https://doi.org/10.2352/ISSN.2470-1173.2020.16.AVM-149

Simulating tests to test simulation

***Patrick Mueller*†, *Matthias Lehmann*†, *and Alexander Braun*†**
†University of Applied Sciences Düsseldorf; Münsterstraße 156, 40476 Düsseldorf

Abstract

Simulation is an established tool to develop and validate camera systems. The goal of autonomous driving is pushing simulation into a more important and fundamental role for safety, validation and coverage of billions of miles. Realistic camera models are moving more and more into focus, as simulations need to be more then photo-realistic, they need to be physical-realistic, representing the actual camera system onboard the self-driving vehicle in all relevant physical aspects – and this is not only true for cameras, but also for radar and lidar. But when the camera simulations are becoming more and more realistic, how is this realism tested? Actual, physical camera samples are tested in laboratories following norms like ISO12233, EMVA1288 or the developing P2020, with test charts like dead leaves, slanted edge or OECF-charts. In this article we propose to validate the realism of camera simulations by simulating the physical test bench setup, and then comparing the synthetical simulation result with physical results from the real-world test bench using the established normative metrics and KPIs. While this procedure is used sporadically in industrial settings we are not aware of a rigorous presentation of these ideas in the context of realistic camera models for autonomous driving. After the description of the process we give concrete examples for several different measurement setups using MTF and SFR, and show how these can be used to characterize the quality of different camera models.

Introduction

Because the maturity of detection algorithms for autonomous driving is increasing, the need for physical-realistic camera simulation becomes more important as well. In the last year especially (even though some results have been known for years) both academia and the automotive industry are increasingly taking physical effects into account. Camera sensors are modelled following EMVA1288[3], and modulation transfer function (MTF) curves are applied to simulate optical aberrations[1]. In our group we've modelled the point spread function of a measured lens, including all physical-optical aberrations[6, 7]. Every stakeholder devises its own quality criteria with which these camera simulations are optimized and measured. What is missing is a generally applicable framework that allows for an independent evaluation of those camera simulations.

In a somewhat larger context it has also become clear over the last year that every single detection algorithm has to be treated as an individual observer with distinct likes and dislikes. Tuning and optimization (lens and sensor selection, ISP tuning) need not only be tailored to the Viewing/Computer Vision dichotomy, but actually on the CV side every algorithm is its own category[8]. End-to-end optimization is an often aspired-for goal, bypassing image quality altogether by looking only at final system performance. This is the right goal, but not always applicable, and thus our work highlights the need for a clear definition and measurement of image quality as a conceptual framework, such that the requirements of individual or classes of detection algorithms can be characterized. In other words: having a clear-defined test to quantify the sensor simulations we are able in simulation alone to both see what level of physical realism is needed, and at the same time define those properties of image quality that are relevant for different groups of detection algorithms.

Our idea is symbolized in Fig. 1, where we compare a real image with a simulated image. This resembles the current practice of comparing rendering results from a graphics engine with real-world recordings of the same or similar scenes. The goal of the simulated image is of course to match the real image as closely as possible. In part (a) a real image is shown, taken with an industry camera and an off-the-shelf lens. In (b) we show a simulation, where a ground-truth image is degraded with

(a) Real camera

(b) Simulation

Figure 1: Two images of the same scene. a) Real image, taken with an industry camera, b) Simulation based on a high-end DSLR image as ground-truth with numerical degradation that reflects an optical model .

a numerical optical model to imprint the optical properties onto the image. The ground-truth image could be either a rendered synthetic scene from a graphics engine and a virtual camera model, or a high-quality DSLR camera shot of the same scene. In (b) it is the latter, while for remainder of this publication the simulated images will be completely synthetic.

Differences in the images are immediately apparent. The simulated image appears sharper, and the dynamic range does not match exactly. Both field of view and distortion are also not aligned perfectly. This is the same as in any comparison of drive scene simulations with its real counterpart. But the question we are interested in here is: how much *exactly* do the images match, how much do they differ? Only looking at these images is not good enough to quantify theses differences. Therefore, we propose to use the same method that is used on real-world cameras during production and calibration and use these methods on the synthetic images as well.

Conceptual framework to test camera simulations

A real-world camera for automotive use cases – both ADAS and autonomous driving – undergoes many different tests during development, production and calibration. Well-known examples include the intrinsic and extrinsic calibration of the distortion, or the measurement of the resolution (MTF or SFR) during alignment. The camera module is measured, the results are compared to the specification, and the camera is delivered and used if it is within specification.

Every simulation includes a camera model, from the simple but frequently used pinhole camera to more elaborate models that include distortion, vignetting and sensor noise. A camera model is a basic requirement of image formation, regardless of z-buffer style graphics or raytracing engines. But all these effects, even if they are well done and look quite real, lack an essential property: They do not represent the actual camera that is used in the real-world. As an example relevant to this paper, the resolution of the simulations camera model is almost never defined in terms of the MTF or SFR curve of the real camera system. In most cases it is completely ignored, sometimes a Gaussian blur filter degrades the images. Both approaches are inadequate in representing real-world optical properties, as will be detailed in the results.

Our proposal therefore is to reproduce in simulation the exact same test stands that were used to qualify the real-world camera. Applying the same evaluation metric used during production and testing we can then compare the test results of the real-world camera with the results of the simulated camera, and hence conclude and *quantify* from this comparison how real the simulation really is. In this article we will concentrate on the SFR measurement as an example, but the proposed framework applies to all relevant optical properties of real camera systems. Focal length, field-of-view, distortion, vignetting, sensor noise, sensor MTF and even the ISP performance are just a sample of all the properties of camera system that are both measured and specified in real life.

The procedure is as follows: first, a real camera is measured in a laboratory following a well-defined norm, in this case ISO12233. Second, the vector graphics that served as template for the lab charts is rasterized and used as input for the simulation. The simulation then applies an optical model to this input image to numerically simulate some form of optical aberrations. Finally, a software algorithm determines the SFR curves from the images of both the lab test and the numerical simulation. These results are compared to quantify the likeness or differences between the real world and the simulation.

Simulating ISO12233

The ISO12233 norm defines measurements of the resolution of a camera system as SFR (spatial frequency response) evaluated by analyzing specific features e.g. slanted edges or Siemens stars [5]. The automotive industry has been using this norm for decades, using these tests during development and as end-of-line (EOL) tests to check against specification limits. The standard itself has undergone several important changes in the last twenty years.

Today three main targets are in use: The (harmonic) Siemens star, the slanted edge target, and the Dead Leaves Pattern (also known as Spilled Coins) for texture loss analysis [2]. In this publication we concentrate on the 3×3 Siemens star chart, where nine Siemens stars are arranged in a matrix pattern of three rows and columns. This allows for checking the spatially-varying SFR over field, which we find is one of the most often overlooked yet most important aspects of a physical-realistic simulation.

Laboratory measurement

For the real-world tests we used the testing laboratory of the company Image Engineering. The laboratory comes with three blackly painted walls and a molleton separated section of the room to prevent reflection and allows for absorbing of all light reaching the room. Further, the room is air-conditioned to provide stable and adjustable temperature conditions during measurements. Within the room there is a magnetic blackboard where the target can be placed. We used several different reflective targets (Siemens star, slanted edge, Dead Leaves), and will present results for the 3×3 Siemens star measurement. Further, a rail forces the tripod mounting of the camera to be within a specified distance and orthogonal to the placed target. The remaining axes are equalized by built-in spirit levels at the tripod. To ensure homogeneous illumination and reproducible conditions during the measurements the illuminance is measured at multiple positions on the target.

The device unter test (DUT) in this case was an *Allied Vision Manta G201* industry camera body, with a CCD sensor with a pixel size of 4.4 µm, and which has at $1624(\mathrm{H}) \times 1234(\mathrm{V})$ pixels a size of 7.15×5.43 mm. The objective lens is a *Kowa LM6NCL* and was mounted by a C-mount flange. F-number was 2.0. The focal length of $f = 6\,\mathrm{mm}$ yielded a field of view (FoV) of $61.6°$ and $48.7°$ respectively.

Using this setup we measured a through-focus sweep by manually adjusting the focus in approximately 10 µm steps, while at each position ten measurements were taken to denoise the SFR by averaging.

Simulating the measurement

The following simulations were performed in Matlab, using both established toolboxes as well as our own programming. Starting with a high resolution synthetic test chart we re-scale the image to match the resolution from the measurement in the laboratory, and match the position of the chart within the image. Since the algorithm calculates PSFs according to its position within the image, it is important to have the synthetic chart well aligned to the captures from the laboratory. The resulting image provides the input for the following simulations of optical aberrations.

First, we used a simple distortion model from Matlab to warp the images. In a second step the intensity was scaled over field by a $\cos^4$ fall-off to simulate vignetting. Finally, the image was blurred by one of three different optical point spread function (PSF) models. This step represents the actual quality we wanted to investigate with the SFR measurements, as these optical models simulate the blur due to limited resolution of the optical system. It is here where we simulate the test to test the simulation, as the SFR measurement allows us to tell the quality of one optical simulation model from another.

In this article, we use three different PSF models and apply them by superposition on the synthetic

Siemens test chart.

The first model is a simple rotationally symmetric Gaussian blur, assumed to be constant over the whole field of the image.

The second model still uses a rotationally symmetric Gaussian blur kernel, but improves on the first model by including a spatial variance. The width of the Gaussian kernel is increased going from the center of the image towards the edge of the image. In this case we used PSFs of size $[21 \times 21]$.

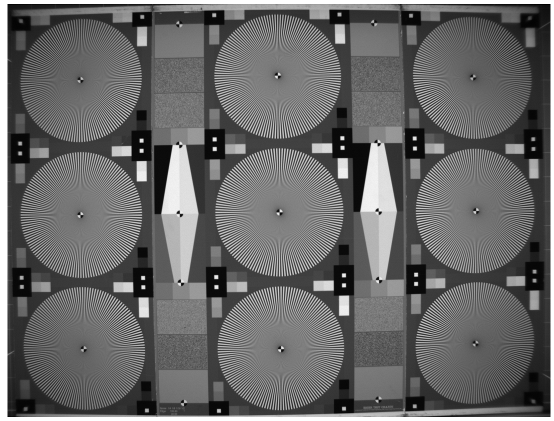
The third model is inspired by real optical aberrations that can be analytically described by the so-called *Zernike* polynomials[9]. We simulate a wavefront aberration as phase error in the pupil plane, which then is mapped by Fourier transform into the image space yielding the PSF. Using a finite number of distinct aberrations (defocus, coma, astigmatism, spherical aberration) we vary the value of the respective coefficients over field, mixing the different aberrations in different compositions to arrive at a realistic looking PSF. We rely on experience and measurements of real lenses to approximate sensible PSFs for this case. The goal of a numerical model for a real, measured lens as proposed in [6] and [7] is left for future work.

Results

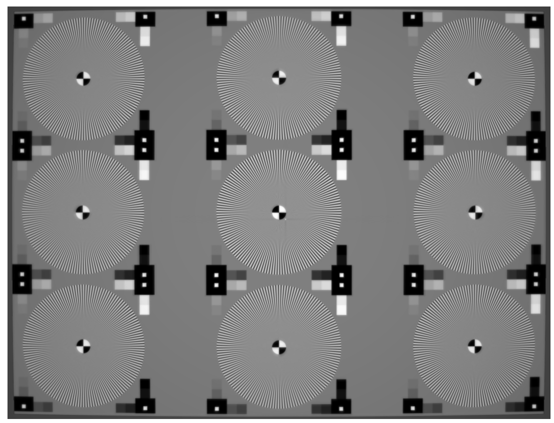
Figure 2a shows a capture of the modulated Siemens Star Target (TE253 9x) from the test laboratory with a slightly defocused objective lens, and fig. 2b a simulation based on a synthetic Siemens star chart blurred with the Zernike model. The two other images for the Gaussian and Gaussian variable model were omitted here for brevity, but are included in the detailed analysis. All test charts have nine harmonic Siemens stars with 144 cycles and OECF-patches. The simulations includes barrel distortion and illumination falloff ($\cos^4$-law).

Visual analysis

Both test images in fig. 2 overall have a similar sharpness, where the blur increases from the center towards the edges of the image. While the blur model is only dependent on the radius, the real camera breaks the radial symmetry due to production tolerances, and produces an image which looks slightly sharper at the very left corners than at both right corners. Thus, the real camera blur depends locally on both azimuth and radius, but mainly on radius.



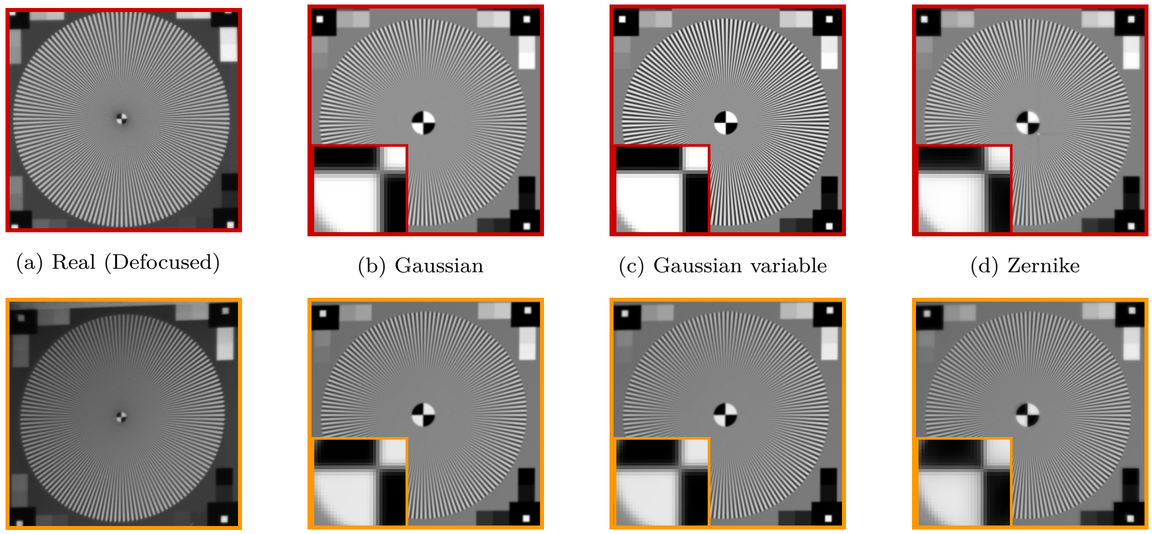
(a) Real blurred



(b) Zernike blurred

Figure 2: Comparison of laboratory measurement and simulation.

The upper and lower row in fig. 3 zoom into the different versions of the star 0 (center) and star 4 (upper left corner) respectively for the measurement (left column) and the three optical models (columns two to four). The insets further zoom into the checkerboard circle in the center of each Siemens star. The sharpness in the center of the image (upper row) is always higher than in the upper left corner (lower row), as is expected. The center stars all exhibit a similar sharpness, except for the Gaussian variable center star (Fig. 3c) which appears with a slightly higher contrast than the others. The stars in the upper left corner (lower row) all have a very similar sharpness. The insets here show another difference of the optical models.



(a) Real (Defocused) (b) Gaussian (c) Gaussian variable (d) Zernike

(e) Real (Defocused) (f) Gaussian (g) Gaussian variable (h) Zernike

Figure 3: Harmonic Siemens Star from the center (upper row) and the upper left corner (lower row).

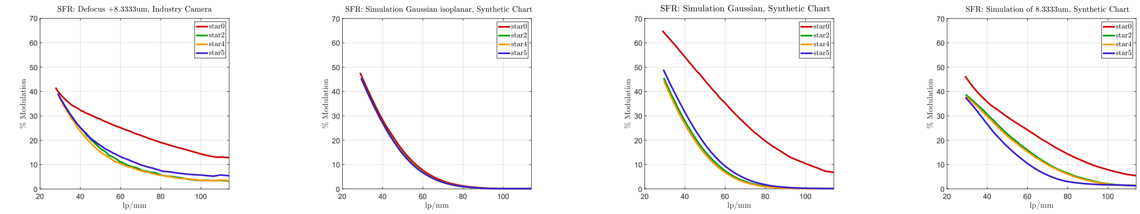
Whereas the two Gaussian PSF models yield rotationally symmetric blur close to the center cross, the Zernike model has more blur in the x-direction than in y-direction, as would be expected from the asymmetric blur kernel.

SFR analysis

However, such a visual analysis remains subjective and time-consuming when comparing several ROIs, or unfeasible when examining the display of differently structured sizes. It is here where our proposal to simulate the tests provides a new approach. Because the simulated scene is the test scene – i.e. here the Siemens Star target – we can now move on to determine the SFR curves from the simulated images. These SFR curves provide objective, quantitative and relevant results, when judging resolution as a key image quality parameter. They allow for a quantitative comparison between the real camera system and the different optical models, thus enabling us to select the best optical model based on measured features, and not on looks alone. The MTFs presented in the following were obtained with the iQ-Analyzer Version 6.1.9[4] in semi-automatic mode to ensure valid star feature detection, and with applied distortion correction.

Figure 4 shows several SFRs from the the real defocused camera sample and from the three simulations. Each plot displays SFRs from four different analyzed ROIs: The center star (star 0), star 2 (upper right corner), star 4 (upper left corner) and star 5 (center left column). Each curve is the average across all segments of a single Siemens star and the unit is in lp/mm. The Nyquist frequency at 114 lp/mm refers to the 4.4 µm pixel width of the camera's image sensor. All plots only start at approximately 30 lp/mm, as per usual due to the finite nature of the Siemens star.

The real camera sample in fig. 4a displays the expected behaviour, where the resolution in the corners of the image (green, yellow and blue curve) is much lower than in the center of the image (red curve). Looking at the SFRs from the simulated images several aspects stand out. First, the SFRs of the Gaussian isoplanar simulation in fig. 4b clearly show the isoplanar behaviour, as all four curves coincide, whereas the spatially variable blur models in fig. 4c and fig. 4d reproduce partly the behaviour of the real lens: The center is distinctly sharper than the corners of the image. The slightly sharper appearance of Fig. 3c is corroborated by the SFR in fig. 4c, where the red curve is higher than the red curve of the real camera almost over the entire frequency range.



(a) Real (Defocused) (b) Isoplanar Gaussian (c) Gaussian variable (d) Zernike Simulation

Figure 4: SFRs: Defocused camera and simulations using a synthetic chart.

Second, from both Gaussian models it is apparent that the form of the curves deviates from the forms of the real camera SFRs. The values at 40 lp/mm are too high, the values at 80 lp/mm are too low. Accordingly, both the slope at lower frequency as well as the tail-off to higher frequency does not qualitatively reproduce the behaviour of the real camera. Here, the Zernike model does a better job of approximating the real camera system.

Finally, note that – as mentioned before – the curves are mean values over all directions, and therefore the astigmatism presented in Fig. 3h is not visible in these curves. Nonetheless, it is expected that an according numeric evaluation within sections of the stars or alternatively with a slanted-edge target at two distinct orientations will readily exhibit these effects.

Discussion

Our proposed conceptual framework to simulate the tests that are actually used on real-world cameras clearly shows the ability to quantitatively evaluate different camera models for simulation, thus testing the simulation itself. We have demonstrated this ability using the established ISO12233 norm for resolution measurement of camera systems and three different optical blur models. Every model is readily distinguishable by its resulting SFR curves. The two key take-aways from the SFR curves in fig. 4 are:

1. Any optical model should include spatially variance of the blur kernel.
2. Rotationally symmetric Gaussian blur kernels neither quantitatively nor qualitatively reproduce optical systems.

Our use of the ISO12233 norm and the selection of the 3×3 Siemens star chart are an example only in that *every* test performed with real camera systems during development and production is a suitable candidate for our framework. Other charts for SFR measurements like Slanted Edge or Dead Leaves can be used, as well as completely different optical properties like distortion calibration or sensor noise models. It doesn't matter which metric is selected – what matters is that numerical simulations of drive scenes for ADAS and AD are evaluated in exactly the same way that their real counterparts are tested. It is a current industry trend to move more and more test drives from the real world into simulation. To ensure safety and to reliably validate the camera systems using those simulation the standard for the numerical simulations certainly should be not lower than that for the real systems themselves.

Acknowledgements

Thanks go to Image Engineering GmbH & Co. KG, Kerpen-Horrem, Germany for the access to their measurement laboratory and software.

References

[1] Rogier van Aken. "Validation of a Physics Based Camera Simulation." In: *AutoSens Brussels*. 2019.

[2] Uwe Artmann. "Image quality assessment using the dead leaves target: experience with the latest approach and further investigations." In: *Digital Photography XI*. Vol. 9404. International Society for Optics and Photonics. 2015, 94040J.

[3] Joyce E Farrell et al. "A simulation tool for evaluating digital camera image quality." In: *Image Quality and System Performance*. Vol. 5294. International Society for Optics and Photonics. 2003, pp. 124–131.

[4] *Image Engineering - Professional Image Quality Analysis*. URL: `https://www.image-engineering.de/products/software`.

[5] *ISO12233:2017, Photography — Electronic still picture imaging — Resolution and spatial frequency responses*. Standard. Geneva, CH: International Organization for Standardization, 2017.

[6] Matthias Lehmann et al. "Modeling realistic optical aberrations to reuse existing drive scene recordings for autonomous driving validation." In: *Journal of Electronic Imaging* 28 (Jan. 2019).

[7] Matthias Lehmann et al. "Resolution and accuracy of nonlinear regression of point spread function with artificial neural networks." In: *Optical Engineering* 58 (Apr. 2019).

[8] Ethan Tseng et al. "Hyperparameter Optimization in Black-Box Image Processing Using Differentiable Proxies." In: *ACM Trans. Graph.* 38.4 (July 2019). URL: https://doi.org/10.1145/3306346.3322996.

[9] F. Zernike. "Beugungstheorie des Schneidenverfahrens und seiner verbesserten Form, der Phasenkontrastmethode." In: *Physica* 1.7 (1934), pp. 689–704. URL: http://www.sciencedirect.com/science/article/pii/S0031891434802595.

Author Biography

Patrick Mueller received his B.Eng. in sound and video engineering in 2016 and his M.Sc. in Electrical and Information Engineering in 2018 from the University of Applied Sciences in Düsseldorf. In his master thesis he applied image quality metrics from optical design to quantify different implementations of space-variant image degradation and restoration algorithms. He is currently pursuing his PhD in image quality, validation and examining physical-realistic models of optical systems.

Matthias Lehmann received his B.Eng. in 2015 and his M.Sc. in 2017 from the University of Applied Sciences in Düsseldorf. His bachelors thesis examined the influence of noise on MTF algorithms and measurements. In his master thesis he developed a SiL environment for correlating optical quality with detection performance. He is currently pursuing his PhD with a focus on an universal optical model, and its application to test and validate algorithms for autonomous driving.

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.