

Review article

Vladan Blahnik* and Oliver Schindelbeck

Smartphone imaging technology and its applications

<https://doi.org/10.1515/aot-2021-0023>Received May 6, 2021; accepted June 29, 2021;
published online August 10, 2021

Abstract: Thanks to their portability, connectivity, and their image performance – which is constantly improving – smartphone cameras (SPCs) have been people's loyal companions for quite a while now. In the past few years, multicamera systems have become well and truly established, alongside 3D acquisition systems such as time-of-flight (ToF) sensors. This article looks at the evolution and status of SPC imaging technology. After a brief assessment of the SPC market and supply chain, the camera system and optical image formation is described in more detail. Subsequently, the basic requirements and physical limitations of smartphone imaging are examined, and the optical design of state-of-the-art multicameras is reviewed alongside their optical technology and manufacturing process. The evolution of complementary metal oxide semiconductor (CMOS) image sensors and basic image processing is then briefly summarized. Advanced functions such as a zoom, shallow depth-of-field portrait mode, high dynamic range (HDR), and fast focusing are enabled by computational imaging. Optical image stabilization has greatly improved image performance, enabled as it is by built-in sensors such as a gyroscope and accelerometer. Finally, SPCs' connection interface with telescopes, microscopes, and other auxiliary optical systems is reviewed.

Keywords: computational imaging; digital camera; multicamera; optical design; photography; smartphone.

CONTENTS

1 Evolution of the mobile phone imaging system for the mass consumer market	146
1.1 From mobile phone to smartphone	146
1.2 Smartphones today	147

1.3 Tomorrow's smartphones	147
2 Mobile imaging	147
2.1 Market development	148
2.2 Supply chain	149
3 Brief history and milestones of smartphone imaging technology	150
4 Physical properties and requirements of Smartphone photography	150
4.1 Camera form factor and image sensor size ...	150
4.2 Image sensor resolution	151
4.3 Optical resolution and required aperture	154
4.4 Portrait photography: Perspective, bokeh, and depth of field	155
4.5 Étendue and photographic exposure	159
4.6 David versus Goliath: The pros and cons of miniaturization	160
4.7 SPC lenses: How good are they actually?	162
5 The multicamera system in modern smartphones	164
6 Optical System Design	165
6.1 Optical design structure of a smartphone lens	165
6.2 Optical design imaging performance	172
6.3 Extreme wide-angle lenses	175
6.4 Tele lenses	176
6.5 Periscope tele lenses and alternative tele concepts	176
7 Zoom	177
7.1 Hybrid zoom in multicamera systems	178
7.2 Optical zoom systems	180
8 Opto-mechanical layout and manufacturing	182
8.1 Plastic lenses: Key for miniature opto-mechanical layout	182
8.2 Opto-mechanical layout	182
8.3 Active optical assembly	183
8.4 Tolerancing and yield analysis	184
8.5 Wafer-level manufacturing	184
8.6 Anti-reflection coating for plastic lenses	186
9 Image sensor	186
10 Image processing	189
11 Noise and noise reduction	190
12 Focusing	191

*Corresponding author: Vladan Blahnik, Carl Zeiss AG, Carl-Zeiss-Straße 22, 73447 Oberkochen, Germany,
E-mail: vladan.blahnik@zeiss.com

Oliver Schindelbeck, Carl Zeiss AG, Carl-Zeiss-Straße 22, 73447 Oberkochen, Germany, E-mail: oliver.schindelbeck@zeiss.com

12.1	Autofocus methods: Contrast and phase detection	193
12.2	Optical System changes focus position	194
12.3	Focusing mechanisms: Voice coil motors and other concepts	196
13	Image stabilization	197
13.1	Hand-shaking and image blur	197
13.2	Optical image stabilization implementations	200
14	Dynamic range	55
14.1	HDR imaging	201
14.2	Lens flare and ghosts	204
15	Portrait mode	206
15.1	3D depth acquisition technology	209
15.2	Simulation of lens bokeh: Camera 3D point spread function	68
15.3	Portrait look: a quality evaluation	214
16	Image performance specification and test	216
16.1	Lab evaluation during R&D	218
16.2	Evaluation of image quality in the imaging pipeline	219
17	Smartphone camera interface with telescopes, microscopes, and accessory lenses	220
18	Summary and outlook	224
	References	226

1 Evolution of the mobile phone imaging system for the mass consumer market

1.1 From mobile phone to smartphone

“A **smartphone** is a **mobile phone** that includes advanced functionality beyond making **phone** calls and sending text messages. Most smartphones have the capability to display photos, play videos, check and send e-mail, and surf the Web.” [1].

As per this definition, one could posit that the Personal Communicator launched by IBM in 1992, which was capable of sending and receiving emails and faxes, was in fact the first-ever “smartphone.” The subsequently launched Nokia Communicator and the Blackberry devices from RIM offered a kind of mobile office for one’s pocket, and became an essential part of every manager’s briefcase. In 2005, Nokia offered its 770, a mobile computer with a large, capacitative touchscreen which, in hardware terms, is very close to what we call a smartphone today.

So when we think of smartphones, why do we usually associate them with Apple’s iPhone if it wasn’t launched until 2007? The multimedia features of the first iPhone were not incredibly impressive. It featured only a 2 MP camera, with no autofocus or flash, and no front-facing camera. The display only achieved a resolution of 480×320 pixels and it did not even support the current third generation (3G) mobile standard. Many smartphones that had come out years before actually outranked it. However, what was new and appealing about it were the smartphone’s controls: it featured a single button and a capacitive touchscreen offering precise and intuitive control. The slide-out electronic keyboard replaced the many buttons that mobile phones previously had, allowing for a much larger screen. And the phone itself responded to users’ actions – thanks to the ambient light, acceleration and position sensors. But even more important than the multimedia features was the interaction between man and machine – a phenomenon the iPhone rendered “smart.” But it was not until later iPhone generations that the technical specifications were slowly brought into line with the state of the art.

In 2007, Apple accompanied the release of the iPhone with its dedicated App Store. This gave many small companies and individual developers the chance to contribute to programs and market themselves without having to worry about running their own sales operations. A service that probably helped Apple earn more than its hardware did. This created a major dynamic in the software landscape. Furthermore, a key strategic move that proved crucial for Apple’s success was its decision to partner with mobile network operators and offer data packages alongside its devices. This gave users the feeling that they were not making a costly mistake by purchasing from Apple. It was only then that users could really start surfing the mobile web largely worry-free and actually start using their smartphones to their full advantage.

With its open operating system Android, Google was Apple’s only competitor who proved to be a match for its iOS operating system in the long run. The Android operating system requires and enables a highly standardized hardware environment which, in turn, allows numerous other manufacturers to easily develop and offer smartphones running on the Android system. While Android’s market share is far higher than that of iOS, it is very fragmented and split between multiple manufacturers, each of whom have made their own adjustments to it. During the initial phase in particular – but also to this day – this fragmentation makes it difficult to ensure reliable security concepts and a consistent user experience, which is where Apple comes out on top.

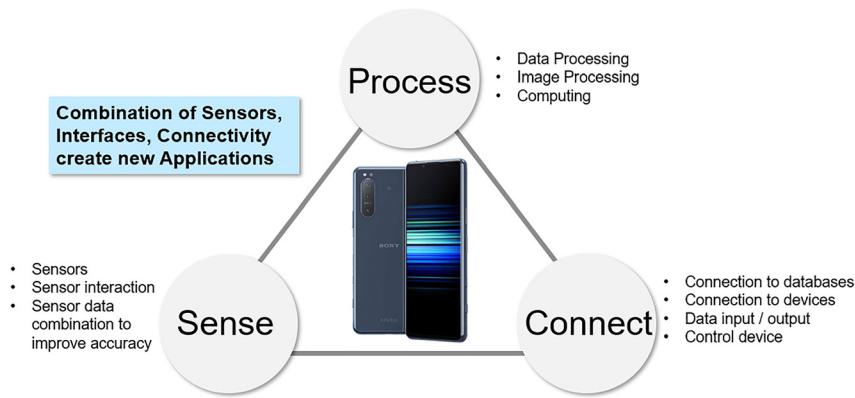


Figure 1: The combination of sensors (e.g. gyroscope, proximity sensor, GPS, microphone, and camera), processing units (CPU, GPU, etc.) and connections (e.g. Wi-Fi and Bluetooth) create a whole world of applications.

1.2 Smartphones today

Today, smartphones are much more than just compact pocket knives boasting a series of nifty functions: their mix of sensors (“sensing”), data processors (“processing”), and their connectivity (“connecting”) create whole new functions and processes that no small device had managed to offer before – even in conjunction with several different devices and plenty of time and effort (Figure 1).

Many new smartphone applications – such as navigation systems and augmented reality – are now possible thanks to a combination of sensing, processing and connectivity. Details on the status and evolution of these fundamental technologies can be found in the literature: For “miniature smartphone camera (SPC) sensors” see refs. [2, 3]; for processing and system on a chip: [4, 5] and connectivity [6].

Nowadays, it is much more likely for people to leave the house without their wallet or car keys than without their phone. That’s because the smartphone has redefined the lives of a whole generation of people. Some have even grown up with smartphones, which has completely changed their values and habits. 30 years ago, most young people wanted a driver’s license and their own car. As well as being a status symbol, a car was their way of getting from A to B and staying in touch. Today, young people often view cars as a necessary evil, while the smartphone is seen as a gateway to the world and to one’s friends. The modern-day status symbol is pocket-sized. If you ask people today what item they couldn’t live without, it’s not just young people who would choose their smartphone. The impact of smartphone’s on society has been discussed in several papers (e.g. [7, 8]).

1.3 Tomorrow’s smartphones

So what might tomorrow’s smartphone be like?

New, faster mobile standards are allowing for ever more data-intensive and storage-intensive functions to be

transferred to external computing centers in the cloud. This means smartphones are becoming less of a high-performance minicomputer – and more a kind of platform to manage how data is distributed to numerous subsystems like screens, eyeglasses, watches, cameras, health sensors, and audio playback functions, thereby making all of these devices smart, too.

So the smartphone itself is likely to take a backseat to them all. For instance, smartphones will no longer need their own (large) screen, so it can be integrated into another device, such as a smartwatch. And the lower processing power requirements mean that miniaturization can continue advancing.

Essentially, users will prefer to interact with these subsystems, and smartphones will fade into the background. This will lead to the different spheres of our lives being seamlessly integrated, i.e. the time we spend at work, at home or on the go.

2 Mobile imaging

The first cameras appeared in around 1840, and the poor sensitivity of the film used back then made them very bulky indeed. They were anything but portable. Nevertheless, continuous improvements in the sensitivity of photographic film by several orders of magnitude enabled significant reductions in the size of the camera, which went some way toward improving its portability. But the real breakthrough to making them portable came with the launch of roll film in compact cameras. This opened up photography to large swathes of the population. The original Leica camera released in 1924 became the icon of this new technology. And while Leica, ZEISS, and other premium manufacturers turned their attention to sophisticated photography equipment, others set their sights squarely on the mass market – like the affordable Brownie cult camera from Kodak and other low-cost point-and-

shoot cameras that featured plastic lenses as early as the 1950s. Alongside these were 1930s manufacturers like Minox, who advanced camera miniaturization even further [9]. But it was not until mobile phone photography that the market began seeing new developments left, right, and center, and the technology was brought to the masses. 2005 saw a very small number of mobile phones with built-in cameras, and five years later people could not imagine their phones without them. Today, more than half the world's population owns a camera that they almost never go anywhere without.

We have grown accustomed to being able to take photos anytime, anywhere. And now, in 2021, some 2 trillion (2000 billion) photos will likely be taken each year – which is twice as many compared with five years ago, and five times as many compared with 10 years ago. And upwards of 90% of them are taken using a smartphone.

The image quality is actually very similar to that achieved with full-frame single-lens reflexes (SLRs) or system cameras across many photography scenarios. Some photographs can even boast resolutions of over 50 megapixels, and films a resolution of up to 8K (8192×4608 pixels) at frame rates of 48 photo/s, which is much higher than those on modern, high-resolution smartphone displays. Smartphones with 512 GB of memory can hold hundreds of thousands of photos, however, the standard practice is to save a large volume of the images centrally, in a cloud, and then access them via a computer, tablet or other device on demand. Modern smartphones are thinner than 1 cm, weigh less than 200 g and easily fit into our pockets.

A key factor for the success of camera phones – “the camera is fully integrated into the phone” – is something many manufacturers have tried to circumvent time and again, usually for technical reasons, by offering clip-on additions, slide-on lenses, etc. This has resulted in these systems becoming a niche product used by only a small number of people.

And when it comes to the many sensors integrated into a smartphone, in recent years in particular the camera system and 3 dimension (3D) image acquisition have improved in leaps and bounds. In 2016, the dual camera made inroads into the smartphone world with Apple's iPhone 7+. Alongside the option of calculating high-resolution depth maps using two or more cameras, lenses were released with double the focal length compared with the long-established standard wide-angle lens. In recent years, even longer focal lengths as well as ultra-wide-angle lenses have been launched. They are known as “zoom lenses” and can continuously zoom in on subjects by interpolating the camera images. The multicameras

released in recent years, the connection with the integrated sensors, and highly advanced image processing have ensured that the image quality delivered by SPCs matches that achieved with full-frame cameras [10, 11].

2.1 Market development

Around the turn of the millennium, starting with the transition from analog to digital cameras, the entire camera market has continued to grow exponentially. Since 2010, the compact camera market has been shrinking considerably, and by 2020 it had dwindled by more than a factor of 10. Sales of top-quality photo cameras with large image formats (full-frame, APS-C) with interchangeable lenses have fallen by around half (Figure 2). Conversely, sales of camera-equipped smartphones have continued to rise. The quality of SPCs has been increasingly enhanced with high-resolution image sensors, image stabilization, and multi-camera systems.

Figure 3 shows how smartphone sales have developed (cf. Figure 3; data source: [12, 13]. The dashes indicate the smartphones with a built-in camera. From 2002 to 2011, the number of smartphones sold quadrupled. And since 2011, that number has continued to rise each year, but is distinctly more moderate than it was for a few years after its initial release. Today, around 200 times more smartphones are sold per year than photo cameras with interchangeable lenses. A similar trend can be seen in the compact camera segment. Over the past decade, smartphone sales have risen by more than a factor of 20 over those of compact cameras – so SPCs have largely replaced the traditional point-and-shoot variety. This is less true if we consider the ratio of SLR or system cameras with interchangeable lenses



Figure 2: Camera sales since 2004. Data source: ref. [206].

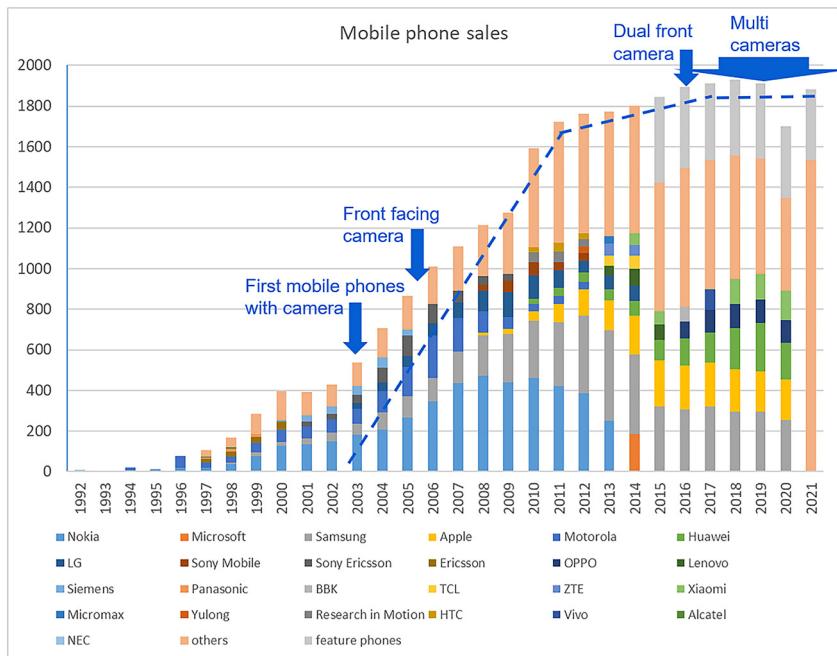


Figure 3: Worldwide smartphone and feature phone sales with breakdown of suppliers (source: Gartner). Since 2015, smartphones and feature phones have been separated in statistics. Feature phones have been added here, and the vast majority of them also have cameras. Estimated number of SPCs shown by blue dashed trend line. For 2021 a forecast is given.

as compared with SPCs. For technological reasons, system and SLR cameras still outdo smartphones in a number of features, which has resulted in a far lower substitution with smartphones for this segment. Here, the smartphone is not a veritable competitor, but rather a welcome addition as the “camera that is always to hand.”

Recently, traditional digital system cameras (DSCs) and system cameras have adopted a number of functions from smartphones (e.g. network connection and digital workflows), which appears to have allowed the segment to somewhat stabilize and recover.

The smartphone business is under tremendous pressure to innovate. The latest features that customers want need to be implemented in just a few months. So the market is highly dynamic. In a few short years, many companies are seeing their revenue grow exponentially, but shrink again just as quickly. Nokia, Motorola, and BlackBerry Limited (previously Research in Motion, and most recently TCL) are examples of such fallen giants.

Today, the major image sensor suppliers are Sony (revenue in this segment approx. \$7 bn), Samsung (\$2.5 bn) and Omnivision (\$1.5 bn). The majority of lens manufacturers is based in China, Taiwan, and Japan, and include

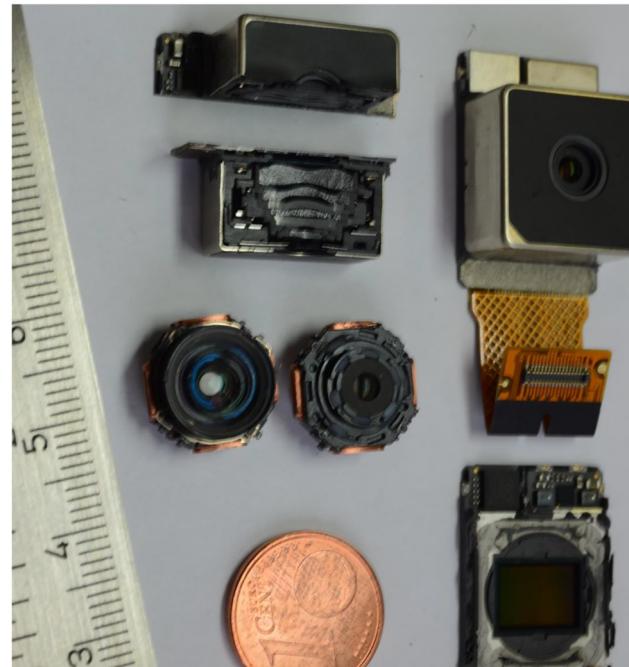


Figure 4: Smartphone camera (SPC) module: Top left: Camera module cross-section, below: Lens barrel including voice-coil motor, bottom right: Image sensor module.

2.2 Supply chain

The majority of smartphone manufacturers does not normally develop and build their own camera modules, which usually comprise an image sensor, lens, actuators, and sensors (Figure 4). They rely on a large network of specialist optics and module developers with whom they specify, develop and then supply new camera modules.

Largan (\$2 bn), Sunny Optical (\$1 bn), and Kantatsu. Integrators (module assembly) include LG Innotek, Semco, and Foxconn (Sharp) (approx. \$3 bn in segment revenue each) (Data from ref. [14]).

The market is growing all the time, buoyed by the trend toward multi-camera systems. While the figure stood at around 1.5 billion camera modules in 2010, this number had almost quadrupled by 2020, to reach 5.5 billion. In addition to being built into smartphones, tablets and laptops, small camera modules are also developed and produced for automobile, security/surveillance, medical, drones, and industrial applications, as well as in robotics. Smartphones hold the majority of the market share, at over 80% [15]. Total global revenue for mobile device camera modules currently stands at around 30 billion dollars. At the same time, the manufacturers have been under tremendous price pressure, which has increased considerably in recent years.

In modern smartphones, the camera system is often the most expensive function in a smartphone [16]. All camera modules, from high-end smartphones together (multicamera system and front-facing camera on screen side), cost between 20 and 80 dollars. That equates to between 6 and 20 dollars per camera. 40% of those costs can be attributed to the image sensor, 30% to the optics incl. image stabilization and autofocus, and 30% to the module (module and assembly).

Besides the rising number of multicamera systems, and those on the front and back of a smartphone, modules are increasingly being integrated for the purposes of human identification (or “biological recognition”): optical fingerprint modules, iris recognition systems, 3D structured light, 3D time-of-flight (ToF) and lidar systems [17] all contain one or several camera imaging modules. They are usually smaller in size and simpler – i.e. with 3 or 4 lens elements – compared to a smartphone’s main camera module, e.g. the standard wide-angle camera on the rear usually contains 6 or 7 lens elements.

3 Brief history and milestones of smartphone imaging technology

In Table 1 we give a chronological list of milestones in mobile phone imaging. One could extend this list indefinitely. The aim is more to give some examples of which developments have had a lasting impact on mobile imaging system evolution than to claim completeness or to

name the actual pioneers. It is not uncommon for developments to be repeated until, after little initial success, they eventually established themselves in the long term. Examples of this include 3D dual cameras from 2011 (LG Optimus 3D, HTC Evo 3D, Sharp Aquos SH80F), some even with a 3D display as a forerunner of the general trend since 2016 in a slightly different form (smaller stereo base and calculation of depth maps instead of stereo images), or the Nokia Lumia 1020 from 2012 with a very large 41 MP image sensor and pixel binning, which has been revived in many high-end camera systems since 2018.

In very simple terms, one can say that we are currently in the third era of smartphone imaging developments: 2002 to around 2010 was the era of the “MegaPixel Race”, and 2008 to around 2016 the era of coupling the camera system to advanced software algorithms and smartphone sensors like the gyroscope or accelerometer (Imaging Apps, Noise Reduction, HDR; Image Stabilization). As of 2016 we have been in the MultiCamera and Computational Imaging Era with 3D sensing heading towards Augmented Reality.

4 Physical properties and requirements of Smartphone photography

We consider practical photographic requirements to derive the basic technical specifications of a SPC. We then compare the physical dimensions of these miniaturized optics with a SLR or system camera for a large-image format.

4.1 Camera form factor and image sensor size

Consumer demand for the smartphone that is “always to hand” means that the camera should be encased in a very flat housing 7–10 mm thick, so it requires a miniaturized camera. Minus the dimensions for the housing and the sensor board, the depth of a cell phone optic must not exceed 5 or 6 mm.

The “relative flatness factor” (r) of the optical design

$$r = \frac{L}{\phi_{\text{im}}} \quad (1)$$

together with the overall length (L) determines the still feasible full image diagonal ϕ_{im} .

Table 1: Milestones in smartphone imaging.

1994	Camera phone patent by Nokia employees ref. [213]
1997	P. Kahn connected a digital camera with his mobile phone and wirelessly shared a picture of his newly born daughter with 2000 people around the world
2000	In November the world's first mobile camera phone is released, the Sharp J-SH04 with 0.11 MP; only distributed in Japan
2001	Launch of third-generation wireless mobile communication standard 3G with UMTS (Universal Mobile Telecommunication System) with data rates of 384 kBit/s enabling picture transfer in mere seconds
2002	Nokia 7650: first mobile phone with camera in Europe: 0.3 MP color camera with first color display Sanyo SPC-5300 first camera mobile phone (0.3 MP) in US market
2003	NTT DoCoMo Mova 505iS: first mobile phone autofocus camera and exceeding 1 MP (1.28 MP)
2003	Sony Ericsson Z1010: first mobile phone with front-facing camera supporting video calls (front and back camera both 0.3 MP)
2005	Nokia N90 with 2 MP camera (ZEISS) with autofocus, LED flash , video; unique swivel design with camcorder feeling
2006	Sony Ericsson K800i with 3.2 MP and Xenon Flash; Nokia N73 also launches 3.2 MP; Nokia launches integration of Flickr to upload photos
2007	Apple's iPhone revolutionizes operating concept via seamless touchscreen and one button only; birth of the “smartphone” ; becomes new standard system layout copied by all other suppliers
2007	Launch of app platform , and open development system iOS Android becomes non Mac platform
2007	Many phone cameras with 5 MP; Samsung SCH-B600 with 10 MP
2009	Consumer mass produced backside-illuminated CMOS from Sony (Exmor R) improves low light performance by a factor of two
2010	Apple's iPhone 4 with gyroscope for precise orientation determination
2012	Nokia PureView 808 followed by Lumia 1020 feature 41 MegaPixel Sensor with Pixel Binning to 5 MP “ high-end digital zoom ”
2012	Nokia N920 first smartphone featuring optical image stabilization
2016	Rise of Dual Cameras (beginning of multicamera systems in smartphones): iPhone 7+ with different focal lengths: 28 mm, and 56 mm equivalent focal length for hybrid zoom and portrait mode Huawei P9 (Leica) with identical FOV, RGB/monochrome sensor for increased resolution, lower noise and portrait mode LG G5: Standard wide angle combined with extreme wide angle (135° FOV) for hybrid wide-angle zoom
2016	Samsung Galaxy S7 with Dual Pixel Autofocus
2017	HEIF compression format in Apple iPhone 8 + halves space required by pictures and film
2017	Samsung Galaxy Note 8 dual cam with synchronized optical image stabilization for both cameras
2017	Sony Xperia ZX1 with Super Slow-Motion Video > 1000 frames/s
2017	LG V30 with real Color Grading (instead of simple color filters)
2017	iPhone X with 3D structured light depth sensor , e.g. for face recognition
2018	Huawei Honor View 20 3D Time-of-Flight depth sensor
2019	Huawei P30 Pro and OPPO Reno 10× feature 120 mm periscopic tele lens to extend hybrid zoom range
2020	Apple iPhone 12 and other flagship smartphones released with wireless communication standard 5G

The image sensor should be as large as possible so that as much light as possible falls on a pixel. This reduces fundamental disadvantages such as image noise, a reduced dynamic range or longer exposure times and thus motion blur. The factor r depends essentially on the field of view (FOV) of the lens and its layout (standard upright, periscopic, etc.), which we will discuss in more detail in Section 6. For wide-angle lenses, $r = 0.83$ is typical, which means that an image diagonal of 6 mm is obtained with an overall length of 5 mm. For example, the image sensor of the Apple iPhone 6 has a sensor diagonal of 6 mm. With a slightly larger construction depth of around 8 mm, i.e. with slightly thicker smartphones or mostly a slightly protruding camera housing, as well as complex 7-lens designs, for which $r = 0.65$ is possible, an image diagonal of more than 12 mm can be achieved. Such image sensor sizes of the main camera (standard wide-angle) are integrated into

several high-end smartphone models (e.g. Huawei P40, Xiaomi Mi 10+, vivo X60 Pro +).

4.2 Image sensor resolution

Smartphones typically offer a standard image resolution of 12 megapixels (12 MP). SPC image sensors have an aspect ratio of 4:3, in contrast to the aspect ratio of 3:2 for full-format ($36 \times 24 \text{ mm}^2$) or the APS-C format (approx. $23.6 \times 15.7 \text{ mm}$, depending on camera manufacturer). The more square format is a compromise so that a picture taken in landscape format does not appear too narrow when viewed in portrait format.

The image sensor formats are usually referred to as “inch values.” The aforementioned image sensor that measures 6 mm in the diagonal would be referred to as

1/3". Unfortunately, the conversion of the unit 1" = 2.54 mm does not itself lead to the actual dimensions of the sensor: $25.4 \text{ mm} \cdot 1/3 = 8.3 \text{ mm}$ does not correspond to any of the actual side dimensions, i.e. 6 mm in the diagonal, 4.8 mm in width, or 3.6 mm in height. The inch specification was taken from old Vidicon video tubes from the 1950s and corresponded to the outer glass diameter of the photoelectric front surface. The sensor diagonal corresponds to about 2/3 of the inch value, but only roughly, and sensor sizes are unfortunately not exactly proportionally scalable according to the inch values. In other words, the absolute inch sizes are antiquated and misleading. So it is better to determine the exact sensor dimensions in millimeters.

Mobile phone camera image sensors were developed between 2002 and 2010, raising pixel counts from 0.3 MP to around 12 MP. Keeping the same image sensor size of about 6 mm in the diagonal, this corresponds to a reduction in pixel pitch from about 6 μm down to 1 μm . The pixel race had come to an end by around 2012 and even began partly moving in reverse in the following years since many SPC suppliers had recognized that pixel amounts far greater than 12 MP are not beneficial due to practically unused resolution. What is more, handling with reasonable amount of image data and better image noise and dynamic range were consequences of the reversed pixel race [18].

Nonetheless, it should be mentioned that as of 2018 an increasing number of main cameras in modern SPCs have featured image sensors with a high number of pixels: 48 MP, 64 MP or even 108 MP. This suggests a very high resolution, especially, since these sensors are often referred to as "high resolution" in product marketing. However, this is misleading: The arrangement of the color pixels is usually not made up of adjacent Bayer patterns. The number of pixels in the final image is typically around 12 MP as standard, which is also effectively supported by the optical performance of the lens. Most of these image sensors are not standard Bayer sensors, but rather "multicell sensors" as they are called by the

manufacturers, under the names "Tetracell technology" (Samsung), "Quad-Bayer" (Sony), and "4-Cell" (OmniVision). These are arrangements with pixel clusters or "macro-pixels" of 4 or 9 pixels (Figure 5). The standard output images are $48 \text{ MP}/4 = 12 \text{ MP}$ for a typical 48 MP sensor with a 4-pixel cluster or $108 \text{ MP}/9 = 12 \text{ MP}$ for the 9-pixel cluster at 108 MP (e.g. in the Oppo Reno 3). Macropixels do of course have enhanced light sensitivity compared to a single pixel. This could also be obtained by just using larger pixels. However the benefit of using a pixel cluster is flexibility and additional advanced features. The image data of the multicell sensors can be read out in various ways: By selecting different sensitivities or exposure times of the pixels, the dynamic range can be increased (see Section 14) or the noise can be reduced through pixel binning, or a very high-resolution image can be output. A clear distinction should therefore be made between sensor pixels, i.e. the pixels available when recording, and the number of pixels used to display the images.

Figures 6 and 7 show the trends of pixel numbers of main rear camera and front camera. In the years after 2016, pixel shrinking was pushed further down, reaching 0.7 μm by 2020: First, so that the additional cameras of different focal lengths that are being launched can achieve the desired resolution of 8 or 12 MP (standard Bayer Pattern). Second, the aforementioned multi-cell sensors of the main camera: Mastering the production of the image sensors of such small pixels in such high volumes through mass production with a high yield is an additional challenge compared to the small sensors with the same pixel pitch.

Even the "typical image resolution of 12 MP" in today's images captured using a smartphone is seldom exhausted in terms of the physiological limits of the human eye under typical viewing conditions. Most of the images are viewed directly on the smartphone and rarely exceed the size of a PC monitor or TV. With the resolution of the eye of around $\theta_{\text{res}} = 0.3 \text{ mrad}$ or 1 arc s [19], the result for the typical

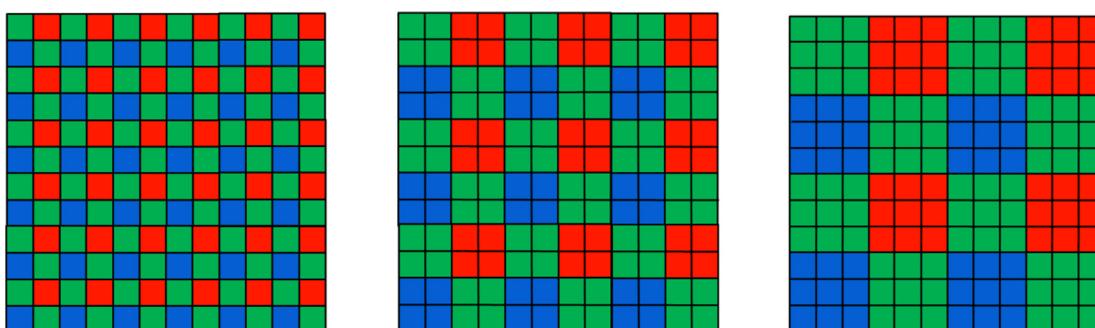


Figure 5: Image sensor architectures: Standard Bayer pattern and multi-cell configurations.

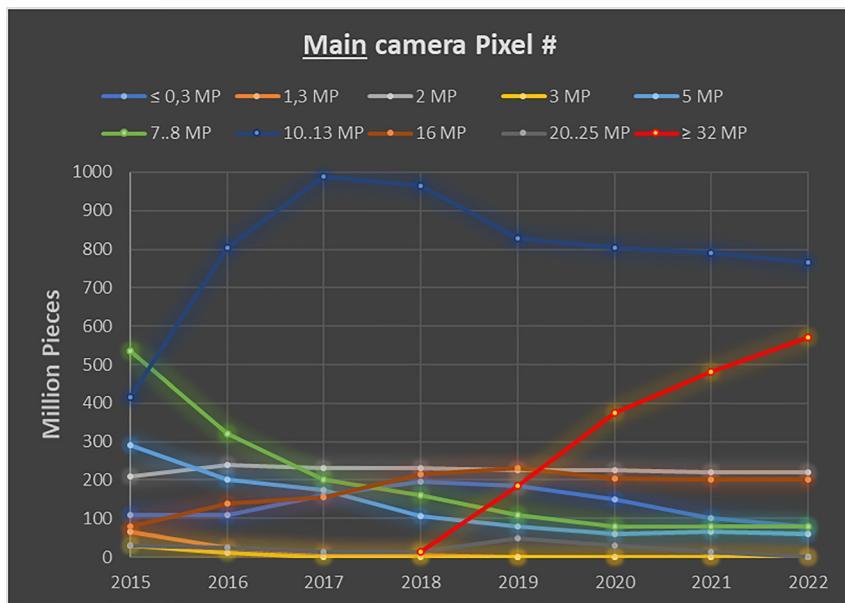


Figure 6: Main camera pixel count trend: As of 2016 most cameras have had about 12 MP (10–13 MP line). Currently >32 MP sensors are increasingly used: These are mainly Quad-Bayer pattern sensors, so while the spatial resolution does not increase, they do offer improved HDR capabilities.

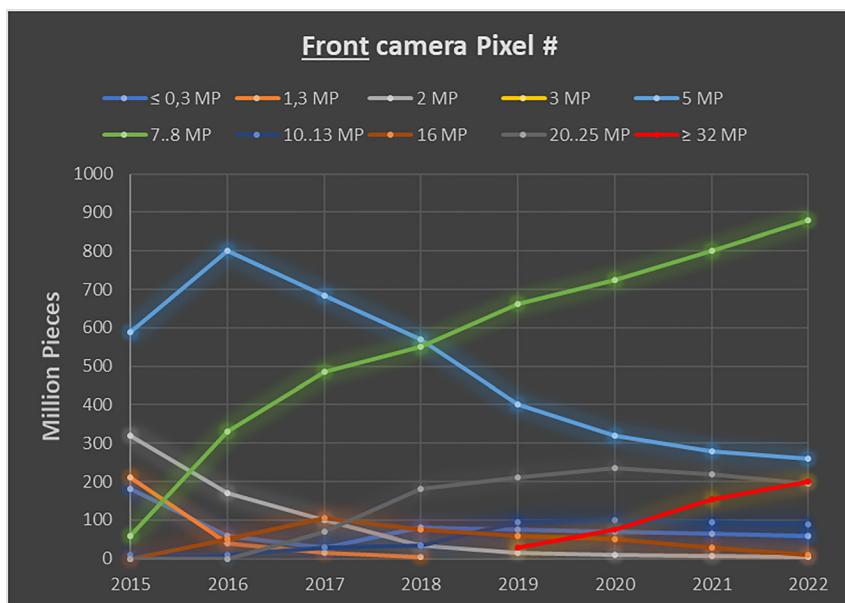


Figure 7: Corresponds to Figure 6, but regarding the front camera. The main trend of 5–8 MP cameras is being supplemented by > 32 MP sensors (mainly larger Quad-Bayer pattern, which improves HDR and other properties).

observer distance of around $s = 25\text{--}40 \text{ cm}$ under optimal conditions (perfectly still image) produces an object resolution that is just barely distinguishable, of:

$$\Delta x_{\text{res}} = s \cdot \theta_{\text{res}} \approx 330 \text{ mm} \cdot 0.3 \text{ mrad} = 0.1 \text{ mm.} \quad (2)$$

This corresponds to a smartphone display measuring $140 \times 70 \text{ mm}$, so even with a large 6.2". display, only 1400×700 pixels can be effectively resolved. With optimal viewing conditions – i.e. bright ambient light and a completely stationary smartphone without hand tremors, and even closer eye relief to the display – at best no more

than 2000×1000 pixels are required. Experimental studies confirm this (e.g. [20]). The resolution of smartphone displays corresponds roughly to this HD resolution, typically around 2300×1080 pixels; higher-resolution displays like that of the Sony Xperia with 3840×1644 pixels are the absolute exception here. Significantly more pixels are only necessary if you want to create a poster from the picture and look at it up close or if you enlarge image sections. This also applies to VR applications in which scenes on the smartphone screen are viewed significantly enlarged “under a magnifying glass” (Samsung Gear VR, ZEISS VR One, etc.).

Such observations on the screen – which are enlarged by a factor of 3 or separately on a screen with a 4K projector – benefit from a resolution of beyond 12 MP.

4.3 Optical resolution and required aperture

In order for the resolution of the image sensor to be used, the quality of the optical image as viewed through the lens must be good enough. Let us now consider the number of pixels of the image sensor and the resolution required for the optical image. Let us continue with the example of the image sensor measuring 6 mm in the diagonal and 12 MP: The side lengths in the aspect ratio 4:3 are 4.8×3.6 mm. At 12 MP, i.e. 4000×3000 pixels, this corresponds to a pixel pitch of 1.2 μm.

The point spread function diameter produced by the lens must be close to the pixel pitch of 1.2 μm in order not to limit the performance of the optics. The diameter of the Airy diffraction disk of the ideal image is

$$\phi_{\text{Airy}} = 2.44 \lambda K \quad (3)$$

where λ denotes the wavelength, that is approx. 0.4–0.7 μm for the visible light range, and the *f*-number K is related by $K = 1/(2NA')$ to the numerical aperture NA' on the image. An appropriate relationship between the diameter of the Airy spot ϕ_{Airy} and the sensor pixel pitch p is to choose

$$\phi_{\text{Airy}} = 2 \cdot p \quad (4)$$

Then, a significant portion of the light distribution, namely a relative encircled energy of 0.73, is inside a square-shaped pixel. Further, if we consider that the intensity is transferred to a grey value distribution by the photo conversion curve and opto-electronic conversion function, the “effective encircled brightness” to give a name to the gray value distribution, is in fact >0.8. We set the corresponding “critical *f*-number” K_{crit} as a requirement for the lens:

$$K_{\text{crit}} = \frac{p}{1.22 \cdot \lambda} \quad (5)$$

This would be $K_{\text{crit}} = 1.8$ (or the aperture ratio $f/1.8$) with the pixel pitch $p = 1.2$ μm. Most of today’s standard wide-angle lenses for smartphones meet this requirement. However, “telephoto lenses” (normal focal lengths and short/long portrait focal lengths) taken using modern dual or multiple camera systems fall below this requirement, with *f*-stops of 2.4 or more and often even smaller pixels of usually less than 1 μm, e.g. $p = 0.8$ μm.

With SPCs, and all digital cameras in general, the image quality is specified via the contrast of (quasi) periodic structures, which in turn is related to the modulation transfer function (MTF). The concept of this linear transfer function opens up the possibility of combining the components of the digital imaging chain (optics, sensor, and image processing – as long as they are not correlated) in order to calculate the overall contrast transferred (for a comprehensive analysis see ref. [21]). From the pixel pitch of the image sensor we have:

$$\text{Nyq} = \frac{1}{2p} \quad (6)$$

which is the Nyquist frequency (Nyq) in the horizontal and vertical direction; in our example, $p = 1.2$ μm, we have $\text{Nyq} = \frac{1000 \frac{\mu\text{m}}{\text{mm}}}{2 \cdot 1.2 \mu\text{m}} \approx 416 \frac{\text{lp}}{\text{mm}}$. This is the smallest period that the image sensor can in principle still resolve or sample as such. However, this only applies to a limited extent: As shown in Figure 8, a periodic intensity distribution close to the Nyq is mapped with a vastly differing contrast depending on the exact position relative to the pixel grid; in the worst case even with no contrast at all. For this reason, it does not make sense to specify the image performance for Nyq itself, but rather starting at around $\text{Nyq}/2$ and higher spatial frequencies, where the deviations of the integrated sampled signal from the original signal become increasingly smaller (Figure 9). The system contrast is often displayed simultaneously with different fine structure periods, such as $\text{Nyq}/8$, $\text{Nyq}/4$, $\text{Nyq}/2$. For $p = 1.2$ μm, these would be the spatial frequencies 52, 104, and 208 lp/mm.

The product of the optics MTF and sensor MTF has an increasingly statistical character near the Nyq because of the position dependence.

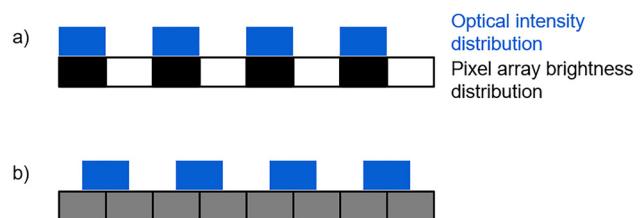


Figure 8: Illustration of capturing a periodic optical intensity distribution with a periodic pixel array. The periodicity of the intensity corresponds to the Nyq. The grid period is exactly 2 pixels here. The periodic intensity distribution (blue) generates a signal on the pixel array, which is represented here by the gray values of the individual boxes (“white pixels”: strong light signal, “black pixels”: no light). In (a) the pixel array distribution perfectly matches the intensity distribution. (b) A structure shift of half a pixel results in a completely homogeneous distribution on the image sensor.

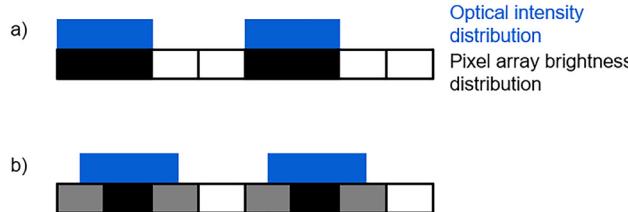


Figure 9: Same as Figure 8, but spatial frequency is $Nyq/2$.

In contrast to most full-frame reflex or system cameras, a SPC does not contain an optical low-pass filter, which suppresses the moiré effect, i.e. disruptive low-frequency beats caused by periodic structures, e.g. a finely checked shirt, which are insufficiently scanned by the equidistant pixel grid. This is due, first, to the fact that optical low-pass filters are usually implemented using birefringent structures with a thickness of around 2 or 3 mm, which is unacceptable for SPCs for reasons of space, and which are also relatively expensive, and second, because the point spread function of the SPC lens is already larger than a pixel and therefore has a low-pass effect.

An ideal incoherent optical system transfers information up to a limited spatial frequency of

$$\nu_{\max} = 2 \frac{NA'}{\lambda} = \frac{1}{\lambda K} \quad (7)$$

For $f/1.8$ and green light of wavelength $\lambda = 0.55 \mu\text{m}$, the transferred spatial frequency limit is $\nu_{\max} = \frac{1}{0.55 \mu\text{m} \cdot 1.8} \approx 1000 \text{ lp/mm}$. The contrast of a periodic signal decreases according to the MTF of the optical system with increasing spatial frequency. With ideal incoherent imaging with a homogeneous circular pupil corresponding to ref. [22]

$$\begin{aligned} MTF_{\text{ideal}}(\nu) &= \begin{cases} \frac{2}{\pi} \left[\arccos\left(\frac{\nu}{\nu_{\max}}\right) - \frac{\nu}{\nu_{\max} \sqrt{1 - \left(\frac{\nu}{\nu_{\max}}\right)^2}} \right], & \nu \leq 2\nu_{\max} \\ 0, & \text{else} \end{cases} \end{aligned} \quad (8)$$

This function falls in a monotonous, almost linear way, towards larger spatial frequencies (Figure 10). The same figure also shows the contrast for spatial frequencies relevant for SPC (pixel pitch of 0.8 or 1.2 μm , corresponding to about 200 or 300 lp/mm) compared to the finest structures typically observed with full-frame cameras (40 lp/mm). The corresponding number of pixels on the image sensor is similar for these values, approx. 12 MP. This makes it clear that SPCs are physically limited by their size alone. They

are “diffraction-limited.” This means that with the same image resolution (relative to the respective pixel size) and the same lens aperture with an ideal aberration-free lens, the contrast is weaker.

The aberrations of SPC lenses are so small that stopping down would lead to a weaker contrast. Therefore, and to minimize the complexity by dispensing with moving parts, SPC lenses are not even equipped with iris diaphragms. The exposure is adjusted solely via the exposure time and the International Organization for Standardization (ISO) sensitivity via the read-out amplifier on the image sensor. But with full-frame lenses, the contrast is usually weaker, with a more open aperture than when you stop down. Often the maximum contrast is obtained at around $f/4$, $f/5.6$, or $f/8$, before the contrast decreases with further stopping down to the diffraction limit. The loss of contrast with an open aperture is due to the fact that larger aberrations are allowed in favor of a simpler, more compact design. Exceptions are premium lenses such as the ZEISS Otus or SIGMA Art series, which already have maximum contrast at $f/1.4$ and $f/2$, but at the expense of a significantly more complex optical design, size, and weight.

Resolution is not independent from signal-to-noise ratio of the image sensor; a more rigorous analysis can be found in ref. [23].

4.4 Portrait photography: Perspective, bokeh, and depth of field

A popular “allrounder lens” is the classic standard wide-angle lens with a focal length of 35 mm for the full-format (abbreviated “ff”) $36 \times 24 \text{ mm}^2$, i.e. a full image diagonal of $\phi_{\text{im, ff}} = \sqrt{36^2 + 24^2} \text{ mm} = 43.3 \text{ mm}$. An equivalent focal length of 35 mm was also typical in the early days of smartphone imaging, but this changed after a couple of years to about 28 mm (more detailed data are given in Section 6.1). This corresponds to a full diagonal field of view (FOV), related to focal length and image field size by

$$\tan(\text{FOV}/2) = \frac{\phi_{\text{im, ff}}/2}{f'} \quad (9)$$

of $\text{FOV} = 2 \arctan(\phi_{\text{im, ff}}/(2f')) = 2\arctan(43.3 \text{ mm}/56 \text{ mm}) \approx 75^\circ$. The 28 mm focal length equivalent for the image sensor with a 6 mm image diagonal is an actual focal length of the SPC of

$$f' = \frac{f'_{\text{eq}}}{\phi_{\text{im, ff}}} \phi_{\text{im, SPC}} \quad (10)$$

so $(28 \text{ mm}/43.3 \text{ mm}) \times 6 \text{ mm} = 3.9 \text{ mm}$ or about 4 mm.

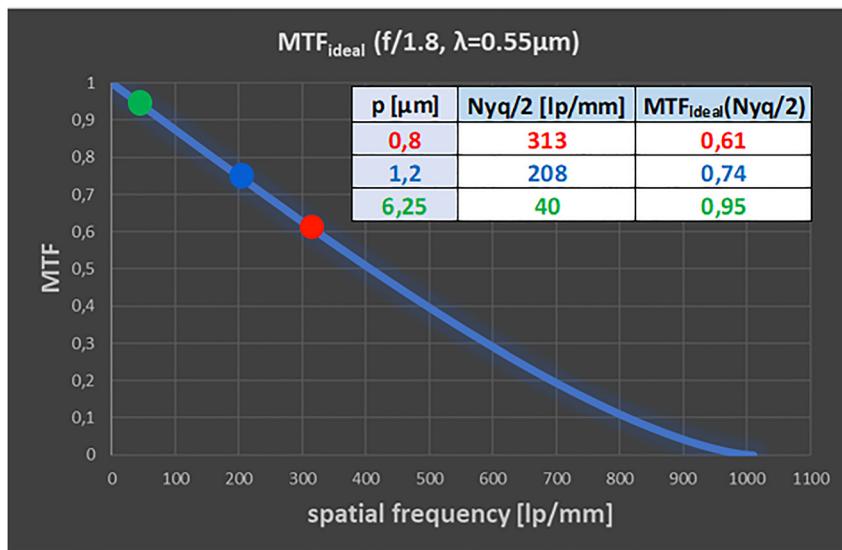


Figure 10: MTF of ideal lens. The achievable contrast in the Nyquist/2 spatial frequency is smaller for miniature SPC systems.

For a “close-up portrait”, in which a face (height from chin to crown approx. 30 cm) almost fills the vertical image field, with a 36 cm vertical section, the distance to the object (measured from the entrance pupil) is about

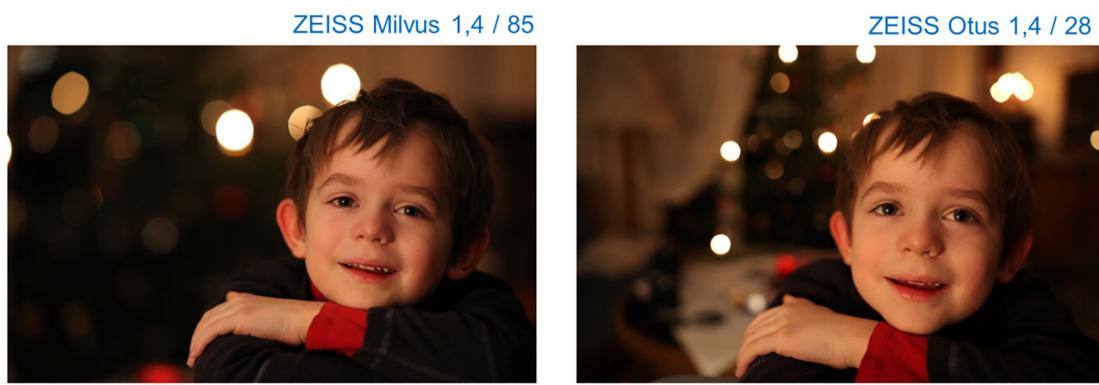
$$s = \frac{2y_{\text{ob}}}{2y_{\text{im, SPC}}} f \quad (11)$$

where $2y_{\text{im}}$ is the total length of the side of the vertical, that is the “short direction” of the $4.8 \times 3.6 \text{ mm}$ image sensor. The object distance is $s = 360 \text{ mm} / 3.6 \text{ mm} \cdot 4 \text{ mm} = 400 \text{ mm}$, i.e. 0.4 m.

Since 0.4 m is also a typical distance at which one holds a smartphone, including for video calls or selfies, one can often choose an equivalent focal length of 28 mm for the front camera.

Typical normal portrait distances with a vertical object side length of 0.72 m are about twice as large, i.e. 0.8 m. To have people completely in the picture, $y_{\text{ob}} = 2.16 \text{ m}$, again about three times as much, i.e. 2.4 m object distance. So you can still see a lot in the picture in most cramped situations such as indoors, in a group photo or when dining at the same table with others.

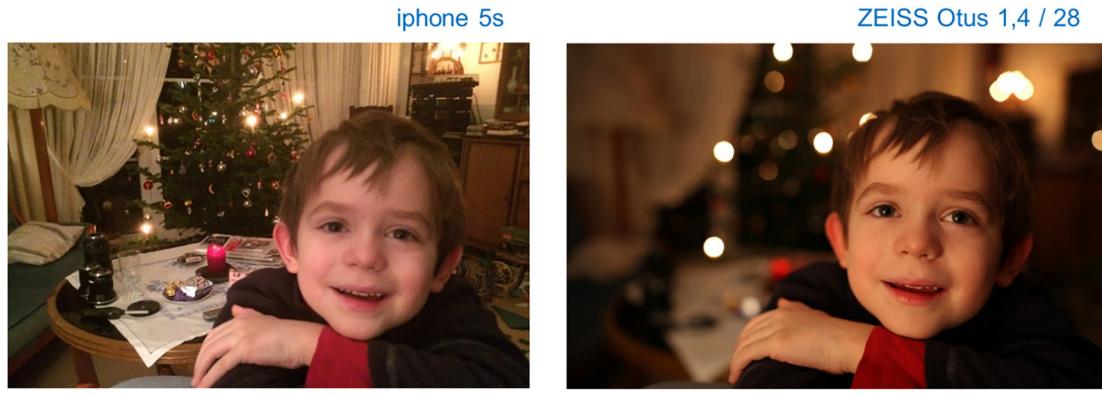
For reasons of perspective, wide-angle lenses are not well suited for portraits: To fully depict a person, one has to get very close up, at $f_{\text{eq}} = 28 \text{ mm}$ – the aforementioned 0.4 m. Then the nose will be 10–20% of the object distance in front of the ears, so it is imaged magnified and leads to deformations on the face shown. Classic portrait focal lengths have an equivalent focal length of approx. 85 mm that is about three times longer. A face looks much more pleasant in a perspective with a portrait lens (see Figure 11).



$f = 85\text{mm}$, diagonal field of view 28° , aperture $f/2.2$
Full frame sensor ($36 \times 24\text{mm}$), focus dist. 1.4m

$f = 28\text{mm}$, diagonal field of view 75° , aperture $f/2.2$
Full frame sensor ($36 \times 24\text{mm}$), focus dist. 0.45m

Figure 11: Portrait with “portrait lens” of $f_{\text{eq}} = 85 \text{ mm}$ (left) shows much less “perspective deformation” compared to a $f_{\text{eq}} = 28 \text{ mm}$ wide-angle lens (right).



$f = 4\text{mm}$, diagonal field-of-view 75° , aperture $f/2.2$
Image sensor diameter 6mm (smartphone camera)

$f = 28\text{mm}$, diagonal field-of-view 75° , aperture $f/2.2$
Image sensor diameter 43.3mm (36mm x 24mm)

Figure 12: Image taken with a smartphone (left) and a DSLR (right) with a lens with the same field of view (FOV) ($f_{eq} = 28\text{ mm}$ or FOV 75°) and same f -number ($f/2.2$).

The pictures in Figure 11 were shot with a full-frame digital single-lens reflex (DSLR): At a large aperture (here $f/2.2$) the person is detached from the background due to a shallow depth of field. A direct comparison of the two lenses in Figure 11 shows that a light source in the background, the defocused point spread function, is larger with a longer focal length. It can be shown that the spot diameter with the same f -number and the same image format scales approximately in the ratio of the focal lengths.

With SPC lenses, however, the depth of field is large. The whole scene, including the background, looks sharp (see Figure 12).

The size of the point spread in the depth – and from this the depth of field – can be calculated using a geometrical model: According to Figure 13 the variables s , s' denote the object and image distance with respect to an out-of-focus object point, relative to the entrance and exit pupil of the optical system, respectively; s_F and s'_F denote the corresponding distances for which the lens is focused, and f the focal length.

By comparing the triangles in the image space, the ratio of the radius of the defocused point image r_{spot} to the position of the defocus in the image space ($s'_F - s'$) is obtained:

$$\frac{r_{\text{spot}}}{|s' - s'_F|} = \frac{\emptyset_{AP}}{2s'_F} = \frac{1}{2K} \quad (12)$$

The spot radius is thus

$$r_{\text{spot}} = \frac{|s' - s'_F|}{2K}. \quad (13)$$

For an optical system, represented by its entrance and exit pupils, the focusing conditions are from ref. [24], where m_p denotes the pupil magnification:

$$-\frac{1}{m_p s} + \frac{m_p}{s'} = \frac{1}{f} \quad (14)$$

and

$$-\frac{1}{m_p s_F} + \frac{m_p}{s'_F} = \frac{1}{f} \quad (15)$$

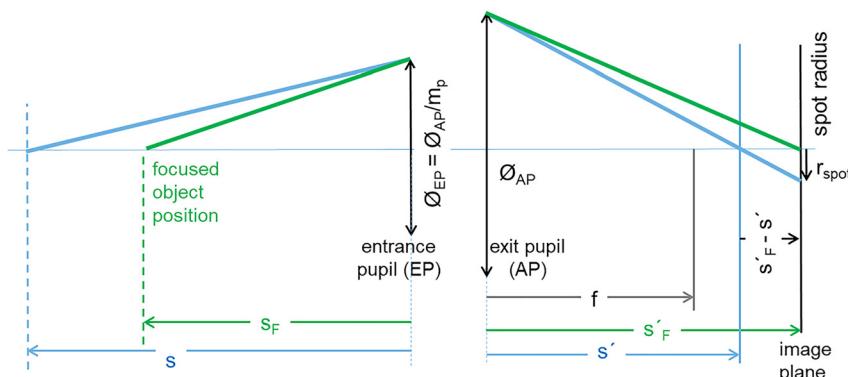


Figure 13: Lens represented as black box by entrance and exit pupil and definition of parameters with respect to focus (F) and out-of-focus (not indexed) distances to calculate spot diameter $2r_{\text{spot}}$.

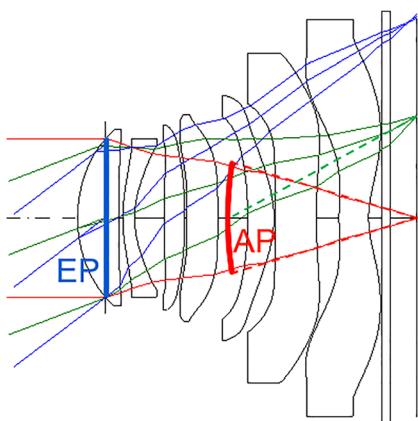


Figure 14: Entrance pupil (EP) and exit pupil (AP) of a standard wide-angle smartphone lens (drawn with respect to on-axis imaging). The position results from the intersection of the (extended) chief ray in the object or image space with the optical axis; the size of the pupil in each case is obtained by lengthening the marginal rays in the object or image space up to this position.

respectively. Replacing the image distances s' , s'_F in Eq. (13) by object distances s , s_F with Eqs. (14) and (15) yields:

$$r_{\text{spot}} = \frac{f^2}{2K} \frac{|s - s_F|}{(f/m_p + s)(f/m_p + s_F)} \quad (16)$$

Pupil magnification m_p depends on the specific optical design. For wide-angle SPC lenses (Figure 14) the value of m_p is typically between 0.5 and 1.

If the object distances are much larger than the focal length, which applies to SPC lenses, since s is at least about 20 times as large as f , then the term f/m_p in the denominator of Eq. (16) can be neglected. The diameter of the image's circle of confusion $\mathcal{O}_{\text{spot}} = 2r_{\text{spot}}$ is then:

$$\mathcal{O}_{\text{spot}} = \frac{f^2}{K} \frac{|s_F - s|}{s_F s} \quad (17)$$

In order to compare the imaging of lenses with different image formats, we define the relative circle of confusion (based on the image diagonal) instead of the absolute circle of confusion in units of length:

$$\mathcal{O}_{\text{rel.spot}} = \frac{\mathcal{O}_{\text{spot}}}{\mathcal{O}_{\text{im}}} \quad (18)$$

This value directly indicates the spot size in the defocused area as it appears in the photo.

For a very distant background, the relative circle of confusion will be:



Figure 15: Portrait: Proportions and sizes.

$$\begin{aligned} &\mathcal{O}_{\text{rel.spot}, \infty} \\ &= \lim_{s \rightarrow \infty} \frac{f^2}{\mathcal{O}_{\text{im}} K} \frac{|s - s_F|}{(f/m_p + s)(f/m_p + s_F)} = \frac{f^2}{\mathcal{O}_{\text{im}} K s_F} \end{aligned} \quad (19)$$

With the definition of the f -number, $K = \frac{f}{\mathcal{O}_{\text{EP}}}$, and the magnification (approximated for large object distances) $m = \frac{f}{s_F + f} \approx \frac{f}{s_F}$ we obtain:

$$\mathcal{O}_{\text{rel.spot}, \infty} = \frac{\mathcal{O}_{\text{EP}}}{\mathcal{O}_{\text{im}}} m \quad (20)$$

For a portrait motif as in Figure 15, the diameter of the object field, i.e. the distance between opposite corners, is around 700 mm (more precisely 666 mm for 4:3 format, 722 mm for 3:2 format):

$$\mathcal{O}_{\text{ob, Portrait}} = 700 \text{ mm} = \frac{\mathcal{O}_{\text{im}}}{m} \quad (21)$$

Inserting (21) in (20) yields that the relative diameter of the circle of confusion only depends on the diameter of the entrance pupil:

$$\mathcal{O}_{\text{rel.spot}, \infty} = \frac{\mathcal{O}_{\text{EP}}}{\mathcal{O}_{\text{ob, Portrait}}} = \frac{\mathcal{O}_{\text{EP}}}{700 \text{ mm}} \quad (22)$$

With this formula you can immediately estimate the size of the highlights far in the background for our standardized portrait situation: With a long full-frame portrait lens of 2/135 mm, the diameter of the entrance pupil is $\mathcal{O}_{\text{EP}} = \frac{f}{K} = \frac{135 \text{ mm}}{2} = 67.5 \text{ mm}$, and thus the relative circle of confusion $67.5 \text{ mm}/700 \text{ mm}$, i.e. about 10%.

With a wide-angle SPC lens, the entrance pupil is only $\mathcal{O}_{\text{EP}} = \frac{f}{K} = \frac{4 \text{ mm}}{1.7} = 2.3 \text{ mm}$ and the relative circle of confusion is only 0.3%.

Inserting the expression

$$\mathcal{D}_{EP} = \frac{f}{K} = \frac{\mathcal{O}_{im}}{2K\tan(\text{FOV}/2)} \quad (23)$$

and replacing $\text{NA}' = \frac{1}{2K}$ and using the abbreviation $c_{\text{FOV}} = \frac{1}{\tan(\text{FOV}/2)}$ in (22) yields:

$$\mathcal{D}_{\text{rel.spot},\infty} = \frac{c_{\text{FOV}}}{700 \text{ mm}} \mathcal{O}_{im} \text{NA}' \quad (24)$$

For a given FOV, the relative diameter of the circle of confusion is therefore directly proportional to the product of the image field diameter and the numerical aperture of the image. For a constant transmission within the pupil and field, the product of field area and squared NA,

$$G = \frac{\pi}{4} \mathcal{O}_{im}^2 \text{NA}'^2 \quad (25)$$

is the optical systems etendue. Other names for etendue are “throughput,” “collecting power,” or “ $A\Omega$ product.” This means that the relative circle of confusion is proportional to the root of the etendue value:

$$\mathcal{D}_{\text{rel.spot},\infty} = \frac{c_{\text{FOV}}}{700 \text{ mm}} \sqrt{G} \quad (26)$$

In other words: To achieve the same background circle of confusion, the etendue of the small lens must be equal to that of a full-format camera lens. For an image diameter that is $7\times$ smaller, the NA would have to be a factor of $7\times$ larger. Of course, this is no longer possible compared to high-aperture camera lenses, e.g. with $f/1.4$, i.e. $\text{NA}' = 1/(2 \cdot 1.4) = 0.36$, the SPC lens would have to be $f/0.2$ or $\text{NA}' = 2.5$.

Depth-of-field formulas are obtained from the equations for the size of the geometric spot (circle of confusion) if a threshold value $\mathcal{O}_{\text{thres}}$ is defined for the spot size for which the image still appears “sharp” and resolved according to the object distances [25, 26].

Specifically, one obtains the hyperfocal distance to

$$s_{F,\text{hyp}} = \frac{f^2}{K\mathcal{O}_{\text{thres}}} + f \quad (27)$$

With regard to this focus distance, the image is sharp from $s_{F,\text{hyp}}/2$ to infinity. For many normal viewing conditions, the threshold value of the circle of confusion is chosen as $1/1500$ of the image diagonal: $\mathcal{O}_{\text{thres}} = \mathcal{O}_{im}/1500$. With $\tan\left(\frac{\text{FOV}}{2}\right) = \frac{\mathcal{O}_{im}}{2f}$ we obtain:

$$s_{F,\text{hyp}} = \frac{f}{2K \frac{1}{1500} \tan\left(\frac{\text{FOV}}{2}\right)} + f \quad (28)$$

The hyperfocal distance of two lenses with the same FOV but different focal lengths, i.e. lenses with the same equivalent focal length, scales directly with the focal length (or with the sensor format). For an SPC standard wide-angle lens with a focal length of 4 mm and an aperture of $f/2$ for a sensor with a diagonal $\mathcal{O}_{im} = 6 \text{ mm}$, $s_{F,\text{hyp}} = 2 \text{ m}$, i.e. the image is sharp from 1 m to infinity. With the equivalent full-frame lens, on the other hand, from 7 m to infinity. The autofocus is therefore only required for close range, e.g. to photograph documents. The typical close-range distance of standard wide-angle lenses in smartphones is approx. 80–100 mm with a magnification of approx. 1:20. The object area of $96 \times 72 \text{ mm}$ shown on a $4.8 \times 3.6 \text{ mm}$ is only slightly larger than a standard $85 \times 55 \text{ mm}$ business card. The depth of field is then only about $\pm 3 \text{ mm}$. An autofocus with the appropriate accuracy is required to achieve this (see Section 12).

The physical – and for the photographer, the creative – limitations imposed by the large depth of field are overcome in modern smartphones in “portrait mode” with the depth determined stereographically and/or with high-resolution 3D sensors via the image. And then, according to the depth distance from the focal plane, the image is “calculated out of focus” (see Section 15).

4.5 Étendue and photographic exposure

In the previous Section 4.4 we established a direct connection between the bokeh and the etendue. Étendue is very important for photography because it defines the exposure control. In addition to the amount of light described by the etendue that reaches the image plane from the object space, the photographic exposure H is controlled by the sensitivity of the film or the image sensor and the exposure time T . Exposure Control in practical photography is described as follows:

$$H \sim \text{ISO} \times \left(\frac{1}{K}\right)^2 \times T \quad (29)$$

“ISO” refers to the sensitivity of the image sensor. Essentially, the sensitivity depends on the “area of a pixel” because this determines how many photons strike each time. For a 4:3 aspect ratio sensor the horizontal side length

of the image sensor is $\frac{4}{\sqrt{4^2+3^2}}\phi_{\text{im}} = \frac{4}{5}\phi_{\text{im}} = 0.8\phi_{\text{im}}$ and the short side length correspondingly $\frac{3}{\sqrt{4^2+3^2}}\phi_{\text{im}} = 0.6\phi_{\text{im}}$, such that the total sensor area is $0.48 \cdot \phi_{\text{im}}^2$. Division by the total number of pixels yields the surface area of a pixel to $0.48 \cdot \phi_{\text{im}}^2 / (\# \text{Pixel})$. The ISO value also indicates how efficiently a pixel absorbs light and converts it into an electrical voltage (see Section 9). The desired exposure time in an SPC is achieved by pure electronic control of the image sensor, while in larger-format cameras a mechanical shutter is usually available, sometimes integrated with the iris aperture, and combined with electronic read-out of the image sensor. A high level of sensitivity enables a shorter exposure time, which means that it is useful for capturing fast-moving subjects. On the other hand, high sensitivity leads to increased image noise.

In contrast to traditional photography, an SPC does not use an iris to vary the *f*-number, because this would lead to loss of resolution (see previous Section 4.3).

A classical photography rule of thumb is the “sunny 16 rule”: On a sunny day with an ISO100 film and *f*/16 stop the required exposure time is about 1/100 s.

All three contributions are scaled logarithmically to base 2, i.e. ISO100, ISO200, ISO400, ... or the *f*-stop numbers according to the square root of 2 between the *f*-stops, i.e. $K = 2, 2.8, 4, \dots$ etc. so that $1/K^2$ doubles per step and finally for the exposure time $T = 1/100, 1/50, 1/25, \dots$ etc. Then it is easy to deduce from the “sunny 16 rule” that an increase of, say, 3 *f*-stops in aperture to *f*/5.6 enables a $2^3 = 8$ -times shorter exposure – that is, 1/800 s – or an ISO of 50 and 1/400 s.

Meanwhile, for an SPC the number of photons per pixel is inversely proportional to the square of the crop factor, c^2 , assuming both the full-frame camera and SPC have an equal number of pixels. For a crop factor of 7, an SPC pixel therefore receives about 50 times fewer photons. This corresponds to 5–6 exposure values (EV), 6 EV corresponds to $2^6 = 64$, which is a significant disadvantage when shooting in low light and/or with fast-moving objects. Note that due to the photographer’s shaking hand, long exposure times are also critical in still photographs. This contribution has been significantly reduced in SPC through optical and electronical image stabilization (Section 13). Improvements in the image sensor technology – like “binning,” “deep trench isolation,” “dual conversion gain,” and “higher quantum efficiency” have also helped to close the gap to some extent in terms of these fundamental disadvantages compared to full-frame cameras [27].

4.6 David versus Goliath: The pros and cons of miniaturization

From the photographic imaging equations above, we may derive scaling rules for a direct comparison between “large,” e.g. a full-format ($36 \times 24 \text{ mm}^2$) DSLR or system cameras and a much smaller SPC (Figure 16). The comparison shall be done for the same photographic situation: A scene, e.g. a portrait of a person or a landscape, taken with the same field of view from the same position (the position of the entrance pupil, to be precise). The same content of the 3D scene is projected in an identical perspective (we assume the lens distortions to be small).

The crop factor

$$c = \frac{\phi_{\text{im, ff}}}{\phi_{\text{im, SPC}}} = \frac{43.3 \text{ mm}}{\phi_{\text{im, SPC}}} \quad (30)$$

between a full-format camera sensor and an SPC sensor is between approx. $c = 3.5$ to 12, depending on the SPC sensor size, which in turn specifically depends on the lens FOV (Section 6). We assume that the image sensor contains an equal number of pixels, e.g. 12 MP. Consequently, the pixel pitch is larger by a factor of c for the full-format sensor.

Furthermore, we assume the lens *f*-number K (or the NA of the image) to be equal as well. We can think of the same lens and image sensor, but just scaled down by a factor of c . For the scaled system (denoted by a bar) all angular quantities remain the same, namely the lens *f*-number (or NA) and field of view

$$\bar{K} = K \quad (31)$$

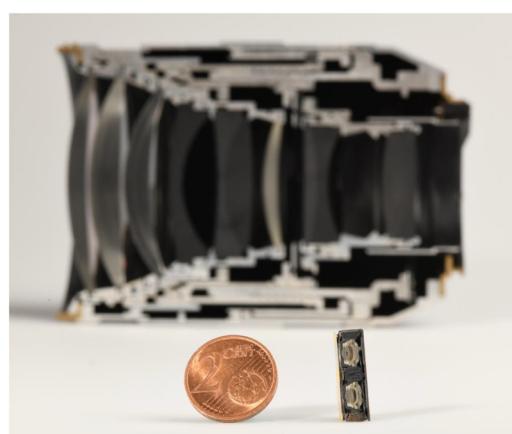


Figure 16: Cross-section of a smartphone dual-camera module versus a DSLR camera lens.

$$\overline{\text{FOV}} = \text{FOV} \quad (32)$$

while all lengths (e.g. focal length, sensor, and lens diameter) scale inversely with the crop factor:

$$\bar{f} = f/c \quad (33)$$

$$\overline{\phi}_{\text{im}} = \phi_{\text{im}}/c \quad (34)$$

$$\bar{L} = L/c. \quad (35)$$

The geometrical scaling on optical system lengths, diameters, surface area and volume/weight are straightforward (Tables 2 and 3). The weight and volume reduction in particular are immense, e.g. for a crop factor of 7 by a factor of 343; reducing weights of full-format camera equipment will be transformed from a few kilograms to just a few grams.

Another aforementioned benefit of miniaturization, also due to length scaling, is the decreased minimum optical distance (assuming the identical focusing mechanism of the original and miniature lens). This enables the user to focus on much smaller objects.

The major disadvantages of miniaturization arise from the fact that a pixel sees much less light, in our assumed model by a factor of $1/c^2$ proportional to the pixel area, resulting either in c^2 larger exposure time or higher ISO sensitivities, the latter resulting in increased noise.

Depth-of-field scales linearly with focal length (see equation on hyperfocal distance) meaning that the hyperfocal distance is a factor of c further away from the miniature lens. For practical photography this is an advantage, since the probability increases that the image quality will not deteriorate due to focusing errors. However, as mentioned previously, miniature lenses are not capable of creating an

artistic shallow depth of field, which most passionate photographers see as a significant disadvantage.

The scaling of optical resolution is a little more complex. For this consideration, lens aberrations should be included, since for full-format lenses at large apertures those tend to significantly limit optical resolution. Now it is straightforward to see in a spot diagram, which describes the ray deviations on the image plane, and that aberrations scale down linearly due to length scaling. That is according to $1/c$ on an absolute scale, when a lens is downscaled. But since pixel size also scales down by the same factor $1/c$, ray aberrations remain unchanged on the scale of a pixel. However, an ideal lens is limited by diffraction, that is the Airy spot size, namely $2.44 \cdot \lambda \cdot K$. Since the f -number K does not change, the resolution does not change either on an absolute scale, but the Airy spot size does change on the scale of the smaller pixel when downscaled to miniature size. Consequently, as soon as lens performance becomes diffraction-limited while downscaling the size of a lens, the resolution will drop linearly according to $1/c$.

Lohmann [28] proposed how the spot size of geometric contribution – describing the lens aberration, and an (ideal) diffraction contribution – can be approximately combined to describe the effect of lens scaling on resolution:

$$A'_p = \lambda^2 K^2 + (1/c)^2 \bar{\xi}^2 \quad (36)$$

A'_p denotes the spot area and $\bar{\xi}^2$ the second moment of the ray deviation distribution. In Figure 17 this scaling rule is exemplified together with a wave-optical calculation of the point spread function (PSF) (including both aberrations and diffraction).

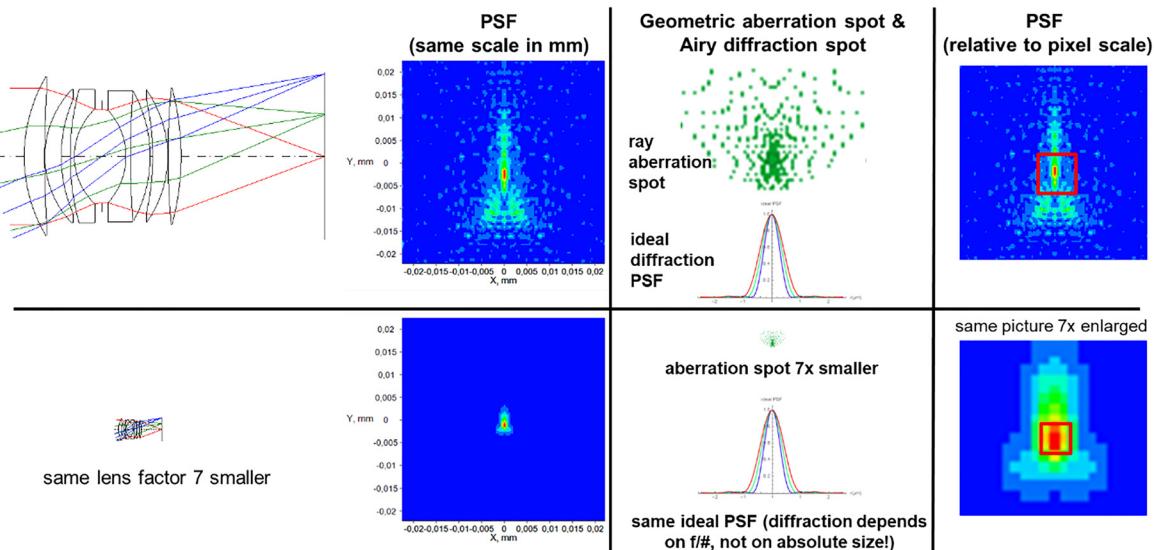


Figure 17: Scaling down the same lens by a factor of 7 and the resulting PSF on the scale of a pixel (red box in graphs on right-hand side).

Table 2: Scaling of physical parameters with crop factor. Advantages of miniature cameras in blue; advantages of large-format cameras in black; depth of field can be seen as both an advantage and a disadvantage.

Parameter	Scaling factor
Length	$1/c$
Diameter	$1/c$
Sensor area	$1/c^2$
Volume	$1/c^3$
Weight	$1/c^3$
Minimum optical distance	$1/c$
Etendue/throughput	$1/c^2$
Exposure time	c^2
Low light noise	c
Object resolution (PSF size relative to pixel size)	$(1 + (k \cdot c)^2)^{0.5}$
Out-of-focus spot diameter	$1/c$
Depth-of-field	c

Multi-camera systems containing many lenses in parallel are a straightforward concept for smartphone imaging [29, 30]: They are thin in order to fit into the tight camera housing, while the effective image sensor area increases. Combining the images of these cameras in order to improve image performance in different directions (e.g. noise reduction, HDR) is a computational

task which to some degree is realized in current smartphone multi-camera systems [31, 32]. Multicamera systems are capable of performing stereographic 3D depth acquisition. We can thus conclude that an attempt has been made to compensate for all disadvantages marked in Table 2 by means of shooting multiple pictures (with the same or several different cameras) and computationally combining those images.

4.7 SPC lenses: How good are they actually?

We want to follow up on the discussion of aberration performance of the previous Section 4.6 by asking: How good are SPC lenses corrected actually? How good would they be if scaled up to the size of a full format ($36 \times 24\text{mm}^2$) lens? Conversely, what happens if they would be scaled further down in size?

We take a SPC lens $f/1.9$ wide-angle (FOV 75°) design for a $1/3.3''$ image sensor. The image diagonal of 5.4 mm is almost exactly $8\times$ smaller compared to a full format image sensor (diagonal 43.3 mm). In Figure 18 the polychromatic MTF is shown of the original SPC lens, denoted as “ff/8, original” (“ff” means “full format”), and scaled up in size

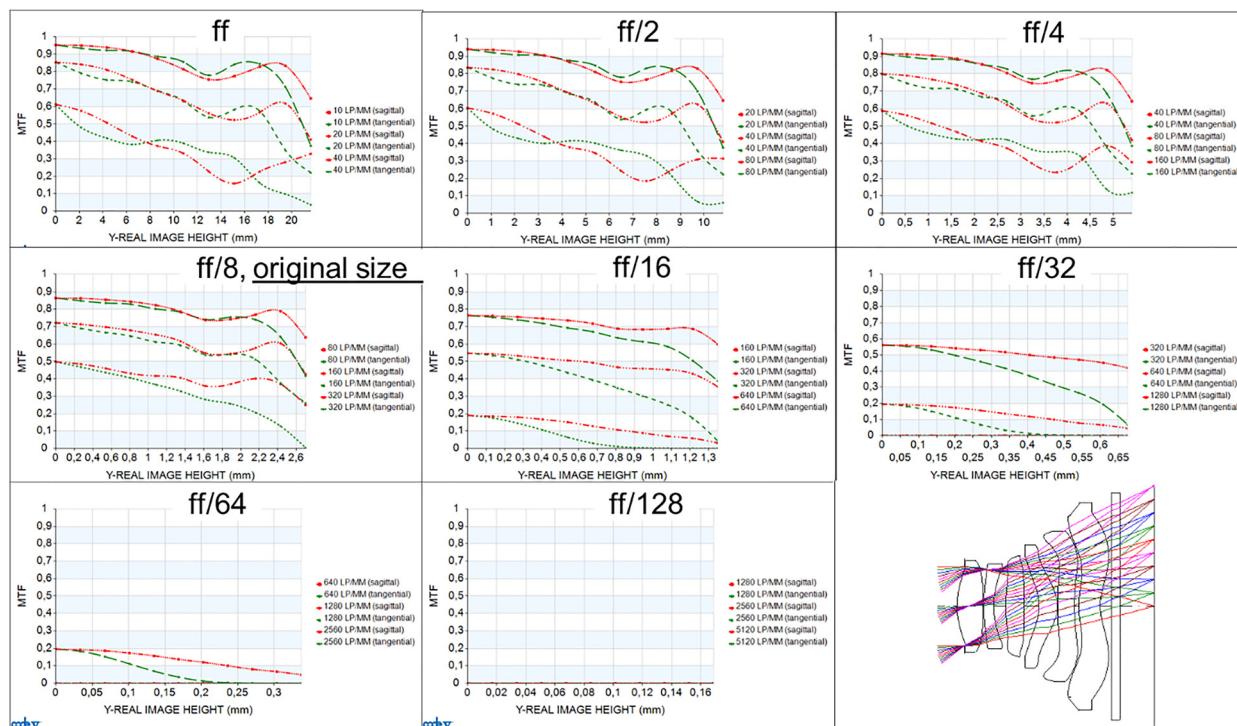


Figure 18: MTF of SPC lens “ff/8, original” scaled up (up to a factor of 8 to full format “ff”) and down (down to a factor 16 to “ff/128”) consecutively by a factor of 2 in each step. The image field size is a factor of 2 different in each step (see scale on abscissa). Spatial frequencies are scaled accordingly. The spectral relative weights of the polychromatic MTF are 1 (405 nm), 3 (436 nm), 12 (486 nm), 24 (546 nm), 24 (587 nm), and 11 (656 nm).

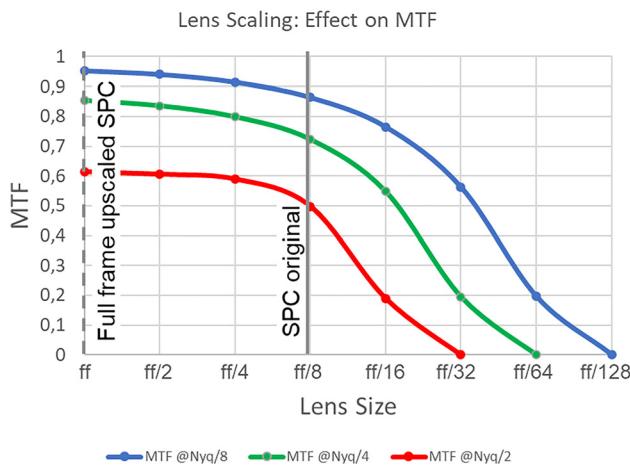


Figure 19: Summary of Figure 18. MTF-performance of original SPC (“ff/8”) and upscaled and downscaled lens design evaluated at Nyq/2, Nyq/4, Nyq/8 for each size. While at original size diffraction limits the performance only moderately (moderate performance improvement by upscaling), downscaling lead to severely limited performance due to diffraction limitation.

by a factor of 2 (“ff/4”), 4 (“ff/2”) and 8 (“ff”). In order not to compare apples and pears the spatial frequency is correspondingly scaled down in each consecutive step by a factor of 2. A common choice (at ZEISS and other lens suppliers) of spatial frequencies for MTF evaluations at full format is 10, 20 and 40 lp/mm. Correspondingly the 8 times smaller SPC lens is evaluated at 8× larger spatial frequencies 80, 160, 320 lp/mm. (If these spatial frequencies correspond to Nyq/8, Nyq/4, Nyq/2 the pixel pitch $p = 1/(2 \text{ Nyq})$ of the SPC is $p = 0.78 \mu\text{m}$ and for the full format camera $p = 6.25 \mu\text{m}$, respectively).

In Figure 18 the MTF performance is shown and summarized in Figure 19 and Table 3. As can be seen the MTF at original size is only slightly smaller compared to the upscaled version of the lens. According to the Lohmann’s scaling law presented in Section 4.6 this means that the overall performance at the original size (5.4 mm image diagonal) with respect to imaging with $0.78 \mu\text{m}$ pixel pitch is only moderately limited by diffraction compared to the aberration level of the lens design.

Table 3: Summary of MTF according to Figure 19 and corresponding pixel size ($p = 1/(2 \text{ Nyq})$) and image field radius y_{\max} .

	Full format	ff/2	ff/4	ff/8	ff/16	ff/32	ff/64
Scale rel. to SPC	8	4	2	1	1/2	1/4	1/8
MTF @ Nyq/8	0.953	0.941	0.915	0.864	0.765	0.562	0.196
MTF @ Nyq/4	0.854	0.836	0.799	0.723	0.548	0.195	0
MTF @ Nyq/2	0.613	0.605	0.589	0.498	0.19	0	0
Pixel size (μm)	6.250	3.125	1.563	0.781	0.391	0.195	0.098
y_{\max} (mm)	21.6	10.8	5.4	2.7	1.35	0.675	0.3375

The performance level of the full frame upscaled version of the SPC is excellent. The MTF performance is comparable with excellent full format lenses like the ZEISS Otus 1.4/28 mm (when compared at SPC aperture f/1.9) [33].

Looking on the other side of lens scaling: When the lens size is scaled down the overall performance severely drops. The diffraction contribution in Lohmann’s scaling law becomes dominant. This is obvious when the lens is scaled down by a factor of 2 (“ff/16”). The Strehl ratio is a common measure to evaluate lens performance in comparison to the diffraction limit [34]. Figure 20 shows the Strehl ratio for the PSF in the center of the field: At original size (“ff/8”) the SPC lens is rather diffraction-limited ($S = 0.88$), but not when scaled to full frame format ($S = 0.12$). The Strehl ratio approaches its maximum value 1 for scaled down versions but as mentioned is not capable to support the corresponding sensor resolution: This means the performance is principally limited for yet smaller lens sizes. Indeed for actual lenses in current SPC multicameras the diffraction limitation is actually limiting the performance: Unlike in the current analysis with aperture f/1.9 and field diameter 5.4 mm both aperture and field are smaller than that, e.g. f/3.4, $\phi_{\text{im}} = 4.2 \mu\text{m}$ for tele lenses, and therefore predominantly diffraction-limited.

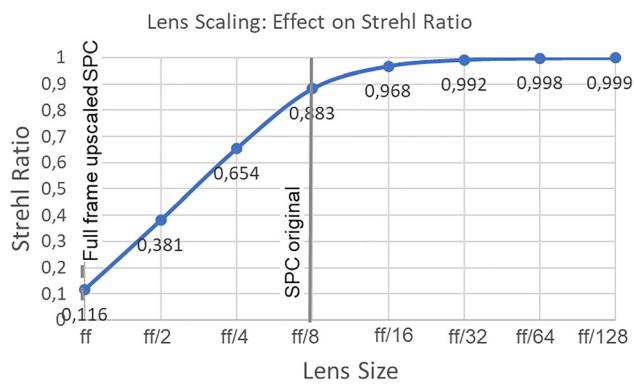


Figure 20: Strehl ratio at the center of field for (the scaled versions of) the SPC lens.



Figure 21: Sony Xperia 1 III multi-camera system. Cameras (top to bottom): ultrawide angle 16 mm, F2.2, 12 MP, Tele 70 mm, F2.4, 12 MP, 3D iToF for depth acquisition, standard wide-angle 24 mm, F1.7 12 MP (main camera). Courtesy of Sony.

Conversely, it can be said that it makes little sense to achieve pixel resolutions far below 0.7–0.8 μm , which is the current state of the art for image sensor CMOS technology. To do this, it would have to be possible to implement optical designs with an even larger aperture such as about f/1.

5 The multicamera system in modern smartphones

For a long time, only one standard wide-angle lens was used in mobile phone photography. Up to about 2006 it typically had an equivalent focal length of 35 mm ($\text{FOV} \approx 60^\circ$), and later of 28 mm ($\text{FOV} \approx 75^\circ$). Today the standard wide-angle camera as a component of the multi-camera system is often called the “main camera.” Besides its importance in everyday photography, this is mainly due to its relatively superior performance compared to other camera lens focal lengths. This is based on the feasibility of achieving extremely flat form factors (see Section 6) at high apertures. Since 2016 and the launch of the Apple iPhone 7+, the number of rear cameras has been steadily increasing from dual, triple, quad, and now penta. In high-end smartphones, the standard wide-angle lens is supplemented by lenses with a shorter focal length (corresponding to an FOV of around 120°) and longer focal lengths such as 55 mm, 70 mm or even 125 mm. In addition, there is a 3D depth sensor, e.g. based on ToF measurement, which is capable of generating real-time depth maps of a scene over a wide FOV (see Section 15.1). Figure 21 shows an example of a multicamera system.

Today, almost all smartphones are equipped with two rear cameras as standard, including mid- and low-end smartphones. The number of cameras continues to grow, both on the rear and the front. The multicamera market

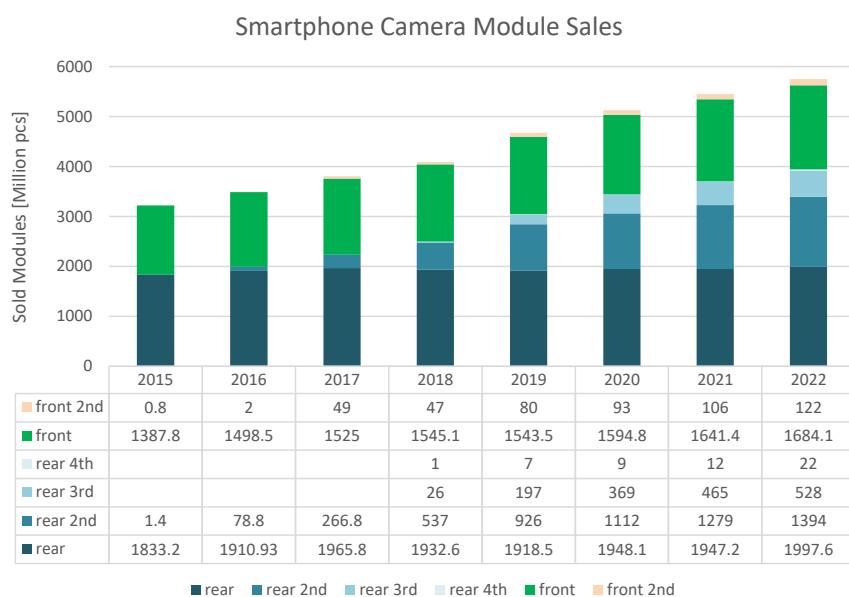


Figure 22: Camera sales per module in million units (Source: TSR; 2021/22 are predictions).

Table 4: Sample data from a high-end SPC system consisting of four rear cameras and one front camera.

	Front cam	Extreme wide cam	Wide cam	Tele cam	Folded tele cam
Diagonal field-of-view	80°	120°	80°	42°	20°
Equivalent focal length	26 mm	14 mm	26 mm	56 mm	125 mm
Sensor size (sensor diagonal)	5.7 mm (1/2.8")	8 mm (1/2")	11.3 mm (1/1.4")	5.7 mm (1/2.8")	4.5 mm (1/3.8")
Sensor pixel #	32 MP	48 MP	48 MP	32 MP	10 MP
Pixel pitch	0.8 µm	0.8 µm	1.12 µm	0.8 µm	1 µm
f-Number	2.45	2.2	1.6	2.1	3.4
Focal length	3.7 mm	2.6 mm	6.8 mm	7.5 mm	13 mm
Minimum optical distance	35 cm; fixed focus	3 cm	12 cm	50 cm	100 cm
Image stabilisation	/	OIS	OIS	/	OIS

trend is shown in Figure 22. In the past the front “selfie camera” usually offered a low resolution, as a fixed focus camera with a small sensor and small aperture to increase the depth of field. In recent years, however, the standard wide-angle lens has evolved into a larger camera with a high-resolution image sensor (often multicell to improve HDR) and an autofocus. High-end smartphones usually have a dual front camera. Often, the standard wide-angle lens is provided in tandem with an ultrawide camera for handheld panorama selfies. Since about 2018 there has usually been a 3D sensing camera for face recognition on the front, just next to the visual cameras.

Table 4 shows sample technical data from a 4-rear-camera and 1-front-camera system.

Figure 23 shows the FOV of each camera, together with the minimum optical distance (MOD) and the depth of field at MOD. The object field diameter ($\phi_{ob} = 0.15$ m) is also shown for the front camera’s MOD. As you can see, the tele lenses’ captured object diameter at their corresponding MODs is larger (about $\phi_{ob} = 0.5$ m). Similarly, for the front camera the fixed focus distance is $\phi_{ob} = 0.6$ m, such that a person’s face is well within this object region.

6 Optical System Design

6.1 Optical design structure of a smartphone lens

The opto-mechanical design of and the manufacturing technology for smartphone lenses are very different from those of classic photo lenses. The use of highly aspherical plastic lenses, the miniaturization, and the completely automated production in quantities of several millions can hardly be compared with classic optics manufacturing.

The structure of the optical system (Figure 24) is largely determined by the required miniaturization. The SPC must fit into a flat case about 8–9 mm thick. Subtracting the housing and image sensor thickness, the overall length of the lens must therefore not be longer than about 5–6 mm. The diagonal of the image sensor should be as large as possible in order to reduce the disadvantages of small image sensors described above (image noise, dynamic range, etc.). At the same time, the aperture of the lens must be relatively large, about $f/2$ or larger, so that the system resolution is not limited with image pixel sizes of around 1 µm.

With an FOV of around 80°, SPC wide-angle lenses achieve a form factor of construction length (from front lens to image plane) to the image sensor diagonal of

$$r = \frac{L}{\phi_{im}} \approx 0.65 \dots 0.85. \quad (37)$$

With an overall length of $L = 6$ mm, this enables an image sensor diameter of up to $\phi_{im} = 9$ mm. If all means are exhausted, such as minimum board and housing thickness, and the camera is protruding slightly from the housing, some current high-end models exceed the image sensor diameter of the wide-angle main camera by 12 mm. Despite the highly miniaturized design, they are only a crop factor of around 4 away from full-frame image sensors.

With spherical lens shapes, there is no classic lens type with such a large aperture of around $f/2$ that has such a small overall length-to-image-diameter ratio. The Biogon is a well-known classic spherical ultracompact high-aperture wide-angle lens. Steinich and Blahnik [35] compare the Biogon with a mobile phone lens with the same aperture, and FOV, scaled to the same image diagonal: The image performance of the SPC optics turns out to be even better than that of the Biogon, despite the Biogon being about twice as long. Figure 25 shows this with another example: For the 2/35 mm Biogon made of spherical lenses, $r = 1.58$,

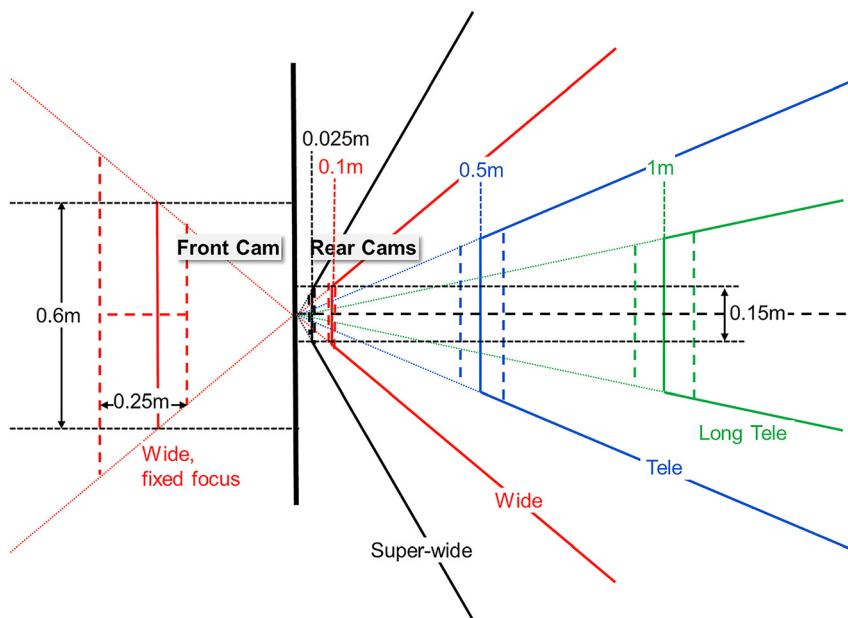


Figure 23: Multi-camera system according to data in Table 4: FOV, MOD, DoF @MOD of the corresponding camera lens. The object field diameter for the main rear and front cameras has also been drawn (0.15 and 0.6 m, respectively.).

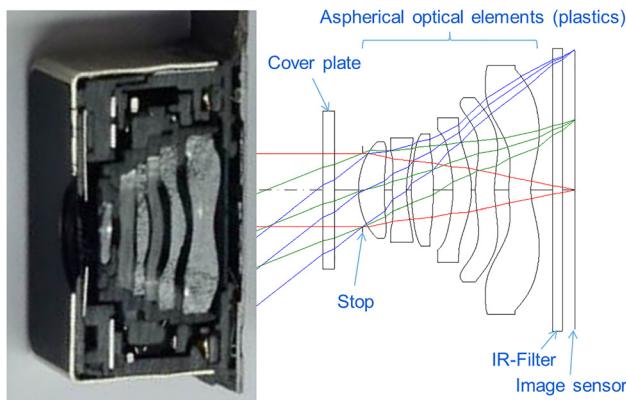


Figure 24: Structure of a wide-angle lens: The light coming from a faraway object enters the lens at a cover glass (approx. 0.2 mm thick), then passes through the plastic lens elements and an infrared (IR) filter (thickness approx. 0.2–0.3 mm) before finally arriving at the image sensor. The fixed lens stop is usually placed at the lens entrance.

while the 1.9/28 mm aspherical plastic lens achieves $r = 0.83$. The contrast of the SPC lens is even higher, the peripheral light intensity drop is lower (both due to the lack of vignetting with the SPC lens) and the distortion is comparably very good (<1%), as are the chromatic aberrations (<1.5 μm color fringes).

The key to the flat form factor is to use extreme plastic aspheres. The use of plastic lenses did not emerge for SPC until after 2000, but has already been in use since

the early 1950s in the form of simple meniscus lenses or doublets, and a correspondingly small aperture of $f/15$, as in the Kodak Kodet [36]. In 1957 Eastman Kodak used molded plastic lenses in the rangefinder of their Signet cameras and since 1959 an $f/8$ triplet lens made of plastic. Kodak established a standard in the consumer camera market for the following decades. Until the 1970s, the technological advantages of plastic had not yet been exhausted: the lenses remained spherical, due in part to the still-low computing power of mainframes for lens design calculations. That changed in the 1970s. The innovative Polaroid SX70 camera from 1972 even had a freeform surface in the rangefinder [37]. Precisely manufactured molded aspheres have been used in many projectors and photo cameras since the 1970s, such as the triplet shown in Figure 26 or the Kodak Ektar 1.9/25 mm built in the Kodak Ektramax, also with a plastic asphere. These lenses can be regarded as early fore-runners of modern SPC lenses, although the aspheres still had low asphericities, which was in part due to the limited computing power available at the time for optical design [38].

From the doublets or triplets in the early days of smartphone photography, the number of lens elements was continuously increased to further increase the numerical aperture and thus the optical resolution, as dictated by pixel shrinking. In 2020, there were often 7 lens elements in the standard wide-angle lens, which were

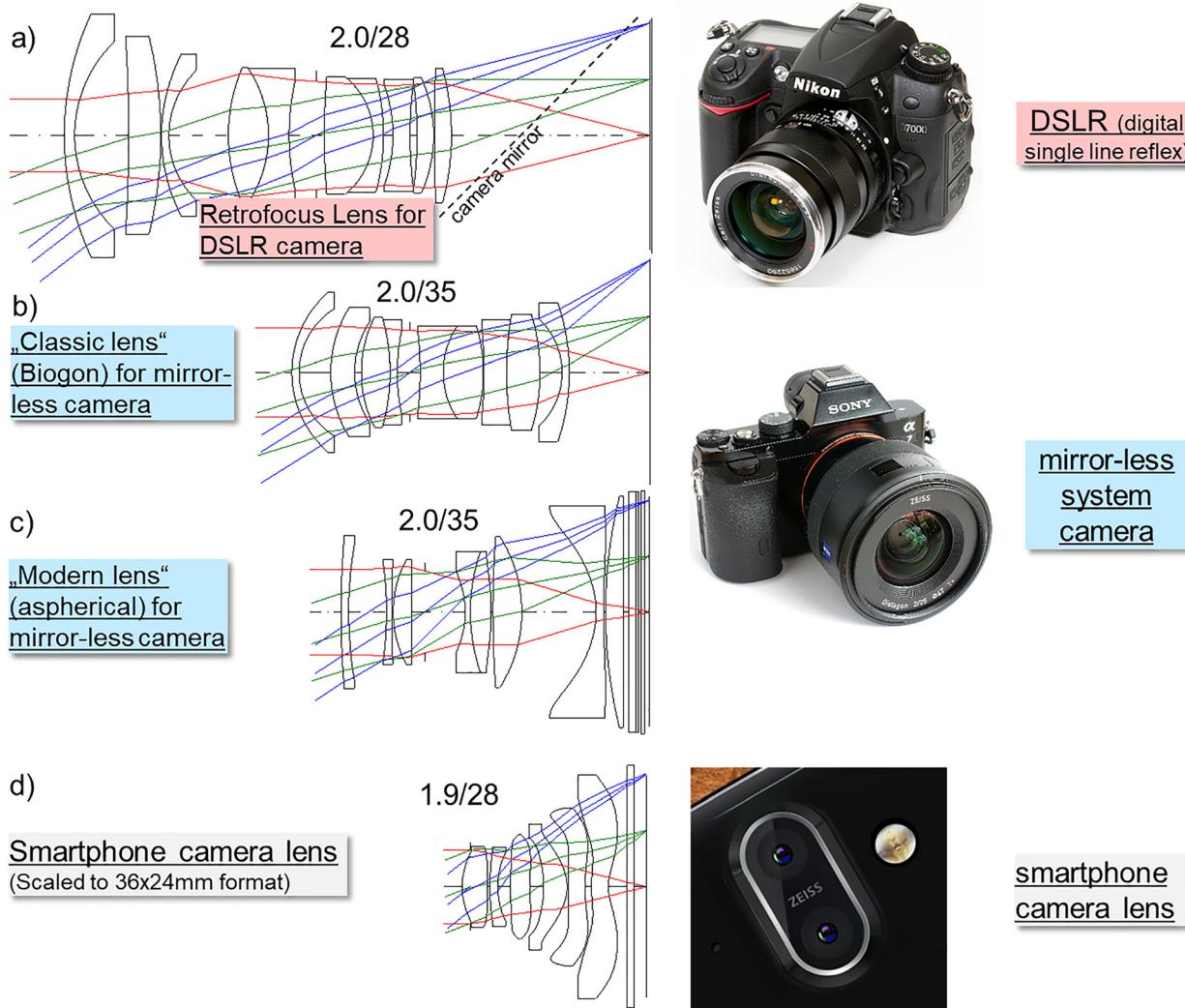


Figure 25: Comparison of different wide-angle lenses: (a) 2/28 for SLR camera, (b) classic 2/35 mm for (mirrorless) rangefinder camera, (c) modern 2/35 mm for mirrorless system camera, and (d) SPC lens; all systems are scaled to the same image size (to true scale the SPC lens d) is by a factor of 4–10 smaller in length and diameter). The ratio of the overall length (vertex of the first lens to the image plane) to the sensor diagonal is (a) $r = 2.63$, (b) 1.58, (c) 1.39, and (d) 0.83.

closely packed one after the other. Figure 29 shows a chronological history from 2004 to 2020. Until 2016, SPCs were practically exclusively equipped with standard wide-angle lenses. Since 2016 lenses with other focal lengths

started being added. A study of SPC lens patent literature [39] shows this same trend shifted forward in time: For many years only FOVs of around 60° and 75° were considered; but now the optical designs are very diverse in

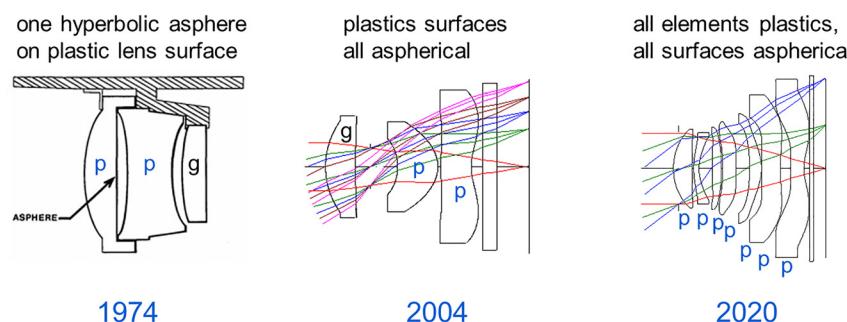


Figure 26: Comparison of optical designs: Early aspherical plastic consumer camera lens from 1974 with one aspherical lens element and one early (2004) and modern (2020) mobile phone camera wide-angle lens. The material type is abbreviated as “g” for glass and “p” for plastics.

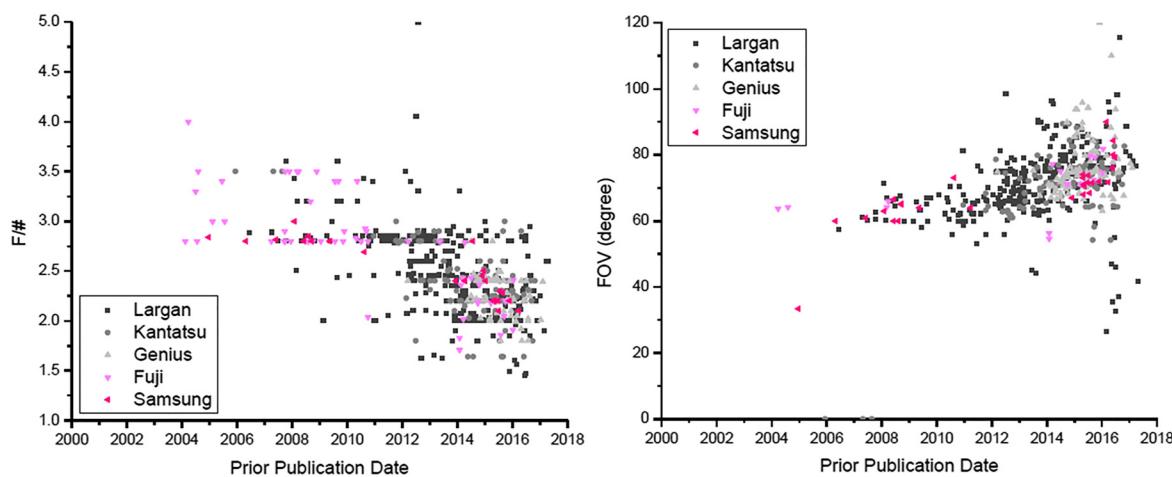


Figure 27: Patent literature on mobile phone camera optical design [39]: F-number and FOV of about 750 mobile phone camera lens patents versus patent publication day. Courtesy of Luxin Nie.

an FOV range of about 20° – 150° . At the same time the f-number decreased to support pixel shrinking (Figure 27). Consequently, optical design has become more and more complex, with the number of lens elements continuously increasing (Figure 28). Meanwhile there are optical design patents which feature 9 lens elements [40].

In Figure 29 the evolution in the diversity of optical designs is shown in terms of typical configurations of multicamera system SPCs.

The optical design of SPC lenses amounts to a paradigm shift in the fundamental layout of camera lenses. Despite its widespread use today, the optical design of SPC

lenses is rarely dealt with in the literature, as in refs. [41–44] or in the work done by the group of José Sasian [45].

The optical designer David Shafer wrote the following in his study: “A new theory of cell phone lenses” [46]: “My conclusion from this study is that the usual cell phone designs with very wiggly aspherics are using extremely high-order aberrations to compensate for uncorrectable lower-order aberrations – due to a nonoptimum third and fifth order distribution inside the designs.”

In contrast to classic camera lens designs consisting of spherical lens elements there are excellent configurations, e.g. the Double-Gauss type and Triplet variants, (e.g. [47–49])

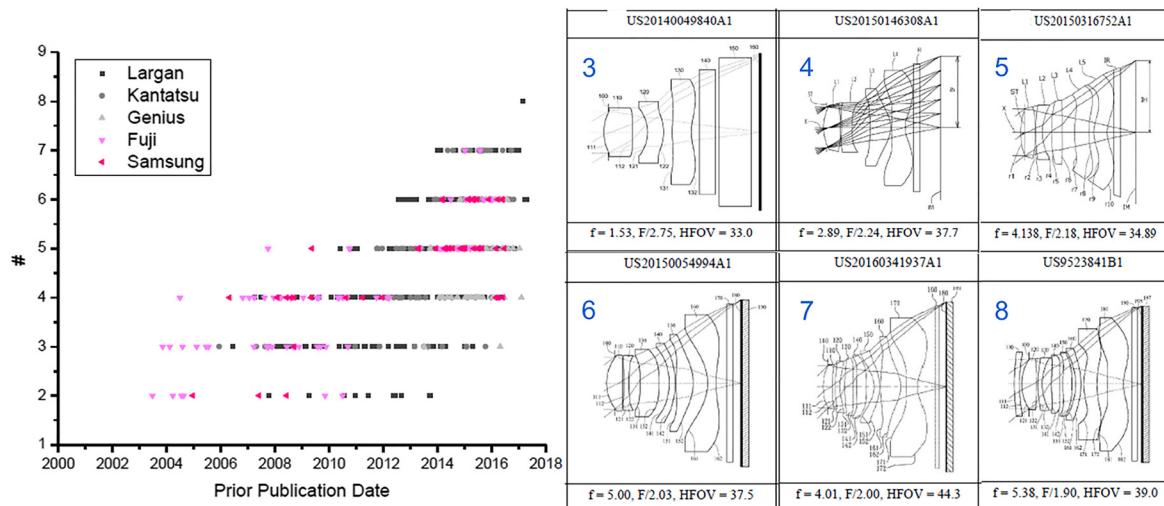


Figure 28: Number of lens elements versus patent publication day (left); wide-angle lens designs with 3–8 lens elements (right). Courtesy of Luxin Nie.

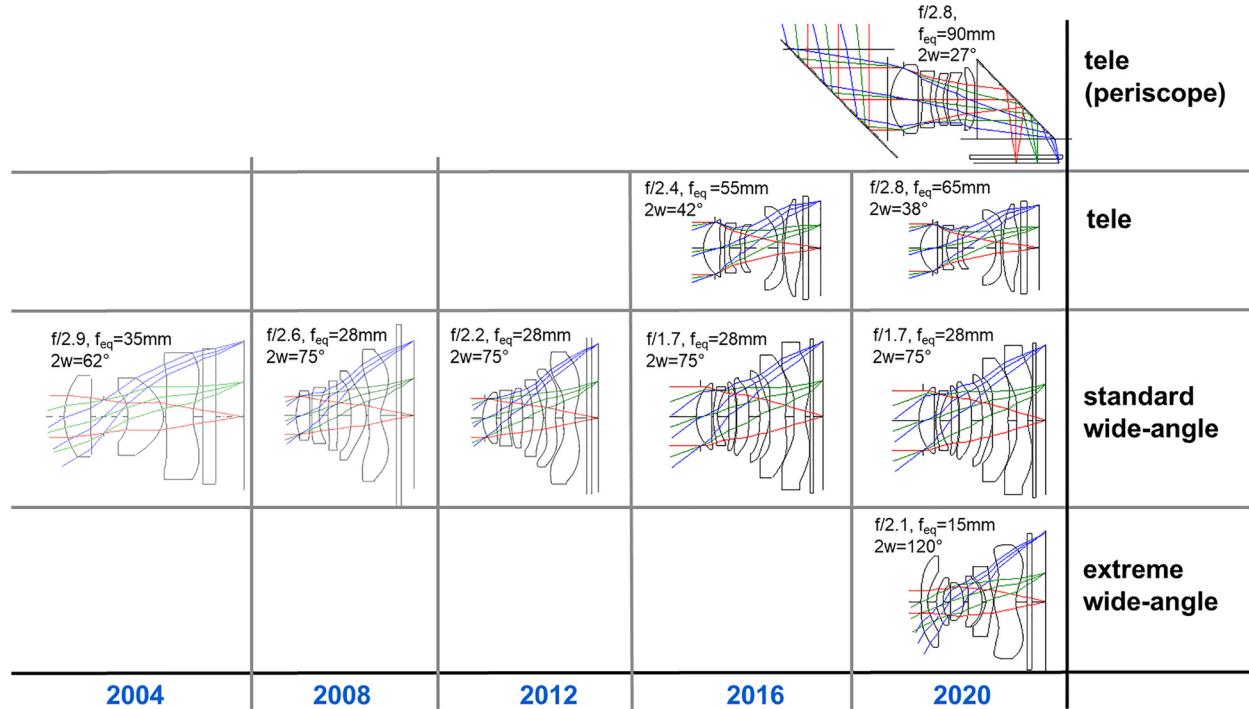


Figure 29: Evolution of smartphone optical design beginning with standard wide-angle lenses of increasing aperture and complexity. FOV was extended towards the extreme wide-angle and tele ranges.

which are able (within a certain parameter range of f-number and FOV) to completely remove third and fifth-order aberrations. Let us have a look at the optical design structure of the Biogon and then compare it with the design of a corresponding SPC lens with the same FOV: The Biogon is based on a strictly concentric structure with respect to both the outer lens elements and the inner positive lens group (Figure 30a). The stop is placed in the center of this system, which results in the rays entering the system at a certain angle of incidence and also exiting the system at this same angle. A strictly concentric system delivers consistent image quality from a spherically curved object surface to a spherically curved image surface. For a curved image sensor with simple monochromatic lenses, excellent image performance is feasible [50]. For SPC lenses it has been shown that curved image sensors can result in lenses with one f-stop superior f-number and comparable aberration performance [51]. In particular, extreme wide-angle lenses would benefit significantly from curved image fields. For a plane image sensor, however, the Petzval condition must be fulfilled,

$$\sum_j \frac{\varphi_j}{n_j} = 0 \quad (38)$$

where the lens component's refractive powers are φ_j and refractive indices n_j . Consequently, the lens must consist of both positive and negative components. The Biogon lens structure consists of negative outer lens elements (meniscus shape) and a positive inner group (Figure 30b). With this structure the chief ray is bent to a smaller angle inside the lens, which is beneficial because it leads to smaller aberration contributions. Another advantage of the negative outer elements is that the off-axis entrance pupil becomes larger, improving the relative illumination. Asymmetrical aberration types, namely distortion, coma, and lateral chromatic, are eliminated by the quasi-symmetric arrangement around the stop in the center of the system, because the aberration contributions in the front lens part occur with opposite signs at the corresponding position at the rear and thus cancel each other out (Figure 30c). Longitudinal and higher-order chromatic aberrations are corrected by an arrangement of low-dispersion glasses for the outer negative lens elements and achromats of the inner positive elements (Figure 30d). Spherical aberration and astigmatism are the remaining aberrations to be corrected by fine-tuning all lens parameters through optimization, especially the

lens radii. At the time of the invention [52] these calculations were extremely time-consuming since computers were not available yet.

Now let us compare the optical designs of the Biogon and SPC lens: both have a field of view of about 80° (Figure 31) and are similar in that the ray angle at the entrance of the lens is about the same as at the lens exit. Conversely, for SPC half of the system structure from the aperture to the image plane is sufficient: the aberrations, in particular distortion and coma, are corrected by the aspheres. With spherical elements in such an arrangement with the diaphragm in the front area, the distortion would be difficult to correct; with the strongly aspherical design though, relaxation of the distortion specification (typically <1%) and subsequent digital correction would bring almost no advantage. Within the SPC lens the chief ray path generally runs along a straight line. Despite a much larger aperture ($f/1.7$ compared to $f/4.5$), the SPC lens is considerably shorter.

With wide-angle lenses for SLR cameras, one is forced to use a retrofocus type because of the space required for the folding mirror between the last lens element and the image sensor, i.e. negative refractive power is required in the front and negative power in the rear of the lens [53]. This enables a certain symmetry and the lens construction becomes considerably more complex (see Figure 25a). Modern photo lenses for mirrorless cameras are also

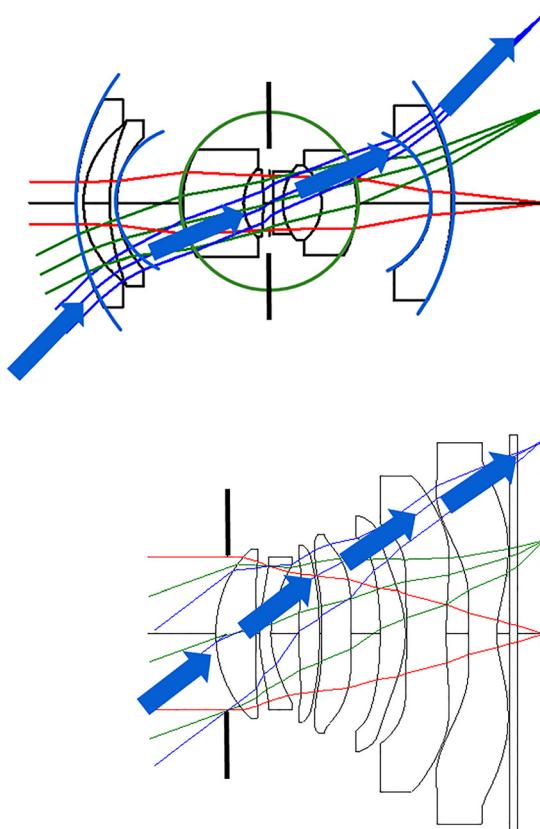


Figure 31: Comparison of a Biogon (4.5/21 mm) with an SPC lens (1.7/25 mm) scaled to same image size.

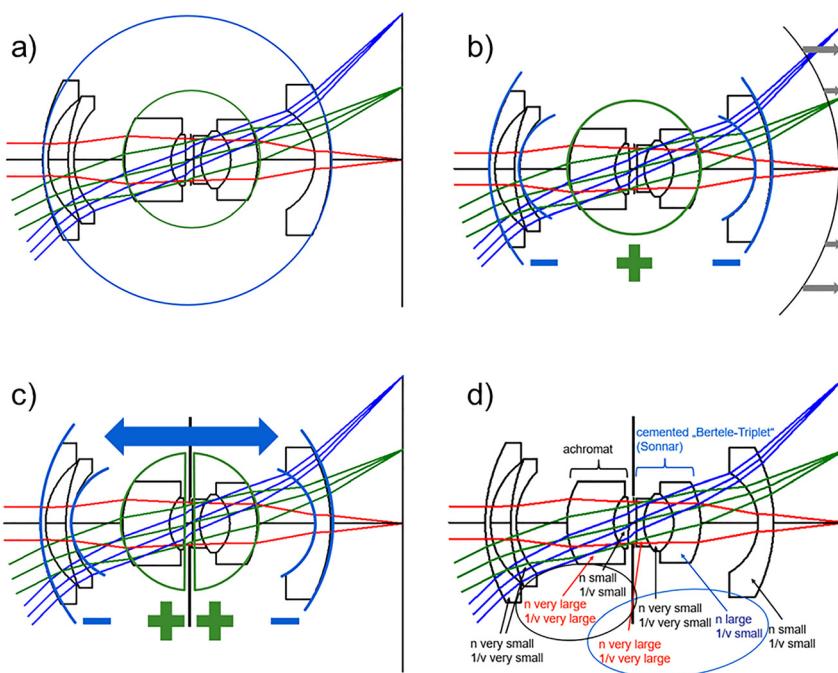


Figure 30: Structural features of the classic Biogon lens type: (a) Monocentric layout, b) $(-, +, -)$ refractive power to correct for field curvature, (c) symmetry to the central stop, and (d) achromatization.

increasingly asymmetrical (see Figure 25c). To correct distortion, curvature of field and astigmatism, aspherical lenses are often placed directly in front of the image plane. These very compact systems are favored thanks to the availability of inexpensive aspheres through molding processes and also by the considerably greater computing power needed for optimization in optical design. Nevertheless, the classic symmetrical lens types, all of which were created before computers were used in the 1950s, can still be found in many of today's products.

For the optical design of SPC optics, the classic design rules according to Seidel's third order theory [54] are no longer applicable. High-order aberrations are used here in order to reduce low-order aberrations. All lens surfaces are aspherical. The spherical basic shape corresponding to the radius of curvature in the center of the surface is shown in blue in comparison. On the rear, the deviations from the spherical shape are very large. On the last surfaces it is obvious that this asphere cannot be obtained from manufacturing techniques like grinding and polishing starting from a spherical basic shape (Figure 33). The standard surface description of aspheres is:

$$z = \frac{cr^2}{1 + \sqrt{1 - (1+k)c^2r^2}} + a_4 r^4 + a_6 r^6 + \dots \quad (39)$$

Here, c denotes the curvature at the apex of the surface, k the conical constant and r the radial distance from

the optical axis. The first term alone results in different conic shapes depending on the value of the conical constant: $k = 0$: sphere, $-1 < k < 0$: ellipsoid with main axis on the optical axis (prolate spheroid), $k = -1$: paraboloid, $k < -1$: Hyperboloid. In Figure 32 the description of the asphere of a typical SPC "w-shaped" lens element surface is shown.

In the case of SPC lenses, usually even polynomial coefficients up to about the 10th order are used for the first 2 to 3 moderate aspherical lens elements and up to the 14th or 16th order for the extremely strongly aspherically curved near-field lenses, i.e. a_4, a_6, \dots, a_{14} or a_{16} . Some optical designers also use odd orders: a_3, a_5, \dots (odd polynomials). In addition, there are surface descriptions with orthogonal function systems such as the Forbes polynomials [55], which have advantages in the convergence of the optimization [56] and also in the desensitization of the system [57] (Figure 33).

The low aberration orders are largely compensated for with high-order aspheres, but residual high-order aberrations remain (as can be seen in the aberration graphs in Figure 36). When optimizing SPC lenses, the pupil and the field must therefore be sampled sufficiently, otherwise the image performance between the optimized field points threatens to drop sharply.

In contrast to a lens for a large-format camera, SPC optics must hardly have any vignetting by lens edges or other

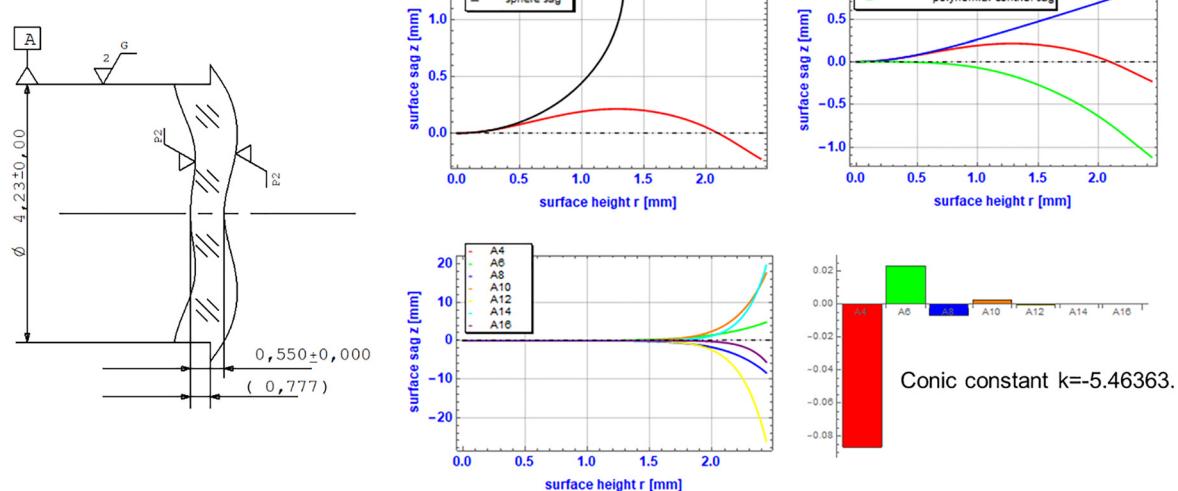


Figure 32: Right surface of a "w-shaped" SPC lens element. The coefficient values are given on the bottom right. Middle top: The actual asphere shape deviates strongly from the corresponding lens center spherical curvature. Top right: Conic constants are often used for SPC asphere description with considerable contribution. Middle bottom: Typically the separate functions (monomials) take on large values at the edge.

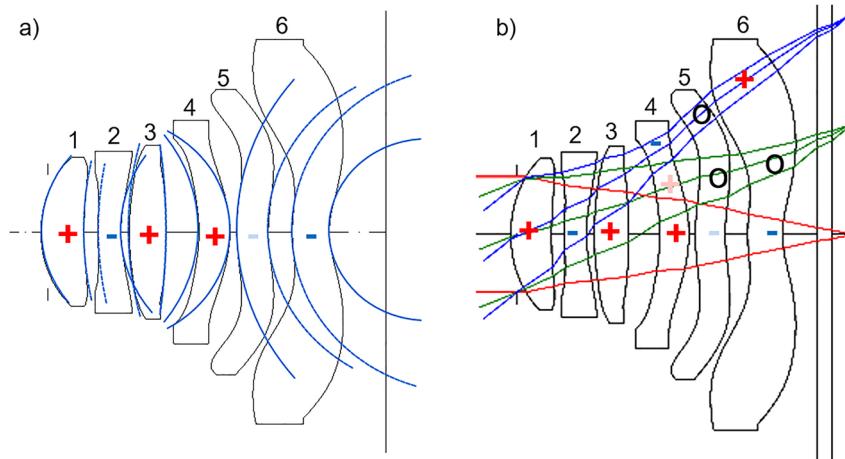


Figure 33: SPC optical design. (a) The local radius of curvature in the center of the lens is shown in blue for comparison. In particular, the last lens elements in front of the image plane deviate significantly from the spherical shape. The refractive power of a lens in the center of the lens, whether positive or negative, is also shown. However, the last two lens elements have a strong field-dependent effect with (locally) very different refractive powers and deflections on the beam. (b) The local curvature variation causes a change in the refractive power along each field-dependent light path. Positive (+), neutral (o), and negative (-) power is indicated in the figure for the field near lens elements 4, 5 and 6.

apertures. This would be tantamount to reducing the aperture and would lead to a loss of resolution towards the image corners. In the case of an SPC, and unlike in DSLR lenses, the light is usually not blocked by any field stop or lens edge.

With mobile phone wide-angle lenses, the geometric light path to the corner of the image is more than 20% longer than it is to the center of the image. The design structure ensures that the size of the numerical aperture remains almost the same up to the edge of the image, which is necessary to keep the diffraction-limited resolution almost constant up to the corner of the image: The lens elements in the middle of the lens between the pupil and field considerably change their refractive power depending on the image field height. This can be seen particularly well in Figure 33 on lens element 4 due to the more negative refractive power at the edge of the field, the light bundle initially becomes more divergent. On the rear of the cell phone optics, especially with the “w-shaped” last lens element, the refractive power at the edge of the field is positive, while it is clearly negative in the center of the lens. As a result of these differences in refractive power, the aperture at the field edge becomes significantly larger than it would be with a corresponding conventional spherical optic. With these extreme asphericities of the last lens elements of the SPC optics, one can no longer speak of positive or negative lenses: the refractive power varies over the field height. Zhuang et al. [58] propose a systematic design approach based on different types of such significantly curved field lenses.

A characteristic of the standard wide-angle lenses is that the chief rays strike the image plane at a similarly high

angle as they enter the lens: With a standard wide-angle lens, this is approx. $\pm 35^\circ$ – 40° when entering the lens; the typical chief ray angle in the image corner on the image plane is also around 35° . The mechanism described in the previous paragraph (same refractive power in the anterior pupil area, then more negative and then positive for the field edge compared to the center of the image) ensures, in addition to the relative aperture size, that the chief rays are bent considerably more. The result is a nonlinear course of the angle of incidence towards the field edge: Rising from the center of the image, stagnating in the field zone and edge. This means that the position of the exit pupil is not in the same place for all field points, but rather for field points at the edge of the field much further in the front.

In Figure 34 lenses of different focal lengths are shown: “Super wide-angle” (FOV 120°), wide-angle (75°), normal tele (38°), and periscope tele (27°). The chief ray in the standard wide-angle lens essentially runs along a line through the lens. For the telephoto and extreme wide-angle types, however, there is global bending of the chief ray passing through the lens (Figure 35). This characteristic is a reason why standard wide-angle lenses are the most compact lens type among SPC lenses. Therefore, the image sensor used for a standard wide-angle lens is the largest within the multicamera system.

6.2 Optical design imaging performance

Optical image performance evaluations of camera lens design commonly include:

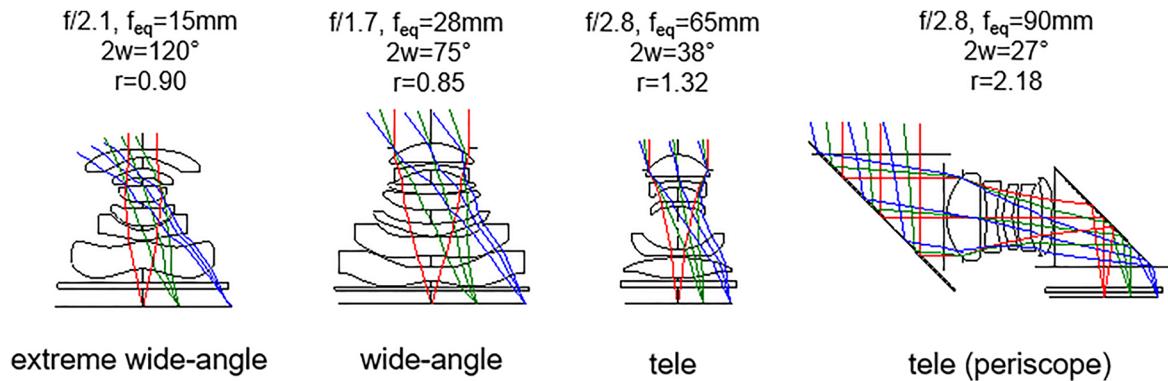


Figure 34: State-of-the-art optical design of today's typical SPC lenses with different focal lengths (Courtesy of ref. [59]: (a) Extreme wide-angle [60], (b) standard wide-angle of 28 mm [61], (c) Tele 65 mm [61], and (d) periscope Tele 90 mm [62].

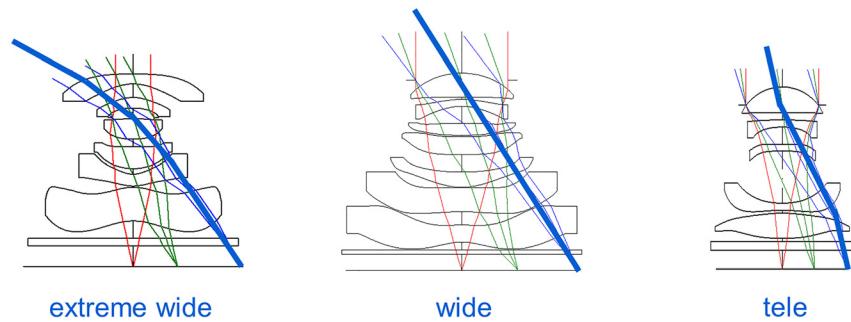


Figure 35: Comparison of chief ray path through extreme wide, wide, and tele lenses.

- MTF (vs. field or FOV and vs. spatial frequencies and vs. distance with respect to image plane)
- Distortion (radial distortion, TV distortion, and distortion grid plot)
- Relative illumination
- Aberration (e.g. spot diagrams, ray aberration curves, chromatic focus, and lateral shift)
- Angle of incidence at image sensor

In Figure 36 the image performance of the 4 designs are shown: “Extreme wide,” “wide,” “tele,” and “periscope tele.” The corresponding module sizes fit into a 6 mm available length as constrained by smartphone thickness (the standard wide-angle lens is extended slightly further outside the housing). The spatial frequencies for the first three graphs on MTF are chosen with reference to a 12 MP image sensor, that is 4000×3000 pixels. Since the image sensor sizes are different (5.8; 10; 4.4; 4.4) mm the “effective pixel pitches” are also different (1.16, 2.0, 0.88, and 0.88) μm and therefore so too is the corresponding Nyq. As explained earlier “effective pixel pitches” are not actual pixel pitches on the image sensor but refer to the “macro-pixel” size of multicell sensors. The first graph shows polychromatic MTF versus image field for spatial frequencies Nyq/2, Nyq/4, and Nyq/8. The second graph

shows MTF versus spatial frequencies up to the cut-off frequency $2 \text{NA}'/\lambda = 1/(\lambda \cdot \text{f-number})$, the third MTF through image focus for Nyq/4.

From these graphs it follows that the image performance of all lenses is diffraction-limited near the image center and drops off for the wide-angle lenses (but not dramatically as can be seen in the MTF versus field graphs for the relevant spatial frequencies). The through-focus region with very high contrast is roughly only about $\pm 10 \mu\text{m}$ for all lenses.

The next graph is a distortion graph on the 4:3 aspect ratio image field showing that distortion of all lenses is not noticeable (typically <1%). The barrel-type distortion of the extreme wide-angle lens is about 20%. In many camera modules this distortion is not corrected by software, although there are apps available that can correct this distortion. Note that it depends on the scene in question whether there is a need to correct this for the barrel distortion: Just photographing a 2D plane like a chess board fish-eye distortion is of course undesired; but when photographing a group of people barrel distortion compensates for the “egghead effect,” also known as “perspective distortion” [63].

Relative illumination drops more towards the corner of the image field for wide-angle lenses. The intensity drop-off, sometimes called “shading,” is corrected by software. Although shading is not apparent in pictures taken using

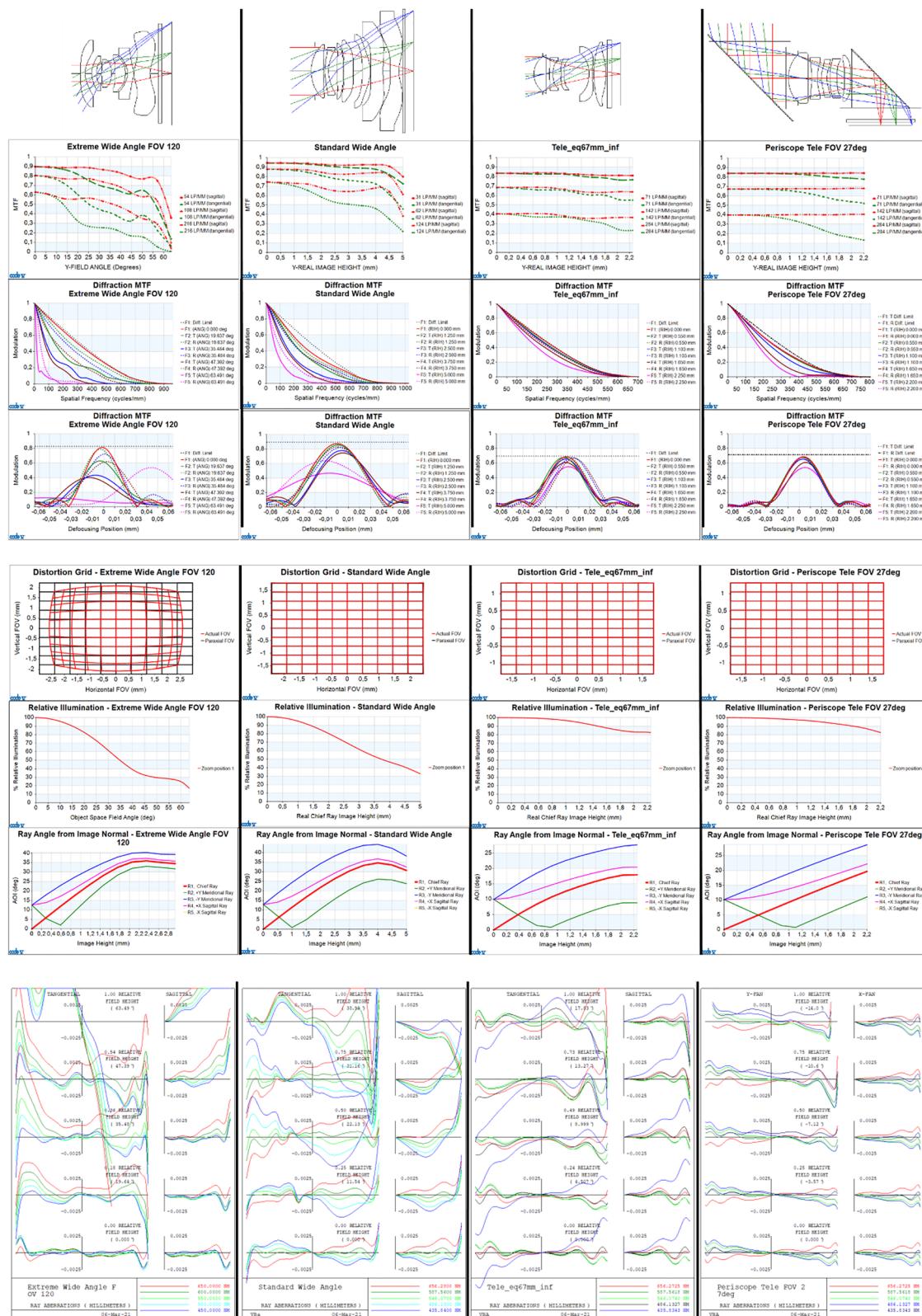


Figure 36: Optical design performance for lenses given in Figure 34: extreme wide-angle $f/2.1$, FOV 120° , wide-angle $f/1.7$, FOV 75° , tele $f/2.8$, FOV 38° , and periscope tele $f/2.8$, FOV 27° .

smartphones it gives rise to increased noise sensitivity by 1 or 2 EV for wide-angle photography.

The angle-of-incidence graph shows the chief ray and marginal ray angles in the image plane. In order to avoid further light loss the chief ray incidence angle is usually limited to about $<35^\circ$: In addition to the lenses' relative illumination, which includes a natural geometrical loss according to approximately $\cos^4(\text{AOI})$, namely $\cos^4(35^\circ) \approx 0.45$ (which is already included in the relative illumination graph), the oblique incidence on the image sensor results in further intensity losses (included in the software-corrected shading). In addition to the improvements obtained by back-side-illuminated image sensors and specific architecture, e.g. deep-trench structures (see Section 9), those losses can be further minimized by slightly shifting the micro lenses according to the incidence angle (that is a slightly smaller micro lens array compared to the pixel array).

The ray aberration graphs in the final line of Figure 36 show the specific aberrations on the tangential and sagittal image plane for different wavelengths versus field (image center on bottom, image corner on top). The scale is only $\pm 2.5 \mu\text{m}$. The aberration curves are very “wiggly” as residual aberrations of the compensation principle of lower-order aberrations by higher-order aberrations through usage of high-order aspherical lens surface deformations. Chromatic aberrations can be physically evaluated e.g. by image simulations of edge or line spread functions versus field and through focus. From that the “color fringe widths” can be evaluated defining the “number of colored pixels along a

high-contrast edge” under photographic worst-case situations [64].

6.3 Extreme wide-angle lenses

Extreme wide-angle lenses have been used in many high-end SPCs with FOVs of around 120° – 130° since around 2018. A distortion of around 20% is permitted. In this way, a short flatness factor of $r = L/\phi_{\text{im}} < 1$ is achieved. Due to the additional negative “bending lens elements” at the front, r is larger compared to a standard smartphone wide-angle lens. At the expense of a slightly larger r factor of about 1, it is possible to remove the distortion completely (examples for FOV 116° in ref. [65]). In the front part of the lens there are usually one or two lens elements in front of the diaphragm, which are strongly aspherical with a lens curvature that rises rapidly towards the edge of the lens and which bend the steep angles at 60° such that the angles at the diaphragm are reduced to less than 40° . Starting here, the beam path is comparable to that of a compact standard wide-angle lens. Accordingly, the structure of the system and the chief ray angles of approx. 35° on the image plane are similar.

Optical designs of extreme fish-eye wide-angle lenses with an even larger FOV of up to around 150° [66] and also over 160° are possible with about 50% distortion and a similar form factor [65]. Such camera lenses may well enter the smartphone market soon.

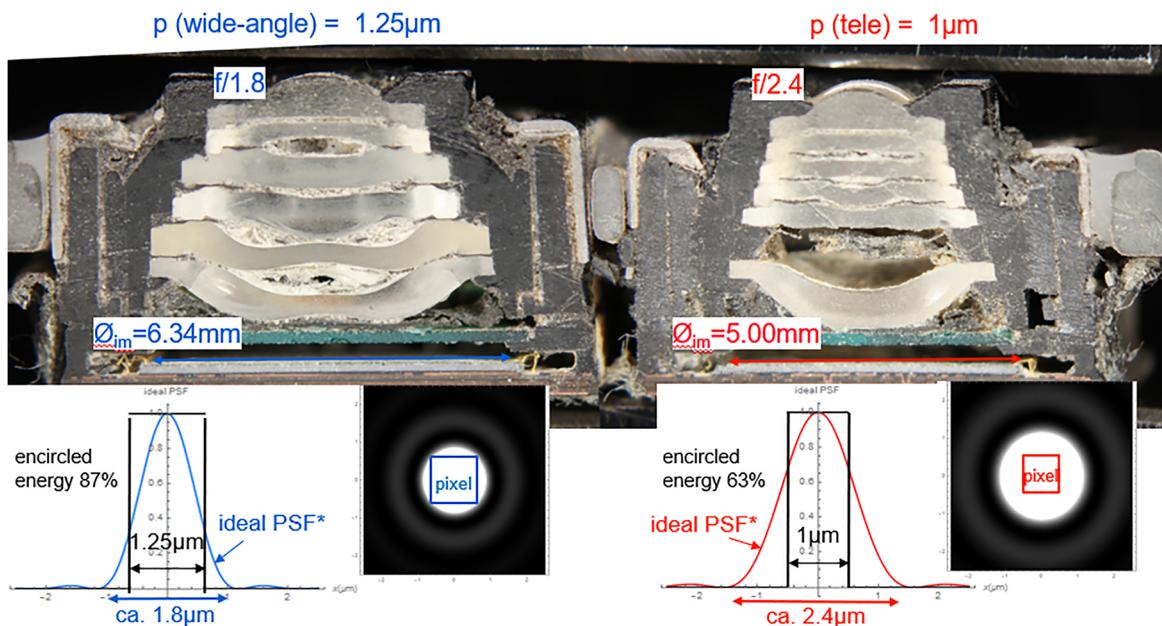


Figure 37: Cross-section of dual SPC with standard wide-angle lens (1.8/28 mm) and “tele lens” (2.4/56 mm). Due to the smaller aperture and smaller sensor size the optical resolution of the tele lens is weaker, i.e. the ratio of Airy spot diameter to pixel size is larger (bottom graph).

6.4 Tele lenses

Fulfilling the overall length requirement is even more difficult for longer focal lengths, such as those on the market for dual systems with a (hybrid) “optical zoom.” Therefore, a compromise is usually made with these lenses by using smaller image sensors and smaller apertures – both at the expense of optical resolution (Figure 37). The difficulties of compact telephoto lenses are due to the following: The longer the focal length, the smaller the required telefactor (TF):

$$\text{TF} = \frac{L}{f} \quad (40)$$

$\text{TF} < 1$ can only be achieved if the refractive power is positive in the front part of the lens and negative in the rear part. The smaller the TF, the more positive or negative refractive power is required, and the larger aberrations are introduced. Greater refractive powers lead to greater lens curvatures and these in turn lead to greater aberrations.

With this layout, the optical performance is severely limited as the focal length increases (Figure 38).

The ratio of the overall length to the image diameter $r = L/\varnothing_{\text{im}}$ increases and the size of the image sensor inevitably decreases due to the limitation of the overall length. According to the lens data in the example in Figure 34 with a 65 mm equivalent focal length $r = 1.32$, which means that the image sensor size is $0.85/1.32 \approx 0.64$ smaller than the standard wide-angle lens with the same overall length. That means that if we assume the image sensor size of the 28 mm wide-angle lens is 6 mm, then that of the 65 mm lens is only 4 mm. In addition, the telephoto only has an aperture of $f/2.5$ compared to $f/1.7$ of the wide angle: This means that the diffraction-limited resolution is weaker by a factor of $1.7/2.5 = 0.68$. Overall, this leads to a smaller number of “optical pixels” (meaning Airy spots per sensor area) by a factor of $(0.64 \times 0.68)^2 \approx 0.19$.

6.5 Periscope tele lenses and alternative tele concepts

The problem of the short telephoto lens can be avoided by rotating the optics by 90° in the housing. This can be achieved with a 45° mirror or with a 45° prism mirror (see Figures 34, 39 and 40): Then the optics can be much longer.

However, this “periscope layout” has another limitation: The mirror must of course be smaller than the depth of the housing and thus also the entrance pupil of the lens: This means that the longer the focal length of the lens, the f -number $K = f/\varnothing_{\text{EP}}$ increases and thus the diffraction-limited resolution becomes weaker. An approximately realistic entrance pupil diameter of $\varnothing_{\text{EP}} = 4$ mm means that even with a relatively small aperture of $f/3.4$ a focal length of only $f = K \cdot \varnothing_{\text{EP}} = 3.4 \cdot 4$ mm = 13.2 mm can be achieved. For a still realistic image diameter of 4.4 mm, this corresponds to an equivalent focal length of 130 mm. These data are close to what is feasible with this concept and can be found with similar data in some current multi-lens cameras and

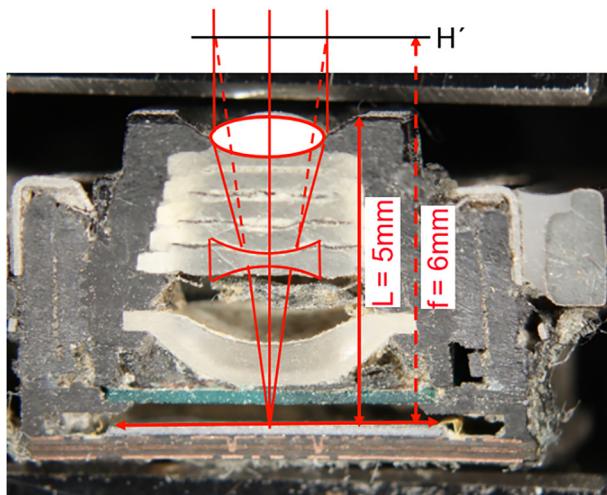


Figure 38: For lenses with an equivalent focal length of 56 mm, $f > L$, a telephoto design is required with a refractive power distribution of $(+, -)$.

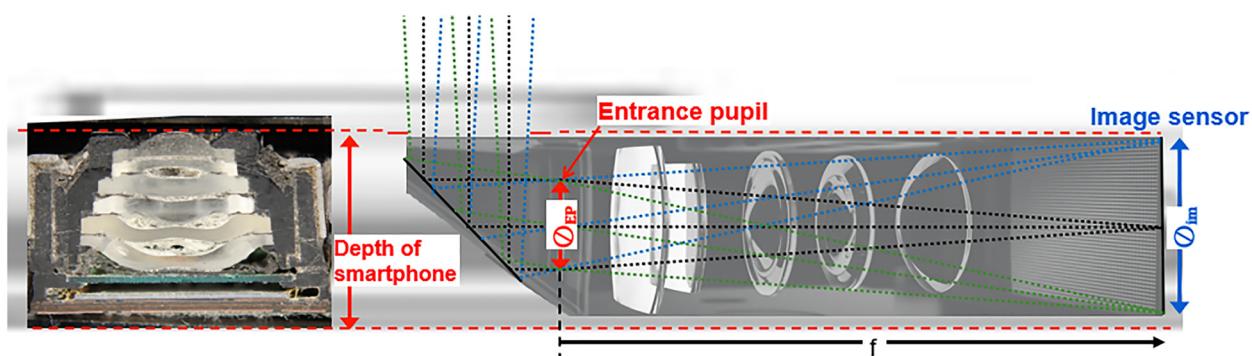


Figure 39: Standard tele lens (left) and periscope tele lens (right).

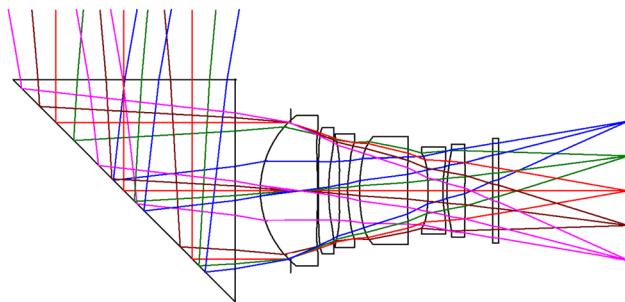


Figure 40: Periscope Tele lens [67]. $f/3$, $f = 14.55$ mm, $\phi_{\text{im}} = 5$ mm, corresponding to a FOV = 19.8° .

patent literature [67] (Figure 40), [68–70]. Alternative, but catadioptric periscope layouts enabling the desired form factor were proposed by Araki et al. [71]. An intermediate image is used to get sufficient space for folding and enabling a large aperture.

Carles and Harvey [72] take advantage of the fact that much more space is available perpendicular to the narrow housing depth, so that the entrance pupil can be made rectangular with a larger f -number (and therefore deliver better resolution) in the long direction. The contrast can be azimuthally homogenized by using several systems or rotating the system in combination with a fusion of the images of these subsystems.

Another alternative are catadioptric layouts with two mirrors in the front part [73, 74] or with several reflections between the mirrors [75, 76]). This layout allows for very small telefactors of $\text{TF} < 0.5$ (Figure 41) or even less for layouts with multiple reflections at the front mirrors. In principle, this allows a very large entrance pupil diameter to be achieved and therefore a high aperture ratio. Aperture ratios as large as $f/1$ are possible, but only with a correspondingly large central obscuration, i.e. the objective pupil is a narrow ring. This leads to a distinct loss of contrast in the lower spatial frequencies. In principle, very fine structures would be displayed with a higher contrast, which is irrelevant for smartphone dimensions because this high resolution cannot be used due to the available pixel sizes. Catadioptric lenses are known in photography, especially as telephoto lenses with very long focal lengths. They are extremely short compared to standard telephoto lenses. In addition to the loss of contrast, many photographers do not appreciate the “donut bokeh” caused by the obscuration as much, i.e. the noticeable ring-shaped out-of-focus highlights. This would play less of a role with SPC lenses, but out-of-focus lines tend to take the form of double lines.

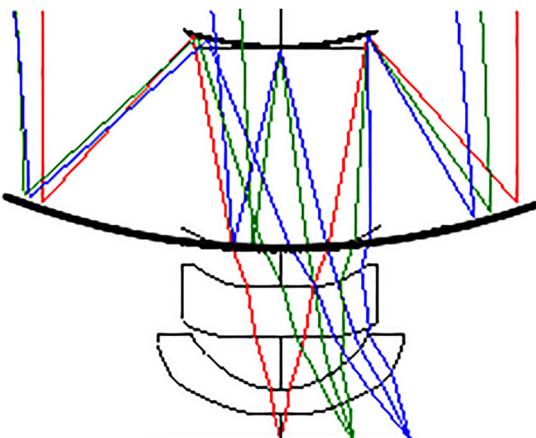


Figure 41: Rotationally symmetric catadioptric design [73], $f/2.4$ with central obscuration of $1/3$ of the pupil. For an overall length $L = 6$ mm, $f = 14.4$ mm (that is tele factor $\text{TF} = L/f = 0.42$), $\phi_{\text{im}} = 3.6$ mm, corresponding to a FOV = 13.6° .

7 Zoom

Due to space limitations, it is very difficult to implement good zoom systems in smartphones. There were various approaches to implementing zooms in cell phones early on. Figure 42 shows an overview based on some products. Since there were very good optical zoom systems for compact digital system cameras (DSC) as early as 2000, attempts have been made again and again to integrate these in various forms in the mobile phone, e.g. in a compact, foldable layout as in the Nokia N93 or as an extendable zoom as in DSCs (e.g. Samsung Galaxy s4). Of course, digital zoom was also available very early on in cell phones, albeit in a very modest resolution due to the image sensors, which were already limited in resolution: The Nokia Pureview 808 and subsequently the Nokia 1020 were pioneering and ahead of their time. They featured a huge 41 MP “Supersensor” (standard for high-quality smartphones at that time were 5–8 MP) for their time, and therefore also very high image quality even with a $3\times$ zoom.

And finally, since 2016, hybrid zooms through multi-camera systems using lenses of different focal lengths have become a standard. Every year, smartphones are equipped with more and more individual cameras, thus increasing the zoom range. This trend is likely to continue, but that does not mean that alternative concepts are dead: After the “super digital zoom” from Nokia from 2012/13 had no successor for a few years, wide-angle modules with very bulky image sensors have been around since 2018. They measure around 12 mm in diameter and are fitted as



Figure 42: Evolution and dead ends: Different zoom concepts in Smartphones.

standard in high-end smartphones. With a multicell sensor architecture, these image sensors usually contain more than 50 MP, sometimes even more than 100 MP.

7.1 Hybrid zoom in multicamera systems

The combination of different fixed focal lengths listed in the previous sections is referred to as the “zoom” in modern smartphone multicamera systems [77].

As explained in the previous sections, the achievable optical resolution of SPC lenses is heavily dependent on the specific optical design, which in turns depends on an equivalent focal length. To simplify matters, we assume a diffraction-limited optical resolution of $\text{res}_{\text{optics}} = 0.5 \cdot \phi_{\text{Airy}}$ (see Section 4.3), that is:

$$\text{res}_{\text{optics}} = 1.22 \cdot \lambda \cdot K \quad (41)$$

and counts all these resolved areas on the image sensor surface (4:3 aspect ratio with full diagonal ϕ_{im})

$$A_{\text{im}} = 0.48 \cdot \frac{\pi}{4} \cdot \phi_{\text{im}}^2 \quad (42)$$

Then the number of “optical pixels” (i.e. resolved areas) NP is

$$\text{NP}_{\text{optics}} = \frac{A_{\text{im}}}{\frac{\pi}{4} \text{res}_{\text{optics}}^2} \quad (43)$$

so

$$\text{NP}_{\text{optics}} \approx 0.32 \frac{\phi_{\text{im}}^2}{\lambda^2 K} \quad (44)$$

Figure 43 shows this “optical resolution in Mega-Pixels” $\text{NP}_{\text{optics}}$ for the aforementioned patent lens examples.

In the following simplified consideration, we include the digital zoom and the image sensor resolution in order to roughly estimate the total resolution of a hybrid multicamera zoom system over the entire focal length range. The hybrid multicamera system consists of several cameras with lenses of different focal lengths: f_1, f_2, \dots . Digital zoom reduces the resolution according to the cropped FOV. Since this FOV directly scales with the focal length (f), the value of $\text{NP}_{\text{optics}}$, the resolution drops according to:

$$\text{NP}_{\text{optics digital zoomed}} = \text{NP}_{\text{optics}, 0} \frac{f_0^2}{f^2} \quad (45)$$

where f_0 denotes the focal length of the lens corresponding to $\text{NP}_{\text{optics}} = \text{NP}_{\text{optics}, 0}$.

Figure 44 shows NP versus the equivalent focal length for the camera module consisting of four cameras with $f_{\text{eq}} = 15, 28, 56, \text{ and } 125 \text{ mm}$, as defined in Table 4: The blue curve shows the contribution of optics to the resolution (as just discussed), together with digital zoom from a shorter focal length to the adjacent longer one. The red curve shows the number of “macro pixels” $\text{NP}_{\text{sensor}}$ on the image sensor that is e.g. a 4-pixel of a Quad-Bayer sensor taken as one macropixel. The number of macropixels $\text{NP}_{\text{sensor}}$ is

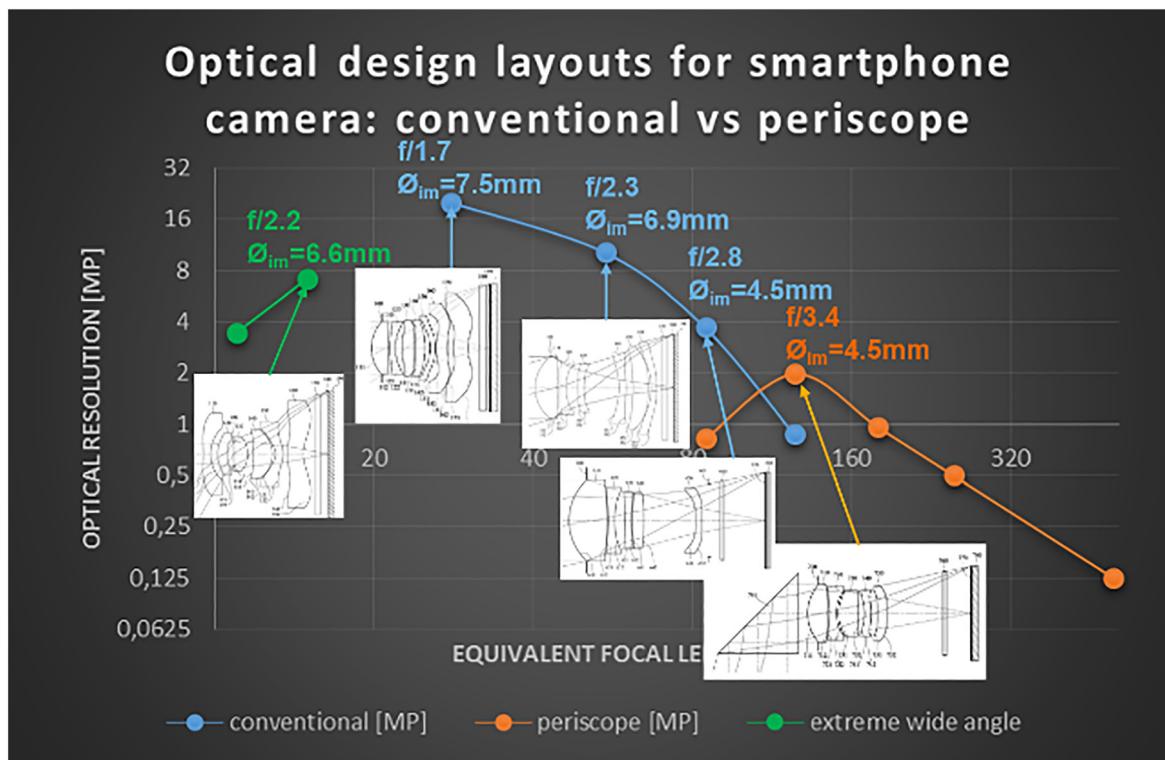


Figure 43: NP_{optics} , the number of optical resolution areas on the image sensor, for different lenses from the patent database with FOV 20–144° corresponding to equivalent focal lengths 8–125 mm. Beyond $f_{\text{eq}} = 125$ mm we can assume the same entrance pupil and image sensor size, but the corresponding f_{eq} , due to the fundamental periscope space constraint.

given by the total number of pixels on the image sensor $NP_{\text{sensor, total}}$ divided by the number of pixels within the cluster of pixels building a “macropixel” $NP_{\text{macro-pixel}}$:

$$NP_{\text{sensor}} = \frac{NP_{\text{sensor, total}}}{NP_{\text{macro-pixel}}} \quad (46)$$

The number of pixels per macropixel is namely $NP_{\text{macro-pixel}} = 1$ for a standard Bayer sensor and $NP_{\text{macro-pixel}} = 4$ or 9 for 2 × 2- and 3 × 3-multi-cell sensor, respectively. A well-balanced system should have about the same resolution of optics and image sensor. This is roughly the case for extreme wide-angle and standard tele camera lenses. However, for the wide main camera the optical resolution is clearly better than the image sensor resolution, so the sensor limits the overall resolution. The opposite is true for the periscopic long tele camera lens: Here the optical performance of the lens clearly limits the overall resolution. Note that the optical performance of the wide-angle lens is also better than the optical performance of tele and periscope tele lenses even at the same effective focal length, that is when zoomed in digitally. However, the actual resolution of the complete wide-angle camera

system is worse than the corresponding resolutions of the tele cameras, since the digitally zoomed-in pixel resolution is clearly worse for $f_{\text{eq}} > 56$ mm. So overall the tele lenses improve the image resolution for the considered setup. The white line in Figure 44 shows the minimum “optical pixel number” and “sensor pixel number,” that is the effective pixel number in this simplified consideration:

$$NP_{\text{effective}} = \min(NP_{\text{optics digital zoomed}}, NP_{\text{sensor}}) \quad (47)$$

According to the graph the effective resolution in the zoom range 15–125 mm varies between about 2 and 12 MP.

A more detailed analysis requires a simulation of all the steps of the digital imaging chain, which also includes demosaicing, image enhancement, the lens aberrations and the specific algorithms how images are fused by multiple cameras [78]. To a certain degree image performance can be improved through image fusion in the common FOV of both cameras. However, this requires some cumbersome joined camera module calibration and is increasingly difficult at close range, due to the parallax of the images caused by the spacing of the camera modules.

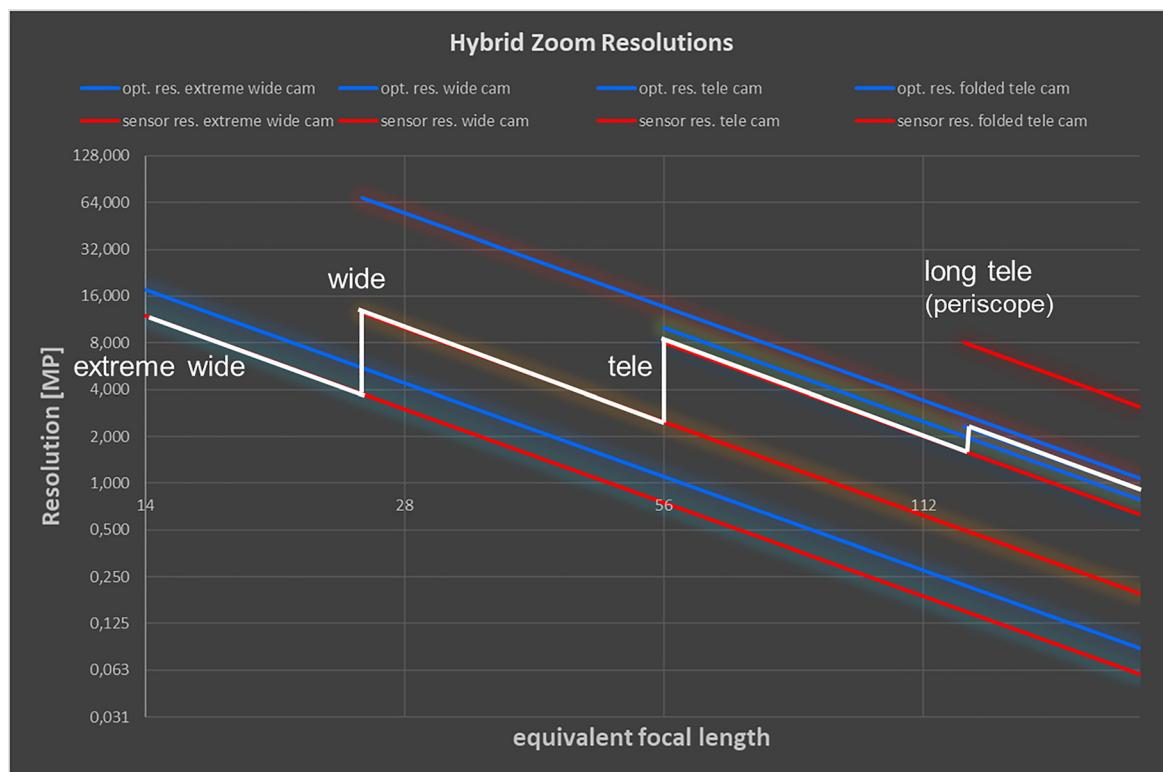


Figure 44: Optical (blue curves) and sensor (red curves) resolution of the multi-camera system according to Table 4 while zooming in with digital zoom until longer focal length camera takes over (without image fusion).

7.2 Optical zoom systems

Optical zooms have never truly caught on in the smartphone market. Since major manufacturers such as Nokia (N93) and Samsung (G810, Galaxy K Zoom, Galaxy S4 Zoom) have made some attempts on the market in the past and manufacturers continue to make announcements on optical zooms, we will deal with this topic in this section.

A classic optical zoom changes the imaged object frame, i.e. the FOV, by changing the focal length (f) of the system. In most zoom systems, the change in focal length is done by changing the distance between the lens elements or groups of lens elements. The total focal length changes because it depends not only on the individual focal lengths of its subgroups, but also on their distances. For example, in the case of two optical groups with focal lengths (f_1, f_2) the total focal length (f) is given by:

$$\frac{1}{f} = \frac{1}{f_1} + \frac{1}{f_2} - \frac{d}{f_1 f_2} \quad (48)$$

In contrast to “digital zoom,” “optical zoom” is often associated with “lossless image performance” over the entire focal length range. However, this is not necessarily the case. Especially for the tight space constraints of a

smartphone, the realized image performance of an optical zoom system can decrease significantly at long focal lengths, as shown in the following example:

In 2006 Nokia joined forces with ZEISS to launch a true optical zoom in the N93 model (Figure 42a). The mobile phone was designed to be “camcorder-like,” with a rotatable display and lens arrangement alongside the housing. Figure 45 shows the type of system used: The almost linearly moved “variator” changes the overall focal length and the front group of optics, the “compensator,” compensates for the focus deviation with a small nonlinear movement. The required movements of the optical groups of around 10 mm driven by voice coil motors (VCMs) took a couple of seconds. With an overall length of 20 mm and a 6 mm image sensor diagonal, the aperture drops from $f/3.3$ in the wide angle to $f/5.9$ with a long focal length and thus also the image sharpness due to the diffraction limitation.

For similarly tight space constraints, Kreitzer and Moskovich [79] describe a 28–85 mm zoom lens design with a continuously high $f/2$ aperture in a periscope design. This aperture for the entire zoom range is made possible with an intermediate image, which becomes smaller and smaller with increasing system focal length, but at the price of a length of almost 60 mm and 17 lens elements.

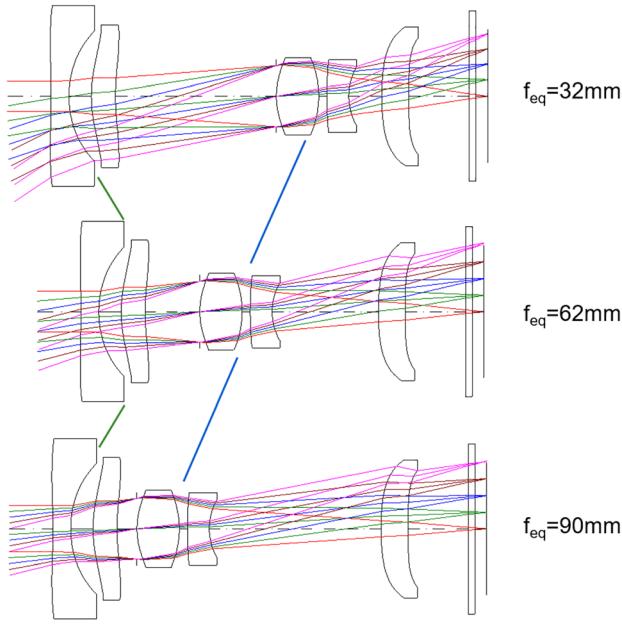


Figure 45: 3× standard zoom 3.3–5.9/32–90 mm.

Excellent image performance can be achieved with the relatively loose space constraints of a compact digital camera (DSC). In the 2000s, due to the strong DSC consumer market, the R&D activities on optical zoom lens design were quite significant. DSC zoom systems are highly complex and make use of digital-optical co-optimization. There is a large base of patent literature on compact camera zooms (e.g. [80, 81]).

The ratio between the overall length and the image sensor size in the example in Figure 46 is approx. $r = 4.5\text{--}5.5$, i.e. much larger than with the fixed focal lengths, but can be mechanically reduced to a ratio of $r = 2\text{--}3$. The opto-mechanical design of the camera lens including the zoom curves is complex and time-consuming. A significant proportion of lens elements, sometimes about half of all lens elements, are aspherical. In addition, digital aberration corrections are made: In the wide-angle range, a distortion of approx. 20–30% in the design is permitted and digitally corrected, which significantly reduces the size; the image is cropped somewhat depending on the zoom. Finally, as is usual with many zooms with a large zoom factor, the aperture in the long focal length area is reduced to limit the diameter in the front area, and the overall length.

Until around 2010, the compact digital camera market was steadily growing (Figure 2). The image quality of digital system cameras (DSCs) was clearly superior. In particular, the zoom functionality of compact camera zooms of DSCs was completely missing in mobile phones, respectively, and

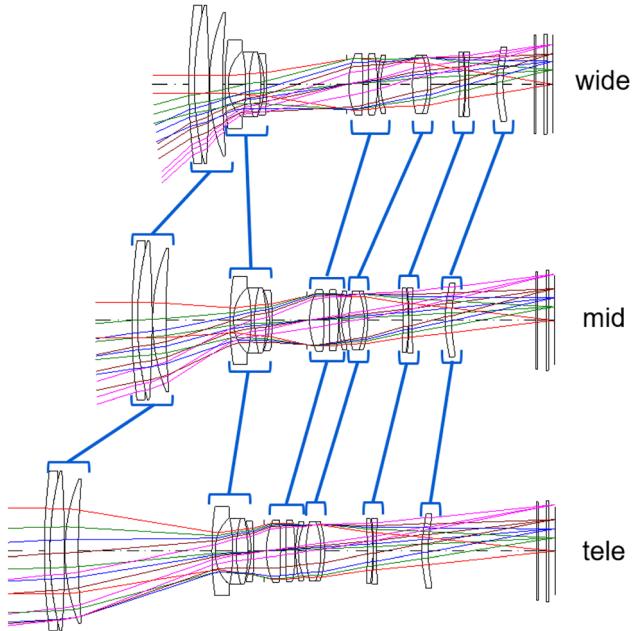


Figure 46: Zoom lens design of a typical compact camera, an equivalent 2.8–4.7/24–200 mm zoom. The structure has six movable groups, some of which have both zoom and focus functions, and is highly complex [82].

was far inferior with regard to its digital zoom. While in the 2010s image quality with SPC steadily improved – and in many everyday situations hardly compromised vis-a-vis DSCs – the DSC market declined. In response to this, several DSC camera makers substantially increased the zoom range of DSCs to more than 20× and even up to 50× – which is practically unachievable for SPCs – in an attempt to defend their market position. These attempts were unsuccessful. The large R&D departments of smartphone manufacturers put a lot of effort into further improving their cameras, which for most everyday situations became at least equal to the image quality of DSCs.

Although DSC sales today are smaller than SPC sales by about a factor of 200, several smartphone makers have implemented DSC-like optical zooms in their smartphones. However, attempts to sell a “DSC with phone functionality” proved unsuccessful due to the bulky shape of the smartphone.

To conclude this chapter: in principle, opto-mechanical zooms are feasible and brought to market by large manufacturers such as Samsung and Nokia, but without a sustainable trend. The hybrid multicamera zoom is widely accepted today (2021) and this is expected to remain so in the years to come, and to be supplemented by further camera modules, e.g. fish-eye wide-angle lenses.

8 Opto-mechanical layout and manufacturing

8.1 Plastic lenses: Key for miniature opto-mechanical layout

As already shown in comparison to classic design forms, plastic has been used to create the distinctly aspherical lens shape and it is the key to the small depth of SPC optics (for a detailed analysis, see ref. [83]). In addition, the high complexity of the structural shapes that can be manufactured with plastic makes it possible to implement not only the complex lens shape but also the mechanical mount in the same component (Figure 47). The reproduction accuracy of the components is in the sub-micrometer range. Special noncontact interferometric measurement technology is required to measure the small, complex components, often at steep angles of incidence [84]. Besides noncontact metrology, contact-mode measurement devices are also in use, e.g. the Panasonic UA3-P or Werth VideoCheck UA. Due to the very good reproducible component shape accuracy, the mechanical mount concept is also very simple: The plastic lenses can be stacked directly on top of each other in a barrel.

8.2 Opto-mechanical layout

The lenses are pressed into the plastic barrel either directly one behind the other or with spacers, and many are almost fully automated [85]. No adjustments have been made here, i.e. no measurements are carried out during the assembly process. MTF measurements are only carried out once assembly is complete. To improve quality, the

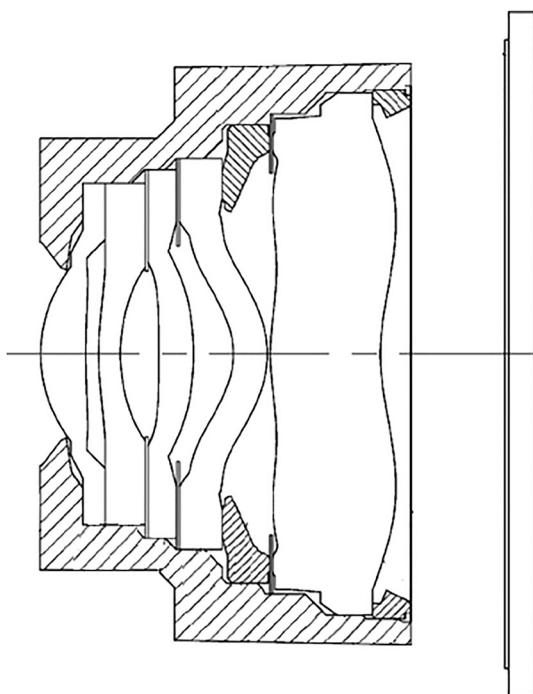


Figure 48: SPC lens mount concept [86].

individual lens elements are often matched to one another from the injection molding cavities.

The combination of “optics” and “lens-bearing mount” in a single component is a key to mastering the extreme manufacturing tolerances in mass production. Figure 48 shows an example of a mount concept. The lens elements are positioned directly on top of each other on the flat plastic mounting rings, often in a Sandwich with ring stops to prevent straylight.

The accuracy requirements of SPC lenses are extremely high. First of all, for very small lenses, the sensitivities also scale directly with the size difference of the system: For the same lens that is 7 times smaller for a camera with the same number of pixels, all position and geometry tolerances, measured in units of length, are also a factor of 7 smaller. In addition, due to their extreme shape, especially those in the rear part of the SPC lens, the aspheres are even more sensitive to decentering or tilting of lens elements. Critical tolerances are typically decentered and surface accuracy with tolerances for high-end modules of about $\pm 1 \mu\text{m}$, sometimes even less. Detailed analysis of sensitivities, tolerancing, and yield can be found in the references given in Section 8.5.

SPC modules are mostly used in several different smartphone models, sometimes over several years, so that quantities are in the range of millions, often tens of millions. This mass production enables low manufacturing costs of the aspherical lenses. The lens elements are

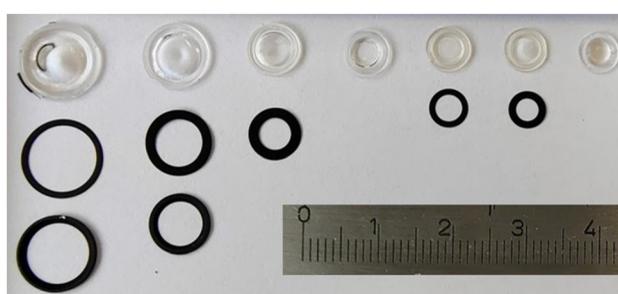


Figure 47: Lens elements of a standard wide-angle lens measured against a millimeter scale: First line: 7 aspherical plastics lens elements, the largest “w-shaped” lens element close to the sensor (pictured left), the first lens element close to the pupil (pictured right); below: Very thin straylight discs; last line 2 mounting elements positioned in between the last lens elements.

usually all made of plastic. They are manufactured using an injection molding process at temperatures of around 90–170 °C. Depending on the type of plastic, it is either injected or pressed into precise molds while still in liquid form. The melting temperatures are lower than in the corresponding aspheric pressing process for glasses. Plastics are not as rigid and stable as glass [87]. A lens is manufactured after about a minute. The cost of making the precise molds is in the tens of thousands of dollars. If the production runs with high yield, then the tool costs quickly pay for themselves.

The disadvantages of plastics are the relatively low refractive indices with a large dispersion (Figure 49), which complicates the optical design somewhat [88]. Precise information can be obtained from plastic optics suppliers, e.g. OGN, Zeonex ([http://www.ogc.co.jp/e/products/fluoren\(okp.html](http://www.ogc.co.jp/e/products/fluoren(okp.html), <http://www.zeonex.com/Optics.aspx.html>). The disadvantageous dispersion and refractive index properties could be mitigated in the future with the use of nanocomposites [89, 90].

There are some technological limitations, for example there are no cemented lenses like there are photo lenses made of glass. Another general disadvantage is the thermal

sensitivity, which is one to two orders of magnitude higher, both in terms of the refractive index dn/dT and the expansion coefficient. However, because of the small size, these high sensitivities are much less significant than with a full-frame lens: the thermally induced wavefront deformation scales directly with the size, i.e. it is about a factor of 7 smaller than with a full-frame lens. The dominating aberration type caused by temperature change is usually a focus shift which is compensated for by the camera modules with autofocus. To a lesser extent, other aberrations such as field curvature can also arise with temperature changes. Systematic passive thermal compensation strategies for plastic lenses have been around for some time now [91]. In practice, like when skiing in low-temperature environments, the situation is often somewhat relaxed, as the smartphone is usually worn close to the body, and due to the waste heat from the electronic components in the closed housing.

8.3 Active optical assembly

With active optical alignment (AOA, sometimes abbreviated to AA), the lenses are aligned with the image sensor and glued in with UV adhesive [92, 93]. The AOA runs fully automatically and takes just a few seconds. Today the cycle takes around 2–5 s per module, having been continuously reduced over the years. During the assembly process on the sensor, the barrel is aligned in the degrees of freedom: centering x/y , tilt x/y and focus distance (Figure 51) until the specified spatial frequency response (SFR) values are reached simultaneously over the entire image field (Figure 50). The assembly accuracies of the robotic machines are in the order of $\pm 1 \mu\text{m}$ for $x/y/z$ position and about $\pm 0.005^\circ$ for the angular position $\theta_x, \theta_y, \theta_z$. Modules that have not reached the specification within a certain period of time are marked as “scrap” and rejected. The quality, i.e. the measured MTF values, is continuously monitored over

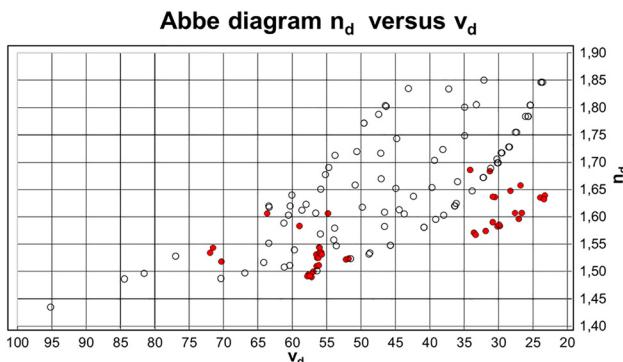


Figure 49: Abbe diagram (refractive index n at $\lambda = 587.6 \text{ nm}$ versus Abbe-number $v_d = (n_d - 1)/(n_F - n_C)$) for optical glass (white circles) and optical plastic (red circles).

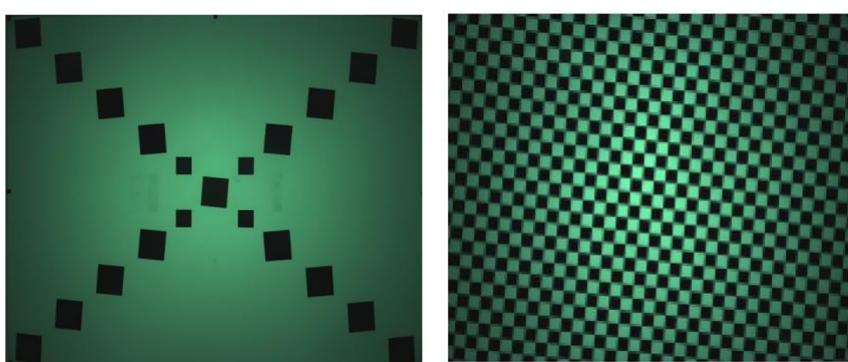


Figure 50: Typical test chart for MTF evaluation during active optical alignment: slanted edge charts [94]. The contrast transfer function of the individual module and the function of the autofocus or the setting of the fixed focus are also checked separately using these charts.

long periods of time and the process is interrupted if the yield threatens to fall short of the target.

Leading suppliers of these robotic machines include ASM, HVS, IsMedia, and Pioneer, who usually supply different types of machines for the complete smartphone packaging process (electronics assembly), among other equipment. There are also dedicated machines for the assembly of dual- or multi-camera modules which adjust, in 6 axes, the optical axis between the camera modules and then glue the modules into a single housing.

8.4 Tolerancing and yield analysis

The manufacturability and desensitization of the manufacturing and assembly tolerances of a lens requires detailed knowledge of the manufacturing technologies on the part of the optical designer. "Yield optimization" is a central topic of product development in close coordination with technologists in production, often at different production locations, e.g. in China or Taiwan at the same time. If you set the permitted deviations from the theoretically achievable image performance too low, you run the risk of disappointing customers with poor-quality products. If you set the limits too tight, you risk large rejects or long delivery times.

Finally, during the production process the quality of the optics is qualified on the basis of MTF values [95]. Tri-optics ref. [212] is a supplier of MTF measurement equipment. This happens during the final inspection done by the optics module supplier for the system integrator, and partly during the receiving inspection performed by the system integrator. The optics are not yet connected to the image sensor. The system integrator then actively adjusts the optics to the image sensor (previous Section 8.3). The final image performance of the SPC is qualified with SFR measurements. SFR is the common notation of the MTF of

the complete system optics/sensor, while the term MTF is commonly used for the optics qualification only. The yield for the optics module output inspection is typically around 50%, while for active optical alignment of optics to image sensor, a yield greater than 95% or more is aimed for because of the significantly higher component costs.

In the case of SPC lenses, reject analyses, or to put it in the common positive term, yield analyses, are an essential part of the development work [96, 97]. These are done in the final stages of optical design with Monte Carlo analyses: The optical designer uses his/her sensitivity analyses to set tolerances for the lens elements (radius, thickness, aspherical deviations or deviations in the refractive index, and dispersion of the plastic) and their relative positioning errors in assembly (decentering, tilting, spacing deviations, etc.) and probability distributions of those individual errors (e.g. Gaussian or "top hat" distributions matching typical statistical distributions of the manufacturing data). Then he/she "rolls" many different systems, where each component's tolerance assumes a random value according to the assumed probability distribution. The result is an ensemble of many different realizations of the system with slightly different radii, aspherical deviations, refractive indices, etc. For all these systems MTF data gives a statistical distribution which can be evaluated for production yields with respect to a system MTF specification (Figure 52). With respect to the final active optical alignment of the mounted lens to the image sensor the compensation of aberrations is taken into account [98] (Figure 53). In the optical design process, these analyzes are preceded by various methods of optimizing as-built performance by desensitizing lens aberration [99–101].

8.5 Wafer-level manufacturing

SPC wafer-level manufacturing has been seen as a very promising endeavor for low-cost mass production, especially when it comes to thin devices [102, 103]. Although it has been believed for some time that wafer-level manufacturing will replace assembly-based manufacturing [104] the breakthrough is still pending. Nevertheless, in recent years with the rise of new modules like 3D acquisition systems, some parts have now been produced through wafer-level mass manufacturing. It will be interesting to see whether the technology will take off in the next few years.

Related to lithographic production are microelectromechanical systems (MEMS)-based sensors (e.g. gyroscope, accelerometer) as well as photonic chips (photonic integrated circuits (PICs)), which might play a role in consumer smartphones in the future. Multiple photonic functionalities in the form of building blocks

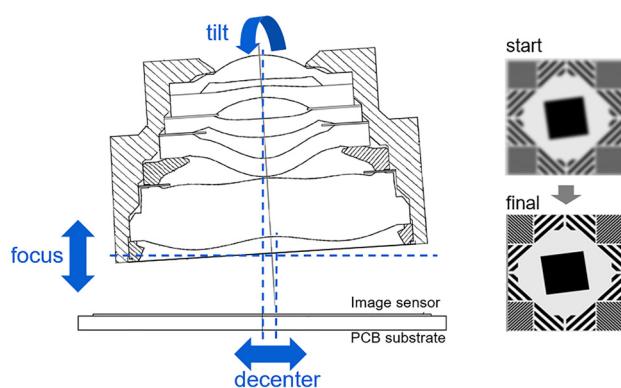


Figure 51: Active optical alignment of lens module to image sensor.

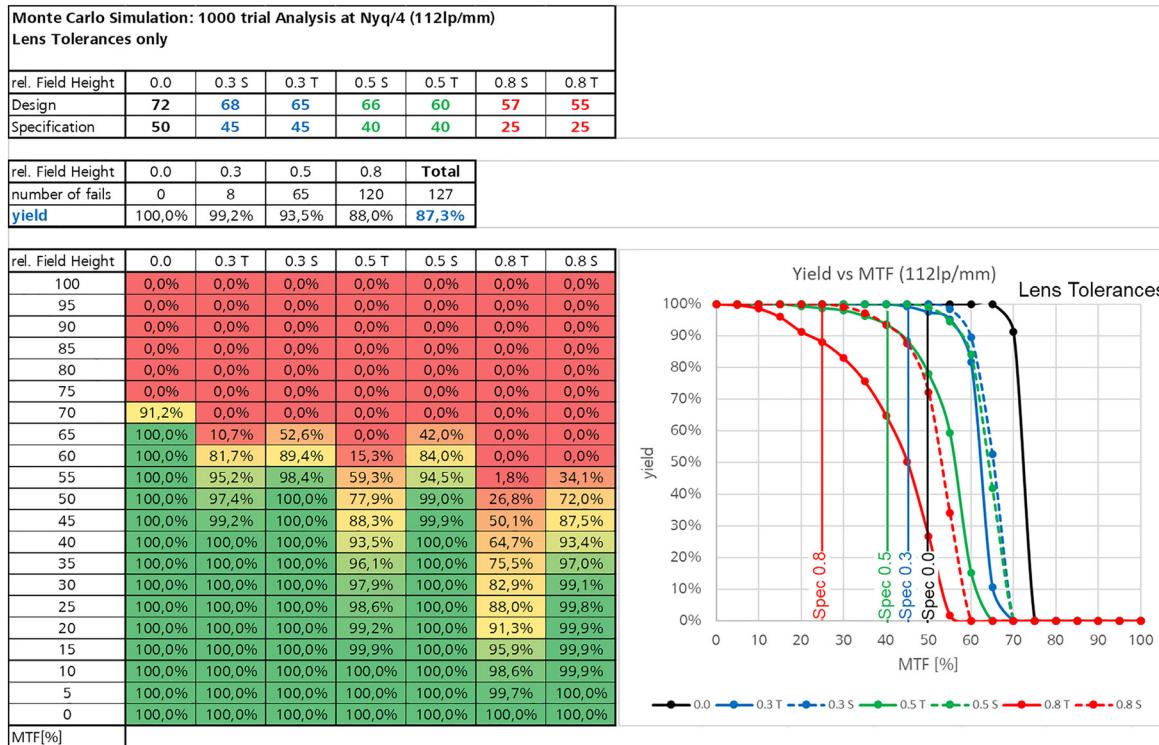


Figure 52: MTF vs. yield at different field positions and azimuths.

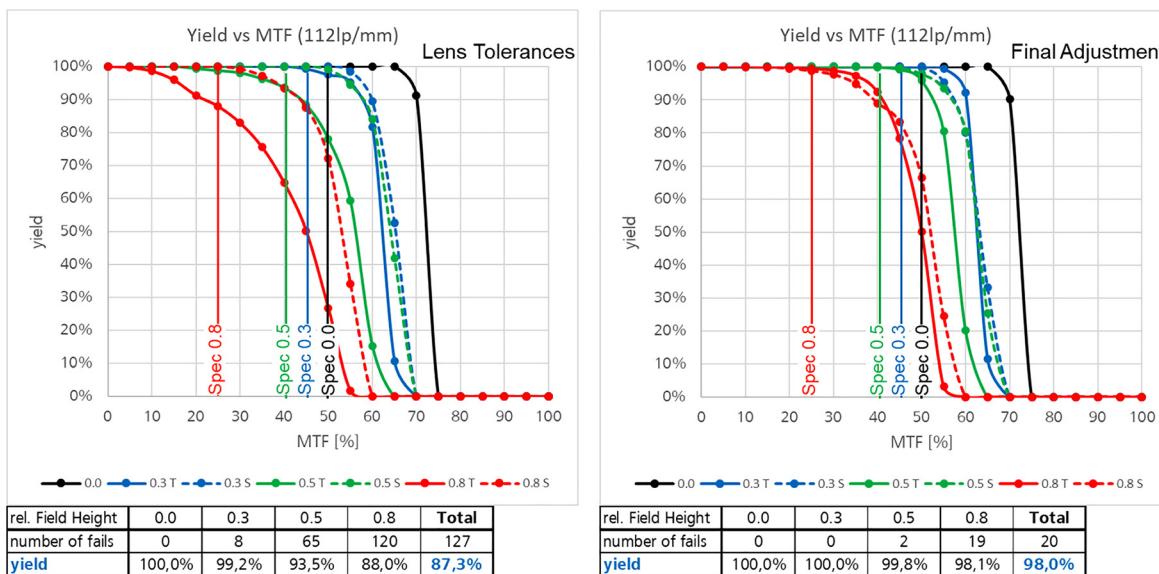


Figure 53: Yield-analysis incl. lens tolerances only (left graph); compensation by active alignment improves the yield from 87.3 to 98% for the final camera lens qualification.

are integrated into a (usually semiconductor) substrate. Semiconductor manufacturing technology like lithography, epitaxial growth and etching is used for production. Scalable, low-cost mass production makes PICs a promising candidate, e.g. for lidar sensors in selfdriving

vehicles. Optical functionalities (e.g. switchable light sources or sensors) can be co-integrated and combined with electronics and MEMS. As for electronically integrated chips, PICs promise to be robust and energy efficient [105].

8.6 Anti-reflection coating for plastic lenses

Anti-reflection coatings are especially important in scenes with a large dynamic range, especially in the presence of bright light sources. Residual light reflections at lens and image sensor surfaces can result in unwanted straylight or “ghosts” on the image plane (see Section 14.2).

According to the principles of the invention of Smakula [106] at Carl Zeiss in Jena, reflections on optical surfaces can be reduced through the destructive interference caused by a single or multilayer coating by properly choosing the layer thicknesses and material refractive index. In today’s camera lenses, as well as SPC lenses, the AR coatings are multilayer coatings typically consisting of a succession of 2 or 3 materials of thicknesses of about 10 to a few 100 nm. Usually a low ($n < 1.5$) and a high ($n > 2$) refractive index material is used in succession. A multilayer AR coating typically consists of 6–8 layers.

Because of different adhesion, lower melting temperatures and other physical differences, the coating process for plastic cannot easily be transferred to glass [87]. Many low-index materials, like magnesium fluoride, are not suited to AR coatings due to poor mechanical properties of layers deposited without substrate heating and the high-tensile growth stress. Alternative materials of standard materials like MgF_2 ($n = 1.38$) include SiO_2 ($n = 1.45$), leading to minor compromises in performance or greater effort, e.g. more layers being required. Nevertheless, comparable reflectivities, such as in glass AR coatings like $<0.5\%$ or less within the visible spectrum and a wide range of incidence angles, are possible and common in mass production.

And when it comes to glass, vapor deposition or sputter processes are preferred for plastic lens AR coatings [107–109]. Due to the significantly deformed surface geometries in SPC lenses, for a large range of different angles of incidence, a uniform coating thickness is even more difficult. Of course, complex machining kinematics such as planetary gearings are out of the question in the mass production of the small lenses massively parallelized and certain inhomogeneities in the coating thickness are accepted. Recently an alternative new process, atomic layer deposition (ALD), was applied to the vivo X60 Pro for performance improvement due to more uniform coating thicknesses [110].

9 Image sensor

The digital photography era started at the end of the last century, when charged coupled device (CCD) image

sensors, which were already being used in scientific applications from the 1970s on, became widely available. This invention [111] was honored with the Nobel Prize in 2009. From the beginning, SPC contained CMOS sensors, and now almost all digital consumer cameras also feature them (as of around 2010 onwards). The metal-oxide-semiconductor (MOS) active pixel sensor (APS) was developed by Tsutomu Nakamura at Olympus in 1985 [112]. The CMOS active pixel sensor (CMOS sensor) “camera-on-a-chip” was later developed by Eric Fossum and his team in the early 1990s [113–116]. A detailed description of digital image sensors can be found in refs. [27, 117].

A CMOS sensor is a matrix of semiconductor photodiodes that detects the irradiance distribution on the sensor surface. According to the irradiance distribution on the sensor chip and the exposure time (T), electrons are generated as charge carriers in the individual photodiodes and converted by capacitors a voltage is finally generated [118]. The voltage is amplified and digitized resulting in a digital value, e.g. a number between 0 and 255 for an 8-Bit image.

The probability of whether a photon, which enters the sensor finally generates an electron in the photo-electric layer is the quantum efficiency (QE) (often also denoted “ η ”), that is $0 \leq \text{QE} \leq 1$. QE depends e.g. on transmission (coating and material absorption) and geometry of micro lens above the pixel and the entire light path past the 3D electrical circuits. In addition to the sensor architecture, QE also depends on the wavelength of the light, the angle of incidence and the numerical aperture of the incident light.

Novel technologies have been introduced to overcome the problems of CMOS sensors with very small pixel sizes. Over the past 10–15 years, there has been an improvement in quantum efficiency of miniature CMOS sensors, especially in the chip architecture [119], from $\text{QE} = 0.3\text{--}0.4$ to $\text{QE} > 0.7$. For example the photon absorbing active silicon layer (epitaxic layer) was increased by about a factor of 2 over the past 10 years. With “deep trench isolation” (TSMC patent [120]; Omnisvision patent [121]) walls are built between the pixels enhancing the QE and reducing cross-talk. With stacking technology the light sensitive rear illuminated photodiode array is separated from the electronics. A valuable resource of trends and developments on digital image sensors is the “Image Sensors World” Blog by Vladimir Koifman ref. [209]. For a comprehensive summary of very small pixel pitch CMOS evolution see ref. [122]. The majority of the image sensors have been made as back-side illuminated sensors (BSI) [123] since around 2010, which have a higher sensitivity than a front-side illuminated sensor because the light at the BSI travels along a shorter, undisturbed path from the micro lens to the photoelectric layer (Figure 54).

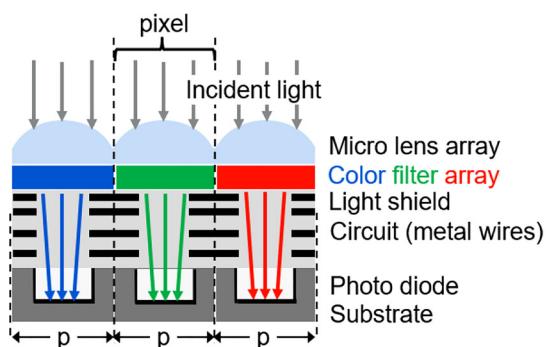
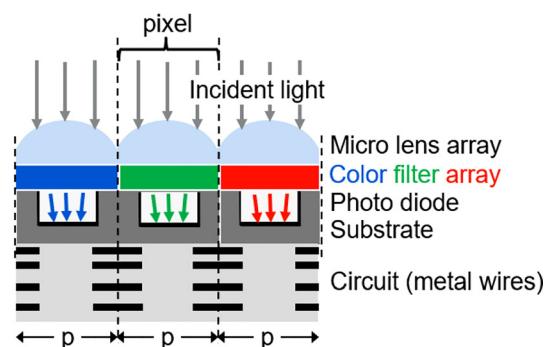
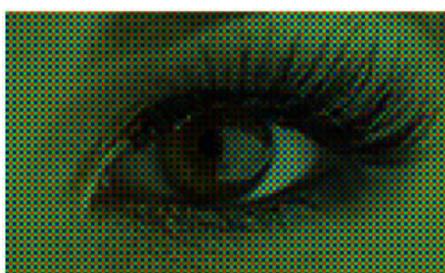
Front-side illuminated CMOS Sensor**Back-side illuminated CMOS Sensor**

Figure 54: Front-side and back-side illuminated CMOS sensors.

Each individual photodiode is provided with its own electronic circuit, specifically a readout amplifier, and can be read out individually at each XY coordinate. Together, the cell and the electronic circuits form a pixel (**picture element**). And with the wiring, it represents the cell of a pixel, which is why only a portion of the cell surface is sensitive to light. A microlens is attached in front of each cell, i.e. a microlens array over the entire sensor. This collects as much of the incident light as possible, including that which would otherwise hit the electronics on the light-sensitive photodiode, and also avoids shadowing within the cell structure. The voltage level on the individual photodiodes, and thus the image signal, depend solely on the respective brightness and the exposure time.

A color signal is obtained by placing a color filter directly below the microlens in front of each individual cell. Similar to the human eye, color information is detected with 3 different types of color sensors. Bryce Bayer developed and patented this concept while working at Kodak in the 1970s [124]. Each pixel transmits only a limited spectral range, for example in the red–green–blue (RGB) color model with a red, green or blue filter. The spatial arrangement of the individual color filters is often implemented using the Bayer mask as RGGB. Some cameras also use different arrangements and

sometimes different color models such as CYGM or, more recently, RYYB. Since each pixel is now only sensitive to one color, more precisely a specific spectral range, the missing color information must be estimated from the signal of the neighboring pixels (Figure 55). There are different interpolation models for this. These are considerably more complex than a bilinear interpolation of neighboring pixels of the same color, with case-sensitive weighted averaging over larger areas, also considering the brightness of pixels of different colors [125]. The interpolation routines can also vary in a context-sensitive way in the image field, e.g. on high-contrast edges compared to quasi-homogeneous image areas. Of course, an interpolation calculation cannot reliably determine the missing signal, e.g. the red and green components at the position of a diode behind a blue filter, and thus the correct color and brightness value. In practice, clearly visible artifacts are seldom recognizable, and then only when the image is greatly magnified. One consequence is that color cameras with a Bayer or other color mask in the individual spectral ranges – and thus overall – have a lower resolution than monochrome cameras. Especially since only 25% of all pixels are sensitive to blue and red, respectively, and 50% are sensitive to green. For this reason, there are monochrome camera modules for some SPC multicamera systems with



Raw Image

Interpolation
("Demosaicing",
"De-Bayering")



Interpolated ("De-Bayered") Image

Figure 55: Interpolation (demosaicing) of a raw image.



Figure 56: Rolling shutter effect on fast-moving objects: Picture taken from inside a moving car; rolling shutter effect significantly tilts road sign toward the foreground (left); moving propeller appears highly deformed due to rolling shutter effect (right). Images taken with SPC cameras. (Left-hand image courtesy of Richard Gerlich.)

subsequent fusion of the high-resolution image from the monochrome camera with the color information from the second camera (e.g. Huawei P9).

In contrast to high-quality DSLRs or mirrorless system cameras, in which a mechanical shutter leads to the fast, almost simultaneous exposure of all pixels, exposure control with SPC is carried out purely electronically. This means that here, in contrast to DSLR, exposure is done during the readout. In the case of CMOS sensors, this means that the individual photodiodes are not exposed and read out at the same time, so runtime effects do occur (rolling shutter effect, Figure 56). As a result, rolling shutters are especially limiting in slow motion mode. However, there are new CMOS image sensor developments, especially for industrial machine vision applications, that

already feature an electronic “global shutter” (for high-speed slow-motion pictures in smartphones: Sony (2017)).

Silicon has a monotonically increasing sensitivity from blue towards IR. Due to this strong absorption of silicon in IR and in order to limit the complexity of the RGB filters, an additional IR band pass filter is used to block the residual significant portion of light in IR.

Figure 57 shows the spectral response of the sensor’s RGB filters, including the transmission of the IR filter for the standard cameras working in visible-shortwave-infrared (VIS). Please note that for special camera modules like for 3D face or iris recognition, the IR cut-off may be at, e.g., 840 nm or 950 nm. In general, the lens transmission and the external window glass absorbs some light and varies with wavelength. The transmission of these

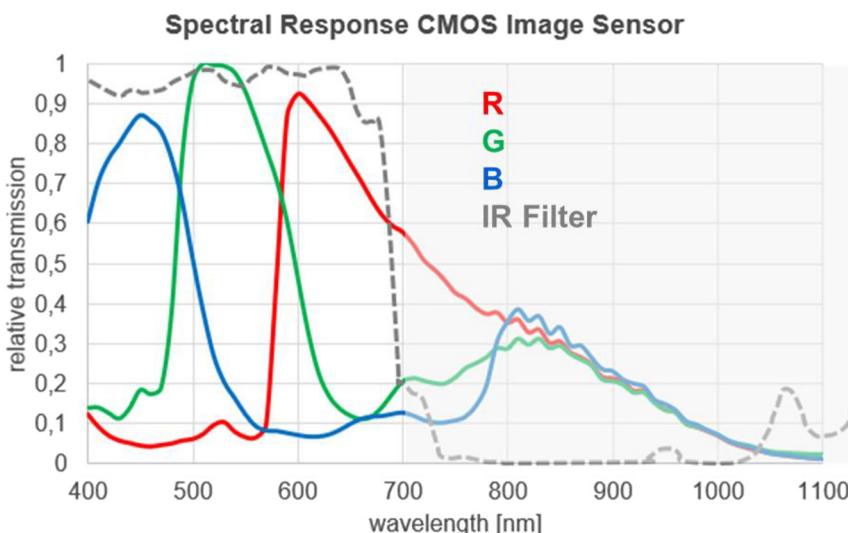


Figure 57: Relative sensitivity of a diode behind a red, green or blue filter of the CMOS sensor of an SPC. The transmission of a typical IR filter is also shown.

contributions depends especially on the specific coatings as well as on the spectral material transmission of the lens elements [126].

The relation between the physical intensity of the light (horizontal axis) received by the image sensor and the output digital numeric value (vertical axis) is described by the so-called opto-electronic conversion function (OECF, [127]). The OECF is often simply designated as ‘response curve’ or ‘characteristic curve.’ The exact term for “intensity” in the image plane is “irradiance” as used in radiometry, which is the radiation power per area, measured in Watts per square meter (W/m^2). For visible light, the Commission Internationale de l’Eclairage (CIE)-standardized luminous efficiency function of the human eye at daylight is used to define the corresponding photometric variable called “luminance”, which is measured in Candelas per square meter (cd/m^2).

Let us consider a single photodiode: To generate a measurable signal, there must be a minimum brightness on the pixel (minimum signal). This corresponds to the noise limit. The relationship between the number of electrons generated in the diode and the luminance is described by the photoconversion curve (see Figure 58). The diode signal rises from the noise limit. With current high-end smartphone CMOS sensors this is between 1 and 2 electrons until it saturates (maximum signal). The number of generated electrons is called full well capacity (FWC) and thus represents the capacity of the potential well of the diode. Typical values for the current CMOS sensors under consideration are 4000–5000 electrons. The FWC of good SLR cameras is more than one order of magnitude higher.

The abscissa is shown in exposure values (EV). The zero value was chosen here so that it is close to the saturation value, so that the noise value is 10 EV, i.e. a factor of 2^{10} below saturation. Here the exposure levels result as the logarithm of two of the number of incident photons divided by their value when saturation occurs. The change by +1 EV corresponds to an increase in exposure by a factor of 2. The ratio of maximum and minimum signal is called the dynamic range.

An analog-to-digital converter converts the voltage generated by the diode into a digital value. Simple, and especially older, SPCs use 8-bit converters to display signals from 0 to 255. Newer and high-quality SPCs and camera models also use 10-, 12-, or 14-bit converters. However, a higher bit value in the analog-digital converter does not necessarily lead to better image quality. This depends entirely on factors such as the noise behavior of the image sensor and the image motif.

10 Image processing

The basic image processing function on a digital camera is used to create the most natural image possible of the subject, and perhaps enhance it somewhat in terms of contrast or color rendition. In recent years in particular, and promoted by social platforms such as Instagram and Snapchat, the cosmetic improvement and alienation of photos of people taken on smartphones has become established and grown in importance. In Asia, smartphones without “beautifying modes” are practically unsellable. What is more, they even influence the general ideal of beauty today. Moreover, the Association of German Plastic and Aesthetic Surgeons has noted that an increasing number of people are opting for cosmetic surgery in order to recreate this heavily filtered and distorted “Insta look” in real life [129].

Advanced image processing often combines or fuses several images from the same or different camera modules, including 3D sensors, and uses them to generate a better image. This is known as computational imaging, or computational photography. With computational imaging, the raw images can also be “encoded,” e.g. in the form of subfields or subapertures of the light field. Alternatively, coded or phase-distorted apertures can be deployed in combination with deconvolution, e.g. to extend the depth of field (EDoF). (For a broad overview of computational imaging, see ref. [130]. Current research and development in mobile imaging focuses on artificial intelligence, machine learning and other software improvements. Meanwhile, augmented reality applications are driving further

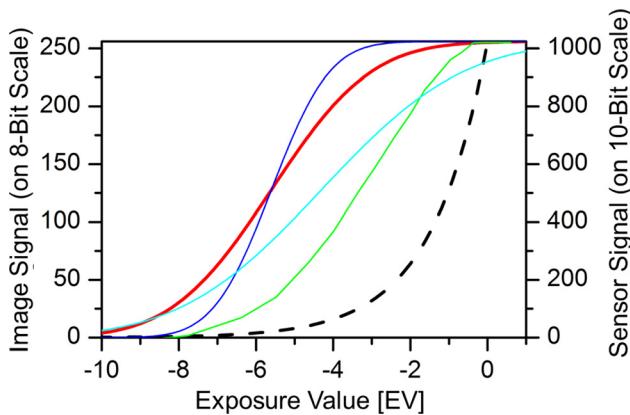


Figure 58: Typical camera response function (sensor signal; dashed black) and various tone value curves (image signal; other curves) (Courtesy of ref. [128]). All these tone value curves provide very different images. Please note: Due to the definition of EV, the x axis is a logarithmic axis and the almost linear photoconversion curve is shown as “curved.”

improvements in 3D acquisition and image recognition. Great hope and large investments, e.g. from companies like Magic Leap, continue to be poured into superior interfaces like augmented reality (AR) glasses [131].

Before an image is saved, a great deal of work needs to be done by the camera's image processor. In addition to demosaicing, the photoconversion curve is fitted into an 8-bit brightness scale (for each of the 3 color channels) by means of a tone value curve, as is required for the image to be displayed on the usual screens (smartphone screen, TV, PC monitor, etc.) or for the purposes of a photo print. As with all photographic film, the tone value curve considers the logarithmic sensitivity curve of the eye. A simple linear rescaling, especially of a large dynamic range, would not approximate human perception. The distinction between brightness tones, as well as other human senses such as hearing, touch, weight estimates, etc., scales according to the Weber–Fechner law [132, 133] rather than according to a logarithmic scale, from which one can also mathematically (formally) show that this minimizes the maximum errors in distinguishing between signal differences [134].

Furthermore, other color adjustments are usually made, including white balance. The image is sharpened and the noise is reduced. In principle, these are competing operations, where the trick is to apply greater sharpening to medium or low spatial frequencies to which humans are particularly responsive and to remove noise (convolution) moderately, on high-frequency areas according to the noise amplitudes. This has its limits, especially with complex motifs for low-light shots.

In recent years an increasing number of SPCs has been outputting the images in RAW format with a typical bit depth of 10, 12, or 14 bits for further processing, e.g. on a computer. As a rule, however, the images are compressed.

The image is then usually output in an 8-bit JPG file. So, only 8 bits of information are available per pixel, whereby the associated color component is still known due to the Bayer mask.

Different tone value curves are generally used for different photographed scenes to optimize the image. These can consider an overall increase or decrease in the contrast or brightness, or only in the dark areas of the image. The image quality has been significantly improved by adapting specific image processing algorithms to the respective exposure parameters (ISO, exposure time) as well as the properties of the recorded scene (brightness, color, structure, etc.). To do this, operations and algorithm parameters must be saved for a large parameter space.

The quality of the individual algorithms, such as demosaicing, tone value adjustments, noise removal, etc. (e.g. [135]) alone are no guarantee of a good image result. Since most of the image processing operations influence each other, the order of the individual steps and the individual weights of the operations in the consecutive steps (Figure 59) are at least as important, if not more so.

11 Noise and noise reduction

Small pixels are less sensitive to light than larger ones. As a result, SPCs have a lower signal-to-noise ratio. In addition, there is the low full well capacity value for very small pixels, which reduces the dynamic range. Compared to DSLRs, these are all significant disadvantages, but with some SPCs they are partly compensated for by image processing and, recently, also by improved sensor technology, e.g. deep trench isolation, binning, dual conversion gain, and higher quantum efficiency [27].

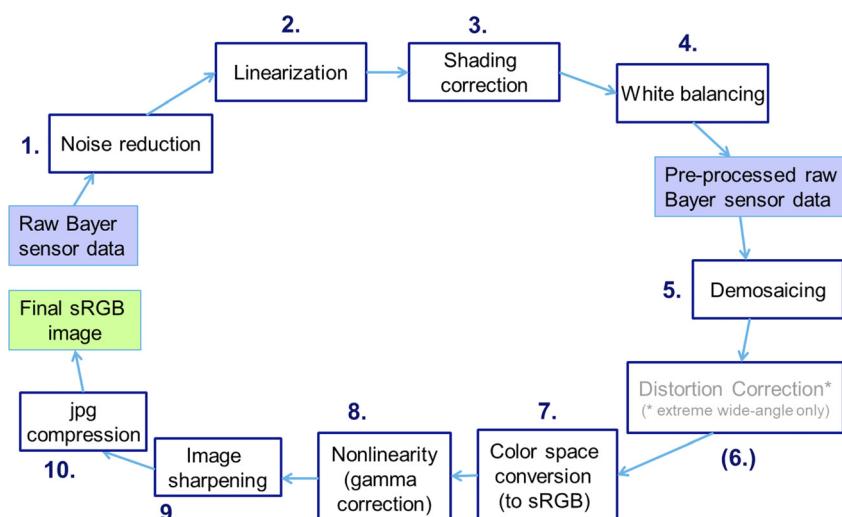


Figure 59: Example of (basic) image processing in SPC. The diagram is simplified as often several iterations are done, image data is exchanged with data bases, different images are composed (different times, different modules, ...), etc.

The noise of the electronics plays a minor role in today's contact image sensor (CIS) in reasonably good lighting conditions. Photon noise dominates, which affects all cameras. However, the user often does not notice increased noise directly, as the image processing in the camera results in software-based noise reduction. Such image corrections irrevocably smooth out small image details, i.e. the resolution becomes smaller and structures with less contrast, such as human skin (Figure 60), appear unnatural. This loss in image quality cannot be compensated for in subsequent image processing. However, the image recognition processes – which are getting better and better – improve the image quality thanks to different noise filters, which are used depending on the image content, and sometimes at individual points on an image.

12 Focusing

The depth of field achieved by mobile phone lenses is significantly greater than that of full-frame camera lenses. For a long time, the rule of thumb was that with SPCs for object distances of >1 m, i.e. for most everyday situations, no focusing is required. This has changed somewhat in recent years due to the larger focal lengths available in multi-cameras. A high-end standard wide-angle lens, $\text{FOV} = 75^\circ$, for an image diameter of $\phi_{\text{im}} = 12$ mm has a focal length of $f = 7.82$ mm. For an aperture ratio of $f/1.7$ the hyperfocal distance is 4.5 m. That is, if the lens is focused to an object distance of 4.5 m, the image is sharp for all object distances between 2.25 m to infinity.

To implement a focusing mechanism, the key parameters are the required focusing range and accuracy. As

mentioned earlier, a typical focusing range for a standard wide-angle camera is an object distance of between 0.1 m and infinity. The requirement as regards the focusing accuracy can be deduced from the number of "depth ranges" within the complete focusing range. The first depth range is the aforementioned hyperfocal depth range, i.e. 2.25 m to infinity. And the second depth range can be determined by setting 2.25 m as the far distance and calculating the corresponding focusing distance, and from that deriving the lower limit of the depth range. The hyperfocal distance can be approximated very accurately as (Section 4.4; Eq. 27), neglecting the very small second summand):

$$s_{F,\text{hyp}} = :h = \frac{f^2}{K\phi_{\text{thres}}} = \frac{f^2}{K\phi_{\text{im}}} \quad (49)$$

The relative size of the PSF spot size where the image can be considered "sharp" is about $r = 1/1500$. The near and far distances (s_{near} and s_{far}) corresponding to the focusing distance s_F are:

$$s_{\text{near}} = \frac{s_F}{1 + \frac{s_F}{h}} \quad (50)$$

$$s_{\text{far}} = \frac{s_F}{1 - \frac{s_F}{h}} \quad (51)$$

We introduce the index (j) to define the successive depth ranges, starting with $j = 1$ from the hyperfocal depth region from infinity to $h/2$. We can find the focus positions $s_{F,j}$ of the adjacent depth regions by setting

$$s_{\text{near},j} = s_{\text{far},j+1} \quad (52)$$

Specifically, that is $\frac{s_{F,j}}{1 + \frac{s_{F,j}}{h}} = \frac{s_{F,j+1}}{1 - \frac{s_{F,j+1}}{h}}$ and solving for $s_{F,j+1}$ yields:



Figure 60: Image of a human hand in weak ambient light, taken with a DSLR (left) and an SPC (right): As a result of the distinct, software-enabled noise reduction, the skin's structure is reproduced with low contrast and thus appears unnatural.

$$s_{F,j+1} = \frac{s_{F,j}}{1 + 2\frac{s_{F,j}}{h}} \quad (53)$$

Recursive calculation yields the following focusing distances for adjacent depth ranges:

$$\begin{aligned} s_{F,0} &= h, \quad s_{F,1} = \frac{1}{3}h, \quad s_{F,2} = \frac{1}{5}h, \quad s_{F,3} = \frac{1}{7}h, \dots \\ s_{F,j} &= \frac{1}{2j+1}h, \dots \end{aligned} \quad (54)$$

Table 5 shows the depth ranges within the complete focusing range of 0.1 m to infinity. The depth range quickly decreases towards small focusing distances: at $s_F = 0.9$ m the region (-15 cm, +23 cm) to the foreground and background, respectively, is still sharp; at minimum focusing distance $s_F = 0.1$ m the depth range is extremely small, only (-22 µm, +23 µm).

Expressing the consecutive depth ranges instead of the distance (s_F) in terms of the magnification

$$m_F = \frac{s'_F}{s_F} = \frac{f}{s_F + f} \quad (55)$$

we find with the expression we just obtained $s_{F,j} = \frac{1}{2j+1}h$: that for the nonmacro distance ($s_F \gg f$), that is $m_F \approx \frac{f}{s_F}$, the depth range number (j) is linearly related to the magnification

$$m_{F,j} = \frac{f}{\frac{1}{2j+1}h} = (2j+1)\frac{f}{h} \quad (56)$$

As can be seen in Table 5 the magnification grows nearly linearly with j .

From $s_{F,j} = \frac{1}{2j+1}h = s_{\text{MOD}}$ we get the total number of depth ranges:

$$J = \frac{h}{2s_{\text{MOD}}} \quad (57)$$

For the considered SPC lens data we get $J = \frac{h}{2s_{\text{MOD}}} = \frac{1}{2} \left(\frac{4500 \text{ mm}}{100 \text{ mm}} \right) \approx 23$.

A fairly reasonable accuracy for focus positioning is having about three focus steps per depth range (instead of one focus step, just to allow some margin for a safe specification in continuous operation). Consequently, within the complete focusing range about $3J \approx 70$ focus positions should be resolved. As we will see in the following section the total movement of the lens actuator to focus an object from infinity to MOD (0.1 m) is about 0.28 mm. So the required positioning accuracy is equal to $0.28 \text{ mm} / 70 = 4 \text{ } \mu\text{m}$. Specifically, with $\Delta m_F = m_{F,j+1} - m_{F,j} = \frac{f}{h}$ and $h = \frac{f^2}{Kr\varnothing_{\text{im}}}$ the required focusing accuracy is:

$$\Delta m_{F,\text{acc}} \approx \frac{f}{3h} = \frac{K\varnothing_{\text{im}}}{3f} = \frac{K}{2250\tan(\text{FOV}/2)} \quad (58)$$

For $K = 1.7$ and $\text{FOV} = 75^\circ$ the required accuracy is $\Delta m_{F,\text{acc}} \approx 0.001$.

Note that the required accuracy – as expressed with respect to magnification – depends on the *f*-number K and FOV only, and not on size-related quantities of the camera lens (e.g. on focal length f). This is a consequence of relating accuracy to magnification only: As we will see in Section 12.2 the required actual distances that optical groups need to move in order to bring an object into focus depend very much on the actual size of the lens.

Not all mobile phone camera modules have a focusing device: Often front cameras will have a fixed focus, which is positioned at a typical face distance toward the person operating the smartphone. There are also “depth map cameras” which, with a smaller image format and a larger *f*-number, provide greater depth of field (but poorer resolution) in order to create stereoscopic depth maps in conjunction with a main camera module. Until around 2005, many cameras with poor resolution in early generations of SPCs still offered no focusing, but this all changed with increasing resolution.

Table 5: Depth ranges of an SPC standard wide-angle lens ($f/1.7$, $\text{FOV}=75^\circ$, $\varnothing_{\text{im}} = 12 \text{ mm}$).

Depth range #	S_{far}	S_F	S_{near}	m_F	$1/m_F$
1	inf	4.500	2.250	0.002	576.5
2	2.250	1.500	1.125	0.005	192.8
3	1.125	0.900	0.750	0.009	116.1
4	0.750	0.643	0.563	0.012	83.2
5	0.563	0.500	0.450	0.015	64.9
6	0.450	0.409	0.375	0.019	53.3
7	0.375	0.346	0.321	0.022	45.3
8	0.321	0.300	0.281	0.025	39.4
9	0.281	0.265	0.250	0.029	34.9
10	0.250	0.237	0.225	0.032	31.3
11	0.225	0.214	0.205	0.035	28.4
12	0.205	0.196	0.188	0.038	26.0
13	0.188	0.180	0.173	0.042	24.0
14	0.173	0.167	0.161	0.045	22.3
15	0.161	0.155	0.150	0.048	20.8
16	0.150	0.145	0.141	0.051	19.6
17	0.141	0.136	0.132	0.054	18.4
18	0.132	0.129	0.125	0.057	17.4
19	0.125	0.122	0.118	0.060	16.6
20	0.118	0.115	0.113	0.063	15.8
21	0.113	0.110	0.107	0.067	15.0
22	0.107	0.105	0.102	0.070	14.4
23	0.102	0.100	0.098	0.073	13.8

“Focusing” comprises three aspects:

- An autofocus system, i.e. the automatic determination of the target focus position
- Modification of the optical imaging system so that an object focuses at a different distance
- Mechatronic implementation of the focus drive

There is no manual focusing option on a smartphone (except for the special creative modes on a few smartphones or some apps enabling this feature), unlike there is on most cameras for large image formats. This opens up the possibility of “creative photography,” in which image areas outside the main subject can be selectively focused. For the sake of simpler operation and because of the large depth of field of the SPC, manual focusing is not very useful. In addition, “creative photography” has also been added to the portrait mode function in recent years, and even enables users to change the focus position and depth of field of a scene, even after the picture has been taken.

12.1 Autofocus methods: Contrast and phase detection

We will first consider autofocus systems; see also ref. [136]. With classic cameras, a distinction is made between reflex cameras and rangefinder cameras. Autofocus systems for SLR cameras have been around since the 1980s. With SLRs, the focus is measured using a “phase contrast measurement.” The position of light rays is measured on a separate image sensor that is positioned conjugate to the camera image sensor, which it reaches via a mirror attached to the beam splitter. The position of this light beam in the objective pupil is precisely known and the deviation from the focus can be quantitatively determined using triangulation. The great advantage of phase contrast measurement is that you know exactly what the target shift is from a single measurement in order to get to the optimal focal point. The disadvantage is that you need a folding mirror between the lens and the image plane. The space required for this results in a bulkier, heavier camera, and in larger lenses. Wide-angle lenses in particular are becoming considerably larger and more complex because retrofocus lenses are needed in order to ensure the required distance between the last lens element and the image plane (Figure 25a). Due to their intrinsically asymmetrical structure (negative front group and positive rear group), these lenses are much more difficult to correct than lenses a small distance away from the image plane.

In the digital age, the availability of image data as a number matrix for structured objects makes it possible to directly evaluate the contrast based on the brightness variations in the read-out image. In order to ensure the optimal image position, the focus position must be changed, and the contrast evaluated using the focusing mechanism until the contrast is at its maximum. The large number of measurements required, and the associated time is the disadvantage of contrast autofocus. In addition, the focus initially runs beyond the optimal focus point and then swings back again before finally settling at the optimal focus point, which is perceived as unpleasant, especially when making films (“overshooting oscillations”). It is also possible for the focus movement to initially move in the wrong direction (“bad direction move”) (Figure 61). Another disadvantage is that contrast autofocus becomes error-prone or fails for low-contrast objects, especially in low-light conditions. The big advantage is the compact design of the camera due to the elimination of the folding mirror and the lenses due to the simpler correction, thanks to the close distance to the image plane (see Section 6.1). Typical shooting lags can range anywhere from 0.5 s under good conditions to about 2 s in low-contrast conditions.

Due to the very small depth required by thin smartphones, the DSLR principle is of course out of the question for SPC. From the very beginning, practically all AF mobile cameras have had a contrast autofocus. In recent years, the contrast autofocus in high-end SPCs has been supplemented by types of “phase detection auto focus” (PDAF) pixels, and in some cases even completely replaced. “Phase contrast pixels” as split or dual pixels first appeared in around 2008 with the rise of professional mirrorless system cameras like the Sony Nex and, later, the Sony a7. This was later extended to groups of 4 pixels, with PDAF aiming to detect the actual focus position. The principle can be seen as a “light version” of SLR phase contrast: Instead of scanning a very small area of the lens pupil as with an SLR camera, the amount of light that passes through a part, about half of the pupil, is evaluated here. In 2014 Sony applied masks in front of the photoelectric layer in the image sensor [137], mostly in 4 different orientations, e.g. to differentiate between “left, right, above, and below” (or rotated by 45° to this arrangement) which partially blocks the light. Choi et al. [138] present a geometrical model and analysis. With this PDAF principle with masked pixels, approx. 5–10% of the total pixels are used as focus pixels and are not available as normal sensor pixels but leave behind “blind spots” that have to be interpolated. More recent solutions use two separate neighbored pixels, a “photodiode twin”, below a common microlens [139]. This concept was developed by ON

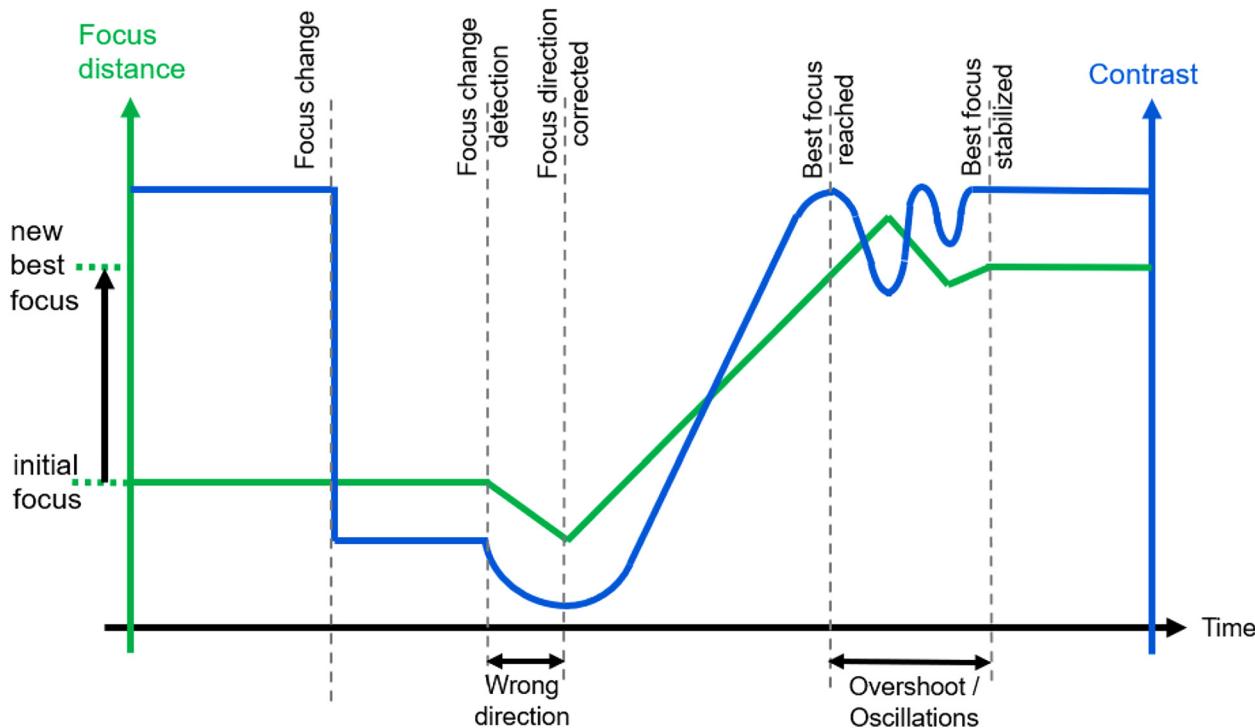


Figure 61: Typical curve of actual focus position (green) and contrast (blue) for contrast autofocus process when focus is changed to another best focus position.

Semiconductor (Aptina) and applied, e.g. to iPhones, starting in 2016 with Sony image sensors. Here all pixels can be used for imaging as well as PDAF. It does not suffer from blind spots. The entire image sensor can consist of these “Dual Pixel AF” such as in the Samsung Galaxy S7 and, in principle, distances can be measured over the entire image field. A detailed analysis is given by Kobayashi et al. [140]. Towards outer field regions, however, PDAF becomes increasingly problematic due to the oblique incidence of light since the angle of incidence is about 35° in the image corners. A cooptimization of pixel architecture, microlens design and data processing can be supported by ray-tracing- and wave-calculation-based image simulations [141].

In addition to the focus detection mechanisms mentioned on the actual image sensor, high-end SPCs usually also contain active distance measuring systems, ToF, or lidar, with continuously improving spatial resolution. By combining different distance measuring systems, accuracy can be improved, especially in situations in which a certain measuring system falls short.

We must also mention one reason why SPC focusing has become faster still: In most SPCs, when the trigger is released, instead of a single picture, a whole stack of pictures is taken in a somewhat reduced resolution. This is essentially a short film in HD or 4K resolution. As soon as the object is in focus, the SPC switches to a high resolution,

e.g. 12 MP, and the image is saved. The image stack is then removed from the buffer memory. Alternatively, the “animation” can also be saved as a “live image” or “motion still” at the expense of storage space.

12.2 Optical System changes focus position

There are different ways to implement optical focusing. With almost all camera lenses, either the entire lens (“total lens focusing”) or one or more individual optical groups (“floating element focusing”) is moved along the optical axis. With SPC lenses, however, with the exception of a few long periscope telephoto lenses, focusing is almost exclusively done through total lens focusing.

Another optical focusing concept involves changing the focal length of the lens either by deforming the lens – usually achieved with liquid lenses – or by using Alvarez–Lohmann manipulators. The latter consist of a pair of aspherical components that can be moved laterally toward one another [142, 143]. A MEMS-driven implementation is described by Zhou et al. [144]. A liquid lens can be realized by electrowetting, whereby two immiscible liquids with different refractive indices are placed in a cell (e.g. a cylindrical volume) and the curvature of the boundary between them is varied electrically [145]; alternatively, in a

liquid crystal lens, cells of birefringent liquid crystal material form a variable gradient index lens [146, 147]. Liquid lenses were first commercially integrated in a smartphone in the Xiaomi Mi Mix in 2021.

The required lens focusing distance for focusing between two object distances is determined using the imaging equation

$$\frac{1}{s} + \frac{1}{s'} = \frac{1}{f} \quad (59)$$

where s denotes the object distance and s' the image distance from the front or rear principal plane: For an infinite object distance the focal length (s') on the image is equal to the focal length (f), so $s'_{\infty} = f$. The minimum optical distance (MOD) is obtained from (59): $s'_{\text{MOD}} = \frac{s_{\text{MOD}}}{1+s_{\text{MOD}}/f}$. The difference between these image distances, as expressed by object distances, is

$$\Delta s'_{\text{MOD}} = s'_{\text{MOD}} - f = \frac{s_{\text{MOD}}f}{s_{\text{MOD}} + f} - f = -\frac{f^2}{s_{\text{MOD}} + f} \quad (60)$$

This difference is the distance that the lens and the image plane have to move in relation to one another so that the image remains sharp. In practice, both with SPC and with SLR cameras, the image sensor is fixed and the lens is moved forward in order to focus on an object that is closer to the lens (Figure 62). Therefore, the entire lens is moved. With many modern DSLR or system camera lenses, more complex focusing mechanisms are used, whereby one or more individual optical groups are moved within the lens [148].

It is noteworthy that, according to this equation, the distance required for focusing scales almost quadratically with the focal length, and not in an approximate linear fashion. This means that two lenses with the same equivalent focal length (or the same FOV) but differently sized image sensors have to move significantly different distances in order to focus from infinity to the same close distance. We consider this particular photographic situation for an SPC and a DSLR:

A typical close-range distance for an SPC is about 100 mm. Thus, for the standard wide-angle lens of an SPC

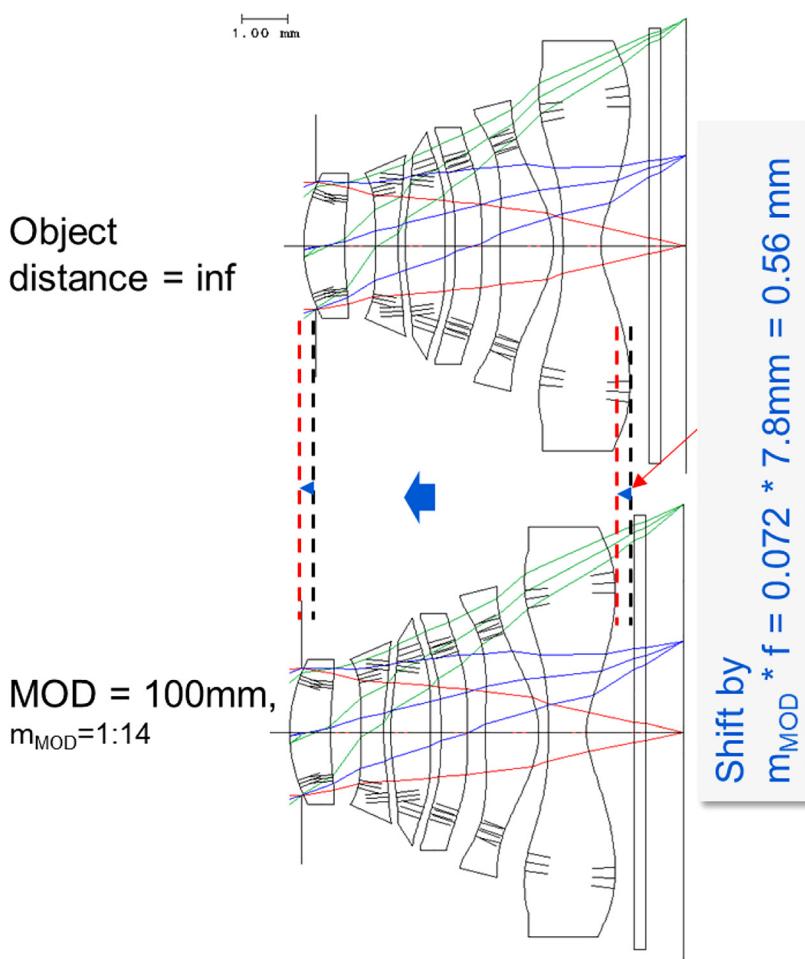


Figure 62: To focus on a close object, the entire lens is moved forward. The movement range is relatively small, only about 0.56 mm.

($f = 7.8 \text{ mm}$, equivalent on an image sensor with a diagonal of 12 mm with $f_{\text{eq}} = 28 \text{ mm}$ in full format)

$-\Delta s'_{\text{MOD}} = \frac{f^2}{s_{\text{MOD}} + f} = \frac{7.8^2}{100+7.8} \approx 0.56 \text{ mm}$, while with the full-frame camera $-\Delta s'_{\text{MOD}} = \frac{28^2}{80+28} \approx 7.2 \text{ mm}$. In this example, a path distance almost 13 times longer is required for a crop factor of only 3.6. (A DSLR with a magnification of approx. 1:3, the user is already in the “macro photography range”; meanwhile, for more moderate close-range distances, the path scales almost quadratically with the focal length).

Alternatively, the focus offset can also be specified as a function of magnification instead of the object distance. Now, with the help of the imaging Eq. (59), the magnification is

$$m = \frac{s'}{s} = \frac{f}{s+f} \quad (61)$$

or $s = f \left(\frac{1}{m} - 1 \right) = f \frac{1-m}{m}$ and inserted above $\Delta s'_{\text{MOD}} = \frac{-f^2}{f \frac{1-m}{m} + f} = \frac{-mf^2}{f(1-m)+mf} = -mf$.

That means that for a lens with a focal length (f) the focus movement distance directly scales with magnification:

$$\Delta s'_{\text{MOD}} = -mf \quad (62)$$

In other words: the focus scale is linear with respect to magnification.

From this equation it also follows that for a given magnification the required lens movement focusing distance directly scales with focal length. For many “normal” DSLR lens series, i.e. with the exception of “close focus” or macro lenses, the typical close distance magnification is about 1:10. Accordingly, for longer focal lengths you have to go back a long way, e.g. 10 times as much for a 280 mm telephoto lens as compared with a 28 mm wide angle. This applies accordingly to floating element focused lenses, but of course also specifically – depending on the focusing

mechanism used: For longer focal lengths, more space is required for the focus movements.

In the comparison above between the SPC and full-frame system camera with the same FOV, a close distance of 100 mm (measured from the entrance pupil and not, as is commercially common, but not useful for this direct comparison from the image plane) corresponds to a magnification of approx. 1:15 for smartphones, and for system cameras about 1:4. This means that a much larger distance range for the lens design must be optimized, which significantly increases the opto-mechanical complexity of the system. This necessitates internal focusing, preferably with two moving groups to achieve sufficient image performance, whereby the movement of one group is usually nonlinear.

12.3 Focusing mechanisms: Voice coil motors and other concepts

VCMs are almost exclusively used in SPCs as drives for the focus movement of the lens relative to the image sensor. Voice coil actuation exploits the interaction between a current-carrying coil winding and the field generated by a permanent magnet. The coil and the magnet, one in front of the other, are attached to two sides of the camera module’s housing (Figure 63). When a current is applied to the coil the interaction between the fixed and the magnetic fields electrically generated by the coil creates a Lorentz force that enables the camera body to move by a distance that is directly proportional to the current applied. Detailed explanations are given in refs. [149–151, 210, 211].

Alternative actuator types for focus movement are stepper motors, piezo motors [152], and MEMS. A comparison of different actuator concepts is offered by Murphy

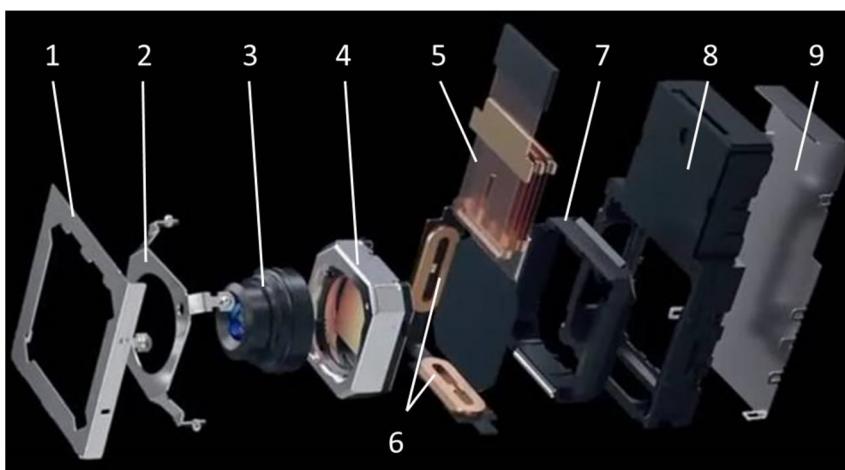


Figure 63: Camera components: 1. Housing top cover, 2. spring element for gimbal movement, 3. lens Barrel, 4. AF VCM and housing with image sensor, 5. multiply folded connector cable, 6. gimbal voice coils, 7. gimbal housing, 8. camera housing, and 9. housing cover. Courtesy of vivo. Gimbal components are related to a new generation of image stabilization; see Section 13.

et al. [153]; and novel VCM concepts by Hsieh and Liu [154]. An implementation of MEMS in a SPC and comparison with VCM is given in ref. [155]. Silicon-based MEMS actuation could be a viable alternative for VCM [156] in the future.

VCM have issues with predictability in its location due to directional hysteresis, typically approx. 10 µm – that is, in the order of image depth of field, temperature dependencies, and coil resistance variation. This necessitates open-loop control with multiple adjustments before the focus is correct.

EDoF by computational imaging was also considered as a promising option for smartphones [157]. According to the method proposed by Dowski and Cathey [158] a 3rd order profile aspherical (alternative phase functions are also possible) extends the DoF at the expense of contrast, which can, essentially, be recovered by deconvolution with the *a priori* data of the phase mask. With this method a factor of about 2 DoF extension in image space is achievable, which in turn translates to a significant extension of DoF in object distance equal to about the minimum optical distance: MOD = 30 cm to infinity. That is less than a standard autofocus with barrel shift (MOD approx. 10 cm). Autofocus achieved by moving the lens to a desired focus position is obsolete. However, deconvolution struggles with noise, which is omnipresent in low light: A significant noise level results in unrecoverable contrast. EDoF also tends to produce image artefacts. There were a few EDoF mobile phone cameras on the market like the Nokia E52 with a 3.2 MP camera in 2009. Nokia's marketing department named it the “full-focus camera.” Even though sophisticated computational imaging methods were used, consumers regarded the camera as rather a cheap alternative to a real autofocus camera.

13 Image stabilization

As the image resolution of mobile phone cameras improved, it became more and more obvious that their imaging performance was fundamentally limited by human hand-shaking during image exposure [159]. Particularly in low light but also indoors, photography quality suffered significantly, especially when compared with full-frame cameras. This was due to the intrinsically low etendue and a lower FWC, and therefore much longer required exposure times of the squared crop factor (approx. factor 50 for same *f*-stop). A compensation by increasing ISO sensitivity results in additional noise. Furthermore, in smartphone photography, camera handling is much less stable. Often the lever is large, as there is a long distance between the trigger pressure point

and the camera's position. Unlike in a DSLR the trigger direction is the same as the camera orientation. SLRs are designed for the user's hands, while a smartphone is primarily designed to fit in the pocket. The weight of the SLR also helps to reduce hand-shaking.

Electronic image stabilization (EIS) and optical image stabilization (OIS) are integrated in most high-end smartphones. Often OIS and EIS are used simultaneously. Some of the modules of one camera system may contain OIS while others do not, e.g. front cameras may not contain OIS. The first mobile phone with EIS was the LG Viewty in 2007 (a 5 MP high-end camera). For EIS the image frame is slightly cropped and, according to a photographer's hand-shaking as measured with the on-board gyroscope and accelerometer, compensated for by frame shift. EIS results in much better moving image performance. However, the movement during one-frame exposure is not compensated for, which results in blur. This blur can be reduced somewhat through deconvolution of the integral PSF for a frame exposure [160], which is nevertheless a computationally intensive task. Another disadvantage of EIS is the reduction of FOV and the fact that the image stabilization range is limited to the FOV portion, which is taken into account for EIS, such that any larger hand-shaking amplitudes result in residual errors.

All these disadvantages do not arise with OIS, at least in an idealized consideration that all hand-shaking is *in situ* ideally compensated for – that is, that all movement is instantaneously compensated for such that the camera does not move at all. In fact, there are some residual aberrations since the degrees of freedom used in the OIS system (e.g. relative movement of the image sensor) are not the same as the actual movement of the camera (e.g. an angular pitch or yaw displacement of the complete camera). The first OIS in smartphone imaging was released in 2012 by Nokia, in its N920.

13.1 Hand-shaking and image blur

A tremor is not a pathology; it is a common physiological phenomenon present in all humans. It's an involuntary oscillatory movement of body parts directly caused by muscles contracting and relaxing repetitively [161].

Everyone has a typical movement pattern, plus unsystematic variations. Simon [162] empirically analyzes the typical type and amount of hand-shaking in test subjects and their effects on the drop in image quality. Due to the omnipresent human tremor, both video performance and photography worsen significantly, especially in low light (Figure 64).



Figure 64: Picture taken with the same SPC without (right) and with OIS (left).

The key for EIS and OIS is the availability of on-board MEMS-based miniature sensors. Smartphone gyroscopes and accelerometers are lithographically manufactured as complex 3D structures on silicon wafers that are only about 0.5 mm in diameter. Since differential capacities are used in a comb-drive actuator design, the sensitivities of these sensors are linear – unlike standard capacitors [163] – and have a high-voltage sensitivity, leading to low power consumption [164].

In order to model the effect of tremors on image performance, 6 degrees of freedom of camera movement need to be considered: 3 translation directions denoted (x, y, z) , and 3 angular directions (pitch, yaw, roll) denoted $\theta_x, \theta_y, \theta_z$, respectively (Figure 65).

For a hand-shaking translation (dx, dy, dz) and a hand-shaking tilt $(d\theta_x, d\theta_y, d\theta_z)$ of the camera during exposure time, the image point (x', y') of an object point (x, y) at a distance (s) from the lens with a focal length (f) is displaced by following the distance (dx', dy') (see Figure 66):

$$dx' = f'd\theta_y + \frac{f'}{s}dx + yd\theta_z \quad (63)$$

$$dy' = f'd\theta_x + \frac{f'}{s}dy + xd\theta_z \quad (64)$$

The first components ($f'd\theta_x$ and $f'd\theta_y$) are due to tremor pitch and yaw, respectively. They do not depend on the distance of an object to the camera. The sensitivity increases towards a larger focal length (f'), that is towards tele lenses. Those components are measured by the gyroscope.

The second components ($\frac{f'}{s}dx$ and $\frac{f'}{s}dy$) are due to hand-shaking translation of the camera and are measured by the accelerometer. They are larger at close distances to the camera, e.g. an object distance of 0.1 m is 10× more sensitive than an object at 1 m.

The third components ($x d\theta_z$ and $y d\theta_z$) are due to the roll: this is effectively a rotation around the optical axis and results in azimuthal-oriented blur, which increases linearly in a radial direction from the image center. Since the hand-shaking components $(d\theta_x, d\theta_y, d\theta_z)$ are typically similar in amplitude, the roll-induced blur in the image corner is purely geometrically related to the yaw- and pitch-induced blur offsets: Due to the relation $\tan(\text{FOV}/2) = \phi_{\text{im}}/(2f')$ a standard wide-angle lens with an FOV of 75° has a factor of about 0.75 smaller blur in the image corner compared to yaw- and pitch-induced blur. Correspondingly, for a tele of FOV = 25° it is 0.25× smaller. However, note that although the sensitivities are smaller this component is highly

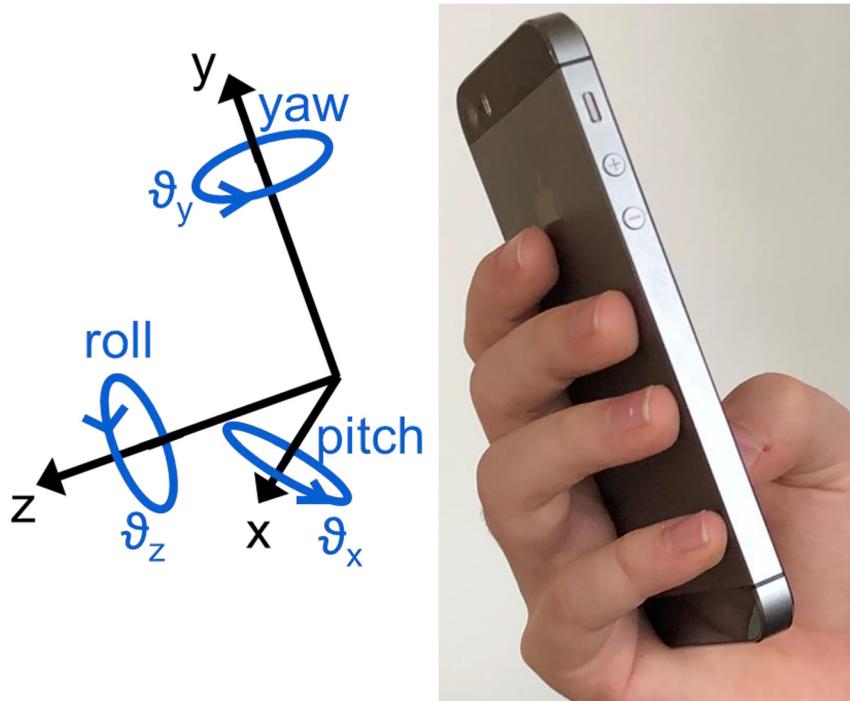


Figure 65: Definition of parameters:
Translation (x, y, z) and tilt components ($\vartheta_x, \vartheta_y, \vartheta_z$).

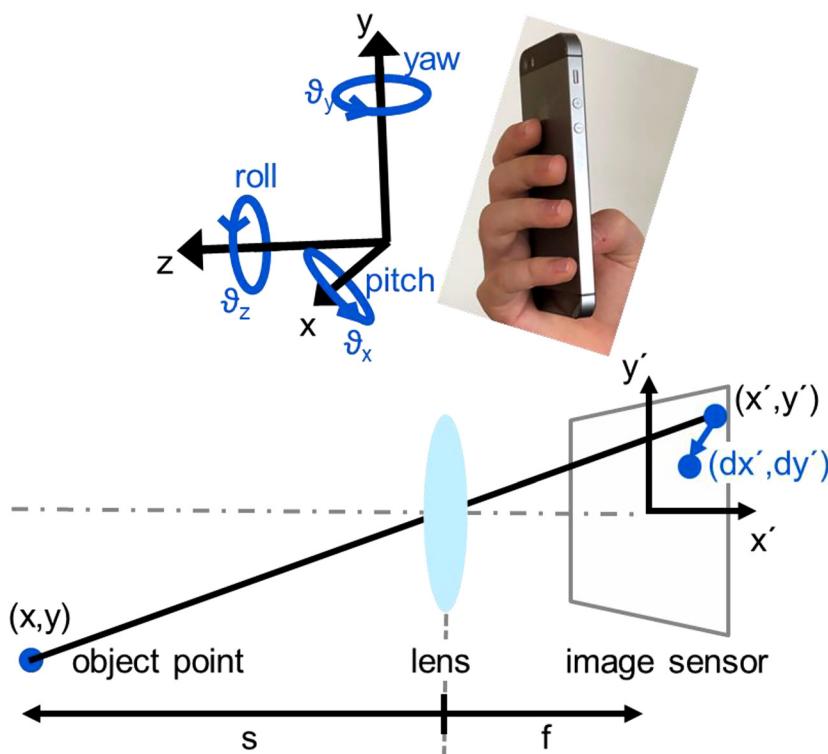


Figure 66: Definition of distance of camera to object point (s) and image distance (approximately equal to focal length (f) for nonmacro distances) to formulate the equations for tremor-induced image shift (dx', dy').

relevant, since most OIS systems are not able to compensate for this component.

Table 6 shows typical hand-shaking according to Simon's study for an exposure time of $T = 0.4$ s (considerably simplifying the statistical analysis performed in ref. [162]

and their effect on the image blur component $\Delta x'$ in micrometers, assuming that $f' = 4$ mm and $\text{FOV} = 75^\circ$. According to the detailed statistical data, also collected by other authors [165, 166], the tremor-induced tilts of any person in any “usual” (non-action) situation is clearly

Table 6: Typical hand-shaking-induced errors and their effect on image blur for mobile phone photography with an exposure time of 0.4 s (Data from ref. [162]).

Mean handshake		$\Delta x' [\mu\text{m}]$, $s = 1 \text{ m}$	$\Delta x' [\mu\text{m}]$, $s = 0.1 \text{ m}$
$\Delta\vartheta_x$ 0.32°	Pitch, yaw	22.3	22.3
Δx 1.7 mm	Decenter	6.8	68.0
$\Delta\vartheta_z$ 0.32°	Roll	17.4	17.4

below 1°. Accordingly, the range of most OIS is in the range of 0.5–1°.

The y components ($\Delta\vartheta_y$, Δy) are typically similar in magnitude to the x components (statistically, not individually). For “normal” object distances like 1 m the angular hand-shaking components (pitch, yaw, roll) clearly dominate, whereas for close-distance photography (0.1 m) decenter components dominate (Table 6). At about 0.3 m the contributions are comparable.

Obviously, according to these data – and without any image stabilization – the hand-shaking-induced blur by far exceeds the nominal camera lens performance and significantly limits image performance. The effect is smaller at smaller exposure times, as is typical in bright light environments: Simon has also given an empirical formula for the dependence of integrated hand-shaking tilt rms

$$\Delta\vartheta_{x,\text{rms}}(T) = \int_0^T \frac{d(\vartheta_x - \langle\vartheta_x\rangle)}{dt} dt$$

on exposure time in seconds:

$$\Delta\vartheta_{x,\text{rms}}(T) \approx 0.37 \cdot T^{0.62} \quad (65)$$

According to this equation’s exponent, the typical hand-shaking blur is not a purely random walk (exponent = 0.5), it also contains some systematic components (exponent = 1). Figure 67 shows the corresponding effect of exposure time on image blur with $d\vartheta = fd\vartheta_y$ and $f = 4 \text{ mm}$.

For interior images, exposure times with SPCs are typically in the order of around 1/10 s, night scenes around 1 s (often ISO sensitivity is automatically increased significantly to prevent this): These situations clearly benefit from image stabilization, as image blur is in the order of 10 pixels (Figure 67). On the other hand, daylight photography with exposure times of 1/100 s or less is far less critical.

13.2 Optical image stabilization implementations

There are several different compensation mechanisms (Figure 68): In a first-order approximation a tilt and shift of the barrel is equivalent and therefore also the correction of a hand-shaking tilt or decenter. For a $f' = 4 \text{ mm}$ a 1° tilt corresponds to about a 70 μm image shift. In most SPCs, OIS is implemented by a compensating movement of the entire lens barrel in x, y with a voice coil motor [167]. In a

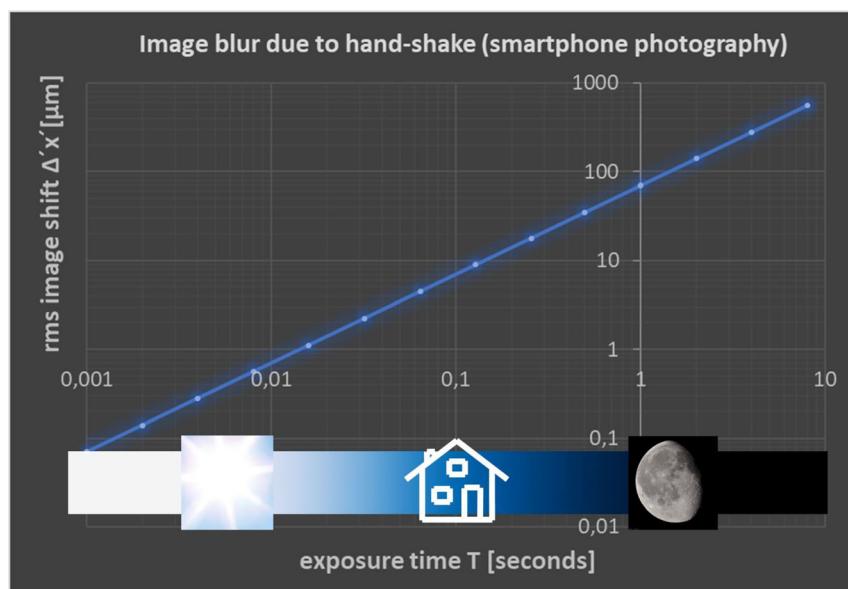


Figure 67: Typical image blur due to angular hand-shaking versus exposure time.

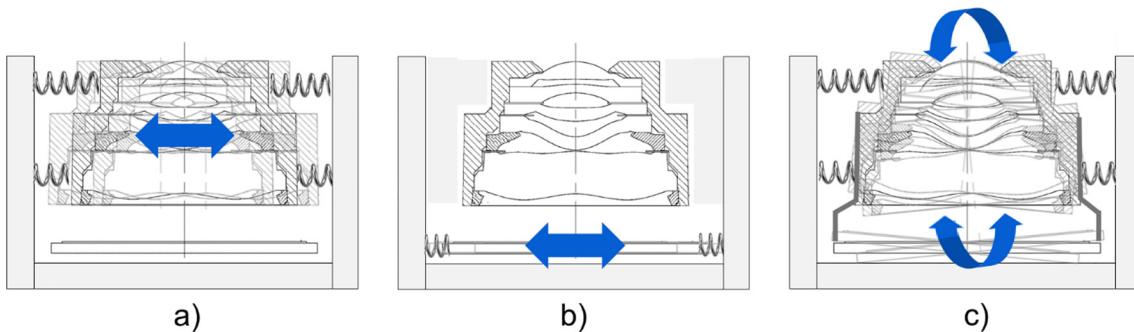


Figure 68: Different OIS systems: (a) Barrel decenter (most common), (b) image sensor decenter (in the SLR world this is referred to as body image stabilization [BIS]), and (c) “gimbal.”

few systems the barrel is tilted. Recently, alternative movements have been introduced. In the Apple iPhone 12 Pro released in 2020, the image sensor is actively moved. Also in 2020, vivo launched a “gimbal system” that tilts the complete barrel together with the fixed-to-barrel image sensor (vivo X51). In combination with the wide-angle camera with a large 3° amplitude, it also supports “action cam video shooting” situations.

Several mechanical design parameters of OIS, e.g. the most common one with a VCM-driven barrel decenter, are critical for OIS performance as well as eliminating unwanted side-effects: The stiffness of the springs determines the correctable amplitude frequency distribution, e.g. a very stiff spring enables high-frequency corrections but exhibits smaller amplitudes at low frequencies. The mechanical adjustment of yoke to moving barrel also needs to be balanced: A tight adjustment increases friction and hysteresis, while a loose one – that is, a large clearance – may lead to parasitic tilts. A more detailed description of recent optical image stabilization technology can be found in a white paper from ST Microelectronics [168].

objects and brightness variations in the interior indistinguishable. If the exposure is increased to show the interior, then the outer area is completely overexposed, that is, it is homogeneously white and therefore no longer resolved in the dynamic range shown (Figure 69).

The reason behind this deficiency is the limited dynamic range of image sensors in relation to the range of irradiances in a scene.

As long as there are no very bright light sources (like the sun) or spotlights in the field of view or its immediate vicinity, the dynamic range of typical outdoor scenes is in a range of about 9–12 EV and rarely lies above 14 EV (Figure 70). However, with strong light sources, the dynamic range may lie significantly above 20 EVs [170].

Dynamic range is usually referred to as low ISO sensitivity (i.e. ISO 100). It is important to mention that DR decreases as the ISO value increases, typically with 0.6–1.1 EV per ISO step. For example, the dynamic range with a high ISO sensitivity of 6400 mostly ranges between 6 and 9 EV, with smartphones achieving even lower values. The DxOMark website provides a comprehensive database of many cameras with measured data of the dynamic range. Figure 71 shows a development in the dynamic range of different cameras over the past years, according to measurements taken by DxOMark showing that the dynamic range of image sensors is constantly increasing. For today’s high-end, full-frame digital consumer cameras, the dynamic range, i.e. the range of distinguishable luminance, is about 14–15 EV, while SPCs range from about 10–12 EV.

Moreover, quick sequences of exposure times are realized; they can be subsequently processed in HDR image processing programs in ever better quality, and also converted to HDR images on the camera itself. Meanwhile, many digital cameras, including mobile phone cameras, calculate HDR images from an automatic sequence of different exposure times. The algorithms on how to obtain an extended dynamic range image from a sequence of

14 Dynamic range

14.1 HDR imaging

In Section 10, we looked at the display of brightness, e.g. as an 8-bit image, that is as $2^8 = 256$ different brightness levels. Now, different irradiances of the real physical object space can only be recognized as distinguishable if the camera is able to record all brightness values of a scene simultaneously, as distinguishable values. This is usually not the case. If you take a picture outside on a sunny day from within a room and set the exposure to the outside area, then the inside area will be completely dark and make all



Figure 69: Indoor and outdoor areas are not simultaneously captured within the limited dynamic range of the camera. (Source: Reference [169]).

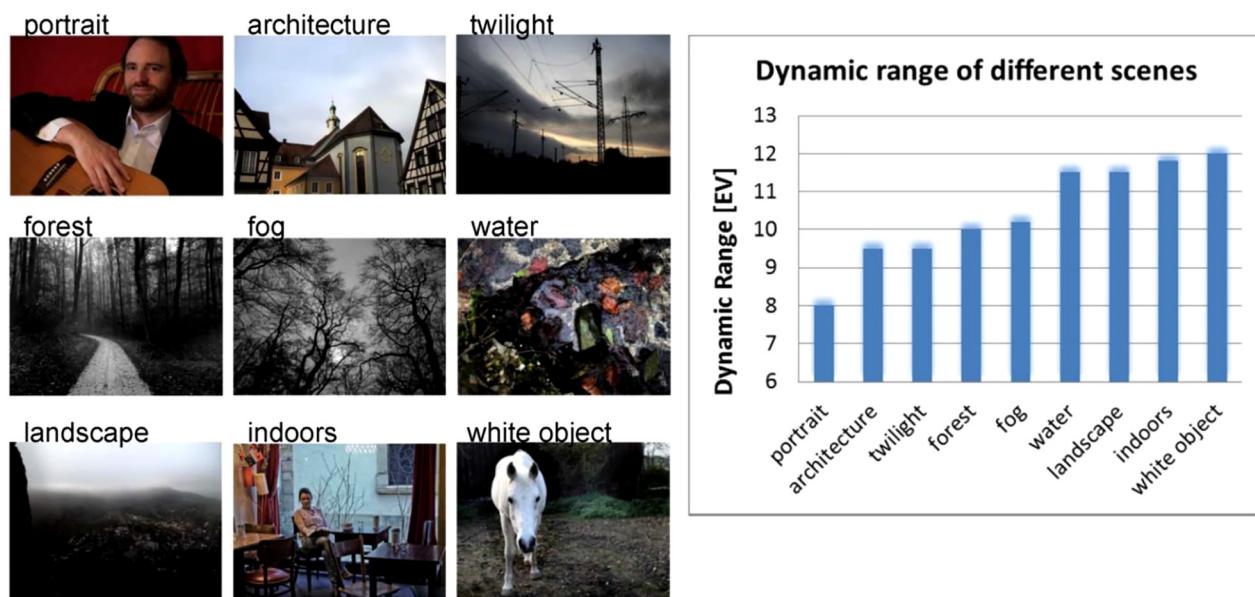


Figure 70: Typical dynamic ranges of real-life scenes [170].

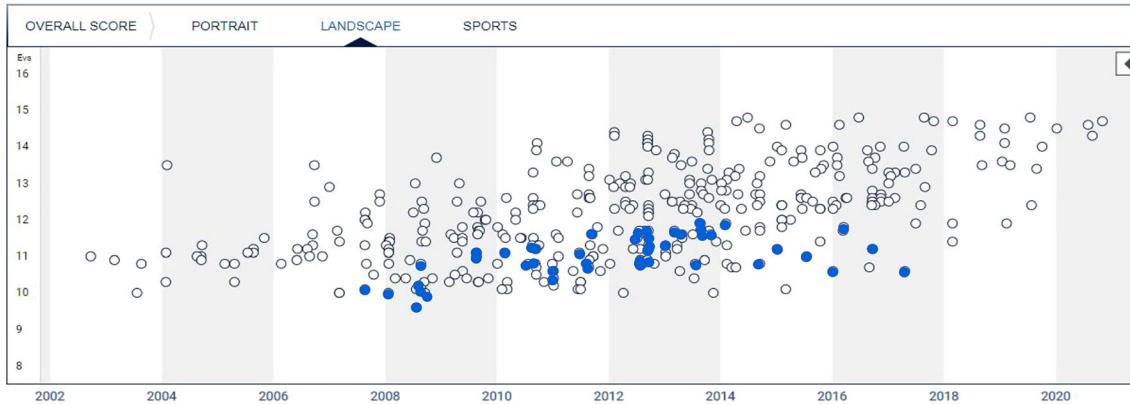


Figure 71: Dynamic range of digital cameras. SPCs are marked in blue in this graph. Source: DXOMark.

images taken with different exposures have been developed in the 1990s [171, 172]. This replaced methods of manual local tone mapping, as practiced by the famous photographer Ansel Adams for his illustrated book “The Print” by locally exposing negatives, with a holistic process [173]. The tonal values of these HDR recordings, which

extend over a very large value range (e.g. 16 bits, i.e. $2^{16} = 65536$ tonal values, or more) are reduced to a much smaller brightness range (often 8 bits, i.e. 256 tones). In the HDR process, several images with different exposure values are recorded to represent a situation with a high contrast between light and dark areas, each of which then



Figure 72: From a series of 4 pictures taken at different exposures (left) and a calculated HDR image (right).

contain information from overexposed or underexposed areas (Figure 72). This information is artificially combined to form a new image in which the very high tonal range is compressed to a lower one [174–176].

When it comes to smartphone photography, it is helpful that the brightness scale can be quickly adjusted by tapping on the main subject on the display. The dynamic range can be increased using the HDR function, which can usually be selected separately. The scene is then recorded several times with different exposure times and combined for the extended dynamic range according to the brightness range shown.

Due to the much longer exposure time required, this technique is unsuitable for fast-moving subjects and more prone to loss of resolution due to hand shaking. The problem with dynamic situations has been mitigated with multicell sensors because they enable exposure times of different lengths to be parallelized with different pixel clusters (Figure 73).

14.2 Lens flare and ghosts

A common HDR situation is when the sun or another very bright light source is in or just outside the frame. In this case the dynamic range often exceeds 25 or 30 EV (factor of 1,000,000 or much higher). Then, in addition to the image sensor, the lens quality is crucial: residual light reflection on lens surfaces from a bright light source may superpose or even cover up parts of the image. This may even happen for a multilayer coated surface (<0.2% reflectivity) if the reflection is considered in combination with the image sensor (approx. 5% reflectivity): The double reflex is a factor of $0.002 \times 0.05 = 0.00001 = 1/10,000$ weaker than the bright light source, which corresponds to about 2^{-14} or -14 EV. Since the maximum irradiance within a normal scene is also about 15 EV smaller than the bright light source, the reflected ghost of the light source may appear very bright on the picture. Whether or not the ghost is apparent depends very much on if it is almost in focus or out of focus (Figure 74).

normal mode



HDR mode

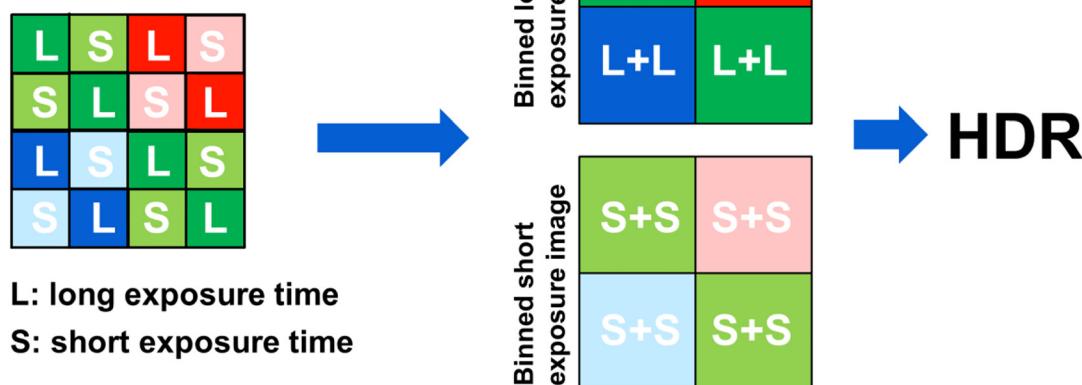


Figure 73: Multi-cell sensor in normal binning mode and HDR mode.

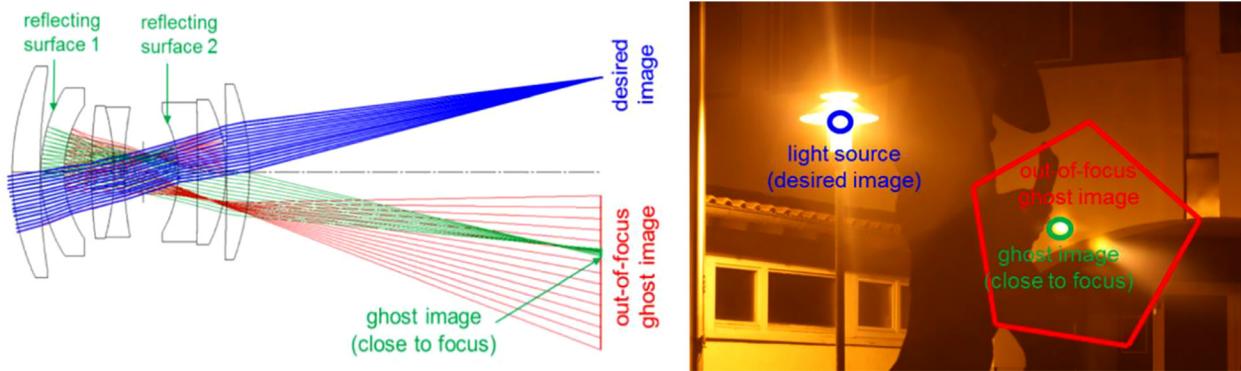


Figure 74: Ghosts' paths of out-of-focus and almost in-focus ghost images. The in-focus section appears very bright on the image (Source: Reference [169]).

The number of lens surface-surface reflections inside the lens with n surfaces is

$$n + (n - 1) + \dots + 2 + 1 = \frac{1}{2}n(n + 1). \quad (66)$$

The exact position, shape, color, and intensity of the individual ghost images produced at the various optical lens surfaces are each dependent on the optical design of the lens. With large apertures, local light concentrations usually have aberrated shapes: Caustics can occur and may be crescent-shaped (Figure 75, left). In smartphone imaging, sometimes “egg-shaped” in color, often purple, structures around very bright light sources are especially characteristic (Figure 75, right). These are due to reflections that occur in between the micro lens array above the image sensor combined with the IR filter and possibly also with other lens surfaces. The purple color arises from higher coating reflectivity in blue and red compared to green.

The reflections within the lens and in interactions with the image sensor not only affect images with local highlights, but also every scene with a large dynamic range: The false light from the bright areas in the image overlays the darker areas and reduces the macro contrast there.

In addition to using good hardware, i.e. AR coatings or straylight-blocking rings, the straylight performance of a lens can be optimized as early on as the optical design phase (e.g. [169, 177, 178]) to avoid the aforementioned in-focus ghosts. With ghost ray trace analysis these ghosts can be identified and potentially eliminated or reduced later on in the design process through modifications on the optical surface shapes (see Figure 76). With powerful computer-aided design (CAD) software like ref. [207] or ref. [205]; straylight analysis can be extended to include the detailed opto-mechanical layout of the system. For a comprehensive analysis of the straylight effects of smartphone lenses, see reference [179].

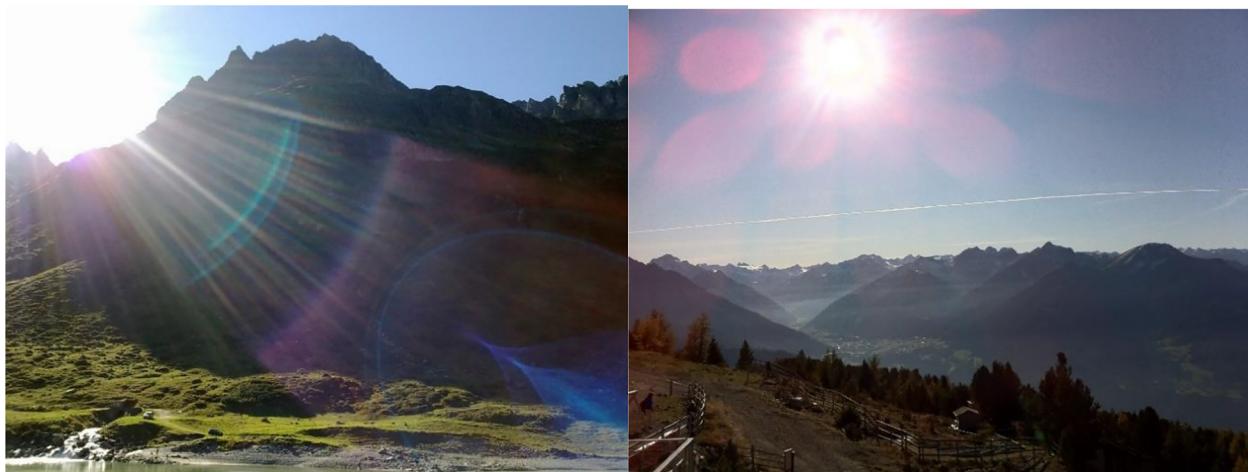


Figure 75: Straylight in smartphone photography. Left: Radial structures produced by roughness on the lens stop, and ghosts created by internal reflections. Right: Ghosts created by internal reflections from the image sensor micro lens array combined with IR filter. Courtesy of Richard Gerlich.

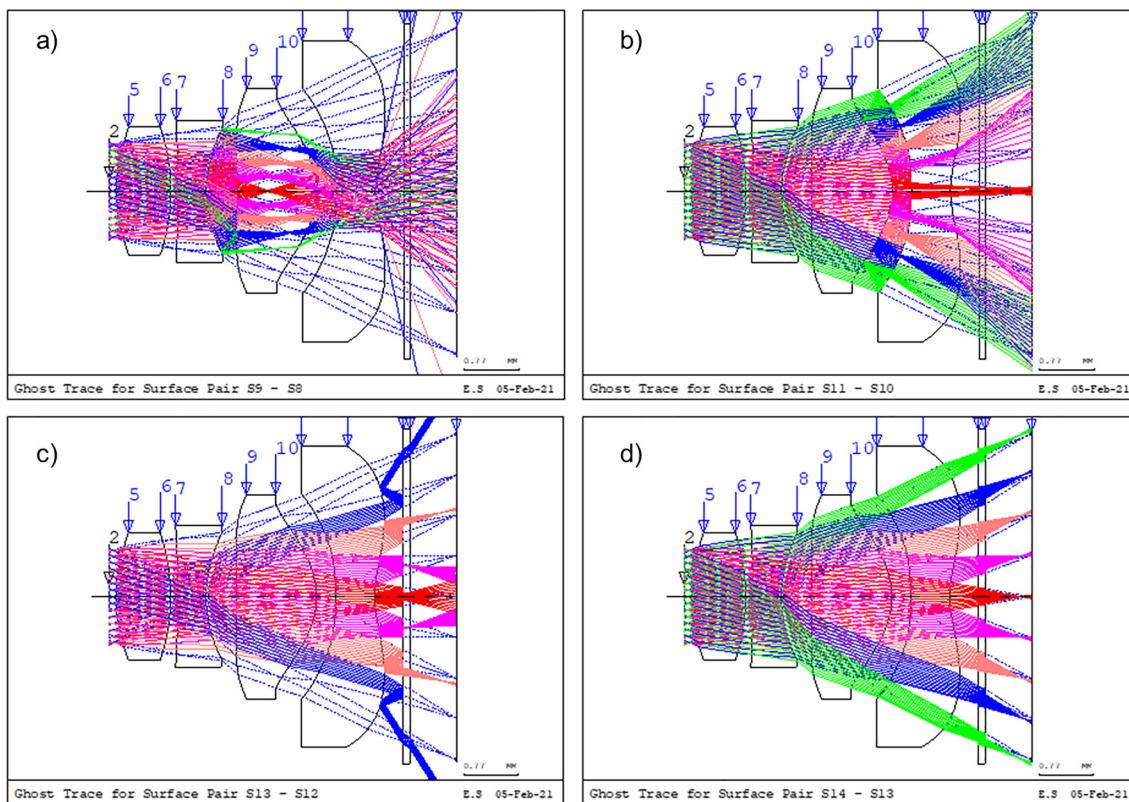


Figure 76: Reflections on different pairs of optical surfaces. Most of the reflex paths are rather uncritical because they are severely defocused or do not reach the image plane at all, such as 9a). Critical are those reflections that arrive almost in focus on the image plane from a larger aperture area (b), (c). (d) Shows that the IR filter must be well coated, since otherwise – with very bright light sources – ghost images adjacent to the desired lens PSF will appear.

For SPC lenses' ghost ray trace analysis the fields must be sufficiently well sampled, since bright ghosts may appear inside a very small portion of the full FOV. This is due to the distinct local differences in reflection direction at the wiggly lens surfaces. In particular, surfaces exceeding the total reflection angle can cause problems if they accidentally reflect the light near to the focus. The local variation of straylight can easily be observed when rotating the lens relative to a very bright light source (Figure 77). This is essentially also the situation in which SPCs are tested in the lab under standardized conditions.

15 Portrait mode

The dual or multicameras and 3D depth sensors installed in smartphones in recent years make it possible to determine depth information in high resolution, i.e. depth maps. Software development, e.g. through machine learning, coupled with access to image databases, has made great strides in recent years and is used in subject recognition,

such as in portraits of people, but also for portraits of pets, etc. This is done to generate purely on a context-by-context basis, and is based in particular on the shape of edges (face contours), to create simplified depth maps. Or it is combined with 3D acquisition to improve depth maps.

From the normal photo, which is naturally sharp over a large depth range on smartphones, and the depth map created at the same time (Figure 78), a “shallow depth of field look” is generated by computationally blurring out-of-focus regions according to the depth map data. Mathematically, this corresponds to a (local) convolution with a depth-dependent point spread function. The depth dependent PSF is stored in the memory either as a function or as a comprehensive data set. In principle, these calculated images could resemble the real images of a DSLR if these 3D PSF data – which is dependent on both the depth and the selected focus – match the physical PSF data of the DSLR lens. The depth map should accurately reproduce the scene in all its details: Incorrectly recognized depths can lead to very unpleasant artifacts. In addition to imitating the DSLR camera look, one can do something in this way that one

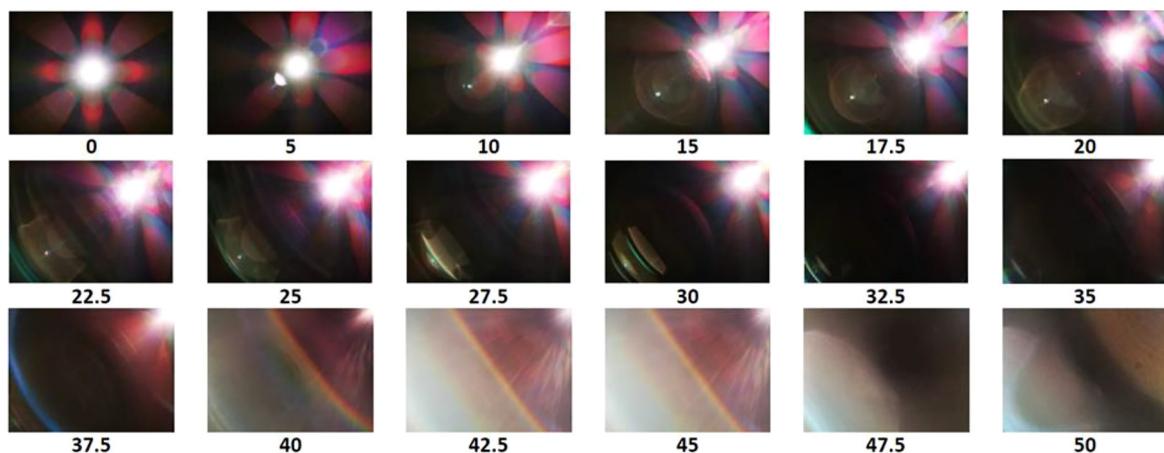
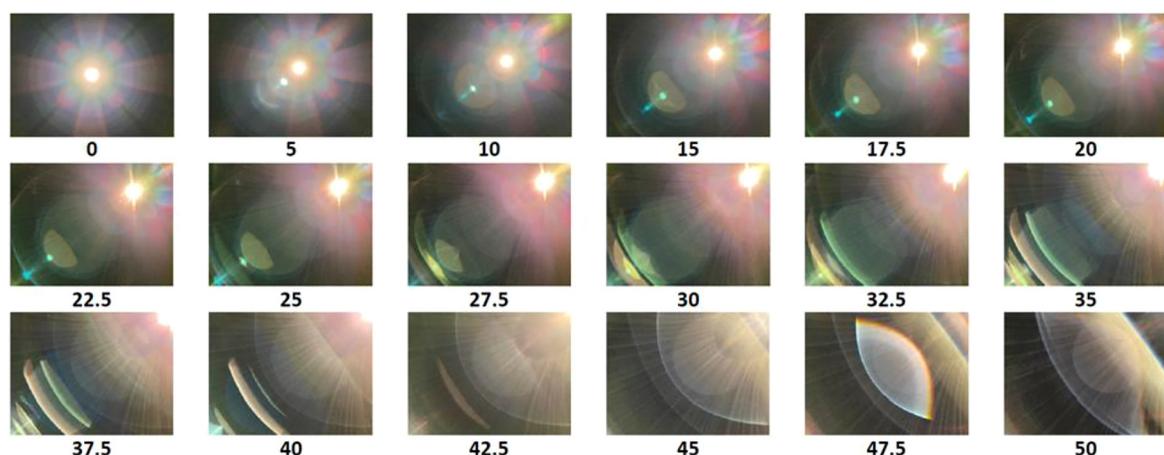
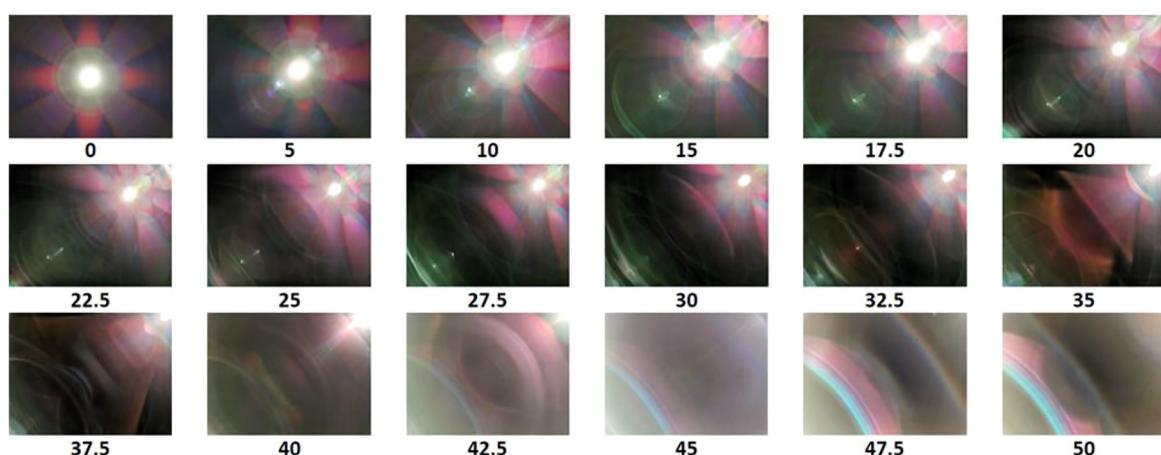
SPC Lens A**SPC Lens B****SPC Lens C**

Figure 77: Straylight and ghost lab analysis performed by rotating a camera relative to a bright light source (relative angle indicated in the graph) for three different SPCs. All SPCs feature standard wide-angle lenses with an FOV of about 80°, which can be seen as the light source in the corner of the image at 40° rotation. The ghost distribution depends on the specific optical design: overall quality is very different, and the strength and color of the ghosts depends very much on the coating's spectral and angular-dependent reflectivity.



Figure 78: Photo and associated depth map of an SPC.



Figure 79: Synthesis of “DSLR-like” shallow depth of field and change in focusing distance (not feasible in DSLRs “after the fact”).

cannot of course do with a DSLR: change the sharpness level with a single image after the fact (Figure 79).

The (artistic) “isolation” of a person against a blurred background is very popular among ambitious photographers using large image sensors. With a DSLR, portraits are usually taken with lenses of medium or long focal lengths with a high aperture. The background of the image taken with an SPC is hardly blurred, as the depth of field of an SPC is many times greater than that of the DSLR due to the much smaller image format.

We derived the relative size of the circle of confusion, the diameter of the PSF, in the out-of-focus background earlier, in Section 4.4; Eq. (22) as follows:

$$\varnothing_{\text{rel.spot},\infty} = \frac{\varnothing_{\text{EP}}}{\varnothing_{\text{ob,Portrait}}} = \frac{\varnothing_{\text{EP}}}{700 \text{ mm}}. \quad (67)$$

A portrait photo with SPC multicamera realistic lens data at an equivalent focal length of 85 mm requires $f/3$ with $f = 8.8 \text{ mm}$ ($85 \text{ mm}/\text{CF}$, where the crop factor is about $43.3 \text{ mm}/4.5 \text{ mm} = 9.6$), resulting in an entrance pupil diameter $\varnothing_{\text{EP}} = 8.8 \text{ mm}/3 \approx 3 \text{ mm}$. This gives a relative background blur spot diameter of only $\varnothing_{\text{rel.spot},\infty} \approx 0.42\%$.

On the other hand a classic DSLR portrait lens (full-frame $1.4/85 \text{ mm}$) has an entrance pupil diameter of $85 \text{ mm}/1.4 \approx 60 \text{ mm}$ and a relative background blur spot diameter of

$\phi_{\text{rel. spot},\infty} \approx 8.7\%$, which is more than 20 times as large as for the SPC.

15.1 3D depth acquisition technology

Since SPCs have been equipped with several cameras, the distance can be determined stereoscopically on the common image portion of the cameras. With current multi-cameras, this is often the given FOV of the standard telephoto camera ($f_{\text{eq}} \approx 55\text{--}75\text{ mm}$) and the corresponding crop of the wide-angle camera.

With dual cameras, distances to objects can of course only be determined for the directional component of the neighboring cameras and not in the direction perpendicular to it. By using more cameras, like the 5 lenses of the Nokia N9 Pureview from 2019 (Figure 80), the disparity can be captured along different directions, which improves the quality of the depth map.

When calculating images from several cameras, the amount of occluded areas (Figure 81) and areas with unusable image information – e.g. due to reflections – is reduced. In addition, the correlation analysis to determine the disparity becomes more robust as the amount of overlapping image data increases [181–183]. Classic multiview stereo processing is computationally expensive and usually done on a desktop PC. It is very challenging to perform this on a mobile platform such as a smartphone. It involves image feature extraction, feature matching, camera calibration and pose estimation, dense depth estimation, surface reconstruction, and texturing. Another multi-aperture depth-acquisition method is lightfield imaging, as realized in the Lytro Illum Camera [184] or by Raytrix [185]. Here, a conventional lens is combined with a micro lens array, which is positioned closely in front of an image

sensor such that the raw image consists of many small, partly overlapping, fields of view of a scene. From correlation analysis between adjacent fields of view, disparity – and therefore depth – can be determined. The setup can be chosen such that either only a few or many capture the same object details within the micro image frames. The more images are involved the better the quality of the depth estimation, which however is at the expense of spatial resolution. A consequence of combining micro lenses with a much smaller focal length compared to the main lens results in small achievable disparities (typically <1%) and therefore inferior depth resolution compared to multi-cameras separated by base lengths in the order of the image sensor size or higher. Another feature of such lightfield cameras is an extended depth of field – especially for the Raytrix system that uses micro lenses with different focal lengths [185].

From the similar triangles in Figure 82 we derive the following relations between the lenses' focal lengths (we assume both lenses have both same f and FOV), the object points distances s_1 and s_2 and the base distance (b) between the cameras:

$$\frac{s_1}{b} = \frac{f}{y_1} \quad (68)$$

$$\frac{s_2}{b} = \frac{f}{y_2} \quad (69)$$

The difference in image position on the second camera's image plane is the disparity $d = y_2 - y_1$:

$$d = fb \left(\frac{1}{s_2} - \frac{1}{s_1} \right) \quad (70)$$

Now with $f = \frac{y_{\max}}{\tan(\text{FOV}/2)}$ we express the disparity relative to the image frame as d/y_{\max} :



Figure 80: Miniature multi-camera systems. Left: Nokia N9 with 5 camera lenses with equal focal lengths; right: PiCam from Pelican Imaging [180].



Figure 81: Raw images from the 70 mm cameras (marked in red) of the Light L16 during a close-up indoor shot. The disparities of the individual images are large here, greater than disparities typical of SPC multicameras, as these are usually placed close together. As a result, there are large occluded areas in the background, i.e. not visible by both cameras at the same time.

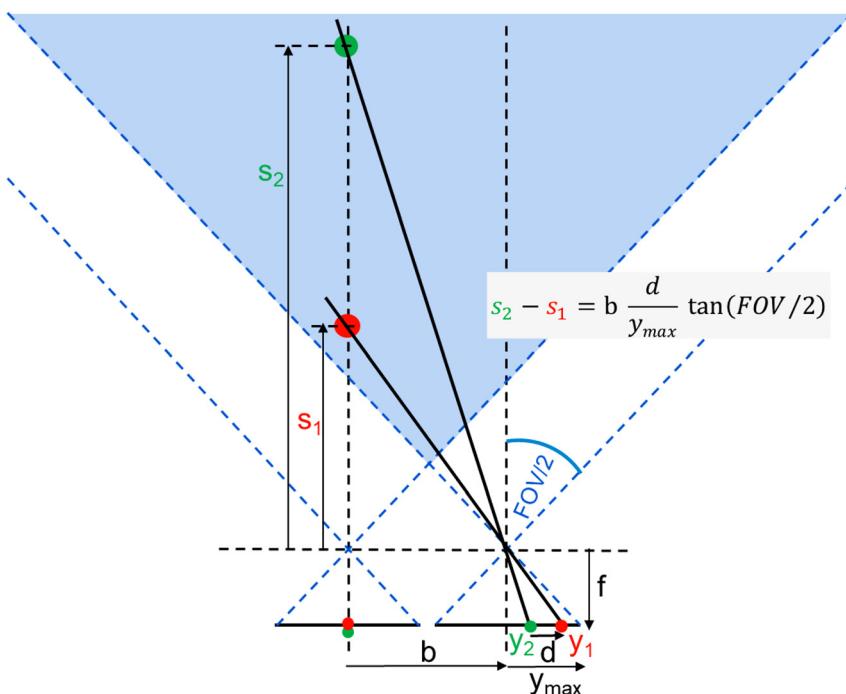


Figure 82: The disparity is the change in the relative distance on the image sensor of two objects at different distances, as seen from two different perspectives.

$$\frac{d}{y_{\max}} = \frac{b}{\tan(\text{FOV}/2)} \left(\frac{1}{s_2} - \frac{1}{s_1} \right) \quad (71)$$

A very far (“infinite”) distant object point appears at the same position on both image sensors. In this case $1/s_1 = 0$ and the distance $s = s_2$ is determined by:

$$s = \frac{b}{\frac{d}{y_{\max}} \tan(\text{FOV}/2)} \quad (72)$$

Since b , d , y_{\max} and FOV are all known from the lens and camera data and the cameras are calibrated (positioning, axis-orientation), inside the common FOV distances can be determined by the disparity evaluation in absolute units.

A typical baseline distance (b) between SPC modules is about 1–2 cm. For a baseline distance of $b = 10$ mm and an FOV 44° normal tele lens a relative disparity of $d/y_{\max} = 0.354$ (that is, 3.54% of the distance between image center and corner) corresponds to an object distance of 700 mm, which is a common portrait distance for this FOV. Correspondingly, a 2.5-m object distance gives a relative disparity of $d/h_{\text{im}} = 0.01$, that is 1% of the relative image height which is still in the order of 20 pixels. A roughly 10 m distance disparity is in the order of 2 pixels, so depth becomes indistinguishable, which is of no practical concern since the through depth PSF of a lens is practically invariant – this also applies when photographing at usual portrait distances (see next subchapter). With multi-camera systems, depth resolution decreases as distance from the lens increases. That is, the distance between the “depth planes” increases (Figure 83). In general, depth resolution depends on the depth of field of the lenses and on their actual focus position as well (Figure 84).

All stereoscopic distance measurement methods fundamentally require structured objects in order to be

able to compare (correlate) them with one another in different image fields. A depth estimation is impossible for unstructured objects, e.g. a clear blue sky or periodic objects. The depth of these image areas can only be assigned using assumptions or trained by machine learning: For example, “evenly light-blue areas in the upper part of the image are probably sky and are assumed to be somewhere in the distance. The quality of the depth map, determined stereoscopically or multiscopically, is generally highly dependent on the contrast and therefore also on the light intensity and spectral distribution of the illumination in the scene (Figure 85).

Incorrect depth data lead to unnatural blurring on the image and at the edges between the foreground and background. The quality of the portrait calculated in the smartphone ultimately depends on the quality of the depth map. So, if the depth of the image areas is incorrectly estimated, the blurred image appears unnatural in this area. Figure 86 shows a portrait: At first glance, the person looks nicely “isolated” from the background. If one takes a closer look, one can see that areas at the edge of the head are incorrectly assigned to the background. Even very fine object structures in the foreground, such as fine hair, often cannot be accurately depicted. For normal viewing on a smartphone screen, these errors are mostly inconspicuous and completely sufficient for the purpose.

Sound face recognition requires the acquisition of three-dimensional depth data. For this purpose, ToF sensors are mostly used in SPC, in which distances are measured simultaneously at different positions, measuring the propagation time of IR light from the light source and back to the receiver. ToF does not encounter the same problems as stereoscopic distance measurement, i.e. the dependence on high-contrast structures and occlusions. On the other hand, the resolution is currently much

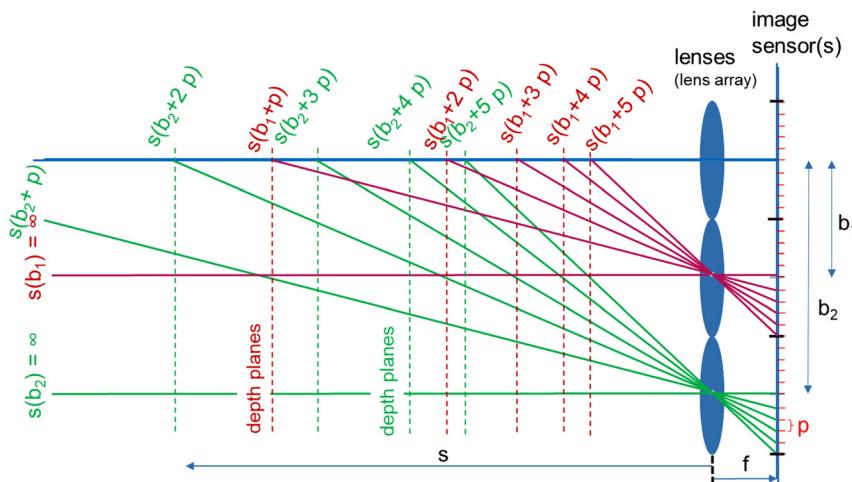


Figure 83: The position of the “depth planes” depends on the base distance (b_1, b_2, \dots). More cameras lead to more depth planes and will increase the depth resolution.

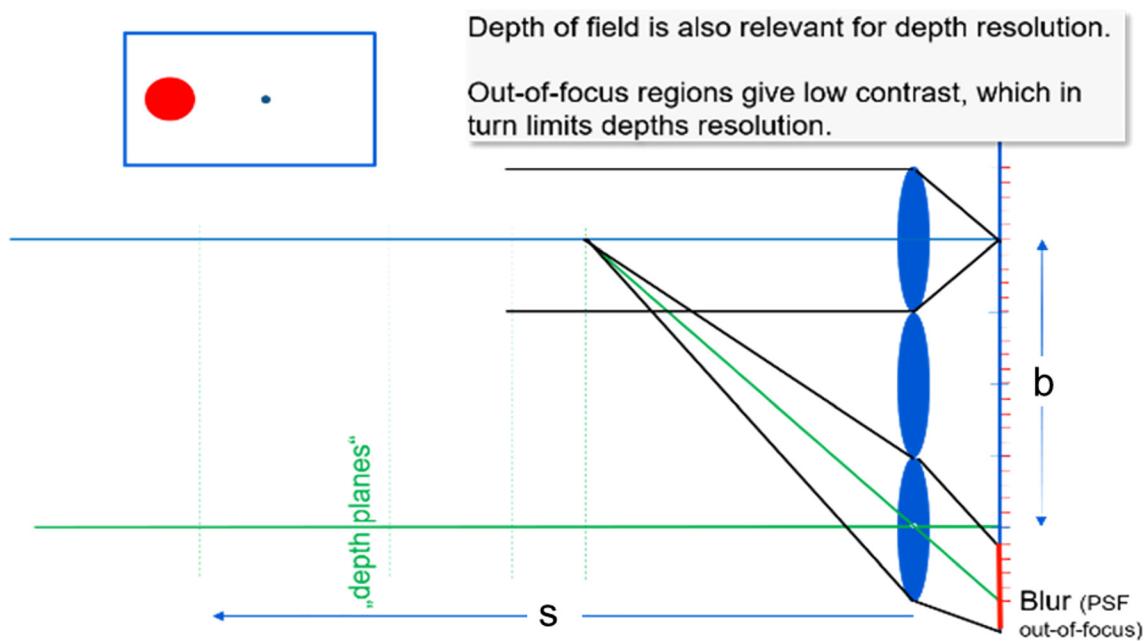


Figure 84: Depth resolution depends on the f -number, as well as on the sensor size and pixel resolution on the actual focusing position: For regions near the focus plane, the depth resolution increases due to better lateral resolution.

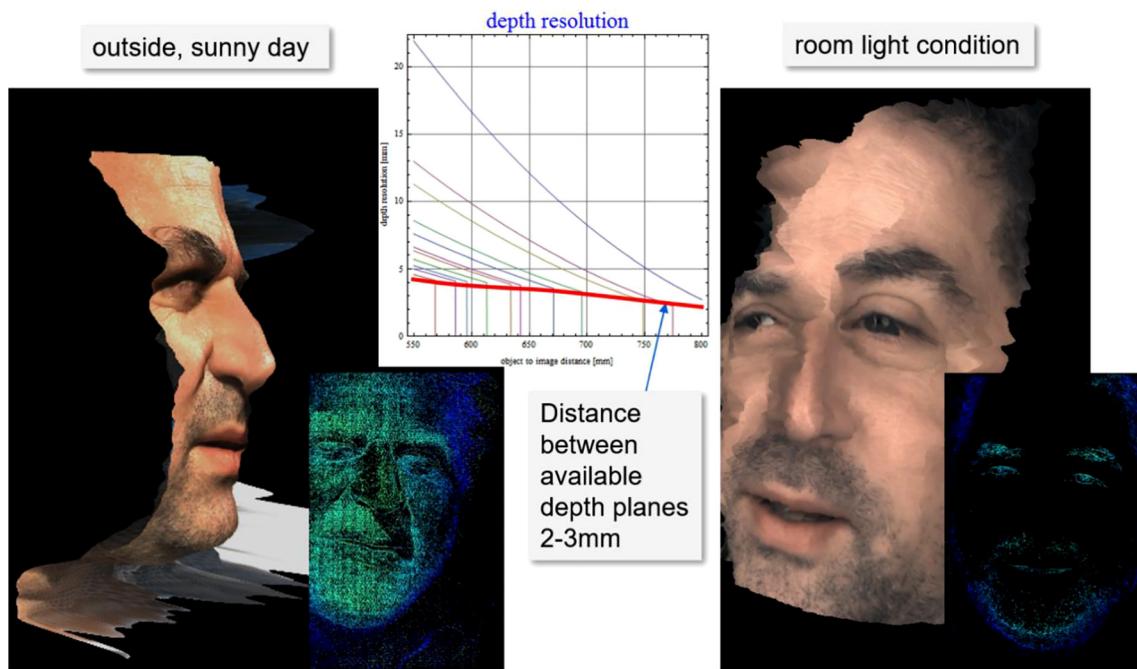


Figure 85: The quality of the stereoscopic depth estimate is highly dependent on the contrasts in the image and thus also on the intensity, direction and spectral characteristics of the illumination: Despite nominally sufficient depth resolution, there are problems with depth detection in the example shown for the moderately illuminated indoor shot (right), while the outdoor depth map shows very good, high-resolution depth detection (left).

smaller, at around 240×180 pixels. Figure 87 shows the system layout and the principle [186].

3D depth acquisition cameras that augment the traditional front-facing camera are now fitted as standard in

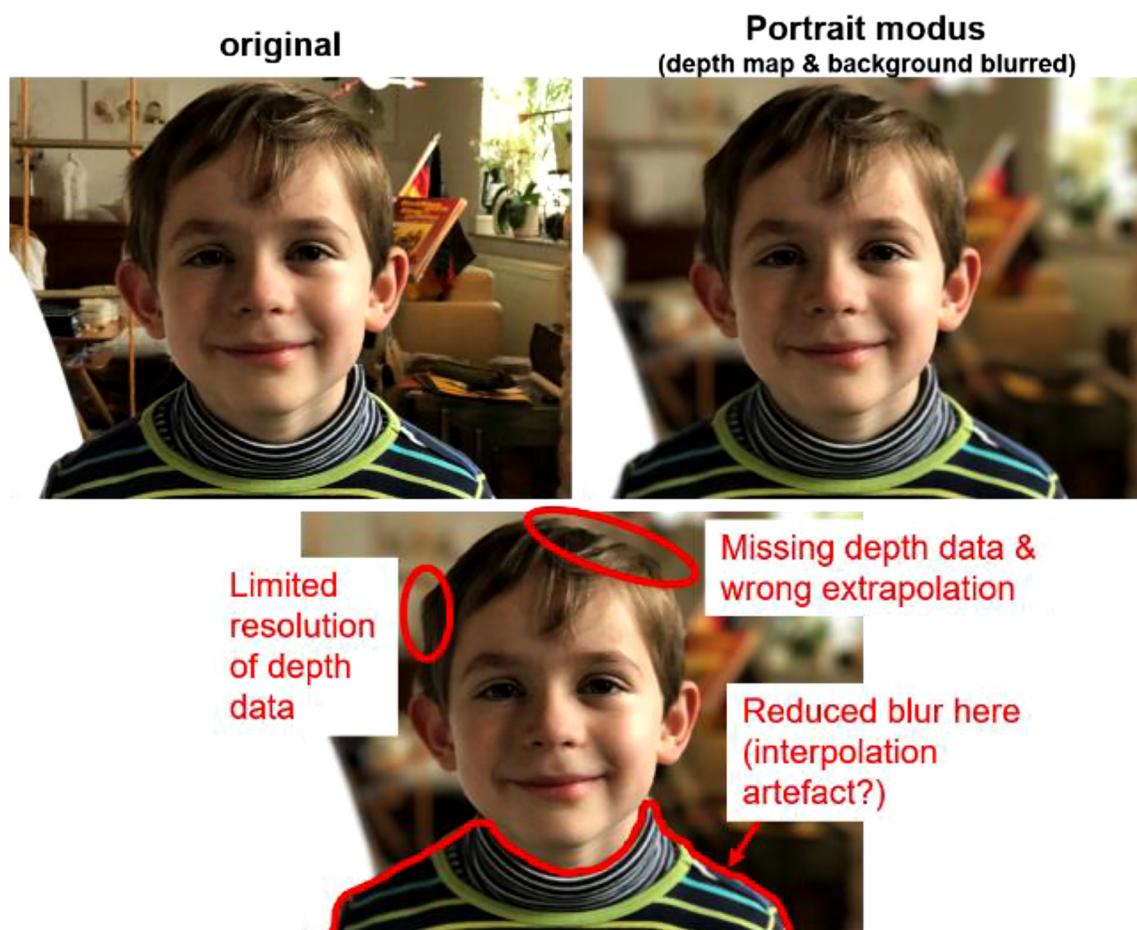


Figure 86: Failures with depth detection (picture taken with the iPhone 7+ in portrait mode).

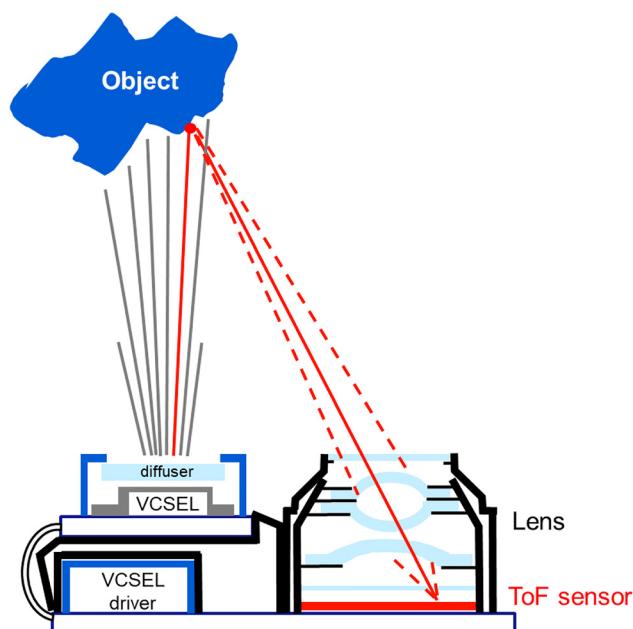


Figure 87: Time-of-flight sensor for 3D acquisition.

high-end smartphones. Some smartphones use a structured light camera system for face recognition. In most cases, a time-of-flight system is used. This component provides the depth data of objects in the distance, between approximately 0.15 and 1 m in real time. 3D acquisition systems are also used on the rear side in combination with the rear camera system. The light source is typically a vertical cavity surface-emitting laser (VCSEL); for a concise presentation on principle, research and applications see ref. [187] at a power of about 150 mW in near IR light, e.g. $\lambda = 940$ nm. In order to deduce the time of flight that is temporally pulsed or continuous-waved, modulated light is required and a multiple pixel system (e.g. consisting of 2 or 4 pixels) for a single direction within the FOV, where the pixels detect in a mutually phase-shifted way in order to recover the travel time of light [188, 189]. The dimension of one of these pixels is about 3 μm . As for usual imaging, a lens is required to focus the light on the sensor. The lens' field of view is usually similar to the standard wide-angle imaging lens, e.g. a 75–80° fully diagonal FOV. The

aperture is also similar, e.g. $f/2$ in order to support the required lateral resolution. The sensor is about 4.5 mm in diameter. The working distance range is about 0.2–2 m.

15.2 Simulation of lens bokeh: Camera 3D point spread function

The equation of the diameter of the geometrical PSF, the “circle of confusion,” through depth was given in Section 4.4; Eq. (17). The diameter of the circle of confusion relative to the image diagonal with respect to the focus distance (s_F) and the distance from the lens (s) is:

$$\varnothing_{\text{rel.spot}} = \frac{f^2}{\varnothing_{\text{im}} K} \frac{|s_F - s|}{s_F s} \quad (73)$$

This equation is true for “normal imaging.” That is, if we do not consider macro distances. The spot diameter dependence is shown in Figure 88 for different focusing distances.

The imperfections of the bokeh created by a lens reveals much more information about a lens than the sharpness it delivers when in focus. In photography the term “confused,” as “bokeh” translates from Japanese, relates to light beams which no longer come together at a single point, and so are out of focus [190]. A non-uniform intensity distribution within an out-of-focus highlight source is due to lens aberrations, i.e. overcorrected low-order spherical aberration gives rise to an “donut shape” that has highlighted edges at the spots in the image foreground whereas in the background the distribution is brighter in the spot center (Figure 89).

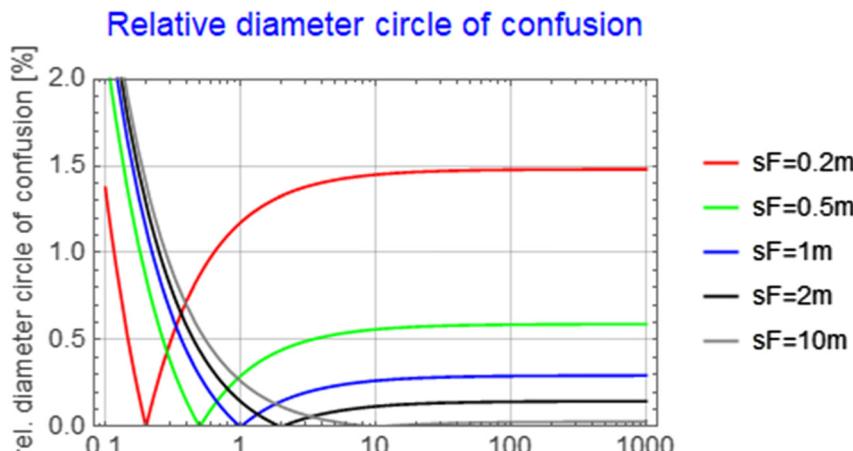


Figure 88: Relative diameter of the circle of confusion versus object distance from the lens for different focusing distances s_F according to $\varnothing_{\text{rel.spot}} = \frac{f^2}{\varnothing_{\text{im}} K} \frac{|s_F - s|}{s_F s}$. The lens is a $f/1.7$ standard wide-angle, $f_{\text{eq}} = 28$ mm (real $f = 7.8$ mm), for an image sensor diameter of $\varnothing_{\text{im}} = 12$ mm.

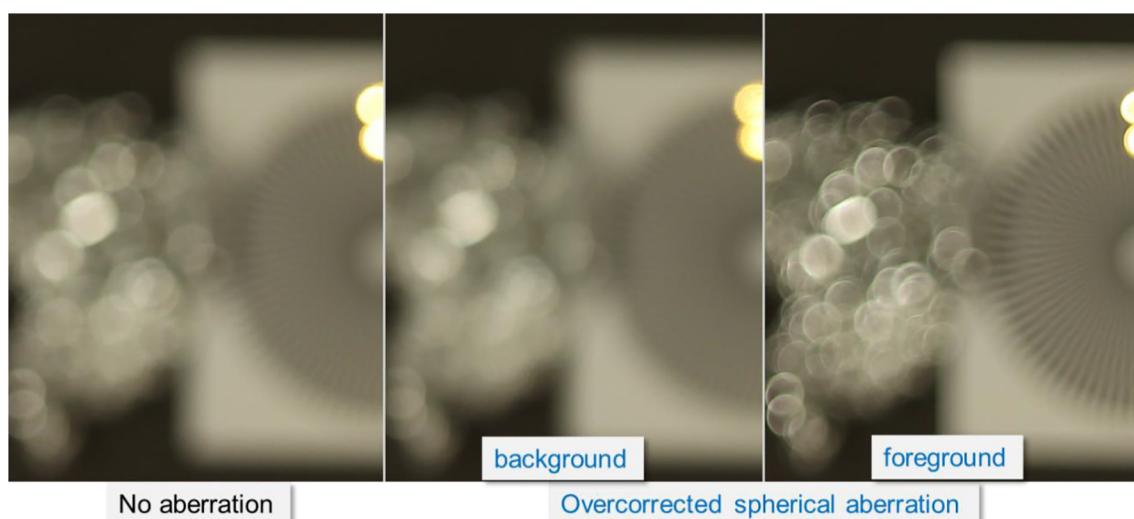


Figure 89: Effect of spherical aberration on out-of-focus spot distribution.

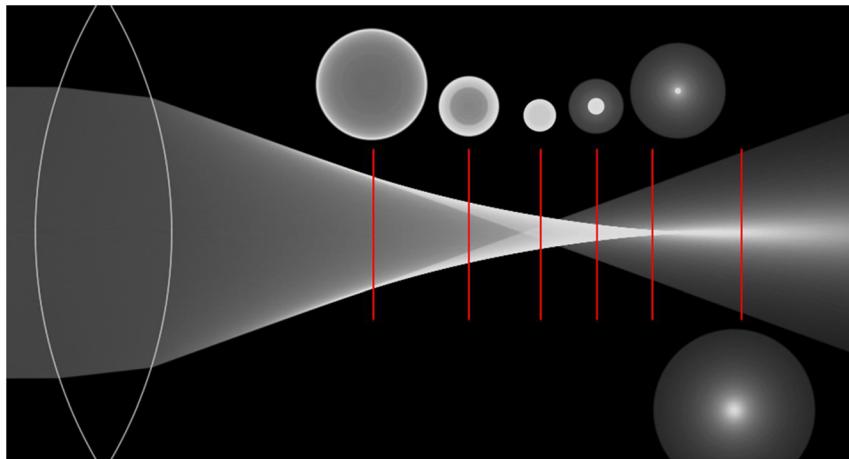


Figure 90: Through-focus intensity distribution in the presence of overcorrected spherical aberration.

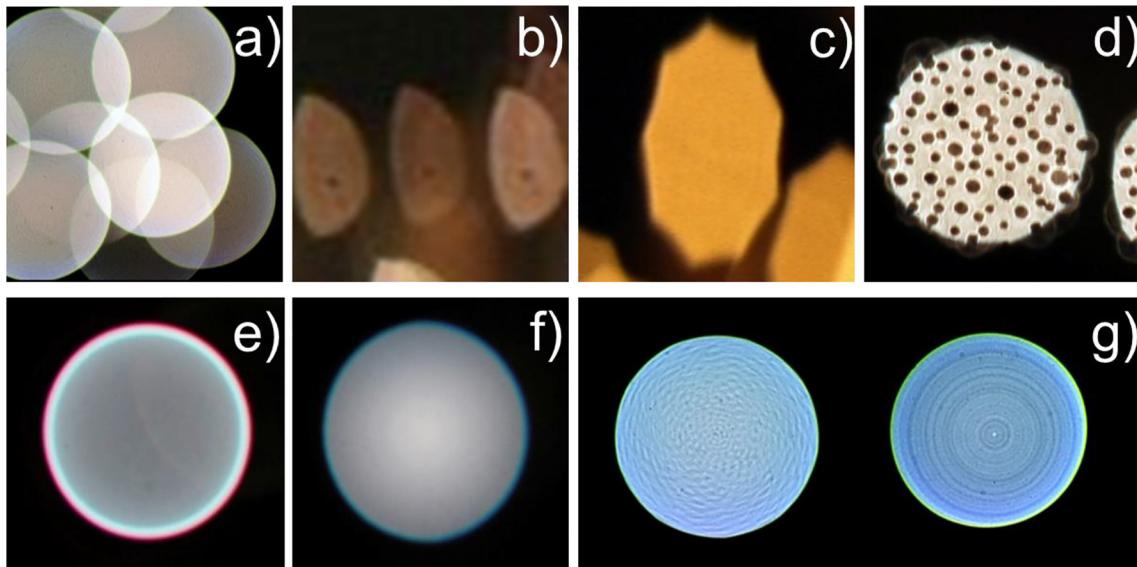


Figure 91: Out-of-focus highlights of lenses: (a) Overlapping out-of-focus highlights, (b) “cat’s-eye bokeh” due to vignetting, (c) “edgy bokeh” due to iris stop (diaphragm with 9 blades), (d) Softar bokeh (Softar filter scatters light out of nominal light path), (e) “donut bokeh” (spherical aberration, also red color fringe due to chromatic aberration), (f) “Christmas ball bokeh” (spherical aberration), and (g) “fine structured bokeh”, bokeh on the right sometimes called “onion ring bokeh” (optical manufacturing induced residual surface deformations).

The effect is due to the redistribution of intensity in the presence of aberrations (Figure 90). Furthermore, local lens surface defects, e.g. polishing artefacts, may appear as spatially high-frequent structures.

As high-aperture camera lenses vignetted due to mount constraints [148] the bokeh spots of these lenses looks increasingly like a cat’s eye (Figure 91c)), i.e. towards the image corners. This “cat’s eye” is created by the lens’ field stops in the front and rear part of camera lenses. There are many more attributes, like chromatic aberrations, apodization, and the digital filter components, which lend the bokeh of a specific lens a characteristic look (Figure 91).

In addition to the PSF, depending on the depth and FOV for the respective focusing position, yet another parameter is important for the appearance of the bokeh (especially in comparison to the “natural bokeh” of a DSLR lens): the light intensity of the respective source point. This is not a problem as long as the dynamic range is not exceeded, but as soon as it is exceeded, that is to say overexposed, it is unknown whether it is exceeded only slightly or by several orders of magnitude. While for a real lens the irradiance of a (local) light source is physically redistributed inversely proportional to the surface area of the out-of-focus spot, for the synthetic bokeh of an SPC the



Figure 92: Left: Normal shot captured with a smartphone; right: Shot captured with a full-frame camera and a highly stopped-down lens ($f/16$) to give the photo a rich depth of field.



Figure 93: Left: Shot captured with a smartphone (iPhone X) in portrait mode; right: Shot captured with a full-frame system camera (Sony α7 with ZEISS Batis 2/40).

spot brightness must be guessed as the relative irradiance is unknown.

15.3 Portrait look: a quality evaluation

The quality of synthetic bokeh has been continuously improved for SPCs in recent years. However, it is very difficult to approximate full-frame camera quality due to the issues we discussed above (depth map quality, realistic bokeh is computationally expensive) and because of the inferior low-light capabilities and reduced dynamic range. Figures 92 and 93 show a challenging low-light scene for computational bokeh with bright local light sources.

Obtaining a high-quality bokeh computationally in SPCs is especially challenging in low-light scenes with very bright background highlights for several reasons:

The face is usually noisy due to the low ambient light. And the background is initially just as noisy in the real recorded image but is smoothed out by convolution to create the bokeh effect. The side effect is that there is a difference in noise between the person portrayed and the background, which is perceived as unpleasant. You can compensate for this difference in noise by artificially adding noise to the background and/or removing noise from the face area.

In addition, very bright highlights are problematic because they are overexposed in the real photo. As soon as all color channels (R, G, and B), are overexposed, a colored light source will appear white. The color of the source, that is a difference between RGB as it would appear within the dynamic range, cannot be recovered. Consequently, the computed blurred background will be incorrectly displayed as white (Figure 94). This may not happen when taking a picture with a full-frame camera: The overexposed area of a bright light spot is distributed

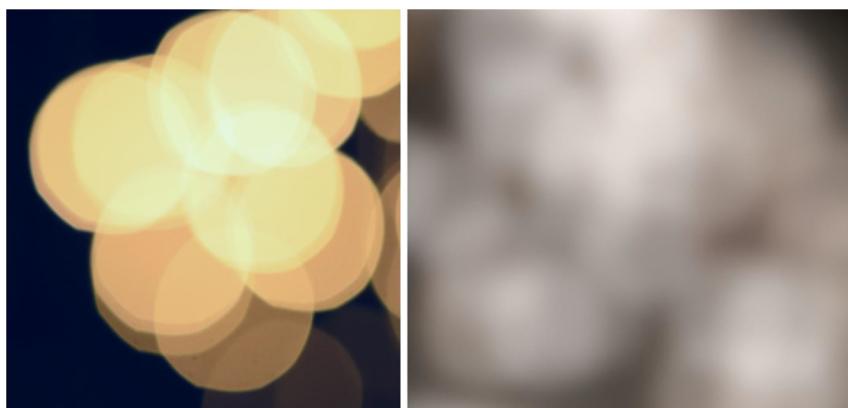


Figure 94: Detailed look at a bokeh (here: Out-of-focus spots of several bright light sources): Physical bokeh of a full-frame camera and high-aperture lens (left) versus the synthetic bokeh of a typical SPC (please note: there are significant quality differences between different SPC suppliers).



Figure 95: Lens bokeh of a classic camera lens for a modern SPC's portrait mode.
Image Credits: vivo.

over a much larger area due to the defocusing, and decreases in intensity according to the area it occupies. If the resulting intensity is less than the saturation value of the dynamic range, the bright light spot no longer appears oversaturated as a uniformly bright spot, but as a distribution and additively overlaps with the structures in the area, possibly also with neighboring, defocused highlights.

However, the quality of synthetic bokeh has improved significantly at some smartphone manufacturers in recent years. Figure 95 shows a portrait image with the vivo X60 Pro+ with a “vintage look” based on real lens data. The original optical construction data of a famous classic 35 mm format camera lens, the Carl Zeiss Biotar 2/58 mm from 1936, are used to create the bokeh.

Bokeh quality tests have helped to push and guide smartphone suppliers to systematically improve synthetic bokeh. As of 2018, DXO offers a test setup with critical objects for high-resolution depth acquisition in topologically enclosed areas like holes or tiny eyelets (Figure 96).

With structured stripes, which are placed diagonally into the depth, it must be checked whether the natural sharpness transition in the depth is reproduced continuously and naturally.

The evaluation criteria are based on the ideal of the natural bokeh of a full-frame photo camera. The test criteria are as follows: 1. subject background segmentation; 2. repeatability; 3. blur gradient smoothness; 4. bokeh shape; 5. equivalent aperture; 6. noise consistency. Criteria 1 to 3 relate to depth map quality, while 3 to 5

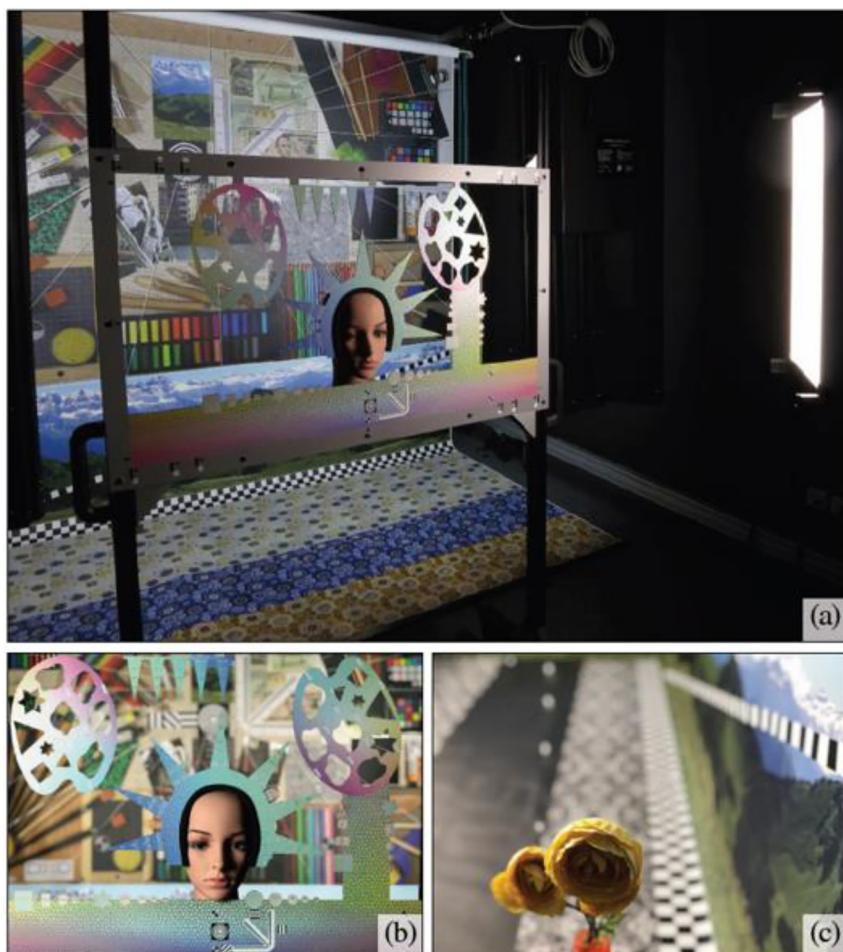


Figure 96: DXO computational bokeh test
(Source: Reference [191]).

relate to the 3D PSF model, and criterion 6 to the aforementioned “noise” uniformity by any additional image processing.

16 Image performance specification and test

Image quality tests are extensive and done at different stages in the development and production processes:

- Lab evaluation during R&D
- Qualification in mass production
- Qualification of the image quality, including signal and image processing in the smartphone

We mentioned qualification in mass production earlier (in Section 8.3). We will now briefly summarize other tests performed during R&D and production.

16.1 Lab evaluation during R&D

The focus of these tests is to objectively check measurable optical and sensor properties. A typical specification for the development of a camera module can be between 20 and 120 pages. Most of the requirements are related to electronic functions. The image quality of the module is assessed using the following tests:

- Resolution and contrast
- Color reproduction
- Field of view and distortion
- Dynamic range
- Auto exposure (AE)
- Autofocus (AF)
- Auto white balance (AWB)
- Lens shading
- Color shading
- Flare
- Ghosts

Figure 97 shows several standardized test charts for these evaluations.

In addition to these optical performance parameters, numerous other tests and inspections must be carried out by the module manufacturer. These include:

- Material tests (RoHS)
- Dust and environmental tests
- Shock and vibration
- Continuous run, lifetime
- Electromagnetic compatibility
- Electronic tests e.g. software and sensor interfaces to the camera, drives (VCM, OIS), etc.

16.2 Evaluation of image quality in the imaging pipeline

During development of the signal processing and imaging software, the camera modules and the entire imaging chain are evaluated and verified.

Not only are the objectively measurable and assessable criteria that have already been tested during module development checked, numerous subjective quality parameters are also taken into account. In addition to the technical test charts, test images of natural objects, special image arrangements, and people (Figure 98) are created for evaluation purposes.

In order to make the subjective evaluations as objective as possible, the respective evaluations are performed by several trained persons according to precisely specified test and evaluation procedures (lighting, viewing time, etc.) and converted into key figures.

Since the development of the image processing software takes place practically until the device is sold (and beyond), it is extremely important to be able to assess as early as possible which deficiencies can still be remedied during regular development and which weak points still require special treatment. The camera device must then be released on key date X, even if the final quality can only be estimated at this point in time.

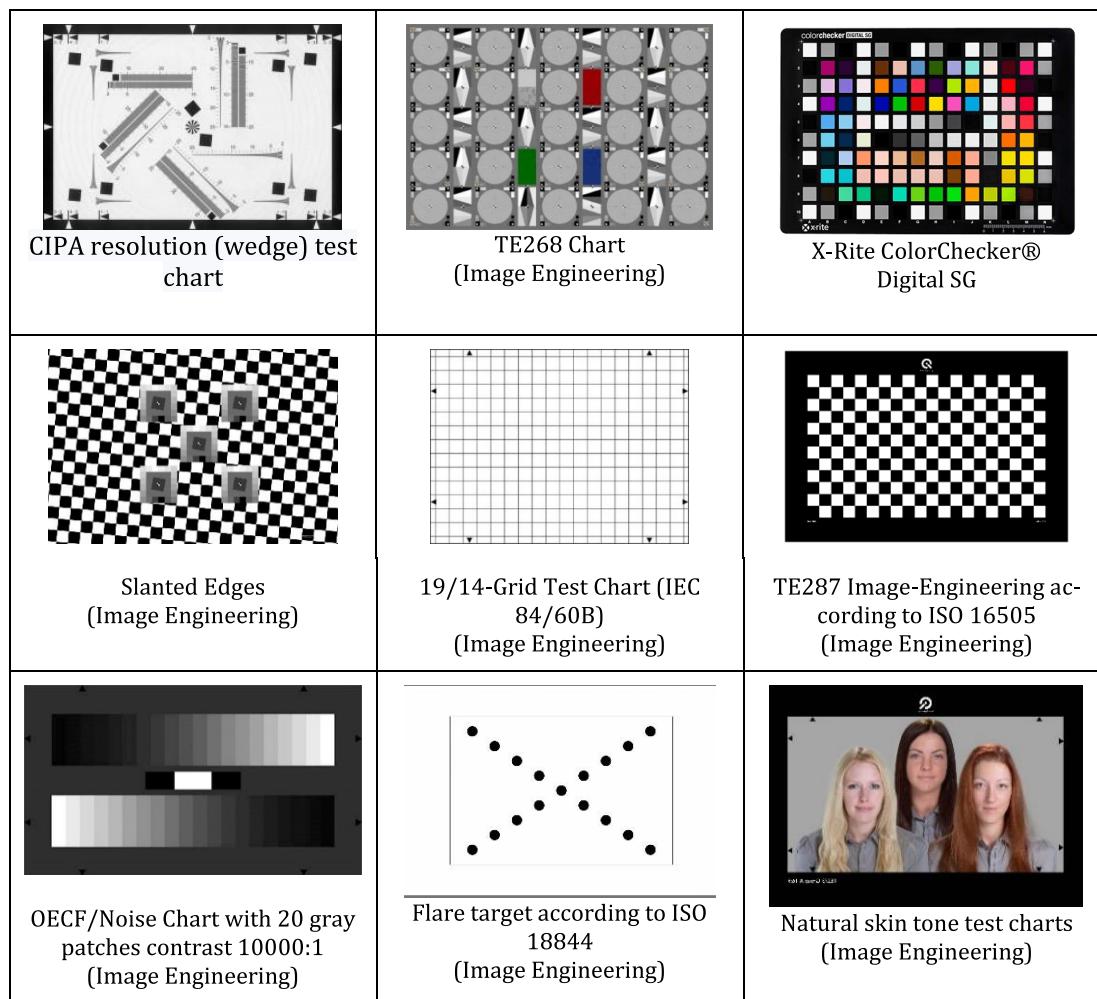


Figure 97: Examples of test charts. Courtesy of Image Engineering ref. [208].

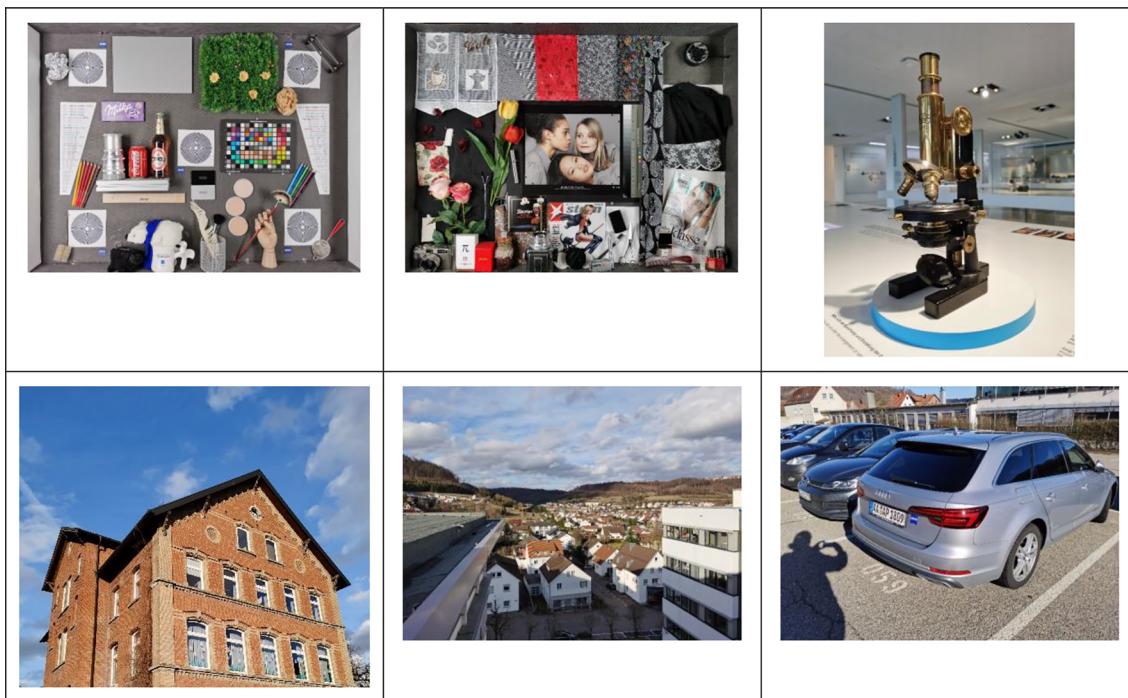


Figure 98: Examples of subjective motifs.

While there are different and mostly standardized methods for measuring and evaluating individual parameters, there are no recognized standards for the overall evaluation of a camera system.

In recent years, DXOMark (<https://www.dxomark.com/>) has established itself in the broad public perception as a commercial provider of tests. DXOMark is an excellent resource for tests for many different (SPCs and for numerous reports on image quality evaluation. Unfortunately, the test procedures, evaluation matrix and weighting, especially when evaluating subjective image parameters, are not freely accessible and therefore only partially comprehensible and reproducible from the outside.

The nonprofit organization VCX (<https://vcx-forum.org/>) aims to bridge this gap across multiple manufacturers and providers by developing a comprehensive test methodology and adapting it to the constantly changing requirements [192]. Great importance is attached to transparency and traceability. Any manufacturer or supplier can join this consortium, work on further developing the test procedures, and subsequently apply the test procedures themselves.

In order to guarantee the objectivity and neutrality of the evaluations, the implementation of “official” tests with ranking is exclusively reserved for manufacturer-independent test laboratories certified by VCX.

17 Smartphone camera interface with telescopes, microscopes, and accessory lenses

Let us consider another practical aspect: With the SPC one can very easily take pictures through a telescope (Figure 99), binoculars or a microscope, simply by placing the SPC on the eyepiece instead of the eye. Additional optics for connecting a smartphone and telescope, binoculars or a microscope to one another are not required.



Figure 99: Image of the moon taken with SPC through the telescope at the Aalen observatory.

This is possible because the FOV and entrance pupil diameter of an SPC are similar to that of the human eye and thus also the eyepieces that are designed for it (Figure 100): With the 28 mm-equivalent SPC wide-angle lens, the entrance pupil diameter is around 2–4 mm, like that of the day vision of the human eye (night vision approx. 5–7 mm depending on age). The captured image field angle of typical high-quality eyepieces on telescopes, spotting

scopes, binoculars, or microscopes is often 60–70° (image circle diameter) and is therefore completely captured by the SPC wide-angle optics (FOV approx. 70–80°).

However, good freehand photos are a matter of patience and luck: The tolerances for the positioning (centering and distance) of the SPC entrance pupil are smaller than one 10th of a millimeter, otherwise we have to accept a loss in sharpness. There are some precise adapter

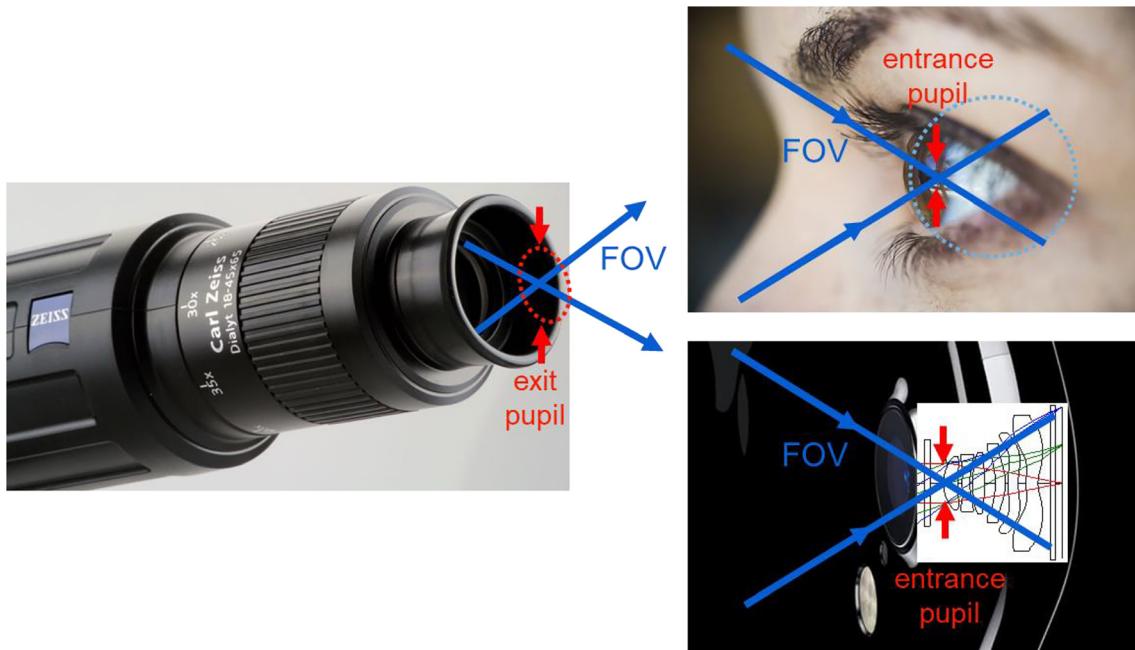


Figure 100: The smartphone's exit pupil diameter in daylight conditions of approx. 2–4 mm is similar to the entrance pupil diameter of the human eye. As the entrance pupil of the smartphone is also at the front of the device, it can be placed right at the desired position of the optical instruments' eyepieces' exit position. Typical fields of view of high-end spotting scopes or astronomical telescopes or microscopes are very similar to the FOV of the smartphones' standard wide-angle lens of approx. 75°.



Figure 101: Adapter for attaching a smartphone to a spotting scope.

solutions available (Figure 101). Usually these adapters are only suitable for a limited number of eyepieces and only for certain smartphones.

It is important that the entrance pupil of the camera lens can be brought exactly to the position of the exit pupil of the eyepiece. This is almost always the case with the SPC because the entrance pupil is at the front. However, the entrance pupil on full-frame cameras is usually located deep inside the lens. If the distance from the front lens to the entrance pupil exceeds the eye clearance of the eyepiece – in the case of eyepieces designed for spectacle wearers this distance is about 18 mm – then the image field is vignetted. This means that one can only see a limited area of the image field. For direct observation with an eyepiece, a certain amount of field curvature can be tolerated because this can be compensated for by the accommodation of the eye while the fovea wanders through the field of view. The amount of compensable focus of a 50-year-old human is about 2–3 diopters, but strongly depends on the humans age [193]. Now a field curvature corresponding to 1–2 diopters clearly deteriorates off-axis image quality when viewed through a camera on an image sensor.

Sometimes a connection to a camera with a longer focal length would be more attractive because the entrance pupil is larger, and the image field appears more magnified. It is not uncommon for the internal image processing of the camera system to thwart the bill. This is because the “telephoto range” also requires the wide-angle camera to be active and not covered, as the software needs the image from the wide-angle camera to display the two images in the telephoto FOV in order to merge them.

The modern integrated multicamera systems have replaced the accessory lenses which were on the market for many years. These auxiliary optics, or converter lenses, cause a moderate change in the focal length: Typical converter lenses achieve a telephoto factor of 2–3× at angular magnification or 0.6× for distortion-free wide-angle conversion. Even at these moderate conversion factors quite significant effort is required, e.g. usage of aspherical lens elements, to achieve high-contrast and distortion-free image performance (Figure 104). Furthermore, there are converters for extremely wide FOV fish-eye images. It is practically impossible to implement larger angular magnification or demagnification in good quality: The afocal design [194] must be in

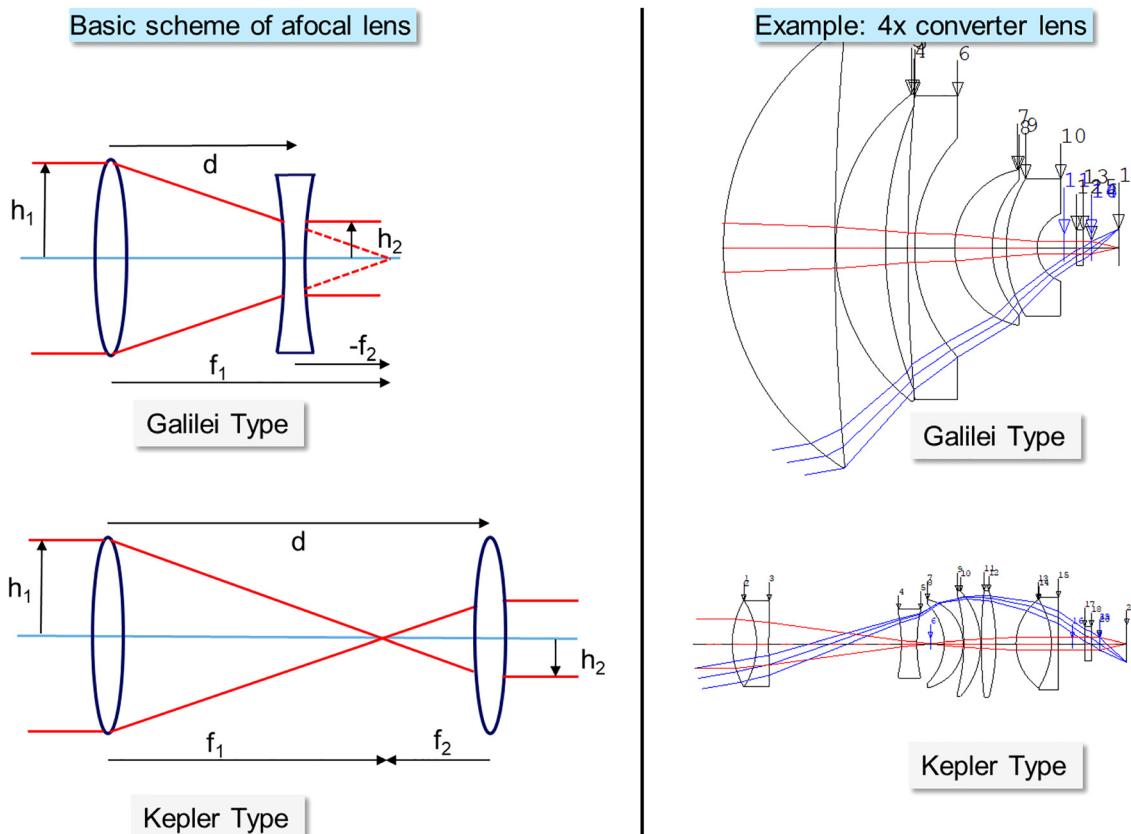


Figure 102: Scheme of afocal lens: Galilei and Kepler Type (left) and optical design example of a 4× converter Galilei and Kepler Type (right).

the Galileo design, i.e. with a positive and negative lens group but without an intermediate image, so that the image on the display does not appear upside down, as with a Kepler type, which contains an intermediate image (Figure 102).

The optical lenses available on the market were not connected to the SPC with any electronic interfaces to enable the system to recognize that the image would have to be turned upside down electronically when connecting. The diameter of a Galileo attachment lens increases almost proportionally to the TF and is unacceptably large – at around 4× – and delivers good image performance at the same time (Figure 102 (right)). For larger telefactors it is cheaper to use Kepler systems [195].

While afocal lens attachments have practically been replaced by integrated camera lenses with the corresponding focal lengths, smartphone-integrated macro lenses are still a rarity, such as in the Nokia 8.3. Such macro lenses achieve magnifications of around 1:6 to 1:8, i.e. they image an object field of around 25–35 mm and can resolve

object structures of around 10 µm. With this magnification, it is inevitable that the object is very close, only about 20–30 mm in front of the cover glass, in front of the smartphone, and that the smartphone housing covers a lot of ambient light. The locally mounted, laterally offset flash is unsuitable for uniform and natural lighting. This is a major disadvantage compared to macro lenses for large image formats, with a magnification of up to 1:1, i.e. an object field of 43.2 mm in diameter and an approx. 10-µm resolution, where the working distance between lens and object is much larger. Here, the external lens attachment for SPC in Figure 103 (left) is still a good compromise: Ambient light reaches the object from the side through the diffusely scattering glass ring. The thickness of the glass ring corresponds exactly to the necessary focus distance, so the system is placed on it when taking a picture. The layout with two lens groups enables high-contrast, distortion-free images. Furthermore, via an adjustment mechanism for the distance between the two lens groups an additional manual focusing mechanism can be implemented, which



Figure 103: Auxiliary optics for SPCs. Macro close-range optics, and wide-angle and telephoto converters.

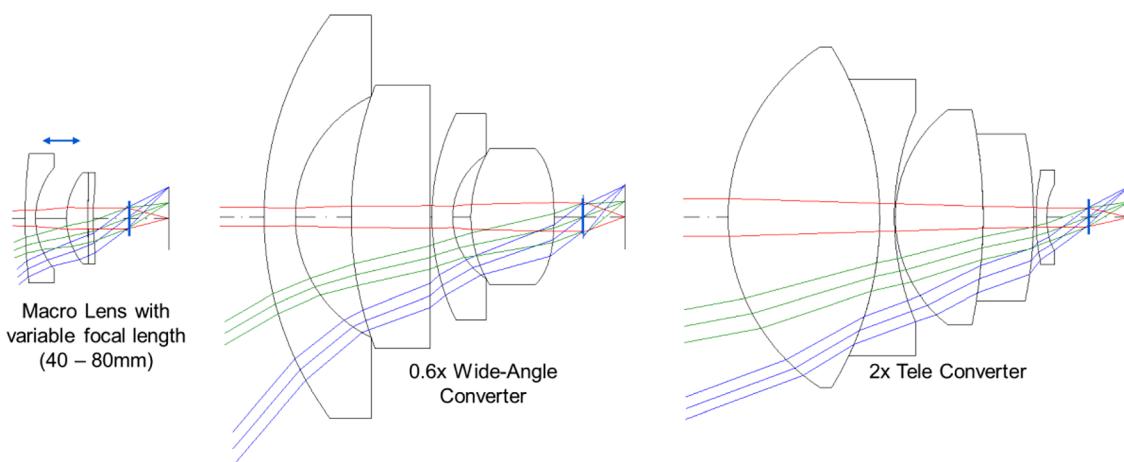


Figure 104: Optical design of accessory lenses shown in Figure 103: Vario-Proxar 40–80, Mutar 0.6×, and 2×, respectively, (drawn on same scale) [196–198]. The SPC lens is represented by an ideal lens with a 4 mm focal length (thin blue lines).

enables the photographer to shoot sharp pictures within a certain magnification range [198].

Microscope attachment systems go one step further in terms of object resolution: Such systems are described by Switz et al. [199–201]. There are also some commercially available systems, e.g. <https://diple.smartmicrooptics.com/>.

One can easily achieve very good imaging performance by placing a smartphone lens directly in front of the main camera (see Figure 105). This is possible because the diaphragm is in front of the first lens. A nearly diffraction-limited image is then obtained over a large part of the image field, because each lens is well corrected, and the beam path is collimated as in the original system. Then one has a 1:1 magnification, the resolution of which is essentially only limited by the sensor pixel size due to the high optical quality. Since the pixels are about 1 μm in size, the object detail resolution is about 2 μm , which is close to the resolution of an ordinary professional light microscope (Figure 106). The object field is about 6 mm in diameter, which is comparable with commercial light microscopes. Smartphone microscopy is not exclusively an add-on feature anymore, as beginning of 2021 in OPPO Find x3 Pro a camera module with a magnification around 1:1 has been integrated.

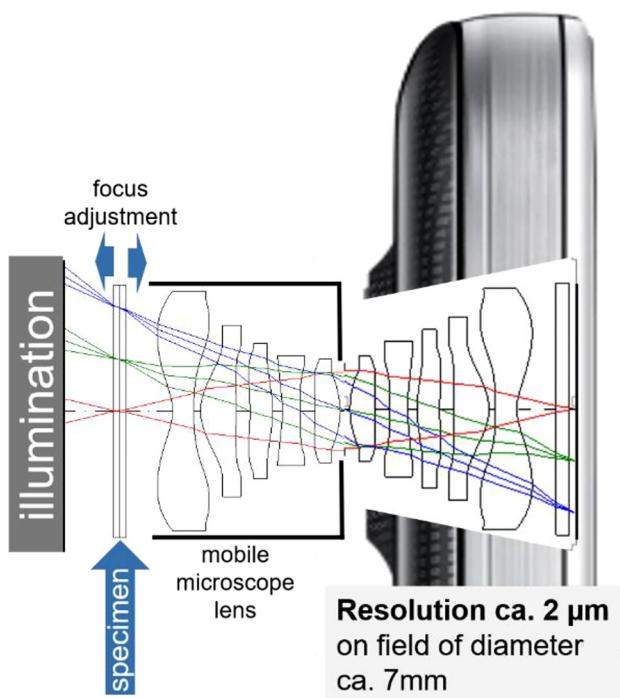


Figure 105: Mini microscope consisting of two smartphone lenses, image sensor, specimen sample with cover glass, in front of a light source (e.g. area light emitting diode (LED)) for uniform illumination of the specimen.

Structure recognition in microscope applications, e.g. when analyzing bacteria and parasites in drinking water or clinical pictures in blood samples, has significantly improved in recent years thanks to machine learning. Based on sample data, and possibly also on system settings such as lighting mode or depth stack, neural networks learn relationships with which features in image data can be better identified.

There are several accessory products for healthcare, monitoring and diagnosis, like Shack–Hartmann sensors for eye defect analysis, spectrometers, and microscopes [202, 203]. Ultimately, finding everyday applications for a large number of people will decide whether those functionalities will be integrated into smartphones or not. During the COVID-19 pandemic, some smartphones – like the Huawei Honor Play 4 Pro 5G – have been equipped with IR thermal sensors for quick human body temperature measurement. Overall, medical and healthcare applications benefit from the improved capabilities of smartphones [204]. FLIR offers thermal IR images that can be used to identify heat leaks in buildings. However, this functionality is limited to a niche market.

18 Summary and outlook

Almost nobody would have expected the excellent quality of today's SPCs to be achievable 20 years ago when the first miniature cameras were integrated into mobile phones. They have almost completely replaced compact cameras and have slowed professional camera system sales (Section 2.1). Today the quality of the recordings with SPCs is hardly distinguishable from professional cameras in many everyday situations, especially when there is plenty of ambient light. In more difficult conditions such as low-light situations or to capture fast-moving subjects, however, there are still significant differences, but these will continue to become more minor in the future. The basic physical disadvantages of the miniaturization of a photo camera are (Section 4):

- (1) For optical image formation diffraction limits to achieve very high pixel resolutions
- (2) Very large depth of field, which is undesirable in creative photography (e.g. portraits)
- (3) Small pixel size results in low light per pixel: increased noise or longer exposure times
- (4) Due to space limitations in a smartphone body, tele lenses, and optical zoom systems are often not possible

In Sections 4.3, 4.4 and 4.5 we show that the first 3 disadvantages can be traced back to a single parameter, the

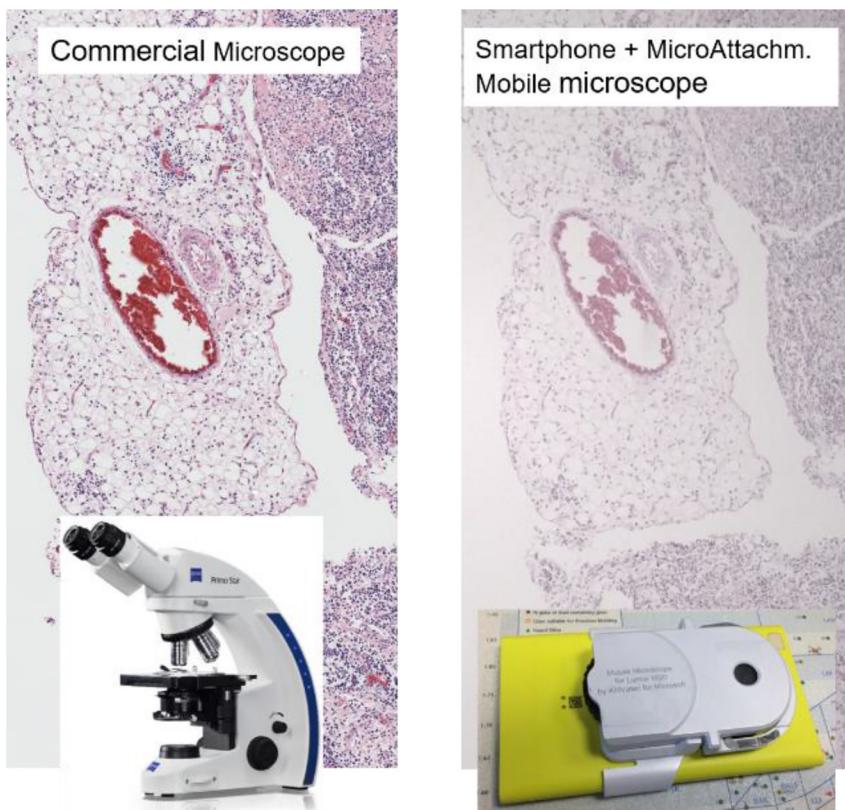


Figure 106: Comparison of the image quality of a smartphone microscope with that of a commercial microscope. The smartphone resolves structures just a few micrometers in size.

optical systems etendue. In addition, starting from the current system architecture with scaling laws, we discuss the consequences of a further reduction in the size of the systems or of pixel sizes (Sections 4.6 and 4.7). Accordingly, a further reduction in the pixel size below the current level of $0.7\text{ }\mu\text{m}$ would have to increase the complexity of the optical system considerably due to the necessary increase in numerical aperture. But, this is contrary to miniaturization. With the current system concept, it is therefore not to be expected that significantly better optical performance will be achievable in the future.

The efforts of a branch of industry with strong growth that is closely interlinked (Section 2.2) led to dynamic technology development (briefly summarized in Section 3) to reach the limits of the physical limitations mentioned and to circumvent them in some cases. Computational imaging in particular was responsible for shifting the physical limits, which was spurred on by the considerable increase in computing power on mobile platforms, greatly improved algorithms, and hardware extensions such as multicamera systems and 3D acquisition systems (e.g. ToF sensors). Identity verification and augmented reality applications will be a motor to further improve imaging including 3D acquisition and position recognition sensors in the years to come.

Without the user noticing, a single image is created through multiple recordings, being processed in various ways depending on the recording situation or camera mode. In the meantime, artificial intelligence algorithms identify image content through the connection to image databases linked with position recognition via global positioning system (GPS) and change or insert them virtually into the recorded image, or remove disruptive objects. In portrait mode, the large depth of field, as dictated by physics, is computationally reduced by convolution of the 3D data of the scene with a depth-dependent point spread function according to the depth map of the scene (Section 15). Dynamic range can be extended by exposure bracketing, that is combining multiple images with different exposure times to form an HDR image. This function has been further improved in recent years with multicell image sensors, in which individual pixels are controlled in different ways, e.g. with regard to the exposure time (Section 14). Optionally, the effective pixel area can also be increased in order to reduce noise in low-light situations. Advances in the architecture of CMOS image sensors (Section 9) or context- and situation-sensitive noise suppression algorithms in image processing (Sections 10 and 11) also help in a low-light environment. The subjects of image sensors and image processing are presented here rather succinctly with reference to the very extensive literature.

We go into more detail about the optical system and optical technologies (Sections 6–8). The wide-angle lens of the main camera has become more and more complex and powerful over the years and, with a ratio of track length to sensor diagonal of only about 2/3, achieves a resolution that is almost diffraction-limited with high apertures of about $f/1.6$ and thus supports image sensor resolutions of almost $1\text{ }\mu\text{m}$. This very compact size, which is far superior to comparable classic optical designs made of glass, is made possible by optical designs with extreme plastic aspheres. In the past few years, more and more lenses with other focal lengths have been integrated and supplement the standard wide angle as a super wide angle up to around FOV 120° or tele lenses either of conventional and periscope layout with field-of-view up to approx. 16° . Due to a smartphone's space constraints, both extreme wide-angle and tele systems do not achieve the same high optical resolution as a standard wide-angle lens. Correspondingly, in the hybrid zoom systems consisting of several lenses of different focal lengths, the camera resolution constantly decreases towards very large or very small fields of view (Section 7.1). This is a disadvantage compared to many real optical zoom systems for SLR or system cameras which, unfortunately, cannot be scaled to the space restrictions of smartphones (Section 7.2).

The enabler of miniaturization is the production technology of the complex components, i.e. the strongly aspherical lenses with integrated mounts, which can be mass produced in plastic with high precision and reproducibility (Sections 8.1 and 8.2), paired with a fully automated assembly and adjustment process for the lens and the active alignment of lens to image sensor with precise robotics and fast measurement technology (Section 8.3) based on extensive yield optimization in the optical design process (Section 8.4). The costs for high-end SPCs are well below \$20 per camera module. Some parts and components, such as time-of-flight cameras, are increasingly being manufactured lithographically using wafer-level technology. It is possible that even more lithographically manufactured components, including photonic integrated circuits, will be used in smartphones, and linked to electronic circuits in the future.

In Section 12 requirements for focusing accuracy and distance determination are derived. The entire lens, driven by voice coil motors, is moved forward for close-up focus. Motion blurring is reduced with EIS and OIS. OIS has improved the image quality of lower-light images by compensating for hand tremors by enabling longer exposure times (Section 13). OIS was made possible by integrating MEMS gyroscopes and

accelerometers in smartphones as well as in the refinement of the actuators.

Finally, in Section 17, SPCs' connection interface with telescopes, microscopes, and other auxiliary optical systems is reviewed. Due to the similarity of the SPC with the human eye in terms of pupil position, size, and field of view, the connection is usually possible without significant loss of quality. With the connection to powerful image processing and extensive image databases, the images can be processed and interpreted and then sent or shared via the connectivity interfaces.

Acknowledgement: The authors would like to thank Dr. Norbert Kerwien for his valuable comments.

References

- [1] Techterms, 2021. Available at: <https://techterms.com/definition/mobile-phone>.
- [2] M. B. del Rosario, S. J. Redmond, and N. H. Lovell, "Tracking the evolution of smartphone sensing for monitoring human movement," *Sensors (Basel, Switzerland)*, vol. 15, no. 8, pp. 18901–18933, 2015.
- [3] S. S. Saini, A. Sridhar, and K. Ahluwalia, "Smartphone optical sensors," *Opt. Photonics News*, vol. 30, no. 2, pp. 34–41, 2019.
- [4] K. C. Ginny and K. Naik, "Smartphone processor architecture, operations, and functions: current state-of-the-art and future outlook: energy performance trade-off," *J. Supercomput.*, vol. 77, pp. 1377–1454, 2021.
- [5] P. M. Singh and M. Kumar, "Evolution of processor architecture in mobile phones," *Int. J. Comput. Appl.*, vol. 90, 2014, <https://doi.org/10.5120/15564-4339>.
- [6] M. E. Khaddar and M. Boulmalf, "Smartphone: the ultimate IoT and IoE device", in *Smartphones from an Applied Research Perspective, Chapter 7*, N. Mohamudally ed. 2017, <https://doi.org/10.5772/intechopen.69734>.
- [7] M. Rather and S. Rather, "Impact of smartphones on young generation," *Libr. Philos. Pract.*, vol. 2384, 2019.
- [8] M. Sarwar and T. R. Soomro, "Impact of smartphone's on society," *Eur. J. Sci. Res.*, vol. 98, pp. 216–226, 2013.
- [9] M. Moses and J. Wade, *Spycamera – The Minox Story*, 2nd ed, Denver CO, Hove Foto Books, 1998.
- [10] DXO, *Disruptive Technologies: Mobile Imaging taking Smartphone Cameras to Next Level*, 2018a. Available at: <https://www.dxomark.com/disruptive-technologies-mobile-imaging-taking-smartphone-cameras-next-level/>.
- [11] DXO, *Smartphones versus Cameras: Closing the Gap on Image Quality*, 2020. Available at: <https://www.dxomark.com/smartphones-vs-cameras-closing-the-gap-on-image-quality/>.
- [12] Counterpointresearch, 2021. Available at: <https://www.counterpointresearch.com/global-smartphone-share/>.
- [13] Gartner Newsroom. Available at: <https://www.gartner.com/>.
- [14] P. Cambou and J.-L. Jaffart, *Status of the Camera Module Industry 2019*, Yole Développement, 2019. Available at: <https://www.systemplus.fr/wp-content/uploads/2019/04/>

- YD19006_Status_Camera_Module_Industry_2019_WLO_Yole_Développement_Sample.pdf.
- [15] P. Cambou, *Smartphone Camera Trends, Yole Développement*, 2020. Available at: <https://www.youtube.com/watch?v=fR4KXZ1zjw0>.
 - [16] D. Yang, S. Wegner, and A. Cowsky, *Apple iphone 11 Pro Max Teardown*, 2019. Available at: <https://www.techinsights.com/blog/apple-iphone-11-pro-max-teardown>.
 - [17] Yole, *Rear ToF to Become Main Smartphone Camera Growth Engine*, 2019. Available at: <http://image-sensors-world.blogspot.com/2020/02/yole-rear-tof-to-become-main-smartphone.html>.
 - [18] R. L. Baer, “Resolution limits in digital photography: the looming end of the pixel wars,” in *Imaging Systems, OSA Technical Digest (CD) (Optical Society of America, 2010)*, paper ITuB3, 2010.
 - [19] G. Westheimer, “Visual acuity and hyperacuity,” in *Handbook of Optics, Volume III Vision and Vision Optics*, M. Bass, Ed., New York, McGraw-Hill, 2010.
 - [20] S. Triantaphyllidou, J. Smejkal, E. Fry, and C. Hsin Hung, “Studies on the effect of megapixel sensor resolution on displayed image quality and relevant metrics,” in *Proceedings of the IS&T International Symposium on Electronic Imaging*, 2020, p. 170. Available at: <https://t1p.de/MegaPixelEffect>.
 - [21] R. D. Fiete, *Modeling the Imaging Chain of Digital Cameras*, Bellingham, SPIE Press, Tutorial Text, 2010.
 - [22] J. W. Goodman, *Introduction to Fourier Optics*, Greenwood Village, CO, Roberts & Co., 2005.
 - [23] C.-L. Tisse, F. Guichard, and F. Cao, “Does resolution really increase image quality?,” *Proc. SPIE 6817, Digital Photography*, vol. IV, 2008, Art no. 68170Q.
 - [24] M. Berek, *Grundlagen der praktischen Optik*, Berlin, de Gruyter, 1930.
 - [25] S. Ray, *Applied Photographic Optics*, 3rd ed. Waltham, Focal Press, 2002.
 - [26] D. A. Rowlands, “Physics of digital photography,” in *IOP Series in Emerging Technologies in Optics and Photonics*, R.B. Johnson, Ed., IOP Publishing Limited, 2020.
 - [27] U. Teubner and H. J. Brückner, *Optical Imaging and Photography*, Berlin/Boston, De Gruyter Verlag, 2019.
 - [28] A. W. Lohmann, “Scaling laws for lens systems,” *Appl. Opt.*, vol. 28, no. 23, pp. 4996–4998, 1989.
 - [29] D. J. Brady and N. Hagen, “Multiscale lens design,” *Opt. Express*, no. 13, pp. 10659–10674, 2009. <https://doi.org/10.1364/oe.17.010659>.
 - [30] M. W. Haney, “Performance scaling in flat imagers,” *Appl. Optics*, vol. 45, no. 13, p. 2901, 2006.
 - [31] A. Brückner, “Multiaperture cameras, chapter 8,” in *Smart Mini-Cameras*, T.V. Galstian, Ed., Boca Raton, FL, CRC Press, 2014.
 - [32] G. Carles, G. Muyo, N. Bustin, A. Wood, and A. R. Harvey, “Compact multi-aperture imaging with high angular resolution,” *J. Opt. Soc. Am. A*, vol. 32, pp. 411–419, 2015.
 - [33] ZEISS Lens Data Sheet Otus 1.4/28, 2021. Available at: <https://www.zeiss.com/content/dam/consumer-products/downloads/photography/datasheets/en/otus-lenses/datasheet-zeiss-otus-1428.pdf>.
 - [34] M. Born and E. Wolf, *Principles of Optics*, Cambridge, Cambridge University Press, 2019.
 - [35] T. Steinich, and V. Blahnik, *Optical Design of Camera Optics for Mobile Phones*, 2012. Available at: <http://www.degruyter.com/view/j/aot.2012.1.issue-1-2/aot-2012-0002/aot-2012-0002.xml>.
 - [36] P. L. Ruben, “Design and use of mass-produced aspheres at Kodak,” *Appl. Opt.*, vol. 24, no. 11, 1985. <https://doi.org/10.1364/ao.24.001682>.
 - [37] W. T. Plummer, “Unusual optics of the Polaroid SX-70 land camera,” *Appl. Opt.*, vol. 21, pp. 196–202, 1982.
 - [38] R. Altman and J. Lytle, “Optical design techniques for polymer optics,” *Proc. SPIE*, vol. 237, p. 380, 1980.
 - [39] L. Nie, “Patent review of miniature camera lenses and a brief comparison of two relative design patterns,” Master Thesis, College of Optical Sciences, University of Arizona, 2017.
 - [40] W.-Y. Chen, “Optical lens system, image capturing unit and electronic device,” US 2020/0393653 A1, 2020.
 - [41] J. Bureau and P. Clark, “The optics of miniature digital camera modules,” in *International Optical Design, Technical Digest (CD) (Optical Society of America, 2006)*, paper WB3, 2006.
 - [42] R. Bates, “The modern miniature camera objective: an evolutionary path from the landscape lens,” *Adv. Opt. Technol.*, vol. 2, no. 1, pp. 13–20, 2013.
 - [43] P. Clark, “Mobile platform optical design,” *Proc. SPIE, International Optical Design Conference*, vol. 9293, p. 92931M, 2014.
 - [44] P. Clark, “Lens design and advanced function for module cameras, chapter 1,” in *Smart Mini-Cameras*, T.V. Galstian, Ed., Boca Raton, FL, CRC Press, 2014.
 - [45] J. Sasian, *Introduction to Lens Design*, Cambridge, Cambridge University Press, 2019.
 - [46] D. Shafer, *A New Theory of Cell Phone Lenses*, 2019. Available at: <https://de.slideshare.net/operacrazy/a-new-theory-of-cell-phone-lenses>.
 - [47] H. Gross, F. Blechinger, and B. Achtner, “Photographic lenses,” in *Handbook of Optical Systems*, vol. 4, Survey of Optical Instruments 4, Weinheim, Wiley-VCH, 2008.
 - [48] R. Kingslake, *Lens Design Fundamentals*, Cambridge, Academic Press, 1978.
 - [49] W. Smith, *Modern Lens Design*, New York, McGraw-Hill, 2004.
 - [50] I. Stamenov, I. P. Agurok, and J. E. Ford, “Optimization of two-glass monocentric lenses for compact panoramic imagers: general aberration analysis and specific designs,” *Appl. Opt.*, vol. 51, pp. 7648–7661, 2012.
 - [51] D. Reshidko, and J. Sasian, “Optical analysis of miniature lenses with curved imaging surfaces,” *Appl. Opt.*, vol. 54, no. 28, pp. E216–E223, 2015.
 - [52] L. Bertele, “Five component wide-angle objective,” US2721499A, 1951.
 - [53] H. H. Nasse, *Distagon, Biogon und Hologon. ZEISS Camera Lenses Technical article*, 2011. Available at: https://lenspire.zeiss.com/photo/app/uploads/2018/02/en_CLB41_Nasse_LensNames_Distagon.pdf.
 - [54] L. Seidel, “Ueber die Theorie der Fehler, mit welchen die durch optische Instrumente gesehnen Bilder, behaftet sind, und über die mathematischen Bedingungen ihrer Aufhebung,” in *Abhandlungen der naturwissenschaftlich-technischen Commission bei der Königl. München, Bayerischen Akademie der Wissenschaften in München*, 1857, pp. 227–267.
 - [55] G. Forbes, “Shape specification for axially symmetric optical surfaces,” *Opt. Express*, vol. 15, pp. 5218–5226, 2007.
 - [56] G. W. Forbes and C. Menke, “Optical design with orthogonal surface descriptions,” *Proc. SPIE*, vol. 8884, p. 88841C, 2013.
 - [57] B. Ma, K. Sharma, K. P. Thompson, and J. P. Rolland, “Mobile device camera design with Q-type polynomials to achieve

- higher production yield,” *Opt. Express*, vol. 21, pp. 17454–17463, 2013.
- [58] Z. Zhuang, X. Dallaire, J. Parent, P. Roulet, and S. Thibault, “Geometrical-based quasi-aspheric surface description and design method for miniature, low-distortion, wide-angle camera lens,” *Appl. Opt.*, vol. 59, pp. 8408–8417, 2020.
- [59] H. J. Brückner, V. Blahnik, and U. Teubner, “*Maximale Bildqualität aus Minikameras: Smartphonekameras, Teil 1: Kameraoptik*,” *Physik in unserer Zeit*, vol. 51, no. 5, p. 236, 2020.
- [60] C.-C. Lin, C.-Y. Chen, Y.-T. Tseng, K.-T. Yeh, and H.-H. Huang, “Optical Imaging module, image capturing apparatus and electronic device,” US20190250380, 2019.
- [61] M. Dror, E. Goldenberg, and G. Shabtay, “Miniature telephoto lens assembly,” US 9568712 B2, 2017.
- [62] Y.-T. Tseng, T.-Y. Hsieh, and H.-H. Huang, “Photographing lens assembly, image capturing unit and electronic device,” US20190094500A1, 2019.
- [63] B. Hönliger and H. H. Nasse, *Distortion. ZEISS camera lenses technical article*, 2009. Available at: <https://lenspire.zeiss.com/photo/app/uploads/2018/04/Article-Distortion-2009-EN.pdf>.
- [64] V. Blahnik, D. Gaengler, and J.-M. Kaltenbach, “Evaluation and analysis of chromatic aberrations in images,” *Proc. SPIE, Optical Design and Engineering IV*, vol. 8167, p. 81670G, 2011.
- [65] P.-L. Hsu, and S.-Y. Yang, “Optical photographing lens assembly, imaging apparatus and electronic device,” US 20190271831A1, 2019.
- [66] L.-M. Chen and H. H. Huang, “Imaging optical lens assembly, imaging apparatus and electronic device,” US2019/0170970 A1, 2019.
- [67] C.-Y. Liao and W.-Y. Chen, “Imaging lens assembly, image capturing unit and electronic device,” US20210048630A1, 2021.
- [68] P.-L. Hsu and W.-Y. Chen, “Imaging lens assembly, image capturing unit and electronic device,” US20210018725, 2021.
- [69] D. H. Jeong, H. S. Yoo, J. I. Lee, and J. K. Kim, “Portable electronic device, optical imaging system, and lens assembly,” US20200409065A1, 2020.
- [70] Y. Yao, and Y. Shinohara, *Folded Lens System*, Patent US 2019/0196148 A1, 2019.
- [71] K. Araki, T. Tanaka, M. Sekita, K. Kimura, N. Nanba, H. Saruwatari, and T. Akiyama, “Optical apparatus,” US6426841B1, 2002.
- [72] G. Carles and A. R. Harvey, “Multi-aperture imaging for flat cameras,” *Opt. Lett.*, vol. 45, pp. 6182–6185, 2020.
- [73] F. Dai, J. Wenren, and J. Yang, “Optical imaging system,” US 20190187446, 2019.
- [74] L.-M. Chen, and H. H. Huang, “Optical photographing system and electronic device,” US10877244B1, 2020.
- [75] E. Tremblay, R. A. Stack, R. L. Morrison, and E. F. Ford, “Ultrathin cameras using annular folded optics,” *Appl. Opt.*, vol. 46, no. 4, p. 463, 2007.
- [76] E. Tremblay, R. A. Stack, R. L. Morrison, J. H. Karp, and E. F. Ford, “Ultrathin four reflection imager,” *Appl. Opt.*, vol. 48, no. 2, p. 343, 2009.
- [77] I. Yedid, *The Evolution of Zoom Camera Technologies in Smartphones, Corephotonics White Paper*, 2017. Available at: https://corephotonics.com/wp-content/uploads/2017/09/Corephotonics-White-Paper_The-Evolution-of-Zoom-Camera-Technologies-in-Smartphones_Aug-2017.pdf.
- [78] G. Shabtay, E. Goldenberg, O. Gigushinski, N. Cohen, Dual aperture zoom digital camera. US 9.185.291 B1, 2015.
- [79] M. Kreitzer and J. Moskovich, “Optical design of a smartphone zoom lens,” *Proc. SPIE, Zoom Lenses VI*, vol. 11106, p. 111060D, 2019.
- [80] T. Kato, A. Oohata, and H. Hagiwara, “Zoom lens and imaging apparatus,” US9235036B2, 2014.
- [81] H. Yamanaka, M. Kanai, M. Sueyoshi, and M. Hosoi, “Enhanced variable power zoom lens,” US 8.451 549, 2013.
- [82] K. Li and Kato, “Zoom lens,” US 9 176 308 B2, 2014.
- [83] M. P. Schaub, *The Design of Plastic Optical Systems*, Bellingham, SPIE Press, 2009.
- [84] Y. Yu and G. Berger, *Fast and Highly Accurate Metrology Tool for Smart Phone Lens Mold Characterization*, 2016. Available at: https://www.taylor-hobson.com/de/-/media/ametektaylorhobson/files/learning-zone/application-notes/a150—smart-phone-lens-mold_highres_en.pdf.
- [85] J. Y. Song, T. H. Ha, C. W. Lee, J. H. Lee, and D. H. Kim, “Development of minimized-assembly system for camera phone lens module,” in *Proceedings of the 11th euspen International Conference – Como*, 2011. Available at: <https://www.euspen.eu/knowledge-base/ICE11252.pdf>.
- [86] C. F. Lin, C.-H. Tsai, and M.-T. Chou, “Imaging lens module and mobile terminal,” US 2016/0231526 A1, 2016.
- [87] S. Bäumer, *Handbook of Plastic Optics*, Weinheim, Wiley-VCH Verlag, 2005.
- [88] H. Gross, “Fundamentals of technical optics,” in *Handbook of Optical Systems*, vol. 1, Weinheim, Wiley-VCH, 2005.
- [89] K. Weber, D. Werdehausen, P. König, S. Thiele, M. Schmid, M. Decker, P. W. De Oliveira, A. Herkommer, and H. Giessen, “Tailored nanocomposites for 3D printed micro-optics,” *Opt. Mater. Express*, vol. 10, pp. 2345–2355, 2020.
- [90] D. Werdehausen, *Nanocomposites as Next-Generation Optical Materials. Fundamentals, Design and Advanced Applications*, Berlin, Springer Series in Materials Science, 2021.
- [91] K. Straw, “Control of thermal focus shift in plastic-glass lenses,” *Proc. SPIE*, vol. 237, p. 386, 1980.
- [92] K. Bitzer and A. By, “Active alignment for cameras in mobile devices and automotive applications,” in *IEEE Electronics Packaging Technology Conference, EPTC*, 2010, pp. 260–264.
- [93] D. Wilson, *Automating Optical Alignment of Camera Modules*, 2017. Available at: http://www.novuslight.com/automatingoptical-alignment-of-camera-modules_N3530.html.
- [94] Imatest, *Imatest Documentation by Norman Koren*, 2009. Available at: <https://www.imatest.com/docs/Imatest%20Documentation.pdf>.
- [95] D. Winters, “Image quality testing improves as cameras advance,” *Photonics Spectra*, vol. 48, no. 1, pp. 66–70, 2014, Available at: https://trioptics.com/wp-content/uploads/2019/08/Articles_Image_Quality_Testing_improves_as_Camera_Advance_Photronics_Spectra_0114-2.pdf.
- [96] R. Bates, “Performance and tolerance sensitivity optimization of highly aspheric miniature camera lenses. Optical system alignment, tolerancing, and verification IV,” *Proc. SPIE*, vol. 7793, pp. 779302, 2010.
- [97] S. Jung, D.-H. Choi, B.-L. Choi, and J. H. Kim, “Tolerance optimization of a mobile phone camera lens system,” *Appl. Opt.*, vol. 50, no. 23, 2011, <https://doi.org/10.1364/AO.50.004688>.

- [98] F. E. Sahin, "Lens design for active alignment of mobile phone cameras," *Opt. Eng.*, vol. 56, no. 6, 2017, Art no. 065102.
- [99] L. Carrión-Higueras, A. Calatayud, and J. Sasian, "Improving as-built miniature lenses that use many aspheric surface coefficients with two desensitizing techniques," *Opt. Eng.*, vol. 60, no. 5, 2021. <https://doi.org/10.1117/1.oe.60.051208>.
- [100] J. Joo, "Design of a miniaturized imaging system for as-built performance," *Graduate Theses-Phys. Opt. Eng.*, vol. 8, 2020. https://scholar.rose-hulman.edu/dept_optics/8.
- [101] J. P. McGuire, and T. G. Kuper, "Approaching direct optimization of as-built lens performance," *Proc. SPIE, Novel Optical Systems Design and Optimization XV*, vol. 8487, p. 84870D, 2012.
- [102] T. Hayes, "Next-generation cell phone cameras," *Opt. Photon. News*, vol. 23, no. 2, pp. 16–21, 2012.
- [103] E. Wolterink and K. Demeyer, "WaferOptics(R) mass volume production and reliability," *Proc. SPIE*, vol. 7716, p. 771614, 2010.
- [104] M. Zoberbier, S. Hansen, M. Hennemeyer, D. Tönnies, R. Zoberbier, M. Brehm, A. Kraft, M. Eisner, and R. Voelkel, *Wafer Level Cameras – Novel Fabrication and Packaging Technologies*, 2009. Available at: https://www.researchgate.net/publication/266046128_Wafer_level_cameras_-_novel_fabrication_and_packaging_technologies.
- [105] J. Lapointe, M. Gagné, M.-J. Li, and R. Kashyap, "Making smart phones smarter with photonics," *Opt. Express*, vol. 22, pp. 15473–15483, 2014.
- [106] A. Smakula, "Verfahren zur Erhöhung der Lichtdurchlässigkeit optischer Teile durch Erniedrigung des Brechungsexponenten an den Grenzflächen dieser Teile," *Deutsches Reichspatent*, vol. 685, p. 767, 1935.
- [107] L. Martinu and J. E. Klemburg-Sapieha, "Optical coatings on plastics," in *Optical Interference Coatings*, N. Kaiser, and H. Pulker, Eds., Berlin/Heidelberg, Springer-Verlag, 2003.
- [108] U. Schulz, U. Schallenberg, and N. Kaiser, "Antireflection coating design for plastic optics," *Appl. Opt.*, 2002. <https://doi.org/10.1364/AO.41.003107>.
- [109] U. Schulz, "Review of modern techniques to generate antireflective properties on thermoplastic polymers," *Appl. Opt.*, vol. 45, pp. 1608–1618, 2006.
- [110] P. Paul, K. Pfeiffer, and A. Szeghalmi, "Antireflection coating on PMMA substrates by atomic layer deposition," *Coatings*, 2020. <https://doi.org/10.64.10.3390/coatings10010064>.
- [111] W. S. Boyle and G. E. Smith, "Charge coupled semiconductor devices," *The Bell Systems Technical Journal*, vol. 49, pp. 587–593, 1970.
- [112] K. Matsumoto, T. Nakamura, A. Yusa and S. Nagai, "A new MOS phototransistor operating in a non-destructive readout mode," *Jpn. J. Appl. Phys.*, vol. 24, no. 5A, p. L323, 1985.
- [113] E. R. Fossum, "CMOS image sensors: electronic camera-on-a-chip," *IEEE Trans. Electron. Dev.*, vol. 44, no. 10, pp. 1689–1698, 1997.
- [114] E. R. Fossum, "Active pixel sensors—are CCDs dinosaurs?," *CCD's and Optical Sensors III, Proc. SPIE*, vol. 1900, pp. 2–14, 1993.
- [115] S. Mendis, S. E. Kemeny, R. Gee, B. Pain, and E. R. Fossum, "Progress in CMOS active pixel image sensors," *Charge-Coupled Devices and Solid State Optical Sensors IV, Proc. SPIE*, vol. 2172, pp. 19–29, 1994.
- [116] S. Mendis and E. R. Fossum, "CMOS active pixel image sensor," *IEEE Trans. Electron Devices*, vol. 41, no. 3, pp. 452–453, 1994.
- [117] J. Nakamura, *Image Sensors and Signal Processing for Digital Still Cameras*, Boca Raton, CRC-Press, 2005.
- [118] I. Takayanagi, "CMOS image sensors," in *Image Sensors and Signal Processing for Digital Still Cameras*, J. Nakamura, Ed., Boca Raton, FL, CRC Press, 2006, pp. 143–177.
- [119] R. Fontaine, "The state-of-the-art of mainstream CMOS image sensors," in *International Image Sensor Workshop*, 2015.
- [120] S.-Y. Chen, C.-C. Chuang, J.-C. Liu, and D.-N. Yaung, "Image sensor with deep trench isolation structure," US2012/0025199 A1, 2012.
- [121] Y. Qian, H.-C. Tai, D. Mao, V. Venezia, and H. E. Rhodes, "Image sensor having dark sidewalls between color filters to reduce optical crosstalk," US 2012/0019695 A1, 2012, .
- [122] M. Lapedus, *Scaling CMOS Image Sensors*, 2020. Available at: <https://semiengineering.com/scaling-cmos-image-sensors/>.
- [123] H. Yamanaka, "Method and apparatus for producing ultra-thin semiconductor chip and method and apparatus for producing ultra-thin back-illuminated solid-state image pickup device," US7521335 B2, 2009.
- [124] B. E. Bayer, "Color imaging array," U.S. Patent 3 971 065, 1976.
- [125] H. Malvar, L.-W. He, and R. Cutler, "High-quality linear interpolation for demosaicing of Bayer-patterned color images. Acoustics, speech, and signal processing, 1988," in *ICASSP-88., 1988 International Conference on. 3. iii-485*, 2004.
- [126] C.-Y. Lee, S.-W. Hyun, Y.-J. Kim, and S.-W. Kim, "Optical inspection of smartphone camera modules by near-infrared low-coherence interferometry," *Opt. Eng.*, vol. 55, no. 9, 2016. <https://doi.org/10.1117/1.OE.55.9.091404>.
- [127] ISO 14524, *Photography, Electronic Still-picture Cameras, Methods for Measuring Opto-Electronic Conversion Functions (OECFs)*, Geneva, Switzerland, Vernier, 2009.
- [128] H. J. Brückner, V. Blahnik, and U. Teubner, "Tricksen für gute Bilder: Smartphonekameras, Teil 2: Bildsensor und -verarbeitung," *Physik in unserer Zeit*, vol. 51, no. 5, p. 290, 2020b.
- [129] VDÄPC, *Pressemitteilung der Vereinigung der deutschen ästhetisch-plastischen Chirurgen 2019*, 2019. Available at: <https://www.vdaepc.de/pressemitteilung-statistik-2019-trends-der-aesthetisch-plastischen-chirurgie/> [Press release of the Association of German Aesthetic Plastic Surgeons 2019].
- [130] J. N. Mait, G. W. Euliss, and R. A. Athale, "Computational imaging," *Adv. Opt. Photon.*, vol. 10, pp. 409–483, 2018.
- [131] B. C. Kress, "Digital optical elements and technologies (EDO19): applications to AR/VR/MR," *Proc. SPIE, Digital Optical Technologies*, vol. 11062, p. 1106222, 2019.
- [132] G. T. Fechner, *Elemente der Psychophysik*, vol. 2, Leipzig, Breitkopf und Haertel, 1860.
- [133] E. H. Weber, "Der Tastsinn und das Gemeingefühl," in *Handwörterbuch der Physiologie*, vol. 3, Braunschweig, 1850.
- [134] R. Portugal and B. Svaiter, "Weber-fechner law and the optimality of the logarithmic scale," *Minds and Machines*, vol. 21, pp. 73–81, 2011.
- [135] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Upper Saddle River, NJ, Prentice Hall, 2008.
- [136] P. Robisson, J.-B. Jourdain, W. Hauser, C. Viard, and F. Guichard, *Autofocus Measurement for Imaging Devices*, IS & T, 2017.
- [137] Shimoda, K. and Fujii, S.. "Imaging device and imaging apparatus," US 8,817,166 B2, 2014.

- [138] B.-S. Choi, J. Lee, S.-H. Kim, S. Chang, J. Park, S.-J. Lee, and J.-K. Shin, "Analysis of disparity information for depth extraction using CMOS image sensor with offset pixel aperture technique," *Sensors*, vol. 19, p. 472, 2019.
- [139] A. Morimitsu, I. Hirota, S. Yokogawa, I. Ohdaira, M. Matsumura, H. Takahashi, T. Yamazaki, H. Oyaizu, Y. Incesu, M. Atif, and Y. Nitta, *A 4M pixel full-PDAF CMOS image sensor with 1.58 μm 2×1 on-chip micro-split-lens technology*, 2015. Available at: Imagesensors.org.
- [140] M. Kobayashi, M. Ohmura, H. Takashi, T. Shirai, K. Sakurai, T. Ichikawa, H. Yuzurihara, and S. Inoue, "High-definition and high-sensitivity CMOS image sensor with all-pixel image plane phase-difference detection autofocus," *Jpn. J. Appl. Phys.*, vol. 57, no. 10, 2018. <https://doi.org/10.7567/jjap.57.1002b5>.
- [141] G. Chataignier, B. Vandame, and J. Vaillant, "Joint electromagnetic and ray-tracing simulations for quad-pixel sensor and computational imaging," *Opt. Express*, vol. 27, pp. 30486–30501, 2019.
- [142] L. W. Alvarez and W. E. Humphrey, "Variable power lens and system," US Patent 3 507 565, 1970.
- [143] A. W. Lohmann, "A new class of varifocal lenses," *Appl. Opt.*, vol. 9, no. 7, pp. 1669–1671, 1970.
- [144] G. Zhou, H. Yu, and F. S. Chau, "Microelectromechanically-driven miniature adaptive Alvarez lens," *Opt. Exp.*, vol. 21, no. 1, pp. 1226–1233, 2012.
- [145] B. Berge, "Liquid lens," in *Chapter 5 in Smart Mini-Cameras*, T. V. Galstian, Ed., Boca Raton, FL, CRC Press, 2014.
- [146] T. V. Galstian, "Electrically variable liquid crystal lenses, chapter 6," in *Smart Mini-Cameras*, T.V. Galstian, Ed., Boca Raton, FL, CRC Press, 2014.
- [147] T. V. Galstian, O. Sova, K. Asatryan, V. Presniakov, A. Zohrabyan, and M. Evensen, "Optical camera with liquid crystal autofocus lens," *Opt. Express*, vol. 25, pp. 29945–29964, 2017.
- [148] V. Blahnik, *About the Irradiance and Apertures of Camera Lenses*, 2014. Available at: http://www.zeiss.com/content/dam/Photography/new/pdf/en/cln_archiv/cln_en_web_special_irradiance_and_apertures.pdf.
- [149] C.-L. Hsieh, C.-S. Liu, and C.-C. Cheng, *Design of a 5 Degree of Freedom-Voice Coil Motor Actuator for Smartphone Camera Modules, Sensors and Actuators: Physical*, New York, Elsevier, 2020.
- [150] L. Lu, "Voice coil motors for mobile applications. Chapter 3," in *Smart Mini-Cameras*, T.V. Galstian, Ed., Boca Raton, FL, CRC Press, 2014.
- [151] H. Wu, C. Chou, and S. Shyiu, "Voice coil driving auto-focus lens module," US20097612957, 2009.
- [152] Y. Chou, "Camera module with piezoelectric actuator," US20100271541Al, 2009.
- [153] M. Murphy, M. Conway, and G. Casey, "Lens drivers focus on performance in high-resolution camera modules," *Analog Dialogue*, pp. 40–11, 2006.
- [154] C.-L. Hsieh and C.-S. Liu (2020). Design of a voice coil motor actuator with L-shape coils for optical zooming smartphone cameras. *IEEE Access*, vol. 8, pp. 20884–20891.
- [155] A. A. N. Galaom, "Integration of a MEMS-based autofocus actuator into a smartphone camera," Master thesis, Mechanical and Industrial Engineering University of Toronto, 2016. Available at: https://tspace.library.utoronto.ca/bitstream/1807/92573/3/Galaom_Ahmed_201611_MAS_thesis.pdf.
- [156] B. T. Faez and B. M. Ridha, "Autofocus camera using MEMS actuation of image sensor," WO2017106955, 2017.
- [157] H.-Y. Sung, P.-C. Chen, C.-C. Chang, C.-W. Chang, S. Yang, and H. Chang, *Mobile Phone Imaging Module with Extended Depth of Focus Based on Axial Irradiance Equalization Phase Coding*, 2011.
- [158] E. R. Dowski and W. T. Cathey, "Extended depth of field through wave-front coding," *Appl. Opt.*, vol. 34, pp. 1859–1866, 1995.
- [159] B. Golik and D. Wueller, "Measurement method for image stabilizing systems," *Proc. SPIE, Digital Photography III*, vol. 6502, p. 650200, 2007.
- [160] O. Šindelář and F. Šroubek, "Image deblurring in smartphone devices using built-in inertial measurement sensors," *J. Electron. Imag.*, vol. 22, no. 1, 2013, Art no. 011003.
- [161] K. T. Wyne, "A comprehensive review of tremor," *JAAPA*, vol. 18, no. 12, pp. 43–50, 2005, quiz 57–8.
- [162] E. Simon, "Optical image stabilization for miniature cameras, chapter 7," in *Smart Mini-Cameras*, T.V. Galstian, Ed., Boca Raton, FL, CRC Press, 2014.
- [163] C. Acar and A. Shkel, *MEMS Vibratory Gyroscopes*. Berlin, Springer, 2009.
- [164] S. Sinha, S. Shakya, R. Mukhiya, R. Gopal, and B. Pant, *Design and Simulation of MEMS Differential Capacitive Accelerometer*, 2014, <https://doi.org/10.13140/2.1.1074.8809>.
- [165] F. Xiao, A. Silverstein, and J. Farrel, "Camera-motion and effective spatial resolution," *Proc. SPIE*, vol. 7241, pp. 33–36, 2006.
- [166] F. Xiao, J. E. Farrell, P. B. Catrysse, and B. Wandell, "Mobile imaging: the big challenge of the small pixel," *Proc. SPIE, Digital Photography V*, vol. 7250, p. 72500K, 2009.
- [167] C.-W. Chiu, P. C. P. Chao, N.-Y. Y. Kao, and F.-K. Young, "Optimal design and experimental verification of magnetically actuated optical image stabilization system for cameras in mobile phones," *J. Appl. Phys.*, vol. 103, 2008, Art no. 07F136.
- [168] F. La Rosa, M. C. Virzì, F. Bonaccorso, and M. Branciforte, *Optical Image Stabilization*, ST Microelectronics, White Paper, 2017. Available at: https://www.st.com/content/ccc/resource/technical/document/white_paper/c9/a6/fd/e4/e6/4e/48/60/ois_white_paper.pdf/files/ois_white_paper.pdf/jcr:content/translations/en.ois_white_paper.pdf.
- [169] V. Blahnik and B. Voelker, *About the Reduction of Reflections in Camera Lenses. How T*-coating Made Glass Invisible*, 2016. Available at: https://pixinfo.com/wp-content/uploads/2016/04/en_About-the-reduction-of-reflections-of-camera-lenses.pdf.
- [170] M. Steinbach, "Development of a method to measure the Veiling Glare contributions of the lens and the sensor and its influence on the limit of the dynamic range when shooting movies," Bachelor Thesis, Cologne University of Applied Sciences, 2015.

- [171] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *Siggraph Conference, July 1997*, 1997.
- [172] S. Mann and R. W. Picard, "On being 'undigital' with digital cameras: extending dynamic range by combining differently exposed pictures," *Proc. IST*, pp. 422–428, 1995.
- [173] A. Adams, *The Print. New Edition of 1950 from Little, Brown and Company*. New York, Boston, Little, Brown and Company, 1980.
- [174] A. Darmont, *High Dynamic Range Imaging – Sensors and Architectures*, Bellingham, Washington USA, SPIE Press, 2012.
- [175] J. J. McCann and A. Rizzi, *The Art and Science of HDR Imaging*, New York, Wiley, 2012.
- [176] E. Reinhard, E. Ward, S. Pattanaik, P. Debevec, W. Heidrich, and K. Myszkowski, *High Dynamic Range Imaging*, New York, Elsevier, Morgan Kaufmann, 2010.
- [177] R. H. Abd El-Maksoud, and J. Sasian, "Modeling and analyzing ghost images for incoherent optical systems," *Appl. Opt.*, vol. 50, no. 15, pp. 2305–2315, 2011.
- [178] E. Fest, *Straylight Analysis and Control*, Bellingham, SPIE Press, 2013.
- [179] J. Jur, *Straylight Analysis of a Mobile Phone Camera*, 2016. Available at: <https://www.slideshare.net/JordanJur/straylight-analysis-of-a-mobile-phone-camerarevd>.
- [180] K. Venkataraman, D. Lelescu, J. Duparré, and A. McMahon, "PiCam: an ultra-thin high performance monolithic camera array," in *Conference: ACM Transactions on Graphics, Proceedings of SIGGRAPH Asia*, vol. 32, 2014.
- [181] Y. Furukawa and C. Hernández, "Multi-view stereo: a tutorial," *Foundations and Trends in Computer Graphics and Vision*, vol. 9, 2015. <https://doi.org/10.1561/0600000052>.
- [182] B. Krishnamurthy and A. Rastogi, "Refinement of depth maps by fusion of multiple estimates," *J. Electron. Imag.*, vol. 22, no. 1, 2013, Art no. 011002.
- [183] R. Szeliski, *Computer Vision: Algorithms and Applications*, Berlin, Springer-Verlag, 2010.
- [184] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan, "Light field photography with a hand-held plenoptic camera," Technical Report CTSR 2005-02, CTSR, 2005.
- [185] C. Perwass and L. Wietzke, "Single lens 3D-camera with extended depth-of-field," *Proc. SPIE*, vol. 8291, 2012. <https://doi.org/10.1117/12.909882>.
- [186] M. Hansard, S. Lee, O. Choi, and R. Horaud, *Time-of-Flight Cameras: Principles, Methods and Applications*, Berlin, Springer, 2012.
- [187] K. Iga, "VCSEL: born small and grown big," *Proc. SPIE, Vertical External Cavity Surface Emitting Lasers (VECSELs) X*, vol. 11263, p. 1126302, 2020.
- [188] S. Foix, G. Alenyà, and C. Torras, "Lock-in Time-of-Flight (ToF) cameras: a survey," *IEEE Sensors Journal*, vol. 11, no. 3, 2011. <https://doi.org/10.1109/jsen.2010.2101060>.
- [189] L. Li, "Time-of-flight camera—an introduction," Technical white paper, Texas Instruments, 2014. Available at: <https://www.ti.com/lit/wp/sloa190b/sloa190b.pdf>.
- [190] H. H. Nasse, *Depth of field and bokeh. ZEISS Camera Lenses Technical Article*, 2010. Available at: <https://lenspire.zeiss.com/photo/app/uploads/2018/04/Article-Bokeh-2010-EN.pdf>.
- [191] DXO, *Test of Computational Bokeh*, 2018b. Available at: <https://www.dxomark.com/evaluating-computational-bokeh-test-smartphone-portrait-modes/>.
- [192] D. Wueller, U. Artmann, V. Rao, G. Reif, J. Kramer, and F. Knauf, "VCX: an industry initiative to create an objective camera module evaluation for mobile devices," *Electronic Imaging*, vol. 2018, no. 5, pp. 1–5, 2018.
- [193] A. Duane, "Studies in monocular and binocular accommodation with their clinical applications," *Am. J. Ophthalmol. Ser.*, vol. 35, pp. 865–877, 1922.
- [194] W. B. Wetherell, "Afocal lenses," in *Applied Optics and Optical Engineering*, vol. X, R.R. Shannon, and J.C. Wyant, Eds., Orlando, FL, Academic Press, 1987, pp. 109–192 also: Afocal systems, in *Handbook of Optics*, ed. M. Bass, Vol. 1, 2.1–2.23, McGraw-Hill, New York (1995).
- [195] V. Blahnik and O. Schindelbeck, "Optical lens system for a front lens in front of a camera module of an electrical terminal," US20180210173A1, 2018.
- [196] V. Blahnik and H. Mehnert, "Optical lens system for a wide-angle afocal optic in front of a camera module," DE102016004455A1, 2017.
- [197] V. Blahnik and D. Gängler, "Optical lens system for afocal tele-optical attachment in front of a camera module," DE102016004454A1, 2017.
- [198] V. Blahnik, N. Lippok, and T. Steinich, "Optical lens system for a macro attachment in front of a camera module," DE102016004457A1, 2017.
- [199] R. Diederich, R. Wartmann, H. Schadwinkel, and R. Heintzmann, "Using machine-learning to optimize phase contrast in a low-cost cellphone microscope," *PLoS One*, 2018. <https://doi.org/10.1371/journal.pone.0192937>.
- [200] M. B. Schäfer, D. Reichert, K. Stewart, A. Herkommer, C. Reichert, and P. Pott, "Smartphone-based low-cost microscope with monolithic focusing mechanism," *Curr. Dir. Biomed. Eng.*, vol. 4, no. 1, p. 267270, 2018.
- [201] N. A. Switz, M. V. D'Ambrosio, and D. A. Fletcher, "Low-cost mobile phone microscopy with a reversed mobile phone camera lens," *PLOS ONE*, vol. 9, no. 5, 2014, Art no. e95330.
- [202] V. Lakshminarayanan, J. Zelek, and A. McBride, "Smartphone science," *Eye Care Med. Opt. Photonics News*, vol. 26, no. 1, pp. 44–51, 2015.
- [203] S. Majumder and M. J. Deen, "Smartphone sensors for health monitoring and diagnosis," *Sensors (Basel)*, vol. 19, no. 9, p. 2164, 2019.
- [204] B. Hunt, A. J. Ruiz, and B. W. Pogue, "Smartphone-based imaging systems for medical applications: a critical review," *Journal of Biomedical Optics*, vol. 26, no. 4, 2021. <https://doi.org/10.1117/1.jbo.26.4.040902>.
- [205] ASAP, CAD-Software. Available at: <https://www.breault.com/software/about-asap>.
- [206] CIPA, *Camera and Imaging Products Association*. Available at: https://www.cipa.jp/stats/dc_e.html.
- [207] FRED, *CAD-Software*. Available at: <https://photonengr.com/fred-software/>.
- [208] Image Engineering, *Solutions to Test Image Quality*. Available at: <https://www.image-engineering.de/>.
- [209] Image Sensor World Blog. *News and Discussions About Image Sensors*, by Vladimir Koifman. Available at: [Image-sensors-blogspot.com](https://image-sensors-blogspot.com).
- [210] C. S. Liu and P. D. Lin, "A miniaturized low-power VCM actuator for auto-focusing applications," *Opt Express*, vol. 16, no. 4, p. 2533, 2008.

- [211] C. S. Liu, P. D. Lin, P. H. Lin, S. S. Ke, Y. H. Chang, and J. B. Horng, “Design and characterization of miniature auto-focusing voice coil motor actuator for cell phone camera applications,” *IEEE Trans. Magn.*, vol. 45, no. 1, p. 155, 2009.
- [212] Trioptics, *Metrology for Miniature Lenses and Cameras*. Available at: <https://trioptics.com/products/imagemaster-pro-image-quality-mtf-testing/>.
- [213] K.-P. Wilska, R. Paajanen, M. Terho, and J. Hämäläinen, “Device for personal communication, data collection and processing and circuit board,” Finnish patent FI942334A, 1994.

Bionotes



Vladan Blahnik
Carl Zeiss AG, Carl-Zeiss-Straße 22,
Oberkochen, Germany
vladan.blahnik@zeiss.com

Vladan Blahnik studied physics and wrote his PhD thesis on partially coherent optical imaging theory at the Technical University of Braunschweig (Germany) and the Optical Sciences Center in Tucson, Arizona (USA). He joined ZEISS in 2001 as Lead System Engineer and Project Leader for lithography optics, a role that saw him working on the development of projection lenses and illumination systems for lithographic scanners for microchip mass production. Examples include the NA0.93, 193nm optical system for ASML's TWINSCAN XT:1400. In 2008, he became Head of Optical Design at ZEISS'

Consumer Products division and worked on developing lenses for photography, cinematography, industrial applications and later also binoculars and spotting scopes. He worked also on developing digital-optical systems and features such as 3D acquisition & imaging systems, digital-optical co-optimization or machine vision, partly for consumer imaging devices in close cooperation with ZEISS partners like Nokia and Sony. Since 2018, he has been a Senior Optical Systems Engineer at Optical Design & Concepts for ZEISS' Corporate Research and Technology working on different topics, such as LIDAR systems for autonomous driving, high-precision interferometry, and 3D scanning systems.



Oliver Schindelbeck
Carl Zeiss AG, Carl-Zeiss-Straße 22, 73447
Oberkochen, Germany
oliver.schindelbeck@zeiss.com

Oliver Schindelbeck earned his Dipl.-Ing. (BA) in Mechanical Engineering from the Cooperative State University Heidenheim. In 1992 he started working at the Carl Zeiss Eyeglass Lens division as a production technology developer and planner. From 2002 to 2005 he assumed the role of project manager, to optimize the manufacturing processes in the central optics production facility of Carl Zeiss AG. Since 2005 he has been based at ZEISS' Consumer Products division as a project leader, quality assurance manager and innovation manager for digital consumer imaging systems. Today, as Senior Smartphone Technology Manager, he focuses on the development of mobile imaging solutions, collaborating with ZEISS' industrial partners who include Sony and vivo.