https://doi.org/10.2352/EI.2022.34.16.AVM-161

Potentials of combined visible light and near infrared imaging for driving automation

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]Driving Experience, BMW Group, Munich, Germany

Abstract

An extension of automotive imaging from the visible (VIS) to the near infrared (NIR) spectrum is promising for driving automation applications because the technology is readily available and offers potential benefits in low visibility conditions, in low light conditions with active illumination, and by collection of complementary data. We propose the evaluation of VIS-NIR imaging in simulation using an extended version of our camera simulation and optimization framework. Our extended framework generates realistic spectral irradiance data of synthetic scenes in the VIS and NIR spectral range and includes physically based camera models with characteristic increased NIR sensitivity of VIS-NIR CMOS imagers, modified automotive VIS-NIR color filter arrays and adapted image processing. We evaluate the reproduction of potential benefits of VIS-NIR imaging in our simulated camera images using exemplary night time and daylight traffic scenes, and discuss further extensions for creation of a well-balanced VIS-NIR dataset for quantitative evaluation.

Introduction

A large variety of *driver assistance* and *partial driving automation* functions are available in series production vehicles today, and first systems for *conditional* and *high driving automation* are presented and brought into production. Yet, to reach the highest level of automation [1] defined as *full driving automation*, there are still various challenges ahead. Many of the technical challenges can be categorized as either complexity challenges (e.g. navigation in dense urban environments with many traffic participants) or as operational design domain (ODD) limits (e.g. conditions at night time or with low visibility due to rain, fog, etc.). To enable full automation, the imaging system of the vehicle consisting of cameras and a computer vision (CV) system must be capable of replacing the visual system of a human driver in any of these scenarios and conditions [2].

One degree of freedom to increase the performance of the imaging system is to extend the range of the electromagnetic spectrum that it senses. An extension from the visible (VIS) to the near infrared (NIR) spectrum is a promising candidate technology because silicon-based pixels are inherently sensitive to NIR radiation and CMOS imagers with NIR-sensitive color channels are already in use e.g. in surveillance cameras. There are three potential benefits that a combined VIS-NIR imaging system offers for driving automation applications:

- Higher wavelength radiation can be less subject to attenuation and scattering in fog, rain, snowfall, smog and other low visibility conditions. These effects are difficult to model and predict and are outside the scope of this work. [3] presents an experimental evaluation of several spectral bands for automotive detection tasks in low visibility conditions.
- At night time and in other low-light conditions, active illumination with higher output power in the NIR spectrum can be used to achieve higher illumination ranges because NIR eye safety limits are higher than those for VIS light sources. We provide details on eye safety limits in the framework section.
- The additional NIR color channel allows the collection of scene information that is complementary to the VIS information because some materials reflect NIR light different to VIS light. The additional information can resolve metamers and make the CV more reliable. An example is shown in Fig. 1.

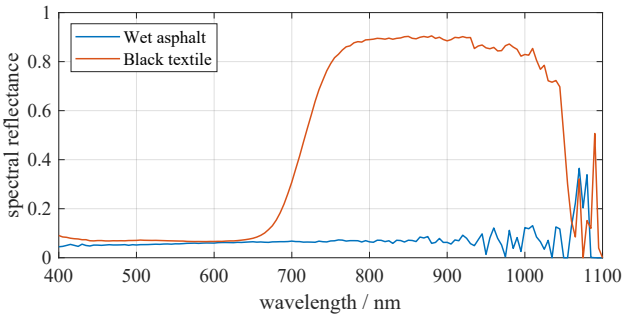


Figure 1. *Spectral reflectance measurements of wet asphalt and a black textile. The materials exhibit a very similar reflectance in the VIS spectrum, but a good contrast in the NIR spectrum. This property can be used to make detection of pedestrians more reliable with VIS-NIR imaging.*

It remains to be researched how these potential benefits can be best applied to solve complexity and ODD limit challenges in driving automation.

An early implementation of NIR imaging for pedestrian detection with a focus on night time conditions can be found in [4]. The authors use a monochrome CCD imager, standard vehicle headlights (at the time of publication) and classic CV algorithms. An implementation of a VIS-NIR imaging system for automotive applications is presented in [5]. The proposed solution is based on a modified RGB imager with an additional panchromatic, i.e. clear pixel. The authors construct separate VIS (RGB) and NIR images from the raw sensor data and subsequently enhance the RGB image for human observers by a fusion of VIS and NIR luminances. [6] describes a physically based camera simulation

system that is suitable for extension to VIS-NIR imaging. The authors motivate the use of synthetic spectral data as physically accurate input to automotive injection hardware-in-the-loop systems, and an extension of this spectral pipeline to the VIS-NIR range for pedestrian detection applications.

Similarly, we present an extension of our camera simulation and optimization framework for simulation of VIS-NIR imaging systems. However, we focus on extending synthetic scenes to reproduce specific VIS-NIR imaging properties, and on modeling a VIS-NIR camera as an extension of a state-of-the-art automotive camera. We include details on the extended database for synthetic scene generation and on the extended camera models. We explore potentials of automotive VIS-NIR imaging systems in daylight and night time conditions using exemplary scenes and discuss the necessary steps to construct a well-balanced VIS-NIR dataset for quantitative evaluation of VIS-NIR imaging.

Framework

We extend our framework to enable the simulation and optimization of VIS-NIR imaging systems. Our framework targets end-to-end optimization of the imaging system using synthetic traffic scenes, physically based camera models, and a benchmark CV system [2]. An overview of the framework's building blocks with extensions for VIS-NIR imaging is depicted in Fig. 2. The

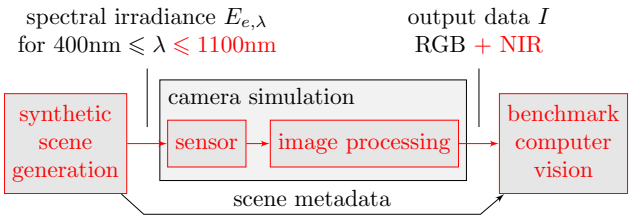


Figure 2. *Framework overview: Extensions for VIS-NIR imaging are marked in red. The original framework is presented in [2].*

following extension are required for VIS-NIR imaging:

- To generate realistic spectral irradiance data in the wavelength range from 400 nm to 1100 nm, synthetic scenes must use hyperspectral material models and spectral radiance distributions of light sources that are defined at least across the same spectrum. Wavelength sampling of the rendering core must be configured to the same spectral range.
- VIS-NIR camera models must include sensor models with state-of-the-art increased NIR sensitivity and adapted color filter arrays (CFAs) for distinction of VIS and NIR signals.
- The image processing of VIS-NIR camera models must be adapted to the imager's color channels, i.e. implement a generalized demosaicing algorithm and custom color channel transformations to process VIS and NIR signals.
- Finally, image pre-processing steps must be implemented that convert VIS and NIR signals to an image format that can be used by an adapted benchmark CV system.

Synthetic scene generation

We collect a library of hyperspectral material models that we can use for a large part of the content of our ISET3d [7, 8] based synthetic automotive traffic scenes. The material models are based on reflectance measurements in the VIS-NIR range, using a spectroradiometer and a Polytetrafluoroethylene (PTFE)-based white reference target equivalent to the measurements presented in [9]. However, we do not resolve the angular reflectance distribution of the materials. Instead, we use spectral reflectances measured at an arbitrary angle for parameterization of existing material models in PBRT [10, 11] using the models' diffuse reflection parameters. This greatly simplifies material reflectance measurements and material models. The simplification comes at the cost of limited material photorealism, but we consider this a reasonable trade-of for our application of evaluating fundamental benefits of imaging systems in different spectral bands. Our hyperspectral material library contains:

- Asphalt, concrete, and cobble surfaces
- car paint (66 shades)
- Bark and leafage
- Textiles and skin
- General materials, e.g. tire rubber

A limitation for the use of hyperspectral material models are RGB textures that are used e.g. for building facades and road surfaces with road markings in ISET3d scenes. These textures mimic geometric features and can hence not be replaced by uniform materials without loss of scene details. We keep the original textures of building facades but replace the road surface textures by hyperspectral material models to achieve realistic contrast of objects against the road. Road markings are not relevant for our object detection benchmark CV.

For realistic light interaction with our hyperspectral material models we model the spectral radiance distributions of two main light sources: For night time scenes, we construct spectral radiance curves for ego vehicle headlights based on eye safety limits in the VIS and NIR spectra. As starting point we use an equal energy radiator in the spectrum from 400 nm to 1100 nm. We then define a threshold wavelength that marks the separation between VIS and NIR illumination, and scale the radiance in both spectra to reach the applicable eye safety limits defined in the appendix of EN 62471 [12]. Solutions of this spectral radiance model are depicted in Fig. 3.

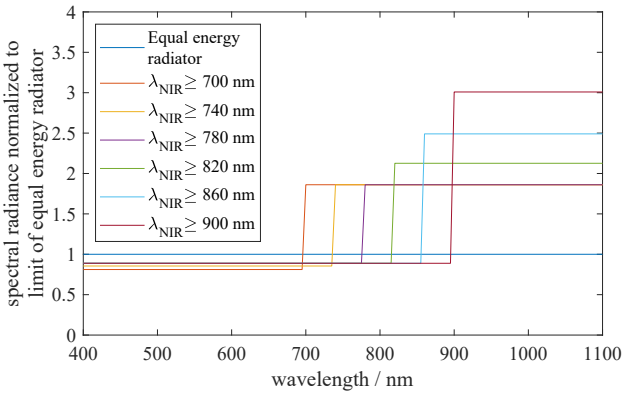


Figure 3. *Solutions of our headlight spectral radiance model based on eye safety limits [12]. Depending on the threshold wavelength, NIR illumination can have a spectral radiance limit that is more than 3x the limit of VIS illumination while reducing the spectral radiance limit of VIS illumination less than 20% compared to the equal energy radiator. We assume a 10 s exposure and a 40 ° wide light source.*

For daylight scenes the primary light source is natural illumination. ISET3d uses PBRT's *infinite* light source model with RGB sky maps to model this source. We construct a solar radiance scale factor based on the ASTM G-173-03 reference spectrum at global tilt [13]. To account for the 3 bin radiance sampling and scaling by the RGB sky map within the VIS spectrum, we set the radiance scale factor in the spectrum from 400 nm to 700 nm to its mean value in this spectral range. Fig. 4 shows the resulting radiance spectrum across the VIS-NIR spectrum. An alternative

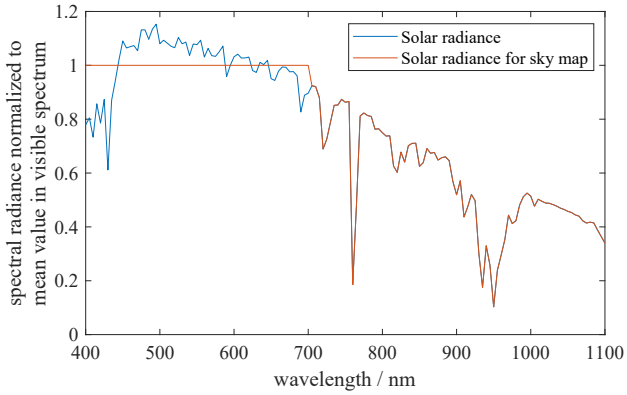


Figure 4. *Solar radiance model based on the ASTM G-173-03 reference spectrum at global tilt [13]. We adapt the reference spectrum to account for the use of RGB sky maps for natural illumination modeling, but preserve the realistic spectral radiance curve in the NIR spectrum.*

would be to use an analytic sky mdel with support of the VIS-NIR spectral range, as described for the VIS range in [6].

Sensor models

To model the characteristics of state-of-the-art VIS-NIR imagers, we modify the quantum efficiency (QE) data of an automotive imager to reproduce the increased NIR sensitivity that can be found in existing CMOS VIS-NIR imagers. As reference for similar VIS-NIR QE data see [5]. We use the unmodified imager QE with RYYCy and RCCB CFAs as well as a global IR cut filter as reference for automotive VIS-only imaging. For the modified VIS-NIR imager we construct adapted CFAs by replacing one of the Y respectively C channels by a modified filter that is sensitive either in only the VIS (Yv, Cv) or the NIR (Yn, Cn) spectrum. Our new CFAs are hence named RYYvCy, RYYnCy, RCCvB, and RCCnB. As examples, the reference RYYCy and new RYYnCy pixel QEs are shown in Figs. 5 and 6. Higher NIR QE can be achieved in future VIS-NIR imagers e.g. using specialized pixel geometries [14] and upper wavelength limits beyond the silicon bandgap become possible on silicon imagers with plasmonic photodetectors as presented e.g. in [15]. For simulation of the exemplary scenes shown in the results section, we run the imager models in single dynamic range (SDR) mode.

Image processing and interface to CV

As first step of our image processing model we implement a generalized demosaicing algorithm based on fast curvature based interpolation [16] of each color plane. Subsequently we implement simple color channel transformations to generate VIS-NIR RGB and separate NIR monochrome output images. In literature,

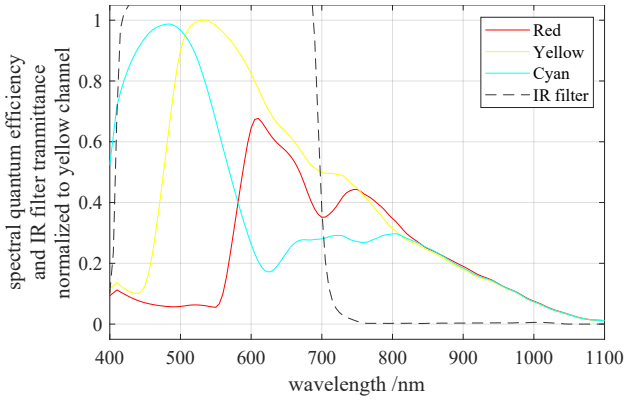


Figure 5. *QEs of all pixels and IR cut filter transmittance for RYYCy imager model. Data anonymized by QE smoothing and normalization.*

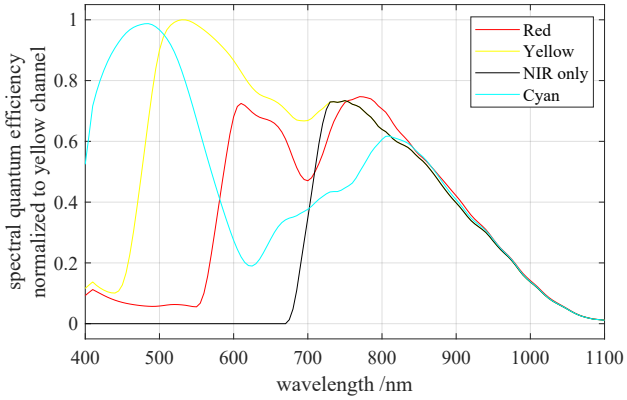


Figure 6. *QEs of all pixels for RYYnCy imager model. Data anonymized by QE smoothing and normalization*

many proposals exist for VIS-NIR color correction and image fusion that are optimized for human vision, e.g. [5, 17, 18]. Which color corrections and VIS-NIR image formats are best suited for CV input data remains future work. The final step of our image processing model is a dynamic range compression.

Results

Using our extended framework we generate VIS-NIR spectral irradiance data of exemplary synthetic traffic scenes and process these with our reference VIS-only, as well as new VIS-NIR camera models. We use these exemplary scenes to evaluate the reproduction of the expected benefits of VIS-NIR imaging.

Results for a night time scene are shown in Fig. 7. The RYYnCy VIS-NIR RGB image does not exhibit a higher range of active illumination than the reference RYYCy VIS-only RGB image even though the ego vehicle headlights emit a NIR spectral radiance at a level of more than 2x the level of the VIS spectral radiance. Moreover, a reduced sensitivity to natural illumination by the night sky can be seen in the periphery of the image.

Results for a daylight scene are shown in Fig. 8. The RYYnCy NIR image shows the increased contrast of pedestrians against the environment when compared to the reference RYYCy RGB image, due to the high NIR reflectance of some textiles. The rest of the scene content is reproduced similarly in both images.

(a) RYYCy - RGB

(b) RYYnCy - RGB

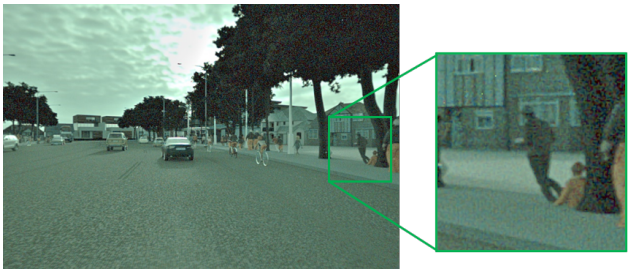
Figure 7. *Simulated camera images of an exemplary night time scene. The spectral radiance model with NIR illumination threshold wavelength of 700 nm was used for the ego vehicle headlights.*

Discussion

Our results for night time scenes do not confirm the expected benefit of VIS-NIR imaging that higher active illumination ranges can be achieved. The similar ranges of VIS and VIS-NIR cameras can be explained by the still lower QE of the imager in the NIR spectrum compared to the VIS spectrum, which compensates the higher NIR radiance of active illumination. The reduced sensitivity of the VIS-NIR camera to night sky natural illumination can be explained by the missing second Y channel in the CFA but should be validated using more sophisticated color correction methods.

Our results for daylight scenes do confirm the expected benefit of complementary information for metamerism resolution in the example of pedestrians' clothing. The similar reproduction of other scene content in VIS and NIR spectral bands also indicates that an adaptation of existing CV systems to NIR image input is feasible.

To better understand the performance of VIS-NIR imaging in night time scenes, more models of artificial light sources should be added to the database. For example, street lights, illuminated buildings and the head- and taillights of other traffic participants should be taken into account. Narrow-band light sources in VIS and/or NIR spectrum should be modeled as alternative ego vehicle headlights, which will allow higher peak spectral radiance compared to our wide-band spectral radiance model shown in Fig. 3. Different angular radiance distributions should also be modeled for ego vehicle headlights. A variety of color correction methods and output image formats should be implemented to optimize utilization of the VIS-NIR information by the CV system. These extensions can help to construct a well-balanced dataset and a variety of VIS-NIR camera models for a quantitative evaluation of VIS-NIR imaging potentials for driving automation.



(a) RYYCy - RGB

(b) RYYnCy - NIR

Figure 8. *Simulated camera images of an exemplary daylight scene with highlight on a pedestrian.*

Conclusion

We extend our camera simulation and optimization framework to enable simulation of VIS-NIR imaging systems. We present details on our extensions of synthetic scene generation, sensor models and image processing models that are required for physically based VIS-NIR camera simulation. We evaluate which properties and potential benefits of VIS-NIR imaging we can reproduce in exemplary simulated camera images using our models, and discuss further extensions for compilation of a well-balanced VIS-NIR dataset.

In future work we want to perform a quantitative evaluation of VIS-NIR imaging using a dataset of our synthetic scenes as input, and output metrics of our benchmark CV system as performance indicators. Additionally, we want to add physically based models of low visibility conditions to include the third potential benefit of VIS-NIR imaging in our simulations.

References

[1] "Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles," SAE International, Warrendale, PA, USA, Recommended Practice J3016, Jun. 2018.

[2] K. Weikl, D. Schroeder, D. Blau, Z. Liu, and W. Stechele, "End-to-End Imaging System Optimization for Computer Vision in Driving Automation," in *Proc. Electronic Imaging Symp.*, 2021.

[3] N. Pinchon, M. Ibn-Khedher, O. Cassignol, A. Nicolas, F. Bernardin, P. Leduc, J.-P. Tarel, R. Bremond, E. Bercier, and G. Julien, "All-weather vision for automotive safety: which spectral band?" in *Proc. SIA Vision Int. Conf.*, 2016.

[4] A. Broggi, R. I. Fedriga, A. Tagliati, T. Graf, and M. Meinecke, "Pedestrian detection on a moving vehicle: an investigation about near infra-red images," in *Proc. Intelligent Vehicles Symp.* IEEE, 2006, pp. 431–436.

[5] D. Hertel, H. Marechal, D. A. Tefera, W. Fan, and R. Hicks, "A low-cost VIS-NIR true color night vision video system based on a wide

dynamic range CMOS imager," in *Proc. Intelligent Vehicles Symp.* IEEE, 2009, pp. 273–278.

[6] R. Molenaar, A. van Bilsen, R. van der Made, and R. de Vries, "Full spectrum camera simulation for reliable virtual development and validation of ADAS and automated driving applications," in *Proc. Intelligent Vehicles Symp.* IEEE, 2015, pp. 47–52.

[7] Stanford Center for Image Systems Engineering, "ISET3d," 2020. [Online]. Available: https://github.com/ISET/iset3d

[8] Z. Liu, M. Shen, J. Zhang, S. Liu, H. Blasinski, T. Lian, and B. Wandell, "A system for generating complex physically accurate sensor images for automotive applications," in *Proc. Electronic Imaging Symp.*, 2019, pp. 53-1–53-6.

[9] M. Colomb, P. Duthon, and F. Bernardin, "Spectral reflectance characterization of the road environment to optimize the choice of autonomous vehicle sensors*," in *Proc. Intelligent Transportation Systems Conf.* IEEE, 2019, pp. 1085–1090.

[10] M. Pharr, "pbrt-v3," 2020. [Online]. Available: https://github.com/mmp/pbrt-v3

[11] M. Pharr, W. Jakob, and G. Humphreys, *Physically Based Rendering*, 3rd ed. Cambridge, MA, USA: Morgan Kaufmann, 2016.

[12] DIN, and VDE, "Photobiological safety of lamps and lamp systems (IEC 62471:2006, modified); german version EN 62471:2008," Berlin, Germany, Mar. 2009.

[13] U.S. National Renewable Energy Laboratory (NREL), "Reference air mass 1.5 spectra: ASTM G-173-03 reference spectra." [Online]. Available: https://www.nrel.gov/grid/solar-resource/spectra-am1.5.html

[14] E. P. Devine, W. Qarony, A. Ahamed, A. S. Mayet, S. Ghandiparsi, C. Bartolo-Perez, A. F. Elrefaie, T. Yamada, S.-Y. Wang, and M. S. Islam, "Single microhole per pixel in CMOS image sensors with enhanced optical sensitivity in near-infrared," *IEEE Sensors Journal*, vol. 21, no. 9, pp. 10 556–10 562, 2021.

[15] B. Desiatov, I. Goykhman, N. Mazurski, J. Shappir, J. B. Khurgin, and U. Levy, "Plasmonic enhanced silicon pyramids for internal photoemission schottky detectors in the near-infrared regime," *Optica*, vol. 2, no. 4, pp. 335–338, 2015.

[16] A. Giachetti and N. Asuni, "Fast artifacts-free image interpolation," in *Proc. British Machine Vision Conf.*, 2008, pp. 13.1–13.10.

[17] S. Jee and M. G. Kang, "Sensitivity improvement of extremely low light scenes with RGB-NIR multispectral filter array sensor," *Sensors*, vol. 19, no. 5, 2019, Art. no. 1256.

[18] Z. Han, W. Jin, L. Li, X. Wang, X. Bai, and H. Wang, "Nonlinear regression color correction method for RGBN cameras," *IEEE Access*, vol. 8, pp. 25 914–25 926, 2020.

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 2019 he joined the BMW Group, initially as doctoral researcher supervised by Prof. Dr.-Ing. Walter Stechele of TUM, and currently as sensor specialist for camera systems. His research focuses on advances in 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).