Repeatability and Reproducibility Considerations for BlackMURA Measurements

Ingo Rotscholl, Tobias Porsch, Udo Krueger

TechnoTeam Bildverarbeitung GmbH, Werner-von-Siemens-Str. 5, 98693 Ilmenau, Germany

Keywords: BlackMURA; Measurement Uncertainty; Reproducibility; Imaging Luminance Measurement Device;

ABSTRACT

The "Uniformity measurement standard for Displays", which is used for automotive applications, describes precise setup and alignment procedures to ensure reproducible measurement results. However, the influences of the tested device and the ILMD are not considered in detail. This contribution shows experiments and simulations to estimate these influences as well.

1 INTRODUCTION

The uniformity evaluation of high quality automotive displays is performed according to [1], which refers to the "Uniformity measurement standard for displays" [2] also known as BlackMURA. This standard describes precise setup and alignment procedures to ensure reproducibility. However, the influences of the device under test and the Imaging Luminance Measurement Device (ILMD) are not considered. After summarizing important terms according to the international vocabulary of metrology [3] and specific ILMD metrics, this work analyses specific influences of typical ILMD via GUM-compliant Monte Carlo simulations [4,5]. Finally the study introduces experimental results of repeatability and reproducibility experiments. This also includes a percentile based black uniformity approach, which was considered in a recent study [6].

2 Basics of Measurement Uncertainty

The most important metrology terms should be known by all whose responsibility is connected to either, performing measurements, conformity assessment or defining measurement procedures and standardization.

Repeatability is defined as "the measurement precision (e.g. standard deviation, variance or coefficient of variation.) under a set of repeatability conditions of measurement" [3]. This means that the measurement procedure, the operator, the measurement devices, setups, operating conditions, location and the device under test remain unchanged. Also, the period between repeated measurements remains short. In practice this means, that the evaluation button of the measurement device is pressed several times and the resulting values are evaluated according to the definition of measurement precision.

Reproducibility is in contrast to repeatability defined as "the measurement precision under a set of reproducibility conditions of measurement [3]. Here the location, operator, devices, setups, and measurement devices may change. Only the device under tests remains the same. The period between measurements can be large as well. Thus, the data acquisition concept is different compared to repeatability. Basically, the setup is done several times by different people and/or devices according to the same or similar procedures and the resulting values are evaluated according to the definition of measurement precision.

The most advanced and complex concept to characterize a measurement is providing the **measurement uncertainty** for a measurement value. It is defined as "a non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand based on the information used" [3]. This concept includes the complete influence of all involved measurement devices, such as luminance camera, power supply, position stages, environmental influences and the properties of the device under test (DUT) and thus also influences from the repeatability, reproducibility and contributions from the traceability chain. Note that it is not possible to state a measurement uncertainty for a measurement device alone. For that purpose there are several index values as described in the next section

3 Basics of ILMDs

There are different sources of imperfection within the ILMD. Currently the CIE TC2-59 and the German DIN FNL3 (future DIN5032-10) aim to provide a document to characterize imaging luminance measurement devices.

In this work only selected attributes will be discussed, namely the spectral mismatch index f_1', the non-uniformity index values f_{21} / f_{22} as well as the detection limit $f_{3,0}$ and the non-linearity index values $f_{3,1}$ / $f_{3,2}$.

The spectral mismatch index f_1' describes how well the spectral sensitivity of the overall system matches to the spectral luminous efficiency function for photopic vision V(λ). Thus, it mainly characterizes the overall quality of the matching filter glass. However, it is hard to estimate the spectral mismatch correction needed for a specific measurement only by the f_1' value because the spectral mismatch correction highly depends on the spectral power distribution of the DUT.

The detection limit $f_{3,0}$ is the least signal level, which can be distinguished from the background noise with a certain probability. In combination with the exposure time and a calibration factor, it determines the least measureable luminance. The detection limit depends on the dark signal and its extended standard deviation (all corrections are applied) and thus on the sensor array.

The non-linearity index values $f_{3,1}$ and $f_{3,2}$ describe the deviation from the ideal linearity between input signal level and the resulting signal after applying the non-linearity correction. Thus, it can be interpreted as the difference introduced by a varying degree of saturation *S* across the individual pixels of each image. The only difference between $f_{3,1}$ and $f_{3,2}$ is the measurement performed in order to achieve the non-linearity index. In the case of $f_{3,1}$, the luminance of the measurement standard source is adjusted to change the degree of saturation, while the exposure time remains constant. In the case of $f_{3,2}$, it is the other way around. The sensor array and its amplification circuit are the main reason for the non-linearity. Figure 1 (top) visualizes an example for the non-linearity. The left y-axis shows the generated electrical signal as a function of the degree of saturation. While the red line shows the ideal case, the green solid line provides the measured signal. The dotted line shows the absolute deviation between the real and ideal case.

The non-uniformity index values f_{21} and f_{22} describe the largest deviation of a measured result as a function of the spatial position on the sensor array. The non-uniformity can be caused by different effects. There can be non-uniformities due to the imaging lens, due to the V(λ) matching, the sensor array sensitivity and the temperature distribution on the sensor array. The index f_{21} describes the maximum deviation of a specific spot of a large uniform source across the camera field of view.

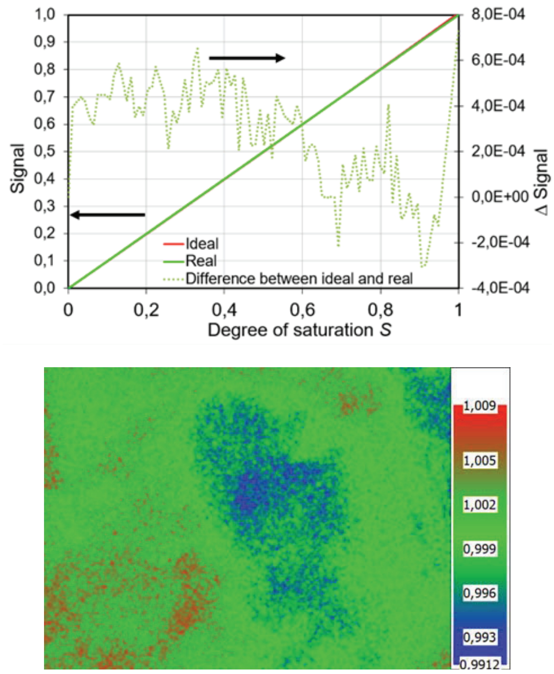

Figure 1: Results of a non-linearity measuremen (Top) and Non-uniformity with a $f_{21}$of 1.8% (Bottom)

The index f_{22} describes the maximum deviation of a small spot source across the camera field of view. They can be slightly different due to differences in stray light caused by the size of the source. The bottom figure shows an example non-uniformity f_{21} of a luminance camera.

4 BlackMURA measurements

Measurements according to the BlackMURA standard do not only imply a specific evaluation algorithm but also a well-defined setup procedure. This procedure consists of three steps: the geometrical alignment, the measurement field angle setup and a reproducible defocusing technique to avoid the Moiré phenomenon. It includes for instance parameters such as angular deviations between the optical axis of the camera and the display normal, spatial deviations of the optical axis and the display center or the measurement distance and focus settings with respect to the focal length and procedures how to achieve them. This is important to ensure the reproducibility of the measurement results

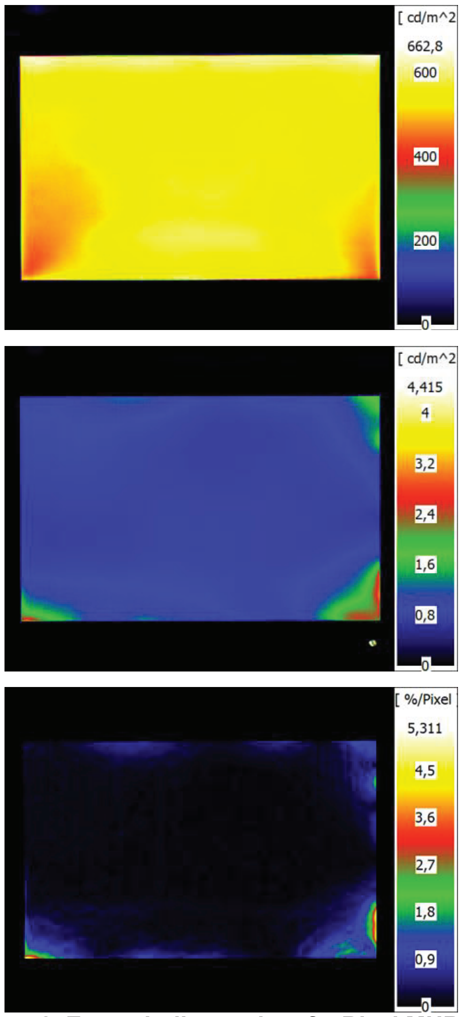

Figure 2. Exemplarily results of a BlackMURA measurement: White and Black luminance image (Top and Middle) as well as gradient image (Bottom)

After completing the setup, two luminance images are captured, which are the “Dark image” and the “Bright image” as exemplarily shown in Figure 2.

Then, the active display area is detected using a threshold operation. Based on the pixel pitch and the magnification, a filter size parameter is calculated. This parameter is based on the maximum of the contrast sensitivity function from [7] applied to a typical viewing distance for an automotive display. For modern automotive displays, it will be approximately 2.24 mm [1].

These 2.24 mm are used to decrease (erode) the detection area. After that a box filter with this filter size is applied to further reduce the noise and thus to enhance the repeatability. Then the luminance values and the non-uniformity, which is defined as the ratio of the minimum and maximum luminance of an image, are extracted.

Recent studies [6] also considered to use a percentile based evaluation instead of the min/max based ratio to further enhance the reproducibility. The physio-psychological study performed in [6] showed that the correlation to the perceived uniformity is slightly increased by using a percentile based approach.

Further, a gradient image of the dark luminance image as shown in Figure 2 is calculated. The mathematical details of this calculation are provided in [2]. The maximum from this gradient image is used to identify local non-uniformities.

5 Repeatability and Reproducibility Considerations for BlackMURA Measurements

The following section describes the results on simulations of ILMD influences as well as experiments that were performed to estimate the influence of the setup procedure. The simulations were performed according to [8,9] and have been performed again to include the percentile based uniformity evaluations.

5.1 Simulation of ILMD Influence

To estimate the influence of the camera attributes on the final measurement results, the example measurement images from Figure 2 are used. Both captured images are changed for each iteration by randomly generated uncertainty images. They are based on predefined values for $f_{2,1}$, $f_{3,0}$ and $f_{3,2}$ as well as the maximal degree of saturation S_{Max} within each image. S_{Max} is considered because it might be useful to decrease S_{Max} in order to reduce the cycle time of a measurement. To systematically differentiate the influence of different camera attributes, we define four different cases of luminance cameras and operating modes, which are shown in Table 1.

The first parameter set describes a typical calibrated ILMD with all corrections applied. The second camera system NU has a higher non-uniformity. The third camera system NL has a lower detection limit and generally a larger non-linearity but a good non-uniformity index value. The camera NL is simulated twice to consider two different degrees of saturation

Table 1 Parameters for Monte Carlo Simulation

Name	$f_{3,0}$	$f_{3,2}$	f_{21}	S_{Max}
Camera A	0.01%	1 %	2 %	90 %
Camera NU	0.01%	1 %	5 %	90 %
Camera NL OM 1	0.1%	5 %	2 %	90 %
Camera NL OM 2	0.1%	5 %	2 %	15 %

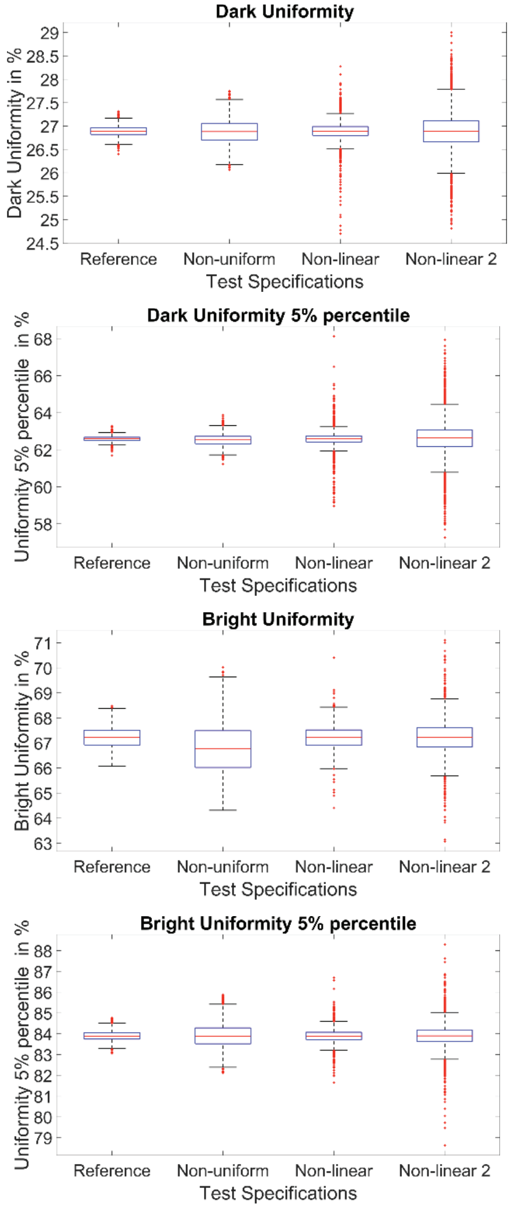

Figure 3. Simulation results of ILMD influence on black and white uniformity metrics

The simulation results of the white and black uniformity, as well as a 5% percentile uniformity are shown as box plots in Fig. 3.

It can be seen that the camera reference always performs best, while the camera NL with the low saturation always shows the worst performance. The influence of $f_{2,1}$ is higher for higher uniformity values [8] but reduced by applying the percentile based approach. The same is true for the camera reference.

5.2 Experiment on noise influence (repeatability)

To test the repeatability, we performed the setup procedure once and waited until the DUT (a randomly chosen LCD display) reached the steady state condition. After that we performed the BlackMURA measurements 30 times and calculated the coefficient of variation of all results in percent. The results are shown in Table 2. It can be seen that the CV is small and thus that the repeatability is high. Further, the percentile based evaluation does not affect the repeatability significantly. Thus, it can be concluded that the evaluation procedure of BlackMURA supports the requirement of repeatable measurements.

5.3 Experiment on setup influence (reproducibility)

To test the reproducibility, we performed the setup procedure 15 times and on several different days. For each setup a mean value of at least 10 BlackMURA evaluations was derived and used as representative value of the setup. Furthermore, the distances were adjusted such that the box filters in camera pixels did either change slightly or remained constant at different reproduction scales. After that, we calculated the coefficient of variation in percent. Again, the steady state condition was ensured.

The results are shown in Table 3. It can be seen that the reproducibility is high as well but lower than the repeatability, which is to be expected. It can also be seen that the percentile based evaluation increases the reproducibility from 2.7% up to 1.3%. It should be noted here that although we tried to minimize influences from the test display itself, they cannot be excluded. However, it would be best practice to replace the display by a more stable uniform and non-uniform light source.

6 Summary, conclusions and outlook

Repeatability and reproducibility are important conditions for the practical application of a measurement method. The current version of the BlackMURA standard considers these concepts by its evaluation algorithms and specific conditions regarding the ILMD specifications and the setup procedure. These concept work for most metrics considered in BlackMURA. However, especially the black uniformity metric reproducibility might still be improved.

Table 2: Repeatability results

Parameter	Image	Unit	Value	CV in %
Uniformity	Dark image	%	27.4	0.07
Uniform. 1%	Dark image	%	39.3	0.07
Uniform. 2%	Dark image	%	47.8	0.06
Uniform. 5%	Dark image	%	62.9	0.06
Maximum W	Gradient image	%/px	0.008	-
Maximum B	Gradient image	%/px	4.951	0.2
Uniformity	Bright image	%	68	0.09

Table 3: Reproducibility results

Parameter	Image	Unit	Value	CV in %
Uniformity	Dark image	%	27.4	2.7
Uniform. 1%	Dark image	%	39.3	2.37
Uniform. 2%	Dark image	%	47.8	1.67
Uniform. 5%	Dark image	%	62.9	1.33
Maximum W	Gradient image	%/px	0.007	-
Maximum B	Gradient image	%/px	4.708	4.5
Uniformity	Bright image	%	68	0.22

We have shown by GUM-complied Monte Carlo simulations and experiments that especially this reproducibility value can be increased by using a percentile based black uniformity evaluation instead of the current extrema based evaluations. A recent study has shown that the correlation to human perception does not change significantly and is even slightly improved by this change Also it should be noted that this concept implicitly ignores the “worst” pixels during a uniformity evaluation. However, all small and strong local non-uniformities are still considered by the gradient metric, which is also part of the standard and the specifications.

In a future work the experiments regarding the reproducibility shall be repeated with a more stable light source as for instance the U4 normal developed by TechnoTeam. That way the influences from the DUT can be minimized and the effects can be assigned to the measurement equipment and the setup procedure.

Also round-robin like test scenarios are possible with these U4 normals. Then these experiments offer a more practical reproducibility test scenario of the critical uniformity values considering a more complete “set of reproducibility conditions of measurement”.

REFERENCES

[1] German Automotive OEM Work Group Displays, Display Specification for Automotive Application V5.1; 2018

[2] German Automotive OEM Work Group Displays, Uniformity Measurement Standard for Displays V1.30; 2018

[3] Bureau International des Poids et Mesures (BIPM), International Vocabulary of Metrology, BIPM, 2012

[4] Bureau International des Poids et Mesures (BIPM), GUM: Guide to the Expression of Uncertainty in Measurement;" BIPM, 2008.

[5] Bureau International des Poids et Mesures (BIPM), Evaluation of measurement data – Supplement 1– Propagation of distributions using a Monte Carlo method," BIPM, 2011.

[6] H.-J. Kwon, J. G. Jo, J. H. Ham, T. Y. Park, D. G. Lee, and B. Y Lee, “Comparison of Black Uniformity Metrics for Automotive LCDs”, Proc. IMID, 2019

[7] J.-L. Mannos and D. Sakrison, "The Effects of a Visual Fidelity Criterion on the Encoding of Images," IEEE Transactions on Information Theory, vol. 20, no. 4, pp. 525-535, 1974.

[8] I. Rotscholl and U. Krueger; „Requirements on the characterization of optical display attributes for automotive applications; Proc. of SID Vehicle Display; 2018.

[9] P. Kärhä, A. Vaskuri, H. Mäntynen, N. Mikkonen and E. Ikonen, "Method for estimating effects of unknown correlations in spectral irradiance data on uncertainties of spectrally integrated colorimetric quantities," Metrologia, vol. 54, 2017.