https://doi.org/10.2352/issn.2169-2629.2020.28.46

Optimization of automotive color filter arrays for traffic light color separation

Korbinian Weikl[1,2], Damien Schroeder[2], and Walter Stechele[1]
[1]Department of Electrical and Computer Engineering, Technical University of Munich, Munich, Germany
[2]Automated Driving, BMW Group, Munich, Germany

Abstract

Traffic light (TL) classification is an important feature for automated driving, and it requires correct color separation of the TL signals captured using cameras. A key camera component for the color separation performance is the color filter array (CFA). For common automotive-specific CFAs, we have observed unsatisfactory performance for TL color separation, which indicates the need for an optimization. Based on typical scenarios for TL classification and a set of recorded TL signals, we evaluate the performance of common automotive CFAs. For a quantitative evaluation, we propose a suitable color distance metric. We also propose a method for optimization of the CFA and show that using this method, reference color separation performance can be achieved, trading in only a small amount of sensitivity.

Introduction

With introduction of advanced driver assistance systems (ADASs) and increasingly also automated driving systems (ADSs), cameras have become essential components of the sensor setup in automotive vehicles. The specific requirements of automotive applications, e.g. a very high dynamic range, but also the need for very cost-efficient sensor components for mass production, have led to custom camera designs. A key component for camera performance is the color filter array (CFA). Figure 1 shows common CFAs for driving automation applications.

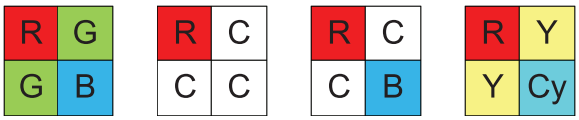

Figure 1. *Common CFA configurations consisting of the red (R), green (G), blue (B), clear (C), yellow (Y) and cyan (Cy) color filters given in Figure 2. RCCC is constructed using the same filters as RCCB.*

We construct the CFAs based on the exemplary, normalized quantum efficiency (QE) data shown in Figure 2. The data is derived from characterization of the ON Semiconductor AR0820 imager [1], which we use as exemplary sensor in this paper. In the form provided by ON Semiconductor [2], it is normalized for the RGGB-G QE. The CFAs are typically used with a global IR cutoff filter, which suppresses the non-negligible QE of all color channels in the near infrared spectrum. For this reason, we consider only the spectral range from 400 nm to 700 nm.

While the well-known RGGB Bayer pattern still remains the reference for visualization applications due to its optimized color reproduction for human vision, the RCCC pattern offers superior sensitivity and still simple recognition of red signals that are especially important for ADAS functions, for example for detection of breaking lights. For more complex applications in driving automation, state-of-the-art filter arrays are RCCB and

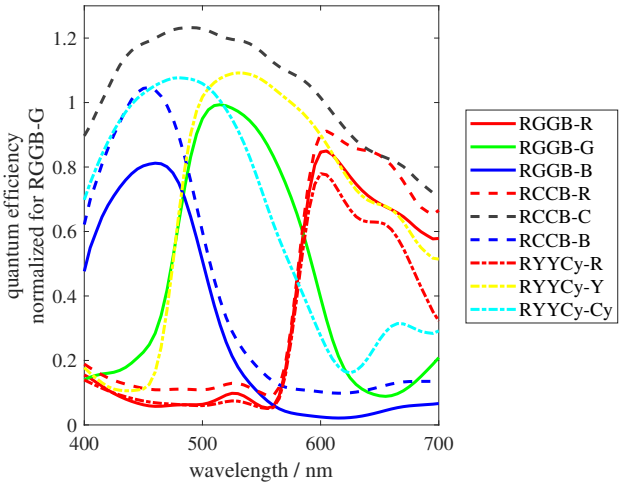

Figure 2. *Normalized QE data based on the ON Semiconductor AR0820 imager [1] for filters of common CFAs, as provided by ON Seminconductor [2]. Color filters are denoted as <CFA configuration>−<filter type>.*

RYYCy, which enable three-channel color vision but still provide extended sensitivity through wideband clear respectively yellow filters. RYYCy is additionally designed to work with a lower resolution blue channel, i.e. more cost-efficient optics can be used which may introduce chromatic aberrations at low wavelength [3]. Other CFA patterns have been proposed for applications that require high sensitivity, using an additional panchromatic [4] respectively wideband [5] channel similar to the clear and yellow pixels in the CFAs described above. However, they have not established themselves in automotive applications.

An inherent limitation of imaging systems that use a small set of filters, which cover an extended range of the light spectrum, is the existence of metamers: Objects that emit or reflect light in a different spectral distribution can still generate the same sensor response because photons of different wavelengths pass through each filter and are integrated in one pixel, making them indistinguishable.

One highly safety-relevant case in automotive applications, in which metamers have to be ruled out, is the separation of traffic light (TL) colors for TL classification. Especially in conditions in which the TL geometry cannot reliably be captured, e.g. in low light conditions at night, the color information is necessary for correct classification of the TL state. This is illustrated in Figure 3.

In this paper, we introduce a color distance metric that we can use to evaluate the color separation performance of CFAs. We also propose a method for optimization of CFAs for color separation, and show that using this method, reference color separation can be achieved for which we only have to trade in a small amount of sensitivity.

Figure 3. *Example images of a crossing from automotive camera, at daytime and nighttime. While the high-contrast TL profiles allow a contour-based state classification in daylight, merely a color-based classification is possible in the dark.*

In the following, Section 2 introduces typical scenarios for automotive TL classification and a set of recorded TL signals, characterizing the conditions of automotive TL classification. In Section 3, we introduce a metric for color separation in order to evaluate the performance of common CFAs with regard to TL color separation. In Section 4, we propose a method for construction of an adapted CFA that enables robust TL color separation on the level of an RGGB CFA while minimizing the loss of sensor sensitivity. The results obtained with such a CFA, compared to the common patterns, are presented in Section 5. Section 6 concludes the findings.

Characterization of scenarios for TL classification

Three scenarios relevant for TL classification shall be considered: TL observation in close proximity of a crossroad, navigation through urban traffic, and far-sighted navigation on a highway. The scenarios are examined based on exemplary detection distance requirements and typical automotive camera optics with a horizontal field of view (FOV) selected for imaging the relevant sections of the environment in each scenario. State-of-the-art automotive sensor setups contain several front cameras to support all these scenarios. Again, as exemplary state-of-the-art automotive imager, we choose the ON Semiconductor AR0820. It has an image resolution of 8 Mpx [1]. We assume camera optics that achieve a uniform pixel density, i.e. a constant number of pixels per degree of FOV. Furthermore, we use a TL that corresponds to the European standard for signal heads [6], having a signal diameter of 200 mm, minimum luminous intensity of $I_{min} = 200$ cd (performance level 2) and an extra wide-angle beam. As compared in [7], the European regulations for TL luminous intensity are also similarly covered in US-American standards.

In Table 1, the resulting object resolutions r_o are given for the considered scenarios and TL diameter. Because of the worst case r_o that is in the same order of magnitude as the size of the filter pattern, i.e. the elemental pattern resolution r_p in a CFA, the problem of TL color separation is generally also a problem of demosaicking. That is, for poorly resolved TLs it has to be considered that the imperfect reconstruction of all spectral channels for each pixel can cause a significantly distorted spectral representation of the TL in the output image. However, as we will show in the following, currently used CFAs may inherently not allow robust color separation: Even for uniform TL signals that are imaged at $r_o \gg r_p$, i.e. a locally constant input signal which does not cause demosaicking artifacts, the sensor responses for different TL signals result in poor color separation. Thus, at this time, we examine the problem without considering demosaicking and evaluate sensor responses using the color channels directly, without spatial interpolation.

The imaging conditions for TL detection do not only depend on the driving scenario, they also vary between different

Table 1. Considered scenarios for TL classification.

scenario	distance	FOV	r_o
Crossroad proximity	20 m	120°	18 px
Urban navigation	100 m	60°	7.3 px
Highway navigation	250 m	30°	5.9 px

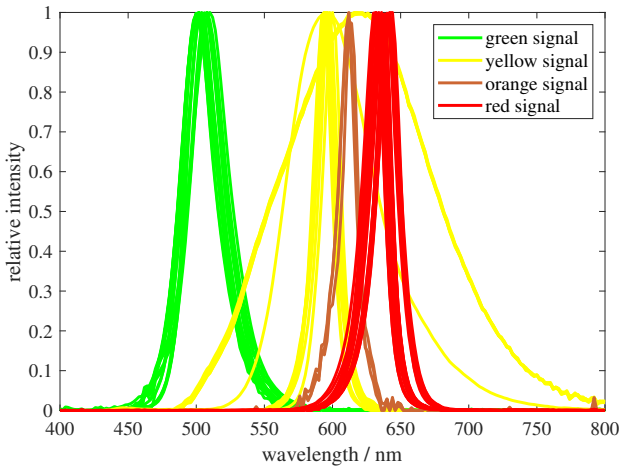

Figure 4. *Light emission spectra of US-American [2] and German traffic light signals.*

TLs, e.g. different used light sources and different color filters on these light sources. To account for these variations, we use provided light emission spectra of eleven TLs in California, USA [2] as well as light emission spectra of five TLs in the Munich area, Germany, of our own data collection. Each TL consists of a green, yellow and red signal. Three of the US-American TLs additionally comprise an orange "don't walk" signal. We use the spectra of all signals, see Figure 4, for construction of a custom color checker chart, see Figure 5.

We then generate the sensor response to the test chart with a camera simulation based on the ISETCAM toolbox [8, 9]. The simulation model comprises of simple optical image generation through equal photon count excitation of the custom color checker, resembling a self-emissive target, and a diffraction-limited optics model, as well as filtering and sensor simulation through the exemplary QE curves and sensor characterization data of the ON Semiconductor AR0820 imager [2]. For simulation, we adapt the normalization of the QE data so that no QE exceeds 1. A sensor model without HDR operation is used and the sensor responses then result as digital numbers in 12 bit integer range, which we refer to as raw sensor data. To obtain consistent sensor responses, we disable ISETCAM's noise simulation capabilities.

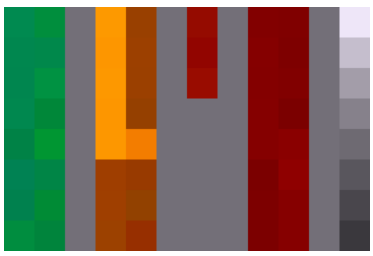

Figure 5. *Custom color checker chart generated from light emission spectra in Figure 4, containing from left to right: green, yellow, orange, and red TL signals; grayscale reference pattern. RGB visualization for illustration only.*

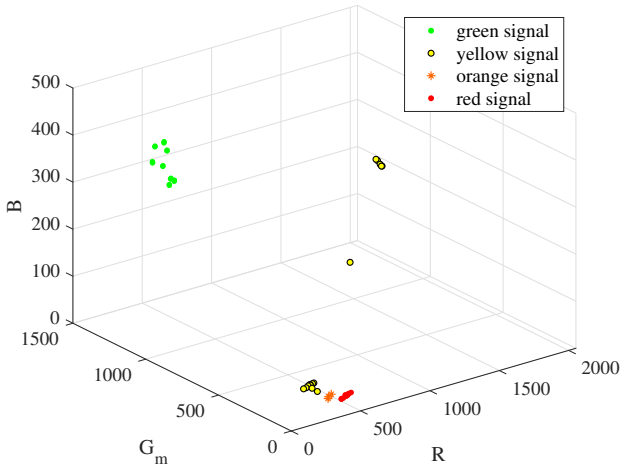

Figure 6. *Point cloud representation of raw sensor responses to TL signals, for RGGB CFA. G_m denotes the mean responses of the two G channels present in the CFA. The representation is based on 12 bit unsigned integer raw data and highlights the relative positions of the point clouds.*

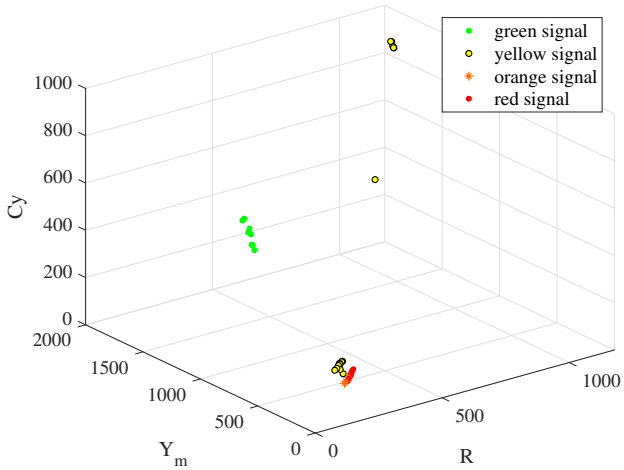

Figure 7. *Point cloud representation of raw sensor responses to TL signals, for RYYCy CFA. Y_m denotes the mean responses of the two Y channels present in the CFA. The representation is based on 12 bit unsigned integer raw data and highlights the relative positions of the point clouds.*

Performance of state-of-the-art color filters

TL classification for driving automation is a task fulfilled by computer vision systems. We examine the color separation based on the raw sensor data, which exhibits how well the CFA produces numerically distinguishable sensor responses to different TL signals. Distinguishable sensor responses are a precondition for color-based classification by computer vision.

Color separation

To quantify the color separation capability, the sensor response can be represented as a point in a space spanned by a set of base vectors, of which each represents the dimension of one of the unique filters in the CFA pattern. Such representations are commonly used for color space coordinates. We denote the number of base vectors as the pattern dimensionality, m_p. Without loss of generality, the vectors shall form an orthogonal basis. Other base vectors can be chosen by applying color space transformations to the sensor responses, e.g. for decorrelation. [10] uses a discriminant analysis method for choosing the best color space for their color separation application. We focus on the raw sensor data for a direct, data-independent evaluation. The sets of measurements of one signal type form point clouds in this space representation. In this paper we use simulated sensor responses to 16 TLs as measurements. Hence, we have four sets for the four signal types, namely green, yellow, orange and red TL signals. The sets contain raw sensor data and form clouds of three points in the case of orange TL signals, and 16 points otherwise. Such representations are shown in Figures 6 and 7. For channels that are present multiple times in a CFA, we use the mean response and count them only once for m_p. That means that for example RCCC has an $m_p = 2$, while RYYCy has an $m_p = 3$.

In this space, we choose the separation distance between two measurements, i.e. between the raw sensor data coordinates, as the L^2 distance between the points that represent them. Using such L^2 distances based on color coordinates for color distance formulae has been an accepted approach since the introduction of the CIE L* a* b* and CIE L* u* v* ΔE formulae, as discussed for example in [11]. Since computer vision systems do not inherently bias some spectral regions for perception of spectral information, however, we choose to not apply weighting factors to the color coordinate differences and hence we cannot use these metrics optimized for human vision directly.

The authors in [12] use a similar metric, called color separation factor, for optimization of color filters. It is based on a sum of squared color coordinate distances. However, they use averages of pixel values, i.e. measurements, taken of the same signal type, and hence measure the average separability. In the application of TL separation, on the contrary, we are interested in the worst-case, i.e. minimum distance between measurements of two signal types. Thus, we also cannot directly use this metric. A new metric is required for evaluation of color separation in the context of TL classification.

For two sets of measurements, we define the metric color separation distance (CSD) as the minimum of the L^2 distances between an element of the first set and an element of the second set. That is, for measurements y of two signal types, given by set $S_1 = \{y_{1,i} : i = 1 \dots l_1\}$ of l_1 measurements of the first signal type and set $S_2 = \{y_{2,j} : j = 1 \dots l_2\}$ of l_2 measurements of the second signal type, we define the CSD as

$$\mathrm{CSD} = \min_{y_{1,i} \in S_1,\, y_{2,j} \in S_2} \sqrt{\sum_{m=1}^{m_p} (y_{1,i,m} - y_{2,j,m})^2}, \qquad (1)$$

where $y_{1,i,m}$ denotes coordinate m of measurement $y_{1,i}$ in the m_p-dimensional space representation.

It must be noted, that in general the CSD may be greater than zero even if the sets of measurements completely overlap in the space representation defined above, since measurements from the two sets still don't necessarily coincide and only in that case would a distance of zero result. However, these insignificant values can easily be detected by defining a lower bound for the CSD, for example given by the noise level present in the measurements. In this paper, we use the metric for relative comparison without a lower bound.

Application to separation of measured TL signals

We use the CSD metric to evaluate the separability of the TL signals introduced in Section 2, and calculate the CSDs based on the simulated sensor responses for the common CFAs. The results are presented in Table 2.

As can be seen in Table 2, the smallest CSDs are found between yellow, orange and red signals. However, orange signals are only used at some TLs in some countries. On a global level,

Table 2. TL signal CSDs for state-of-the-art CFAs; for 12 bit unsigned integer raw data.

signal pair	CSD			
	RGGB	RCCC	RCCB	RYYCy
green / yellow	769	263	316	591
green / orange	888	298	348	657
green / red	949	302	349	644
yellow / orange	70	16	16	26
yellow / red	**138**	**29**	**30**	**32**
orange / red	66	11	11	8

especially yellow / red separation is of high importance. The RGGB CFA performs significantly better in this case than any of the automotive-specific patterns. Therefore, we choose RGGB as our baseline for color separation, and the yellow / red CSD as our relevant performance metric for color separation.

Color filter optimization method for robust color separation

We choose the starting point for optimization of the CFA for TL color separation to be the RYYCy pattern, due to its advantages for automotive imaging: High sensitivity for low light imaging and optimized performance in combination with cost-efficient camera optics. Regarding sensitivity, we choose RYYCy as our baseline. Alternatively, the optimization could be applied to RCCB.

The concept for the optimization is to use the redundancy of the Y channels in RYYCy as one degree of freedom, and introduce a small modification to the filter transmissivity respectively QE curve of one of them, in order to allow robust separation of the considered signals while trading in a small amount of total sensitivity. It should be kept in mind that since this filter optimization implies a hardware modification at design time, it is only appropriate for high priority use cases that exhibit invariant spectral properties. TL color separation in automotive imaging is such a case, due to the significance of TLs for ADAS and ADS functions and the lighting-independent self-emissive signals. The optimization we propose is based on a modification as small as possible, but the performance for other use cases still has to be verified which remains as future work.

Inspecting the RGGB QE curves in Figure 2, we observe that the transition between the green and red filter well corresponds to exclusively sensing the red TL signals which peak at 640 nm, see Figure 4: The RGGB-G QE is nearly minimal and RGGB-R QE nearly maximal at that wavelength.

To achieve the same separation capability using the modified Y channel, we introduce a stopband as present in RGGB-G at 640 nm. We model this stopband in several ways, denoting the modified channel as N_X and the resulting CFA as RYN_XCy. All models have in common, that the stopband is generated by transitions between the RYYCy-Y and RGGB-G QE curves. We choose to construct the modified channel in that way with the intention of reflecting producibility constraints. However, we are aware that for an implementation further effort for tuning of the model will be required.

The modified channels are shown in Figure 8, using the same normalization as in Figure 2. To select one of the models, we consider the requirements of a yellow / red CSD at the level of the RGGB CFA, as well as a minimimal loss of sensitivity. We present the evaluation of the modified channels from Figure 8 regarding these criteria in section 5.

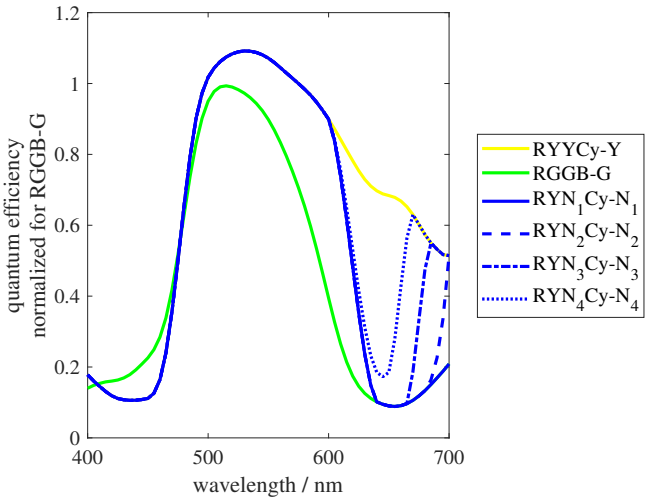

Figure 8. *Normalized QE of modified channels RYN_XCy-N_X, constructed based on RYYCy-Y by introducing transitions to RGGB-G. The transitions form a stopband that attenuates the red TL signals. In the case of N_4, the stopband is very narrow, resulting in reduced attenuation.*

Results

Using the RYN_XCy CFAs adapted to the recorded TL signals, we again simulate sensor responses and calculate the resulting CSDs. They are given in Table 3.

Table 3. TL signal CSDs for proposed CFAs; for 12 bit unsigned integer raw data.

signal pair	CSD, RYN_xCy			
	RYN_1Cy	RYN_2Cy	RYN_3Cy	RYN_4Cy
green / yellow	643	643	643	644
green / orange	750	751	751	748
green / red	814	815	815	800
yellow / orange	54	54	53	47
yellow / red	**168**	**168**	**168**	**142**
orange / red	105	105	106	90

Noticeably, all RYN_XCy CFAs have yellow / red CSDs at least as high as RGGB. To match the yellow / red color separation performance of RGGB, CSD_{yr} = 138, see Table 2, it is sufficient to use RYN_4Cy with CSD_{yr} = 142, see Table 3. RYN_4Cy also has the most narrowband stopband and thus the least loss of sensitivity. We therefore select that CFA for further evaluation.

For comparison to Figures 6 and 7, the space representation of the sensor responses using this pattern is also given in Figure 9. We observe that the point clouds especially of the yellow, orange and red TL signals for RYN_4Cy in Figure 9 are much more clearly separated than for RYYCy in Figure 7, more similar to RGGB in Figure 6.

As our further criterion for the CFA performance, we compare the total QE of the modified filter RYN_4Cy-N_4 to that of RYYCy-Y, our baseline for sensitivity. In the considered wavelength range of 400 nm to 700 nm, N_4 reaches 90.2 % of the reference total QE. Hence, the loss of sensitivity of the channel is less than 10 %. Since we modify only one of the four channels in the CFA, the overall loss of sensitivity of the CFA is lower still. Thus, using the method of introducing a small application-specific modification to one of the wideband channels of a CFA, we can achieve reference color separation performance, trading in only a small amount of sensitivity.

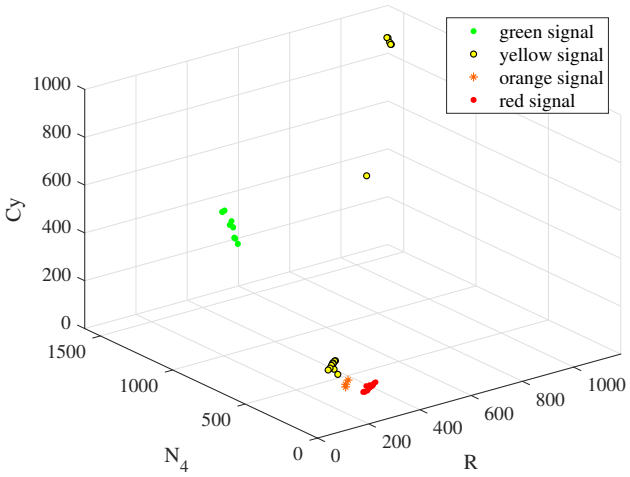

Figure 9. *Point cloud representation of raw sensor responses to TL signals, for RYN$_4$Cy CFA. The dimension that represents the Y channel is not shown. The representation is based on 12 bit unsigned integer raw data and highlights the relative positions of the point clouds.*

Conclusion

In this paper, we have proposed a method for optimization of CFAs for color separation, based on a small application-specific modification of a wideband channel in the CFAs. To quantify the color separation performance of different CFAs, we have also proposed a color separation metric for sets of measurements of different signal types. We have used the metric and optimization method for the application of TL classification in automotive imaging systems, which is an important feature for ADAS and ADS functions.

As future work, we want to consider using small modifications in both redundant, e.g. RYYCy-Y, wideband channels that result in the same mean response as the two original channels. We also want to explore producibility constraints to incorporate them in the optimization. As a holistic approach, we want to include the proposed CFA optimization in a simulation of an automotive imaging pipeline that can measure the performance based on the output metrics of a benchmark computer vision.

Acknowledgments

The authors would like to thank James Tornes, David Jasinski and Paul Kane of ON Semiconductor for providing the traffic light spectral data, and for discussions about metrics for color separation performance of different CFA patterns. ON Semiconductor and the ON Semiconductor logo are registered trademarks of SCILLC or its subsidiaries in the United States and/or other countries.

References

[1] ON Semiconductor, "AR0820: 1/2-inch CMOS digital image sensor," Tech. datasheet AR0820/D rev. 1, Apr. 2020.

[2] Paul Kane, "ON Semiconductor AR0820 sensor characterization and traffic light measurements," Private communication, Nov. 2019 - Apr. 2020.

[3] O. Eytan and E. Belman, "High-resolution automotive lens and sensor," Patent application publication US 2019/0 377 110 A1, 12 Dec. 2019.

[4] M. Kumar, E. O. Morales, J. E. Adams, and W. Hao, "New digital camera sensor architecture for low light imaging," in *Proc. ICIP*, 2009, pp. 2681–2684.

[5] M. Parmar and B. A. Wandell, "Interleaved imaging: An imaging system design inspired by rod-cone vision," in *Proc. IS&T/SPIE Electronic Imaging*, ser. Proceedings of SPIE, vol. 7250, 2009.

[6] "Traffic control equipment – signal heads," German version EN 12368:2015, DIN Deutsches Institut für Normung e. V., Berlin, Germany, Sep. 2015.

[7] J. D. Bullough, P. R. Boyce, A. Bierman, C. M. Hunter, K. M. Conway, A. Nakata, and M. G. Figueiro, "Traffic signal luminance and visual discomfort at night," *Transportation Research Record*, vol. 1754, no. 1, pp. 42–47, 2001.

[8] J. E. Farrell, P. B. Catrysse, and B. A. Wandell, "Digital camera simulation," *Applied optics*, vol. 51, no. 4, pp. A80–A90, 2012.

[9] Stanford Center for Image Systems Engineering, "ISETCAM," Jan. 2020. [Online]. Available: https://github.com/ISET/isetcam

[10] N. Kehtarnavaz and A. Ahmad, "Traffic sign recognition in noisy outdoor scenes," in *Proc. Intelligent Vehicles '95*, 1995, pp. 460–465.

[11] M. R. Luo, "Development of colour-difference formulae," *Review of Progress in Coloration and Related Topics*, vol. 32, no. 1, pp. 28–39, 2002.

[12] K. Heidary and H. J. Caulfield, "Spectral sensitivity design for maximum colour separation in artificial colour systems," *IET Image Processing*, vol. 3, no. 3, pp. 135–146, 2009.

Author Biography

Korbinian Weikl received his B.Sc. from the Hamburg University of Technology in 2014 and M.Sc. from the Technical University of Munich (TUM) in 2017. In 2017 he joined Rohde & Schwarz GmbH & Co. KG as development engineer for signal processing. Since 2019, he has been working as doctoral researcher affiliated with BMW Group and supervised by Prof. Dr.-Ing. Walter Stechele of TUM. His research focusses on imaging technologies for full driving automation.

Damien Schroeder received the Dipl.-Ing. degree in electrical engineering and information technology from the Technical University of Munich (TUM), Germany and the Diploma degree from Supélec, France. He joined the Chair of Media Technology at TUM in 2012, from which he received the Dr.-Ing. degree in 2017. Since 2017, he is a project manager for camera systems for automated driving with the BMW group. His research interests include video coding and camera technology.

Walter Stechele received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from the Technical University of Munich, Germany, in 1983 and 1988, respectively. In 1990 he joined Kontron Elektronik GmbH, a German electronic company, where he was responsible for the ASIC and PCB design department. Since 1993 he has been Academic Director at the Technical University of Munich. His interests include visual computing and robotic vision, with focus on Multi Processor System-on-Chip (MPSoC).