2010 IEEE Intelligent Vehicles Symposium
University of California, San Diego, CA, USA
June 21-24, 2010

Evaluating Sensor Effects on Perception Performance

Stefan Franz
Daimler AG, Ulm, Germany
stefan.franz@daimler.com

Dieter Willersinn
Fraunhofer IOSB, Karlsruhe, Germany
dieter.willersinn@iosb.fraunhofer.de

Kristian Kroschel
Fraunhofer IOSB, Karlsruhe, Germany
kristian.kroschel@iosb-extern.fraunhofer.de

***Abstract*— Performance of perceptive systems depends on the quality of the input data. In this contribution, an approach to evaluate perception performance as a function of quality of the sensor data is presented. Standardized quality metrics support the imaging sensors performance measurement. Several imaging setups are analyzed with real world experiments. The output of each setup is processed offline to track down performance differences with respect to the quality of sensor data. An adapted measurement is calculated to measure the sensor performance with respect to the data quality for the involved perceptive components. The measured performance is assessed by processing the data of different simultaneously recorded imaging setups for the task of feature extraction of road lanes.**

***Index Terms*— image quality, performance evaluation, perceptive components, sensor evaluation**

I. Introduction

Perceptive components of recognition systems including sensors and processing units are the components responsible for acquiring and processing data, as well as extracting features of the environment. In combination with corresponding computational environment models, object generation and situation analysis is performed to generate the final perception result.

The evaluation and assessment of the performance of recognition systems is a permanent challenge for developers. With a constantly increasing number of variables, the required system complexity impedes the identification and measurement of performance parameters.

In this correspondence, an approach is presented to describe the sensor performance as a function of its data quality and relate it to the performance of dependent perception results. With higher performance of the data acquisition component, the data quality is improved and the probability of a successful feature extraction is raised. A better feature extraction improves the perception performance, which ultimately increases the overall performance of the recognition system. The problem is to identify and measure the impact of the sensor performance on the data quality and the perception result. The task is performed for cameras with different combinations of imaging sensors and lenses. With standardized tests their performance difference is measured and used as reference to relate the results of a data quality measurement for the perception task of a road lane recognition system.

Different institutions have proposed concepts for standardized tests and measurements for the performance evaluation of imaging sensors based on periodic signals, random noise and other inputs [1], [2]. From these proposed measurement methods the edge-gradient analysis provides the advantage of simple experimental setups, test charts and is still able to produce comparable results from different test scenarios [3]. The analysis results in an estimate of the modulation transfer function (MTF) of the measured imaging sensor, which is widely used to assess and compare performance. An overview of the state of the art in image quality assessment is given by Angelis [4]. Measures that take the wide range of influences of automotive video footage into account have been proposed, which is useful to check whether selected sensors meet minimal quality criteria for the target applications [5]. For the specific task of lane recognition, the determination of a suitable sensor configuration has been investigated in [6]. Providing a concept and measures for the detailed performance dependent analysis of perceptive components is the target of this work.

II. Sensor performance measurement

The focus of this paper is in the field of camera based systems. The quality of an imaging component can be described as the accuracy to represent scene detail in the resulting image data. Intensity and contrast in a digital image decrease, as the distance between two contrast edges decreases, or the spatial frequency increases. The capability of an imaging device to capture the information content of an object as a function of spatial frequencies is given by the modulation transfer function (MTF).

High spatial frequencies correspond to fine image detail. Therefore, the more detailed an object and the lower the MTF of a system at high frequencies, the lower the ability to resolve these details in the image. From the MTF the limit of resolution can be determined. This limit depends on the characteristics of all the involved optoelectronic components. The system MTF is a combination of the modulations occurring in all the involved components affecting the digital image. From the prominent role of lenses and imagers to small interference factors of cables for example, which can be summarized in a noise term:

$$MTF_{sys} = MTF_{lens} \cdot MTF_{imager} \cdot MTF_{noise} \quad (1)$$

The ISO12233 methodology [1] has been established in order to provide a fast MTF measurement method based on only one image. In such a standardized way, the MTF data from various digital input devices may be easily reproduced and compared. The MTF, in general, is a graph of the intensity measure in gray level percentage versus spatial frequency, resulting in a map of image intensities for various frequencies.

Spatial frequency is typically measured in lp/mm, describing the ability to discern the number of contrast pairs of lines appearing within a millimeter. For digital cameras the normalized unit cycles per pixel (c/p) is more appropriate to account for the variety of sensor sizes. Fine image detail is represented in high spatial frequencies. The more distinct the response, the finer the detail and in the end, the sharper the image. The results of the MTF measurement gives an insight on how well an imaging component performs the task of resolving contrast and detail of a scene. Especially fine details and features with small intervals gain from higher MTF performance. This higher resolving capacity directly affects the performance of perceptive algorithms. The MTF measurement provides a reference to compare and assess the quality measurements of real data for specific perception tasks.

III. Perception performance measurement

To relate imaging system performance to the performance of the targeted perception task, a measure is needed that represents the changes in quality of the input data for the perceptive component that performs the task. Road lanes represent vertically oriented contrast edges in images and can be filtered with classical edge detection algorithms, like sobel filtering or enhanced variants, with respect to orientation and magnitude [7]. In one line of an image a road marking is described as a contrast step. This information can be transformed into a deterministic signal

$$x(t) = x_0 \cdot rect(t/T) \tag{2}$$

with signal length T defined by the width of the marking that should be extracted. Each imaging system has a different response to the carrier signal, based on the different MTF. The quality of the transmitted signal is described by the signal-to-noise ratio (SNR). A matched filter with the impulse response

$$h(t) = c \cdot x(T-t) \tag{3}$$

is able to produce the maximum SNR for a given input signal at time $t = T$. The output of a filter with response function $h(t)$ excited by $x(t)$ is

$$y(t) = \int_{-\infty}^{\infty} h(t-\tau)x(\tau)d\tau \tag{4}$$

The signal is assumed to be superimposed by a white noise process $N(t)$ with the sample function $n(t)$ and the auto covariance function $c_{NN}(\tau) = N_w \cdot \delta_0(\tau)$. The SNR at the output of the filter at time $t = T$ is described with the ratio of the squared source signal and the square mean of the noise process, which equals its variance.This leads to

$$\begin{aligned} SNR &= \frac{\left(\int_{-\infty}^{\infty} h(T-\tau)x(\tau)d\tau\right)^2}{N_w \int_{-\infty}^{\infty} h^2(\tau)d\tau} \\ &= \frac{|\left(h(t) * x(t)\right)^2|_{t=T}}{N_w \cdot \int_{-\infty}^{\infty} h^2(t)dt} \end{aligned} \tag{5}$$

Since the number of pixels describing the marking varies with the viewing distance, the matched filter is adapted to the correct signal length for each image line to compute the maximum SNR. The information about position and width of the marking is provided by the recognition system that is used to process the data.

The SNR provides the needed measurability of the impact of an imaging component on the quality of input data similar to the MTF measurement and is an indicator for the detection probability of a given signal. The higher the SNR at the receiver input, the higher the detection probability of the signal and eventually the performance of the perception algorithm.

IV. Experimental results

This section details the results of the conducted experiments and tests. First, the results of the sensor performance measurement described in section II for several lens and imaging sensor combinations are presented. Thereafter the SNR measurement and perception results on road scenes in front of a moving truck for the same set of combinations is given.

The Slant Edge Feature measurement provides a fast and robust MTF estimation based on gradient edge evaluation of a fine print ISO12233 test target [1]. The process itself is robust with respect to small alignment variations and different contrast ratios, as long as the camera's exposure limit is not reached [8]. To achieve a stable illumination level for all measurements and to prevent saturation, the histogram of the scene was used as an input to control the camera. Image enhancement features of the sensors were deactivated to provide unbiased results. These would produce distorted MTF curves. No compression or other conversions were applied to the data. The capturing requirements for the measuring target were maintained using standard components for lighting and fixation and assured similar conditions for repeated measurements.

sensor	width	height	pixel size	$f_{Nyquist}$
1	$640px$	$480px$	$7.4\mu m^2$	$68lp/mm$
2	$1024px$	$768px$	$4.6\mu m^2$	$108lp/mm$

TABLE I

Parameters of the used grayscale imaging devices

Tables I and II show the tested imaging sensors and lenses. The grayscale CCD sensors have different resolution and pixel sizes. The lenses vary mainly in focal length. To test the variance and accuracy of the process, two lenses of the same model (1 and 3) were tested as well. Each camera was

lens	focal length	type
1	$8mm$	normal
2	$12.1mm$	tele
3	$8mm$	normal
4	$13mm$	tele
5	$12.6mm$	tele
6	$4mm$	wide

TABLE II

PARAMETERS OF THE USED LENSES

equipped with different lenses by turn and the MTF of the imaging system was measured. The result is a measure for the effects of different lenses on the imaging sensors and a rating for the overall performance of one combination.

Fig. 1 and Fig. 2 show the normalized MTF for each lens over the corresponding spatial frequency measured in cycles per pixel, with respect to $F_{Nyquist}$. A good way to interpret the performance with the MTF is the frequency where the contrast has only half its original value. For all investigated systems, lens 4 poses the best performance. Lens no. 6, a wide angle lens with short focal length, shows the lowest performance, with even dropping under a critical 0.1 MTF before reaching half of the sampling rate ($F_{Nyquist}$) and thus limiting the sensors performance. This lens degrades the performance of the imaging system noticeably. Also sensor 2 shows better performance because of its superior resolution which leads to a Nyquist frequency nearly twice as high as sensor 1.

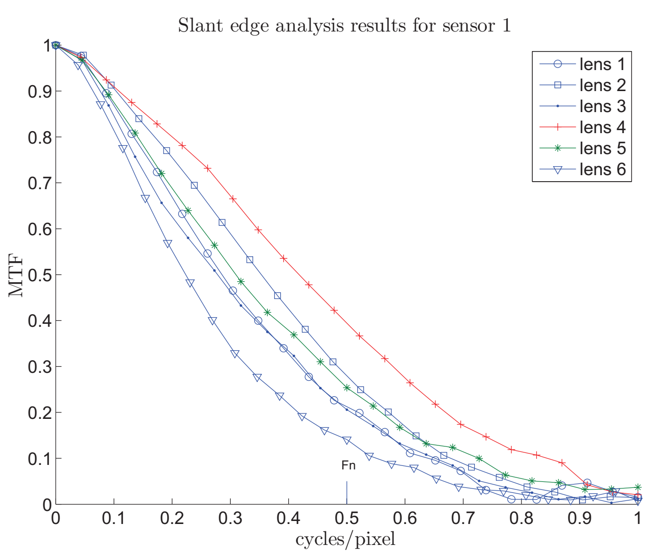

Fig. 1. Normalized MTF for sensor 1 over cycles per pixel (F_n = Nyquist Frequency). Lens 4 shows the best performance and lens 6 the lowest. The structurally identical lenses 1 and 3 show the expected equal performance.

A comparison of the two resulting graphs (Fig. 1 and Fig. 2) shows that the lower resolution of sensor 1 leads to less specifiable graphs and an overall lower performance. Lenses 1 and 3 show comparable MTF performance on both sensors with a visible divergence for sensor 2, which is related to the doubled Nyquist frequency resulting in a higher sensitivity for measurement errors. This leads to the assumption, that small differences between lenses can not be resolved sufficiently with this simple measurement setup. However, the measurement bears a high potential for fast and easy performance estimation and a first impression of how the overall system affects contrast and resolution which directly influences the performance of a given recognition task. Being able to distinguish the performance of the lenses together with the different optical characteristics from the perception results, is the task of the perception evaluation.

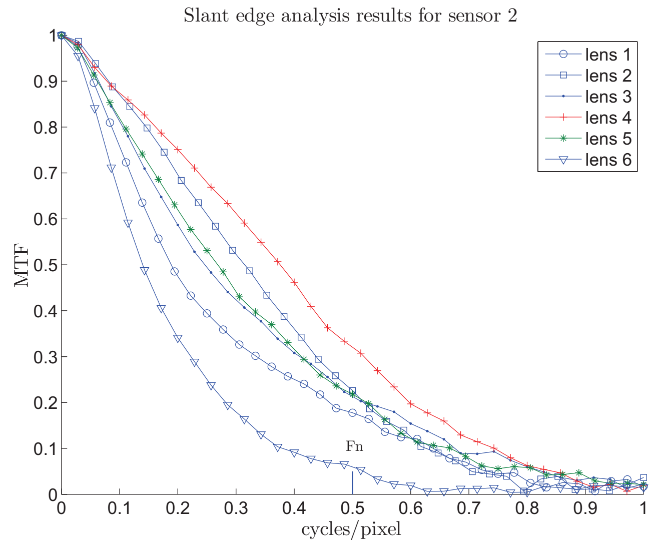

Fig. 2. Normalized MTF for sensor 2 over cycles per pixel (F_n = Nyquist Frequency). Lens 4 shows the best performance, while the low performance of lens 6 even limits the sensors performance.

For the experimental work on the recognition task the influence of the different imaging combinations is evaluated on the scenery in front of a moving truck. The acquired images of freeway roads are processed offline to extract the position and specificity of road lanes. The camera and lens combinations were placed behind the windshield at the center of the vehicle. It is important to align the field of view, to have the effects of different imaging systems pictured in comparable data and not to introduce more scenery than needed. Images of two different combinations were recorded in parallel for offline processing, allowing for a direct frame-by-frame comparison (Fig. 3).

Fig. 3. Scene from real world experimental setup (left: sensor 1, right: sensor 2)

Data of the combinations and the output of an additional built-in lane recognition system were recorded for about $100km$ of freeway driving. The built-in system provides reference data for the final system evaluation. The SNR measurement is performed offline and can be tuned to match the focus of the evaluation. Applying the SNR measurement

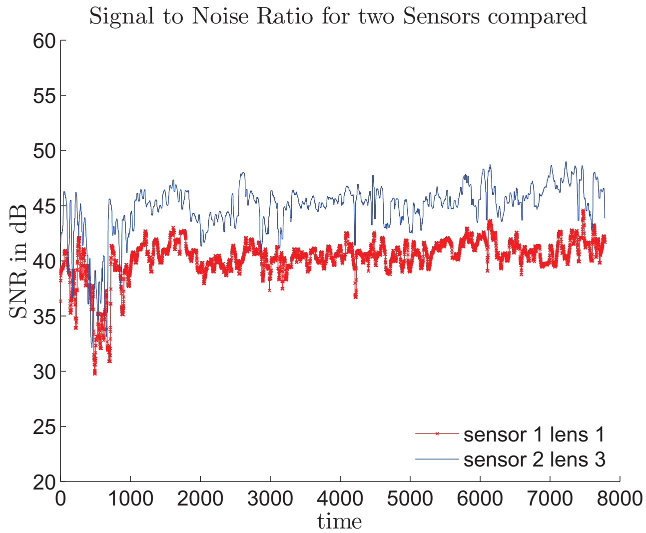

Fig. 4. SNR results of the right lane marking for sensor 1 and sensor 2 equipped with the same lens model for a long sequence of one measuring campaign (ca. $9min$ at 15 frames per second). Sensor 1 is permanently below the performance of sensor 2.

to a mid range measurement window for both sensors equipped with the same lens model produces results that show a clear separation of the two sensors with sensor 2 being permanently superior (Fig. 4). The high variation at the beginning is due to very poor quality of the marking itself resulting in unbalanced signal response. For normal quality, a stable average is maintained with a visible correlation for performance drops, while passing dashed markings of exit lanes. This shows additional use cases for the measurement in terms of describing the impact of external conditions on the quality of input data. Weather and lighting conditions as well as road constitution influence the SNR results and knowing the threshold for successful feature extraction an estimate for the performance in these situations should be given.

For a more detailed evaluation of the different combinations, the SNR was applied to $1m$ sized measurement windows up to $60m$ in front of the vehicle, which covers the region of interest for the recognition task and computed for all recorded scenes. Thereby the operating area for each setup and the performance inside can be determined at once. This provides information on what impact the sensors have on the signal quality in different distances and thus a measure where the probability for successful feature extraction is limited by the sensor.

The results presented in Fig. 5 and Fig. 6 confirm that sensor 2 is superior for almost all distances regardless of the used lens. The performance for operating distances above

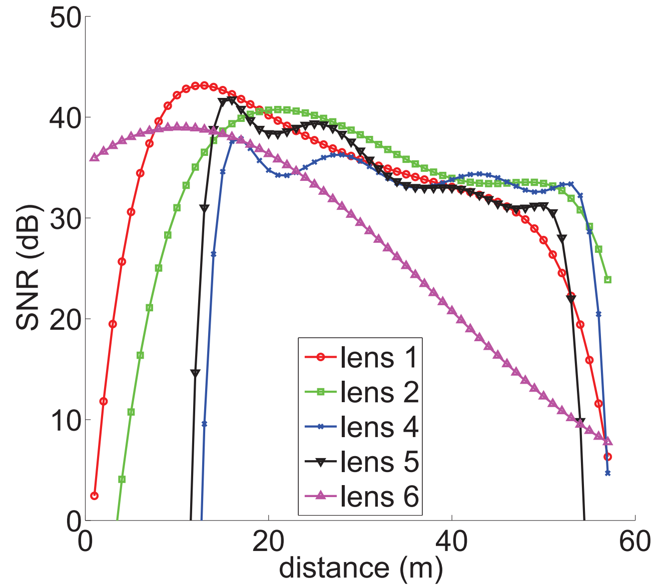

Fig. 5. SNR results of sensor 1 for $1m$ segments up to $60m$ for different lenses. Lens 6 and 1 perform well in the near distance and lens 1 is able to compete with the other lenses above $35m$.

$40m$ for the lenses are increased too.

The differences between the lenses for the same sensor reproduce the performance differences as indicated by the MTF results. Also, lenses that are not well separated in the MTF graph, can not be separated well in these measurements either, which was expected. The variation of environment parameters for the recorded scenes assumedly has a higher impact than the variance of the lenses. Lens 6 is inferior to all other lenses except for its advantage in the near distance due to the wide angle view but lens 1 is able to cover this area too and can provide better SNR in the distance. Lens 4 shows a high and stable performance but due to its zoomed field of view has a small operating area.

The impact of the different sensor performances on the perception result of the lane recognition component is investigated by comparing the results of the tested sensor with the results of the recognition system recorded together with the images during the experiments. The significant measures that show a reliable estimation are the curvature and the width of the road. If the number of extracted features for the far end of the operating distance is not sufficient, the road estimation will not correctly follow upcoming curves and the width estimation produces results for the wrong model parameters.

For lens 1, both sensors can provide sufficient performance for the perception task and no significant difference between the sensors is visible (Fig. 7. This changes when the low performing lens 6 is used. Compared to lens 5 during one of the recorded road sequences with changing turns, the curvature estimation is constantly off target (Fig. 8). The same behavior is true for the setup with switched lenses on a similar road scene. Whereas sensor 1 does not perform as good as sensor 2 with lens 5 equipped, sensor 2 is not able to get better performance out of lens 6 than sensor 1 (Fig. 9).

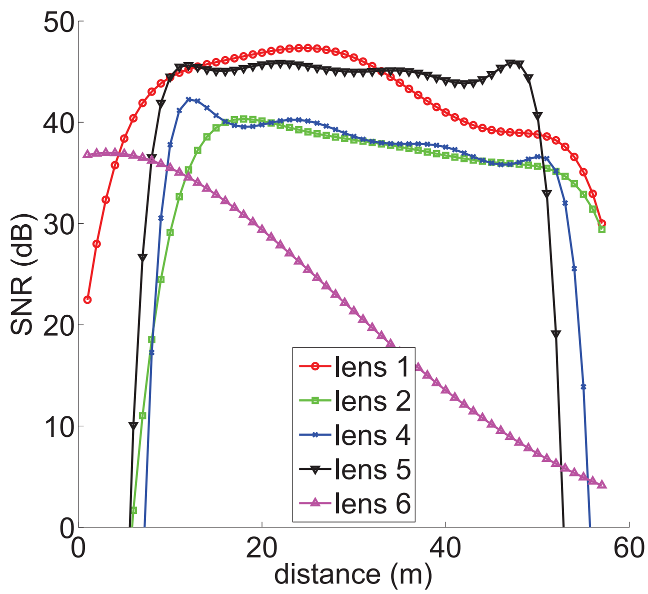

Fig. 6. SNR results of sensor 2 for $1m$ segments up to $60m$ for different lenses. All but lens 6 show better performance than on sensor 1 and good performance even above $40m$.

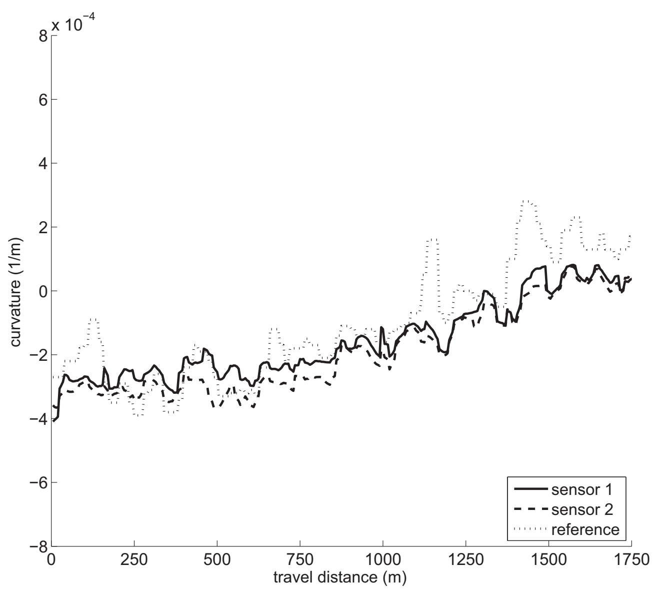

Fig. 7. Lane curvature estimation on sensor 1 and sensor 2 with evenly good SNR performance lenses and the result of the reference system.

For both sequences it is also visible that lens 6 introduces more noise for the perception estimates which complicates the task for possible controlling components that would use this as input.

Analyzing the cumulated width estimation results for the computed road sequences gives an insight on how well the tested devices could perform for tasks where the distance of the vehicle to the road lanes has to be computed.

It is again with lens 6, that both sensors show the expected low performance depicted in Fig. 10 and Fig. 11 with an absolute deviation $\delta = |Width_{ref} - Width_{est}|$ from the reference data of up to $0.4m$, which is not acceptable if this

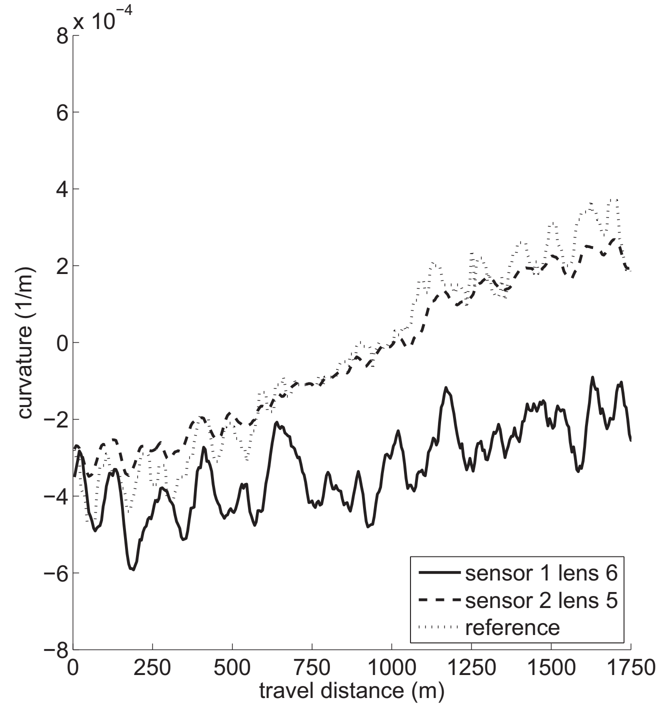

Fig. 8. Lane curvature estimation on sensor 1 and sensor 2 with a poor and a good SNR performance lens and the result of the reference system.

was an input for critical controlling components. Sensor 1 also shows a higher variance in the data which would most likely worsen the results if more data was computed and more complex scenes would be investigated.

V. Conclusion

The investigated perception results for the different sensor setups show that in each performed measuring step, from the first performance evaluation with standardized image quality measures to the applied signal to noise ratio, measurements for significant performance differences could be provided and matched the final overall performance for the targeted perception task.

This method provides an insight on how the quality directly influences the perception performance and that it is possible to give a relative measure for its impact on the perception performance. With lower data quality for far distances, indicated by the lower signal to noise ratio, the probability and amount of successful detected lane features decreases. The connected perceptive components that analyze these features have to cope with the lower detection rate if the system performance may not drop down, too, which is shown in our results.

The SNR provides a relative performance measure for the significance of an explicit feature for captured scenes and conditions. Influences on the feature quality, lightning and changing weather conditions, distort the results respectively. From overcast to rainy or snowy weather a comparably drop in performance could be observed for all investigated setups. By computing the SNR for different segments of the data it is possible to evaluate the performance differences dependent

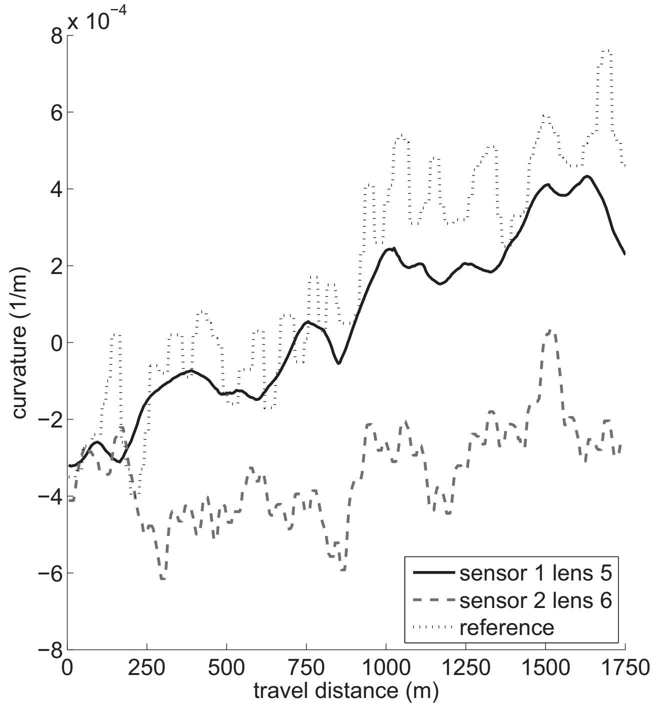

Fig. 9. Lane curvature estimation on sensor 1 and sensor 2 with a poor and a good SNR performance lens and the result of the reference system.

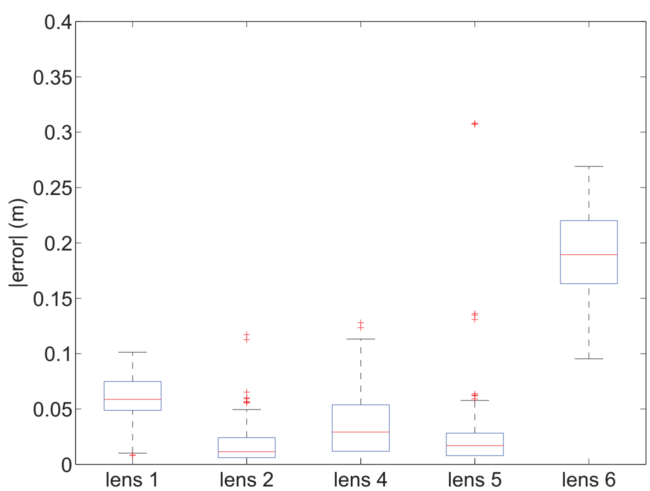

Fig. 10. Overall lane width absolute estimation error δ on sensor 1 versus the reference data.

on parameters like viewing distance and further distinguish the different components. The detection rate of the investigated perception task correlates with the performance of the applied SNR measurement and makes it possible to link the sensor performance to the perception performance.

Especially when the decision for sensor components has to be made with costs in mind the need for a well selected range of performance is apparent. Extending these measurements with additional parameters of interest like costs and application specific variables will help developers to find a balanced operating point.

Image quality measurement using the proposed standards

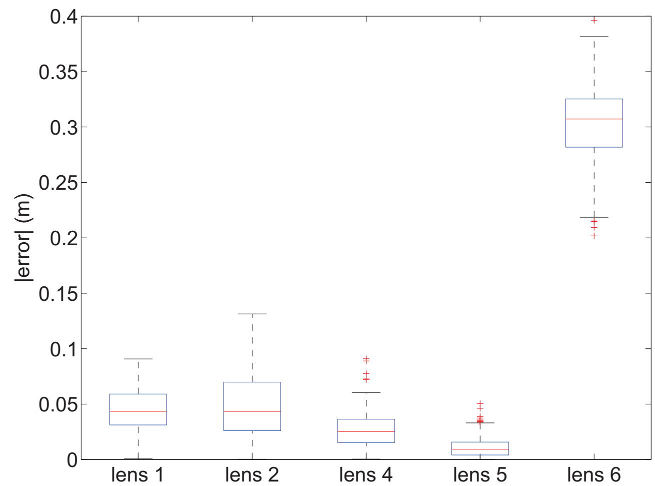

Fig. 11. Overall lane width absolute estimation error δ for sensor 2 versus the reference data.

can be applied to evaluate the performance of imaging systems with their resulting MTF, without the need of cost intensive measurement processes and tools. We assessed the measured performance by processing the data of different imaging setups for the task of lane recognition. The presented measurement of the SNR of an optimal feature signal, is able to rate the performance of different imaging systems when compared to each other and resolves similar performance relations to the MTF analysis of the imaging systems.

To get precise and reliable results it is advised to compare systems that are recording in parallel, or use a large quantity of data from each imaging component, that covers enough scenarios to compute a fair average result.

References

[1] International Standards Organization, "ISO12233 Photography - Electronic still-picture cameras - Resolution measurements," 2000.

[2] European Machine Vision Association, "EMVA Standard 1288: Standard for Measurement and Presentation of Specifications for Machine Vision Sensors and Cameras," 2007.

[3] D. Williams and P. Burns, "Low-frequency MTF estimation for digital imaging devices using slanted-edge analysis," *Proceedings of SPIE*, vol. 5294, no. 1, pp. 93–101, 2003.

[4] A. Angelis, A. Moschitta, F. Russo, and P. Carbone, "Image Quality Assessment: an Overview and some Metrological Considerations," in *IEEE International Workshop on Advanced Methods for Uncertainty Estimation in Measurement*, ser. IEEE International Workshop on Advanced Methods for Uncertainty Estimation in Measurement, 2007, pp. 47–52.

[5] D. W. Hertel and E. Chang, "Image Quality Standards in Automotive Vision Applications," in *Proceedings of the IEEE Intelligent Vehicles Symposium*, ser. Intelligent Vehicles Symposium, IEEE, 2007, pp. 404–409.

[6] H. Kunsoo, P. Jaehak, H. Daegun, D. Cho, and P. Jahng Hyon, "Vision-based lane detection for passenger cars: configuration aspects," in *Proceedings of the 2004 American Control Conference, ACC: June 30 - July 2, 2004, Boston Sheraton Hotel, Boston, Massachusetts*, vol. 1. IEEE Service Center, 2004, pp. 792–797.

[7] J. W. Lee, "A Machine Vision System for Lane-Departure Warning," *Computer Vision and Image Understanding*, vol. 86, no. 1, pp. 52–78, 2002.

[8] M. Estribeau and P. Magnan, "Fast MTF measurement of CMOS imagers using ISO 12333 slanted-edge methodology," *Proceedings of SPIE*, vol. 5251, no. 1, pp. 243–252, 2004.