https://doi.org/10.2352/ISSN.2470-1173.2018.17.AVM-148

Detection Probabilities: Performance Prediction for Sensors of Autonomous Vehicles

Marc Geese, Ulrich Seger and Alfredo Paolillo
Robert Bosch GmbH, Leonberg, Germany

Abstract

The recent established goal of autonomous driving cars, motivates the discussion about safety relevant performance parameters in the automotive industry. The majority of currently accepted key performance indicators (KPIs) do not allow a good prediction over the system performance along a safety relevant critical effect chain. A breakdown of the functional system down to component and sensor levels makes this KPI problem evident.

We will present a methodology for sensor performance prediction by a probabilistic approach, on the basis of significant critical use cases. As a result the requirement engineering along the effect chain especially for safety relevant processes appears transparent and understandable. Specific examples from the field of image quality will concentrate on the proposal of a new KPI, the ***contrast detection probability (CDP)****. This proposal is currently under discussion within the P2020 work group on automotive image quality and challenges known KPIs such as SNR, especially with respect their effects on automotive use cases.*

Introduction and Overview

In the last few years the market for driver assistance systems has been emerging very fast. Starting with anti-blockage-systems (ABS) for braking, the driver assistance systems spread into other applications of the car like automatic light switching, rain sensors for wiper control and ultrasound for parking assistance. While these and other systems are meanwhile standard for new cars, the development of driver assistance system continued.

Especially with the introduction of video cameras into the cars, the driver assistance systems tackled the human sense of sight as it is used for driving cars. Starting with rear view cameras and display viewing applications, these video based driver assistance systems are able to assist the driver with function like automatic head beam control (including the switching from high to low beam), traffic sign recognition, pedestrian detection, detection of obstacles and lane marking detection.

In all these functions the video camera plays a major role, however, the driver assistance systems come to full potential if various parts of the car's functionality are combined into an *advanced* driver assistance system. For example, the traffic sign recognition can be connected to the automatic speed control to enable the car to adapt the speed limit correctly and automatically. The detection of pedestrians and other cars needs to be connected to the break system to allow an independent automatic emergency breaking if the driver is inattentive. Another example is to connect the lane mark detection to the steering wheel actuators to allow the car follow a marked lane autonomously.

The development from isolated functionalities towards a connected and highly complex system is required for the introduction of systems for a fully autonomous cars. Considering the effect chain of these systems the above described functionality relies on a good signal quality of a video capturing device.

Motivation and Goal

As described, advanced driver assistance systems are based on a complex effect chain for the video signal processing. This effect chain has to make ensure that the created output signal guarantees enough information to fulfill the demanded function in all use cases. For a video based ADAS we have for example to consider the effect chain that is composed by a scene in the world, windshield of the car, camera optics, image sensor and parts of the image signal preparation. The complete chain contributes to store an image into memory that can be used by further algorithms and neural networks.

Analyzing the above mentioned effect chain makes it evident that meaningful key performance indicators (KPIs) need to be specified. In this paper we will show by example for the imaging effect chain what KPIs could look like that allow to specify the components needed for an ADAS systems. The used probabilistic approach that derives a detection probability along the chain is suited to be transferred to other ADAS systems than the imaging chain.

The Problem of Requirement Engineering ADAS

As mentioned above the definition of the components for an ADAS system is a complex procedure that has to analyze a component chain and the components that form it. Two main problems arise in this context:

Cross Domain Communication First, all the components in the effect chain are covering their own specific field of expertise. This is the reason why the formation of separate components makes sense. But further, the indiviudal optimization created in each field of expertise a different language. In order to specify a system that enables an ADAS function all these components need to form a perfect fit, and therefore not only the optics and image sensor manufacturers need to focus on fitting their components to each other, but the whole chain needs to be optimized.

Image Quality Definitions The second problematic point in requirement engineering for video based systems concerns the term *image quality*. While in consumer cameras, image quality is driven by the visual image impression demanded by the customer

base, for driver assistance systems other requirements need to be fulfilled. For example, the final system has to work over a wide temperature range of $-40°C$ to $125°C$. If the image quality fails within this temperature range, the cars functionality will degrade which is either not favorable in case of ADAS functionality, or with respect to autonomous cars, simply not acceptable. On the other side in consumer cameras the image quality can be focused to reach its optimum around the typical use case of room temperature. This is a big difference in the definition of image quality with respect to the temperature requirements.

Another example from this field is the preservation of contrasts. While in the visual domain noisy parts of the image can be corrected and adapted, such a step may yield to catastrophic problems in driver assistance systems. This is based on the fact that the image for an ADAS function is usually not viewed by a human observer, but by a machine or a neural network. For example if a pedestrian in the dark is almost fully covered in noise, an improper image processing may *denoise* the image in that region. If the target is to form a visually pleasing image, the noise may be removed and with it the remaining information of the pedestrian.

Switching off the denoising is however not the solution either. The question to answer is rather, how can we make sure that a system guarantees its functionality in all demanded situations and which KPIs do we need to specify the components? In other words and targeting on the above example: How can we set up a denoising operation in a way that it does not destroy valuable information? Having that question in mind, we can investigate it with the analysis of the critical use cases.

Critical Use Cases Analysis

As mentioned in [6], an ADAS has to be functional in all the demanded use cases. For example, a requirement for a road sign recognition is usually to detect all visible road signs of a certain class when the car is driving past them. In this example the car is benchmarking the human observer. If the roadsign is covered or occluded, the system does not need to see it as the human driver would also have missed it. However, if the sign is in the shadows or misplaced so that its illumination is not in favor for the imaging chain, the system has to detect it. Those situations are covered under the so called *corner use cases*, which means that the fulfillment of those use cases guarantees the fulfillment of all other use cases as well.

Use cases and corner cases can be transferred into *critical use cases* which occur if there are one or more external conditions that challenge the fulfillment of the use case. Such conditions could be: fog, snow, heavy rain, dust on the windscreen and high ambient temperature. For example a system could be good enough to detect a pedestrian at night under low illumination conditions but then the system could start to degrade in performance if the same situation is recorded with a car driving from a sunny park-deck into a parking garage. The reason for the degradation could be that the dark current of the image sensor leads to an image quality degradation. Depending on the severity, either the algorithm needs to be adapted or the information has been erased while transmitting through the imaging chain. Only an investigation along the imaging chain gives insight into the root cause.

Some of these critical use cases conditions are already in the ADAS' typical specifications, however, many of them are still demanded implicitly by very general requirements. Fig. 1 and 2 shows a critical use case due to veiling glare in the scene. For the human observer it is possible to detect the car in both situations, however the image quality provided by this imaging chain with a dusty windscreen makes creates here a challenge.

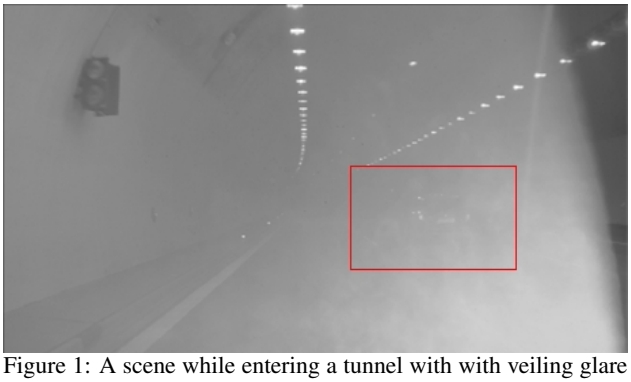


Figure 1: A scene while entering a tunnel with with veiling glare on the windscreen due to dust particles. A human observer may still resolve the car due to a good contrast detection ability of the human eye

Figure 2: The same scene as in fig. 1 but now the veiling glare has disappeared. Inside the video-stream an object has suddenly appeared.

Fig. 3 to 5 shows the contrast loss due to fog in Leonberg, Germany. It is clearly visible that the fog diminishes the ability to detect the objects in $50m$ distance. However for a safe driving of the car, these objects should be detected. Human observers tend to have a superior contrast detection ability compared to the typical imaging chains. Therefore the driver assistance system should be able to resolve the objects even in the heavy fog scene. If a texture analysis is necessary for good algorithmic performance strong fog may become a problem due to the contrast degradation.

And fig. 6 and 7 show a classical detection problem due to a high dynamic range in the scene. With a an imaging chain that does not allow to capture the whole dynamic range, the objects outside the tunnel are not visible and therefore cannot be detected. In this picture the cars outside the tunnel are less than $100m$ away and therefore need to be detected to allow a good ADAS performance. In fig.7 and 8 the same scene is correctly captured with a system capable of handling 140dB of dynamic range. Consequently the scene content is correctly transferred into the image memory.

Given the above examples, it is necessary to analyze the typical use cases with respect to their illuminances and to analyse the relevant objects with respect to their reflectances and contrasts.

Fig. 10 and 9 show recordings with a luminance camera [1]. Here we can observe that the contrast between the border and the center of the traffic sign is expressed as Weber contrast[2]:

$$K_{\text{Weber}} = \frac{E_{\max}}{E_{\min}} - 1 = \frac{\frac{1}{2}(460+900)}{\frac{1}{2}(80+120)} - 1 \approx 580\% \quad (1)$$

The contrast of the cyclist and its background however only:

$$K_{\text{Weber}} = \frac{E_{\max}}{E_{\min}} - 1 = \frac{\frac{1}{2}(49+134)}{\frac{1}{2}(44+100)} - 1 \approx 30\% \quad (2)$$

We can summarize that there are a lot of situations that can be classified either as corner use case, or critical use case in which the imaging chain needs to be tailored in a specific way to achieve the demanded performance. If that step is completed and a database of use cases is available, the problem remains to find meaningful KPIs to describe the image quality in a way that allows to specify the whole imaging chain and its components.

Figure 3: A reference scene showing several cars and traffic lights in a close distance of $\approx 50m$.

Figure 4: The reference scene with typical November fog in Germany, demonstrating the contrast loss of the objects in the box and also a contrast loss of the close by object.

Figure 5: The reference scene with strong fog. The objects in the red box are not detectable any more and even the contrast loss of nearby objects decreased to a level that texture extraction may fail.

A Video Sensing Effect Chain

Before we start to introduce the KPI development, a simplified model of an imaging chain is introduced that describes the components and some of their typical effects onto the image quality. This imaging chain is composed by an illumination, a scene, the windshield of the car, a camera optics, an image sensor and some basic image signal preparation (ISP). Fig. 11 shows an

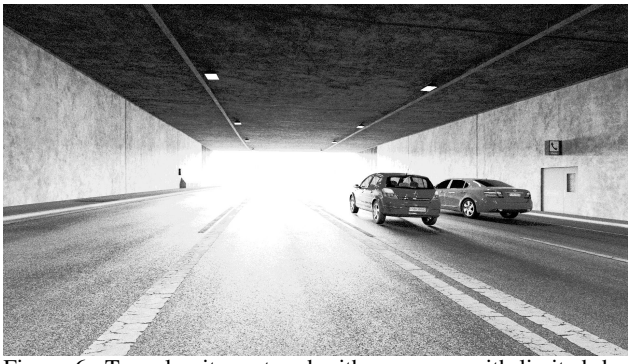


Figure 6: Tunnel exit, captured with a camera with limited dynamic range ability. The objects outside the tunnel are not visible.

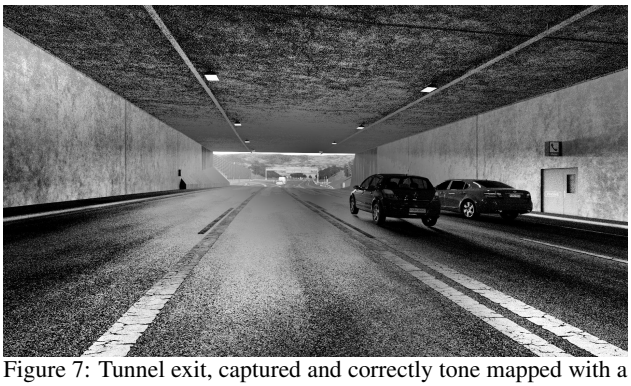


Figure 7: Tunnel exit, captured and correctly tone mapped with a camera with high dynamic range ability. The objects outside the tunnel are clearly visible.

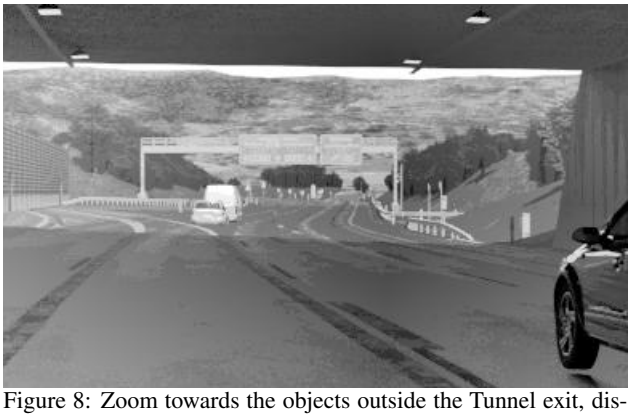


Figure 8: Zoom towards the objects outside the Tunnel exit, distance of the cars from the camera is $\approx 100m$.

overview of this imaging chain and the next subsections introduce how these blocks influencing the signal. The camera part of this imaging chain is developed on basis of the EMVA1288 standard for cameras [3]. Although some of the depicted blocks currently introduce only linear transformations, a simplification shall not take place to allow more elaborate models with nonlinear effects in the future.

Illumination and Scene

The first two blocks of the imaging chain represent the scene and its illumination. A basic property of a surface is its reflectance, which is a constant property that is usually independent from the illumination. Fig. 12a shows the reflectance map of a traffic sign. The reflectances have been deduced from the above analysis conducted with a luminance camera. Fig. 12b shows the corresponding probability function of these reflectances. To simpify the model, Lambertian surfaces are assumed.

Illuminating the scene with $10\frac{cd}{m^2}$ allows to calculate the number of photons that are emitted from the surface. Again, to simplify the model, the light source is limited to $500nm$ photons. Extensions to arbitrary spectral illuminations and also to spectral reflectances can be created by a linear superposition of the individual spectra. Further we assume a luminous efficacy of $L = 1000\frac{\mathrm{lm}}{\mathrm{W}}$ as an average value of daytime and night seeing [5].

Using Planck's constant h and the speed of light c the number of photons per steradians and square meter is given by:

$$1 \cdot \frac{\mathrm{cd}}{\mathrm{m}^2} = 1 \cdot \frac{\mathrm{lm}}{\mathrm{m}^2 \cdot \mathrm{sr}} = 1 \cdot \frac{\mathrm{W}}{L \cdot \mathrm{m}^2 \cdot \mathrm{sr}} = 1 \cdot \frac{\mathrm{J}}{L \cdot \mathrm{s} \cdot \mathrm{m}^2 \cdot \mathrm{sr}} \quad (3)$$

$$= 1 \cdot \frac{\lambda}{hc \cdot L \cdot \mathrm{s} \cdot \mathrm{m}^2 \cdot \mathrm{sr}} = 1 \cdot \frac{\lambda}{hc \cdot L} \cdot \frac{\mathrm{photons}}{\mathrm{s} \cdot \mathrm{m}^2 \cdot \mathrm{sr}} \quad (4)$$

Fig. 13 shows the traffic sign and the statistics for an illumination of $10 \cdot \frac{\mathrm{cd}}{\mathrm{m}^2}$. The photon flux obeys a Poisson distribution [7]:

$$\mathscr{P}_{\mu_n}(n) = \frac{{\mu_n}^n}{n!} e^{-\mu_n} \quad (5)$$

and is also modeled in this way. Fig. 14 shows an illumination with only $5 \cdot 10^{-15} \cdot \frac{\mathrm{cd}}{\mathrm{m}^2}$, resulting in only 4 photons per square meter and steradians. This makes the Poisson distribution of the flux visible.

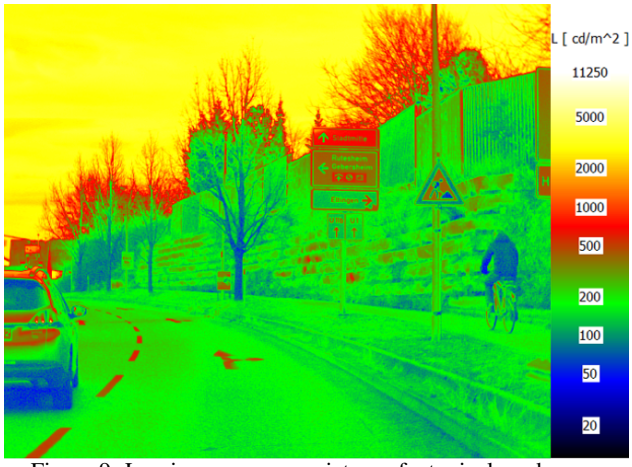


Figure 9: Luminance camera picture of a typical road scene

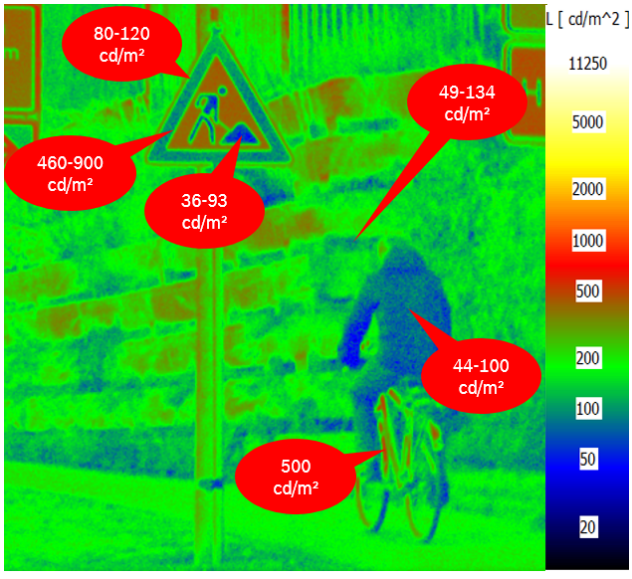


Figure 10: Zoom into fig. 9 for a contrast analysis of the objects

Windshield

After the illuminated scene, the next block in the imaging chain is the windscreen according to fig. 11. As shown in the use case analysis, the wind screen may be the source for stray light or veiling glare (fig. 1) Again the photon fluxes need to be modeled as Poisson events and simplified to a small patch of the image, the veiling glare adds a certain number of uniform distributed photons over the whole image.

Such an effect is easily obtained in a car, if the windscreen is dirty, e.g. by pollen dust. If the car is driving below an illumination, the glare photons are added into the imaging photon flux and reduce the contrast significantly. These added photons are independent of the scene illuminance and can therefore extend the amount of photons that are emitted from the scene. Adding the average photon amount from the scene as veiling glare, the con-

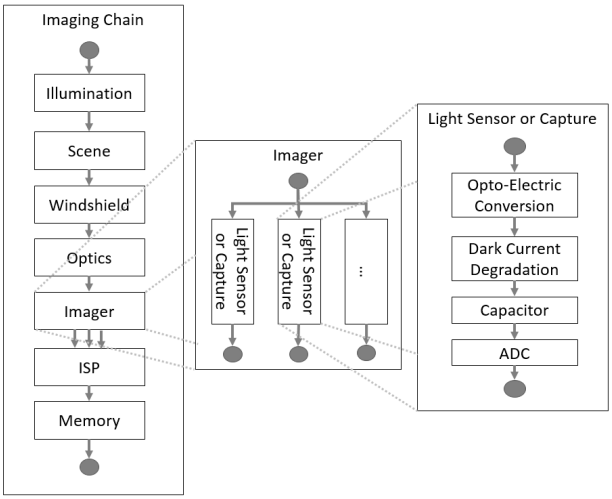
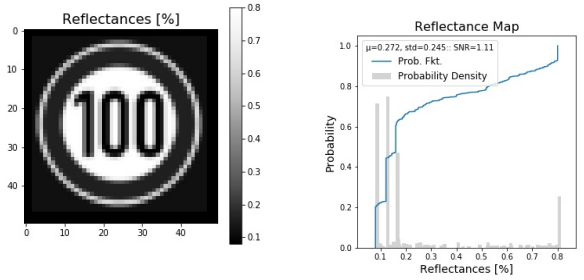
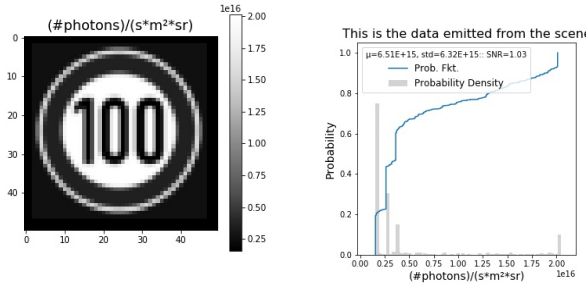


Figure 11: Overview of the used imaging chain, including a zoom view onto the imager, and the light sensor model.



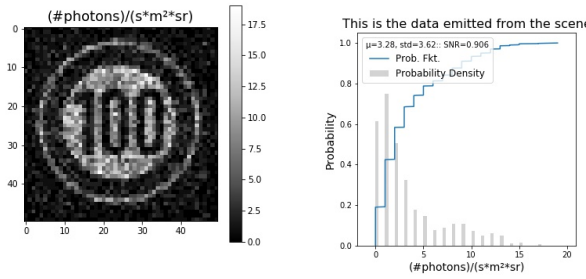
(a) (b)

Figure 12: Reflectance Information. 12a: Reflectance map of a German 100$\frac{km}{h}$ speed limit sign. 12b: Probability density function and probability function for the data from fig. 12a



(a) (b)

Figure 13: Emitted Photons in a $10 \cdot \frac{\text{cd}}{\text{m}^2}$ illumination 13a: Emittance map of a German 100$\frac{km}{h}$ Speed limit sign. 13b: Probability density function and probability function for the data from fig. 12a



(a) (b)

Figure 14: Emitted Photons in a $5 \cdot 10^{-15} \cdot \frac{\text{cd}}{\text{m}^2}$ illumination. 14a:Reflectance Map of a German 100$\frac{km}{h}$ Speed limit sign. 14b:Probability density function and probability function for the data from fig. 14a.

trast reduces already significantly:

$$K_{\text{Weber}} = \frac{E_{\max}}{E_{\min}} - 1 = \frac{\frac{1}{2}(460+900)}{\frac{1}{2}(80+120)} - 1 \approx 580\% \quad (6)$$

$$E_{\text{glare}} \approx 390 \quad (7)$$

$$K_{\text{Weber,Glare}} = \frac{E_{\max}+E_{\text{glare}}}{E_{\min}+E_{\text{glare}}} - 1 \quad (8)$$

$$= \frac{\frac{1}{2}(460+900)+390}{\frac{1}{2}(80+120)+390} - 1 \approx 118\% \quad (9)$$

Fig. 15 depicts the described veiling glare model and fig. 16 demonstrates the impact onto the visibility in case of a low photon count. Please note that the SNR value of this signal is increased over fig. 14a due to the increased photon count. As only Poisson processes are present, the resulting larger expectation value gives again a corresponding Poisson standard deviation. This leads however to an increased SNR value while a reduced visibility of the object is given.

In addition to the veiling glare, a transmission of 96% through the windscreen has been modeled.

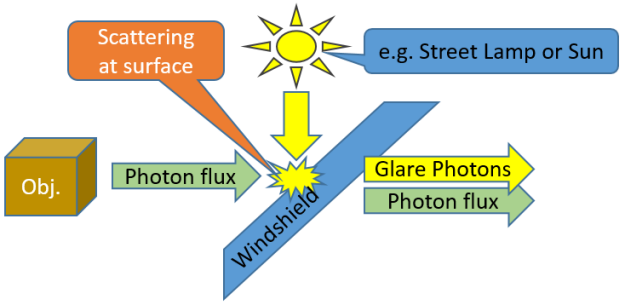
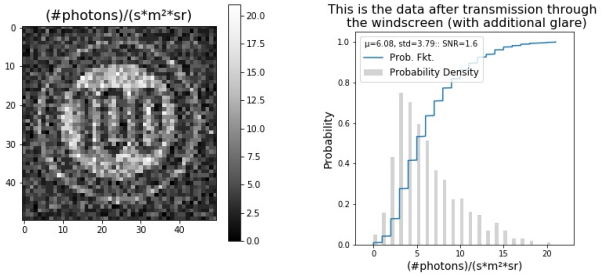


Figure 15: Simplified glare model: An external light source scatters at the windshield and adds to the object's photon flux



(a) (b)

Figure 16: Emitted Photons in a $5 \cdot 10^{-15} \cdot \frac{\text{cd}}{\text{m}^2}$ illumination with 100% additional glare photons that diminish the contrast. 16a: Emitted photons polluted with 100% glare photons. 16b: Probability density function and probability function for the data from fig. 16a.

Optics

The next block in the imaging chain is the optics, which are modelled following [4] p.96:

$$E\left[\frac{\text{photons}}{\text{s}\cdot\text{m}^2}\right] = \frac{t\pi\cos^4(\Theta)}{4f_{\#}^2(1+m_l)^2} \cdot L\left[\frac{\text{photons}}{\text{s}\cdot\text{m}^2\cdot\text{sr}}\right] \quad (10)$$

Here t is the optical transmission, m_l the lateral mapping factor of the optics, Θ the angle of the incoming light with respect to the optical axis and $f_{\#}$ is the f-stop number of the optics.

For the simulations we assume an $t = 90\%$, $f_{\#} = 2.0$, $\Theta \approx 0$ assuming far away objects and for simplification. Then we assume that the distance of the objects is much larger than the focal

distance f, which simplifies m_l to:

$$m_l = \frac{f+d'}{f+d} \tag{11}$$

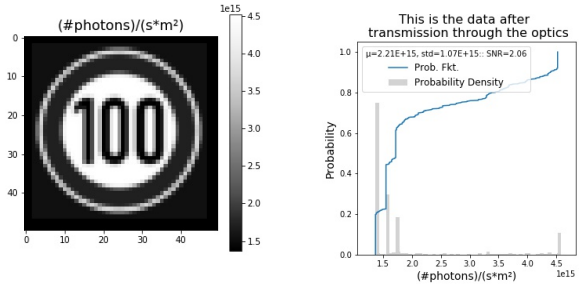
$$d' \approx 0 \quad \text{means image in the focal plane} \tag{12}$$

$$d \gg f \rightarrow d = N \cdot f \tag{13}$$

$$m_l = \frac{f+d'}{f+d} \approx \frac{f+0}{f+Nf} = \frac{1}{1+N} \rightarrow 0 \tag{14}$$

$$\Rightarrow E \cdot \frac{\text{photons}}{\text{s} \cdot \text{m}^2} = \frac{t\pi}{4f_\#^2} \cdot L \frac{\text{photons}}{\text{s} \cdot \text{m}^2 \cdot \text{sr}} \tag{15}$$

After all this is a linear effect onto the photon flux and the resulting photons are focussed onto the imager surface. The effects are depicted in fig.17



(a) (b)

Figure 17: Photon count at the imager surface after transmission of the signal trough the optics. 17a Photons at the imager surface after the optics (polluted with 100% glare photons and $10 \cdot \frac{\text{cd}}{\text{m}^2}$ illumination. 17bProbability density function and probability function for the data from fig. 17a

Other effects of the optics have not been modelled. For an improved simulation model the strong influence of the optical point spread function onto the contrast transfer given different spatial frequencies should be considered.

Imager

The next block in the imaging chain is the image sensor (see fig. 11). Image sensors consists of several million light sensors that are copies of each other at different locations. So each light sensor captures the signal in the same way like its neighbors.

However, the typical ADAS light sensors are able to take several captures of the same light signal, leading to several different captures per image. This is usually realized by either different capture surfaces (e.g. split pixel approaches [8]) or by taking several captures temporally one after the other (staggered HDR approach [9]). The later approach may lead to motion artifacts for fast moving objects, while the split pixel approach consumes more area in the silicon process for comparable pixel performance.

To simulate an image sensor, the light sensors need to be modeled in detail. The here proposed image sensor is derived from typical values for illustrative purposes. It is not a model of a real device.

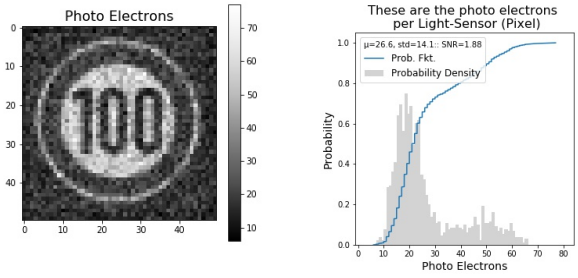
Light sensor

As shown in fig. 11 the light sensors can be roughly simulated by the four steps: Optoelectronic conversion, Dark current pollution, electron aggregation in a capacitor and finally an analog digital conversion. Modern image sensors will have a complicated model that exceeds the abilities of a single paper, but the EMVA1288 camera and light sensor model [3] has been proven to be a good approximation.

Optoelectronic Conversion (OEC): The OEC describes the photon to electron conversion at the surface of the light sensor. Here we assume a pixel pitch of $2\mu m$ leading a photo receptive surface of $A_{\text{lightsensor}} = 2\mu m \cdot 2\mu m = 4 \cdot 10^{-12}\text{m}^2$. The quantum efficiency in average over the whole visual spectra may yield to $\eta \approx 70\% \frac{\text{electrons}}{\text{photons}}$ with peaks reaching around $\eta \approx 80\%$. Typical integration times of the sensor element at daytime will be around $t_{\text{exp.}} = 1ms$ to $t_{\text{exp.}} = 10ms$ to avoid saturation at the bright parts of the scene. This results to

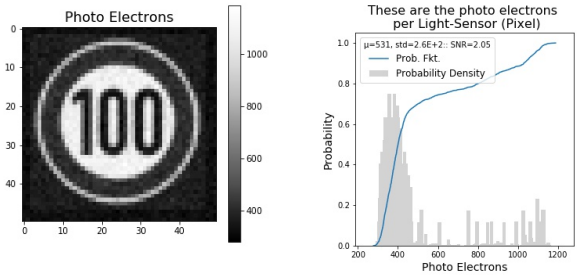
$$e[\text{electrons}] = E \cdot \frac{\text{photons}}{\text{s} \cdot \text{m}^2} \cdot A_{\text{lightsensor}} \cdot t_{\text{exp.}} \cdot \eta \tag{16}$$

Due to the very small size of the photo sensitive area, the number of photons per lightsensor capture decreases to a few dozen as shown in fig. 18 for an exposure of $5ms$. Fig. 19 shows an exposure of $100ms$ which gives 10 times more photons, however in $100ms$ the dark current effects will become visible as shown in the next section.



(a) (b)

Figure 18: Produced photo electrons for a $5ms$ exposure. 18a: Electrons generated at the photo diode with $5ms$ exposure time. 18b: Probability density function and probability function for the data from fig. 18a.



(a) (b)

Figure 19: Produced photo electrons for a $100ms$ exposure, row and column wise effects become visible. 19a: Electrons generated at the photo diode with $100ms$ exposure time. 19bProbability density function and probability function for the data from fig. 19a.

Dark Current Pollution: While the photo active area of the light sensor is producing electrons that are generated by light, unwanted effects produce electrons as well. As these electrons are generated even without light, they are called dark current electrons. These dark current electrons are generated by the thermal energy that is present in the system. The dark current obeys also a Poisson process.

However, each lightsensor has dark current properties due to variations in the manufacturing process, resulting in the following Poisson Processes that create the dark current:

- $\mathscr{P}_{\mu_l}$ as an average dark current which is is identical to all light sensors.
- $\mathscr{P}_{\lambda_{l,i,j}}$ as an individual dark current for each light sensor.
- $\mathscr{P}_{\mu_c}$ as an average column wise dark current, identical to all columns
- $\mathscr{P}_{\lambda_{c,i}}$ an individual column wise dark current.
- $\mathscr{P}_{\mu_r}$ an average row wise dark current that is identical to all rows
- $\mathscr{P}_{\lambda_{r,j}}$ an individual row wise dark current.

The parameters $\mu[\frac{e^-}{s}]$ and $\lambda[\frac{e^-}{s}]$ are hereby given in units of electrons per second. μ represents the overall dark current effects, while the deviation of the parameters λ are responsible for differences that are systematic different between the light sensors. The standard deviation over all parameters of a given type of lambda is called Fixed Pattern Noise (FPN):

$$\lambda_{(\cdot)}: \quad \text{realization of:} \quad \mathscr{P}_{\text{FPN},\{\text{r,c,l}\}} \tag{17}$$

A separation between μ and λ is not done in each standard or investigation, however, due to the physical properties of different electron creation processes, this model is more accurate.

The dark current is a thermal generated process, the parameters μ and λ have a strong temperature dependency. We use the EMVA1288 model to describe it and use reference values $\mu_{\text{ref}}, \lambda_{\text{ref}}$, measured at a reference temperature T_{ref}. This allows to calculate the Poisson parameters for the temperature of interest T by applying a doubling Temperature T_{doubling} in the following way:

$$\mu = \mu_{\text{ref}} \cdot 2^{\frac{T-T_{\text{ref}}}{T_{\text{doubling}}}} \quad \lambda = \lambda_{\text{ref}} \cdot 2^{\frac{T-T_{\text{ref}}}{T_{\text{doubling}}}} \tag{18}$$

In each final measurement of a single light sensor, the different dark current contributions add up. This makes a direct measurement not possible, and to extract the above mentioned components, measurements at darkness and two different temperatures have to be conducted. A linear system of equations can be formed by calculating the row and column wise sums. The sums can be interpreted as estimators of the parameters μ and λ of the random processes:

$$\eta_{\{r(j),c(i),l(i,j)\}}(T) = \mu_{\{r,c,l\}}(T) + \lambda_{\{r(j),c(i),l(i,j)\}}(T) \tag{19}$$

$$\rightarrow \mathscr{P}_{\eta_{\{r,c,l\}}} \quad \text{also a Poisson Process} \tag{20}$$

$$e_{\{r,c,l\}}(T) \quad \text{as the realizations of } \mathscr{P}_\eta \tag{21}$$

$$m_{\{r,c,l\}}(T) \quad \text{as the realizations of } \mathscr{P}_\mu \tag{22}$$

$$l_{\{r,c,l\}}(T) \quad \text{as the realizations of } \mathscr{P}_\lambda \tag{23}$$

$$k = e_l + e_r + e_c \quad \text{as measurements} \tag{24}$$

Given $i, j \in M, N$ as row and column indices, which apply to: $l(i,j), r(j)$ and $c(i)$ we obtain:

$$k = e_l + e_r + e_c \tag{25}$$

$$\underbrace{\frac{1}{MN}\sum_{i,j} k}_{\approx \mu_k} = \underbrace{\frac{1}{MN}\sum_{i,j} e_l}_{\approx \eta_l(T)} + \underbrace{\frac{1}{MN}\sum_{i,j} e_r}_{\approx \eta_r(T)} + \underbrace{\frac{1}{MN}\sum_{i,j} e_c}_{\approx \eta_c(T)} \tag{26}$$

$$\underbrace{\frac{1}{N}\sum_{i} k_c}_{\approx \mu_{k_c}} = \underbrace{\frac{1}{N}\sum_{j} e_l}_{\approx \eta_l(T)} + \underbrace{\frac{1}{N}\sum_{j} e_r}_{\approx \eta_r(T)} + \underbrace{\frac{1}{N}\sum_{j} e_c}_{\approx e_c} \tag{27}$$

$$\underbrace{\frac{1}{M}\sum_{j} k_r}_{\approx \mu_{k_r}} = \underbrace{\frac{1}{M}\sum_{i} e_l}_{\approx \eta_l(T)} + \underbrace{\frac{1}{M}\sum_{i} e_r}_{\approx e_r} + \underbrace{\frac{1}{M}\sum_{i} e_c}_{\approx \eta_c(T)} \tag{28}$$

So far we have M equations of type 27 and N equations of type 28. We can use the definition of variance and the fact that for Poisson processes the variance (σ^2) equals the expectation value μ:

$$\frac{1}{N}\sum_j (\mu_k - \mu_{k_c})^2 = \frac{1}{N}\sum_j (\eta_c - e_c)^2 \approx \sigma(T)_c^2 = \eta_c \tag{29}$$

$$\frac{1}{M}\sum_i (\mu_k - \mu_{k_r})^2 = \frac{1}{M}\sum_i (\eta_r - e_r)^2 \approx \sigma(T)_r^2 = \eta_r \tag{30}$$

With the knowledge of the parameters $\eta_c(T)$ and $\eta_r(T)$ we can use equation 26 to determine $\eta_l(T)$, and then subsequently the equations of type 27 and type 28 to estimate the realizations e_c and e_r. Following with equation 26 we can finally estimate e_l as well from the measurements k.

Given the realizations of a single frame, the capture of multiple frames at different and different Temperatures T allows to extract the wanted parameters:

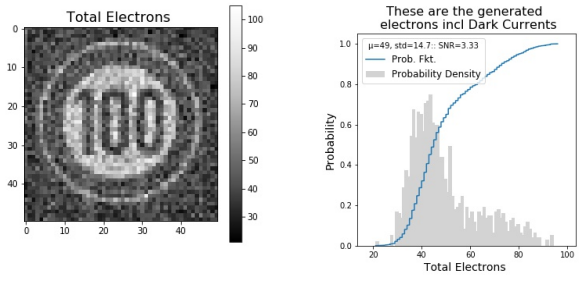
$$e_{\{r,c,l\}}(t,T_a) = m_{\{r,c,l\}}(t,T_a) + l_{\{r,c,l\}}(t,T_a) \tag{31}$$

$$e_{\{r,c,l\}}(t,T_b) = m_{\{r,c,l\}}(t,T_b) + l_{\{r,c,l\}}(t,T_b) \tag{32}$$

Finally the temporal statistics over multiple measurements for the different realizations of types $l_(t)$ yields to the individual Poisson Parameters λ for each row, column and light sensor. The statistics of the different $m(t)$ yields to the Poisson parameters of type μ that are identical for each type of row, column and pixel. When correcting the dark current effects, these parameters form the expectation values and can directly be subtracted from the signal.

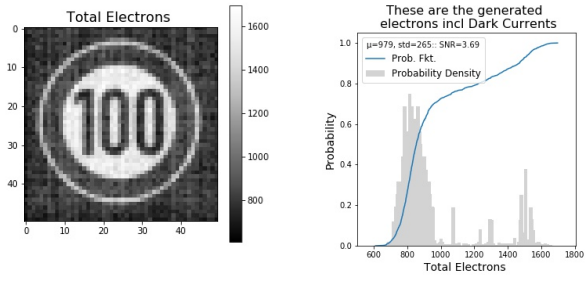
The nature of the dark current makes it almost undetectable at single frames if low temperatures and low exposure times are given. Fig. 20 shows the accumulated electrons including the dark current electrons for an exposure time of 5ms as before. However, in fig. 21 the 100ms exposure time shows first effects of the dark current. Both evaluations have been simulated at 125$^\circ C$, a total dark current of $50\frac{e^-}{s}$, a row wise dark current of $5\frac{e^-}{s}$, a column wise dark current of $10\frac{e^-}{s}$, a total FPN of $20\frac{e^-}{s}$ standard deviation, a row wise FPN of $5\frac{e^-}{s}$ standard deviation, a column wise FPN of $10\frac{e^-}{s}$ standard deviation.

Accumulation in a Capacitor The generated photo electrons and the dark electrons are summed up and stored in a capacitive



(a) (b)

Figure 20: Photo and Dark current electrons for a 5*ms* exposure of the discussed scene. 20a: Photo and dark current generated with 5*ms* exposure time. 20bProbability density function and probability function for the data from fig. 20a.



(a) (b)

Figure 21: Photo and Dark current electrons for a 100*ms* exposure of the discussed scene. 21a: Photo and dark current generated with 100*ms* exposure time. 21b: Probability density function and probability function for the data from fig. 21a.

element inside the device. As the charge storage is limited, the simulation clips all occurrences that are larger than the storage limit. The storage limit is also called the full well capacity (FWC) and forms a well established KPI of light sensors and image sensors.

The storage has been set to $15000e^{-}$, for this simulation. The real imaging chain is more complicated due to the fact that the electrons are usually generated and stored in different locations. Therefore usually a transfer loss during the read processes occurs which is connected to the so called kTC noise. These effects have not been simulated in the current chain.

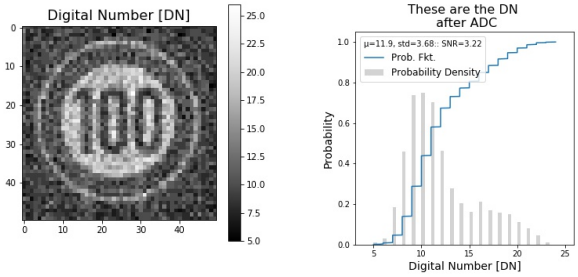
ADC Conversion Having the electrons stored inside the capacitor, they result in a certain voltage, which is then usually amplified and connected to an analog digital converter. To avoid modeling all the effects inside this chain, again the simplified EMVA1288 model has been used.

Here an overall system gain is used to describe how many digital numbers are produced per electron. In use cases where a single electron should be countable at least $K = 2\frac{DN}{e^{-}}$ should be used according to the sampling theorem.

But as in many cases, the lower bound of electrons is polluted by noise anyway and a system gain is chosen that guarantees to reach the full well capacity with the available digits of the ADC. For $15ke^{-}$ a 12 bit ADC should use a system gain of:

$$K = \frac{2^{12}\text{DN}}{15000\text{e}^{-}} \approx 0.25\frac{\text{DN}}{\text{e}^{-}} \tag{33}$$

Of course such an undersampling produces (discretization) artifacts that can be seen in fig. 22.



(a) (b)

Figure 22: Digital numbers (DN) after discretization with 12 bits. 22a: Digital numbers (DN) created from the 5*ms* exposure. 22bProbability density function and probability function for the data from fig. 22a.

Image signal processing (ISP)

As mentioned in the imager section, multiple captures are produced by the light sensors. The ISP is now responsible to recombine these captures into one light signal that corresponds to the input signal of the system in the best possible way.

HDR reconstruction

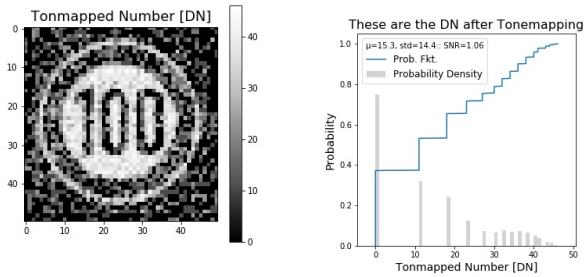
To accomplish this, an HDR reconstruction module extracts out of the multiple exposures the linear gain factor that has been applied onto the photon fluxes. This gain factor allows to normalize all captures onto the longest exposure. Given the number of HDR-bits that are foreseen to capture the whole signal, it may be that these bits do not allow to cover the largest number that is produced by the gained exposure stages. Such a misalignment causes a loss of dynamic range of the sensor's hardware capabilities and is demonstrated in the later section of this paper.

The dynamic range can be adjusted with the overall system gain of the light sensors but sacrifices the accuracy in low light scenarios. For this simulation a 22-bit number has been used to capture the HDR signal.

Tonemapping

After the signal is HDR reconstructed, the output usually disallows to to display and to process the signal in an efficient way. Therefore the HDR number is compressed in a process that is called tone mapping. Tonemapping can be done by analyzing the contained information inside the signal, or simply by design of the demanded information by the next processing steps.

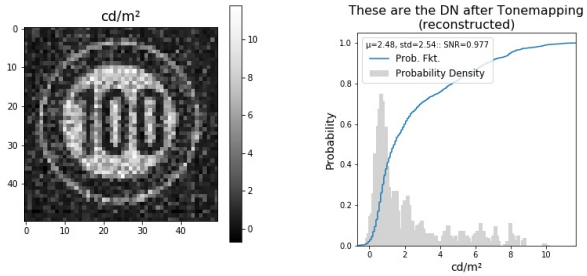
In case of a visual pleasing representation, a logarithmic guarantees a constant contrast encoding. In this simulation an 8-bit logarithmic tone mapping has been used. Despite the constant contrast encoding, we see in fig. 23 that some information of the original data has been lost. This is especially visible in the histogram, that shows the discretization effects of the tone mapping, which include rounding and changes in the intensity steps.



(a) (b)

Figure 23: Digital numbers (DN) after HDR reconstruction and log-compression to 8 bits. 23a: Compressed 8-bit logarithmic digital numbers created from the 5*ms* exposure. 23b: Probability density function and probability function for the data from fig. 23a.

However, such a discretization might not be visible at the imager output. If we calculate back all the effects that have been introduced towards in the effect chain (e.g. the dark current and veiling glare effects) we end up in the $\frac{\text{cd}}{\text{m}^2}$ domain. Due to the corrections, the gaps in the histogram have been filled as shown in fig. 24.



(a) (b)

Figure 24: Reconstruction of $\frac{\text{cd}}{\text{m}^2}$ from the Digital numbers in fig.23. 24a: Spatial impression in $\frac{\text{cd}}{\text{m}^2}$ after inverting the effect chain. 24b: Probability density function and probability function for the data from fig. 24a

Contrast Detection Probability

Without extended simulation and experimental efforts the presented effects in the imaging chain are hard to judge with their impact onto image quality. As shown, it might happen that the object detection in the final image signal gets severely degraded. To specify requirements along the imaging chain a measure of the signal to noise ratio is often used. But still it remains difficult to connect an SNR value to the algorithmic or human ability to recognize an object inside the represented signal.

Given the requirement engineering section above, the contrasts of the objects in question are known and including elaborated optical simulations, these contrasts could also be expressed depending on the object's spatial frequencies. As the photon flux is based on a Poisson process, all measurements along the system chain are realizations of on random variables that are based on the photon flux. Therefore each measurement result can be expressed inside confidence intervals with given probabilities.

Given the requirement engineering, the imaging chain has to guarantee that a specified contrast can be detected in the signal. Concluding, the contrast measurement will also be a realization of a random variable. Given the knowledge of the objects in question, a good requirement for the imaging chain is to guarantee the measurement of a demanded contrast by a given probability. And to specify the probability an established or derived confidence interval has to be given.

Definition

There exist manifold contrast definitions, among them is the Weber Contrast [2] which is based on the Weber-Fechner law on human perception:

$$K_{\text{Weber}} = K_{\text{W}} = \frac{E_{\text{max}}}{E_{\text{min}}} - 1 \quad \text{with:} \quad 0 \leq K_{\text{Weber}} \leq \infty \tag{34}$$

If we assume a relative brightness difference p:

$$E_{\text{max}} = (1+p)E_{\text{min}} \tag{35}$$

$$\Rightarrow K_{\text{Weber}} = \frac{E_{\text{max}} - E_{\text{min}}}{E_{\text{min}}} \tag{36}$$

$$= \frac{(1+p)E_{\text{min}} - E_{\text{min}}}{E_{\text{min}}} \tag{37}$$

$$= \frac{E_{\text{min}}((1+p)-1)}{E_{\text{min}}} \tag{38}$$

$$= p = p_{\text{Weber}} \tag{39}$$

the contrast K_{Weber} corresponds directly to this percentual illuminance difference.

In the field of Optics the Michelson contrast is the more common used contrast and it is defined by:

$$K_{\text{Michelson}} = K_{\text{M}} = \frac{E_{\text{max}} - E_{\text{min}}}{E_{\text{max}} + E_{\text{min}}} \quad \text{with:} \quad 0 \leq K_{\text{Michelson}} \leq 1 \tag{40}$$

and can be transferred into the Weber Contrast by:

$$K_{\text{M}} = \frac{E_{\text{max}} - E_{\text{min}}}{E_{\text{max}} + E_{\text{min}}} \tag{41}$$

$$\Leftrightarrow \frac{1+K_{\text{M}}}{1-K_{\text{M}}} - 1 = \frac{E_{\text{max}}}{E_{\text{min}}} - 1 \tag{42}$$

$$\Leftrightarrow \frac{1+K_{\text{M}}}{1-K_{\text{M}}} - 1 = K_{\text{W}} \tag{43}$$

$$\Leftrightarrow K_{\text{W}} = \frac{2K_{\text{M}}}{1-K_{\text{M}}} \tag{44}$$

Having a contrast definition at hand we can now define the random variable of the measured contrast:

$$\mathbf{K}_{\text{meas.}}: \quad \text{random variable} \tag{45}$$

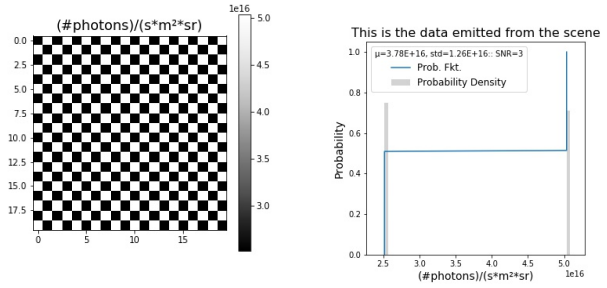
and the derive the **Contrast Detection Probability** CDP by utilizing the probability function $\text{Prob}(\cdot)$ and systems or components input contrast K_{in}:

$$C_{K_{\text{in}}}DP = \text{Prob}\left(K_{\text{in}}(1-\varepsilon) \leq \mathbf{K}_{\text{meas.}} \leq K_{\text{in}}(1+\varepsilon)\right) \tag{46}$$

Example for CDP

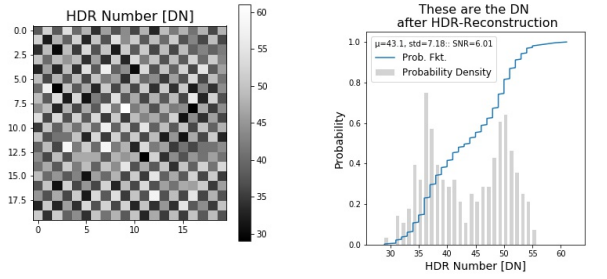
A requirement for a specific use case could be to detect an input contrast of $K_W = 100\%$ with a probability larger 90% in a confidence interval of $\varepsilon = 50\%$. Concluding, the random variable that measures the contrast K_W at the point of interest should then realize to contrast values $K \in (50\%, 150\%)$ with a 90% probability. Fig.25, 26 and 27 illustrate this example for the discussed optical simulation with help of a checker board reflectance map at different illuminations. It becomes visually evident that less photons result in higher noise and consequently in less probability to detect the contrast. The CDP confirms this impression.

It should be mentioned that the CDP is defined in the physical dimension of the input. For this paper this is the reflectance map of the scene, or the derived $\frac{\text{cd}}{\text{m}^2}$ luminous emittance of the corresponding lambertian surfaces. Thus the digital numbers at the end of the evaluation chain have to be transferred back into the $\frac{\text{cd}}{\text{m}^2}$ domain before an evaluation of the CDP is possible. Numbers that are below $0\,\frac{\text{cd}}{\text{m}^2}$ should be clamped to 0 as there is no negative photon count in the physics of this context.



(a) (b)

Figure 25: Checkerboard Reference. 25a: A checkerboard with a contrast of 100% and $20\,\frac{\text{cd}}{\text{m}^2}$. 25b: Probability density function and probability function for the data from fig. 25a .

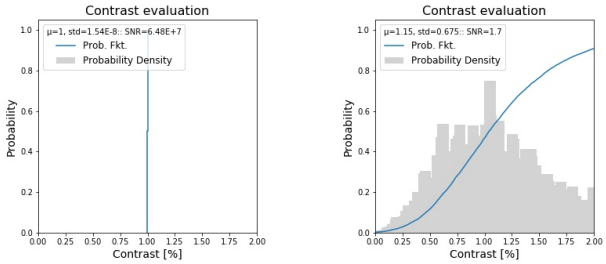


(a) (b)

Figure 26: Checkerboard after processing by the imaging chain, with reduced CDP. 26a: A checkerboard with a contrast of 100% and $20\,\frac{\text{cd}}{\text{m}^2}$ after HDR Reconstruction. 26b: Probability density function and probability function for the data from fig. 26a.

Depended variables and Comparison to SNR

Contrast detection probability can be evaluated against different parameters. For a random variable such evaluations represent the conditional probabilities. In this paper CDP is investigated, given the realization of different intensities in a scene. This



(a) (b)

Figure 27: Evaluations of the Contrast detection Probability for the examples of 100% contrast (see fig.25) and 30% contrast (see fig. 26. 27a: CDP evaluates to $\approx 100\%$, because 100% of the measurements are inside the interval of $50-150\%$. 27b: After HDR reconstruction in the imaging chain, CDP evaluates to $\approx 90\%$ because 90% of the measurements are inside the interval $50-150\%$.

allows to judge the system performance with respect to the dynamic range of the intensity which is a critical use case as shown with the tunnel exit in fig. 6

Other parameters that like the realization of spatial frequencies and the different temperatures are also valuable choices for conditional CDP probabilities. To compare the KPI's suitability we compare the results against the common measure in signal theory: The signal to noise ratio (SNR). We show that SNR can be connected in some cases to the manifold contrast requirements by using the CDP as transfer layer. CDP considers all the events that fall inside the demanded range of measurement, and considers therefore all relevant details of the probability density function. SNR is on the other hand just one parameter that describes the density function just very roughly. However, as SNR does not consider the various shapes of the probability density functions (e.g. due to dark currents and quantization), it can be shown that there is no general mathematical way to connect SNR and CDP without further assumption. Possible connections between SNR and CDP that are based on weak assumptions towards the probability densities are not within the scope of this paper.

Experimental results

For a first insight of the KPI's development the imaging chain according to fig. 11 is evaluated and the SNR and CDP are measured in their temporal order. Fig.28 shows how SNR and CDP develop along the imaging chain and a $500\%, 100\%$ and 30% Weber contrast have been evaluated for CDP, all referring to the same SNR values, which are evaluated by a uniform intensity patch at the average intensity of the CDP contrast intensities.

The figure shows that the SNR decreases from a very large number in the photon domain of 150dB down to $<$ 20dB. One can observe the unwanted effect that the SNR increases at two stages if SNR is measured in the signal domain of the corresponding blocks:

- Adding veiling glare photons at the windscreen, the SNR increases because adding another Poisson random variable increases the signal value, but the noise characteristics remain poissonic.
- Tone mapping increases the SNR as well because here the

standard deviation of the signal gets decreased by quantization effects. Consequently the SNR value increases.

In both cases we have shown in the imaging chain analyses that the image quality in fact degrades. Therefore it is not useful to measure SNR in output domain if the goal is to describe image quality with SNR. False increases of such a wrongly designed KPI may lead to wrongly specified system components.

If we transfer the measured data numbers back into the original input domain (in our example always $\frac{\text{cd}}{\text{m}^2}$) then the SNR behaves as expected in the shown example of $I = 20\frac{\text{cd}}{\text{m}^2}$. But in the next section we will show later that this is unfortunately not always the case.

Considering the CDP, we can observe that for the same SNR values the systems reaction to detect different contrasts differs. Due to the fact that CDP is a probability the values are bound to the interval $CDP \in (0,1)$, which makes a detailed analysis much easier than with the unbound SNR values. Very large SNR numbers don't increase contrast detection probability any more, and useless for the requirement evaluation. CDP decreases with each step of the imaging chain as does the image quality, therefore these CDP results are as expected from the image quality analysis.

Given the direct link towards contrast detectability, CDP is to favor over the SNR when it comes to specify a system performance and the components inside the system.

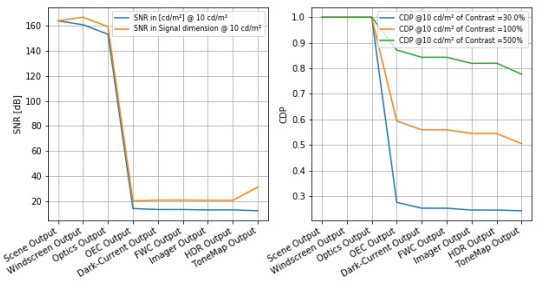


Figure 28: Overview of the KPI development along the effect chain. The SNR measure increases although the signal is degraded, e.g. through veiling glare at the windshield and after tone mapping. SNR in the $\frac{\text{cd}}{\text{m}^2}$ domain behaves as expected, but does not mirror directly the impact onto the imposed requirement. CDP shows the detection probability of a 100% contrast and how it is degraded by the several steps in the imaging chain. The requirement can be judged directly

Fig.28 did show the development of the KPIs given the only one illumination of $20\frac{\text{cd}}{\text{m}^2}$. However to judge an imaging system, the performance over the whole dynamic range of the scene needs to be analyzed.

For the next plots, fig.28 showed that signals before the opto-electrical conversion do not degrade the image quality. Fig. 29 and fig. 30 therefore show the development of the KPIs SNR and CDP starting with the opto-electrical conversion, but evaluate a single exposure imaging chain over all illuminances.

SNR Bad Practice: Fig. 29 emphasizes the above statement to never measure the SNR in the component's signal domain. If this is done effects like in the left hand side of the plot will appear:

- The tonemapped SNR is higher than the SNR achieved by the optical output, which does not correlate to the observed image quality.
- At the capacitor level, the fact that 0 electrons are quantized leads to the SNR values above the optical signal.
- The same happens for the HDR-Output where the SNR might lead to higher values than possible due to quantization effects.

Compared to the observed image quality from the imaging chain analysis the SNR values do not correlate to the image quality.

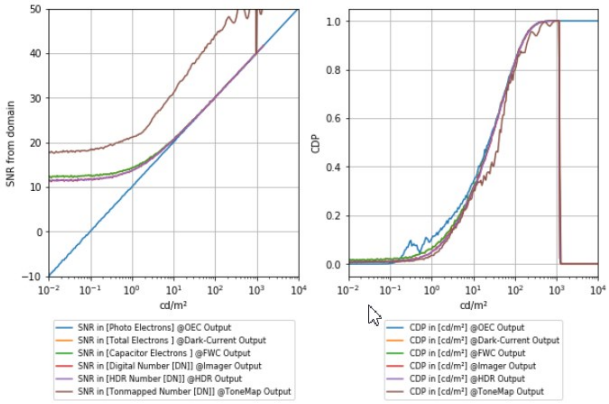


Figure 29: SNR and CDP comparison. CDP for a 30% contrast. SNR correlates not to image quality, as reduced image quality in tone mapped and HDR output is observed, while SNR is increased. From for less than $10\frac{\text{cd}}{\text{m}^2}$ SNR does not drop according to the image quality degradation. CDP however is always below the optical CDP and therefore maps the image quality for the given requirement.

SNR Good Practice: If the SNR is calculated in the domain of the physical system input (e.g. $\frac{\text{cd}}{\text{m}^2}$), the above described negative SNR effects become obsolete (see fig.30). But still at points below $-5dB$ for this setup one can observe a crossing point between the optical SNR and the SNR calculated back from the quantized digital numbers which does not correlate to the demanded image quality description.

In our example this $-5dB$ point is located in the area for very low light performance. There exist techniques to increase the SNR at this points, however their impact onto the contrast detection probability has not been shown. For example a black level cut of the signal might increase the SNR of the signal drastically in this region, leading to improved visual image quality but resulting into a decreased contrast detection probability. Therefore only if the CDP is increased as well by the black level compensation the effect is of value for the system performance.

With respect to the SNR and CDP comparison, we observe that once the capacitor reaches its full well level, the CDP converges to 0% while the SNR converges to infinity. This is due to the fact that more and more pixels achieve saturation and therefore the standard deviation gets reduced when calculating the SNR. For the CDP, the difference between the two contrast levels becomes 0 if the full well level is reached and therefore the CDP drops to 0 this correlates will to the observed image quality, the SNR increase does not.

In the SNR measurements oscillations in the tone mapped signal occur which exceed the SNR value that is possible by the optical signal even before the full well level is reached. This leads to the effect that if only the tone mapped signal is available, the calculation back to the $\frac{cd}{m^2}$ domain results in too high estimates of the image quality when judged by SNR. The CDP shows the mentioned oscillations as well, however they are limited below the contrast detection ability of the optical signal. The reason for these oscillations are quantization effects that arise due a limited count of digital numbers.

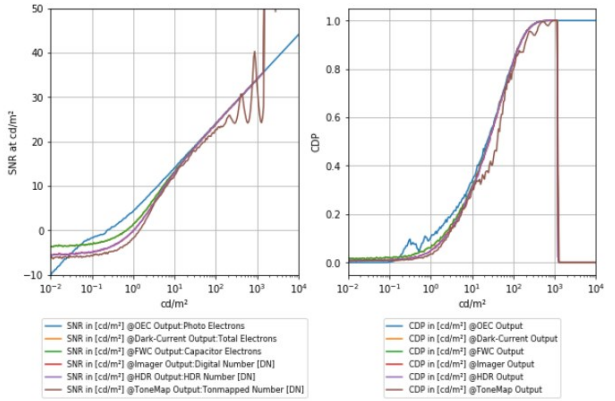


Figure 30: SNR and CDP for 30% contrast comparison. SNR is measured in $\frac{cd}{m^2}$ dimension. SNR stops to decrease for intensities less than $1\frac{cd}{m^2}$. And SNR oscillations exceed the optical SNR values. CDP behaves in accordance to image quality always below the optical CDP. Oscillations are visible according to quantization in the tone mapping

HDR Sensors

A single exposure is not able to capture a high dynamic range scene. Therefore SNR and CDP are evaluated for an image sensor model that uses a staggered HDR approach. Here we use the above mentioned imaging chain and use a 3-fold exposure with expose ratios of 100. For the simulated sensor this is necessary to cover the dynamic range of 140dB needed for scene like the tunnel exit.

Fig. 31 shows the concatenated SNR plots of the three exposures, including the combined HDR output and the tonemapped 8-bit signal. Due to limited bits for the HDR signal it saturates before the third exposure has reached its full well capacity. Such a behavior should be avoided by a good exposure control, but also shows how the ISP setup may limit the sensor's performance. The CDP part of the plot is generated for a 30% contrast and reveals that the SNR drop which go below $20dB$ SNR result in less than 50% contrast detection probability for such contrasts. For a human perception, objects are hard to identify if randomly more than 50% of their features become are wrongly depicted. Therefore the limit of 50% CDP is a good indicator for a threshold of visibility and detectability. As mentioned before such a clear interpretation is not possible in the SNR domain.

To demonstrate the effect of lost detectability in the SNR and CDP drops, the cyclist that has been analyzed in fig. 10 is transferred into a reflectance map and used as input for the imaging chain. Fig. 32 shows the a cyclist with 30% contrast, evaluated at certain illuminations. It can be observed that the cyclist disappears even though the illumination has increased from $500\frac{cd}{m^2}$ to $900\frac{cd}{m^2}$. This is a paradox effect as more light actually guarantees a better detectability. Thus the operation mode if this image sensor makes it impossible to detect the cyclist at certain illumination levels. The covered dynamic range of the sensor is however unchanged at $140dB$.

To emphasize the above discussion, fig. 33 shows the plot for 500% contrast requirement. While the SNR numbers remain unchanged, the CDP does never drop below 80%. Fig. 34 shows the same evaluation as 36 but this time the object of interest with a 500% contrast. The object is detectable throughout the whole dynamic range as predicted by the CDP evaluation.

Concluding, the analysis of the objects in question for the required use cases has to be ombined with the KPI contrast detection probability to judge and forecast the systems ability to detect the object or not.

Low Light Requirement Engineering

For the requirement derivation of low light performances, the $SNR = 1$ point is currently state of the art for specification and classification. Investigating this practice, fig. 33 shows this point at around $1\frac{cd}{m^2}$. The $SNR = 15$ level is reached at around $10\frac{cd}{m^2}$ CDP evaluates at the $1\frac{cd}{m^2}$ point to 40% and for $10\frac{cd}{m^2}$ for the $CDP = 80\%$. Thus we do not expect to detect the 500% contrast traffic sign at the $SNR = 1$ point. At the $SNR = 15$ point this is however possible according to the CDP values and fig. 35 confirms these judgments.

Given the task to detect the 30% cyclist, the CDP evaluates to 10% at the $SNR = 1$ point with $1\frac{cd}{m^2}$ and evaluates to 40% at the $SNR = 15$ point at $10\frac{cd}{m^2}$. Fig. 36 confirms that the cyclist is not detectable in either case.

Concluding the performance specification for low-light scenarios should also be evaluated using object contrasts to obtain a good system performance. If that is not done the wrong SNR values get specified and the image sensors might get tuned towards the wrong goal.

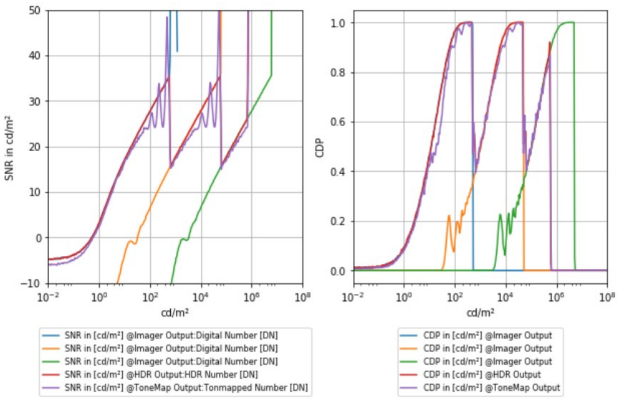
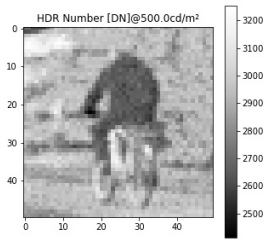
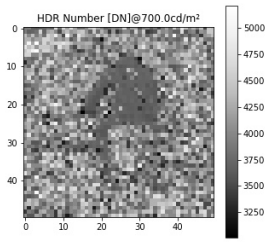


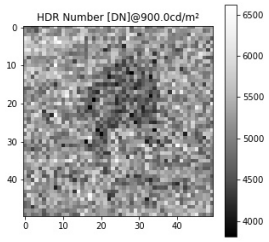
Figure 31: SNR and CDP for 30% contrast and a 3-fold staggered HDR sensor, covering a 140dB illuminance range



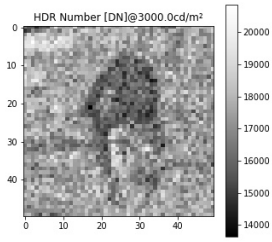
(a) 500 $\frac{cd}{m^2}$, detectable



(b) 700 $\frac{cd}{m^2}$, starting to become noisy



(c) 900 $\frac{cd}{m^2}$, not detectable



(d) 3000 $\frac{cd}{m^2}$, possilbe detection

Figure 32: Verification to fit the requirement analysis of fig. 31 onto a real life reflectance target. Paradox effect: Even with doubled amount of light the object disappears due to the sensor's operation mode. At 900 $\frac{cd}{m^2}$, the cyclist is not detectable and the $CDP < 50\%$ according to sensor analysis.

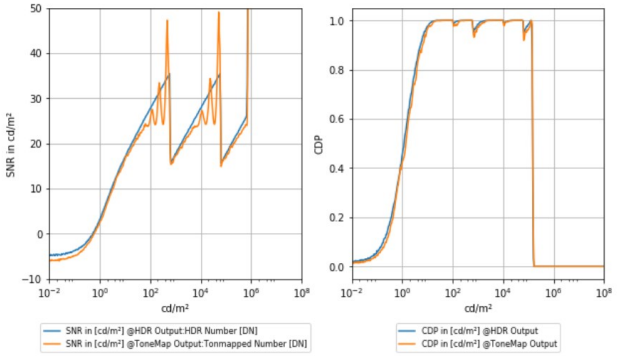
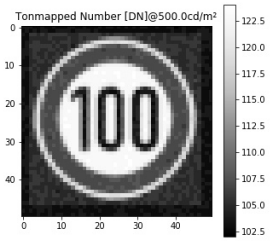
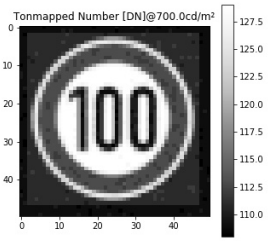


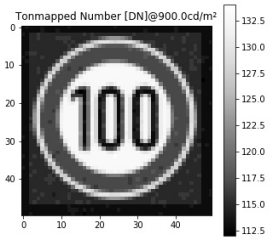
Figure 33: SNR and CDP for 500% contrast and a 3-fold staggered HDR sensor, covering 140dB scene dynamic



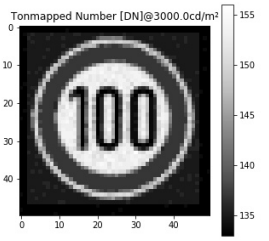
(a) 500 $\frac{cd}{m^2}$, detectable



(b) 700 $\frac{cd}{m^2}$, detectable

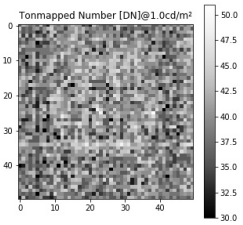


(c) 900 $\frac{cd}{m^2}$, detectable

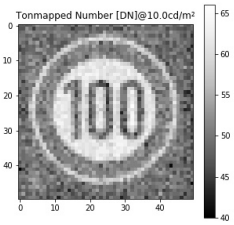


(d) 3000 $\frac{cd}{m^2}$, detectable

Figure 34: Same sensor setup as in fig. 32, however with a 500% contrast in the object: traffic sign. The sign is always detectable, which is confirmed by a CDP¿90% over the whole dynamic range.

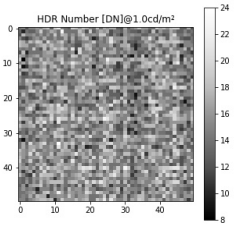


(a)

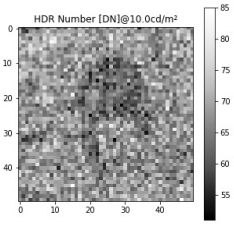


(b)

Figure 35: Low light requirement analysis for a traffic sign. 35a: 500% contrast traffic sign at 1 $\frac{cd}{m^2}$ and $CDP = 40\%$, not detectable. 35b: 500% contrast traffic sign at 10 $\frac{cd}{m^2}$ and $CDP = 80\%$, well detectable



(a)



(b)

Figure 36: Low light requirement analysis for the cyclist. 36a: 30% contrast cyclist at 1 $\frac{cd}{m^2}$ and $CDP = 10\%$: not detectable. 36b: 30% contrast cyclist at 10 $\frac{cd}{m^2}$ and $CDP = 40\%$: not detectable.

Outlook and Summary

In this paper we have emphasized the importance of safety relevant performance parameters for the automotive industry. Specific examples from the field of image quality have been used to derive performance KPIs. Therefore an analysis over the system's performance along a safety relevant critical effect chain has been conducted and we demonstrated that currently accepted KPIs such as SNR do not allow a precise and meaningful requirement derivation. We proposed a probabilistic approach for contrast detection, the contrast detection probability as a new KPI. As a result the requirement engineering along the effect chain especially for safety relevant use cases appears transparent and understandable.

The proposal of CDP as new KPI is currently under discussion within the P2020 work group on automotive image quality. The general approach to use detection probabilities in addition the classic KPIs is also discussed and under development. For example in this paper only one contrast has been judged against its influencing variables. In other fields of image quality, for example for color requirements, a similar approach is developed to specify a color separation probability (CSP) and to judge the systems geometric resolution via a geometric resolution probability (GRP).

The backbone of these ideas lies in the fact that the basic signal is a stochastic process and therefore should be described with KPIs based on random variables. In the near future safety relevant use cases will be pushed in the focus of development and a clear and easily understandable requirement derivation has to be conducted. The presented approach offers a possible solution to face this challenge.

.

References

[1] LK5 Luminance Camera, TechnoTeam GmbH (2017)

[2] Rainer Klinke, Stefan Silbernagl: Physiologie. Hrsg.: Hans-Christian Pape, Armin Kurtz, Stefan Silbernagl. 7. Auflage. ISBN 978-3-13-796007-2, p. 942. (2014)

[3] EMVA1288 3.1a, Url: http://www.emva.org/standards-technology/emva-1288/emva-standard-1288-downloads/

[4] Bernd Jaehne, Digitale Bildverarbeitung, ISBN 978-3-642-04952-1, (2012)

[5] Robert Boyd: Radiometry and the Detection of Optical Radiation. New York: Wiley and Son. (1983)

[6] Marc Geese, Ulrich Seger: Critical use cases for video capturing systems in autonomous driving applications Autosens Brussels (2017).

[7] Dieter Meschede, Optik, Licht und Laser 3.Auflage p.384 (2008)

[8] Omnivision, http://www.ovt.com/sensors/OV10640, retrived (2017)

[9] J. Solhusvik, A comparison of high dynamic range CIS technologies for automotive applications, IISW2013

Author Biography

Marc Geese *received a diploma in applied physics from Frankfurt University and then continued with an MPhil in Electrical Engineering at the University of Manchester (UK). In 2013 he received a PhD in Physics from Heidelberg University and joined the Robert Bosch GmbH where he worked as a software engineer on computer vision algorithms for video based ADAS Systems. Since 2014 he is a system architect for optical systems in the predevelopment for Boschs next generation video systems, optimizing the image quality effect chain and developing relevant KPIs. Since 2016 he actively participates in the IEEE P2020 Working group as the leader of the subgroup "Image Quality for Computer Vision on System and Component Level".*

Ulrich Seger *has experience in camera design and application as well as in digital image processing. In his role as senior manager for next generation video equipment he and his team are responsible to identify innovative optical technologies for applications in the next generation ADAS and autonomous driving systems. Prior to his engagement in Bosch since 1999 he's been a researcher at the Institute of Microelectronics Stuttgart developing the first generation of high dynamic range CMOS sensors. He received his diploma in electrical engineering from University of Applied Science in Konstanz for his work on optical character recognition in 1987.*

Alfredo Paolillo *received his MS degree in electronic engineering and the PhD in information engineering from University of Salerno, Italy, in 2000 and 2004, respectively. He was an Assistant and an Associate Professor of Electronic Measurements with the Department of Industrial Engineering, University of Salerno, Italy, from 2003 to 2016. He has been with Robert Bosch GmbH, Leonberg, Germany, since 2016. His work focuses on the development of methods for the verification of image sensors.*

Free access to this paper is brought to you with the generous support of ON Semiconductor.

All research funding for this paper is referenced in the text; unless noted therein, no research funding was provided by ON.