https://doi.org/10.2352/issn.2694-118X.2021.LIM-83

Impact of Windshield Optical Aberrations on Visual Range Camera Based Classification Tasks Performed by CNNs

Christian Krebs[1], Patrick Müller[2], Alexander Braun[2]
[1]AGP Europe GmbH, Baesweiler, Germany [2]Hochschule Düsseldorf, Düsseldorf, Germany

Abstract

Cameras operating in the visual range of the electromagnetic spectrum are central to advanced driver assistance systems (ADAS). Front cameras, analyzing traffic, are often located behind the windshield to detect and classify objects.Thus, the area of the windshield within the camera's field of view is a part of the optical system. Simple windshields consist of two curved glass surfaces connected by a thermoplastic interlayer. Due to defects present in the raw glass, as well as those introduced during the bending and lamination process, windshields will have optical aberrations. While optical quality may be suitable for human vision, it can fall short of what is needed for machine vision. In this article we investigate how the optical aberrations generated by laminated safety glass (LSG) influence the optical performance of a camera system and based on this, how the classification of image content by a convolutional neural network (CNN) is affected. A method for wavefront measurements of LSG samples is presented, which allows us to parameterize a linear optical model in Zernike Space. From this, we derive space-variant point spread functions (PSFs) and apply those to the dataset to simulate the windshield's impact on the camera image. As a use case, a CNN was trained on the unmodified dataset and compared to the modified versions with the LSG models applied. We measured and modelled two different LSG samples, one with high and the other one with low optical quality. We compare the prediction accuracy of the classification with the unmodified data. The high-quality sample had negligible effect on the overall classification accuracy, while the low-quality sample lowered the prediction accuracy by up to ten percentage points due to the optical aberrations.

Introduction

Advanced Driver Assistance Systems (ADAS) enhance the safety of road users [1]. For example Blind Spot or Lane Keeping Assist, Collision Warning and Mitigation Systems are believed to have had the potential to prevent up to 1.6 million accidents per year in the U.S. in 2015 [2]. ADAS are therefore a crucial means to make our roads safer and achieve the so-called "vision zero", no more fatalities by auto accidents. Therefore, as the field of ADAS grows, assistants such as emergency lane keeping and collision, speed, and backup warning systems for pedestrians and cyclists have been added to regulations such as EU Regulation 2019/2144 [3]. Further, ADAS are often based on the interaction of different sensor types, where camera systems operating in the visual range of the electromagnetic spectrum are crucial. Digital camera systems are continuously being improved. Inexpensive and readily available, they are flexible [4] and can be combined with other sensor systems. These characteristics allow for their use in assisted as well as automated driving. Furthermore, cameras for traffic monitoring are typically placed behind the windshield in a central location close to the top and near the transition to the roof edge. This ensures good viewing angles, and the system is protected by the windshield. The vehicle occupants are protected by the laminated safety glass (LSG) windshield, usually consisting of two layers of glass and a thermoplastic bonding layer.

However, windshields are not an optically ideal for machine vision. Aberrations can arise from the intrinsic material properties and defects present in the raw glass, the bonding interlayer, and coatings. The bending and lamination process also create optical defects. Furthermore, optical errors can be intensified by low camera installation angles. Therefore, it is necessary to recognize and assess a performance drop in the corresponding ADAS caused by the presence of the windshield [5]. One of the main requirements for windshields is currently the refractive power specified in milli diopters [mdpt]. Besides the refractive power, further optical requirements like transmission, haze and double image exist. Zone definitions, some optical requirements and regulations for windshields are specified [6, 7]. Conversely, camera requirements, as an example, are specified in terms of spatial frequency response (SFR) as defined in ISO 12233 [8]. The parameters that camera specifications share with automotive glazing regulations, such as zone definition, filter parameters and limit values, differ from the standard windshield requirements for human drivers and OEM's specifications for maximum optical refractive power, among others. Nevertheless, the main goal of both is the same: flawless operation of the ADAS must always be ensured. The windshield should have no influence on its performance. However, a reliable quality assessment of an ADAS placed behind different windshields would require several hundred million kilometers [9] of real driving tests, which is not practical. Therefore, simulations for ADAS offer many advantages [10]. Approaches to model-based safety analysis [11] and suggestions regarding the design, development of tests and validation of ADAS [12] can be found in recent publications. However, the influence of the windshield is often ignored in this context or described by overly simple representations, e.g. flat monolithic glasses, which are ideal.In this study, a novel approach for the measurement of a LSG, the simulation of its optical aberration, and the influence on a classification task performed by CNNs is presented. The main benefit is to have an optical model of a physical LSG, which can be validated with lab images or real driving tests. The model can be applied to degrade an image database to generalize and evaluate the influence on ADAS tasks. Following this approach, the relationship between measured aberrations, quality requirements and actual ADAS performance can be investigated.

This article is structured as follows: first, we introduce the example application chosen for our investigation, a CNN for street sign classification. The database, the CNN's architecture and prediction results are briefly described. This is followed by an explanation of how we create our optical model of a LSG. The laboratory setup is described, and we outline the wavefront measurements for a LSG followed by the linear optical model, which

we apply to a database. Finally, we show the windshield's effect on the classification of street signs.

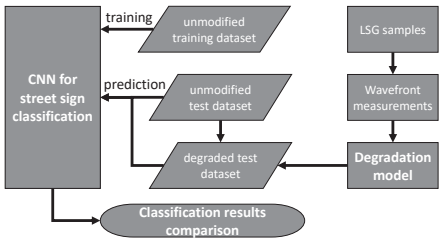


Figure 1. *The core concept: Comparing classification results of a CNN for street sign classification for an unmodified image dataset and a modified version. The modified dataset simulates the influence of a camera system capturing images through a windshield. The model is parameterized by real wavefront measurements of LSG samples. The CNN is trained on the unmodified image dataset.*

A CNN for street sign classification

One out of the many different use cases of an ADAS is the recognition of objects in a driving scene. In this article, we examine the classification accuracy of street signs by a CNN due to optical errors. Aberrations in the windshield can lower the contrast, blur, and distort the objects which need to be classified. The benefit of selecting this task is based on the properties of the street signs. Street signs appear in a moderate range of sizes in images, have distinct classes as well as high- and low-level details ranging from simple shapes to more detailed pictographs. We use the *German Traffic Sign Recognition Benchmark* (GTSRB) from the Institute of Neuroinformatic at the Ruhr University in Bochum, Germany [13]. The GTSRB consists of a large number of training $(39,209)$ and test images $(12,630)$ divided into forty-three classes. The relative number of samples per class in the training and test dataset is similar. Data augmentation is used by applying an additional translation variation of up to $\pm 10\%$ and rotation variation of up to $\pm 10°$ to the complete dataset. A *Contrast Limited Adaptive Histogram Equalization* (CLAHE) [14] is used to preprocess the images.

Architecture and Training

The presented CNN is chosen due to two reasons: first, it is designed to achieve an average classification accuracy of street signs similar to the average human performance of 98.84% [13]. Second, we aimed for a network which is potentially sensitive for the predicted aberrations to create a better understanding of their influence to the performance. The CNN consists of an *input layer* of size $[49 \times 49 \times 3]$ and three *convolutional layers* of size $[43 \times 43 \times 120]$, $[17 \times 17 \times 180]$ and $[6 \times 6 \times 260]$ each with *max pooling*. A *dropout* of 0.5 and *softmax* is used in the final *dense layer*. The training data of the GTSRB database is split into subsets for training and validation with a 15% validation split. It is trained using *stochastic gradient descent* (SGD) on a batch size of ten images per batch over thirty epochs. The *accuracy* and *model loss* are measured after each epoch by evaluating the *Sparse categorial cross entropy*. The presented CNN's overall predicting accuracy is 98.61%.

Image degradation based on an optical model for laminated safety glass

The windshield's impact on an image database was simulated by means of an optical model based on linear system theory and measurement data. The optical data was passed to the model from lab measurements with a Shack-Hartmann wavefront sensor for two different real LSG samples. First, we will describe the LSGs and the required lab setup used to sample the windshields' wavefront at multiple points. Second, we will outline the image degradation using the acquired measurement data from the lab.

Laboratory setup and wavefront measurements of LSG

Two LSG samples (*LSG No. 66* and *LSG No. 84*) were selected for comparison, where No. 66 represented a high-quality sample, and No. 84 a low-quality sample. The location of each measurement on the LSG samples was determined by the camera's field of view (FoV), its distance and orientation to the sample. To validate the model, a reference camera system with an 8.3 MP (2168×3848) sensor, 2.1 µm pixel size and a lens with 32.1°/53.7° vertical/horizontal FoV was used. The general

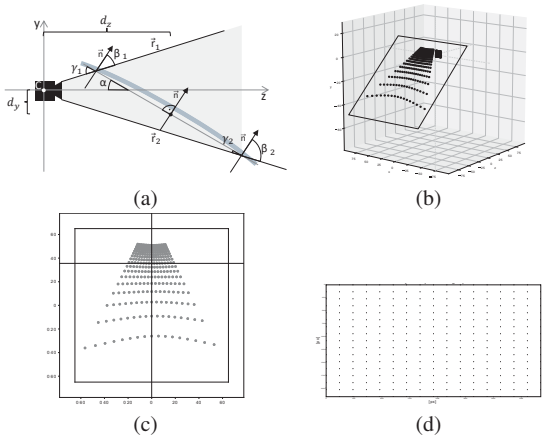


Figure 2. *Measurement setup and sampling: (a) Geometrical setup with reference camera at C. (b-c) Sampling of the tilted LSG with non-regular spaced measurement spots and (d) equidistant spots in the final imager coordinates.*

measurement setup is displayed in Fig. 2a. A windshield was placed within the FoV of the reference camera, where the camera's installation angle α (angle between optical axis z and secant through top and bottom edge of the LSG specimen) was 25°, the horizontal distance d_Z of the camera to the center of the LSG was 35 mm and the vertical distance d_Y of the camera to the center of the LSG sample was 15 mm. The LSG sample size was 130×130 mm^2 with a total thickness of 4.46 mm and consisting of 2.1 mm soda-lime outer glass, 0.76 mm clear PVB (Polyvinylbutyral) layer, 1.6 mm soda-lime inner glass. The nominal vertical radius of curvature (ROC) was 10×10^3 mm, the horizontal ROC was 3×10^3 mm. The LSG's measurement locations as shown in Fig. 2c were obtained from an inverse distortion model, such that all spots were equally distributed in the image space as displayed in Fig. 2d. The PSF was measured at each location at the camera's FoV respective angle. To reduce the number of measurements required, the following simplification and assumptions were made:

- A total of 289 measurements in a 17×17 grid per LSG sample were assumed to be sufficient
- Refraction at both glass surfaces and the interlayer was not considered when calculating the location of the measurement points
- The LSG sample was assumed to be a flat thin surface

- The angle of incidence was only considered for the vertical FoV; furthermore the measurement angles γ_i were reduced to three reference angles 42°, 25° and 8°.
- The mask size diameter for a measurement at γ_i was corrected by the $\sin(\gamma_i)$

For the selected spots the wavefront slopes were measured by a Shack-Hartmann wavefront sensor and output the wavefront in terms of Zernike polynomials. Each measurement point therefore consisted of the coefficient values for the first 24 Fringe Zernike polynomials. In Fig. 3, five are visualized. Zernike Polynomials are important to aberration theory, since each polynomial Z_n^m has its physical meaning, such as Tilt x, vertical Astigmatism and Coma allowing to analyze optical systems, where the amount of the aberration is controlled by its corresponding coefficient A_n^m.

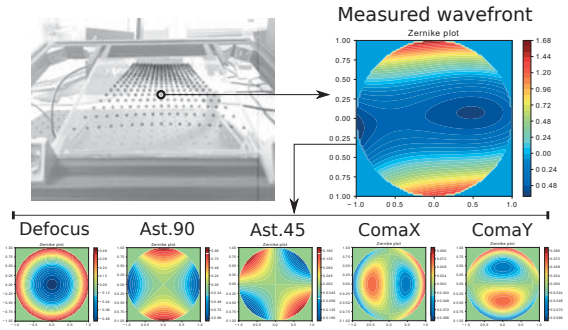


Figure 3. *Wavefront acquisition. The wavefront representing the LSG's optical influence is sampled in 17×17 locations within the reference camera's FoV. Each measurement spot is then decomposed into Zernike polynomials.*

Optical model and image degradation

The effect of the LSGs as a general linear optical system were modeled, and then parameterized with the measurement data. An optical system is described by linear system theory with its characteristic point spread function (PSF). The system's intensity PSF denotes the redistribution of a unit point light source's intensity caused by passing an optical system observed at an observation plane. However, there are some difficulties when measuring the PSF directly, so sometimes there is an advantage when describing the wavefront by other means: Revisiting the diffraction integral, it is also possible to describe this redistribution as phase error in the pupil plane caused by a deformed wavefront [15]. We compute the wavefront error W corresponding to the PSF from a Zernike Polynomials expansion with Zernike coefficients available from the measurements. The Zernike Polynomial expansion is defined on a unit circle in polar coordinates:

$$W(r,\theta) = \sum_{n}^{N} \sum_{m}^{M} A_n^m \cdot Z_n^m(r,\theta) \quad , \tag{1}$$

with Zernike coefficients A_n^m and Z_n^m representing the polynomial depending on pupil coordinates (r,θ) [16]. Since we include measurements from the Shack-Hartmann wavefront sensor outputting Zernike coefficients, we can use these to parametrize our model of a windshield and compute the corresponding PSF at every position in the imager.

In general, this deformed wavefront depends on the viewing angle, and thus the output PSF varies over the image field. We assume that all PSFs can be cropped to a maximum size and evaluate the local PSF by computing the Fourier Transform of the aberrated phase function of the pupil. According to the corresponding measured wavefront for different locations we get monochromatic PSFs, which we scale to the pixel size of the imager, $l1$ normalize and finally apply to an image from the image database. Fig. 4 visualizes the application of the Zernike coefficients to form the specific wavefront error. This two-dimensional function represents the phase function of a complex pupil function, where we assume a circular shape of the pupil. From this, we can compute a PSF including the optical errors caused by the presence of the windshield. We then apply the local PSF to the image as convolution kernel and repeat these steps for all imager locations to form the final degraded image. Note that the degraded image includes the windshield as well as the camera used for capturing the traffic signs. A first validation approach is done by comparing the pixel shift and eSFR of the images captured by the reference camera system through the corresponding LSG sample and the degradation.

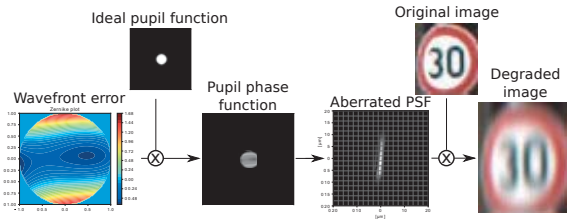


Figure 4. *The LSG's impact on the final image is obtained by transforming the Zernike coefficients into point-spread functions which are then used to degrade the image.*

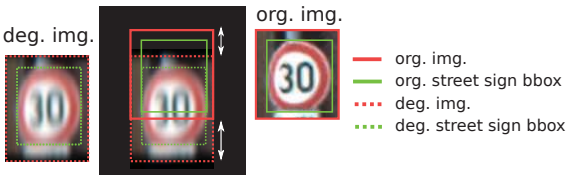
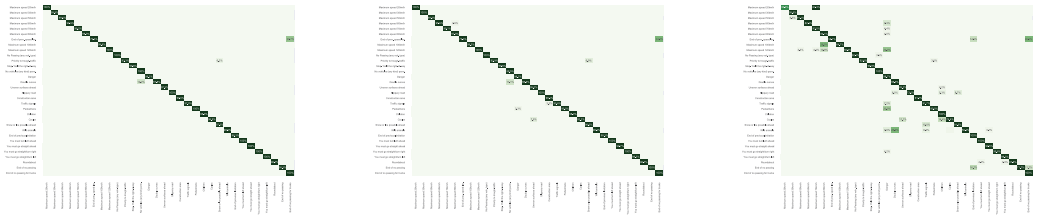


Figure 5. *Image deformation and bounding box (bbox) adjustment.*

The level of degradation is determined by measured and simulated aberrations including local horizontal and vertical pixel shifts, which are caused by the tilt of the glass surfaces. These local shifts deform the image and as a result, sub image areas are stretched or compressed, while e.g. a higher level of defocus and astigmatism blurs the image. Fig. 5 visualizes the modified bounding box for a ground truth image after the application of the degradation model. The original image is shown framed by red solid line on the right-hand side. For each image, a bounding box (solid green line) for the actual street sign is available as ground truth. After applying the degradation, the image (dashed red line) is deformed, and the street sign's bounding box (dashed green line) is adjusted with respect to the simulated pixel shift such that it fits to the degraded sign.

Results: LSG's influence on street sign classification

Table 1 summarizes the classification results for both LSG samples and the original test dataset. The weighted accuracy, precision and recall for each sample and unmodified data (reference) is calculated. As the accuracy loss in LSG No. 66 is lower than 0.5 percentage points compared to the reference dataset, the CNN is not highly affected by the LSG's aberration. In LSG



(a) (b) (c)

Figure 6. *Confusion matrices - classification accuracy for (a) ground truth (b) LSG No. 66 and (c) LSG No. 84 . Note that for better readability only misclassifications above 3% are displayed on a reduced set of classes. Deviations from 100% are due to the reduction.*

Metric	Reference	LSG No. 66	LSG No. 84
Accuracy	98.61%	98.19%	88.54%
Precision	98.77%	98.37%	91.35%
Recall	98.61%	98.19%	88.54%

Table 1. Accuracy, precision and recall (weighted) of unmodified and degraded datasets with different LSG models.

No. 84 the accuracy loss is ten percentage points showing a significant impact on the CNN's performance. Figure 6 shows the confusion matrices for the ground truth (a), LSG No. 66 (b) and LSG No. 84 (c) for thirty-two out of forty-three street signs in which entries greater than 3% are displayed. Figure 7 shows

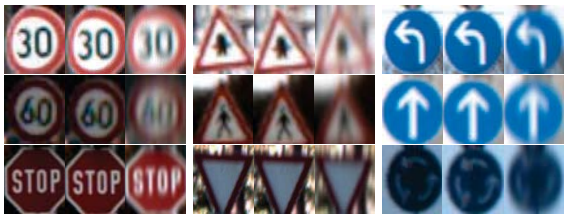
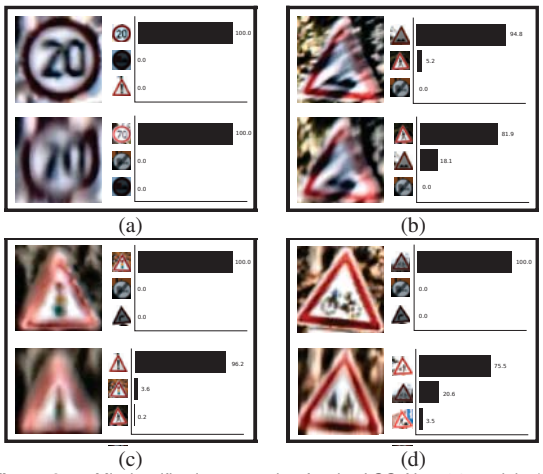


Figure 7. *Comparison of ground truth data (left), samples no. 66 (middle) and no. 84 (right) for different street signs from the GTSRB database. Images are scaled to have equal size.*

the influence on a selected set of street signs. Each street sign is shown in different representations: the unmodified image, the degraded image based on LSG No. 66, and the degraded image based on LSG No. 84. A comparison of the top-3 classification rates for four street signs degraded with the optical model of LSG No. 84 is shown in figure 8. Classification rates for the unmodified images is greater than 94.8% for all signs shown. After the degradation, the CNN predicts the wrong class: *Maximum Speed 20* changes to *Maximum Speed 70*, *Traffic Signal* changes to *Danger*, *Uneven Surfaces Ahead* changes to *Pedestrians*, and *Cycles* changes to *Children*. For a deeper analysis of the accuracy loss in LSG No. 84, the misclassifications are divided into their classes. Increased misclassifications are observed in sets of signs in which the signs are of similar color scheme and shape with different pictographs in the center e.g., speed signs or warnings. However, many street signs are misclassified as the *Danger* sign. If the image content in the center is highly degraded by the aberrations, there is a higher probability that the classification will fail. First investigations lead to the conclusion that similarity between signs and smaller sizes are a reason for these increased misclassifications. A correlation between the false prediction and mentioned attributes are not unexpected, however, it can be shown that aberrations introduced by the windshield can cause this level of degradation and can therefore have a noticeable impact on the performance. As the size of a street sign is connected to its distance, a practical interpretation of this can be summarized as: more distant signs, which are captured as smaller



(a) (b) (c) (d)

Figure 8. *Misclassification examples for the LSG No. 84 model: (a-d) shows the original image (top) and the degraded input image by the LSG model (bottom) passed to the CNN for street sign classification. The achieved top-3 classification scores are shown on the right.*

sub-images by the camera, are more likely to be misclassified, if the windshield introduces strong aberrations. The acceptable level of aberrations and the influence on different ADAS tasks is left for future studies.

Conclusions

An optical model was created for each of the two LSG samples by composing local wavefront measurements, and the derived PSFs were locally applied as a spatially variant convolution. The image degradation based on the optical model was validated by comparing the pixel shift and eSFR for sample images of checkerboard charts. The principal aberrations of the LSG were be simulated with the proposed method. However, further work on the optical model and degradation is needed to improve the fit to real images. The advantage of the presented method is a model based on measurements of a physical sample, which can be verified under real world conditions as well as in test drives. For the low-quality sample, the degradation of the image dataset for traffic sign recognition resulted in a significantly reduced accuracy of the classification by ten percentage points. The loss of accuracy was found in an image region with high Zernike coefficients de-focus and astigmatism 90°, in combination with low resolution of the sub images. Thus, the quality of the windshield as part of the optical system is a non-negligible factor for camera systems for ADAS/AD. Furthermore, this approach is applicable to any other image dataset, to simulate the LSG's effects and investigate their influence on specific tasks.

References

[1] Hermann Winner, Stephan Hakuli, Felix Lotz, and Christina Singer, editors. *Handbook of Driver Assistance Systems*. Springer International Publishing, Cham, 2016.

[2] A. Khan, Corey D. Harper, C. Hendrickson, and C. Samaras. Net-societal and net-private benefits of some existing vehicle crash avoidance technologies. *Accident; analysis and prevention*, 2019.

[3] EU. Verordnung (EU) 2019/2144 des Europäischen Parlaments und des Rates vom 27. November 2019 über die Typgenehmigung von Kraftfahrzeugen und Kraftfahrzeuganhängern sowie von Systemen, Bauteilen und selbstständigen technischen Einheiten für diese Fahrzeuge im Hinblick auf ihre allgemeine Sicherheit und den Schutz der Fahrzeuginsassen und von ungeschützten Verkehrsteilnehmern, zur Änderung der Verordnung (EU) 2018/858 des Europäischen Parlaments und des Rates und zur Aufhebung der Verordnungen (EG) Nr. 78/2009, (EG) Nr. 79/2009 und (EG) Nr. 661/2009 des Europäischen Parlaments und des Rates sowie der Verordnungen (EG) Nr. 631/2009, (EU) Nr. 406/2010, (EU) Nr. 672/2010, (EU) Nr. 1003/2010, (EU) Nr. 1005/2010, (EU) Nr. 1008/2010, (EU) Nr. 1009/2010, (EU) Nr. 19/2011, (EU) Nr. 109/2011, (EU) Nr. 458/2011, (EU) Nr. 65/2012, (EU) Nr. 130/2012, (EU) Nr. 347/2012, (EU) Nr. 351/2012, (EU) Nr. 1230/2012 und (EU) 2015/166 der Kommission (Text von Bedeutung für den EWR), December 2019. Legislative Body: CONSIL, EP.

[4] M. Akhlaq, Tarek R. Sheltami, Bo Helgeson, and Elhadi M. Shakshuki. Designing an integrated driver assistance system using image sensors. *Journal of Intelligent Manufacturing*, 23(6):2109–2132, December 2012.

[5] Alexander Hanel, Ludwig Hoegner, and Uwe Stilla. Towards the influence of a car windshield on depth calculation with a stereo camera system. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B5:461–468, June 2016.

[6] ECE. Regelung Nr. 43 der Wirtschaftskommission der Vereinten Nationen für Europa (UN/ECE) — Einheitliche Bedingungen für die Genehmigung der Sicherheitsverglasungswerkstoffe und ihres Einbaus in Fahrzeuge. page 134.

[7] DIN. DIN 52305:1995-06, June 1995.

[8] ISO. ISO 12233:2017(en), Photography — Electronic still picture imaging — Resolution and spatial frequency responses.

[9] Nidhi Kalra and Susan M Paddock. Driving to Safety: How Many Miles of Driving Would It Take to Demonstrate Autonomous Vehicle Reliability? page 15.

[10] C.-H. Yu, Y.-Z. Chen, and I.-C. Kuo. The benefit of Simulation Test Application on the Development of Autonomous Driving System. In *2020 International Automatic Control Conference (CACS)*, pages 1–5, November 2020.

[11] M. Tlig, M. Machin, R. Kerneis, E. Arbaretier, L. Zhao, F. Meurville, and J. Van Frank. Autonomous Driving System : Model Based Safety Analysis. In *2018 48th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN-W)*, pages 2–5, June 2018. ISSN: 2325-6664.

[12] S. Moten, F. Celiberti, M. Grottoli, A. van der Heide, and Y. Lemmens. X-in-the-loop advanced driving simulation platform for the design,development, testing and validation of ADAS. In *2018 IEEE Intelligent Vehicles Symposium (IV)*, pages 1–6, June 2018. ISSN: 1931-0587.

[13] J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel. Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition. *Neural Networks*, 32:323–332, August 2012.

[14] Stephen M Puer, E Philip Amburn, Roberl Cromartie, and Ari Geselowit. Adaptive Histogram Equalization And Its Variations. page 23.

[15] Joseph W. Goodman. *Introduction to Fourier optics*. W.H. Freeman, Macmillan Learning, New York, fourth edition edition, 2017.

[16] Max Born and Emil Wolf. *Principle of Optics*. The Press Syndicate of the University of Cambridge, The Pitt Building, Trumpington Street, Cambridge, United Kingdom, 7 edition, 2003.

Author Biography

Christian-Benjamin Krebs received his B.Eng. and M.Sc. in Media Technology from the TH Köln. He started as Software Developer and R&D Engineer for Image Quality analysis at Image Engineering and is currently working at AGP Europe GmbH as Project Leader in R&D Optics. His work is focused on the measurement, simulation and optimization of the windshield's camera area for ADAS.

Patrick Müller received his B.Eng. in 2016 and his M.Sc. in 2018. His Master's thesis examined the influence of a Point Spread Function Model to Digital Image Processing algorithms. He is currently pursuing his PhD with a focus on the application of optical models to digital images, their validation, performance and correlation with the performance of Computer Vision algorithms.

Alexander Braun received his diploma in physics with a focus on laser fluorescent spectroscopy from the University of Göttingen in 2001. His PhD research in quantum optics was carried out at the University of Hamburg, resulting in a Doctorate from the University of Siegen in 2007. He started working as an optical designer for camera-based ADAS with the company Kostal, and became a Professor of Physics at the University of Applied Sciences in Düsseldorf in 2013, where he now researches optical metrology and optical models for simulation in the context of autonomous driving. He's member of DPG, SPIE and IS&T, participating in norming efforts at IEEE (P2020) and VDI (FA 8.13), and currently serves on the advisory board for the AutoSens conference.