Conformity assessment with examples for BlackMURA measurements

Bob Liu[1], Ingo Rotscholl[2] and Udo Krueger[2]

1Light-All Co. Ltd, No. 1333 Xinlong Rd.,201101 Minhang District; Shanghai, China

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

Abstract

To decide whether a display works according to the specification it is not only necessary to take the measurement values into account. The measurement uncertainty also plays a significant role in this decision. Besides some basic knowledge about the measurement uncertainty evaluation, the paper shows the main contributions and their influences for the key values of the BlackMURA standard. Using the rules of conformity assessment, the use of tolerance intervals and acceptance intervals is explained in detail to give all parties (OEMs, manufacturer of display and measurement system manufacturer) the chance to discuss their quality metrics in a reasonable manner.

Keywords: *BlackMURA; Photometry; Measurement Uncertainty; Conformity Assessment; Imaging Luminance Measurement Device; ILMD*

1. Introduction

The quality of displays is rated based on measurement results. Besides the specifications for the measurement (e.g. the well-known BlackMURA standard [1]) the OEM specifications [2] limits quantities like luminance, uniformity or the maximum gradient.

Usually, the displays under tests are measured according to a specific procedure, and a value is reported. However, if the conformity assessment, which is the pass/fail decision of a sample display, does not take into consideration a reasonable estimation of the measurement uncertainty the probability of either false acceptances or false rejections may increase significantly. In practice, the tolerance intervals are usually very strict, which on the one hand ensures the display quality, but on the other hand, increases the costs on both sides of the value chain.

This paper addresses this issue. First, we give a basic introduction on measurement uncertainty and conformity assessment including relevant and partly free available references. After that, we focus on the application of selected concepts on BlackMURA evaluations. This includes, on the one hand, simulations of ILMD influences and on the other hand experiments on the repeatability as well as reproducibility of a randomly chosen IPS display.

2. Basics of Measurement Uncertainty and Conformity Assessment

The terminology of metrology can be confusing because of several related terms, which have only slight but important differences. Therefore, the most important terms should be known by all whose responsibility is connected to either, performing measurements, conformity assessment, or defining specifications and tolerances.

2.1 International vocabulary of metrology [3]: Before introducing selected terms, it shall be noted that the BIPM (Bureau International des Poids et Mesures) published the "International Vocabulary of metrology" (VIM) [3], which can be downloaded on their web page free of charge.

Accuracy is the first term that shall be mentioned because it is usually used incorrectly. It is *"the closeness of agreement between a measured quantity values and a true quantity value of a measurand" [3]*. However, note that the concept of accuracy is not a quantity and thus no numerical value can be assessed. Thus, in metrology, it is not possible to state that the accuracy of a measurement or a device is better than a certain value. In the case of any numerical comparison other terms, such as precision needs to be used

Measurement precision is the first applicable term that shall be explained. It is *"the closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions"* [3]. It is usually expressed numerically by measures of imprecision such as standard deviation, variance or coefficient of variation.

Repeatability is defined as "*the measurement precision 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.

The BIPM published guides, which explain how to calculate the measurement uncertainty in general. These are "The Guide to the expression of Uncertainty in Measurement" [4] and its supplements [5,6]. For photometry special guidance is given in [7].

2.2 The role of measurement uncertainty in conformity assessment: The last term that shall be introduced is conformity assessment, which is defined as *"the activity to determine whether specified requirements relating to a product, process, system, person or body are fulfilled"* [6].

This process includes a valid and reasonable estimation of acceptance intervals based on predefined tolerance intervals under consideration of the measurement uncertainty or a good estimation of the uncertainty. The concept of tolerance intervals is shown in Figure 1. The upper sketch shows a lower limit T_{L}. All measured quantities that have a value larger than T_{L} are within the tolerance interval and will pass the test. An example for a lower limit is the Black Uniformity, which has to be larger than 50% according to [2]. The sketch at the bottom shows an upper limit T_{U}, where the tolerance interval is below the threshold value. An example would be the BlackMURA gradient value according to [2].

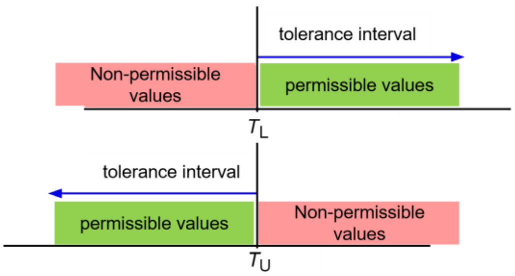


Figure 1: Single sided tolerance intervals

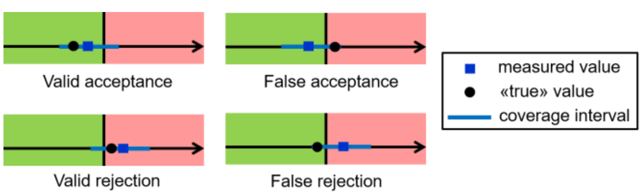


Figure 2: Qualitative influence of measurement uncertainty on conformity assessment. The desired outcomes are valid acceptance and valid rejection. The undesired outcomes are false acceptance (consumer risk) and false rejection (producer risk)

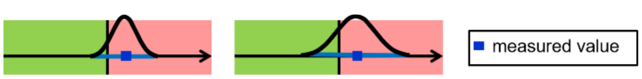


Figure 3: Quantitative influence of measurement uncertainty on conformity assessment.

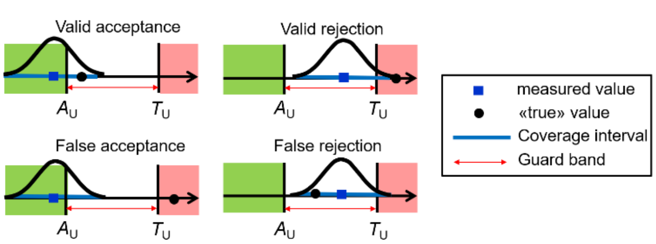


Figure 4: Influence of acceptance interval on conformity assessment: The probability of false acceptance is strongly reduced while the false rejection probability is strongly enhanced

However, due to the always existing measurement uncertainty, the true value of a quantity and the measurement value might differ. Especially near T_{L} or T_{U} this may result in the four different assessment cases "valid acceptance", "false acceptance", "valid rejection" and "false rejection" as shown in Figure 2. Valid acceptance means that the measured value and the true value are within the tolerance interval. Similar, valid rejection, means that both values are outside the tolerance interval. These two valid assessments are the desired outcome of a conformity assessment. In contrast, the false acceptance and the false rejection describe cases in which either only the true value (false rejection) or the measured value (false acceptance) are within the tolerance interval.

A small measurement uncertainty reduces the probability of a false assessment because the probability of occurrence of a false conformity assessment depends on the integral of the uncertainty distribution within the non-permissible values. This is visualized in Figure 3. It shows a rejection event on both the right-hand side and the left-hand side. However, in the case of the right-hand side, the probability that the true measurement value is within the tolerance interval is much larger, which means that the probability of a false assessment is larger as well. On the left-hand side, it is quite unlikely that the true measurement value is still within the tolerance interval.

If the tolerance itself serves as an acceptance threshold, the consumer and producer risk are equally shared. This is also called the shared risk approach. However, in practical applications, the customer wants to reduce the probability of false acceptance. This can be done by defining a separate acceptance interval A_{U} or A_{L}, which lies within the tolerance interval. The distance between the upper or lower limit to the acceptance interval is called the guard band.

Figure 4 visualizes the influence of the guard band on the outcome of the conformity assessment. While the probability of false acceptance is reduced strongly in this example, the probability of false rejection is enhanced. However, at some point, an increasing guard band will not influence the occurrence of false acceptance and only enhance the false rejections with respect to the tolerance interval. In this undesired case, only the consumer costs are enhanced without a real need. Only if there is a well-known and small uncertainty, the guard band can be chosen such that the consumer risk is minimized without unnecessarily enhancing the producer risk and thus the costs.

Thus, it is crucial to know or at least estimate the guard band based on all information, which can be included with respect to the scientific and economic circumstances. At least the reproducibility should be considered. Guidance to consider "The role of measurement uncertainty in conformity assessment" is published free of charge by BIPM as well [6]

3. Conformity assessment for BlackMURA

The "Uniformity measurement standard for displays v1.3" also known as "BlackMURA" describes a procedure to evaluate the luminance and luminance uniformity of displays. The described procedure exists to enhance both the repeatability and the reproducibility of measurement and thus also to reduce the uncertainty. Thus, the standard includes specific requirements for measurement devices, for the test setup and a well described and robust evaluation procedure.

3.1 The measurement device influence on BlackMURA evaluations is simulated in detail in [7]. It shows that there is a large contribution from the non-uniformity index for flat field f_{21}.

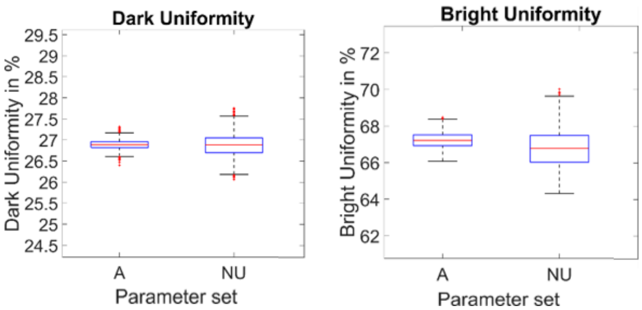


Figure 5: Influence of ILMD non-uniformity on BlackMURA uniformity uncertainty (Monte Carlo simulation result; excerpt from [7])

The BlackMURA standard takes this into consideration by limiting f_{21} below 2%. Further, the spectral mismatch $f_1{}'$ needs to be smaller than 5%. Figure 5 shows a simulated influence of f_{21} on the Black Uniformity of a specific measurement result for two different Imaging Luminance Measurement Devices in form of a box plot. The simulated ILMDs differ only in their uniformity index. The parameter Set A includes a f_{21}, which equals 2% and the parameter set NU equals a f_{21} of 5%. The simulation was carried out with a GUM compliment Monte Carlo simulation and thus also takes the measurement value itself into consideration. This can be seen by comparing the results for the bright and dark uniformity. While the qualitative influence is comparable for both values, the quantitative effects increase with increasing uniformity. Similar estimations are available for the influence of $f_1{}'$ on measurement results [8].

3.2 The setup influence: The main part of the BlackMURA standard is a detailed description of the measurement setup and procedures on how to achieve the setup. The setup description can be roughly divided into three parts, which are:

- The geometrical alignment
- The reproducible defocus
- The limitation of the measurement field angle

Figure 6 shows a BlackMURA complaint test image, which can be used to perform the geometrical alignment, as well as the reproduceable, defocus used to avoid Moiré in a well-defined manner. The limitation of the measurement field angle is a special test, which also takes the properties of the device under test (DUT) into consideration. It helps to distinguish angular and spatial uniformity by selecting an appropriate measurement distance and lens. All these requirements and restriction shall help to ensure a high reproducibility of the measurement results.

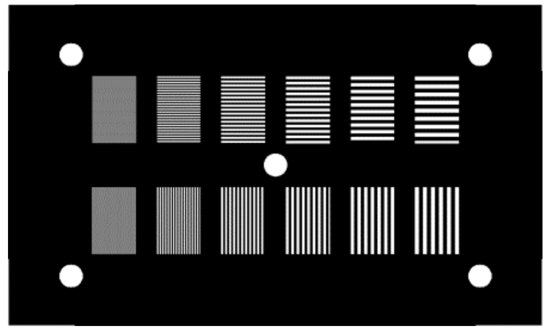
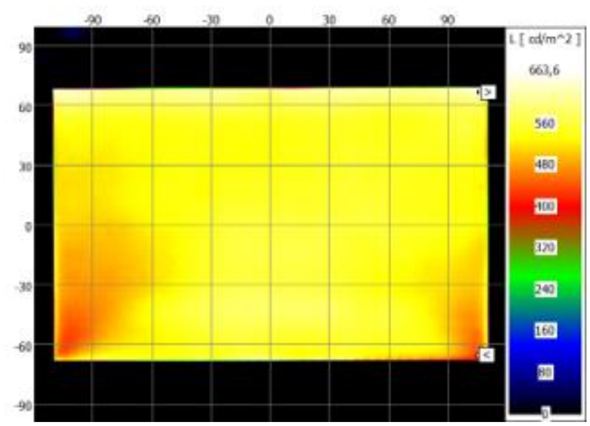


Figure 6: BlackMURA compliant setup image to ensure the geometrical alignment and a reproducible Moiré avoidance

3.2 The evaluation algorithm influence is reduced by clear and openly communicated robust procedures. This includes the chosen algorithms (box operations and gradient filters) to minimize the influence of noise as well as the determination of typical image processing parameters such as relative luminance thresholds to detect the active area, or parameters of erosion operations to consider boundary values. All these definitions shall help to ensure high repeatability of measurements.

4. Experimental results

In order to verify the capabilities of [1] to ensure both high repeatability as well as a high reproducibility, we estimated both values experimentally for a specific test display. In order to exclude the influence of the measurement device, we always used the same LMK5 Color. Figure 7 shows the three BlackMURA evaluation images exemplarily. These are the Dark image, the Bright image and the Gradient image.



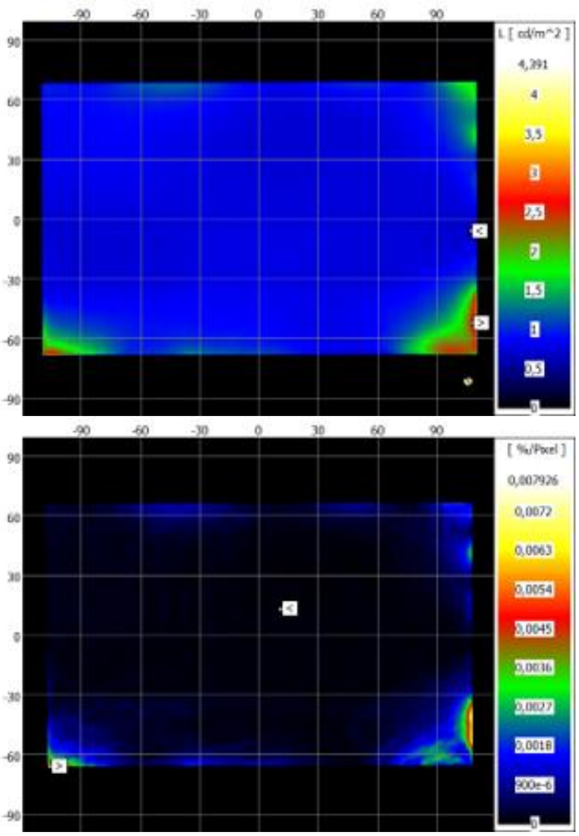


Figure 7: Exemplarily evaluation images

Table 1: Repeatability results

Parameter	Image	Unit	Value	CV in %
Mean	Dark image	cd/m²	0.87	0.04
Minimum	Dark image	cd/m²	0.71	0.05
Maximum	Dark image	cd/m²	2.59	0.04
Uniformity	Dark image	%	27.4	0.07
Maximum W	Gradient image	%/px	0.008	-
Maximum B	Gradient image	%/px	4.951	0.2
Mean	Bright image	cd/m²	543	0.02
Minimum	Bright image	cd/m²	432	0.1
Maximum	Bright image	cd/m²	637	0.05
Uniformity	Bright image	%	68	0.09

To test the repeatability, we performed the setup procedure once and waited until the DUT reached the steady state condition. After that, we performed the measurement 30 times and calculated the coefficient of variation (CV) in percent. The results are shown in Table 1. It can be seen that the repeatability for the directly measured luminance values is very high. The repeatability of derived quantities, which require several luminance values is slightly reduced but still very good. Thus, it can be concluded that the evaluation procedure of BlackMURA supports the requirement of repeatable measurements

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 a 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 (CV) in per cent. Again, the steady-state condition was ensured.

The results are shown in Table 2. It can be seen that the reproducibility for the directly measured luminance values is very high as well, but lower than the repeatability, as expected. Thus, it can be concluded that the setup procedure of BlackMURA ensures the capability of a high reproducibility across different instances of the value chain.

Table 2: Reproducibility results

Parameter	Image	Unit	Value	CV in %
Mean	Dark image	cd/m²	0.86	0.92
Minimum	Dark image	cd/m²	0.71	0.7
Maximum	Dark image	cd/m²	2.59	3.3
Uniformity	Dark image	%	27.4	2.7
Maximum W	Gradient image	%/px	0.007	-
Maximum B	Gradient image	%/px	4.708	4.5
Mean	Bright image	cd/m²	542	0.71
Minimum	Bright image	cd/m²	434	1.02
Maximum	Bright image	cd/m²	639	1.02
Uniformity	Bright image	%	68	0.22

Summary and Conclusions

This work summed up some basics of uncertainty and metrology as well as the concepts of conformity assessment based on the freely available publications of the BIPM. It suggests the incorporation of well-derived tolerance limits and estimations of the DUT-dependent measurement uncertainty (or at least reproducibility) to specify valid and cost-efficient acceptance limits. Furthermore, it summed up the procedures of the BlackMURA standard, which shall ensure valid results with high repeatability and reproducibility and showed supporting simulations and experiments.

However, note that all presented absolute values of both, the simulation, and the repeatability/reproducibility experiments are associated with a specific measurement value. The deviations always also depend on the characteristics of the DUT. This is also supported by a comparison of the achieved values for the "Bright" and "Dark" images.

Therefore, it can be concluded that the definition of an optimal guard band during conformity assessment requires investigations of a specific representative display sample. A initial, cheap and potentially sufficient way is to perform simple repeatability and reproducibility tests as shown in this paper. Based on these, the first estimations of the uncertainty can be achieved. Under consideration of these and openly communicated tolerance intervals, optimized guard bands, which ensure a low consumer risk at an optimized producer risk and thus at optimized costs, can be derived.

5. References

[1] German Automotive OEM Work Group Displays, *Uniformity Measurement Standard for Displays V1.30;* 2017.

[2] German Automotive OEM Work Group Displays, *Display Specification for Automotive Application V5.1;* 2017

[3] Bureau International des Poids et Mesures (BIPM), *International Vocabulary of Metrology – Basic and General Concepts and Associated Terms (VIM 3rd edition);* 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 to the "Guide to the expression of uncertainty in measurement" – Propagation of distributions using a Monte Carlo method*," BIPM, 2011.

[6] Bureau International des Poids et Mesures (BIPM), *Evaluation of measurement data – Supplement 2 to the "Guide to the expression of uncertainty in measurement" – Extension to any number of output quantities;* BIPM, 2011.

[7] Commission Internationale de l'Éclairage (CIE), *CIE 198:2011 Determination of measurement uncertainties in photometry*; CIE, 2011.

[8] Bureau International des Poids et Mesures (BIPM), *Evaluation of measurement data – The role of measurement uncertainty in conformity assessment*; BIPM, 2012.

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

[10] Commission Internationale de l'Éclairage (CIE), *CIE S 025/E:2015 Test Method for LED Lamps, LED Luminaires and LED Modules*; CIE, 2015