



Test Procedure – Whitepaper

VCX – Valued Camera eXperience

Version 2025

2025 07 06

VCX-Forum e.V. www.vcx-forum.org

CONTENT

1	Introduction.....	4
2	Measurement	6
2.1	Approach	6
2.2	Procedure	7
2.2.1	Camera Settings.....	7
2.2.2	Measurement Setup	8
2.2.3	Test Charts.....	8
2.2.4	Test Equipment	11
2.2.4.1	Illumination	11
2.2.4.2	Timing Kit	12
2.2.4.3	iQ-Near Focus device.....	13
2.2.4.4	STEVE.....	14
2.3	Measurement Conditions.....	15
2.3.1	Main Camera (Rear camera).....	15
2.3.2	Video (Rear camera)	16
2.3.3	Selfie Camera (Front camera).....	17
2.3.4	Performance measurements	17
2.4	Viewing Conditions (VC).....	17
3	Metrics.....	18
3.1	Resolution.....	18
3.1.1	Theoretical Pixel Count (TPC).....	18
3.1.2	Theoretical Maximum Resolution (TMR).....	18
3.1.3	s-SFR - Limiting Resolution (LR) Center / Corner	19
3.1.4	Effective Pixel Count (EPC) Center / Overall	21
3.1.5	s-SFR - Acutance Center / Corner	22
3.1.6	e-SFR	23
3.1.7	e-SFR – Acutance	24
3.2	Texture Loss (TL)	24
3.2.1	Texture Loss MTF10 / Limiting resolution	26
3.2.2	Texture Loss Acutance.....	26
3.3	Sharpening Artifacts	27
3.3.1	Overshoot / Undershoot	27
3.3.2	Maximum SFR.....	28
3.4	Tartan SFR.....	29

CONTENT

3.5	Colors	29
3.5.1	Color Reproduction Error.....	30
3.5.1.1	ΔE - Color Error (all)	30
1.1.1.1	ΔE - Color Error (red)	31
1.1.1.2	ΔE - Color Error (green)	31
1.1.1.3	ΔE - Color Error (blue).....	31
1.1.1.4	ΔE - Color Error (skin tones)	31
1.1.1.5	ΔL - Luminance Error (all)	31
1.1.1.6	ΔC - Chrominance Error (all).....	31
1.1.1.7	ΔH - Hue Error (all).....	31
3.5.2	White Balance (WB)	32
3.6	Visual Noise (VN)	32
3.6.1	Visual Noise (mean)	33
3.6.2	Visual Noise (max)	33
3.7	Dynamic Range (DR)	34
3.8	Shading	35
3.8.1	Intensity Shading	36
3.8.2	Color Shading.....	37
3.9	Distortion	38
3.9.1	TV Distortion.....	38
3.10	Video test.....	39
3.10.1	Video Quality	39
3.10.2	Video Performance.....	39
3.11	Extended low light test.....	39
3.11.1	Exposure	40
3.11.2	Resolution.....	40
3.11.3	Texture	40
3.11.4	Noise	40
3.11.5	Color	40
3.12	Response Time	40
3.12.1	Frame Rate.....	41
3.12.2	Shooting time lag.....	41
3.13	Motion Control	43
3.13.1	Visual Noise.....	44
3.13.2	Edge Width	44
3.13.3	Edge Acutance	45
3.13.4	Dead Leaves Acutance	45
4	the VCX Score in THE REPORT.....	46

CONTENT

4.1	The Benchmark	46
4.2	Report.....	47
5	Subjective Study.....	50
6	Appendix	53
6.1	Overview of Numerical Results	53
6.1.1	Overview of All Image Quality Values per Lighting Condition	53
6.1.2	Overview of all Performance/Response Values	54
6.2	Color Patch Groups	54
6.3	Changelog	55
6.3.1	Whitepaper	60
6.3.2	Metrics and Procedure	60
6.3.3	Score	60
6.3.4	Update history	55
6.3.4.1	Version 2025	55
6.3.4.2	Overview from Version 1.0(primary version) to 2020	57
6.4	References	61
6.4.1	Bibliography.....	61
6.4.2	Links	61

1 INTRODUCTION

The VCX Forum is a global consortium of industry leaders and experts who are dedicated to advancing the quality and performance of smartphone cameras, webcam, smart home security cameras & XR devices. The VCX Forum develops and maintains standards and specifications that define the best practices and methodologies for evaluating and improving camera quality on devices under their purview. This whitepaper focuses on the image quality metrics of smartphone cameras.

One of the key outputs of the VCX Forum is the VCX Score, a comprehensive and objective metric that quantifies the quality of smartphone camera images and videos. The VCX Score is based on rigorous testing and analysis of various aspects of smartphone camera quality, such as exposure, color, sharpness, noise, dynamic range, and frame rate. The VCX Score provides consumers, manufacturers, and reviewers with a reliable and consistent way to compare and benchmark smartphone camera performance across different devices and scenarios.

The VCX Score is not a static metric, but a dynamic one that evolves with the changing needs and expectations of smartphone camera users. The VCX Forum regularly updates and enhances the VCX Score to reflect the latest trends and innovations in smartphone camera technology and user behavior.

The VCX Score was created from collaborative and inclusive development that involves the participation and contribution of various stakeholders in the smartphone camera ecosystem, including major OEMs, ISP developers, software and hardware suppliers, and other experts. The VCX Forum's PhoneCam Standard Department is composed of representatives from these stakeholders, who work together to define, develop, and validate the VCX Score. The VCX Forum also welcomes feedback and suggestions from the wider smartphone camera community and strives to incorporate them into the VCX Score as much as possible.

This white paper presents the latest version of the VCX Score, which is the result of four years of intensive research and development by the VCX Forum's PhoneCam Standard Department. This version introduces several major improvements and additions to the VCX Score, which are summarized below:

***User experience focus:** The VCX Score is based not only on technical measurements but also on user feedback and preferences. The VCX Forum has conducted extensive surveys and studies to understand how smartphone camera users perceive and evaluate image and video quality, and what factors influence their satisfaction and dissatisfaction. Based on these insights, the VCX Forum has*

increased the scope of research that contains the user experience by more than 15 times and has tuned the correlation between the technical measurements and the user ratings accordingly. This ensures that the VCX Score reflects the real-world needs and expectations of smartphone camera users, and not just the theoretical or ideal ones.

New test areas: *The VCX Score covers a wide range of test areas that represent the most common and important use cases and scenarios for smartphone camera users. In addition to the existing test areas, such as still image, video, and zoom, the VCX Score has added four new test areas, namely selfie camera video, wide-angle camera of the main camera, 10x zoom test and video zoom test. These new test areas address the growing popularity and demand for these features among smartphone camera users, and the challenges and opportunities they pose for smartphone camera quality.*

More comprehensive and granular evaluation: *The VCX Score evaluates smartphone camera quality based on 24 key performance indicators (KPIs) that are driven by specific consumer complaints in areas such as auto exposure, white balance, spatial frequency response, texture loss, sharpness, noise, color reproduction, exposure/dynamic range, and frame rate. These KPIs are measured in various controlled lighting environments that mimic the environment and conditions where smartphone cameras are mainly used, in consistent and clearly defined scenarios. The VCX Score also provides sub-scores for each test area and KPI, as well as an overall score that aggregates the sub-scores. This allows for a more comprehensive and granular assessment and comparison of smartphone camera quality across different dimensions and levels.*

The rest of this white paper provides more details and explanations about the VCX Score, its methodology, its results, and its implications. The white paper describes the user experience research that underpins the VCX Score, and how it informs and influences the technical measurements and the correlation between them. It also explains the test areas and the KPIs that constitute the VCX Score, and how they are measured and calculated. In the end, it provides some directions and recommendations for future work and improvement of the VCX Score.

2 MEASUREMENT

2.1 Approach

The VCX Score is a single, objective, and comprehensive metric that reflects the user experience of smartphone camera quality and performance. The VCX Score is derived from well-defined and transparent test procedures that follow international standards where applicable. The test procedures measure image quality and performance under different controlled lighting conditions that simulate real-world scenarios where smartphone cameras are used.

The VCX Score is based on 100% objective data, which means that no human judgment or intervention is involved in the analysis or the score generation. The analysis is solely based on the images captured by the smartphone cameras and the algorithms applied to them. The score is calculated using a fixed algorithm that uses the numerical results of the analysis. The score is not influenced by any subjective impression or preference of any individual or group.

The image quality is evaluated for different use cases (see section 2.3 for test conditions) that cover the most important aspects of smartphone camera quality, such as:

- Spatial Resolution: How much detail can the camera capture?
- Texture Loss: How well can the camera reproduce fine and low-contrast details?
- Sharpening artifacts: How much sharpening does the camera apply, and does it cause any artifacts?
- Noise: How much noise does the camera produce, and how does it affect the image quality?
- Dynamic Range: How well can the camera handle high-contrast scenes and preserve details in shadows and highlights?
- Color Reproduction: How accurate and natural are the hues produced by the camera?

The VCX Score provides sub-scores for each use case and each aspect of image quality, as well as an overall score that combines them. This allows the user to see the strengths and weaknesses of each smartphone camera and how they perform in different scenarios. For example, the VCX Score considers the low-light performance of the smartphone camera, which is crucial for smartphone camera users who often use their cameras in dim or dark environments. The VCX Score measures how much the image quality degrades in low-light conditions, depending on the movement of the device. The VCX Score ensures that the smartphone cameras that perform well in low-light conditions can achieve a higher score. *For example, the VCX Score considers the low-light performance of the smartphone camera, which is critical.*

CONTENT

The VCX Score also considers the performance of the smartphone camera, which affects the user experience. The performance is measured by how fast the camera can adjust and then capture images, which is important for capturing spontaneous and fleeting moments.

The VCX Score is a holistic and rigorous metric that evaluates all the relevant aspects of smartphone camera quality and performance simultaneously. The score will penalize shortcuts or trade-offs in smartphone camera tuning that would compromise the user experience. The VCX Score ensures that only smartphone cameras that deliver good quality and performance across the board can achieve a high score.

2.2 Procedure

The VCX Score requires that each smartphone camera is tested under optimal conditions, without any mechanical or software issues that could affect the image quality or performance. Therefore, before the test procedure begins, each device is inspected for any visible damage, such as scratches on the lens or signs of impact. The lens is also cleaned, and the device is rebooted to ensure that no other processes interfere with the camera's operation.

For each test scenario at least four images are captured, the score per image is calculated and then the best one is used for further processing. The camera is also forced to re-focus after each image capture.

2.2.1 Camera Settings

The VCX Score requires that each smartphone camera is tested under optimal conditions, without any mechanical or software issues that could affect the image quality or performance. Therefore, before the test procedure begins, each device is reset to its factory settings, which are the default settings that most end-users do not change. The only exception is the flash and burst mode, which are turned off if they are activated automatically by the device during the test.

The VCX Score analyzes the default images that are produced by the smartphone camera after capturing an image. These images are typically 24-bit JPEGs with 8-bit per channel, and they use the color space that is specified by the device. In most cases, the color space is sRGB, but some devices use other color spaces, such as P3 Display. The VCX Score respects and uses the color space of the images for the analysis.

The VCX Score does not alter any manual settings that the device allows, such as exposure compensation, ISO settings, and so on. These settings remain unchanged during the test, as they are

CONTENT

by default. The VCX Score also does not change the aspect ratio or the pixel count of the images, which are determined by the device's default settings. For example, if the device uses 16:9 aspect ratio by default, the VCX Score tests the device using the same aspect ratio.

2.2.2 Measurement Setup

The VCX Score requires that each smartphone camera is tested under consistent and controlled conditions, without any external factors that could affect the image quality or performance. Therefore, before the test procedure begins, each device is mounted on a tripod that is customized to fit the device's hardware and button positions. The tripod is attached to a rail system that allows adjusting the distance between the device and the test chart. The distance is chosen so that the test chart fills the entire frame of the device's camera. The framing is verified on an external PC screen before capturing the final images. The device is also aligned to the test chart as precisely as possible, to avoid any roll or tilt errors between the chart plane and the sensor plane.

The test environment is also controlled to ensure standard room temperature ($23^{\circ}\text{C} \pm 2^{\circ}\text{C}$) throughout the test procedure. The test charts used are reflective charts that are mounted on a fixture to ensure proper alignment. The illumination of the test charts follows a reprographic approach, which means that two light sources from left, and right are used to provide uniform and consistent lighting on the target and to prevent any reflections on the test chart that could interfere with the measurements.

The test procedure also maintains a constant reproduction scale for each device, which are defined by the device's field of view. This ensures that the test results are comparable and fair across different devices.

2.2.3 Test Charts

The device under test must reproduce reflective test targets under the defined test conditions.

The main chart is the "TE42-LL" (TE42-LL target in A1066 and A 460 (Selfie) in 4:3 and 16:9) developed and produced by Image Engineering. In the current VCX version, we use the "low light" version of this chart.

The reference data for the charts (optical density of the gray patches, XYZ coordinates of the color patches, resolution of structures) is obtained using high-quality measurement devices and comes with the chart. The charts are checked regularly, and if one starts to deteriorate, it is either replaced or the reference information is updated.

The TE42-LL chart is designed as a multi purpose test chart, which means that it contains several different test structures to measure different aspects of image quality. This concept has two main advantages:

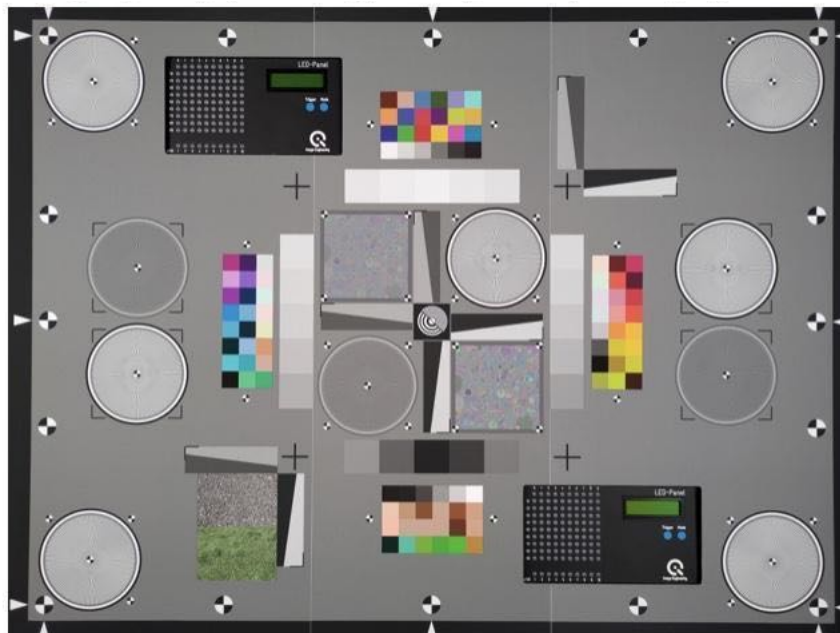
CONTENT

The measurement is time efficient as we do not have to frame, align, and capture several different test targets. All measurements are performed under the same conditions with exactly the same camera parameters.

Since the devices are tested in "fully automatic" mode and modern cameras use adaptive scene-dependent image processing, this chart layout helps to ensure that the different aspects of image quality are measured under the same conditions.

See section 3 for information on which structure is used for which metric.

The TE42-LL test chart is available in two sizes with an aspect ratio of either 4:3 or 16:9. Selecting the correct chart size depends on the aspect ratio of the device under test. The recommended chart is a high-resolution version ("H"), which features a spatial resolution that is high enough to test today's mobile phone cameras.



< The TE42-LL multi purpose test chart >

For the **dynamic range test**, a transmissive TE269B chart with 36 gray patches and a contrast of approx. 1.000.000: 1 is used in front of an integrating sphere to achieve a contrast range that exceeds what a camera can reproduce (see section 3.7).



< The TE269 test target with the LE7 illumination device (integrating sphere) >

Ultra-wide angle chart: The TE42-LL UW is used for measurement of devices that feature a wide lens module which is represented with a zoom factor of smaller than 1 in the user interface. It features the tartan pattern (see section 3.4) in the center and larger areas in the corner. The pattern in the corner is arranged that way so that even with strong barrel distortion, the pattern always covers all areas of the image. This way, the measurement in the image corner can always be be performed and is not limited by possible strong distortion (even though most devices feature distortion correction that limits the remaining distortion to a minimum).



< Chart – TE42-LL UW with Tartan patterns (Ultra-wide) >

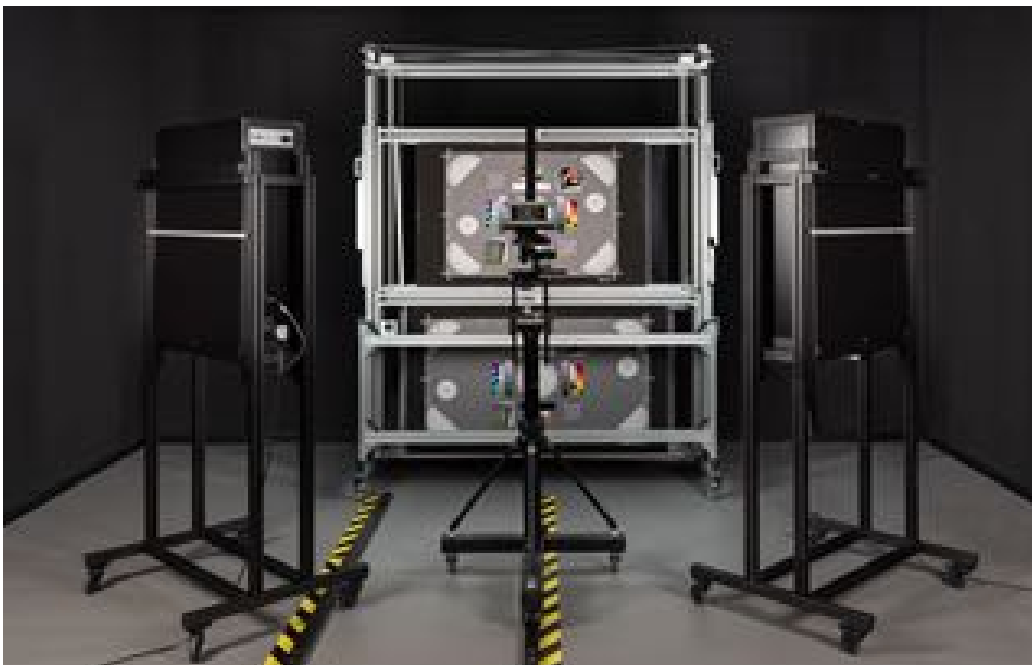
2.2.4 Test Equipment

The VCX Score requires that each smartphone camera is tested using additional hardware that complements the test targets. The test setup is also verified before each test, to ensure that the test conditions are consistent and accurate. The light intensity and uniformity are measured with a lux meter that is calibrated regularly. The device mounting and the tripod are inspected manually for any issues that could affect the device alignment or stability. The spectral distribution of the light sources is also checked monthly to ensure that the color temperature and quality are maintained.

2.2.4.1 Illumination

The TE42-LL test target is illuminated by two iQ-Flatlights, sold by Image Engineering, which are LED light sources that can generate custom spectra to simulate different light sources in a test lab environment. The iQ-Flatlights are ideal for VCX testing, as they can reproduce all the light conditions specified in this white paper. Each iQ-Flatlight has ten iQ-LED modules, which have 20 LED channels each and can provide a maximum intensity of 2000 lx (for both iQ-Flatlights) for standard D illuminants at 1.5 m from the test target.

For the dynamic range test, a different test setup is used: an integrating sphere with Image Engineering's LE7 device, which is a uniform lightbox that also uses LED technology. The LE7 can provide a tunable spectrum and high uniformity for illuminating transparent test charts. The LE7 is used with the TE269-B (V3) chart, which is a transmissive chart with 36 gray patches and a contrast of approximately 1,000,000:1. The details of the dynamic range test are provided in section 3.6.



< An example setup of test chart and illumination (iQ-Flatlight)" >

2.2.4.2 Timing Kit

The VCX Score requires that each smartphone camera is tested using a precise and consistent timing measurement. The timing measurement determines the delay between the user's action and the camera's response, such as the shooting time lag, the shutter release time lag, the startup time, and the burst frame rate. The timing measurement is based on the ISO 15781 standard, which describes the methods and procedures for measuring the timing features of a digital camera.

There are two options for performing the timing measurement:

Option 1: Use the Image Engineering timing kit, which consists of an AF-Box and two LED-Panels. The AF-Box is a device that provides uniform and adjustable illumination for a test chart, such as the TE42-LL chart. The LED panels are devices that act as high-precision clocks by illuminating 100 LEDs in sequence at a selectable speed. The LED panels can measure the time, even if the camera has a long exposure time, by detecting the first LED that appears in the image. The Image Engineering timing kit is a stand-alone setup requiring no additional hardware or software.

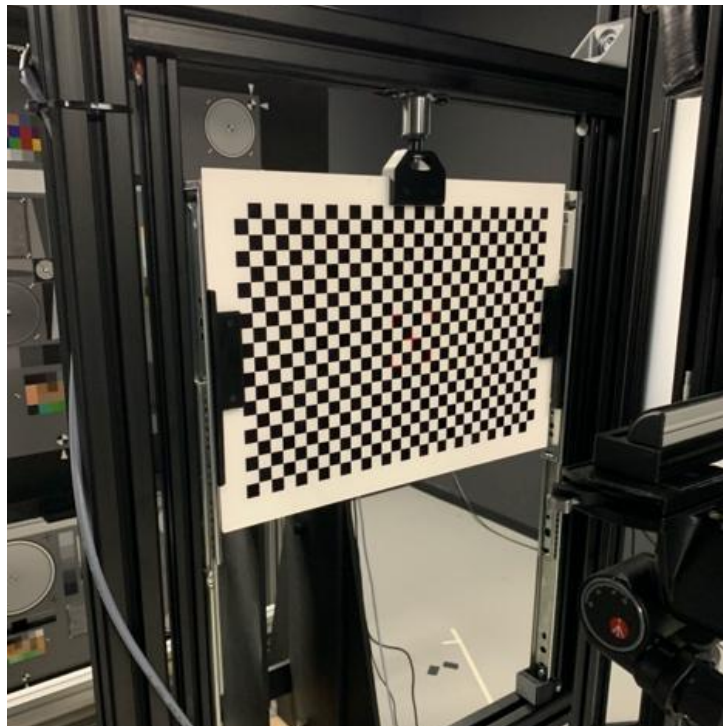
Option 2: Use the TE42-LL-Timing chart, a special version of the TE42-LL chart with two LED panels integrated into the layout. By using the LED panels as clocks, the TE42-LL-Timing chart can measure both the low-light performance and the timing features of a camera in a single image. The TE42-LL-Timing chart can be integrated into the main test stand without requiring a separate working space or an illumination box.

The main component for both options is the LED Panel, which is a device that can accurately measure the timing features of a camera. The LED Panel can be controlled by a USB interface, a manual remote, or a C++ API. The LED Panel can also be combined with the iQ-Trigger or the iQ-Trigger-T, which are devices that can automatically trigger the camera's shutter release button, either mechanically or capacitively. The iQ-Trigger and the iQ-Trigger-T eliminate the human error that could occur when pressing the shutter release button manually.



< The Image Engineering LED-Panel >

2.2.4.3 *iQ-Near Focus device*



< iQ-Near Focus device >

One of the challenges of the timing measurement is to set the autofocus system of the camera to a defined state before each test. The VCX Score uses the iQ-Near Focus device, which is a device that

can control the autofocus state of the camera electronically. The iQ-Near Focus device consists of a small, high-contrast test chart that can be placed at a short distance from the camera. The test chart is mounted on a slider that is held by an electromagnet. When the electromagnet is turned off, the test chart slides down and out of the camera's view, allowing the camera to focus on the larger test chart behind it. This way, the camera starts to focus from a defined state at a defined time, which enables accurate timing measurements. The iQ-Near Focus device can be used with various test scenarios, such as the TE42-LL chart or the LED-Panel, to measure different aspects of the camera's timing performance.

2.2.4.4 STEVE

"STEVE" is an abbreviation for "STabilization EValuation Equipment". The VCX Score requires that each smartphone camera is tested using the STEVE device, which is a device that can measure the performance of the image stabilization system of the camera. The STEVE device is a hexapod that can generate controlled vibrations in six degrees of freedom (pitch, yaw, roll, x, y, z) to simulate the natural human hand tremor that occurs when holding a camera. The STEVE device is used with the TE42-LL test target, which is a multipurpose test chart that can measure various aspects of image quality in low-light conditions. The STEVE device can apply different vibration profiles to the camera, such as the CIPA-compliant profile or the representative handshake profile developed by Apple.

The image stabilization performance of smartphone cameras is challenging to measure, because the camera automatically adjusts its settings, such as the exposure time or the ISO sensitivity, depending on the scene and the lighting conditions. Therefore, the VCX Score monitors and measures all the relevant image quality parameters, such as the resolution, the noise, the dynamic range, and the color reproduction, while testing the image stabilization performance. The VCX Score follows the ISO 20954-2 standard, which defines the measurement method for image stabilization performance for systems that include non-optical elements, such as digital or electronic image stabilization.



< Replicating human hand tremors using "STEVE" >

2.3 Measurement Conditions

2.3.1 Main Camera (Rear camera)

The image quality evaluation is conducted using the TE42-LL test chart while the device is securely mounted on a tripod. This assessment occurs under distinct measurement conditions below:

Bright – This condition serves as the reference point and is performed under a brightness level of 2000 lux, specifically in daylight conditions (D55).

Mid – In this scenario, the illumination level is set to 250 lux using neutral LED lighting. This condition simulates a typical indoor environment without direct sunlight.

Low – Low-light situations pose the greatest challenge for a camera. Here, we reduce the illumination to 10 lux, employing warm LED lighting.

Extended Low light – Under this condition, the device captures images with progressively decreasing illumination. We establish a reference at the bright level (as mentioned above), followed by images taken at 10 lux, 7.5 lux, 5 lux, 3 lux, 2 lux, and 1 lux.

Zoom – Zoom functionality is a critical aspect of modern smartphone cameras. To perform these evaluations, we adjust the distance between the TE42-LL chart and the camera according to the zoom factor. The entire chart is captured to ensure comprehensive assessment. The VCX scoring now considers not only 4x zoom but also wider zoom ranges, including 10x zoom and wide-angle capabilities. Let's delve into the details:

Ultra-Wide: *Wide-angle shots capture expansive scenes, making them ideal for landscape photography, group selfies, and architectural details. We assess how well the camera performs in wider field-of-view scenarios. Indicator is the lowest zoom factor the UI allows the user to define, typically smaller than 1x for this test.*

4x Zoom and 10x Zoom (Optical and/or Digital): *VCX Test applies a 4x digital zoom to capture images of the TE42-LL chart. This is to check their performance at higher zoom levels in still images. VCX Test evaluates devices with 10x zoom at this extended level. If the device supports 10x optical zoom, we evaluate its image quality at this extended zoom level; however, if it does not provide it, it will be excluded in the score calculation rather than having a penalty score.*

2.3.2 Video (Rear camera)

The image quality evaluation of the video is based on the TE42-LL test chart; an extracted image of a frame from the 10 seconds of the video is used for the calculation. The frames for analysis are extracted using FFmpeg after at least 5 seconds in the video file.

VCX Test's initial frame-based analysis provides insights into critical image quality factors related to low-light performance. As camera technology evolves, we adapt our evaluation methods version by version to ensure accurate and relevant results.

This is also performed under different measurement conditions:

Bright – This condition serves as the reference point for video quality assessment. It is conducted at a brightness level of 2000 lux (Daylight D55).

Mid – This simulates a typical indoor environment without direct sunlight. Illumination is maintained at 250 lux using neutral LED lighting.

Low – VCX tries to ensure the most demanding scenario for a camera in the reality of the user's daily situation. It evaluates the camera's performance in challenging low-light conditions. Illumination reduced to 10 lux using warm LED lighting.

Zoom – The device is set to a 3x zoom to record a video of the TE42-LL chart. This is to check performance at higher zoom levels in video recording images.

2.3.3 Selfie Camera (Front camera)

We evaluate the Selfie camera using the same three lighting conditions as for the Main camera (see section 2.3.1 for details)

Video – VCX 2024 version includes video testing to ensure the video quality. As the importance of selfie videos for personal shooting and broadcasting on social media has increased, VCX has applied the tests that were only applied to the main camera to selfies to reflect such trends.

2.3.4 Performance measurements

Timing and Image Stabilization performance measurements are carried out under D55 at 250 lux and 10 lux.

2.4 Viewing Conditions (VC)

When assessing certain metrics, it is essential to define the viewing conditions. The primary metrics include “Visual Noise” (section 3.6), the Acutance analysis of SFR measurement (section 3.1.5, 3.1.7 and 3.2.2), and the evaluation of Undershoot and Overshoot (section 3.3.1).

To account for real-life scenarios, we make assumptions about the geometric setup, specifically the viewing distance and image size. We address three common viewing conditions:

VC 1 – 100% view – This represents the worst-case scenario, where the user can perceive most details. We assume a viewing distance of 0.5 meters and a 100% view on a 96ppi display. In this case, each pixel of the image corresponds to one pixel on the display. The more pixels the image contains, the larger it appears.

VC 2 – Small Print / Smart Phone Display – The entire image is scaled to a height of 10 cm, with the viewing distance set to the natural viewing distance. The natural viewing distance is defined as the diagonal of the image, with a minimum of 25 cm. Therefore, in this scenario, the viewing distance equals 25 cm.

VC 3 – Large Print / PC Display – The complete image is scaled to a height of 40 cm. The viewing distance is determined by the diagonal of the image, accounting for slight variations based on the aspect ratio.

3 METRICS

As explained in section 2.1, all metrics used are strictly objective and evaluated based on the images taken with the device under controlled conditions. This section will explain all metrics and numerical results that are used to create the VCX score. See the appendix 6.1 for an overview.

3.1 Resolution

For the VCX score, we clearly differentiate between pixel count and resolution. While the first one is simply defined by the number of pixels found in the final image, the resolution describes the level of details the camera can reproduce.

A camera system is a highly adaptive system, so depending on the scene content, it will behave differently. High-performance image signal processors (ISP) allow devices to process an edge differently than a uniform area in the image. That way a strong noise reduction in the ISP is possible with the aim to clear uniform areas in the image from image noise without blurring edges.

As we must assume that the device under test will behave differently on different structures in the image, we use several different structures to produce a meaningful and reliable resolution measurement. The resolution measurement is combined with a measurement of the sharpening and the texture loss (see the following sections for details).

3.1.1 Theoretical Pixel Count (TPC)

The theoretical pixel count is simply calculated as the product of image height (picture height, PH) and image width (picture width, PW). This is extracted from the final image the device under test stores to memory and is expressed in Megapixels. The theoretical pixel count is not directly part of the score and a high number of theoretical pixel count by itself does not gain any benefit in the VCX score.

$$TPC = \frac{PH \times PW}{10^6}$$

3.1.2 Theoretical Maximum Resolution (TMR)

A certain number of pixels allow a device to reproduce a certain level of detail. The theoretical maximum resolution (TMR), also called *Nyquist frequency* ($f_{Nyquist}$), is the highest possible resolution based on the sampling frequency.

The ISO 12233:2014 standard defines different units to express the resolution of a device. The unit cycles per pixel express the number of line pairs (cycles) that can be reproduced by the device under test per pixel. The theoretical maximum here is 0.5, as only a half line pair can be reproduced with one line of pixels.

As the unit cycles per pixel do not provide any information about the total amount of details that the device under test is capable of reproducing, the more useful unit is line pairs per picture height (LP/PH).

CONTENT

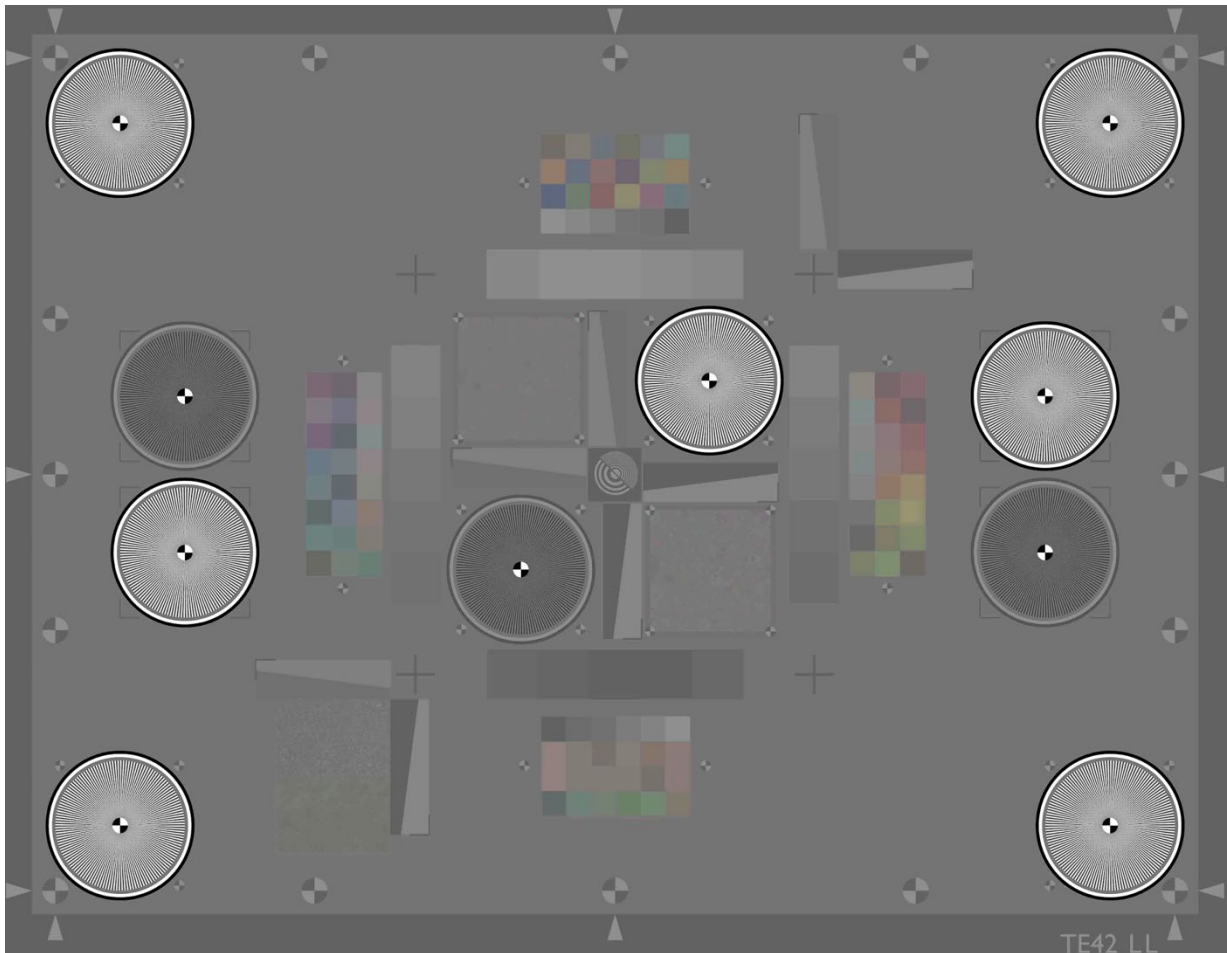
So, the theoretical maximum resolution expressed in LP/PH is calculated as:

$$TMR = \frac{PH}{2}$$

The TMR is used for further calculations and does not impact the VCX score directly.

3.1.3s-SFR - Limiting Resolution (LR) Center / Corner

The limiting resolution describes the maximum level of details the device under test can reproduce. It is measured using the s-SFR method described in ISO 12233:2014. Because the s-SFR method is based on a sinusoidal Siemens star, we use the Siemens stars implemented in the TE42-LL test chart.

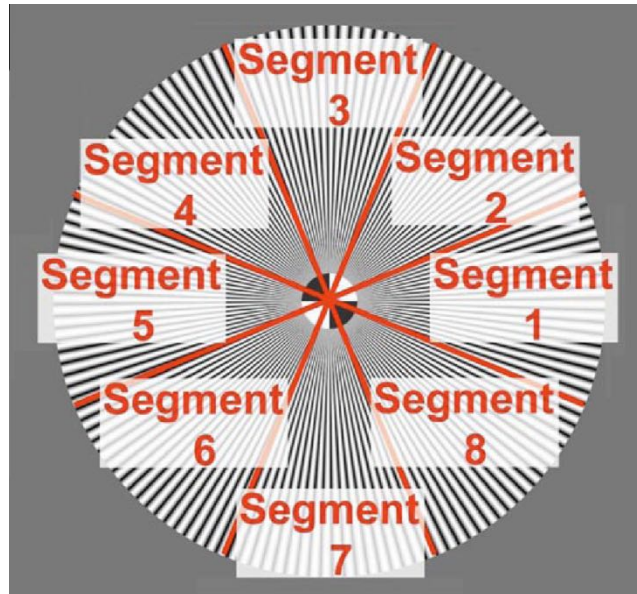


< The sinusoidal Siemens stars within the TE42-LL test chart >

The s-SFR method is known for being less influenced by sharpening and other image enhancement algorithms that might impact the measurement. It is most suitable for the measurement of the so-called limiting resolution. This metric is defined as the spatial frequency (in unit LP/PH) that leads to a modulation transfer of 10%. The analysis is performed based on the final JPG image data. The image is linearized based on the measured tone curve to revert the influence of gamma correction and tone curve optimization on the results.

CONTENT

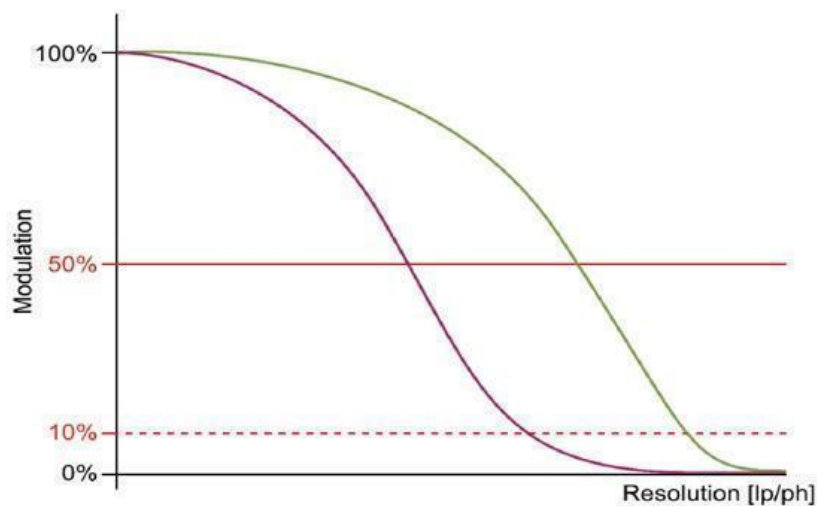
The analysis of the Siemens star provides the possibility to determine the resolution for different orientations. The stars are divided into segments and each segment is analyzed separately.



< The eight segments per star >

The reported limiting resolution for the image center is calculated based on the average SFR from all eight segments of the high contrast center star. So first the average SFR is calculated and based on this SFR the limiting resolution value is derived.

The limiting resolution of the corner is calculated as the average of all available segments in the four corner stars. Each of the corner stars contains three segments.



< Sample SFR of center (green) and corner region (purple). Limiting resolution or MTF10 is the frequency at which the SFR drops below the dashed line (10%) >

If the limiting resolution of the device under test is so low that it cannot be measured by the used test target (e.g. in case of digital zoom) it will result in the lowest possible score for resolution.

In some cases, it might happen that the measured limiting resolution is higher than the theoretical resolution ($LR > TMR$). In this case, the limiting resolution is not calculated based on the average SFR of all eight segments but based on the limiting resolution of each segment.

We make a difference between horizontal, vertical and diagonal segments. Horizontal, and vertical segments (Segment 1,3,5,7) are limited to the Nyquist frequency, which equals the TMR. The diagonal segments (Segment 2,4,6,8) are allowed to have a higher frequency as in an image with square pixels, an ideal diagonal line pair can be narrower than an ideal horizontal line pair. The TMR for a diagonal line can be larger by a factor of the square root of 2 (140%).

$$LR_{all} = \begin{cases} \frac{\sum_{segment=1}^{segment=8} LR_{segment}}{8}, & \text{if } LR_{all} \leq TMR \\ \frac{\sum_{segment=1}^{segment=8} LR_{corr}}{8}, & \text{if } LR_{all} > TMR \end{cases}$$

with

$$LR_{corr} = \{ \min (LR_{segment}, TMR),$$

$$\text{for } segment = 1,3,5,7 \min (LR_{segment}, TMR * \sqrt{2}), \text{ for } segment = 2,4,6,8$$

3.1.4 Effective Pixel Count (EPC) Center / Overall

As the pixel count of a camera system does not describe the level of detail the system can reproduce, it is important to get a meaningful and easy way to communicate a number that can describe this.

Therefore, we use the effective pixel count in contrast to the theoretical pixel count.

The effective pixel count is calculated based on the limiting resolution (LR) as described in section 3.1.3 and states the number of megapixels a camera needs to have to reproduce the level of detail contained in the analyzed image. As the limiting resolution is dependent on the signal processing and therefore on the lighting conditions, the effective pixel count is also dependent on the measurement conditions. The ratio of the limiting resolution (LR) and the theoretical maximum resolution (TMR) generates the performance of the device, which is then multiplied by the theoretical pixel count (TPC).

$$EPC_{center} = \left(\frac{LR_{center}}{TMR} \right)^2 \times TPC$$

For the EPC_{center} the limiting resolution of the image center is used (LR_{center}). For the $EPC_{overall}$ the LR_{center} and LR_{corner} are used.

$$EPC_{overall} = \left(\frac{LR_{center} + LR_{corner}}{2 \times TMR} \right)^2 \times TPC$$

3.1.5s-SFR - Acutance Center / Corner

The measured resolution is a description of the level of detail a system under test can reproduce. It does not describe the *sharpness* of a system. Sharpness is a subjective impression of a human observer, and it depends on the spatial frequency response (SFR) and the viewing condition.

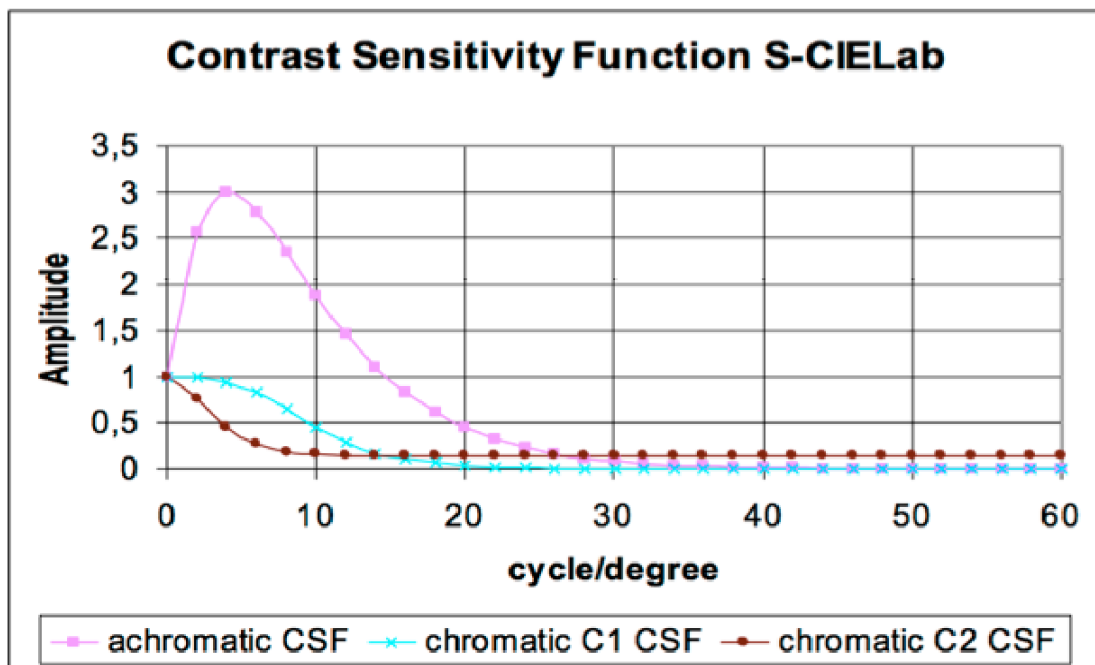
The calculation of the acutance is based on the contrast sensitivity function (CSF). The CSF is a model of the perception of spatial frequencies by a human observer. The CSF is scaled to the defined viewing conditions as described in section 2.4. The CSF used in this context is described in ISO15739:2013.

The acutance is the ratio of two integrals over the spatial frequency f in a range of the minimum spatial frequency that was analyzed in the SFR (f_{min}) and the Nyquist frequency (f_{Nyq}). The first integral A is the product of the measured SFR and the CSF. The SFR can be obtained from different structures; in this case, it is the s-SFR based on the Siemens star. The CSF is calculated and scaled according to the defined viewing condition (VC).

$$Acutance(SFR, VC) = \frac{A}{A_r}$$

$$A = \int_{f_{min}}^{f_{nyq}} SFR(f) \times CSF_{VC}(f) df$$

$$A_r = \int_{f_{min}}^{f_{nyq}} CSF_{VC}(f) df$$



< The Contrast Sensitivity Function used for the acutance calculation >

3.1.6e-SFR

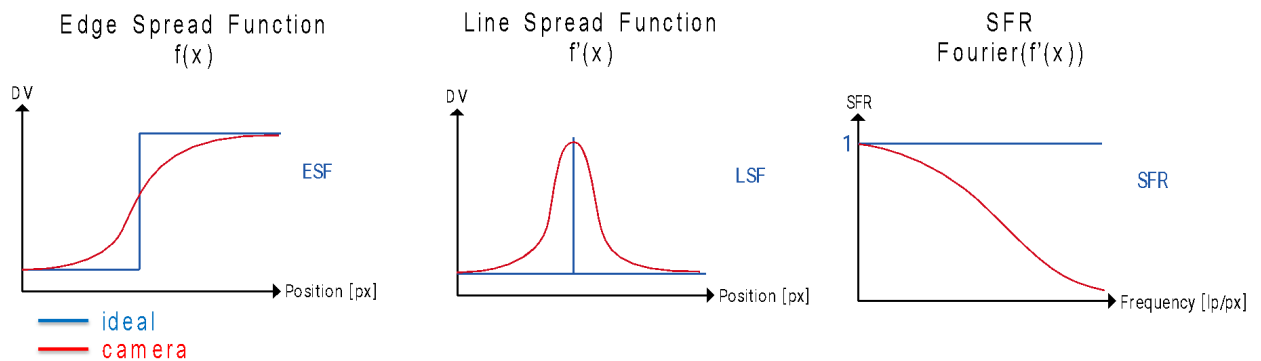
The ISO12233:2014 standard describes different methods to obtain an SFR. In the previous sections, the SFR was obtained from a harmonic Siemens star. The second method to obtain an SFR is the e-SFR based on slanted edges. This method was already part of the previous version of the standard from the year 2000. Signal processing in mobile phones is very powerful and it is very easy for a device to detect edges and to sharpen these. So, the e-SFR method is very useful to measure the sharpening. This dependency on the sharpening is the reason why we use the Siemens star to measure the limiting resolution, as the e-SFR method is influenced by sharpening and in the presence of sharpening it is not possible to obtain a meaningful MTF10 value (limiting resolution) that represents the resolution.

Section 3.3 describes the detailed measurement of sharpening.

The core procedure of the e-SFR method can be described in three steps:

1. Obtain the Edge Spread Function (ESF) from a slanted edge
2. Calculate the Line Spread Function (LSF) as the first derivative of the ESF
3. Calculate the SFR as the Fourier transform of the LSF

The most important part of this analysis is the measurement of the ESF. This function describes how the system under test reproduces an edge. The algorithm uses a super sampling process on all pixels along an edge to provide a high-resolution ESF, which means that the ESF has a four times higher precision than could be obtained from a single line perpendicular to the edge.



< The core analysis steps of the e-SFR method >

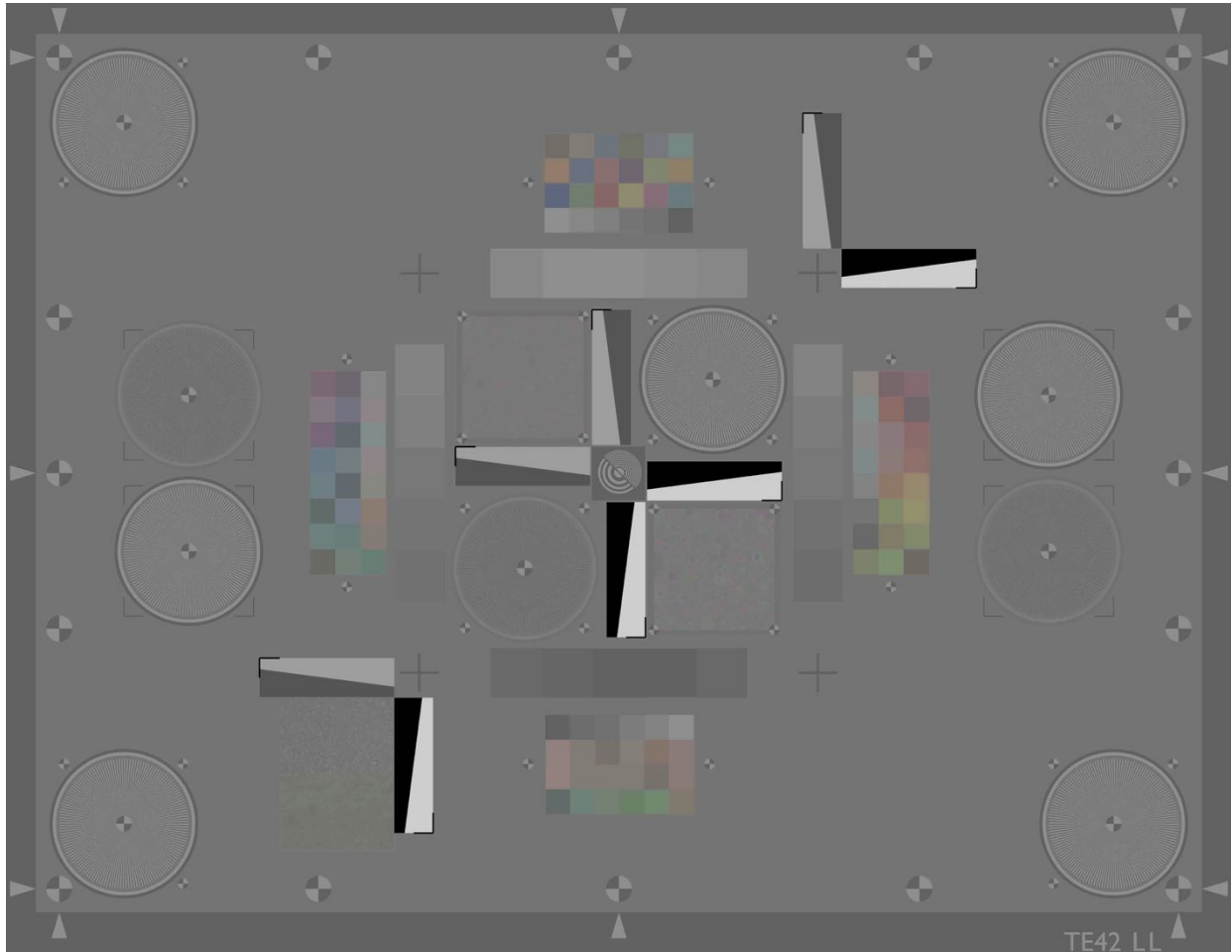
The ISO12233:2014 standard defines the edge contrast in the chart that shall be used for a camera analysis as 4:1, which is equivalent to an edge modulation of 60%. In many devices, the amount of sharpening and therefore the e-SFR is dependent on the edge modulation; so, a different edge modulation leads to the different behavior of the device under test. To measure the different behavior, the test chart used contains two different edges. The low contrast edge has a modulation of 60%, the high contrast edge has a modulation of 80%. Both edges are available in vertical and horizontal

CONTENT

directions; so, in total 8 edges are analyzed. The reported values are based on the average e-SFR for the edges with the same modulation.

3.1.7 e-SFR – Acutance

The acutance for the e-SFR is calculated in the same way as it is calculated for the s-SFR described in section 3.1.5.



< The slanted edges within the TE42-LL test chart >

3.2 Texture Loss (TL)

The image signal processor (ISP) in today's mobile phones works as an adaptive system. So, the processing and optimization of the image are dependent on the image content. A very common and important optimization step is noise reduction. In this case, the ISP analyzes the image content and depending on the analysis, different algorithms are applied to that region of the image. If the algorithm detects a region that does not have any structure (like the sky or other flat areas), the noise is reduced

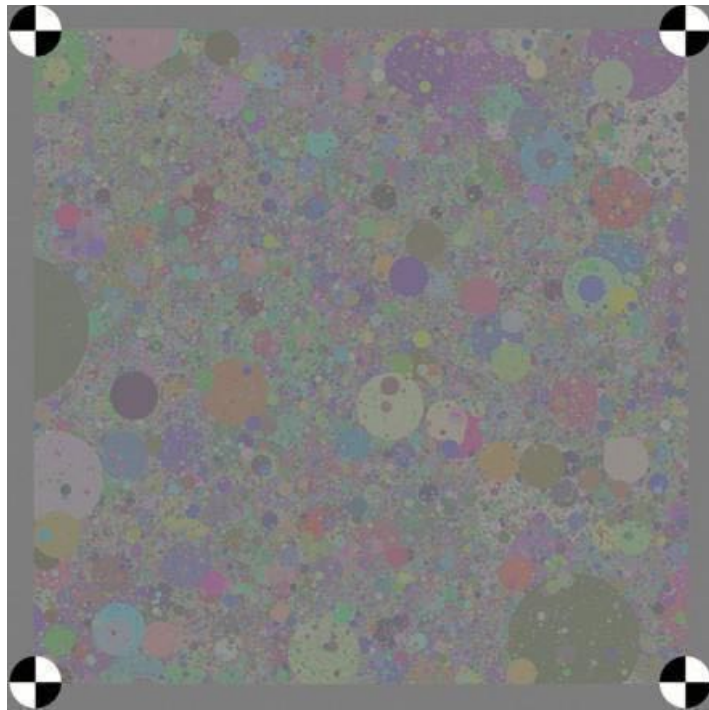
CONTENT

by averaging neighboring pixels. If an edge or another significant structure is detected, the noise reduction is not applied, as it would blur the image content.

The problem for these algorithms is to differentiate between image content and noise. While this is relatively easy for high contrast structures like in the Siemens star or the used slanted edges, it is very hard for low contrast, fine details.

The loss of these low contrast fine details is also called *texture loss*.

To measure the texture loss (TL), we use a colored version of the so-called dead leaves (DL) pattern. The dead leaves pattern is a random pattern built by stacking several thousand circles on top of each other. The size, location, and color of each circle follows a known probability function. In the TE42-LL, we have two versions of the same pattern with a different level of contrast. The analysis is performed on both patterns. The lower the contrast, the harder it is for the device to differentiate between image content and noise.



< dead leaves pattern (low contrast) >

The dead leaves pattern is well known in the industry and different analysis methods have been proposed over the past years. The latest one, described in ISO19567-2, is used for the VCX as it is the only one that is not negatively influenced by image optimization and artifacts, and which is very robust and reliable. The method used is based on cross-correlation between image content and reference data. The reference data is obtained from the original dead leaves pattern and an optimization process to match the reference to the image content. This method provides an SFR measured on the dead leaves structure and is called *DeadLeaves_{cross}*. The method from ISO19567-2

CONTENT

has been slightly modified for this version of the VCX protocol. The normalization using the DC component has been removed, resulting in an SFR that is not necessarily starting from 1 at the lowest frequency. The resulting dependence on the exposure of the image has been compensated by a contrast adjustment in the reference image before the cross correlation is calculated.

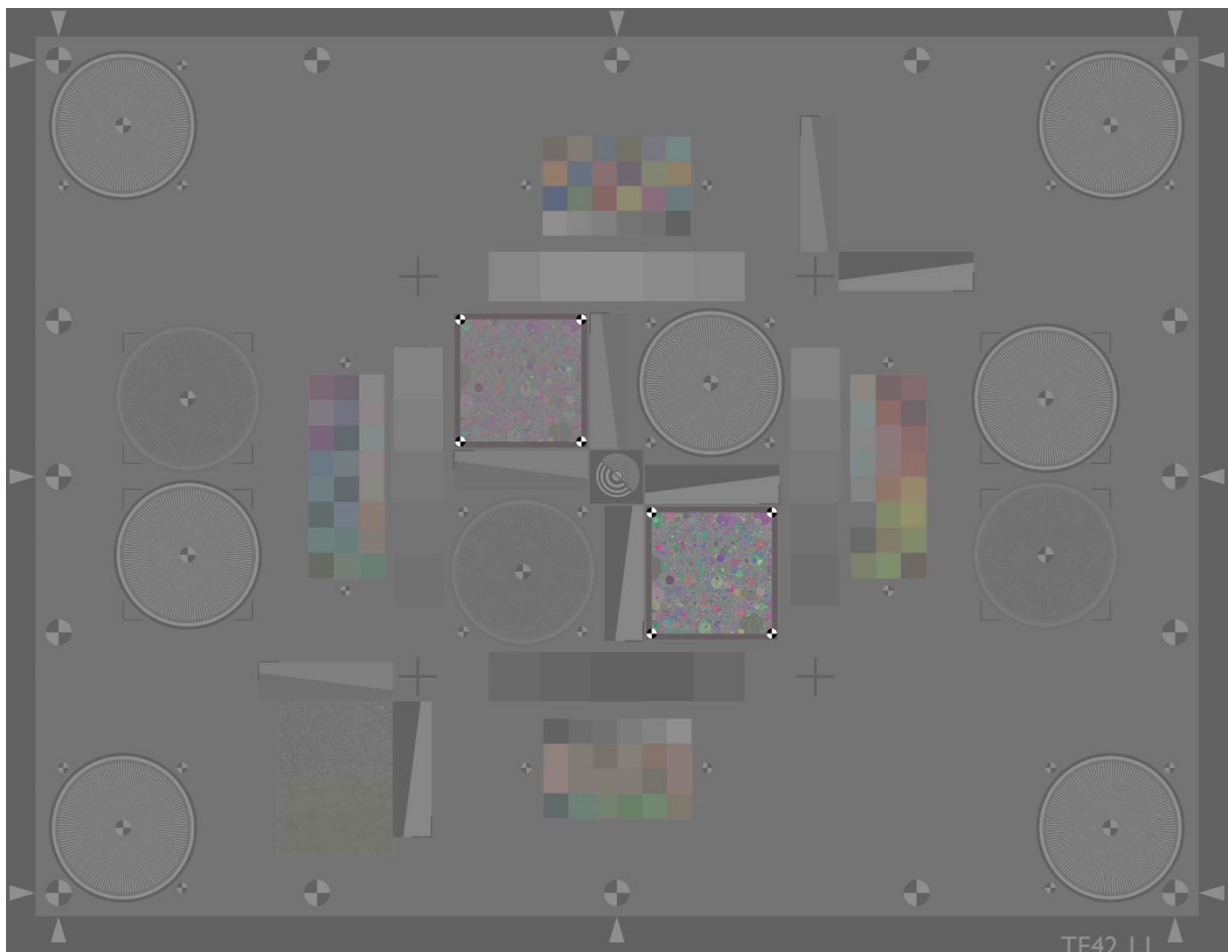
3.2.1 Texture Loss MTF10 / Limiting resolution

The SFR obtained from the $\text{Deadleaves}_{\text{cross}}$ is the root function from which the limiting resolution (defined as MTF10) value is derived. The MTF10 value reflects the level of detail that the device under test can reproduce if the structure contains low contrast, fine details.

3.2.2 Texture Loss Acutance

Acutance measurements based on SFR can quantify perception of a human observer for the given viewing condition.

The calculation is identical to the method described in section 3.1.5.



< Two different versions of the dead leaves pattern within the TE42-LL test chart >

3.3 Sharpening Artifacts

The subjective impression of “sharpness” recognized by a human observer can be increased by using sharpening algorithms in the ISP. Sharpening increases the local contrast, which gets visible along edges in the image.

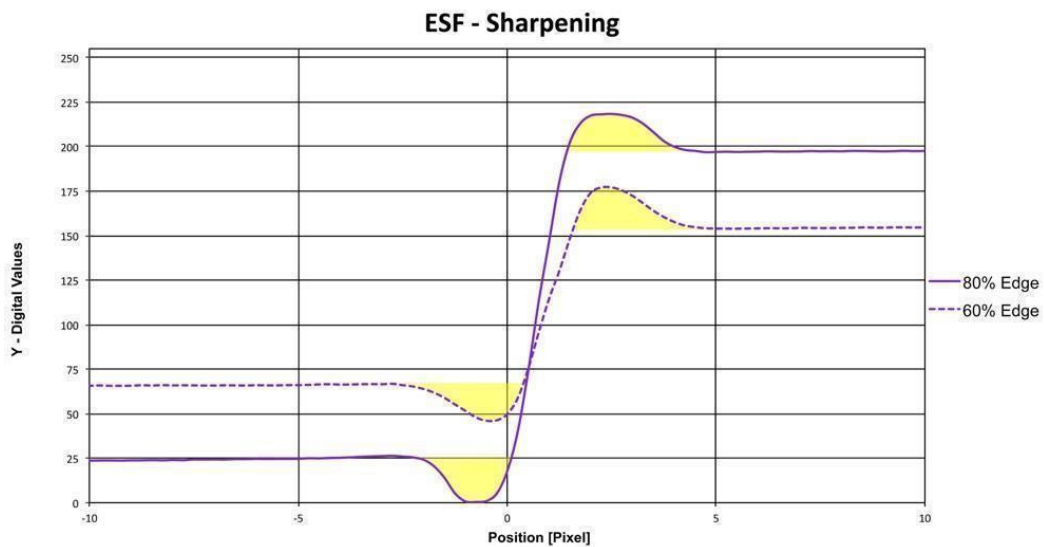
It is important to measure the amount of sharpening applied to the image for two reasons:

First, the influence on the image quality itself shall be checked. A certain amount of sharpening is beneficial to the subjective perceived image quality, as it can, as indented, increase the sharpness of an image. On the other hand, too much sharpening leads to artifacts along edges and the image tends to get an unpleasant artificial look. In extreme cases, edges between two gray areas have such high contrast enhancement that thin black and white lines can be observed along edges.

Secondly, sharpening has an influence on resolution and texture loss measurement. So, the SFR derived from these measurements can increase due to sharpening. This is correct, as the response of the camera to certain spatial frequencies is increased due to sharpening. For an in-depth evaluation, it is important to be able to identify the sharpening as a possible source of good numerical results.

3.3.1 Overshoot / Undershoot

For the e-SFR analysis described in section 3.1.6. the edge spread function (ESF) is determined. This function describes in high accuracy how the device under test reproduces edges. One of the most obvious artifacts that are introduced by sharpening is an undershoot and/or overshoot along edges, which is known as “ringing”.



< Undershoot and Overshoot in the ESF (yellow area) >

CONTENT

The amount of sharpening is measured for all available edges and the number is reported as the average of all edges with the same edge modulation. Undershoot and overshoot is expressed as the integral of that area beneath or above the ESF; represented by the yellow shaded area.

The edge consists of a bright and dark region. The overshoot is the area that is brighter than the bright region of the edge, while undershoot is the area darker than the dark region.

The measured area and the significance for the user are dependent on the pixel count. A device with a high pixel count can have a higher undershoot or overshoot without being noticed by a human observer. Therefore, we do not only report the total area that is measured but also scale it with the assumed pixel density for the given viewing conditions.

So, the scaled Overshoot (OS_{scaled}) is calculated based on the absolute Overshoot (OS_{abs}), the picture height (PH) and the object height (OH) as:

$$OS_{scaled} = \frac{OH}{PH} \times OS_{abs}$$

As an example, for the viewing condition, "VC3 - Large Print" the object height (OH) is 400mm.

3.3.2 Maximum SFR

Another clear indicator of sharpening is the maximum of the measured SFR. So, if the maximum is above 100%, the modulation or contrast in object space is smaller than in the image, so the camera increased this. As this cannot be achieved optically, this must come from the image processing unit. We check the maximum SFR value on all structures (Siemens star, slanted edges, and dead leaves pattern).

3.4 Tartan SFR

In the special case of very wide lenses (used for a zoom factor of smaller than 1 in the user interface), measuring resolution in the corner is challenging. Tests conducted within VCX showed that the performance of lenses drops significantly in the corner, which affects the user perception. As the signal processing within the imaging pipeline will try to compensate for the optical limitations of such modules, we observe higher noise, noise reduction artifacts and other effects that reduce the level of details that can be reproduced. Traditional methods to measure resolution, like the Siemens star and slanted edges, do not work well here, so the VCX workgroup developed a new approach to measure the performance in the corner.

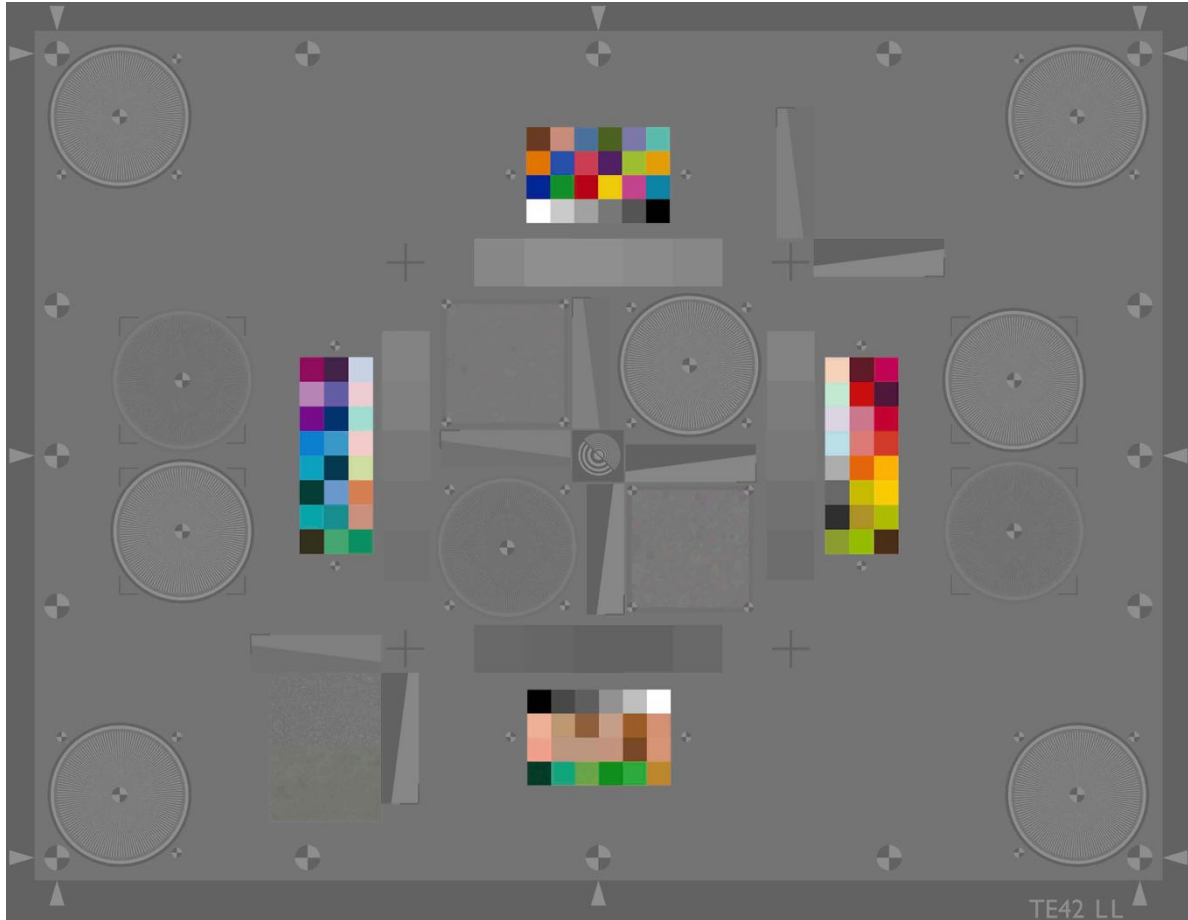
The pattern used for this consists of just a few spatial frequencies in two directions. The resulting pattern looks similar to the pattern seen in Scottish kilts, also called Tartan. When this pattern is reproduced by a camera, the spectral analysis can be used to generate a spatial frequency response (SFR) with a few discrete measurement points. The integral of this SFR is used as the key metric. TE42 LL UW chart (section 2.2.3) has these tartan patterns where SFR can be measured named as TartanSFR_center and TartanSFR_corner.

3.5 Colors

Many test procedures check the color reproduction quality of a device by reproducing a color target following an analysis of how well the rendered colors in the image match the colors in the target. The problem: Other than scanners, digital cameras are not made for a perfect reproduction of object colors, they are made for nice colors. The problem is to decide what nice colors are.

Even though it is well known that a perfect reproduction is not the aim of mobile phone cameras, it is important to check if some colors cause troubles and are not reproduced well. So the score calculation on color is designed that way, so that we identify issues but do not define narrow boundaries in which the device has to stay, so the manufacturer can include user preferences in their tuning without getting penalized.

For this test, the color patches of the TE42-LL are used. These patches are chosen with respect to the well-known X-Rite Color Checker SG color target. The reference data is obtained during chart production using a spectrometer and is expressed in the CIE-XYZ color space. The CIE-XYZ color space is based on the standard observer and can be used as a device-independent color space. Mobile phones typically provide the RGB images in the sRGB color space. This color space is well defined and the conversion from RGB to CIE-XYZ is defined in the sRGB specification IEC 61966-2-1. The CIE-XYZ color space is linear, but the human vision is not. A color comparison in CIE-XYZ would not reflect the experienced color differences by a human observer. For this purpose, the CIE-L*a*b* color space is most suitable. The reference data of the chart and the corresponding R, G, B mean values measured in the image are converted to CIE-L*a*b* color space and the color difference is calculated as ΔE , ΔL , ΔC , and ΔH for different patches. See the details in the following sections.



< The color patches of the TE42-LL test chart >

3.5.1 Color Reproduction Error

The color reproduction measurement is performed in the CIE-L*a*b* color space. The reference data is measured and available as CIE-XYZ values. The image data is transformed from the RGB to CIE-XYZ using a valid sRGB ICCv4 Profile (other profiles in the unlikely case that the image data is not sRGB). Both datasets (image data and reference data) are then converted into CIE-L*a*b* color space for further analysis. The conversion from CIE-XYZ to CIE-L*a*b* requires the definition of a white reference. We use the measured white in the image (patch D4) for the reference and the image data. This way an under-/overexposure is excluded from the analysis.

The color difference is reported for all patches or for just a subset of patches.

3.5.1.1 ΔE - Color Error (all)

For all 96 color patches, the Euclidean distance (ΔE) color difference according to CIE1976 is calculated.

$$\Delta E = \sqrt{(L_{ref} - L_{sample})^2 + (a_{ref} - a_{sample})^2 + (b_{ref} - b_{sample})^2}$$

CONTENT

The reported value is the average over all 96 patches.

1.1.1.1 ΔE - Color Error (red)

The same analysis as in 3.4.1.1, reduced to the reddish color patches only. The ΔE is calculated for the patches as stated in Annex 5.2, the average over these patches is reported

1.1.1.2 ΔE - Color Error (green)

The same analysis as in 3.4.1.1, reduced to the greenish color patches only. The ΔE is calculated for the patches as stated in Annex 5.2, the average over these patches is reported.

1.1.1.3 ΔE - Color Error (blue)

The same analysis as in 3.4.1.1, reduced to the blueish color patches only. The ΔE is calculated for the patches as stated in Annex 5.2, the average over these patches is reported.

1.1.1.4 ΔE - Color Error (skin tones)

The same analysis as in 3.4.1.1, reduced to the skin tone color patches only. The ΔE is calculated for the patches as stated in Annex 5.2, the average over these patches is reported.

1.1.1.5 ΔL - Luminance Error (all)

The L^* component contains the brightness information. The Luminance error (ΔL) is calculated for all patches.

$$\Delta L = L_{ref} - L_{sample}$$

1.1.1.6 ΔC - Chrominance Error (all)

Based on the CIE-a* and CIE-b* component, the saturation or chrominance of each patch can be calculated. The difference in chrominance of reference and sample is reported as ΔC .

$$\Delta C = \sqrt{a_{sample}^2 + b_{sample}^2} - \sqrt{a_{ref}^2 + b_{ref}^2}$$

As shown in the equation, ΔC can be positive or negative. Therefore, a ΔC value of zero does not mean that all patches perfectly reproduce the chrominance of each patch as positive and negative ΔC can sum up to zero. A positive ΔC value means that in average the saturation is higher than in the reference.

1.1.1.7 ΔH - Hue Error (all)

The differences in color tone are expressed as ΔH .

$$\Delta H = \sqrt{(a_{ref} - a_{sample})^2 + (b_{ref} - b_{sample})^2} - \Delta C^2$$

The reported number is the average over all patches.

3.5.2 White Balance (WB)

A human observer adopts his color perception to the assumed light source. That way a color patch looks nearly the same under different light situations, even though the color stimulus is different. A camera cannot change its perception so the adaptation to different lights must be performed as a separate step in the color processing. A camera must estimate the white point of the current scene illumination and perform a white balance according to this information.

The difficult part for a camera is to correctly estimate the illuminant.

As the test chart consists of white patches and a huge amount of gray background, the estimation of the illuminant should be relatively easy for the device under test.

The white balance shall change the color processing in a way that gray patches appear neutral, regardless of the white point of the illuminant. This is true for a high illumination level, but in low light, the camera might introduce some color cast by intention, as this might match the perception of a human observer better.

The quality of the white balance is reported as the average CIE-C* over all gray patches used for the noise analysis (see section 3.6). The CIE-C* is calculated based on the image's RGB data. This data is converted to CIE-XYZ using an appropriate ICC color profile. The resulting CIE-XYZ data is then converted into CIE-LCH which is a representation of CIE-L*a*b* in a polar coordinate system. A C* value of zero means that the color is completely neutral. The higher the number the more chrominance and color is visible in the gray patches.

It is understood and accepted, that for the measurement under low light with warm LED light, the camera will not render the image with the target for neutrality. So, the devices want to keep some of the warmth in the image and not make it perfectly neutral. This is considered in the score calculation.

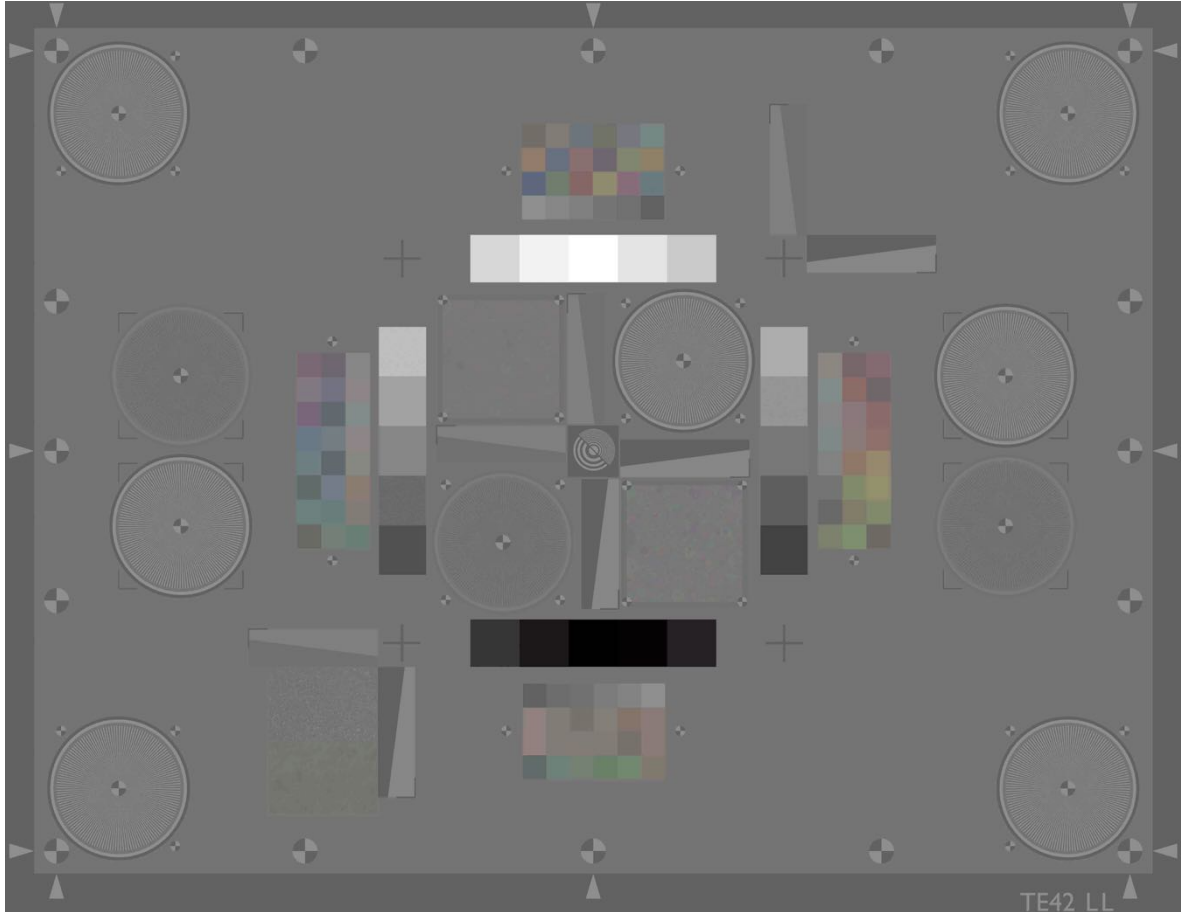
3.6 Visual Noise (VN)

Noise is defined as the fluctuation of measurement due to random processes. As each pixel is basically a measurement of the object luminance, noise introduces random variations in pixel values. In signal processing, the amount of noise is very often expressed as the signal to noise ratio (SNR). In imaging, the SNR is also commonly used, but it has been shown that the SNR value does not correlate well with the human perception of noise in the image.

A much better metric to describe the noise is *Visual Noise*. This metric describes the noise with respect to human perception and the viewing condition. The analysis procedure is defined in ISO15739:2013, and the contrast sensitivity function (CSF) used is also described in ISO15739:2013.

CONTENT

Visual Noise is dependent on the defined viewing condition. The defined viewing conditions for the calculation are described in section 2.4 All reported values reflect the maximum of viewing condition one and two. The metric that produces a higher value from both the Max value and the Mean value will be selected for score calculation.



< gray patches of the TE42 test chart; used for the noise >

3.6.1 Visual Noise (mean)

The visual noise value is calculated for all 20 gray patches. The reported value is the average of 16 patches, excluding the two darkest and two brightest patches from the calculation. This is done to avoid that a low dynamic range is a benefit for the camera in this aspect.

3.6.2 Visual Noise (max)

The noise is changing with the luminance of the object. As digital noise reduction in the ISP is very common in mobile phone cameras, the appearance of noise can significantly change for the gray patches and can also have some peaks, where the noise is very prominent. Therefore, we also check the maximum value and not only the mean value as two cameras might have the same mean value, but one can have more noise in some particular intensities.

3.7 Dynamic Range (DR)

We measure the input-referred dynamic range based on the principles described in ISO15739:2013. This metric describes the maximum scene contrast the device under test can reproduce. A low dynamic range results in clipped highlights and problems with details in the shadows. The dynamic range is calculated based on the Opto-Electronic Conversion Function (OECF). The OECF describes how the device under test reproduces different scene luminance into digital values. To generate the OECF the gray patches in the TE269 chart are used.

To generate the OECF, the luminance of each gray patch (L_i) is calculated based on the known optical density (from the chart manufacturing data), the scene illumination level¹ and the mean value of the pixels found in the region of interest (ROI) in the Y-Channel² (Y_i).

$$L_i = OECF(Y_i)$$

The dynamic range is defined as the ratio of L_{sat} and L_{min} and is reported in the unit “f-stop”. L_{sat} and L_{min} are derived from the measured OECF.

$$DR_{[f-stop]} = \frac{\log_{10}\left(\frac{L_{sat}}{L_{min}}\right)}{\log_{10}(2)}$$

L_{sat} is defined as the smallest luminance level that leads to saturation in the image, so the maximum luminance that can accurately be reproduced. A luminance higher than that would get the same digital output value, so no information is gained any more. For most devices, the maximum digital value is 255, as the image is represented in 8bit sRGB.

The lowest scene luminance (L_{min}) that can be reproduced is defined via the signal to noise ratio (SNR). The SNR is calculated according to ISO15739:2013 per gray patch. If the SNR is too low, the user cannot differentiate between image content and noise, so the information is lost. The threshold for the SNR equals one in the ISO standard. In our experience, this threshold is too low for two reasons:

- Even with a slightly higher SNR than one the signal will not be visible because it is still covered by noise
- Due to strong noise reduction, many devices do not reach the SNR = 1 threshold.

For these reasons, we calculate the L_{min} based on a threshold of SNR = 3. In cases where the SNR=3 threshold is not reached (due to noise reduction), the L_{min} is extrapolated.

¹ For the calculation of the dynamic range, only the relative difference between the luminance of the gray patches is relevant. This is mainly defined by the optical density of the patches.

² The Y-Channel is a weighted sum of the color channels R,G and B.



< The LE7 uniform light box with the TE269 OECF test chart >

The official ISP standard test procedure requires an adjustment of the exposure settings. As this cannot be done with many mobile phone cameras, the test target is captured under three different forced exposures where the operator will touch one region in the image before capture:

- Background** - the gray background, this should typically lead to a well-exposed image.
- Bright** - the brightest patch in the test target.
- Dark** - the darkest patch in the test target.

Per exposure ten images are captured, so in total 30 images are captured per light condition. Three light conditions are used, the same as described in section 2.3.1, bright, mid, and low.

3.8 *Shading*

We use the term *shading* for all kinds of effects that result in a change of intensity or color over the field. For a loss of intensity, this is mainly the result of *lens fall off*, also known as *relative illumination*. The TE42-LL test chart features a uniform gray background. Assuming no shading and perfectly uniform illumination, the background should be rendered to the same digital values over the entire

CONTENT

field. A loss in intensity over the field is called *Intensity Shading*, a shift in color over field is called *Color Shading*.

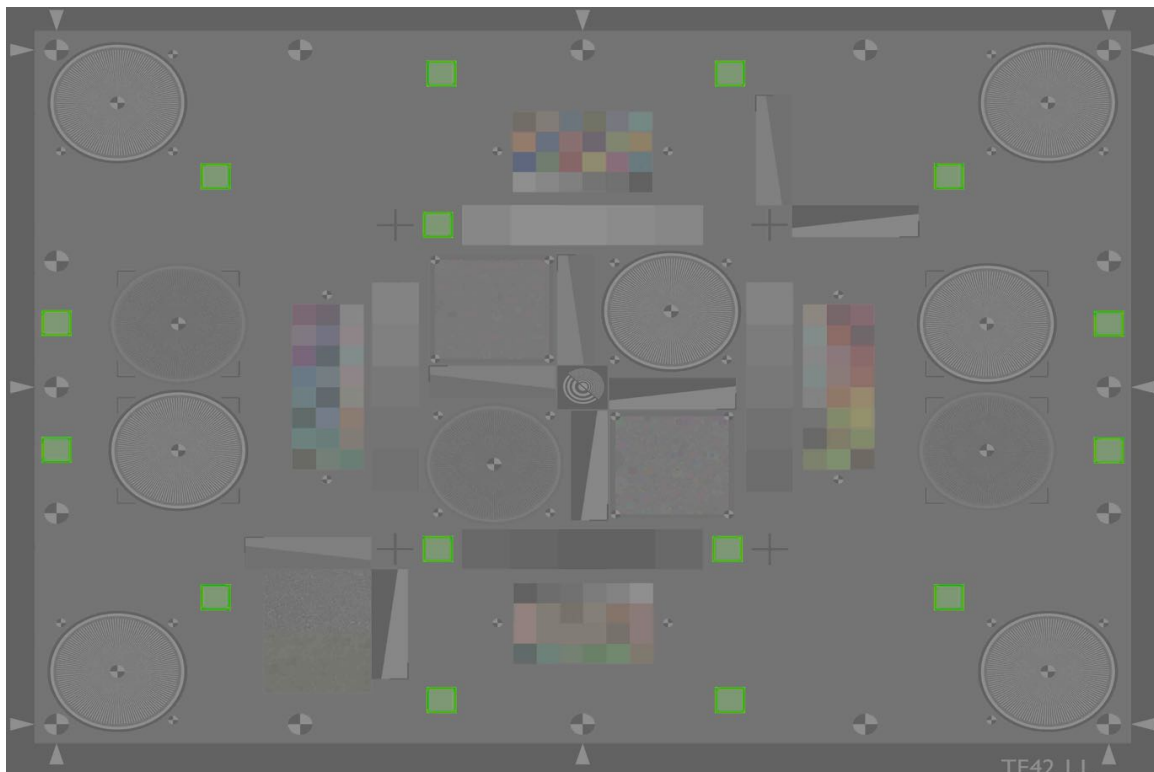
Even though the shading is mainly caused by optical components and should not vary with the illuminance, it is measured for all conditions. All devices perform some kind of shading correction. As this correction process can also increase noise and change color processing, the correction is not the same for different lighting conditions. So, it might happen that the measured shading does change for different lighting situations.

3.8.1 Intensity Shading

For the analysis of the intensity shading, the mean value of 23 ROIs in the Y channel is calculated. A possible non-uniformity in the illumination will result in an increased shading value and should, therefore, be avoided.

The intensity shading is reported in the unit f-stop, based on the difference between the maximum and the minimum equivalent luminance. The equivalent luminance is the corresponding luminance to the measured digital values, described by the measured OECF.

$$Shading_{intensity} = \frac{\log_{10}\left(\frac{L_{max}}{L_{min}}\right)}{\log_{10}(2)}$$



< The ROIs used for shading analysis >

3.8.2 Color Shading

Color Shading is a problem mainly known from mobile phone cameras. It describes an effect where the color changes over the field, resulting in different color shifts in the image center and the image corner. Due to the demand for very small devices, the chief ray angle (CRA) varies significantly on-axis and off-axis. With the CRA, the spectral sensitivity of the sensor and the transmission of other optical components (like the IR cut filter) varies and causes the color shifts.

The color shading is measured on the same patches as the intensity shading. For each patch the color difference (excluding the brightness L^*) is calculated as $\Delta E_{ab i}$

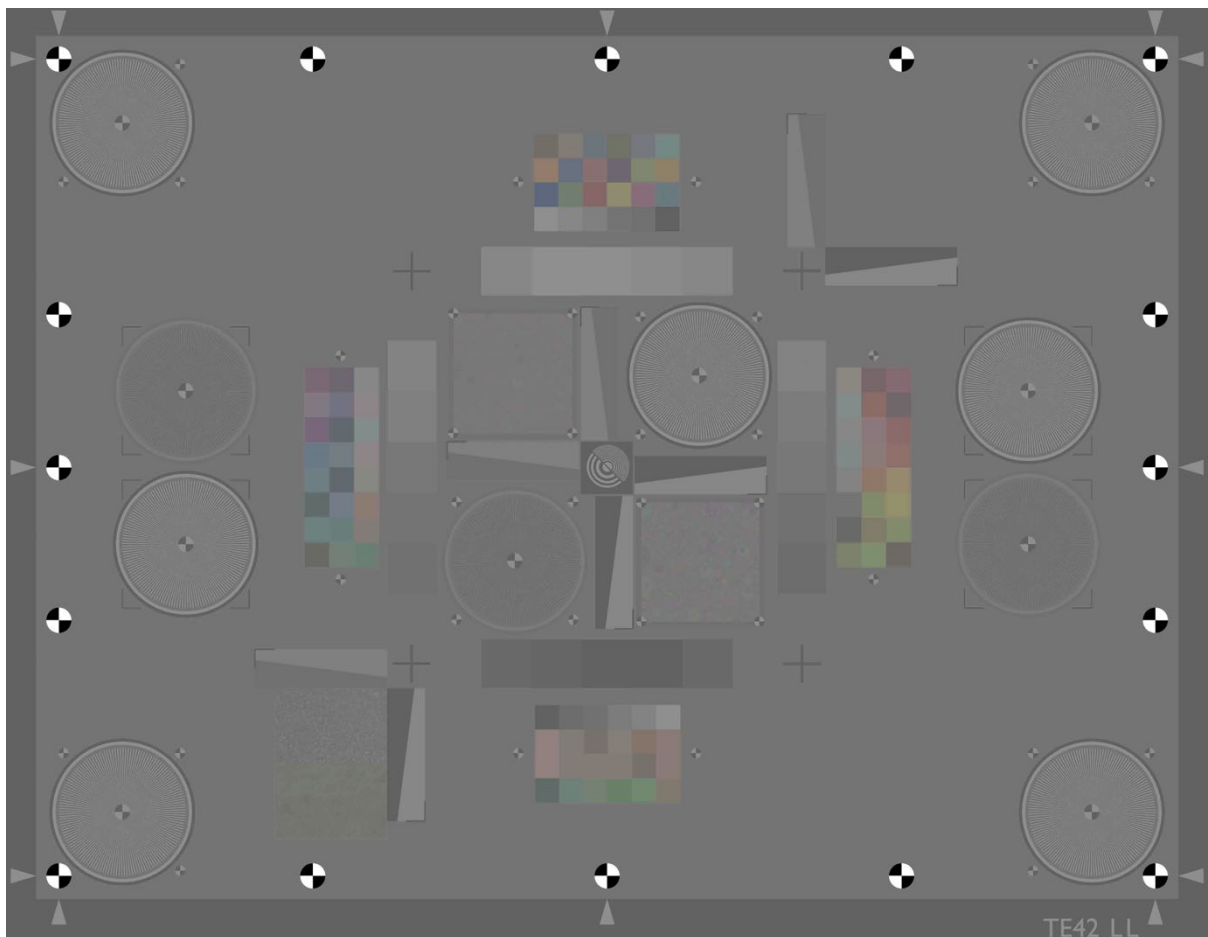
$$\Delta E_{ab i} = \sqrt{(a_i - a_{ref})^2 + (b_i - b_{ref})^2}$$

The reported value is the maximum found in all 23 patches. The reference values a_{ref} and b_{ref} are calculated as the mean value for a^* and b^* over all patches.

3.9 Distortion

Most lenses used for imaging show some amount of geometric distortion. This is caused by a varying reproduction scale over the field. So, the same object is either smaller or larger in the image depending on if it is in the image center or the image corner.

The distortion found in mobile phone lenses can be significant, but most devices apply a distortion correction in the ISP, so the distortion left in the image is not very objectionable.



< The position markers in the TE42-LL for the measurement of distortion >

3.9.1 TV Distortion

We measure the so-called TV Distortion (Definition of EBU). It is calculated using the height of an object in the image center (H) and the difference in height of the same object in the image corner (ΔH). These values are calculated based on the registration marks found in the TE42-LL chart and reported in %.

$$TV\ Distortion = \frac{\Delta H}{H} \times 100$$

The height (H) is the distance between the top center marker and the bottom center marker. The ΔH is calculated based on the difference between H and the average distance between the left and right marker (top and bottom).

3.10 Video test

To capture videos with the mobile phone is an important feature and used by many if not all users. In the VCX test protocol, we perform tests on video image quality and video performance.

3.10.1 Video Quality

The evaluation of Video quality is very similar to the measurement for still images. The device under test captures the test target, frames are extracted from the captured video and these frames are then analyzed for the KPIs also used for still images.

Due to the typical difference in aspect ratio, in most cases the camera has to be aligned to a target with 16:9 aspect ratio in video mode in contrast to the typically used 4:3 for still captures. A minor update in a future version might bring the introduction of a combined target that works for 4:3 and 16:9. For the main camera, video image quality is evaluated without zoom and for 3x zoom.

3.10.2 Video Performance

The current objective measurements of video performance metrics focus on autofocus speed and frame rate consistency. The autofocus speed is measured by using the timing kit as described before. The video capture for the device under test is started while it points to the test chart in far distance. After 10 seconds, the near focus chart moved up into the scene and then removed after another 10 seconds. This way we can measure based on resolution measurement the time it takes for the device to focus.

Another aspect is the frame rate and potential change depending on scene illumination. The video capture of the device under test is started and the scene (TE42-LL chart) is illuminated using the three main conditions (bright, mid, low illumination). Using the LED Panel in the TE42-LL, the frame rate can be checked constantly and any change in the frame rate can be detected.

3.11 Extended low light test

The main principles of extended low light testing follow the ISO standard 19093 (2018) where test images are captured using decreasing illumination levels. The standard defines certain image quality features which are calculated from each illumination level and corresponding threshold limits of each image quality feature. Whenever the threshold criteria are not met, the previous illumination level is recorded as the low light performance of the camera device in the test.

CONTENT

The test environment is the same as in the image quality tests defined in section 2.2. The test chart is TE42-LL and it is illuminated using two “iQ-Flatlights” LED light sources. The color temperature is set to D55.

Extended low light testing contains five different image quality features: exposure value, resolution, texture, noise, and color.

The following decreasing illumination sequence is used, and an image is captured for each illuminant:

- 2000 lux (reference image), 10 lux, 7.5 lux, 5 lux, 3 lux, 2 lux, and 1 lux.

The essential criteria of the test are the threshold values that define the low light performance of the device. For each of the five image quality features thresholds are defined. The reported low light performance is the last illumination level at which at least one of the thresholds is not met. So if a device has a low light performance of 2lux, that means that at 2lux all criteria were still acceptable and at 1lux at least one image quality feature no longer met the requirement.

3.11.1 Exposure

Exposure value is the Y-channel value of the 50% gray patch.

3.11.2 Resolution

MTF10 resolution value is calculated as defined in section 3.2.1. The center Siemens star is used for the calculation.

3.11.3 Texture

MTF10 texture value is calculated as defined in section 3.2.1.

3.11.4 Noise

Visual noise value is calculated as defined in section 3.6.

3.11.5 Color

Chrominance loss is calculated as defined in 3.5.1.7. The chrominance loss is the mean value of color chart patches located on top of the TE42-LL chart.

3.12 Response Time

Besides the pure image quality, the response time of the device under test has a significant influence on the user experience. We objectively measure different aspects of how fast the device reacts to user interaction. All aspects are measured according to ISO15781. A very important element is the LED-Panel, a device featuring 100 LEDs of which only one LED is illuminated at a time. With a known and

well-defined speed the device changes the illuminated LED. This way we have an external reference at which point in time the image was captured.

3.12.1 Frame Rate

The frame rate is normally reported as the number of frames that can be captured within one second. We decided to change this slightly and measure and report the time needed to capture 10 images.

This has two reasons:

- For very slow devices, the number is more intuitive
- Some devices are very fast in capturing two or three images but need a significant time after that to store the images to the permanent memory. Checking for ten images includes this possible delay.

If available, a *burst-mode* of the device is activated. It is checked beforehand if the pixel count is the same in *burst-mode* compared to *single-shot mode*. If we find a difference, the burst mode is ignored.

3.12.2 Shooting time lag

We measure the lag between action (pressing a button) and reaction (capturing an image). The faster the device, the better. The measurements are performed for the more challenging situations in Mid and Low light levels as described in section 2.3.1.

The procedure described is the one mentioned in ISO15781:2019. Mobile phone cameras normally work with a continuous AF, so the devices focus on the objects constantly and do not wait for the user to indicate that he wants to focus (e.g. pressing the release button halfway). In addition, the phone continuously shoots images into the buffer and the exposure button does actually not trigger an exposure anymore, but it selects the last image that has been added to the buffer when the exposure button is pressed. Another issue can arise if the device is using shutter priority, so it will capture an image, regardless if it is in focus or not. So it could happen that a device is measured with extremely short shutter release time lag, but all images are out of focus. To avoid these situations, a new test procedure has been developed since Version 2020.

The setup consists of a near-focus-chart and a far-focus-chart. The device under test will be given the time to focus on the near-focus-chart in order to have a controlled starting point for the autofocus, which is then removed quickly so the device can focus on the far-focus-chart. The near-focus chart is in 20cm distance from the device under test, the far-focus chart is in 1.2m distance. The activation of the timing device (LED Panel) is synchronized with other hardware that will press the release button and control the removal of the near-focus-chart.

CONTENT

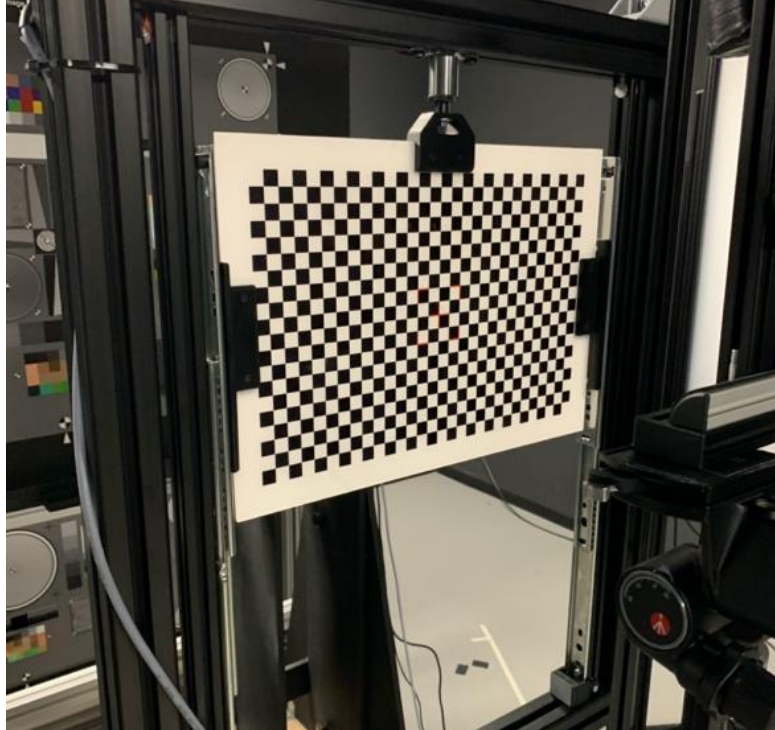
The shooting time lag is the time (in seconds) between pressing the release button and the start of the exposure with the time the camera needs to focus on the far-target. So, it includes the shutter release time lag, and the time needed to focus on the far-target. A release button can be a physical button on the device or a button on the user interface (touch screen).

With the setup described so far, the issue with continuous AF in mobile devices is solved, as we do not rely on a known pre-focus condition (as in ISO15781), nor do we need to modify the device in any way.

The time between removal of the near-focus chart and the event to press the release button can be controlled. So, if the device under test produces too many blurry images, we increase the delay between removing the near-focus chart and pressing the release button and check the images again. However, to minimize the test time this time is set to two levels, 200 ms and 1 s. To see the difference between the devices we look at the number of images that are in and out of focus.

In an ideal case, the whole measurement procedure is an iterative process to find the ideal delay which reflects the shutter release time lag. The delay is increased if more than 80% of the captured images are out-of-focus, the delay is decreased if all images are considered as in-focus. But as stated this measurement is too time-consuming for the daily work and therefore the 2 delay times have been selected.

The criteria of out-of-focus and in-focus are based on a relative SFR measurement, so if the SFR50 value for a measured image is below 50% of a reference image captured, it is considered as out-of-focus. This definition will reveal clearly blurry images, where slight variations in the focus will not change the detection.



< A small near-focus chart is positioned close to the camera to help it focus. It's held in place temporarily, then dropped out of view to allow focusing on a distant chart >

3.13 Motion Control

A digital camera can compensate for the lower illumination level at low light with either a higher gain (increased ISO speed) or a longer exposure time. When increasing the gain, the noise level increases with all the negative effects on the image quality of either noise or noise reduction.

Unlike dedicated camera systems, cameras in mobile phones are mainly used hand-held. So as the human tremor will shake the camera during exposure, the exposure time should not be too long, otherwise, the image will appear blurry due to motion blur. So, the manufacturer has to tune the camera in a way that the gain does not increase too much and at the same time, the exposure time does not get too long. This is only possible if the device is either very sensitive per default (fast lens, high sensor sensitivity) or can compensate the camera shake with an image stabilization (IS) system. As the user should not care what technology the manufacturer used, all devices regardless of their specifications (with or without an IS system) are treated the same way.

The device under test has to reproduce the TE42_LL test target (as used for image quality as well) so that the AF system can work properly, and the performance can be measured on the various test patterns. The chart is illuminated with the same lighting conditions as describes a "mid" and "low" for image quality analysis (250lux and 10lux). Note that these are called "bright" and "low" in this section".

CONTENT

During exposure, the device under test is mounted on a mechanical apparatus, called STEVE (STabilization EValuation Equipment) that can simulate the human tremor.

The following measurements are performed four times:

- Reference bright: The device captures the chart at 250 lux without being shaken
- Measurement bright: The device captures the chart at 250 lux while it is shaken by STEVE
- Reference low: The device captures the chart at 10 lux without being shaken
- Measurement low: The device captures the chart at 10 lux while it is shaken by STEVE

During every measurement with STEVE activated 50 images are captured and evaluated. The reported value is the average of the 50 images. The movement of STEVE is based on a publicly available human handshake trace that has been created based on a study of human handshake.

3.13.1 Visual Noise

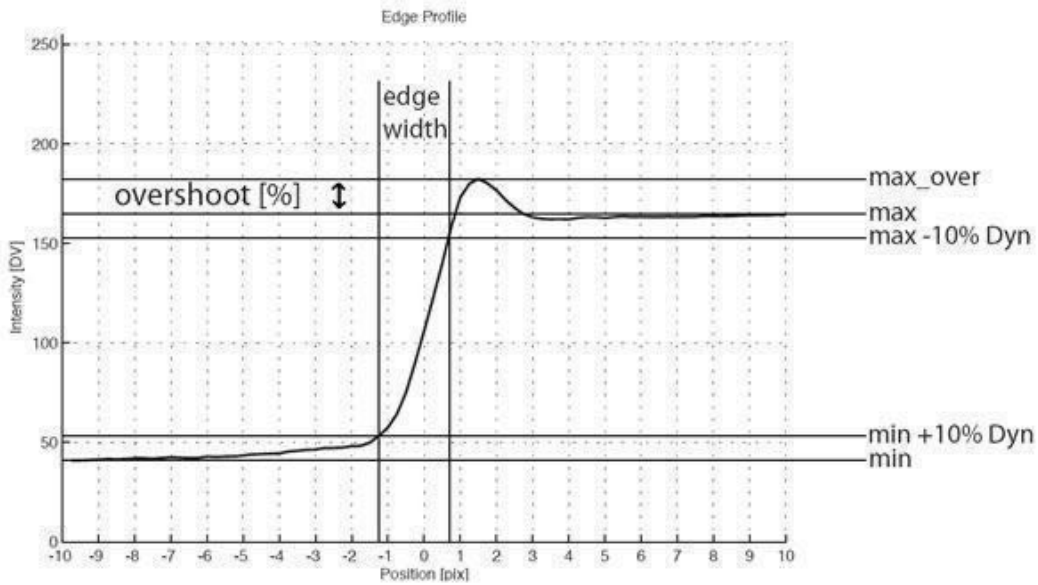
The device under test can potentially detect that it is not on a stable tripod (STEVE on) and change the exposure settings accordingly. As it must reduce the exposure time, it has to increase the gain setting, which results in higher noise.

We measure the Visual Noise (Viewing condition “100% view” (VC1)) on the available OECF patches of the test target (see “Visual Noise” in the section 3.6).

3.13.2 Edge Width

The used chart contains several slanted edges. Based on an e-SFR analysis and the ESF (see section 3.1.6), the edge width is calculated. It is defined as the “10% to 90% rise”, so the distance (in pixel) in the ESF where the intensity reaches 10% and 90% with the definition that the “min” equals 0% and “max” equals 100%. The edge width increases with an increasing motion blur, as the edge is washed out.

The absolute value and the delta between reference and measurement are reported.



< Edge width from an ESF >

3.13.3 Edge Acutance

Based on the same slanted edges, also the acutance as defined in section 3.1.7 is calculated. The absolute value and the delta between reference and measurement are reported.

3.13.4 Dead Leaves Acutance

Based on the high contrast dead leaves pattern, an SFR is computed and from that, an acutance value is calculated. In contrast to the edge acutance, this metric is less prone to sharpening and image enhancement processes and can show loss of details much better.

4 THE VCX SCORE IN THE REPORT

4.1 The Benchmark

The VCX score is derived from objective, numerical results. without any visual assessments or other subjective elements.

The only subjective aspect is the determining process of the relative importance of each metric in overall evaluation during the development. This weighting has been precisely determined by a group of experts from the VCX members and it is uniform across all devices, ensuring that comparisons between devices remain consistent and free from individual biases while the score is produced from a test.

The total score ranges from 0 to 100, with a score of 100 indicating that the device achieves the best possible results in every metric with current camera technology. The standard is expected to be updated every version, with additional updates made as necessary to incorporate significant advancements.

The VCX 2024 score is composed of two main parts: the Main Camera score (75%) and the Selfie Camera score (25%).

- The Main Camera score includes:
- Image Quality (60%), which is split into Still Image Quality (75%) and Video Quality (25%).
- Performance (15%), focused entirely on Still Performance.
- The Selfie Camera score (25%) is based on Image Quality, which is also divided into Still Image (75%) and Video (25%).

Image quality evaluations for both main and selfie cameras are conducted across three lighting conditions: Bright (31%), Medium (29%), and Low Light (41%).



CONTENT

Each metric used in the VCX score is calculated using bespoke algorithms or formulas specifically developed for VCX, based on use-case studies. The total score is a weighted sum of all these individual scores. The weighting of different image quality aspects is derived from case studies on mobile phone usage and internal research, and it aligns well with the results of other independent studies (see 5.4.1).

The metrics are transformed into scores based on theoretical worst and best values, with the scaling varying between these extremes depending on the metric. For some metrics, the correlation between the metric and its influence on image quality is linear, meaning the score is a linear function of the metric, following a "the higher the better" or "the lower the better" principle.

However, this linear assumption does not hold for all metrics. Some metrics require a different approach to accurately reflect perceived quality. Sharpening is an example: no sharpening results in flat images, while excessive sharpening leads to an artificial and unpleasant appearance. Thus, there is an optimal "sweet spot" for sharpening, with scores decreasing if the metric falls below or exceeds this optimal range.

To study this fundamental score correlation, the VCX Forum is conducting a Subjective Study Project. Please check the steps in the next chapter –5 Subjective study.

We continuously monitor the latest developments in the camera industry and update our test procedures and score generation process as necessary to reflect advancements in camera performance or to incorporate new technologies.

4.2 Report

The VCX score report is available in two official formats: a PDF report, which is tailored for marketing purposes and non-technical reviews, and a JSON format, which is intended for development and data management purposes. The JSON report is organized into three main sections:

- Introduction: This section provides general information about the test and includes benchmark scores for each test item.

```
{
  "VCX Version": "2024",
  "Test Date": "",
  "Device Name": "",
  "IMEI": "",
  "OS Version": "",
  "Firmware Version": "",
  "Model Number": "",
  "VCXScore": 53.76076361957615,
  "MainScores": {
    "Main Camera IQ": 58.30205014106626,
    "Main Camera IQ - Still IQ": 62.307077036040255,
    "Main Camera IQ - Still IQ - NoZoom": 68.92057783139947,
    "Main Camera IQ - Still IQ - Zoom4x": 37.59847169846594,
    "Main Camera IQ - Still IQ - UltraWide": 49.76872907720762,
    "Main Camera IQ - Video IQ": 46.28696945614429,
    "Main Camera IQ - Video IQ - NoZoom": 48.719416844315724,
    "Main Camera IQ - Video IQ - Zoom3x": -8.848504675741573,
    "Main Camera Performance": 41.316589419127155,
    "Main Camera Performance - Still Performance": 41.316589419127155,
    "Main Camera Performance - Still Performance - NoZoom": 45.75336264680837,
    "Main Camera Performance - Still Performance - Zoom4x": 19.654696601624764,
    "Selfie IQ": 50.328180488269325,
    "Selfie IQ - Still IQ": 50.276507179055415,
    "Selfie IQ - Video IQ": 50.48320041591104
  }
}
```

- Group Scores: This section presents the performance levels of each KPI, separate from the benchmark scores.

```
"GroupScores": {
  "Visual Noise": 0.9682857142857142,
  "Resolution": 0.31237945178613635,
  "Texture Loss": 0.32236182291666676,
  "Color": 0.9522773109243697,
  "Sharpening Artifact": -0.008225839120370393,
  "Shading": 0.9756166666666666,
  "White Balance": 0.6215384615384616,
  "Distortion": 1.0
}
```

- Body: This section contains detailed metric scores and related information for each KPI, which contribute to both the benchmark scores and group scores.

```
"Main Camera IQ - Still IQ - Zoom4x - Bright": {
  "Score": 44.0215304545319,
  "Filename": "device1_IQ_zoom_bright_d55_2000lx_04_42.xml",
  "Exif": "[ISO 125 - f/1.8 - 1/100s]*",
  "Groups": {
    "Highest (vc1_mean, vc2_mean)": {
      "MetricGroup": "Visual Noise",
      "MetricValue": 1.04,
      "MetricScore": 0.21452,
      "Weight": 0.217,
      "WeightedScore": 0.9885714285714285
    },
    "Highest (vc1_max, vc2_max)": {
      "MetricGroup": "Visual Noise",
      "MetricValue": 1.286,
      "MetricScore": 0.20571599999999998,
      "Weight": 0.217,
      "WeightedScore": 0.948
    },
    "Effective Pixel Count (EPC) overall": {
```

In the Body section, the following terms are defined as:

MetricGroup: Indicates the group to which the metric belongs.

MetricValue: Describes the specific metric.

MetricScore: Represents the outcome of the Look-Up Table (LUT).

Weight: Reflects the importance of the metric.

WeightedScore: How much does it contribute to the score score?

Each metric score is calculated according to the weights and algorithms established by the VCX Forum members and is then converted into the WeightedScore and benchmark score.

5 SUBJECTIVE STUDY

Subjective studies and public engagement play a crucial role in the development of VCX-Forum's objective image quality metrics. Subjective studies involve human observers evaluating image quality based on their perception. These Subjective studies provide valuable insights into how people perceive image quality, which is essential for understanding real-world user experiences. However, subjective results can depend on observers' physical conditions like mood and situation apart from factors like geographical and cultural influences, hence VCX-forum makes every effort to involve as diverse a populace as possible during such studies. Since subjective studies are also time consuming, a fair balance is maintained in terms of frequency, spread and metric revision.

Subjective studies help us identify scenarios where objective metrics could fail. for example, certain artifacts/noise may be considered as objectively disadvantageous, but by engaging with the public, we might get to know otherwise this allows our researchers to address these limitations and enhance the metrics by taking the subjective study results into account.

VCX-Forum uses subjective assessments to validate objective Image Quality Metrics. The subjective study project team within the forum has been conducting a survey to research the correlation between objectively generated benchmarks and human evaluations for each version. By comparing subjective ratings with computed metrics, our team tries to ensure that the objective measures align with human perception. This validation step helps refine and improve the accuracy of our metrics. While objective metrics are essential for automated systems like image compression and sharpening algorithms, subjective results are difficult to incorporate into real-time tests. However, by combining both approaches, we ensure that a comprehensive evaluation of image quality is achieved.

The correlation between subjective tastes and objective measurements can be complex and nonlinear, depending on the Key Performance Indicator (KPI) under study. This complexity significantly impacts the study by illustrating how laboratory measurements translate into real-life experiences.

Some aspects, such as color perception, are highly personal and influenced by factors like light temperature and color memory, which affect perceptions of what is good or right. Another important factor is the continuous evolution of trends, influenced by where you live, what you consume, your age, traditions, and other individual characteristics.

This summary highlights the major challenges of this type of study.

CONTENT

To address these challenges, a complex methodology was developed based on two groups of tasks:

- 1. Definition of study goals / gathering of data.**
- 2. Development of tools to organize and analyze the data.**
- 3. Definition of possible correlations and study outliers & compilation of the conclusions to feed the M2S algorithms.**
- 4. Analysis of the laboratory results and comparison with subjective trends.**

The first stage of the project can be compiled by the following steps:

1. Definition of the study goals - Refine the correlation of the end-user perspective and the objective results, by selecting KPI's to be analyzed.

2. Outline the expected results - Clearly outline the purpose of the study and anticipated outcomes from the gathered data (difficulties and possible conclusions).

3. Capture of real scenes - The same devices used at the laboratory were used to capture real-scene images - By this, we defined a dataset of images with challenging scenes meticulously developed to drive multiple KPI's under test.

4. Image/Question selection: Curate and selection of the images and question from the draft for inclusion in the research.

The images were selected to enable / simplify an evaluation of a specific KPI at once.

The questions were developed to be clear and succinct.

5. Review of expected answers: Evaluation of the anticipated responses to ensure the alignment of the outputs with the research purpose.

6. Research of a company to drive the survey: Identify and secure a suitable research partner.

7. Target selection: Determine the target audience in terms of size, region, segmentation, and survey duration.

The longer the survey time, higher the quality of the target, but on the other hand, with a longer survey we can have issues with less people joining the survey. So, there's a sweet spot of survey time and user's adherence.

8. Questionnaire review: Collaborate with the research company to review the questionnaire, including:

CONTENT

- Total response time and minimum/maximum response time per question.
- Randomizing and grouping the questions.
- Questionnaire screen background and research method (and the response channel).
- Question clarity, order and quantifiable responses. Are the questions intuitive and easy to draw conclusions? Are the responses quantifiable?
- Personal information and environment.

9. Sample creation and pilot test: Trial of the survey environment with a small group of VCX members.

10. Execution confirmation

11. Data collection and delivery: Implement data collection procedures and ensure timely delivery of results.

Evaluation by Expert Users: Conduct an internal survey with VCX members to evaluate outliers and understand unforeseen results. This survey compares expert evaluations with end-user survey responses, providing insights into how different types of users assess the same inputs. The results help clarify end-users understanding of various scenarios.

Compilation of Results and Final Subjective Correlations: Gather all the results obtained so far and summarize them using ranking graphs and other statistical tools. This helps address the main study questions. At the end of this stage, end-user preferences for each KPI are determined, guiding the importance and weight of these KPIs in the metric scoring system.

Comparison of Laboratory and End-User Outputs: Compare the subjective results from the laboratory dataset survey with those from the end-user survey to identify similarities and differences. This analysis reveals strong correlations in some KPIs and sparks discussions on potential explanations for differences in others.

6 APPENDIX

6.1 Overview of Numerical Results

6.1.1 Overview of All Image Quality Values per Lighting Condition

The values in the table below are chosen to create the VCX score and are measured for all five measurement/lighting conditions. For the details of the measurement conditions, see section 2.3. For the details on the values see the sections listed in “Reference”.

Group	Description
Dynamic Range	Dynamic Range (DR)
Visual Noise	Visual Noise 1 (VC1) mean
	Visual Noise 1 (VC1) max
	Visual Noise 2 (VC2) mean
	Visual Noise 2 (VC2) max
Resolution	Effective Pixel Count (EPC) overall
	s-SFR - Acutance overall
Texture Loss	Texture Loss MTF10 high contrast
	Texture Loss Acutance high contrast
	Texture Loss MTF10 low contrast
	Texture Loss Acutance low contrast
Sharpening Artefacts	Overshoot 2 (OS2) high contrast
	Undershoot 2 (US2) high contrast
	Overshoot 2 (OS2) low contrast
	Undershoot 2 (US2) low contrast
	Maximum SFR on deadleaves
	Maximum SFR on deadleaves
Color	dE - Color Error (skin tones)
	dL - Luminance Error (all)
	dC - Chrominance Error (all)
	dH - Hue Error (all)
White balance	White Balance
Shading	Intensity Shading
	Color Shading [ΔE_{ab}]
Distortion	TV-Distortion
Extended Low light	Exposure
	Resolution
	Texture
	Noise
	Color

CONTENT

Tartan SFR	Tartan SFR_center (Integral)
	Tartan SFR_corner (integral)

6.1.2 Overview of all Performance/Response Values

The values in the table below are chosen to evaluate performance and response for the VCX score. For the details of the measurement conditions, see section 3.12 and 3.13.

Group	Description
Frame Rate	Framerate for 10 Pictures
Response	Shooting Time Lag
AF performance	AF-Failure Rate (AFR)
Image Stabilization	STEVE on VN1
	Δ STEVE on VN1-STEVE off VN1
	STEVE on acutance Deadleaves
	Δ STEVE on acutance/STEVE off acutance
	STEVE on MTF10 Deadleaves
	Δ STEVE on MTF10/STEVE off MTF10

6.2 Color Patch Groups

As described in section 3.5.1, the color reproduction quality is reported for different groups of all used color patches.



- Red** A1 K1 L1 F2 K2 L2 F3 H3 K3 L3 K4 L4
- Green** G1 H2 E3 B7 L7 B8 C8 D8 E8 F8 G8 H8
- Blue** F1 E2 B3 C3 D3 A4 B4 J4 A5 B5 A6 B6 A7
- SkinTones** E1 J1 C6 D6 E6 F6 G6 H6 I6 C7 D7 E7 F7 G7 H7 I7

6.3 Changelog

6.3.1 Update history

6.3.1.1 Version 2025

1. Subjective study targets have been increased to 1,500 to reflect real end-user use cases and image evaluations more accurately.
2. Test items added

Selfie includes video, Main camera includes Ultrawide angle, Zoom 3x video, Zoom 4x performance, and removed flash test

3. Test metrics updated

Removed video EX, All test items include the metrics below, except some cases

: Visual Noise, Resolution, Texture Loss, Color, Sharpening, Shading. White Balance, Distortion
Corner blur for Ultrawide angle

4. Score structure changes

VCX score consists of 3 categories: selfie camera image quality, main camera image quality, main camera performance

Selfie camera includes still image test and video test

Main camera image quality includes still image(No zoom, Zoom 4x, Ultrawide angle) and video(no zoom and zoom 3x)

Main camera performance includes Zoom 4x as well as no zoom test.

For the visual explanation, please check the score tree

5. Score weight update

Subjective research for this version found that most image test ratings were largely consistent with those of previous versions. Specifically, based on the lighting conditions' use case, the results aligned well with VCX's previous version, showing no significant discrepancies. Most tests led to slight adjustments in some ratios to maintain overall balance due to adding new test items and emphasizing the importance of leading industry developments, such as video.

6. Modifications in Score Lookup Table Development (VCX 2025)

During the development of the VCX 2025 score lookup tables (LUTs), significant updates were introduced based on comparative analyses of real-scene images, controlled test scenes, and their alignment with subjective evaluation results. These changes reflect a broader structural overhaul and a shift toward end-user relevance. Below are the metric-specific revisions and rationales:

Resolution

The existing evaluation framework remained valid due to consistent performance across test cases. No structural modifications were required in VCX 2025. However, the evaluation has been expanded to include:

CONTENT

- Effective Pixel Count (EPC)-to-Theoretical Pixel Count (TPC) Ratio
- s-SFR-based Acutance

This ensures robust assessment of structural sharpness using realistic imaging parameters.

Texture Loss

Substantial updates were made to reflect real-world usage and prevent artificial score inflation:

- Evaluated under both high- and low-contrast conditions using data from Viewing condition Set 1 and Set 2
- Each Acutance sub-metric carries a weight as a large part.
- The LUT has been redefined to make the drop-off between “good” and “poor” images steeper, increasing scoring precision
- The evaluated MTF10 range was narrowed for greater mid-score differentiation

Visual Noise

The visual noise calculation was refined using updated VC1_mean and VC2_mean values. Rather than averaging metrics within a group, VCX 2025 selects the higher score (Mean or Max group), ensuring a more balanced representation of perceptual quality.

Sharpening → Sharpening Artifact

To avoid confusion in interpretation, the term was updated to Sharpening Artifact.

- The LUT was refined to better capture overshoot/undershoot phenomena
- Weighting was adjusted to reflect its contribution to perceptual sharpness
- User survey data helped define new thresholds, increasing relevance

Color

While color perception is highly subjective, color remains an essential visual indicator.

- The color LUT was updated to minimize undue influence on the overall score
- Quality within acceptable delta ranges is rewarded, with a low total weight of 4%
- The structure accommodates variability in user preferences while penalizing extreme deviations

White Balance

Evaluation thresholds were relaxed, particularly under low-light conditions:

- A separate LUT was implemented for low-light scenarios where perfect neutrality is less critical
- The deltaC tolerance was broadened in bright and mid-tones to reflect user tolerance
- This ensures better fairness without compromising technical rigor

Distortion

CONTENT

No changes were necessary. The current metrics remain intuitive and closely aligned with user perception.

Shading

The good-value range was expanded, as prior thresholds were deemed overly strict. This ensures a more forgiving and realistic evaluation of brightness uniformity.

7. Score tool

Fixed bugs

The function and value names have been modified to reflect the new config file and score structure.

8. WebCam profile has been included to VCX test procedure (please refer to WebCam spec document for the details)

6.3.1.2 Overview from Version 1.0(primary version) to 2020

In order to reflect the user experience better and to make the score of the different shooting condition comparable and more intuitive, the VCX score was revised and has been updated as of 22.12.2016. The update leads to a lower score for all devices compared to version 1.0. As technologies are constantly improving, this also enables keeping the VCX score valid for a longer period.

The main changes/differences to version 1.0 are:

- Changed weighting distribution of the main shooting conditions (bright, mid, low, flash, zoom).
- Clear separation between the main shooting conditions by taking “pinch zoom” out of the VCX score.
- Consistent weighting throughout the main shooting conditions in order to make the individual scores for each shooting condition more intuitive and comparable.
- Adjustment of HGC and LGC and consistent HGC and LGC throughout the main shooting conditions, except for resolution and texture related parameters in the shooting condition “zoom”. The reason is that most devices currently only offer digital zoom and therefore HGC and LGC are set lower for zoom.
- Revision of formulae, including introduction of formulae that lead to point deduction in case the result is way below user acceptance (resolution and texture related parameters only)
- Revision of parameters that are used to create the score
- Correction and adjustment of the whitepaper

The main changes/differences to version 2020 are:

- Tool: When VCX started as a small hobby project on a single PC, little did the team envisage that it would grow to be the industry standard that it is today. The weapon of choice at that time was Microsoft excel which donned the role of a prototyping and modeling tool with

macros and formulae forming the core logic. Although this worked well in a limited set up up to a point, given the complexity and inherent limitations, VCX-Forum has moved to a more open, flexible, expandable, and a robust open-source platform with a multi-team collaborative approach where experts from around the globe can participate. This makes VCX development accessible on all popular operating systems *“Python combines remarkable power with very clear syntax. It has interfaces to many system calls and libraries, as well as to various window systems, and is extensible in C or C++.”* Python being an open-source platform enables anyone who wants to contribute to the code of VCX to do so unencumbered by the burdens of licensing and unnecessary software fees. This might not seem to a big deal for the end customer, but strengthening the platform on which VCX is built is a very important milestone. This platform is planned to be expanded further to facilitate the integration of back-end database, website automation, membership & credential management apart from being secure and state-of-the-art

- Image Quality: updated metric to score with a new approach and formulae. Imaging technology and implementation in the smartphone arena have grown leaps and bounds from whence the VCX 1.5 was formulated. Customer’s expectation from their smartphone has also risen which meant a new approach to metrics and scoring. Over the last several years VCX-Forum has been the central hub where experts in the field have brought their experience back, participated in blind-tests, made recommendations, presented evidence and produced a new set to metrics and scoring that reflects current user expectation. More than 25 experts from all over the world participated in an exercise to finalize these metrics. There were no parameters in VCX V1.5 that were untouched; each and every one of them was relooked, reviewed and updated for Version 2020
- Improved: Dynamic Range: Given the vast improvement in sensor and display technology in capturing and displaying dynamic range, along with user expectation the methods of measuring the ability of a smartphone to capture dynamic range had to be updated. VCX 2020 improved upon the V1.5 by employing a high contrast back-lit target instead of a reflective TE42 chart. This updated methodology closely matches the user expectation and brings forth a more realistic measurement
- Low light performance: large pixels and innovative ways to capture light and reduce noise has given rise to a crop of top-shelf smartphones that capture images in low light that was hitherto the realm of large-bodied SLRs/Mirrorless cameras. The low-light tests and metrics incorporated in v1.5 were not only deemed deprecated to differentiate these top crop cameras but could not match with the evolving user expectation. In V2020, this has been addressed by a new gamut of extended low-light testing and appropriate metrics to reflect the same.
- Video: In the recent past many have used the video feature on their phones to start a revolution. VCX v2020 takes a step forward in that direction to add tests and metrics. As of this version, a single frame is analyzed for imaging metrics that could give the user adequate

guidance into the video performance of the phone. VCX-forum is committed to expanding this in future versions including dynamic lighting conditions and moving objects

- Selfie: use a single chart in size A460 in VCX v2020 the sophisticated tests and metrics have been carried over to the front-facing camera by introducing the same target and lighting conditions. This brings the score and rating on par and in harmony with the rear-facing camera, which was necessary, given the increase in priority of this feature since the last release, thanks to increasing popularity of social media platforms like Instagram and others
- Motion: A key element of any picture-taking process is image stabilization when capturing the picture to this end VCX has built on the latest research by Apple research which is now part of ISO standard #20954-2. This research goes in-depth on the kinesthetic of human hand movement and compensating for that as part of the image-stabilization apparatus in the smartphone. To make testing more effective additional metrics have been added which necessitated the addition of dead-leaves and Siemens-star modules to the test chart. A newly developed TE42-LL was employed apart from adding delta-Acutance Siemens and delta Acutance dead leaves as part of the metric ensemble. Further Optimization was done by removing the Edge width parameter and harmonizing the lighting with other measurements (250 lux + 10 lux)
- Autofocus failure metric has been revamped with automated procedures with no human interaction bringing the potential variation to the minimum and increasing the reliability of results as well as uniformity across labs.
- Shutter lag and Response time parameters have been deprecated as it was found that most modern smartphones register a negative shutter lag due to the use of predictive capture (before the shutter key is pressed)
- Lights used have been revamped for both rear and front camera image quality metrics taking stability, repeatability, and uniformity into consideration. Both rear and front camera testing involves the same lighting conditions giving truly comparative results. The spectral distribution for Bright light has been changed from daylight to white-LED keeping the lux values at the same level. For mid-light/in-door lighting has been changed to 250 lux cool white LEDs from the daylight spectrum. Low light testing sees a change to White LEDs with a color temperature of around 3000K and a lux level of 10 instead of the 64 lux in V1.5
- Score categories:
 - Image quality
 - Rear camera: Bright(default(DR)/video/zoom), Mid(default(DR)/video/zoom), Low(default(DR)/video/zoom/Flash/Extended low light), Zoom in bright, Zoom in mid, Zoom in low, Video in bright, Video in mid, Video in low
 - Front camera(Selfie): Bright, Mid, Low
 - Handling: Motion, Timing

6.4 References

6.4.1 Bibliography

Artmann, U. (n.d.). Image quality assessment using the dead leaves target: experience with the latest approach and further investigations. *Electronic Imaging Conference, Digital Photography XI. 9404*. San Francisco: SPIE.

Susan Farnand, Y. J. (2016). A methodology for perceptual image quality assessment of smartphone cameras. *IS&T International Symposium on Electronic Imaging 2016. Image Quality and System Performance XIII*.

Bucher, François-Xavier; Park, Jae Young; Partinen, Ari; Hubel, Paul. Issues reproducing handshake on mobile phone cameras. Electronic Imaging, Photography, Mobile, and Immersive Imaging 2019, pp. 586-1-586-7(7)

6.4.2 Links

[Paper EIC2015: Image quality assessment using the dead leaves target](#)

[Paper EIC2016: A methodology for perceptual image quality assessment of smartphone cameras; Farnand et.al.](#)

[Paper EIC2010: Differences of digital camera resolution metrology to describe noise reduction artifacts](#)

[The TE42_LL multipurpose test chart made by Image Engineering](#)

[The "AF-Box", made by Image Engineering](#)

[The Timing Kit, made by Image Engineering](#)

[STEVE - STabilization EValution Equipment, made by Image Engineering](#)