



Test Procedure – Whitepaper

VCX – Valued Camera eXperience

Version 2020

21 08 2020

CONTENT

Test Procedure – Whitepaper	1
VCX – Valued Camera eXperience.....	1
1 Introduction.....	5
2 Measurement.....	7
2.1 Approach.....	7
2.2 Procedure	8
2.2.1 Camera Settings	8
2.2.2 Measurement Setup	8
2.2.3 Used Test Charts	9
2.2.3.1 Illumination.....	10
2.2.3.2 Timing Kit.....	11
2.2.3.3 Near Focus Device	12
2.2.3.4 STEVE	12
2.3 Measurement Conditions	13
2.3.1 Main Camera (Rear camera)	13
2.3.2 Video (Rear camera)	13
2.3.3 Selfie Camera (Front camera)	14
2.3.4 Performance measurements	14
2.4 Viewing Conditions (VC).....	14
3 Metrics	15
3.1 Resolution	15
3.1.1 Theoretical Pixel Count (TPC)	15
3.1.2 Theoretical Maximum Resolution (TMR)	15
3.1.3 s-SFR - Limiting Resolution (LR) Center / Corner	16
3.1.4 Effective Pixel Count (EPC) Center / Overall.....	17
3.1.5 s-SFR - Acutance Center / Corner.....	18
3.1.6 e-SFR – MTF50	19
3.1.7 e-SFR – Acutance	20
3.2 Texture Loss (TL).....	21
3.2.1 Texture Loss MTF10 / MTF50	22
3.2.2 Texture Loss Acutance	22
3.2.3 Artifacts	22
3.2.4 Chrominance loss	23
3.3 Sharpening.....	24
3.3.1 Overshoot / Undershoot.....	24
3.3.2 Maximum SFR	25

3.4	Colors.....	26
3.4.1	Color Reproduction Error.....	26
3.4.1.1	ΔE - Color Error (all)	27
3.4.1.2	ΔE - Color Error (red).....	27
3.4.1.3	ΔE - Color Error (green).....	27
3.4.1.4	ΔE - Color Error (blue)	27
3.4.1.5	ΔE - Color Error (skin tones).....	27
3.4.1.6	ΔL - Luminance Error (all).....	27
3.4.1.7	ΔC - Chrominance Error (all)	28
3.4.1.8	ΔH - Hue Error (all)	28
3.4.2	White Balance (WB)	28
3.5	Visual Noise (VN).....	29
3.5.1	Visual Noise (mean)	29
3.5.2	Visual Noise (max).....	29
3.6	Dynamic Range (DR).....	30
3.7	Shading.....	32
3.7.1	Intensity Shading	32
3.7.2	Color Shading	32
3.8	Distortion.....	34
3.8.1	TV Distortion	34
3.9	Extended low light test.....	34
3.9.1	Exposure.....	35
3.9.2	Resolution	35
3.9.3	Texture.....	35
3.9.4	Noise.....	35
3.9.5	Color	35
3.10	Response Time.....	35
3.10.1	Frame Rate	36
3.10.2	Shooting time lag	36
3.11	Motion Control.....	38
3.11.1	Visual Noise	38
3.11.2	Edge Width	39
3.11.3	Edge Acutance	39
3.11.4	Dead Leaves Acutance.....	39
4	Calculating the VCX Score	40
5	Annex	42
5.1	Overview of Numerical Results.....	42
5.1.1	Overview of All Image Quality Values per Lighting Condition.....	42
5.1.2	Overview of all Performance/Response Values	43

5.2	Color Patch Groups	43
5.3	Changelog.....	44
5.3.1	Whitepaper	44
5.3.2	Metrics and Procedure.....	44
5.3.3	Score	44
5.3.4	Overview of Changes – Version 1.5(2019) to 2020	44
5.4	References.....	47
5.4.1	Bibliography	47
5.4.2	Links	47

1 INTRODUCTION

Due to the fast-evolving technologies and the ever-increasing importance of imaging in social media, the camera is one of the most important components of today's smart-phones. Smartphones have essentially replaced low-to-mid end compact cameras, thanks to better imaging apparatus, convenience, and connectivity.

With the vast choice of devices and great promises of manufacturers, there is a demand to characterize image quality and performance in very simple terms in order to provide information that helps users choose the best-suited device.

With a plethora of choices, exaggerated marketing claims, competing technologies, and techno mumbo-jumbo have created a confusing situation for customer planning for device acquisition; A market devoid of a transparent, open, and objective scoring system has exacerbated the situation.

For this very reason, VCX (Valued Camera eXperience) was established to address the challenges.

The mission of the association is the creation and dissemination of a standard for objectively assessing the image quality of cameras in mobile devices such as smartphones, tablets, computers, notebooks, drones, etc. The objective is to provide independent and credible information to consumers and industry alike; For experts who are in the know, an open and transparent system to build, comment and improve upon

The purpose of the white paper is to help understand the VCX standard and score system in detail.

VCX is based on 5 tenets which guarantee results that can be mapped to real-life experience

1. VCX measurements shall ensure out-of-the-box experience
2. VCX shall remain 100% objective
3. VCX shall be open and transparent
4. VCX shall employ/use an independent imaging lab for testing
5. VCX shall seek continuous improvement

Tenet 1. VCX measurements shall ensure out-of-the-box experience: This tenet dictates that the device under test shall ideally be obtained from an unbiased/untainted source i.e., a random sample/s from a store that sells the device under test. This ensures that neither special samples from suppliers nor custom hardware/software are accepted. The results are obtained from a device/devices that are launched onto the market. The device is tested using the default camera application and setting (except for flash/burst mode test cases).

Tenet 2. VCX shall remain 100% objective: The complete process on how the score is created from measurements is based on objective analysis of the device under test, followed by fixed and unbiased processing of the numerical results. No human interaction or subjective scoring is involved when creating the VCX score.

Tenet 3. VCX shall remain open and transparent: The VCX score can be accessed by anyone and not restricted to device vendors or mobile operators. The VCX score is designed to reflect the user experience with a mobile phone camera to make it much easier for end-users to decide on a new device. The VCX score is published on the website www.vcx-forum.org. This white paper details the entire testing and measurement procedure, which is open to critique and scrutiny by the imaging community at large. High-level weighting criteria are published along with the details of high-level components of what makes the final VCX score (performance/response and image quality in various lighting conditions).

Tenet 4. VCX shall employ/use an independent imaging lab for testing: VCX, as a quality improvement process has been adopted by various entities in the mobile imaging industry, but the final results are obtained from an independent trusted lab. Other independent imaging labs are welcome to join the VCX process, although the final results on the www.vcx-forum.org website, are restricted to be measured by trusted labs to ensure the highest quality and consistent results. Processes and procedures for the inclusion of other entities to actively contribute to VCX are underway.

Tenet 5. VCX shall seek continuous improvement: VCX has been developed over several years. Several vendors from the mobile device and chipset industry are now members and have already contributed positively to its improvement. And this cycle of feedback from customers and vendors shall continue.

2 MEASUREMENT

2.1 Approach

The final result of the applied test procedure is a single VCX score that shall reflect the user experience regarding the image quality and the performance of a camera in a mobile phone. To generate the score, image quality and performance are measured under different, controlled light conditions.

The VCX score is generated based on 100% objective data. The objective data is the result of well-defined and transparent test procedures following international standards where possible. Objective data means that at no point in the analysis process does a human observer make a judgment or decide on the performance. The entire analysis is only based on the captured images and the analysis algorithms applied to these images. Based on a fixed algorithm, the score is calculated using the numerical results. Also during the score generation, no subjective influence is applied. So in comparison to other metrics on the market, this metric is not modified based on the subjective impression of only a few people.

The image quality is evaluated for five different use cases (see test conditions in section 2.3) covering the most important aspects like:

- Spatial Resolution – What level of details can I see?
- Texture loss – How does the device reproduce low contrast, fine details?
- Sharpening - Does the device apply so much sharpening that disturbing artifacts occur?
- Noise – How much disturbing noise do I see?
- Dynamic range – What is the maximum contrast in a scene the device can reproduce?
- Color Reproduction – Are there any issues in color processing?

These questions are answered for each of the different viewing conditions, so it is transparent to the user, if the overall performance is bad or if a potentially low score device failed in one test condition only.

It is very important to see that VCX-Forum checks for all different aspects of image quality at the same time. With the currently available devices, we can see different ideas of camera tuning. Some devices show less noise than others but have poor performance in reproducing low contrast, fine details. Only with a strong and reliable setup of different algorithms, we can make sure that there are no easy loopholes in the camera tuning to get a high VCX score. As we could see in a lot of tests, the VCX system does not allow a camera to get a high score without also delivering a good performance. Additionally, the user experience is driven by the responsiveness of the device, so we check how fast it can focus and capture images.

As mobile phone cameras are devices that are used in any condition and mainly hand-held, low light performance is very important. We check how much the image quality changes in low light, depending on the movement of the device.

2.2 Procedure

Each device is checked before the test procedure starts for obvious mechanical problems like scratches on the lens or other signs of mechanical shock to the device. The lens is cleaned and the whole device rebooted to avoid any interference between other processes on the device.

For each measurement, several images are captured and the best image is selected. In case of the TE42_LL test setup (details in the following sections), four images are captured. The image with the highest resolution in the image center is then selected and fully analyzed. The camera is forced to re-focus after each captured image.

2.2.1 Camera Settings

We assume that most end-users do not make changes to the default settings. So, with every device, we perform a “factory reset” before the test and the settings remain “as is” during the test. The only change in the setting is the deactivation of “flash” and burst mode in case the device automatically turns it on during the measurement.

Even if some devices allow RAW shots these days or use a multicamera setup, the complete analysis is performed on the default images that are available after capturing an image. In most cases, this is a 24bit (8-bit per channel) image with JPEG compression. The color space of the images is respected and used. In most cases it is sRGB, but some devices use e.g. P3 Display color space.

Some devices allow manual adjustments like exposure compensation, ISO settings, etc., but these settings remain unchanged during the test.

The default aspect ratio of the image remains unchanged, so if the device uses 16:9 by default, the device is tested using the 16:9 aspect ratio. This also applies to the pixel count of the device. So, if the camera uses a different setting to the maximum pixel count by default, the pixel count remains unchanged.

2.2.2 Measurement Setup

The device under test is mounted on a tripod for all tests. Depending on the hardware and the position of the buttons, the best possible fixture on a tripod is used.

The tripod is mounted on a rail system, so the distance between the chart and camera can be changed easily. The distance between the camera and the test chart is selected so that the chart fits the entire frame in an optimal way. The framing is checked on an external PC screen before the final images are captured. The device is aligned to the chart in the best possible way to minimize roll and non-parallel setup between the chart plane and sensor plane.

The entire lab is temperature-controlled to standard room temperature ($23^{\circ}\text{C} \pm 2^{\circ}\text{C}$).

The tests are performed based on reflective test charts. These are mounted on a fixture to ensure proper alignment. The setup of chart and illumination follows a reprographic approach, which means that two light sources from left and right are used to illuminate the target in a uniform way and to ensure that no reflections on the test chart influence the measurements.

So, the reproduction scale is kept constant and the object distance is defined by the field of view of the device under test.

2.2.3 Used Test Charts

The device under test has to 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 a chart 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:

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

The first point is mainly relevant for the lab that performs the tests, but the second point is very important for the test procedure itself. Since the devices are tested in “fully automatic” mode and modern cameras use adaptive scene dependent image processing, this chart layout is the only way to ensure that the different aspects of image quality are measured under exactly 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 depending 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.

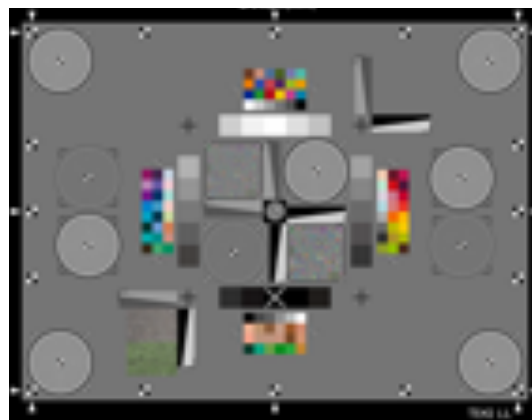


Figure 2-1: The TE42-LL multi-purpose test chart.

For the dynamic range test, a transmissive TE269B chart with 36 gray patches and a contrast of app. 1.000.000 : 1 is used in front of an integrating sphere in order to have a contrast range that exceeds the one that a camera can reproduce (see section 3.6) and figure 2-2.



Figure 2-2: The TE269 test target with the LE7 illumination device (integrating sphere)

2.2.4 Used Hardware

Apart from the test targets described in the previous section, additional hardware is also used. See links in Annex 5.4.2 for further details.

The test setup is checked before each test. Light intensity and light uniformity are checked with a regularly re-calibrated lux meter. The mounting device and the tripod are checked manually for potentially negative influence on the results.

In a monthly checking procedure, the spectral distribution of the light sources is verified.

2.2.3.1 Illumination

The TE42-LL test target is illuminated using two LED light sources. These light sources are “iQ-Flatlights” each use 20 LED-channels providing the capability to generate custom spectra for various testing parameters. These lights are particularly beneficial for VCX testing as they can accurately replicate all of the light conditions outlined in VCX version 2020. Each light is equipped with ten LED modules for maximum intensity of >2000 lx (for standard D illuminants set at a distance of 1.5 m from the test target). The lights also contain two fluorescent tubes with intensity control which are not used in the VCX test case.

An integrating sphere with similar technology as mentioned above is used for the dynamic range measurement. The Image Engineering LE7 device provides a tunable spectrum and high uniformity. See Figure 2-3 for an image.

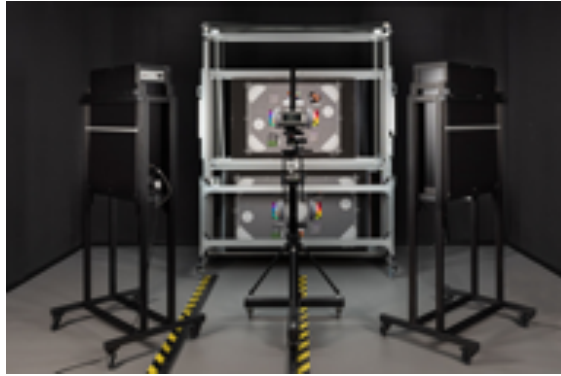


Figure 2-3: An ideal test setup using the LED light sources the "iQ-Flatlights."

2.2.3.2 Timing Kit

Two options are available when it comes to the timing measurement.. One is the Image Engineering timing kit which comes as a stand-alone working space with the AF-Box to illuminate a test chart and two LED Panels for the measurement.

In case the timing measurement shall be integrated into the main test stand the TE42-LL test chart is available in a special version called TE42-LL-T which is equipped with 2 LED Panels. So, this setup does not require a separate working space with an illumination box.

The main component in either hardware setup is the LED-Panel. This device features 100 LEDs, which can be illuminated one after another at a pre-selectable speed, thus acting as a high precision clock. Unlike an 8-segment clock or a mechanical clock with a pointer, the LED-Panel can accurately determine the time even if the device under test has a long exposure time. Long exposure times will only lead to multiple LEDs lighting up, but it is only the first LED that is of interest to a precise timing measurement. It is important to note that an automated solution is used to trigger the camera release button as a human finger might cause a slight delay when actuating the camera. The "iQ-Trigger" can trigger a mechanical release button and the "iQ-Trigger-T" can trigger a soft-button on a touch-sensitive display.

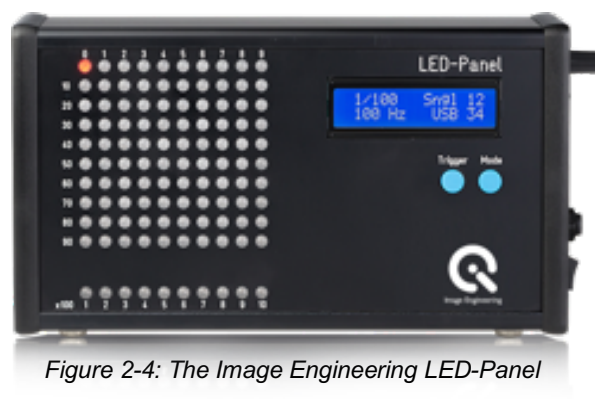


Figure 2-4: The Image Engineering LED-Panel

2.2.3.3 Near Focus Device

Precise timing measurements often require to set the automatic focusing system of the camera to a defined state. The Near Focus Device consists of a small test chart that can be placed into a defined short distance from the camera. The chart is mounted in a slider and held in its place by an electrical magnet. When the magnet is deactivated electronically the chart falls down and out of the optical path, allowing the camera to focus on the bigger test chart behind. That way the camera starts to focus in a defined state at a defined time. Various timing measurements such as shooting time lag can be accurately measured using this device.



Figure 2-5: The Near Focus Device

2.2.3.4 STEVE

"STEVE" is an abbreviation for "**ST**abilization **E**valuation **E**quipment". This device is a mechanical apparatus that can shake the device under test in a defined way. "STEVE" gives us the ability to simulate the motion of human tremors that are introduced while capturing an image. The device features six degrees of freedom (pitch, yaw, and roll, x, y, and z). When using this device, the camera under test has to reproduce a TE42-LL test target.

The performance of image stabilization in cell phones is difficult to measure because of the limited control over the camera. The camera decides which exposure speed or ISO setting it uses. Therefore, all important image quality parameters have to be monitored and measured (see upcoming ISO 20954-2).



Figure 2-6: Replicating human hand tremors using "STEVE".

2.3 Measurement Conditions

2.3.1 Main Camera (Rear camera)

The image quality evaluation based on the TE42-LL test chart is performed while the device is mounted on a tripod under five different measurement conditions:

Bright – This condition is used as the *Reference*. It is performed with a brightness of 2000lux Day light D55.

Mid – which results in an illumination level of 250lux neutral LED. This light condition reflects a normal indoor light situation without direct sunlight.

Low – A low light situation is the most challenging situation for a camera. In this case, we reduce the illumination to 10lux, warm LED.

Extended Low light – the device under test will capture images with decreasing illumination. A reference at bright (see above) and then the images with the illumination of 10lux, 7.5lux 5lux, 3lux, 2lux, 1lux.

Flash – If a situation is too dark, mobile phones mainly use LEDs to illuminate the scene. As a phone is rarely used in cases where there is absolutely no light, the flash is activated for this measurement while the scene is still illuminated with 10lux (Low).

Zoom – To zoom onto a smaller object is a very common action user of mobile phone cameras performs. To evaluate the image quality of a zoomed image, we capture an image of the TE42-LL with 4x zoom at the mentioned conditions bright, mid, and low. If the device provides an optical zoom, we use the optical zoom first and add digital zoom, if needed, to achieve a 4x zoom. For devices that only offer digital zoom, we use a 4x digital zoom. Devices that achieve less than 4x zoom will receive 0 points for this measurement condition. To perform the measurement, the distance between the chart and the camera is increased according to the zoom factor and the entire chart is captured.

2.3.2 Video (Rear camera)

The image quality evaluation of 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 extracted frames for analysis are extracted using FFmpeg after at least 5s in the video file.

This frame-based analysis is the first step in video analysis. A more complex video performance analysis will be addressed in future versions.

This is also performed under the three different measurement conditions:

Bright – This condition is used as the *Reference*. It is performed with a brightness of 2000lux Day light D55.

Mid – which results in an illumination level of 250lux neutral LED. This light condition reflects a normal indoor light situation without direct sunlight.

Low – A low light situation is the most challenging situation for a camera. In this case, we reduce the illumination to 10lux, warm LED.

2.3.3 Selfie Camera (Front camera)

The image quality evaluation based on the TE42-LL A460 test chart is performed under three different measurement conditions:

Bright – This condition is used as the *Reference*. It is performed with a brightness of 2000lux Day light D55.

Mid – which results in an illumination level of 250lux neutral LED. This light condition reflects a normal indoor light situation without direct sunlight.

Low – A low light situation is the most challenging situation for a camera. In this case, we reduce the illumination to 10lux, warm LED.

2.3.4 Performance measurements

For the performance measurements like Timing and Image Stabilization, the Illumination levels of 250 lux and 10 lux are used with a spectral distribution of D55.

2.4 Viewing Conditions (VC)

For some metrics, the viewing condition has to be defined. The main metrics are “Visual Noise” (see 3.5), the Acutance analysis of SFR measurement (see 3.1.5; 3.2.2; 3.1.7), and the assessment of Undershoot and Overshoot (see 3.3.1).

We make an assumption on the geometric setup (viewing distance, image size) how the user views the image. We try to cover the majority of real-life viewing conditions with three setups:

VC 1 – 100% view – This is the worst-case scenario, as the user can see most details. We assume a viewing distance of 0,5m and a 100% view on a 96ppi display. This means that each pixel of the image matches one pixel of the display. The more pixels the image contains, the larger it is displayed.

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

VC 3 – Large Print / PC Display – The complete image is scaled to a height of 40cm. The viewing distance equals the diagonal of the image, so it depends slightly 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 Annex 5.1 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 is able to 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 have to 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 ISO12233:2014 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 to reproduce, the more useful unit is line pairs per picture height (LP/PH). So, the theoretical maximum resolution expressed in LP/PH is calculated as:

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

The TMR is used for further calculations. A high number here does also not gain any benefit in the VCX score.

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

The limiting resolution describes the maximum level of details the device under test is able to reproduce. It is measured using the s-SFR method described in ISO12233:2014. Because the s-SFR method is based on a sinusoidal Siemens star, we use the Siemens stars implemented in the TE42-LL test chart.

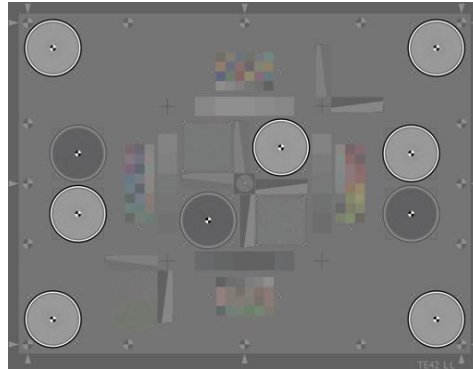


Figure 3-1: The sinusoidal Siemens stars within the TE42-LL test chart

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.

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 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.

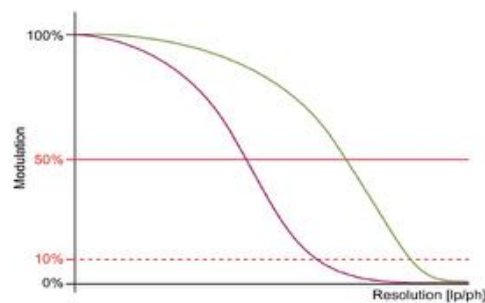


Figure 3-2 A sample SFR. Limiting resolution is the frequency at which the SFR crosses the dashed red line. With the green line representing the center of the image and the purple line representing the corners.

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.

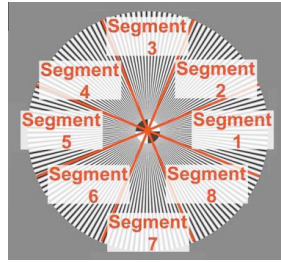


Figure 3-2: The eight segments per star

In case the limiting resolution of the device is extremely low (below the average user acceptance), this value and all depending values will lead to a score deduction.

An example of this can be a very low resolution in case a digital zoom is used. In this case, the resolution is significantly lower than the pixel count.

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. In case the limiting resolution of the segment is higher than the theoretical maximum, it is reduced to the theoretical limit. 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 theoretical win for a diagonal line can be up to 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

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

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 details the analyzed image contains. 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.

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

3.1.5 s-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 used CSF 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).

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

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

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

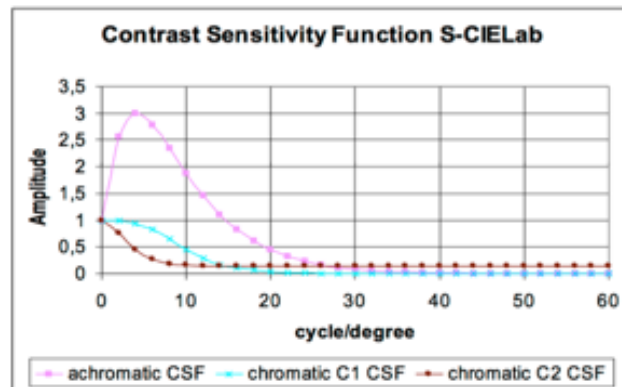


Figure 3-3: The Contrast Sensitivity Function used for the acutance calculation

3.1.6 e-SFR – MTF50

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 it could be obtained from a single line perpendicular to the edge.

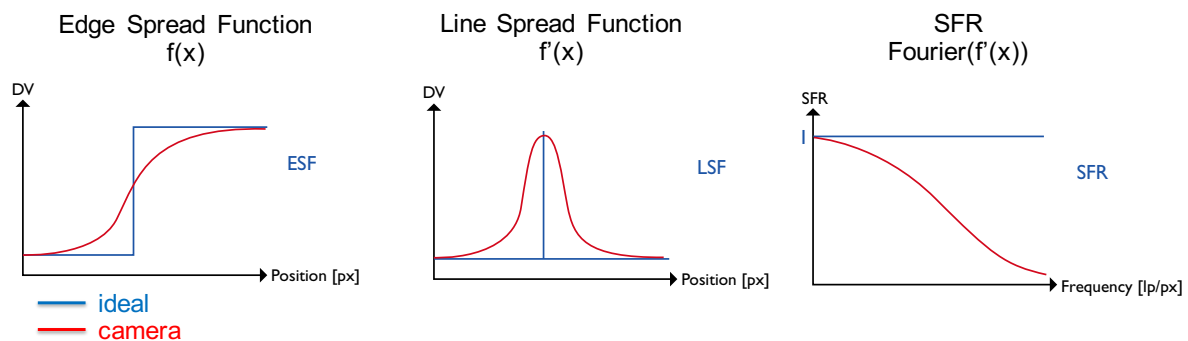


Figure 3-4: The core analysis steps of the e-SFR method.

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 depending 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 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.

The reported MTF50 value is the spatial frequency (in LP/PH) which leads to a spatial frequency response of 50%.

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

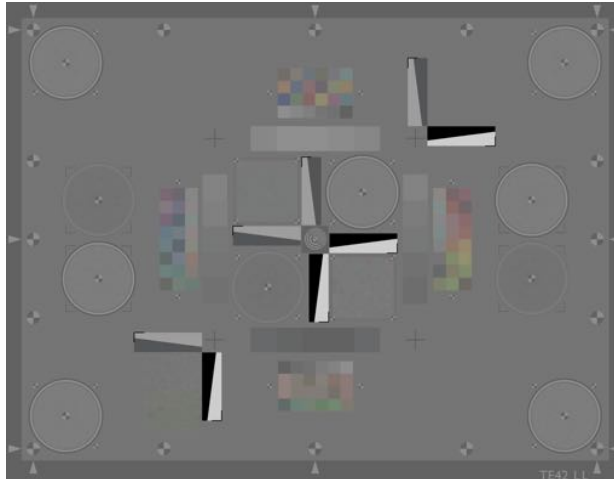


Figure 3-5: The slanted edges within the TE42-LL test chart

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 content are performed depending on the image content. A very common and important optimization step is noise reduction. In this case, the ISP analyses 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 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) structure. The dead leaves structure is a random pattern build by several thousand circles stacked on top of each other. The location and the 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.

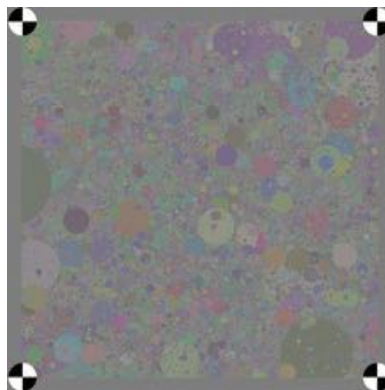


Figure 3-6: The dead leaves pattern (low contrast)

The dead leaves pattern is well known in the industry and different analysis methods have been proposed over the past years. The latest one, developed by Image Engineering, 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}*

3.2.1 Texture Loss MTF10 / MTF50

The SFR obtained from the Deadleaves_{cross} is the root function from which an MTF10 and an MTF50 value are derived. The MTF10 value reflects the level of detail that the device under test can reproduce if the structure contains low contrast, fine details. The MTF50 value is similar to the e-SFR at lower frequencies, which correlates with the perceived sharpness.

3.2.2 Texture Loss Acutance

As we have an SFR, we can also calculate the acutance for different viewing conditions. That way the perception of a human observer for the given viewing condition can be quantified.

The calculation is identical to the method described in section 3.1.5

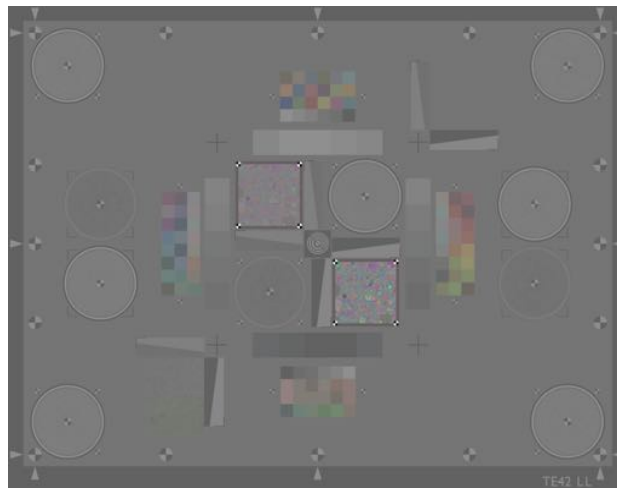


Figure 3-7: Two different versions of the dead leaves pattern within the TE42-LL test chart.

3.2.3 Artifacts

The Artifacts metric shall describe the amount of added image content by the camera. This added image content is mainly considered as artifacts. It might be sharpening artifacts (overshoot/undershoot and ringing) along edges and that part of the noise that is only shown on structured areas, not on flat uniform areas of the image.

It is calculated based on the texture loss acutance derived from Deadleaves_{cross} and Deadleaves_{direct}. While the Deadleaves_{cross} algorithm by design is insensitive to added image content, the SFR based on the Deadleaves_{direct} algorithm is influenced by artifacts and noise. So, the difference between these two SFR is caused by these image distortions.

$$Artifacts = 100 - \left(\frac{Acutance(DL_{cross}, all)}{Acutance(DL_{direct}, all)} \times 100 \right)$$

The Acutance (DLcross, all) in the equation above is the acutance calculated based on the SFR derived from the Deadleaves_{cross} algorithm. In this case, there is no weighting in the viewing condition, so the contrast sensitivity function (CSF) is assumed as a perfect perception of all frequencies. For further information about the differences between the algorithms, see the mentioned paper in the bibliography.

3.2.4 Chrominance loss

Camera noise can be seen as luminance noise and chroma noise. Chroma noise appears as color artifacts in a neutral area. To reduce the visibility of chroma noise, some cameras tend to reduce the saturation in areas that are considered as noise.

This effect can be seen as a loss of saturation when reducing the illumination level. The chrominance loss is reported as a loss between the reference (condition “bright” as describes in 2.3.1) and the other conditions.

For this analysis, the RGB data of each pixel within the dead leaves structure used for the texture loss analysis is converted to CIE-XYZ with the corresponding ICC color profile (in most cases sRGB) and then converted into CIE-L*a*b* color space. The CIE-L*a*b* data is then converted into CIE-L*C*H* data. For each lighting condition, the reported C* is the average C* value over all pixels. This value is calculated for both dead leaves structures separately.

The reported ΔC^* equals the ratio in % between the measurement condition and bright condition, so by definition, it is 100% for bright.

$$\Delta C^*_{condition} = \frac{C^*_{condition}}{C^*_{bright}} \times 100$$

3.3 Sharpening

The subjective impression “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 intended, 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 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 can be followed by “ringing”.

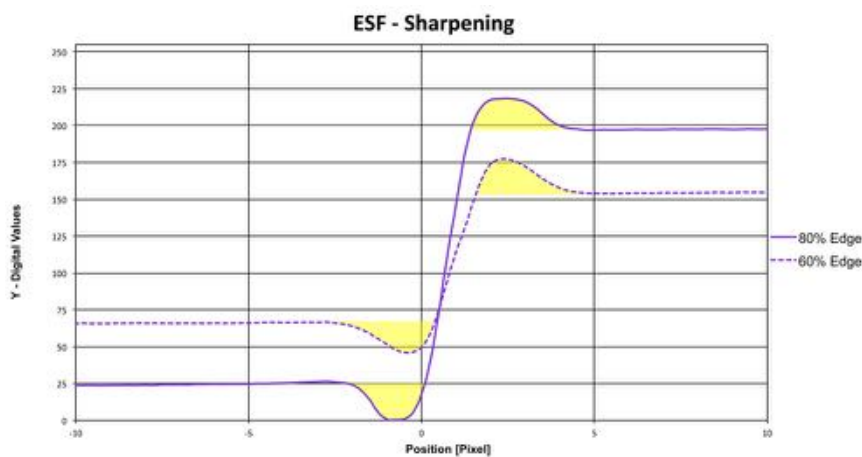


Figure 3-8: Undershoot and Overshoot in the ESF (yellow area).

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, respectively overshoot is expressed as the integral of that area beneath or above the ESF and are marked yellow in Figure 3-8.

The edge consists of a bright and dark region. The overshoot is the area that is brighter than the bright region of the edge, accordingly 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 we obtain an SFR from, so Siemens star, slanted edges, and dead leaves pattern.

3.4 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.

For this test, the color patches of the TE42-LL are used. These patches are chosen with respect to the well-known X-Rite ColorChecker 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.

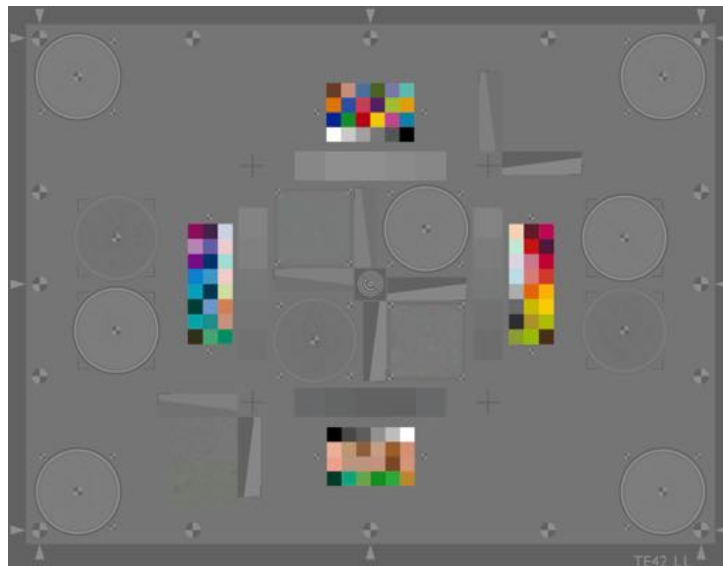


Figure 3-9: The color patches of the TE42-LL test chart.

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.

3.4.1 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.

3.4.1.1 ΔE - Color Error (all)

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

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

The reported value is the average over all 96 patches.

3.4.1.2 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.

3.4.1.3 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.

3.4.1.4 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.

3.4.1.5 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.

3.4.1.6 ΔL - Luminance Error (all)

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

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

3.4.1.7 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.

3.4.1.8 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.

3.4.2 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 adoption to different light has to be performed as a separate step in the color processing. A camera has to estimate what the white point of the current scene illumination is and perform a white balance according to this information.

The difficult part for a camera is to always make a good and correct estimation of 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 3.5). 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.

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.

3.5 Visual Noise (VN)

Noise is defined as the fluctuation of measurement due to random processes. As each pixel is basically a measurement device of the object luminance, noise introduces random variations of 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.

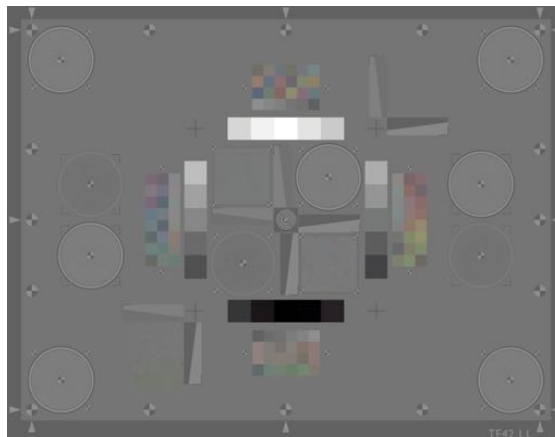


Figure 3-10: The gray patches of the TE42 test chart; used for the noise measurement.

3.5.1 Visual Noise (mean)

The defined viewing conditions for the calculation are described in section 2.4. The contrast sensitivity function (CSF) used is described in ISO15739:2013 and also shown in Figure 3-.

The visual noise value is calculated for all 20 gray patches shown in Figure 3-. 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.

3.5.2 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.

3.6 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.

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

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.



Figure 3-11: 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.

3.7 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 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.

3.7.1 Intensity Shading

For the analysis of the intensity shading, the mean value of 23 ROIs in the Y channel is calculated (see Figure 3-). 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.

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

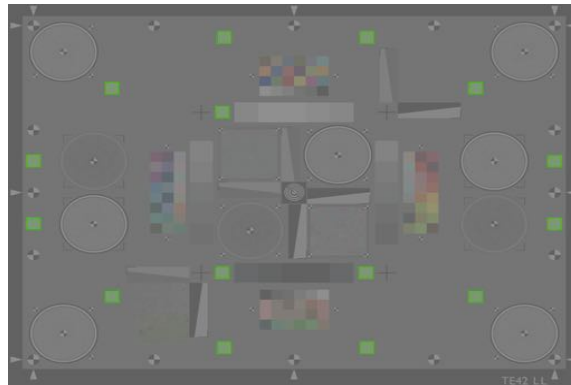


Figure 3-11: The ROIs used for shading analysis.

3.7.2 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.

3.8 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 disturbing.

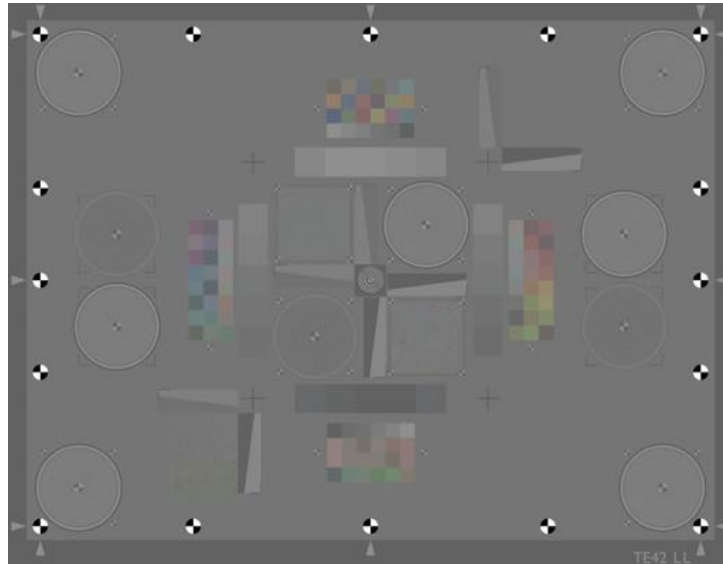


Figure 3-12: The position markers in the TE42-LL for the measurement of distortion.

3.8.1 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 (see Figure 3-) 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).

3.9 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 some threshold criteria do not fulfill, the previous illumination level is recorded as the low light performance of the camera device in the test.

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

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

Following decreasing illumination sequence is used and an image is captured in 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 illumination level, where all 5 thresholds are not reached. 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 did not meet the requirement anymore.

3.9.1 Exposure

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

3.9.2 Resolution

MTF10 resolution value is calculated as defined in chapter 3.1.6. The middlemost Siemens star is used to the calculation.

3.9.3 Texture

MTF10 texture value is calculated as defined in chapter 3.2.1.

3.9.4 Noise

Visual noise value is calculated as defined in chapter 3.5.1.

3.9.5 Color

Chrominance loss is calculated as defined in 3.2.4. The chrominance loss is the mean value of Macbeth color chart patches located on top of the TE42-LL chart.

3.10 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.

3.10.1 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.

3.10.2 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 Mid and Low 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 for 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.

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, slight variations in the focus will not change the detection.

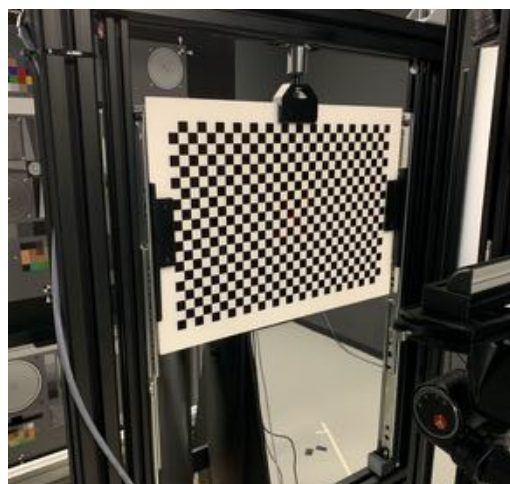


Figure 3-13: The near focus chart is a relative small chart placed close to the camera so that the camera can focus on it. A mechanism holds the chart in the optical path. When released the chart falls down and out of the optical path so that the camera can focus on the chart further away.

3.11 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 more or less 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”. 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:

1. Reference bright: The device captures the chart at 250 lux without being shaken
2. Measurement bright: The device captures the chart at 250 lux while it is shaken by STEVE
3. Reference low: The device captures the chart at 10 lux without being shaken
4. 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

3.11.1 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 has to 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 image quality section).

3.11.2 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% (see Figure 3-). The edge width increases with an increasing motion blur, as the edge is washed out.

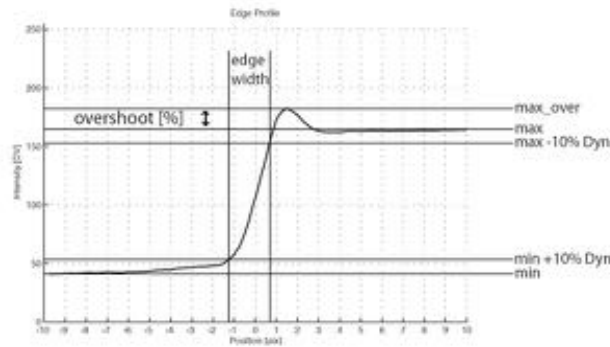


Figure 3-13: Edge width from an ESF

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

3.11.3 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.

3.11.4 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 CALCULATING THE VCX SCORE

The VCX score is calculated based on objective, numerical results. It does not contain any visual assessment or other subjective components.

The only subjective component is the weighting, so the decision which metric is more important for the overall performance compared to others. But this weighting has been accurately determined by a group of experts from the VCX members and it is identical for every device, so the comparison between devices is fixed and not influenced by individual opinions.

The total score range is between 0 and 100. The range is designed in a way, that a value of 100 means, that the device meets the best possible result in every metric that is achievable with today's camera technology. An update to the standard is expected on a yearly basis. Whenever a crucial update to the standard demands, it can be added.

The VCX score is the sum of the main camera score (80%) and the selfie camera score (20%). The main camera score consists of Image Quality (60%), Performance (25%), and Video (15%) score. Both Main camera and selfie camera consist of Image Quality score in three illuminants: Bright (36%), Medium (29%), and Low (35%).

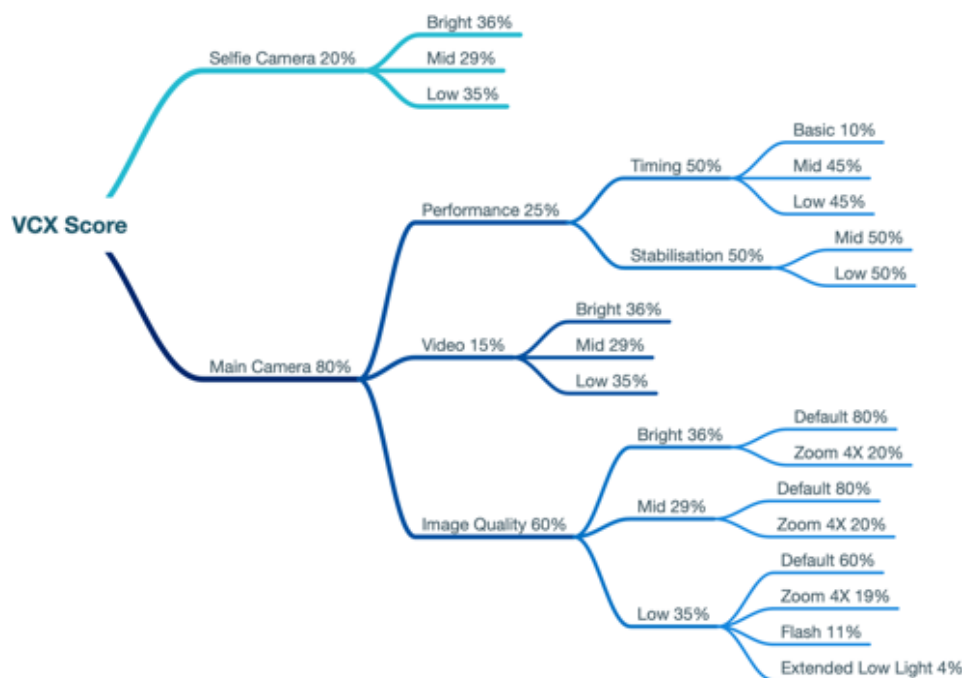


Figure 4-1: Weighting of the different test conditions for the total score.

For each of the used metrics, a single score is calculated via bespoke algorithms/formulae developed specifically for VCX derived from use-case studies. The total score is a weighted sum of all scores.

The weighting of the different aspects of image quality is the result of a case study on how mobile phones are used as well as internal research and studies. It correlates well with the outcome of other independent studies (see 5.4.1).

The transformation of metrics into scores is performed under the definition of a theoretical worst and theoretical best value. The scaling is performed in different ways between the extreme points, depending on the metric itself. For some metrics, the correlation between “metric” and “influence in image quality” is linear, so the score is a linear function of the metric. This would be in the case of a simple “the higher the better” or “the lower the better” assumption.

For others, this assumption is not true. Some metrics require a different approach to the one previously mentioned because it would not reflect the perceived quality. Sharpening is a good example of this behavior. No sharpening is not beneficial for the image quality, as an image would appear flat. But at the same time, a very high sharpening very quickly results in an artificial and unpleasant look of the image. So there is a “sweet spot” and below or above this leads to a reduction in the score.

We regularly check the latest development in the camera industry and will update the test procedure and the score generation process as soon as we see that it does not reflect the improvements in camera performance or if new technologies need to be included in the procedure.

5 ANNEX

5.1 Overview of Numerical Results

5.1.1 Overview of All Image Quality Values per Lighting Condition

The values in the table below are were chosen to create the VCX score and are measured for all five measurement/lighting conditions. For the details of the measurement conditions, see Chapter 2.2.3.4. For the details on the values see the sections listed in “Reference”.

Group	UID	Description
Dynamic Range	OECF-DR	Dynamic Range (DR)
Visual Noise	VN-Display	Visual Noise 1 (VC1) mean
	VN-Max-Display	Visual Noise 1 (VC1) max
	VN2	Visual Noise 2 (VC2) mean
	VN2-Max	Visual Noise 2 (VC2) max
Resolution	Res-EPC_Overall	Effective Pixel Count (EPC) overall
	Res-Contrast_Overall	s-SFR - Acutance overall
Texture Loss	TL_DL_HC_MTF10	Texture Loss MTF10 high contrast
	TL_DL_HC_vMTF1	Texture Loss Acutance high contrast
	TL_DL_LC_MTF10	Texture Loss MTF10 low contrast
	TL_DL_LC_vMTF1	Texture Loss Acutance low contrast
	TL_DL_Artifacts_HC	Artifacts high contrast
	TL_DL_Artifacts_LC	Artifacts low contrast
	TL_DL_C-Star_HC	Chrominance (C*) high contrast
	TL_DL_C-Star_LC	Chrominance (C*) low contrast
Sharpening	Edge_Data_HC_OverShoot_A2	Overshoot 2 (OS2) high contrast
	Edge_Data_HC_UnderShoot_A2	Undershoot 2 (US2) high contrast
	Edge_Data_LC_OverShoot_A2	Overshoot 2 (OS2) low contrast
	Edge_Data_LC_UnderShoot_A2	Undershoot 2 (US2) low contrast
	SFR_MAX_HC	Maximum SFR on deadleaves
	SFR_MAX_LC	Maximum SFR on deadleaves
Color	DE-Skin	ΔE - Color Error (skin tones)
	DE-Luminance	ΔL - Luminance Error (all)
	DE-Chrominance	ΔC - Chrominance Error (all)
	DE-Colourtone	ΔH - Hue Error (all)
White balance	Shading-WB	White Balance
Shading	Shading-IS_Fstop	Intensity Shading
	Shading-CS_Deab	Color Shading [ΔE_{ab}]
Distortion	Distortion-TV	TV-Distortion
Extended Low light	Y value of neutral patch #17	Exposure
	MTF10 center star	Resolution
	MTF10 of low contrast DL	Texture
	VN1 Mean	Noise
	delta C % compared to Bright image deltaC	Color

5.1.2 Overview of all Performance/Response Values

The values in the table below are were chosen to evaluate performance and response for the VCX score. For the details of the measurement conditions, see Chapter 3.9 and 3.10.

Group	UID	Description
Frame Rate	FrameRate-10pics	Framerate for 10 Pictures
Response	Shooting_Time_Lag	Shooting Time Lag
AF performance	AF_Failure_rate	AF-Failure Rate (AFR)
Image Stabilization	STEVE_ON-VN	STEVE on VN1
	STEVE_ON-DeltaVN	Δ STEVE on VN1-STEVE off VN1
	STEVE_ON-Contrast	STEVE on acutance Deadleaves
	STEVE_ON-DeltaContrast	Δ STEVE on acutance/STEVE off acutance
	STEVE_ON-MTF10 Deadleaves	STEVE on MTF10 Deadleaves
	STEVE_ON-DeltaMTF10	Δ STEVE on MTF10/STEVE off MTF10

5.2 Color Patch Groups

As described in section 3.4.1, the color reproduction quality is reported for different groups of all used color patches.

1												
2												
3												
4												
5												
6												
7												
8												
	A	B	C	D	E	F	G	H	I	J	K	L

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

5.3 Changelog

5.3.1 Whitepaper

Version	Date	Author	Comment
1.0	19.09.2016	Image Engineering	Initial Document
1.5	21.12.2016	Image Engineering	Update
2020	20.08.2020	Uwe Artmann(IE) Petteri Valve(Nomicam) Benjamin Pak(VCX) Vijay Kishan Rao (Vodafone)	Update for version 2020, details described 5.3.4

5.3.2 Metrics and Procedure

Version	Date	Author	Changes
1.0	19.09.2016	Uwe Artmann (IE)	First public Version
2020	20.08.20	Uwe Artmann(IE) Petteri Valve(Nomicam) Benjamin Pak(VCX) Veli Tapani Peltoketo(Huawei)	Update for version 2020, details described 5.3.4

5.3.3 Score

Version	Date	Author	Changes
1.0	19.09.2016	Vijay Kishan Rao (Vodafone)	First public Version
1.5	15.12.2016	Vijay Kishan Rao (Vodafone) and Image Engineering	Second public Version
2020	20.08.20	Petteri Valve(Nomicam) Benjamin Pak(VCX)	Update for version 2020, details described 5.3.4

5.3.4 Overview of Changes – Version 1.5(2019) 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 of time.

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 upto 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 ensample. 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 day-light 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

5.4 References

5.4.1 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)

5.4.2 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 multi purpose test chart made by Image Engineering

The "AF-Box", made by Image Engineering

The Timing Kit, made by Image Engineering

STEVE - STabilization EValuation Equipment, made by Image Engineering