Asian Journal of Physics
Vol. 33, Nos 3 & 4 (2024) 161-166
*Available on: www.asianjournalofphysics.com*
ap

# Fundamental imaging limits of smartphone cameras

J C Dainty

*School of Natural Sciences, University of Galway, Ireland*

Dedicated to Professor Anna Consortini for her significant contributions and pioneering works in the field of atmospheric turbulence and her continuous commitment to promote optics at global level

---

I discuss here the design parameters that limit the performance of smartphone cameras, and conclude that the single most critical quantity is the entrance pupil diameter D. Some emphasis is given to the assumptions intrinsic in this result, and suggestions for future studies are made. © Anita Publications. All rights reserved.

**Doi**: 10.54955/AJP.33.3-4.2024.161-166

**Keywords**: Imaging, Cameras, Resolution.

---

## 1 Introduction

Over 5 billion still photographs are uploaded to the internet each day [1], 93% of them taken using smartphone cameras. In addition, considering only YouTube, 720,000 hours of video are uploaded per day [2], which add up to at least another 60 billion individual image frames.

The growth of consumer imaging in the past decade is linked directly to the increased availability of smartphones, although it is interesting to observe that when the iPhone was first announced in 2007, its photographic capabilities were not particularly emphasised. Today, all smartphones have at least one rear-facing camera and one front-facing camera for selfies; some smartphones have 3 rear-facing cameras of differing focal lengths and an additional front-facing camera for 3D face recognition, i.e. a total of 5 cameras. Smartphone models are typically updated yearly, and each year the camera systems have to be "improved" over the previous models. So it is reasonable to ask, can this "improvement" continue indefinitely? What governs the limits to image quality?

In this paper, I shall discuss how the laws of physics set a limit to the quality of recorded images. This does not comprehensively address the first question above, since "improvement" involves more than image quality as defined by physics, but it does help to set the boundary between the captured real information or data, and what can be inferred or (sometimes) invented by machine learning or other algorithms.

The following discussion makes a number of assumptions which are listed here:

1. It is assumed that the photon statistics follow a Poisson random process. This applies to all current consumer imaging with either smartphones or more conventional cameras recording images of scenes with natural or artificial lighting.
2. We assume that the scene being imaged is in the far-field of the camera, which means that we can employ Fraunhofer diffraction theory and the classical resolution limit imposed by that theory. It is well-known that near-field imaging can exceed this limit, but typically this requires the imaging device to be within one optical wavelength of the sample.

---

*Corresponding author*
*e mail: chris.dainty01@gmail.com* (J C Dainty)

3. We assume that the illumination of the object is not specially arranged either by geometry or to excite atomic or molecular states in the object, as in super-resolution microscopy [3]. In practice, we assume that the illumination is natural daylight or artificial light.
4. Finally, we assume that the interaction of the image intensity with the image sensor follows a geometrical model. For pixel dimensions greater than a wavelength, this is a good approximation.

Assumptions 1-3 are valid in all consumer imaging, and assumption 4 is reasonable for the current technology employed in cameras but will not be valid for future ones: this is discussed further in Section 4.

## 2 Smartphone Camera Design

Because of commercial sensitivity, it is hard to find detailed descriptions of the design of smartphone cameras. Fortunately, a complete and authoritative open-access review has been published recently by Blahnik and Schindelbeck [4], and parts of the following description are taken from that paper: the reader is strongly encouraged to read the original review.

From the engineering or technological point of view, a phone camera module has a large number of components, each of which is essential to its performance. The whole design is constrained and optimised for high-yield mass production, and is assembled to very tight mechanical tolerances. For example, a key part is the optical image stabilisation mechanism, which relies on the on-board gyroscope in the phone for control and feedback signals. Another key technological aspect is the mechanism for sensing colour, which at present consists of a Bayer filter array; other diffractive solutions are possible which do not suffer from the intrinsic light loss of absorbing filters [5]. In addition to the many key engineering parts, there is both firmware (e.g. for de-Bayering the sensor signal) and software for processing the raw image.

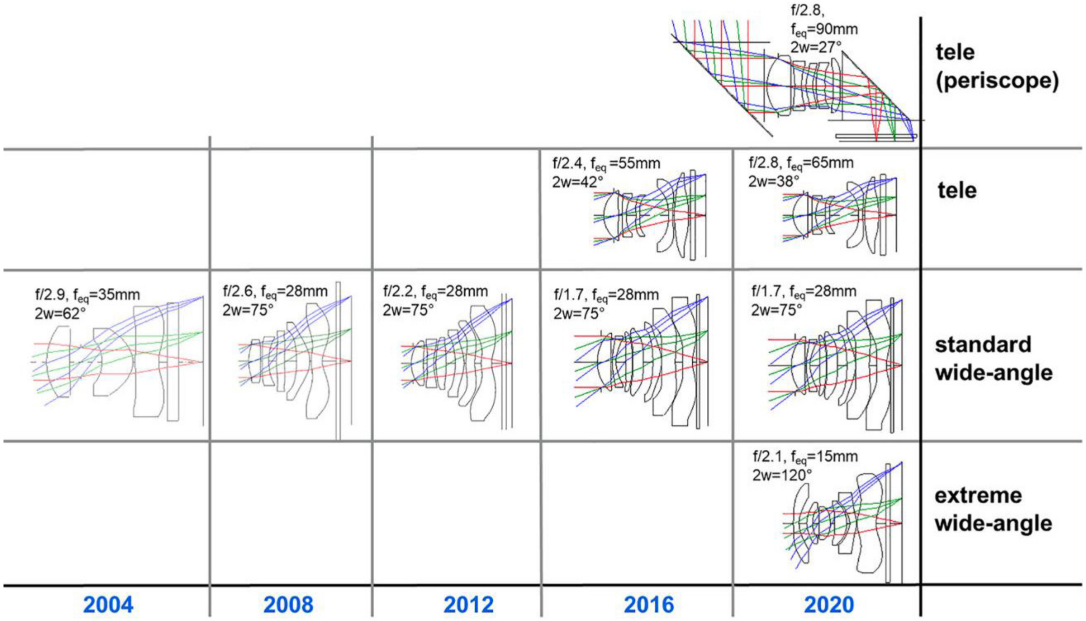


Fig 1. This shows the evolution of phone camera lenses from 2004 to 2020 (from Ref 4 with permission).

However, when considering the fundamental limits of performance, only two components matter: the lens and the detector. Figure 1 (from Ref 4) shows the evolution of smartphone camera lenses between 2004 and 2020, from 3 elements to 6 (currently 7) for the "standard" (actually wide-angle in traditional photography) lens. The entrance pupil of these lenses typically lies at the front element, and over time the F-number (F#) of the standard (wide-angle) lens has decreased from around 2.9 to 1.7, and at the time of writing, values as low as 1.5 can be found in the market.

The design and manufacture of these lenses are the result of meticulous attention to detail by optomechanical and production engineers. The on-axis MTF of standard (wide-angle) lenses approaches the diffraction limited MTF (even more so for the telephoto lenses), although off-axis performance deviates increasingly for larger field angles, particularly for the super-wide angle lenses.

Much emphasis is given in the marketing of these lenses to the F-number, but it alone does not represent a fundamental limit to the whole system, either as regards resolution or noise. Rather, the F-number together with any aberrations determines the extent of the point spread function in image space, and thus the interface to the pixelated detector. As a first approximation, it is not unreasonable to assume that with future advances, diffraction-limited performance can be achieved, at least for smaller field angles (of course, this ignores issues of cost). In this circumstance, the F-number directly influences the pixel sampling through the sampling theorem. For example, an F/1.8 lens at a mean wavelength of 0.55 microns has a diffraction limited cut-off spatial frequency of 1000 mm$^{-1}$, implying a maximum sampling interval of 0.5 microns.

The role of the detector (currently silicon based CMOS for visible light) is to record every incident photon, and its spatial location, ideally with no additional noise. The proper parameter to quantify its performance is the Detective Quantum Efficiency, which macroscopically (zero spatial frequency) can be written as [6,7]

$$DQE = \frac{(S/N)_{out}^2}{(S/N)_{in}^2}$$

where, for a Poisson process, the squared input signal-to-noise ratio is the mean photon flux $q$, measured in photons per unit area, and the output signal and noise are also measured in detected photons per unit area. For an ideal detector DQE = 1, and for a detector whose added noise is very small compared to the detected signal, DQE ≈ ε, the quantum efficiency. For modern back-thinned CMOS detectors, the quantum efficiency can exceed 0.7, so that the DQE at high light levels can also exceed 0.7. It is, therefore, not unreasonable to assume that in the future, detectors with a DQE approaching 1.0 at all but the very lowest light levels will be available. The concept of DQE can be extended to cover all spatial frequencies [7].

## 3 Fundamental Limits

The fundamental limits of phone camera lenses are of course no different to that of any other optical imaging system. One way to understand this is to consider scaling laws, as done in the seminal paper by Lohmann [8]. This approach is very useful for understanding the practical trade-offs in the design of an iteration of camera modules, for example, trading off camera dimensions, weight, etendue, noise, resolution, depth of field, and so on, but it does not cut to the chase of determining the single limiting factor.

In order to isolate the limiting quantity, it is instructive to consider other imaging systems whose performance is required to be as perfect as the laws of physics allow, and a suitable example is the astronomical telescope. Just like phone cameras, in order for telescopes to achieve their specified performance, a very large number of engineering and technological systems need to function well together; however, like the phone camera, these are technical aspects which are all necessary but are not a fundamental limit.

Over time, starting with the first telescopes in the early 1600s, astronomers have found it necessary to increase the collecting area of their telescopes, or for a single aperture telescope, the telescope entrance pupil diameter, D. Why has there been this relentless pursuit of ever large diameter telescopes? The reason is simple:

*The pupil diameter D uniquely determines the fundamental limit to performance of any imaging system, including all telescopes and the smartphone camera. The larger the diameter, the better the angular resolution and the lower the noise in the detected image.*

*Resolution*

The resolution $r$ of a diffraction-limited lens in the image plane is given by the well-known Rayleigh formula, $r = 1.22\lambda F\#$, or alternatively the cut-off spatial frequency is $1/(\lambda F\#)$ $mm^{-1}$. As stated earlier, this tells us how to sample the image intensity with no loss of information. However, the parameter of relevance is not the image plane resolution, but rather the resolution in the object space, and for that we need to consider the angular resolution, $\alpha$, where $\alpha = 1.22\lambda/D$, or alternatively the cut-off angular frequency is $D/\lambda$ $rad^{-1}$. Accordingly, the larger the diameter (of the entrance pupil), the smaller is the finest detail that can be resolved in the object, for a diffraction-limited lens.

For example, considering the "standard wide-angle 2020" lens shown in Fig 1, which has an F-number of 1.7 and a true focal length of ≈ 4 mm, the entrance pupil diameter is $D \approx 2.35$ mm, and for a mean wavelength of 0.55 microns, the angular resolution is $\alpha \approx 3\times10^{-4}$ rad, equivalent to approximately 0.3 mm for an object at a distance of 1m. To put this into perspective, the diameter of a human hair is 0.08 to 0.12 mm, smaller than the phone camera's resolution at this distance (see Section 4 for further discussion of this).

*Noise*

For a Poisson process, the mean-squared noise is proportional to the number of photons collected within a specified area, for example, in image space, the area of one pixel. For an ideal detector, the number of photons detected is the same as the number of photons incident, with no added noise (for example, no read noise, or no dark current fluctuation noise). In image space, the number of incident photons per unit area for a fixed object intensity depends only on the F-number; this is one reason why photographers use the F-number so frequently, as the exposure required to record the same scene uses the same exposure time and F-number for all detector formats/sizes. However, we are interested in the noise projected to object space, and so, as with resolution, it is the lens diameter $D$ (not the F-number) that is the fundamental limiting parameter; the number of photons collected from unit area in object space is proportional to $D^2$. The larger the diameter, the lower the noise in unit area of the object, as projected from image space.

It is clear that the quest for improved angular resolution and lower noise has driven astronomers to larger and larger diameter telescopes, for centuries, but the key message here is that the same considerations apply to all imaging systems, including phone cameras. To improve the resolution and lower the noise, the laws of physics require larger and larger entrance pupil diameter $D$.

A consequence of a large entrance pupil diameter is an increase in the hyperfocal distance, i.e. a reduction in the depth-of-field in object space. This is because the hyperfocal distance $H$, derived using the Rayleigh criterion for image plane depth-of-focus of $2\lambda F\#^2$, is given by $H \approx D^2/\lambda$, for objects relatively distant from the imaging system (i.e. object distance >> focal length). A smaller depth-of-field is generally considered aesthetically pleasing for separating the subject from the background in photography.

## 4 Discussion

The conclusion that the entrance pupil diameter $D$ is the single most important parameter determining the physical image quality of a smart phone camera system appears questionable at first, and this is often confirmed by (occasionally heated!) discussions with those involved in the detailed design of these cameras. Those working to design and build the next phone camera, or indeed one 2-3 years in the future, are constrained by dozens, or possibly hundreds, of practical constraints, especially the issues of the thickness of the camera module, cost, and power consumed (in video mode or with extensive post-detection computational processing). The issue of camera module thickness is determined by the overall phone thickness, and until now the only negotiable parameter for the camera module designer is the height of the "camera bump". Unless, there is a changed approach to smartphone design, we are stuck with small camera modules, small entrance pupil diameters $D < 3$ mm and limited image quality and noise properties (compared to larger camera systems). An exception could be the increased use of folded designs for the main camera (not just long focus modules), which might allow entrance pupil diameters up to approximately 5 mm.

How, in the future, will the images in the coming generations of smartphone cameras be "improved"? As discussed above, the scope for improving the physical resolution and noise is very limited, without a radical change in the form factor of the camera modules, and we can expect only marginal gains (small improvements in the lens and detector). Other physically-based improvements will include adding depth to the image data, using lidar, stereo or structured light, probably motivated by the widespread use of virtual reality/ augmented reality (VR/AR) headsets.

The main "improvements" in image quality will stem from the widespread deployment of machine learning (ML) algorithms. These are already incorporated in many smartphones and are potentially a powerful tool for creating more attractive images, albeit ones that may be physically inaccurate. For example, ML is regularly used to reduce noise to less than the Poisson value in the low-light modes ("night-mode") of some smartphones; reducing the noise to below that imposed by physics involves some level of prior knowledge and guesswork that is used (but is typically hidden) in the ML algorithm. Similarly, image resolution smaller than the optical diffraction limit is possible using ML. This should not be surprising, as prior knowledge (e.g. using Bayes' theorem) has been the pillar of statistical inference for the past 100 years. Machine learning is already widely used to artificially reduce the depth-of-field (blurring the background), to adapt the colour balance to the photographers preference (via "likes") and for face beautification, for example.

The key issue with the application of machine learning is that the enhancements are not real, but rather are based on statistical inference, and are critically dependent on the architecture and training of the machine learning model. Consider the example given earlier of imaging a person's hair with a (diffraction limited) phone camera at a distance of 1 m from the subject, where we showed that the blur in object space was around three times the diameter of a hair (i.e. it is difficult or impossible to distinguish individual hairs). We can ascertain from the image that the person does indeed have hair, we can record the hair colour and from the face can infer ethnicity, so, as we zoom into the image ("pinching" the screen) ML can indeed reveal sharp hairs which suit that person. At the same time, we can "beautify" the image so that wrinkles and blemishes are removed from the person's face.

## 5 Future Studies

The concept of a fundamental limit itself raises questions as to the underlying assumptions. In Section 1, four assumptions were stated, and these may not be valid as we look to the future. Here, we discuss some of the issues that might influence the limits of imaging in the future.

### *Synthetic aperture and lensless imaging*

In this paper, we have only considered a single imaging aperture, with a lens (or mirror) producing a compact, diffraction-limited point spread function in the image plane, i.e. conventional imaging. Synthetic aperture imaging has been used in radio-astronomy for > 50 years and in optical astronomy, the use of multiple telescopes is already a reality (e.g. [9]), achieving the goal of higher angular resolution. There are also many research studies (e.g. [10]) on lensless imaging, which can also to some extent provide object depth information.

*The image intensity – detector interface*

We have assumed that the intensity formed by the lens in the image plane can be partitioned geometrically between pixels, which is reasonable when the pixel dimensions are greater than the mean wavelength of the light. But increasingly this is not the case, with the smallest CMOS pixels currently being ≈0.64 microns square [11]. In this case, one needs to model the intensity/detector interface using electromagnetic theory [12].

*The ideal detector*

In addition to counting every arriving photon, and its location, one might consider that the ideal detector should also count the time-of-arrival, wavelength and polarisation of each photon, with a precision permitted by the uncertainty principle.

*Quantum effects*

Finally, there is the possibility that using sub-Poissonian light sources may enable improved image quality [13]. At the time of writing, this seems far-fetched for consumer imaging in natural light, but the current intense research effort in quantum technology may bear fruit in the future.

## Acknowledgements and Dedication

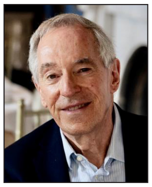
This paper has drawn on material contained in a magazine article [14] by the author.

The paper is dedicated to Professor Anna Consortini, who through her long and active research career performed many ground-breaking experiments on the propagation of light through turbulence, thus contributing significantly to optical communication through the atmosphere and to adaptive optics.

## References

1. https://photutorial.com/photos-statistics/ (retrieved 5th July 2023)
2. https://www.statista.com/statistics/259477/hours-of-video-uploaded-to-youtube-every-minute/ (retrieved 5th July 2023).
3. Prakash K, Diederich B, Heintzmann R, Schermelleh L, Super-resolution microscopy: a brief history and new avenues, *Phil Trans R Soc A*, 380(2022)20210110; doi.org/10.1098/rsta.2021.0110.
4. Blahnik V, Schindelbeck O, Smartphone imaging technology and its applications, *Adv Opt Techn*, 10(2021)145–232.
5. P Wang and R Menon, Ultra-high sensitivity color imaging via a transparent diffractive-filter array and computational optics, *Optica*, 2(2015)933–939.
6. Rose A, A unified approach to the performance of photographic film, television pick up tubes, and the human eye, *J Soc Motion Pic Eng*, 47(1946)273–294.
7. Dainty J C, Shaw R, Image Science, (Academic Press), 1974.
8. Adolf W Lohmann, Scaling laws for lens systems, *Appl Opt*, 28(1989)4996–4998.
9. https://www.eso.org/sci/facilities/paranal/telescopes/vlti.html (retrieved 5th July 2023).
10. Boominathan V, Robinson J T, Waller L, Veeraraghavan A, Recent advances in lensless imaging, *Optica*, 9(2022)1–16.
11. https://news.samsung.com/global/samsung-brings-advanced-ultra-fine-pixel-technologies-to-new-mobile-image-sensors (retrieved 5th July 2023).
12. Catrysse P B, Zhao N, Jin W, Fan S, Subwavelength Bayer RGB color routers with perfect optical efficiency, *Nanophotonics*, 11(2022)2381–2387.
13. Berchera I R, Degiovanni I P, Quantum imaging with sub-Poissonian light: challenges and perspectives in optical metrology, *Metrologica*, 56(2019)024001.
14. Dainty C, Film photography is dead: long live film, IEEE Consumer Electronics Magazine, (Jan 2012), 61–64.

[*Received*: 08.07.2023; *revised recd*: 14.07.2023; *accepted*: 15.07.2023]



Chris Dainty is currently an Emeritus Professor at several universities. His research interests were mainly in imaging, speckle and the propagation of light through turbulence. During his research career at The University of Rochester NY, Imperial College London, The University of Galway Ireland and Xperi (Galway) his group published around 180 peer-reviewed papers and more than 300 conference presentations. He was advisor to 65 PhD students and more than 75 post-docs. Prof Dainty received a number of awards, including the ICO Prize (1984) and the C E K Mees Medal and Prize (2003), and in 2008 was elected a member of the Royal Irish Academy. He has also served as President of The International Commission for Optics (1990-1993), The European Optical Society (2002-2003) and Optica (formerly OSA, 2011) and served on the Boards of SPIE (1994-1996) and the UK Institute of Physics (1996-1997).