from the reconstructed specimens. An additional step is included to subtract the best fit spheres in a least-squares sense using the Levenberg-Marquardt algorithm. Comparing the results while neglecting the dominating spherical part can be of interest because spherical contri-

Table 1: The root-mean-squared deviations (RMSD) of the reconstructed difference topographies, before and after subtracting the best fit sphere (BFS), are presented for the state-of-the-art method as well as for the deep learning based hybrid approach without (hybrid method 1) and with prior misalignment correction (hybrid method 2). The fourth row shows the radii of the best fit spheres, where the mean is taken over the ensemble members for the hybrid methods.

	state-of-the-art	hybrid method 1	hybrid method 2
RMSD (in nm)	489	494	476
RMSD after subtracting the BFS (in nm)	239	192	211
radius of BFS (in m)	59.7	54.7	58

Table 2: The first row shows the root-mean-squared error (RMSE) between the reconstructed difference topographies from the state-of-the-art method and the deep learning hybrid approach before and after subtracting the best fit sphere (BFS). Hybrid method 1 refers to the deep learning hybrid approach without prior misalignment correction, while hybrid method 2 includes it. The second row presents the root-mean variances and the third row shows how much of the reconstructed form from the state-of-the-art method is covered by the 95% symmetric credible interval of the deep learning results.

	before subtracting the BFS		after subtracting the BFS	
	hybrid method 1	hybrid method 2	hybrid method 1	hybrid method 2
RMSE (in nm)	135	81	133	82
root-mean variance (in nm)	113	77	107	70
coverage (in %)	92	93	91	89

butions may arise from a remaining misalignment of the topographies along the optical axis [41]. In addition, the spherical form error is not of interest for some applications because it can sometimes be compensated if an optical system is mounted in a certain way. The steps of the comparison procedure are summarized as follows:

1. Restrict the data of the reconstructed specimens to the largest common aperture radius.
2. Rotate and shift point clouds of the reconstructed specimens into the design.
3. Subtract the design topography from the reconstructed specimens and repeat step 1.
4. Subtract the respective best fit spheres from the reconstructed difference topographies.

The results are compared after steps 3 and 4. The reconstructed topography and the model uncertainty of the deep learning model are propagated by performing each step of the comparison procedure for the individual ensemble member predictions, and by taking the mean and uncorrected sample variance only afterwards. The achieved resolution of the deep learning hybrid approach (64×64) is much smaller than the resolution of the measured data on the camera (2048×2048). Therefore, Zernike polynomials are fitted to the reconstructed difference topographies up to the first 78 coefficients before plotting the results. The chosen polynomial degree could be higher or lower, but the chosen order ($n, m = 11$) [36] is sufficient for

the purpose of this work. In this way, only the lower spatial frequencies of the form reconstruction are compared while the higher frequency structures are neglected. The presented results are produced by evaluating the Zernike polynomials on a fixed equidistant grid.

It is difficult to compare the computing efficiency of the two approaches due to their different requirements. Once the deep neural networks are trained for a design topography, and the measurement data of the specific specimen are acquired and preprocessed, the form reconstruction and the described uncertainty quantification are performed instantaneously. In contrast, state-of-the-art TWI methods do not need any training for different design topographies. However, their form reconstruction time depends on different hyperparameters, such as the chosen resolution, the amount of rays to be traced through the optical system, and the number of steps used for the iterative form reconstruction procedure [9].

4.2 Results

A summary of the results is displayed in Tables 1 and 2. The reconstructed difference topographies (the results obtained after step 3) are shown in Figure 4 for the deep learning hybrid approach with and without prior misalignment correction. For comparison, the reconstructed difference topography from the state-of-the-art method is depicted in Figure 5a. The vertical (left) and horizontal (right)

Figure 5: **a)** The reconstructed difference topography is plotted for the state-of-the-art method for optical form measurement. **b)** Result from **a)** after subtracting the best fit sphere. **c)** Result from Fig. 4a) after subtracting the best fit sphere. **d)** Result from Fig. 4c) after subtracting the best fit sphere.

Figure 6: The vertical (left) and horizontal (right) cross sections of the reconstructed difference topographies plotted for the state-of-the-art TWI method (red) and the deep learning hybrid approach (blue) together with the latter's model uncertainty, i. e., the 95% symmetric credible interval: **a)** without prior misalignment correction; **b)** after subtracting the best fit sphere; and **c)** with prior misalignment correction; **d)** after subtracting the best fit sphere.

cross sections of the reconstructed topographies are directly compared in Figures 6a and 6c, respectively, with and without the misalignment correction prior to applying the trained deep learning model. Most of the reconstructed surface from the state-of-the-art method (red) lies in the 95 % symmetric credible interval of the deep learning results (blue). More precisely, 92 % of the reconstructed form from the state-of-the-art method is covered by the described model uncertainty of the deep learning prediction when no prior misalignment correction is carried out, while 93 % of the topography is covered when the prior correction of the topography position is included. The root-mean-square deviation of the reconstructed difference topography is 489 nm for the state-of-the-art method, 476 nm for the deep learning method with prior misalignment correction, and 494 nm without such correction. The root-mean-square error between the two methods is 81 nm with, and 135 nm without the prior misalignment correction included in the deep learning approach.

Figures 5b–d show the results after subtracting the corresponding best fit spheres (step 4). The result from the state-of-the-art form reconstruction method (Fig. 5b) has a root-mean-square deviation of 239 nm. The subtracted best fit sphere has a radius of 59.7 m. Next, the result from the deep learning hybrid approach without prior misalignment correction (Fig. 5c) has a root-mean-square deviation of 192 nm, and a root-mean-square error of 133 nm compared to the state-of-the-art method, which lies to 91 % in the area covered by the model uncertainty. The subtracted best fit sphere has a mean radius of 54.7 m, ranging from 48.7 m to 60.5 m for the individual ensemble member predictions. Finally, the result from the deep learning hybrid approach with prior misalignment correction (Fig. 5d) has a root-mean-square deviation of 211 nm and a root-mean-square error of 82 nm compared to the state-of-the-art method which it covers to 89 %, and the subtracted best fit sphere has a mean radius of 58 m, ranging from 50.9 m to 66.7 m for the individual ensemble member predictions.

Again, the cross sections of the topography profiles from the state-of-the-art method (red) are plotted together with the deep learning hybrid approach and its model uncertainty (blue) without and with prior misalignment correction, respectively, in Figures 6b and 6d.

The results show that the deep learning hybrid approach produces reliable results for measured data since the state-of-the-art measurement system generates similar results that lie within the model uncertainty range of the deep learning model. We expect the hybrid method to cope successfully also with other kinds of design topographies and specimens which can be measured by the TWI. A more comprehensive comparison study, including such cases, is, however, beyond the scope of this paper and referred to future work.

5 Conclusion

It was shown in [15] that the described deep learning hybrid approach produces reliable results, which also include a quantification of the model uncertainty, for in-domain as well as out-of-distribution test data in a simulated environment. In this paper, we demonstrate how to use the deep learning hybrid approach in a real-world environment and apply it to measurement data from a real specimen of an asphere. Given that the ground truth for real specimens is not known, however, we show that the results are still reliable by comparing the reconstructed topography to the topography measured by a state-of-the-art method for the optical form measurement of optical aspheres and freeform surfaces [34]. High frequency structures are neglected because the deep learning model was trained on, and produces, low resolution images. Also, no high frequencies or positioning errors were included in the generated topographies of the training data. The two methods produce similar results, especially when a prior misalignment correction is performed for the input data of the deep learning model. Furthermore, the form reconstructed by the state-of-the-art method lies well within the estimated model uncertainty of the deep learning hybrid approach.

Acknowledgment: The authors thank the company Mahr GmbH for providing the Zemax® model of the interferometer and for insights into the procedures of the MarOpto TWI 60.

References

1. J. P. Rolland, M. A. Davies, T. J. Suleski, C. Evans, A. Bauer, J. C. Lambropoulos, and K. Falaggis, "Freeform optics for imaging," *Optica*, vol. 8, pp. 161–176, Feb 2021.
2. R. Bergmans, H. Nieuwenkamp, G. Kok, G. Blobel, H. Nouri, A. Küng, M. Baas, M. Tevoert, G. Baer, and S. Stuerwald, "Comparison of asphere measurements by tactile and optical metrological instruments," *Measurement Science and Technology*, vol. 26, no. 10, p. 105004, 2015.
3. R. Schachtschneider, I. Fortmeier, M. Stavridis, J. Asfour, G. Berger, R. Bergmann, A. Beutler, T. Blümel, H. Klawitter, K. Kubo, et al., "Interlaboratory comparison measurements of aspheres," *Measurement Science and Technology*, vol. 29, no. 5, p. 055010, 2018.
4. I. Fortmeier, R. Schachtschneider, V. Ledl, O. Matousek, J. Siepmann, A. Harsch, R. Beisswanger, Y. Bitou, Y. Kondo, M. Schulz, et al., "Round robin comparison study on the form measurement of optical freeform surfaces," *Journal of the European Optical Society-Rapid Publications*, vol. 16, no. 1, pp. 1–15, 2020.
5. E. Garbusi, C. Pruss, and W. Osten, "Interferometer for precise and flexible asphere testing," *Optics Letters*, vol. 33, no. 24, pp. 2973–2975, 2008.
6. G. B. Baer, J. Schindler, C. Pruss, and W. Osten, "Measurement of aspheres and free-form surfaces with the tilted-wave-interferometer," in *Fringe 2013*, pp. 87–95, Springer, 2014.
7. G. Baer, J. Schindler, C. Pruss, J. Siepmann, and W. Osten, "Calibration of a non-null test interferometer for the measurement of aspheres and free-form surfaces," *Optics express*, vol. 22, no. 25, pp. 31200–31211, 2014.
8. S. Mühlig, J. Siepmann, M. Lotz, S. Jung, J. Schindler, and G. Baer, "Tilted wave interferometer-improved measurement uncertainty," in *58th Ilmenau Scientific Colloquium*, 2014.
9. I. Fortmeier, M. Stavridis, A. Wiegmann, M. Schulz, W. Osten, and C. Elster, "Analytical jacobian and its application to tilted-wave interferometry," *Optics express*, vol. 22, no. 18, pp. 21313–21325, 2014.
10. G. Baer, J. Schindler, J. Siepmann, C. Pruss, W. Osten, and M. Schulz, "Measurement of aspheres and free-form surfaces in a non-null test interferometer: reconstruction of high-frequency errors," in *Optical Measurement Systems for Industrial Inspection VIII*, vol. 8788, p. 878818, International Society for Optics and Photonics, 2013.
11. G. B. Baer, *Ein Beitrag zur Kalibrierung von Nicht-Null-Interferometern zur Vermessung von Asphären und Freiformflächen*. No. 86, Institut für Technische Optik, Universität Stuttgart, 2016.
12. I. Fortmeier, M. Stavridis, C. Elster, and M. Schulz, "Steps towards traceability for an asphere interferometer," in *Optical Measurement Systems for Industrial Inspection X*, vol. 10329, p. 1032939, International Society for Optics and Photonics, 2017.

13. J. Schindler, *Methoden zur selbstkalibrierenden Vermessung von Asphären und Freiformen in der Tilted-Wave-Interferometrie*. No. 105, Institut für Technische Optik, Universität Stuttgart, 2020.
14. L. Hoffmann and C. Elster, “Deep neural networks for computational optical form measurements,” *Journal of Sensors and Sensor Systems*, vol. 9, no. 2, pp. 301–307, 2020.
15. L. Hoffmann, I. Fortmeier, and C. Elster, “Uncertainty quantification by ensemble learning for computational optical form measurements,” *Machine Learning: Science and Technology*, 2021.
16. I. Fortmeier, *Zur Optimierung von Auswerteverfahren für Tilted-Wave Interferometer*. No. 82, Stuttgart: Institut für Technische Optik, Universität Stuttgart, 2016.
17. R. Schachtschneider, M. Stavridis, I. Fortmeier, M. Schulz, and C. Elster, “Simoptdevice: a library for virtual optical experiments,” *Journal of Sensors and Sensor Systems*, vol. 8, no. 1, pp. 105–110, 2019.
18. L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu, and M. Pietikäinen, “Deep learning for generic object detection: A survey,” *International journal of computer vision*, vol. 128, no. 2, pp. 261–318, 2020.
19. G. Barbastathis, A. Ozcan, and G. Situ, “On the use of deep learning for computational imaging,” *Optica*, vol. 6, no. 8, pp. 921–943, 2019.
20. D. W. Otter, J. R. Medina, and J. K. Kalita, “A survey of the usages of deep learning for natural language processing,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 2, pp. 604–624, 2020.
21. S. Kuutti, R. Bowden, Y. Jin, P. Barber, and S. Fallah, “A survey of deep learning applications to autonomous vehicle control,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 2, pp. 712–733, 2020.
22. S. Yuan, Y. Hu, Q. Hao, and S. Zhang, “High-accuracy phase demodulation method compatible to closed fringes in a single-frame interferogram based on deep learning,” *Opt. Express*, vol. 29, pp. 2538–2554, Jan 2021.
23. X. Fei, J. Zhao, H. Zhao, D. Yun, and Y. Zhang, “Deblurring adaptive optics retinal images using deep convolutional neural networks,” *Biomed. Opt. Express*, vol. 8, pp. 5675–5687, Dec 2017.
24. L. Zhang, C. Li, S. Zhou, J. Li, and B. Yu, “Enhanced calibration for freeform surface misalignments in non-null interferometers by convolutional neural network,” *Opt. Express*, vol. 28, pp. 4988–4999, Feb 2020.
25. Y. Nishizaki, M. Valdivia, R. Horisaki, K. Kitaguchi, M. Saito, J. Tanida, and E. Vera, “Deep learning wavefront sensing,” *Opt. Express*, vol. 27, pp. 240–251, Jan 2019.
26. J. Gawlikowski, C. R. N. Tassi, M. Ali, J. Lee, M. Humt, J. Feng, A. Kruspe, R. Triebel, P. Jung, R. Roscher, et al., “A survey of uncertainty in deep neural networks,” *arXiv preprint arXiv:2107.03342*, 2021.
27. W. Samek, G. Montavon, S. Lapuschkin, C. J. Anders, and K.-R. Müller, “Explaining deep neural networks and beyond: A review of methods and applications,” *Proceedings of the IEEE*, vol. 109, no. 3, pp. 247–278, 2021.
28. O. Sagi and L. Rokach, “Ensemble learning: A survey,” *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 8, no. 4, p. e1249, 2018.
29. B. Lakshminarayanan, A. Pritzel, and C. Blundell, *Simple and scalable predictive uncertainty estimation using deep ensembles*, Advances in Neural Information Processing Systems, vol. 30, 2017.
30. L. Hoffmann and C. Elster, “Deep ensembles from a bayesian perspective,” *arXiv preprint arXiv:2105.13283*, 2021.
31. D. P. Kingma, T. Salimans, and M. Welling, “Variational dropout and the local reparameterization trick,” *Advances in neural information processing systems*, vol. 28, pp. 2575–2583, 2015.
32. Y. Gal, J. Hron, and A. Kendall, “Concrete dropout,” in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pp. 3584–3593, 2017.
33. Y. Gal, “Uncertainty in deep learning,” 2016.
34. “MarOpto TWI 60.” <https://www.mahr.de/en-us/Services/Production-metrology/Products/MarOpto---Measuring-Devices-for-Optics-Industry/MarOpto-TWI-60/>. Accessed: 2021-07-26.
35. E. Garbusi and W. Osten, “Perturbation methods in optics: application to the interferometric measurement of surfaces,” *JOSA A*, vol. 26, no. 12, pp. 2538–2549, 2009.
36. J. Cook, “The zernike polynomials,” *Journal of Modern Optics*, vol. 23, no. 8, pp. 679–680, 1976.
37. O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241, Springer, 2015.
38. B. Braunecker, R. Hentschel, and H. J. Tiziani, *Advanced optics using aspherical elements, equation 2.2.2.1*, vol. 173, Spie Press, 2008.
39. P. Hariharan, B. F. Oreb, and T. Eiju, “Digital phase-shifting interferometry: a simple error-compensating phase calculation algorithm,” *Applied optics*, vol. 26, no. 13, pp. 2504–2506, 1987.
40. R. M. Goldstein, H. A. Zebker, and C. L. Werner, “Satellite radar interferometry: Two-dimensional phase unwrapping,” *Radio Science*, vol. 23, no. 4, pp. 713–720, 1988.
41. I. Fortmeier, M. Stavridis, A. Wiegmann, M. Schulz, W. Osten, and C. Elster, “Evaluation of absolute form measurements using a tilted-wave interferometer,” *Optics express*, vol. 24, no. 4, pp. 3393–3404, 2016.

Bionotes

Lara Hoffmann

Physikalisch-Technische Bundesanstalt (PTB), Abbestr. 2-12, 10587 Berlin, Germany
lara.hoffmann@ptb.de

Lara Hoffmann studied mathematics at the TU Berlin and started her PhD at the PTB in 2019. Her work focuses on the intersection between deep learning approaches and computational optical form measurements, as she is part of the working groups “data analysis and measurement uncertainty” and “asphere metrology”.

Ines Fortmeier
Physikalisch-Technische Bundesanstalt
(PTB), Bundesallee 100, 38116
Braunschweig, Germany
ines.fortmeier@ptb.de

Ines Fortmeier currently leads the working group “asphere metrology” at PTB. Her main research interests are in the field of asphere and freeform metrology, optical form measurement methods, measurement comparison methods, the development of reference methods, and new data analysis methods in that field.

Clemens Elster
Physikalisch-Technische Bundesanstalt
(PTB), Abbestr. 2-12, 10587 Berlin, Germany
clemens.elster@ptb.de

Clemens Elster currently leads PTB’s working group “data analysis and measurement uncertainty.” His main topics of interest are statistical data analysis and evaluation of measurement uncertainty.