Autofocus measurement for imaging devices

Pierre Robisson, Jean-Benoit Jourdain, Wolf Hauser
Clément Viard, Frédéric Guichard

DxOMark Image Labs, 3 rue Nationale
92100 Boulogne-Billancourt FRANCE

Abstract

We propose an objective measurement protocol to evaluate the autofocus performance of a digital still camera. As most pictures today are taken with smartphones, we have designed the first implementation of this protocol for devices with touchscreen trigger. The lab evaluation must match with the users' real-world experience. Users expect to have an autofocus that is both accurate and fast, so that every picture from their smartphone is sharp and captured precisely when they press the shutter button. There is a strong need for an objective measurement to help users choose the best device for their usage and to help camera manufacturers quantify their performance and benchmark different technologies.

Keywords: *Image quality evaluation, autofocus speed, autofocus irregularity, acutance, shooting time lag, smartphone*

Introduction

Context and motivation

The primary goal of autofocus (AF) is to ensure that every single picture taken by the user has the best possible sharpness regardless of subject distance. This AF accuracy is very important for a digital camera because blurry pictures are unusable, regardless of other image quality characteristics. Defocus cannot be recovered in post-processing. Image quality assessment must therefore take AF into account along with other attributes such as exposure, color and texture preservation. The secondary goal is to converge as fast as possible, so that the picture is taken exactly when the user hits the shutter button. Camera manufacturers might have to make trade-offs between accuracy and speed.

A camera is in focus when all optical rays coming from the same object point reach the sensor at the same point in the image plane. For an object at infinity, this is the case when the lens is placed at its focal length from the sensor. For objects closer than infinity, the lens must be moved further away from the sensor. In most smartphones this motion is done using a voice coil motor (VCM) [1]. The biggest challenge and differentiator in smartphone AF technologies is the ability to determine and reach the correct focus position very quickly.

Autofocus technologies

The most widely used AF technologies for smartphone cameras are contrast, phase detection (PDAF) and laser. Contrast and PDAF are both passive technologies in the sense that they use the light field emitted by the scene. Laser AF is an active technology; it emits a laser beam toward the scene.

Contrast AF is very widely used in digital cameras. It uses the image signal itself to determine the focus position, relying on the assumption that the intensity difference between adjacent pixels of the captured image increases with correct focus [3], [2]. One image at a single focus position is not sufficient for focusing with this technology. Instead, multiple images from different focus positions must be compared, adjusting the focus until the maximum contrast is detected [4], [5]. This technology has three major inconveniences. First, the camera never can be sure whether it is in focus or not. To confirm that the focus is correct, it has to move the lens out of the right position and back. Second, the system does not know whether it should move the lens closer to or farther away from the sensor. It has to start moving the lens, observe how contrast changes, and possibly switch direction when it detects a decrease in contrast. Finally, it tends to overshoot as it goes beyond the maximum and then comes back to best focus, loosing precious milliseconds in the focus process.

Phase detection AF acts as a through-the-lens rangefinder, splitting the incoming light into pairs and comparing them. The shift between the signals received from the left and right side of the lens aperture, respectively, can be used to determine the distance of the subject from the camera. As a consequence, the AF knows precisely in which direction and how far to move the lens [4], [5]. This technology was developed at the age of film cameras and implemented utilizing specific AF sensors sitting typically below the mirror of a DSLR [4]. Recently it became possible to place phase detection pixels directly on the main CMOS image sensor [6, 7], which allows the usage of this technology in mirrorless digital cameras as well as in smartphones.

Laser AF measures the travel time of light from the device to the subject and back, to estimate the distance between the subject and the camera [8]. Even though the technology is totally different, it is comparable to PDAF in that it provides precise information on the subject distance.

Most digital single lens reflex (DSLR) cameras and digital still cameras (DSC) focus on demand, typically when the user begins pressing the shutter button. Depending on user settings, the camera will focus only once or continuously, tracking the subject, but in any case it is the user who triggers the focus. Smartphones, one the other hand, focus continuously, trying to always keep the subject in focus and always be ready for the shot. This AF strategy is part of the zero shutter lag (ZSL) technology found in recent devices [9]. Moving a small smartphone lens via a VCM is less power consuming than moving around big DSLR lenses. Nevertheless, the smartphone does not want to focus all the time, especially when it uses contrast AF, where focusing involves moving the lens out of the correct position and back. Therefore, smartphones observe the scene content and contrast and will typically trigger AF only when something changes. The scene change detection delay adds up to the total time of focusing.

A common smartphone AF behavior is described in Figure 1. It is composed of the following steps:

1. Scene change
2. Scene change detection
3. Focus direction change
4. Best focus reached
5. Stable on best focus

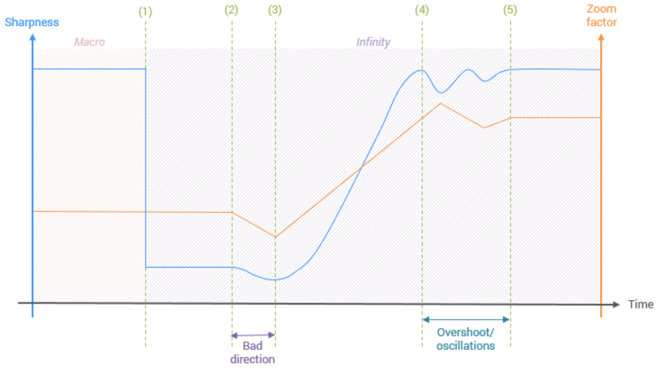


Figure 1. *Common autofocus behavior.*

The scene change corresponds to the user switching between an object at 30 cm and an object at 2 m. When the device detects the scene change, it reacts by starting to focus. Depending on the technology used, some devices do not focus in the right direction resulting in a more blurry image. Then, it focuses in the right direction to finally reach the best focus. Some oscillations may occur at this step. Ideally, a good autofocus must react quickly, it must start its convergence in the right direction and must reach the best focus quickly and smoothly without oscillations.

Our goal is to measure AF performance following a scene change, regardless of the AF technology used. Our approach can give information about the causes of bad autofocus behavior.

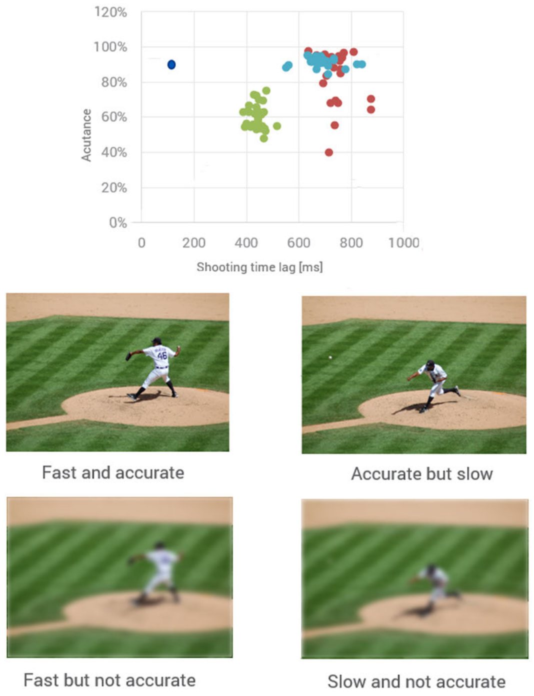
Autofocus quality

The two main criteria a user can expect from an AF are sharpness and speed. We propose with our approach to measure the acutance and the shooting time lag because these two metrics match the user experience best. We also provide information about repeatability of those metrics.

Figure 2 illustrates how the two criteria evaluated by our method translate into image quality and user experience. The acutance is a metric representing the sharpness, described in [10] and [11]. The shooting time lag is the time taken by the system to capture an image, described in [12], [13] and [14]. These two metrics will be defined in more detail later. The ideal case is a fast and accurate AF (top left in figure 2) while the worst results in a blurry image that was not captured when expected (bottom right). The top right picture shows an accurate AF, but too slow to capture the right moment while the bottom left picture has the opposite behavior.

Structure of this paper

First we will describe the state of the art and explain what approaches are currently available to assess AF. Then we will describe our proposed method: the goal, the hardware setup, the measurement and the quality metrics. Finally we will show the results provided by our method, make comparisons between several devices and then conclude.



Fast and accurate

Accurate but slow

Fast but not accurate

Slow and not accurate

Figure 2. *Different autofocus behavior results.*

State of the art

While autofocus hardware components and the computation of focus functions for contrast AF have been widely discussed in scientific literature, there are no scientific publications on the assessment of autofocus systems. Additional relevant information have been published in photography magazines and websites. In addition, the ISO standardization committee [15] is working on a draft standard on autofocus measurement that will not be discussed in this paper because it is not published yet. We hope that this paper will contribute to the dialog between the AF systems technology providers and the teams who evaluate AF quality.

Phase detection vs contrast AF

When live view mode on DSLRs and mirrorless cameras arrived on the market, the main image sensors did not have any phase detection pixels. In live view mode, The only way for focusing using the main image sensor was contrast AF. While photographers complained that this was not as fast as the phase detection AF they were used to, some camera testers pointed out that contrast AF was more accurate [16, 17]. The majority of pictures taken with phase detection AF showed the same sharpness as those shot with contrast AF, but quite a few pictures were slightly or totally out of focus. Contrast AF was much more reliable. This difference of perception between photographers and testers illustrates that AF assessment really must take into account both speed and accuracy.

At the time of these tests, a DSLR could do either phase detection AF (mirror down) or contrast AF (mirror up). Today's cameras have phase detection integrated on the main image sensor, enabling them to do both at the same time. This allows hybrid approaches where slight uncertainties in the distance estimated with by phase detection can be compensated by observing image contrast. As a result, the newest generation of of camera devices has more reliable AF than DSLRs had five or ten years ago.

Commercial image quality evaluation solutions

DxO, Imatest and Image Engineering have commercial solutions for AF measurement described on their websites.

DxO Analyzer 6.2 proposes a timing measurement that includes the shooting time lag measurement, which is very important to measure the autofocus speed. In addition, the video measurement on texture chart provides an acutance and a zoom factor measurement for each frame of a video stream, for an analysis of the dynamic performances of video autofocus by looking at the convergence curves for sharpness as well as the "lens breathing" behavior by looking at the zoom factor change for each frame.

These analyses have also been combined with automated change of the lighting conditions in color temperature and intensity using the automated lighting system proposed with DxO Analyzer.

Imatest propose two AF related measurements in their software suite, one for "AF speed" and one for "AF consistency". The first consists in measuring the MTF for every frame of a video [18]. For simplifying the MTF into a single scalar value, they propose the MTF area, i.e. the normalized sum of the measured MTF values over all frequencies, which they then plot over time. The resulting curve provides precise information on the behavior and speed of a video autofocus. The setup does not, however, provide timing information for still images. As the autofocus algorithms are usually different between photo and video because of different performance criteria, still image autofocus performances cannot reliably be assessed from video measurements.

Their second measurement aims at evaluating the accuracy and consistency of still image AF [19]. This consists in capturing images at different distances from a target, multiple images at each position, and in measuring the MTF50 (which is the frequency for which the MTF reaches 50%) for each image. Then these MTF50 values are plotted in function of the position to make visible the autofocus performance for different distances

The mean values for each position give an idea about AF accuracy at various object distances. It must be recalled, however, that the MTF50 values result from a combination of optical performance, AF and image processing (sharpening). A low value for a certain position might result either from intrinsically low optical performance at that object distance or from AF errors. The individual MTF50 values allow to identify outliers, which can provide valuable information about potential problems in the AF system needing investigation. It is also possible to visualize the deviation at each position, as a metric for AF repeatability.

Sharpness and its consistency are very important metrics for users. But they do not give a complete picture of the autofocus system and are not close enough to the user experience. The other important criterion for smartphone users, the time to focus in case of still image photography, seems not to be addressed by Imatest's offerings.

Image Engineering propose a combination of their "AF Box" and their "LED-Panel" lab equipment, which allows to measure both sharpness and shooting time lag [20], i.e. the time between pressing the shutter button and the beginning of exposure, in different lighting conditions. The photography magazine ColorFoto, who works with Image Engineering for their tests, describes their protocol as follows [21]: mount the camera at 1 m from the chart, (manually) focus at infinity and then trigger the shot. The shooting time lag includes the time to focus at 1 m and can be compared to the shooting time lag obtained with manual focus, which does not include any focusing delay. They test in two lighting conditions: 30 and 1000 lux, repeating the test ten times for each. We have no precise information on how they measure resolution and how they compute their final scores, but we suppose that they compute MTF50 on the slanted edge and compare it to a reference value obtained using manual focus. This protocol allows them to assess both focus accuracy (at a single distance) and timing and gives very comprehensive information about the AF performance of a digital camera.

The method described by Image Engineering and ColorFoto cannot directly be applied to smartphones because it requires manual focusing for both the reference MTF measurement and for the following measurements. More generally, their setup relies on the fact that the camera does nothing before the shutter button is pressed—which is not the case for smartphones. A smartphone, placed in front of an object at 1 m, will already be in focus before the shutter is touched.

DxOMark

The dxomark.com website publishes a mobile camera image quality benchmark that includes an autofocus measurement for smartphones. Like the other proposals, it consists in measuring the MTF on several images. There are some differences however:

First, the test chart is different. The other test charts, even if they differ between Imatest, Image Engineering and the ISO working draft, are mainly composed of (slanted) edges. DxOMark uses the Dead Leaves target described in [10]. While the MTF is in both cases measured on a slanted edge according to ISO 12233 [22], the texture on the Dead Leaves target is more representative of real-world use cases since its statistics follow a distribution with spatial frequency statistics closer to natural images.

Second, for simplifying the MTF into a single scalar value, rather than using the MTF50, DxOMark computes the acutance, which is a metric defined by the IEEE CPIQ group [23]. It is obtained by weighting the Modulation Transfer Function (MTF) by a Contrast Sensitivity Function (CSF), is independent from the sensor resolution and gives a quantitative measurement of the perceived sharpness and therefore represents the user experience more closely than the MTF50.

Finally and most importantly, the DxOMark setup was designed to test smartphones with continuous AF that cannot be switched to manual focus. The Dead Leaves chart is placed at a fixed distance from the device. Then, before every shot, an operator inserts a defocus target between the chart and the camera, waits until the device focuses on this defocus target and then removes it again. The acutance measurement is performed in auto mode (when the device decides itself where to focus) and in trigger mode (when an operator taps on the slanted edge to focus on it).

Proposed method

Rationale

We propose a protocol that provides information about both the AF consistency and the shooting time lag of a device. A Beta version of DxO Analyzer 6.3 was used as the main tool for this analysis.

For evaluating sharpness, we measure the MTF on a slanted edge of the Dead Leaves target and compute the acutance. We use the Dead Leaves target since its texture is close to real-world scene contents. We observe indeed that some devices have better focus performance on the Dead Leaves target than on an MTF target.

For the timing measurements we use the setup and method proposed in [12]. The shooting time lag contains both the time to focus and the processing time before the device captures the image. Measuring only the bare focusing time of a smartphone is not the most relevant information for system level performance assessment because the user will never observe the bare focusing time. Furthermore, it seems to be technically unfeasible without support from the manufacturer. Assessing the shooting time lag seems to be the best solution.

Measuring the shooting time lag requires a LED timer to calculate timestamps, e.g. the DxO Universal Timer Box [13]. The DxO Universal Timer Box is composed of five lines of LEDs that turn on and off at different times. Each line has only one LED illuminated at a time. The next led is illuminated and so on until the complete line is covered in a given time.

Finally, we test the camera in tripod and hand-held conditions. Hand-held conditions are a very common case, so the results are closer to the user experience. For testing the hand-held condition in a repeatable way, we use a hexapod platform to simulate a human holding the device. Hexapod platforms are used for moving and precise positioning along six degrees of freedom.

Hardware and lab setup

Our AF target is composed of a Dead Leaves target and a DxO Universal Timer Box. It is placed at 2 m from the device, which corresponds roughly to 70 times the 35-mm equivalent focal length of most smartphones. Figure 3 shows diagrams about the setup.

The principle is to place a defocus target at macro distance, force focus when necessary and then remove the defocus target to let the device under test focus on a Dead Leaves target at 2 m. Focusing at macro is done with a defocus target, shown in Figure 4. No measurement is performed here, so no specific target is needed, but there must be a texture helping the device to focus on it (text for instance). The defocus target is placed in front of the device to cover its entire field of view as shown in Figure 3. We let the device enough time to focus on its. This target is then quickly moved down outside the field of view to provide a fast switch between macro and 2 m as shown in Figure 3. The removal of the defocus target triggers a scene change detection in the device, which will then start focusing on the Dead Leaves target.

In our current setup, the defocus target is removed manually by an operator. To prevent the device from focusing while the target is still within its field of view, the time for the defocus target to disappear shall be less than 100 ms. The presence and the disappearance speed of the defocus target are measured by two infrared sensors. In order to simulate the device's field of view, the two red dots of the sensors have to be at the top and the bottom of the device screen preview, as shown in Figures 4 and 5. It ensures the device field of view is well represented by the system. These validations are useful for benchmarking as they allow a higher repeatability.

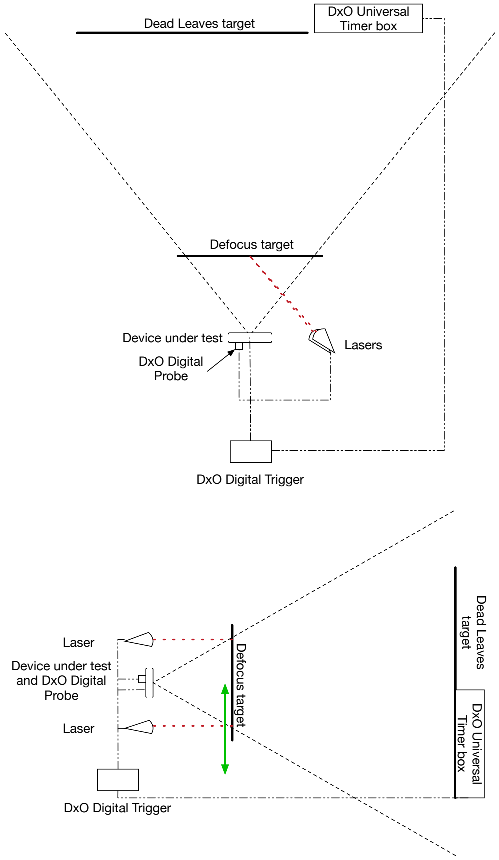


Figure 3. *Diagram of the autofocus measurement setup from the top and the side.*

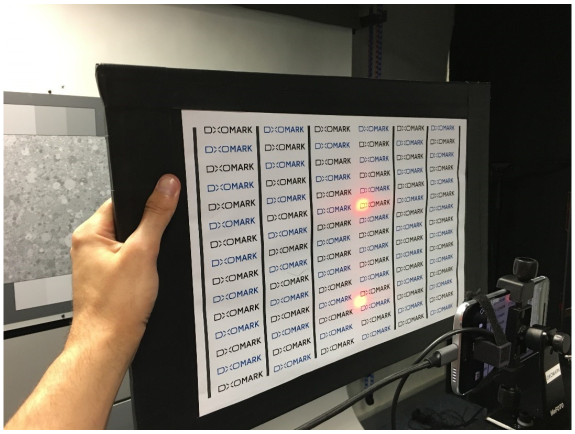


Figure 4. *Defocus target and laser detection.*

Figure 5. *Red dot positions on the device's screen when the defocus target is ahead.*

When the sensors detect disappearance, the system gets the LED positions from the DxO LED Universal Timer. It then waits a short time t_{wait} to simulate the human reaction time lag. After t_{wait}, the digital probe (which simulates a finger on the touchscreen) is used to command capture. By detecting the LED positions on the image finally taken, we can determine precisely the time lag between the trigger and the beginning of the exposure. This is the shooting time lag which is a very important part of AF user experience.

Figure 6 summarizes the different setup components and their connections.

- Camera device under test: must have a capacitive touchscreen to work properly with the DxO Touchscreen Probe.
- DxO Touchscreen Probe: electronically simulates a human finger on a capacitive touch screen. It is attached to the touch screen using a hook-and-loop fastener and must be plugged into a DxO Digital Trigger.
- DxO Digital Trigger: remotely controls a DxO Touchscreen Probe and simultaneously sends synchronization signals to a DxO Universal LED Timer. It sends the LEDs position to the computer when the shot is triggered.
- Dead Leaves target: used to measure the sharpness of a picture. It is placed at 2 m from the device. See Figure 3.

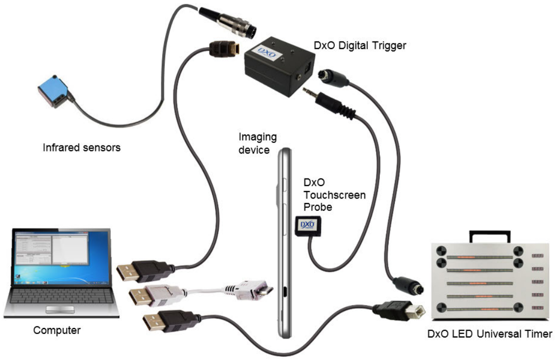


Figure 6. *Components of the autofocus measurement setup and their connections.*

- DxO LED Universal Timer: device composed of several LED lines used to measure multiple timings such as shooting time lag or rolling shutter. It is placed in the same plane as the Dead Leaves target. See Figure 3.
- Defocus target: placed in front of the imaging device to let it focus at a macro position; then moved down to let it focus on the Dead Leaves target. See Figure 4.
- Infrared sensors: used to detect the presence of the defocus target. They are plugged into the DxO Digital Trigger to send a signal when the defocus target disappears, which is when the device starts to focus. It is placed near the imaging device and the laser are facing the defocus target.

The timing diagram of our setup is summarized in Figure 7.

- t_{sensors} is the time between the deactivation of the two infrared sensors when the defocus target is moved down. This time must be less than 100 ms to ensure that the device does not focus while the target is still in its field of view. A sensor is activated when an object (the defocus target in this case) is in front of it.
- t_{wait} corresponds to the time between defocusing and triggering.
- t_{push} represents how long the DxO Digital Probe pushes the trigger. In this case, the synchronization is done on the push down meaning the beginning of the exposure is considered at the push down. It can also be done on the push up, depending on the device tested. The LEDs positions recording is synchronized with the beginning or the end of the push time.
- t_{lag} finally represents the time between pressing the exposure button on a mobile device and the beginning of the exposure, which is the shooting time lag.

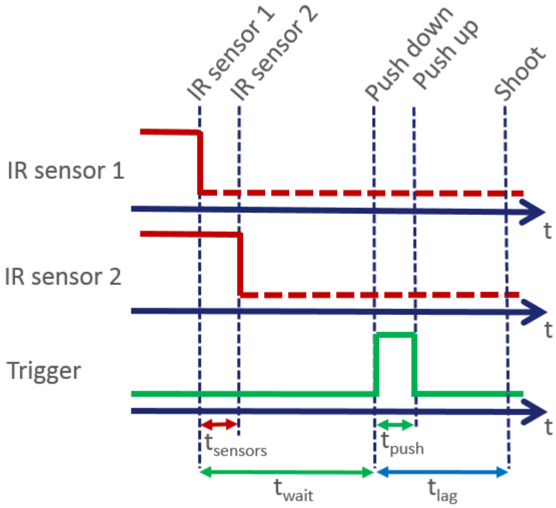


Figure 7. *Timing diagram.*

In order to avoid to stress the device and let it enough time to process the image or frames for multi-images algorithms, we wait a few seconds between each shot.

Measurements

The acutance is computed from the Dead Leaves target's edges as illustrated in Figure 8, following the ISO 12233 method [22] to compute the MTF. We compute the MTF from eight slanted edges (red circles on the picture) and then the mean is used for computing the acutance.

$$Acutance = \int_0^\infty MTF(\upsilon) \cdot CSF(\upsilon) \cdot d\upsilon \qquad (1)$$

Equation (1) shows that a contrast sensitivity function (CSF) is used to weight the values of the MTF for the different spatial frequencies. The CSF is defined in ISO Standard 15739 [24] for visual noise measurement. The CSF is given in Equation (2) where $a = 75$, $b = 2$, $c = 0.8$, $K = 34.05$ and υ is in cycles/degrees.

$$CSF(\upsilon) = \frac{a \cdot \upsilon^c \cdot e^{-b \cdot \upsilon}}{K} \qquad (2)$$

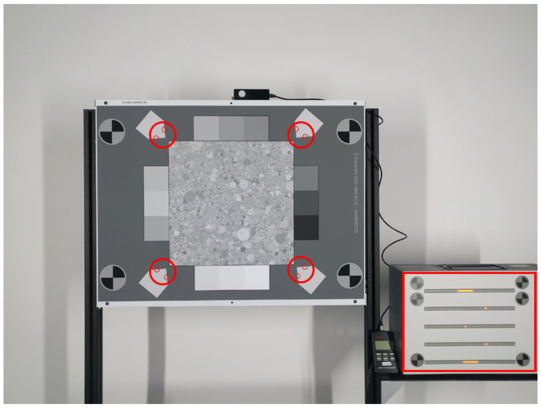


Figure 8. *Dead Leaves target and DxO Universal Timer Box*

The acutance result depends on the viewing condition of the image, the size (be it printed or on-screen) and the viewing distance. For instance, if an image is viewed on a small smartphone screen, we will not have the same perception of sharpness than if it is printed on a large format. The parameters composing the viewing conditions are the following:

- Distance
- Pixel pitch (for computer display)
- Print height (for print)

The measurement algorithm uses these viewing conditions to determine the coefficient for converting the spatial frequency of the CSF of the visual field, expressed in cycle/degree, into cycle/pixel as measured on the image. The effect of the viewing conditions is to stretch the CSF along the frequency axis. If you look at an image from afar, the CSF will narrow on low spatial frequencies, giving more weight to these frequencies and less weight to the high ones. Although the pictures are first seen on the smartphone screen, we are choosing a more challenging viewing condition, such as looking at the pictures on a notebook screen (height 20 cm at a distance of 50 cm), which allows to benchmark and differentiate autofocus performance of different devices.

The shooting time lag is computed with the DxO Universal Timer Box as illustrated in Figure 8, by subtracting the LED positions recorded when triggering from the LED positions observed on the picture. With one LED bar, the minimal measurable time is one LED. The LED calibration is the period of a line. To increase measurement accuracy, one could use a shorter line calibration. But if the line calibration is too short, there can be one or more periods during the time lag, and these would not be visible. So with only one bar, the accuracy of the measurement is severely limited. By using several LED bars at different periods or calibers, it is possible to accurately calculate the capture beginning with maximum accuracy (about 1/100 of the fastest line): the slowest line permits calculating a rough estimate of the time lag, and a faster line permits calculating a better estimate from this value. This is why the periods of the DxO Universal Timer Box lines are set to 100, 1000, 8000, 1000 and 100 ms.

These measurements are performed on several images to asses the repeatability of the AF performance (sharpness and shooting time lag) in identical shooting conditions. That is why the measurement accuracy depends on the number of shot used.

Quality metrics

The work presented in this article combines the acutance and the shooting time lag to provide a simple and relevant AF measurement assessing both sharpness and speed of the AF, which are the two major components of AF quality.

The final result is a graph with acutance plotted against shooting time lag. As you can see in Figure 9, it contains a point for each image taken.

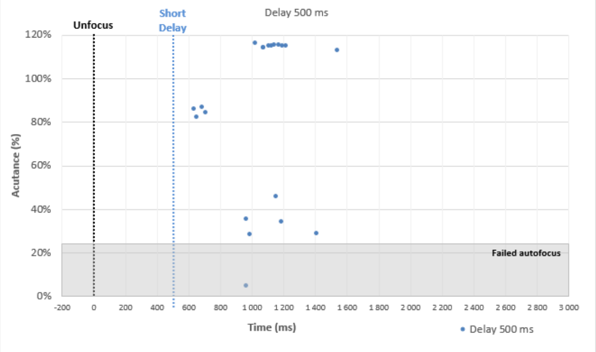


Figure 9. *Device A - Autofocus performances at 1000 lux with proposed measurement*

The dots above 100% are the result of over-sharpening and their values are clipped to 100% to compute the metrics. Indeed, a picture cannot be more precise than the reality.

AF failures are represented in the graph with an acutance of 5%. In fact, these pictures are often too blurry to compute both the acutance and the shooting time lag. The default value for acutance is set to 5% (representing a completely blurry image), but we did not want to penalize the shooting time lag. Indeed, even if the image is blurry the device can be fast to capture it. In order to clearly see the different failures (dots are not overlaid) without much influence on the mean shooting time lag, we choose to assign a random value to the shooting time lag, included in the normal distribution of the successful pictures.

To summarize the AF performance, we propose to compute the following two key metrics:

- **Average shooting time lag** gives a general idea of the capacity of the AF to adapt quickly to a scene change.

- **Autofocus irregularity** provides information about AF repeatability, this is defined as the average acutance difference between the highest acutance in a series and the acutance for each shoot.

We use the highest acutance in a series since most smartphones do not allow us to manually find the focus position that yields the best acutance. We therefore use the highest acutance that the smartphone has reached. As the example of Figure 9 suggests, this is usually equivalent. We are also computing two additional metrics that can be useful for further analysis:

- **Shooting time lag standard deviation** measures the repeatability of AF convergence speed.
- **Average acutance** gives a general idea about the perceived sharpness of the images that a certain device takes. However, this result depends on the lens MTF, the degree of sharpening applied in image processing and on the autofocus.

Limitations and future work

While our proposed quality metrics and most of our method apply to all types of digital still cameras, our setup was designed for smartphones. Its extension to DSLRs is more complicated than simply replacing our touchscreen trigger with a mechanical finger. For instance, letting the device under test focus from macro to a target at 70 times its 35-mm equivalent focal length would require a huge lab for long focal lenses. Image Engineering's approach, to let the device focus from infinity to a close target seems more practical—supposed that the device can be forced to defocus at infinity.

In a more general manner, evaluating an AF at a single distance does not necessarily result in a complete picture of its performance. It might be useful to place our Dead Leaves target, like proposed by Imatest, at several different distances. It might even be useful to place the defocus target at different distances. Currently we place it close to the closest macro distance. This might aid a contrast AF algorithm that has to guess its initial focusing direction. A defocus target placed farther away from the camera might increase the probability that a device chooses the wrong direction, which would result in a significantly longer shooting time lag. We also consider putting a Dead Leaves chart and slanted edges on the defocus target, to assess focusing from far to close. These kind of tests will become possible as we continue to automate our setup.

Finally, our setup does not yet assess the ability of a device to track a subject in motion. Neither does it test the AF reaction to face detection and the ability of the device to keep the subject in focus while it is moving before the command of the shoot.

Results

Plotting the acutance in function of the shooting time lag provides an intuitive visual representation of the detailed information about the AF performance. Not only does this allow to determine instantly the two most important criteria, sharpness repeatability and speed, the plot also allows to analyze the AF strategies of the different devices.

The proposed measurement was used to test the influence of various test conditions such as lighting condition, trigger delay and camera motion. Once the influential parameters were identified and defined, the measurement was used to build a benchmark of more than 20 devices providing very important insights into the performance of various AF technologies.

Autofocus performance comparison

The performance in bright light of two smartphones released in 2016 can be compared by looking at Figures 9 and 10.

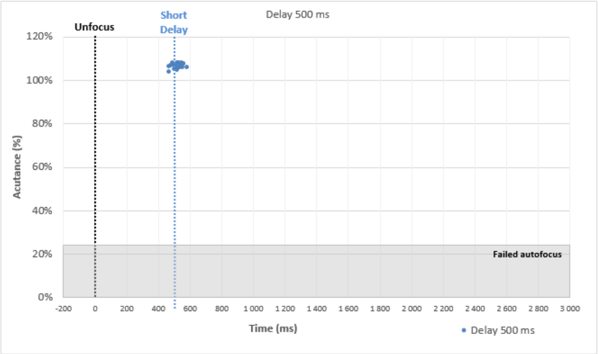


Figure 10. *Device B - Autofocus performances at 1000 lux with proposed measurement*

Differences between AF system performance or between different testing conditions are immediately visible on the chart. AF acutance irregularity is 21.4% for device A against 5.0% for device B. We can conclude that device B is significantly more accurate than device A. In addition, with an average shooting time lag of only 18 ms, device B takes the picture exactly when the user triggers the shutter, whereas device A introduces a notable lag of 546 ms on average. In conclusion, device B has superior performances compared to device A in both acutance repeatability and speed.

The chart intuitively illustrates these metrics from the scattering of the dots. It also enables deeper analysis that can help camera manufacturers and tuning teams to improve performance: The AF results of device A can be divided into three categories. In the first category, the device favors accuracy over speed, these are the dots with acutance $> 100\%$, but with shooting time lag scattered between 500 and 1100 ms. Then in the second category, the device favors short shooting time lag over precision and captures quickly between 100 and 200 ms. With an acutance over 80%, these images are slightly out of focus, but still usable on a smartphone screen. Finally the third category has some strong AF failures resulting in very blurry images having acutance lower than 50%. The device manufacturer could use this information to gain insight into the different failure modes to improve their AF algorithm.

It is interesting to notice that, in Figure 10, there are also some points before the command of the capture (blue dotted line called Short Delay). Some devices continuously save pictures in an internal memory. When the user presses the trigger, the device is able to select the sharpest picture in that buffer. So the device can provide an image captured just before the user pressed the trigger. Ideally, a device must tend toward a zero shutter lag if it has the ability to continuously focus on the scene, thus providing sharp images with zero lag.

Lighting conditions

The test results confirmed that lighting condition is a very influential parameter. For some devices, the results can be completely different in bright and in low light. We can see an example of this behavior by comparing the results obtained with the same test device in bright light and low light conditions. In Figure 11 the AF is fast and accurate. However, in low light conditions shown in Figure 12, the AF is slow, the shooting time lag becomes less predictable and there are even some failures. For $t_{\text{wait}} = 500$ ms, we measure AF irregularity of 20.9% and average shooting time lag of 978 ms—compared to an irregularity of only 5.0% and a lag of only 76 ms in bright light.

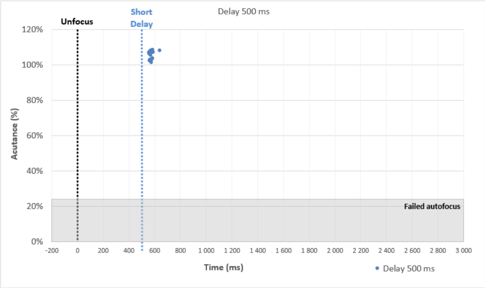


Figure 11. *Device D - Autofocus performances in bright light*

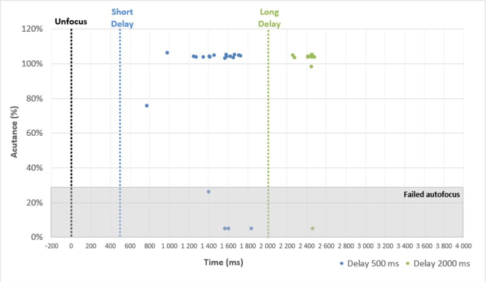


Figure 12. *Device D - Autofocus performances in low light*

Delay between scene change and trigger

In defining the test conditions, the setting of the delay between scene change and trigger is very important to highlight the performances of a continuous autofocus system.

The most challenging condition would be a delay of 200 ms corresponding to the human reaction time lag including the processing of the scene change by the human brain as well as the lag between the decision to press the trigger and the exact time when the finger is touching the screen. The time lag can also be increased up to 500 or 2000 ms to reflect a usage case where the photographer would wait between the scene change and the decision to press the trigger. The relative results for different t_{wait} will depend on the speed and effectiveness of the continuous autofocus.

If the device has a continuous autofocus that manages to focus before the user hits the trigger, it can simply and instantly take the picture. Otherwise, if the image is not in focus yet, the autofocus algorithm has two options as it has to make a trade-off between letting the AF fully converge (preferring accuracy) and taking the picture as fast as possible (preferring short shooting time lag). Different manufacturers may chose different strategies in this case. The user can favor accuracy by waiting longer before hitting the shutter button and thus avoiding to put the AF under pressure. Figure 12 illustrates such a case. We can see that the AF is more repeatable when waiting for 2000 ms instead of 500 ms because the green points are less scattered than the blue ones. There are less AF failures and the AF irregularity metric improves from 20.9% to 6.0%. Average shooting time lag also improves from 258 ms to 58 ms.

Figure 13 shows that even the best device currently tested for AF cannot achieve the same performances with 200 ms than with 500 ms delay. In this example, the average shooting time with a 200 ms delay (in red) is 288 ms while it is 65 ms with a 500 ms delay (in blue). The assumption is that, as the autofocus convergence time of the device is 500 ms and the device autofocus convergence strategy is to favor accuracy, it tends to capture the image 500 ± 50 ms after the defocus event, whether the trigger is pressed 200 ms or 500 ms after defocus.

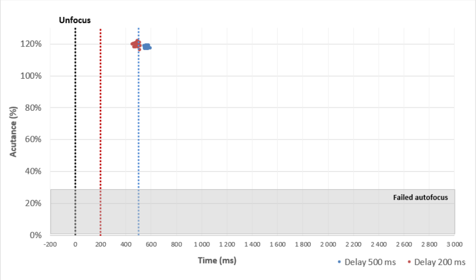


Figure 13. *Device C - Autofocus performances with* $t_{wait} = 200$ *ms*

Hand-held vs tripod

Our test results confirmed that autofocus performances decrease when tested in hand-held conditions (Table 1) compared to tripod conditions (Table 2).

Table 2 shows that the best devices have almost the same performances on tripod and hand-held in bright light conditions.

Table 1: Device A: performances comparison in bright light

	Tripod	**Hand-held**
Average Acutance	90.1 %	71.0 %
Autofocus irregularity	14.8 %	34.2 %
Average shooting time lag	319 ms	390 ms
Standard deviation shooting time lag	202 ms	290 ms

We have observed that the shooting time lag decreases in hand-held conditions as the images will be subject to motion blur that may affects the focus measurement of the device. Therefore, the device may shoots before reaching the best focus resulting in blurry images captured faster hand-held that with a tripod. An analysis of the images confirmed that there is some non-directional blur confirming that the sharpness loss is caused by autofocus failure and not by motion blur in bright light.

Table 2: Device C: performance comparison in bright light

	Tripod	Hand-held
Average Acutance	118.0 %	114.0 %
Autofocus irregularity	5 %	5 %
Average shooting time lag	81 ms	104 ms
Standard deviation shooting time lag	13 ms	12 ms

AF technology benchmark

More than 20 smartphone cameras with different autofocus technologies have been tested with this AF measurement. We are reporting the results from four devices with different AF technologies that are summarized in Table 3. The Figures 14 and 15 illustrate our results with a time delay of 500 ms for both bright light and low light conditions.

Table 3: Technologies used for the devices under test

	Contrast	PDAF	Laser
Device B	X	X	X
Device C	X	X	
Device D	X	X	
Device E	X		

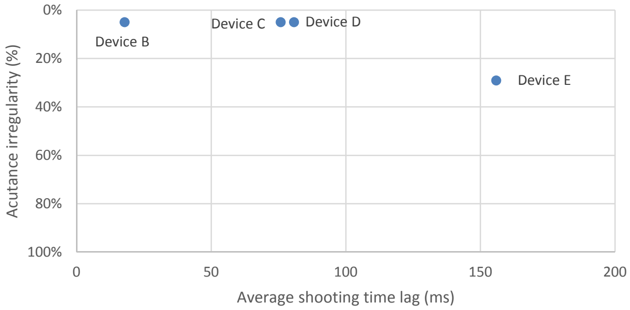


Figure 14. *Autofocus performances in bright light*

The analysis of the bright light from Figure 14 illustrates the following results: With an irregularity of 30%, the device E with only contrast autofocus has the least repeatable results of all four devices but it achieves an acceptable shooting time lag of 150 ms. The best bright light performances are achieved by the device combining both PDAF and contrast (devices B, C and D) as they all have very small acutance irregularities lower than 5% and average shooting time lag smaller than 100 ms. Although all three devices are very good, the device B that also has a laser technology is the best of the three with a shooting time lag smaller than 20 ms.

In low light, the combination of PDAF, laser and contrast embedded in the device B clearly has the best results with a gap compared to other technologies that is even stronger than in bright light. The device B is the only device that achieves a zero shooting time lag with an acutance irregularity lower than 5%. The device C and D are both using PDAF and contrast technologies and are the 2015 and the 2016 versions from the same smartphone manufacturer. It is very interesting to highlight the performance improvement from this technology between two devices released one year apart. On one hand, the device C, which is the 2016 version achieved performances that are very close to device B in acutance irregularity despite a longer average shooting time lag of 200 ms that remain fast although the lag can be perceived by the photographer. On the other hand, the device D, which is the 2015 version using the same PDAF and contrast technologies has a lower 30% acutance irregularity, but more importantly has a very poor average shooting time of almost 1000 ms. The performances of the device E with only contrast autofocus were already low in bright light and decrease further in low light with an acutance irregularity of 50%, meaning that several images are significantly blurry and an average shooting time lag of more than 600 ms that will be perceived as very unpleasant by most end users. This test clearly highlight the benefit of laser and PDAF technologies that provide information about the shooting distance enabling faster and more accurate autofocus performances.

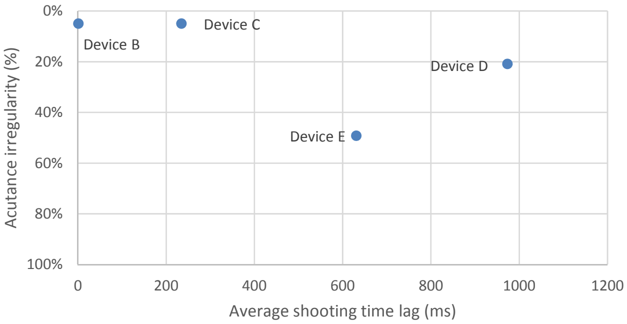


Figure 15. *Autofocus performances in low light*

Conclusion

Everyone has a collection of images that are either blurry because of autofocus failure or taken too late once the scene has changed. An autofocus failure makes an image useless for the user even if all other image quality attributes were to be perfect. With the ever increasing number of pictures taken in the world driven by the raise of image quality in smartphones, it becomes very important to have an autofocus measurement reflecting the experience of the user who is looking for consistently sharp image taken at the precise time he or she presses the trigger.

Although there are no publications related to autofocus measurement, several commercial solutions offer extensions of traditional sharpness measurement for still images to evaluate either the repeatability of the autofocus for photo mode, or assessing the sharpness for every frame of the video thus providing useful information on video autofocus.

Our method is the first one to establish a measurement that will assess both timing and sharpness performances of devices with continuous autofocus such as smartphones.

The method is using together the edge acutance measurement of a textured chart and the time lag measurement with the LED Timer. The automated capture and analysis also enables measurement on large number of shots for each tested camera tested and each relevant lighting condition. This large sample size is very important to have repeatable results since the autofocus systems that we test are not. The method also defines the relevant statistical metrics used to summarize the measurement of dozens of pictures in four metrics.

The method has been tested on more than 20 mobile cameras and has already allowed to establish the difference in performance between the different technologies used for smartphone autofocus. The contrast autofocus is slow and not repeatable and this becomes even stronger in low light. The addition of the PDAF brings a significant improvement in bright light, and our measurements were able to highlight the progress of this technology in low light as it became more mature. We hope that the availability of new autofocus evaluation technologies will help camera manufactures to design and test their product faster, as well as reach better performances for the users.

References

[1] Myung-Jin Chung, "Development of compact auto focus actuator for camera phone by applying new electromagnetic configuration", *Proc. SPIE 6048 Optomechatronic Actuators and Manipulation*, 2005.

[2] John F. Brenner, Brock S. Dew, J. Brian Horton, Thomas King, Peter W. Neurath and William D. Selles, "An Automated Microscope For Cytologic Research: A Preliminary Evaluation", *The Journal of Histochemistry and Cytochemistry*, Vol. 24, No. 1, pp. 100-111, 1976.

[3] A. Santos, C. Ortiz De Solorzano, J. J. Vaquero, J. M. Pena, N. Malpica, F. Del Pozo, "Evaluation of autofocus functions in molecular cytogenetic analysis", *Journal of Microscopy*, Vol. 188, Pt 3, pp. 264272, 1997.

[4] Goldberg, N., *Camera Technology: The Dark Side of the Lens*, Academic Press, 1992.

[5] Sidney F. Ray, *Applied Photographic Optics*, Focal Press, 2002.

[6] Ray Fontaine, "Innovative Technology Elements for Large and Small Pixel CIS Devices," *International Image Sensor Workshop*, 2013.

[7] Ray Fontaine, "The State-of-the-Art of Mainstream CMOS Image Sensors," *International Image Sensor Workshop*, 2015.

[8] Ralph Jacobson, Sidney Ray, Geoffrey G Attridge, Norman Axford, *Manual of Photography: Photographic and Digital Imaging*, Focal Press, 2000.

[9] *Breakthrough mobile imaging experiences*, whitepaper, Qualcomm Technologies, Inc., 2014.

[10] Frédéric Cao, Frédéric Guichard, Hervé Hornung, "Dead leaves model for measuring texture quality on a digital camera", *Proc. SPIE 7537*, Digital Photography VI, 75370E, 2010.

[11] Donald Baxter, Frédéric Cao, Henrik Eliasson, Jonathan philips, "Development of the I3A CPIQ spatial metrics", *Proc. SPIE 8293*, Image Quality and System Performance IX, 829302, 2012.

[12] Franois-Xavier Bucher, Frédéric Cao, Clément Viard, Frédéric Guichard, "Electronic trigger for capacitive touchscreen and extension of ISO 15781 standard time lag measurements to smartphones", *Proc. SPIE 9023*, Digital Photography X, 90230D, 2014.

[13] Lucie Masson, Frédéric Cao, Clément Viard, Frédéric Guichard, "Device and algorithms for camera timing evaluation" *Proc. SPIE 9016*, Image Quality and System Performance XI, 90160G, 2014.

[14] ISO 15781, *Photography – Digital still cameras – Measuring shooting time lag, shutter release time lag, shooting rate, and start-up time*, 2015.

[15] ISO/NP 20490, *Measuring autofocus Performance of a Digital Camera*, Under development.

[16] Malte Neumann, "Scharf gestellt: Phasen- gegen Kontrast-Autofokus", *ColorFoto 9/2011*, p. 26-32.

[17] Roger Cicala, *Autofocus Reality Part 1: Center-Point, Single-Shot Accuracy*, https://www.lensrentals.com/blog/2012/07/autofocus-reality-part-1-center-point-single-shot-accuracy/, 2012, consulted on 2016-11-28.

[18] *Autofocus Speed*, Imatest, http://www.imatest.com/docs/autofocus-speed/, consulted on 2016-11-28.

[19] *Autofocus-Consistency (Post Processor)*, Imatest, http://www.imatest.com/docs/autofocus-consistency/, consulted on 2016-11-28.

[20] *AF BOX: Measure shutter delay*, Image Engineering, http://www.image-engineering.de/products/equipment/measurement-devices/381-af-box, consulted on 2016-11-28.

[21] Uwe Artmann, "Der neue ColorFoto-Kameratest – Testversion 1.6", *ColorFoto 4/2011*.

[22] ISO 12233, *Photography Electronic still-picture cameras Resolution measurements*, 2014.

[23] IEEE 1858, *IEEE Approved Draft Standard for Camera Phone Image Quality (CPIQ)*, 2016.

[24] ISO 15739, *Photography – Electronic still-picture imaging – Noise measurements*, 2013.