ISSN 1756-753X

AARGnews

Number 59 October 2019

Contents

Resolving some spatial resolution issues – Part 2: When diffraction takes over

Geert Verhoeven

In part 1 "Between line pairs and sampling distances" of this split entry on spatial resolution, it was mentioned that this second part would focus on some fundamental laws of electromagnetic radiation to shed more light on the concepts of spatial resolving power and spatial resolution. The numbering of chapters, illustrations and equations simply continues so that one can easily merge both parts into one larger text. However, and in contrast to what was initially stated, this entry will not be the final part about spatial resolution. Because there is quite some, but still basic, ground to cover once the physical views on spatial resolution start to complement the geometrical ones, there will also be a third part. This approach also prevents any of these entries from becoming too lengthy.

3 Beyond pure geometry

To generate a sharp image with lots of details, it does not suffice to use a long focal length, get close to the scene or use a detector with a small detector pitch (as seen in part 1); one also needs a lens that can resolve fine object detail. This is of the utmost importance since the lens elements (and as we will see in part 3, even the sensor) blur the incoming radiation in unavoidable ways. To understand this blurring principle and its integration within the purely geometrical relationships introduced before, it is essential to know that an image is a visual representation of a specific physical object. Ideally, every object point is represented by an infinitesimally small spot in the image. However, the physical phenomenon of **diffraction** prevents this by turning every object point into a little blob.

3.1 Diffraction and the optical PSF

Every lens-based imaging system suffers from specific optical errors or defects. The generic name given to these lens defects is **aberration**. Longitudinal/axial and transverse/lateral chromatic aberrations occur when imaging a spectrum of different wavelengths. These **polychromatic aberrations** contrast with the **monochromatic aberrations** which occur even for incident electromagnetic radiation of only one wavelength. In total, there are five primary monochromatic aberrations, also known as the **Seidel aberrations**: spherical aberration, coma, astigmatism, Petzval field curvature, and distortion.

Even if one could make a lens devoid of all **aberrations** (i.e. a **perfect lens**), the diffractive nature of electromagnetic radiation would still put a physical limit on the smallest resolvable object. **Diffraction** is a phenomenon that comes into play because propagating electromagnetic waves bend in the neighbourhood of tiny obstacles and spread out when passing apertures. When imaging a single, distant point source of electromagnetic energy (such as a star), the resulting image is therefore never a perfect point but a diffraction pattern.

The spatial energy distribution of this image spot is called the **Point-Spread Function** (**PSF**). It describes the smear or spread in the sensor plane introduced by the optical chain. If the imaging system is aberration-free and only limited by diffraction, the three-dimensional PSF of a perfectly focused point is a so-called **Airy diffraction pattern** (Figure 7A). This Airy pattern geometrically describes the best-focused spot of electromagnetic radiation that a perfect lens with a circular aperture can generate. In the plane of the image sensor, the irradiance distribution of the Airy pattern looks like a bright circular central patch (the **Airy disc**) and a series of dimmer concentric rings, each ring separated by a circle of zero irradiance (Figure 7B & 7C). The pattern and disc were named after George Biddell Airy, a nineteenth-century English astronomer who was the first to calculate this irradiance distribution.

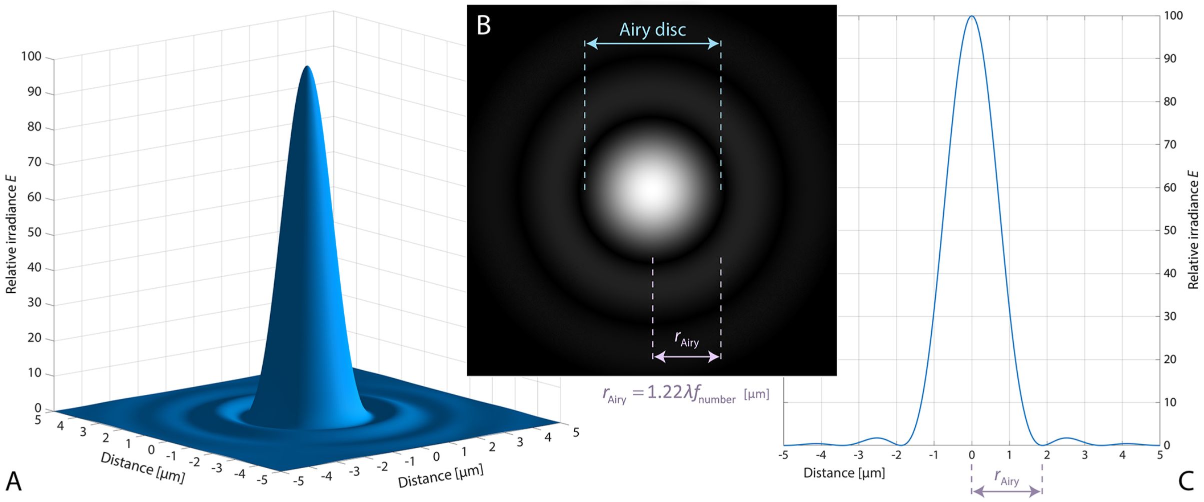


Figure 7 Inset A depicts the three-dimensional Airy diffraction PSF (at 550 nm and *f*/2.8), whereas B shows the two-dimensional PSF (slightly contrast-boosted) as seen in the image plane. Notice the high-irradiance circular Airy disc. Inset C illustrates a profile through the centre of the PSF. In B and C, the radius of the Airy disc is indicated.

The Airy disc contains about 84 % of the Airy pattern's irradiance, and its linear radius *r* equals:

$$r_{\text{Airy}} = 1.22\frac{\lambda f'}{D} \text{ [µm]} \qquad \text{<11>}$$

The spread of the imaged object point thus depends on:

- the wavelength of the electromagnetic radiation (λ);
- the lens' focal length (f');
- and the diameter of the lens aperture (D).

For the lens of any photo camera, the latter two quantities f' (or simply f) and D are always combined in a **relative aperture** (e.g. *f*/2.8 or *f*/8). Photographers use these relative aperture statements simply because it would be hard to remember the exact physical dimensions of every lens aperture.

Using an aperture statement relative to the focal length also has the substantial advantage that any lens using that particular aperture (e.g. *f*/8) will transmit the same amount of electromagnetic radiation, irrespective of its focal length. For instance, a 130 mm lens with a relative aperture *f*/2 has a physical aperture diameter of 65 mm (130 mm / 2), while a 50 mm lens at that setting features a 25 mm aperture diameter (50 mm / 2). Those specific combinations of physical lens aperture size and focal length create the same angular aperture size, which in turns explains why the light transmitted by any *f*/2 configuration will result in the same exposure (that is, ignoring minimal differences in the lens' efficiency to transmit electromagnetic radiation) (see Figure 8 for a graphical explanation).

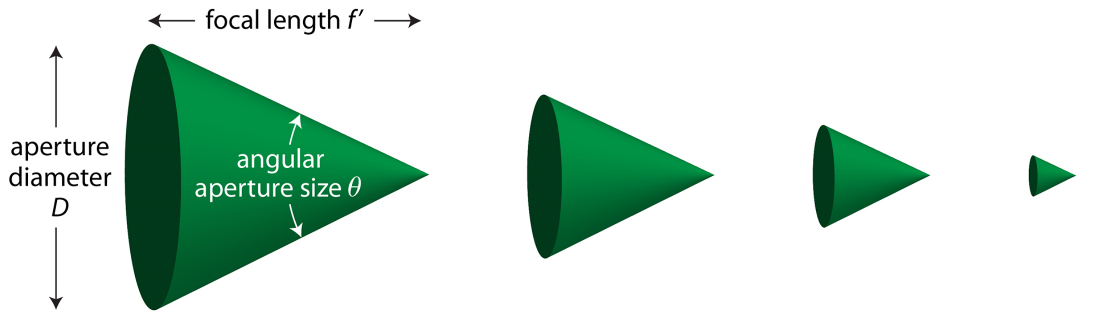
The numeric part of such an aperture statement is the **relative aperture number** or ***f*-number**. One can find a standard aperture series (*f*/2.8, *f*/4, *f*/5.6, *f*/8, *f*/11) engraved on many photographic lenses. This information now allows us to rewrite equation <11> for normal photographic purposes:

$$r_{\text{Airy}} = 1.22\lambda f_{\text{number}} \text{ [µm]} \qquad \text{<12>}$$

The diameter of the Airy disc is then just twice the Airy radius or:

$$d_{\text{Airy}} = 2.44\lambda f_{\text{number}} \text{ [µm]} \qquad \text{<13>}$$

Note that smaller wavelengths (e.g. ultraviolet radiation) and large lens apertures yield smaller Airy discs, and that both variables can be used to minimise the size of the image spot.



Focal length *f′*	130 mm	80 mm	50 mm	20 mm
Aperture diameter *D*	65 mm	40 mm	25 mm	10 mm
Aperture angular size θ	26.6°	26.6°	26.6°	26.6°
***f*-number (*f′* / *D*)** **Relative aperture number**	2	2	2	2
(Relative) aperture	*f*/2	*f*/2	*f*/2	*f*/2

Figure 8 The concept of relative aperture. If lenses with different focal lengths feature the same angular aperture size, their relative aperture number and *f*-number are identical.

Let us illustrate this with some examples. If we photograph something using a lens at aperture *f*/2.8 and the scene is illuminated with visible electromagnetic radiation (i.e. light) which runs from 400 nm to 700 nm (so 550 nm is the central wavelength), the radius of the Airy disc becomes:

$$\text{radius Airy disc} = 1.22 * 0.550\ \mu\text{m} * 2.8 = 1.9\ \mu\text{m},$$

and its diameter 3.8 µm (see Table 2 and Figure 7). Reducing the aperture diameter fourfold to *f*/11 yields an Airy disc diameter of 14.8 µm. Switching the illumination for a near-ultraviolet source with a central wavelength at 360 nm decreases the Airy disc diameter to 9.7 µm (see Table 2 and Figure 9).

	Wavelength	Aperture	Airy disc radius	Airy disc diameter	$RP_{spatial}$
Starting situation	550 nm	*f*/2.8	1.9 µm	3.8 µm	526 LP/mm
Smaller aperture	550 nm	*f*/11	7.4 µm	14.8 µm	135 LP/mm
Shorter wavelength	360 nm	*f*/11	4.8 µm	9.7 µm	207 LP/mm

Table 2 Different apertures and wavelengths yield dissimilar Airy discs and various spatial resolving powers.

Please be aware that the numbers mentioned above, and the plots of Figure 9, are only for **diffraction-limited** lenses (i.e. perfect lenses whose performance is limited only by diffraction and not by any aberration) that image a distant point that is perfectly in focus. So even under these very hard to achieve circumstances, very small apertures (indicated by larger *f*-numbers; e.g. *f*/16 or *f*/22) will always yield big Airy discs, thus lowering the spatial resolving power (see next section). In more realistic imaging conditions with aberration-limited lenses, the diffraction-induced blur is even worse.

In many situations, the smallest spot one can image (given a particular part of the electromagnetic spectrum and a specific lens aperture) surpasses the sensor's detector pitch. For example: an Airy disc diameter of 14.8 µm at *f*/11 is triple the size of the 4.9 µm detector pitch of the Nikon D810. Section 4.4 will explain how and when this size difference can lead to a loss in resolving potential. The important message is thus that the physical nature of electromagnetic radiation sets a fundamental limit by blurring the image, preventing an exact point-for-point copy of the real-world scene. [Note that the next sections will often provide two equations: one applicable to all optical imaging systems, and a simplified version for photo camera lenses. Although standard photo cameras are optical imaging systems, we can simplify the equations because their whole operation revolves around *f*-numbers.]

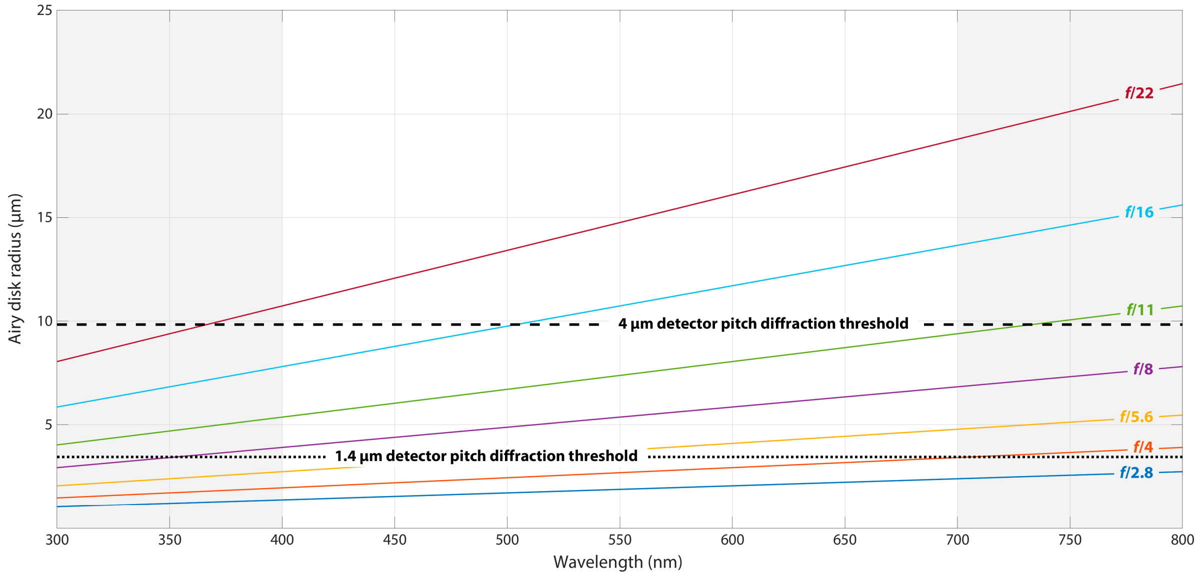


Figure 9 The influence of aperture and wavelengths on the Airy disc radius. The detector pitch diffraction threshold for two common detector pitches is indicated as well. The grey zones indicate invisible electromagnetic radiation.

3.2 Meet Baron Rayleigh

A scene is considered a collection of countless object point sources, and each of these sources generates a diffraction pattern with its amplitude proportional to the source´s radiance. The total radiance arriving at a specific point in the focal plane is nothing but the sum of contributions from all optical PSFs in the neighbourhood of that point. The unavoidable diffraction blurring by the lens has thus a direct effect on the amount of detail one can record and observe in an image, because it influences every imaged object point.

Obviously, smaller Airy discs yield a higher theoretical spatial resolving power of every optical system. However, even if the optical imaging system generates tiny Airy discs, diffraction patterns will always overlap in an image, and at a certain point, it will become more and more challenging to separate them (i.e. detect the presence of distinct objects – see the upper part of Figure 10). A prevalent metric to quantify the smallest spatial separation of two object points that still allows them to be unambiguously resoluted as two different points in the image is called the **Rayleigh criterion** (Figure 10).

Proposed by John William Strutt (Third Baron Rayleigh) in 1879, this criterion states that an aberration-free lens with a uniform circular aperture can spatially just resolve two adjacent point sources of electromagnetic radiation (that are incoherent, unpolarised, and of equal radiance) when the centres of their Airy discs are separated by the radius of the Airy disc r_{Airy} (Figure 10). When the separation distance between adjacent Airy patterns gets smaller than their Airy disc radii, the PSFs of the object points blend into an indecipherable blur, thereby rendering the point sources indistinct. Mathematically, this means that the superimposed Airy patterns must be separated by at least Δx_{min}:

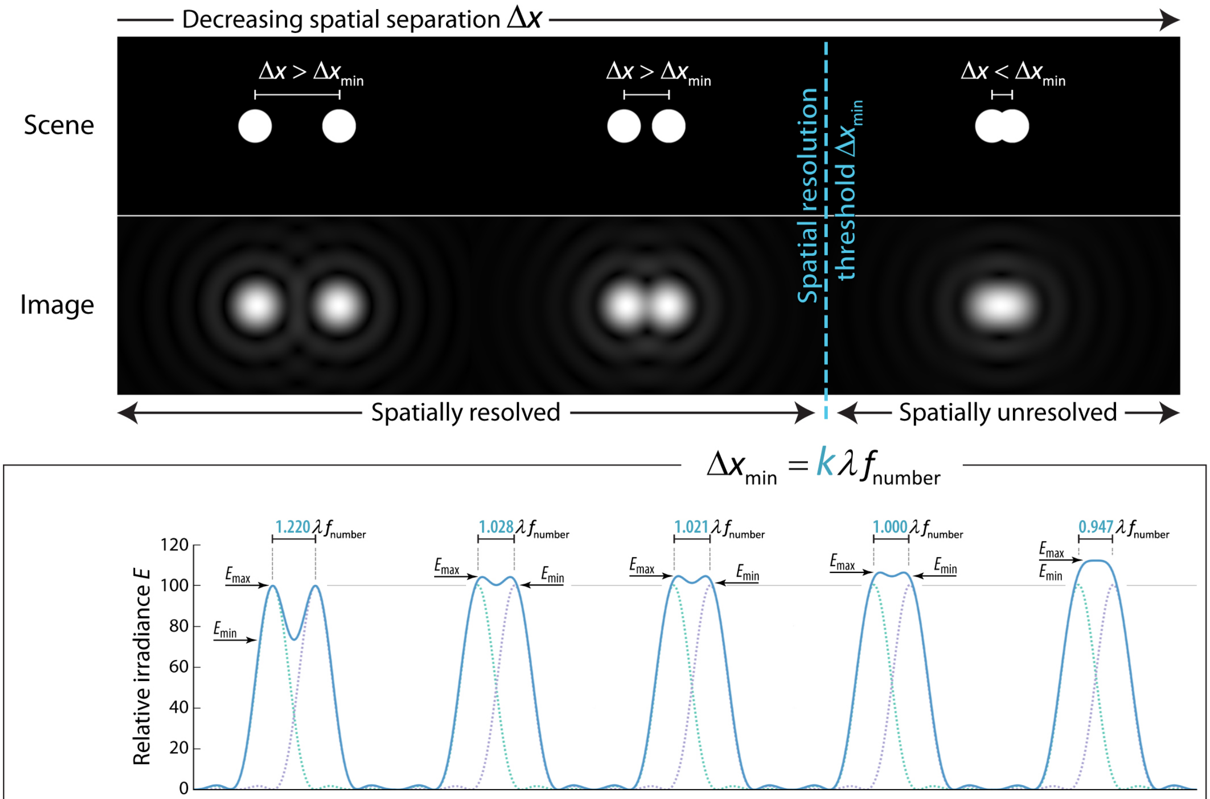
$$\Delta x_{\text{min}} = 1.22\lambda f_{\text{number}} \text{ [µm] (for photo camera lenses)} \quad \text{<14>}$$

$$\Delta x_{\text{min}} = 1.22\frac{\lambda f'}{D} \text{ [µm] (for optical imaging systems).} \quad \text{<15>}$$

This separation is known as the **optical limit of spatial resolution**. With the help of equation <3>, the **angular optical limit of spatial resolution** can be expressed in radians as:

$$\Delta\theta_{\text{min}} = 1.22\frac{\lambda}{D}\ \text{[rad]}\,.$$ <16>

In other words: two point sources (or more general, objects) can be distinctly rendered by a perfect optical instrument only if their images on the sensor feature an angular separation that surpasses $1.22\lambda/D$. This short equation indicates that if one wants to obtain a smaller angular separation to render the final image more detailed, either the captured wavelengths λ must be shorter (e.g. blue light instead of red light), or the lens aperture D should be bigger (e.g. $f/2.8$ instead of $f/11$).



Criterion	Rayleigh criterion	Houston criterion	Dawes criterion	Abbe criterion	Sparrow criterion
Publication year	1879	1926	1867	1873	1916
Author	John William Strutt, 3rd Baron Rayleigh (1842-1919)	William Vermillion Houston (1900-1968)	William Rutter Dawes (1799-1868)	Ernst Abbe (1840-1905)	C.M. Sparrow
k-value	1.220	1.028	1.021	1.000	0.947
Contrast $E_{max} - E_{min}$	26.5 %	3.9 %	3.3 %	1.8 %	0 %
$(E_{max} - E_{min}) / E_{max}$	26.5 %	3.8 %	3.2 %	1.7 %	0 %
$(E_{max} - E_{min}) / (E_{max} + E_{min})$	15.3 %	1.9 %	1.6 %	0.9 %	0 %

Figure 10 Optical diffraction prevents a point-like irradiance distribution of imaged scene points. Resolving two scene points with a lens is only possible when the spatial separation between these fuzzier image renderings is large enough. This figure provides the five most used criteria to quantify the minimum spatial separation between the object points before their PSFs become unresolvable. They all differ by a **multiplication factor *k*** that quantifies the PSFs' spatial separation in terms of the wavelength-aperture product. The figure also indicates that one cannot define spatial resolution without involving the term contrast. For instance, the most popular criterion by Rayleigh states that two identical point objects will be distinct as long as the difference between the two irradiance maxima and the minimum in between them is not smaller than 26.5 %. Imaging detectors and the human visual system can easily distinguish such an irradiance contrast, provided that those maxima and minima are above their own threshold. That is why other scholars have provided smaller values for k.

Strictly speaking, most perfect point sources could still be observable at a slightly smaller separation. This led to the development of other two-point spatial resolution criteria (see Figure 10), many of which were proposed after considering the geometry of the PSF. Figure 10 illustrates that all these criteria can be written in a very general form. One just has to replace the number '1.22' from equations <14>, <15> or <16> by a **multiplication factor *k***. The difference between these criteria then boils down to a different value for *k*. The development of various measures to quantify spatial resolution notwithstanding, Rayleigh gave a very acceptable and general condition that is frequently relied on in the imaging industry. Now, how can we use the dimensions of the Airy disc and this Rayleigh criterion to come up with a statement of spatial resolving power in LP/mm?

3.3 LP/mm quantification

The Airy disc radius can thus be used as an approximate measure of the smallest spatial detail that an image theoretically can contain. An *f*/2.8 imaging system that uses mainly visible light (i.e. a central wavelength of 550 nm) produces Airy discs with radii of 1.9 µm. Using Rayleigh's criterion, Δx_{min} or the absolute minimum spatial resolution of an image equals, therefore, 1.9 µm. Since we know that spatial resolving power is the latter value's reciprocal, this setup features a maximum theoretical spatial resolving power of 526 LP/mm.

$$RP_{spatial} = 1 \text{ LP} / 0.0019 \text{ mm} = 526 \text{ LP/mm}$$

More values are given in Table 2. If we were to use another spatial resolution criterion, like the one from Sparrow, we would arrive at higher spatial resolving power values. Whereas Rayleigh states *k* = 1.22, Sparrow defined *k* to be 0.95. In other words, Sparrow said that two object points can be much closer together than the Rayleigh criterion before they are spatially unresolvable. How close? Until their Airy discs yield a cumulative irradiance distribution without any contrast between the peaks (see Figure 10). For an *f*/2.8 aperture and 550 nm light, this means a separation of 1.46 µm, equalling 667 LP/mm (i.e. a more than 20 % increase compared to the 526 LP/mm according to the Rayleigh criterion):

$$\Delta x_{min} = 0.95 * 0.000550 \text{ mm} * 2.8 = 0.0015 \text{ mm}$$
$$RP_{spatial} = 1 \text{ LP} / 0.0015 \text{ mm} = 667 \text{ LP/mm.}$$

Again, it is imperative to understand that real photographic lenses can never obtain these spatial resolving power values, simply because they always feature a varying amount of aberrations which results in image spots that are fuzzier and more irregular than the Airy pattern (see Figure 11). Only at small apertures, a lens might approach a perfect lens (because smaller apertures reduce the influence of specific aberrations). But even then, imperfect focus blurs an image spot considerably, so that the practical PSF of an imaging lens strongly deviates from (and exceeds!) this ideal diffraction-limited Airy pattern. To indicate that every object point corresponds to a blurry spot in image space due to the combined effect of diffraction, aberrations and focus, the field of photography has coined the quantifiable terms **blur circle** and **circle of confusion**.

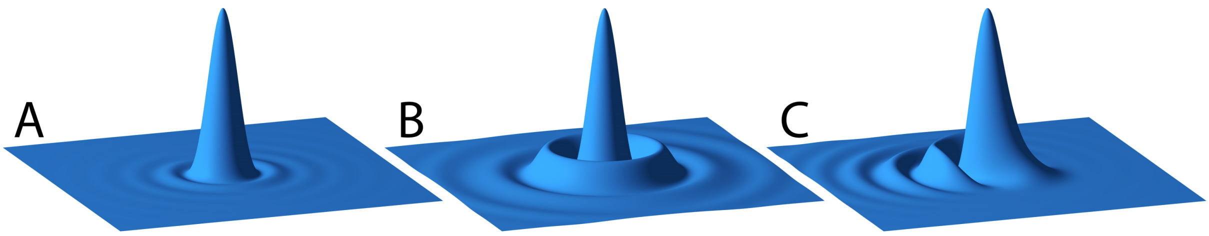


Figure 11 Inset A displays the three-dimensional PSF of a perfect, aberration-free lens. B and C depict the PSF for a lens with spherical aberration and coma, respectively. All plots have their irradiance normalised to their maximum.

4 Combining diffraction and geometry

4.1 Rayleigh in object space

Despite being purely theoretical, these Rayleigh-defined numbers can provide useful insight into the smallest objects one could spot from a remote vantage point. To do so, we have to combine some of the primary geometrical relationships with the diffraction-limited nature of electromagnetic radiation.

The Rayleigh-based computations of the previous section expressed the minimum spatial resolution Δx_{min} in image space. To state the same quantity in scene-specific units, we need to resort to the magnification factor. As explained in part 1, image scale is computed as the ratio of focal length to object distance (*f′*/*s*), whereas the scale factor or magnification *m* is its reciprocal (see equation <6>). Mathematically, the combination of these elements is written as follows:

$$\Delta x_{\text{min}} = 1.22 \lambda f_{\text{number}} \frac{s}{f'} \text{ [mm] (for photo camera lenses)} \qquad <17>$$

$$\Delta x_{\text{min}} = 1.22 \frac{\lambda f' s}{D f'} = 1.22 \frac{\lambda s}{D} \text{ [mm] (for optical imaging systems).} \qquad <18>$$

These equations allow computing the smallest, theoretically possible spatial resolution of any image, irrespective of the imaging sensor, the intervening atmosphere, the contrast of the object, the aberrations of the lens (which are assumed zero), and any object or camera motion.

Note that from everything we have said so far, it should be clear that it does not make sense to use the common term ‘(very) high-resolution image’ to refer to an image in which many scene details are visible. The term is a *contradictio in terminis* since an image with many spatial details features a small resolvable distance, and thus a small spatial resolution. The imaging system that creates such an image is characterised by a high spatial resolving power. Therefore, so-called ’high-resolution’ images do not convey small spatial details, and are produced by devices with low spatial resolving abilities.

4.2 Thermal and spaceborne imagers work against you

To illustrate equations <17> and <18>, consider an airborne camera flown at 350 m using a 50 mm lens at *f*/8. The exact type of camera is not important, as we will compute the smallest possible resolvable distance based on the lens aperture and the electromagnetic radiation imaged. We now get:

λ = 550 nm (0.000550 mm)
f-number = 8
s = 350 m (350 000 mm)
f′ = 50 mm
Δx_{min} = 1.22 * 0.000550 mm * 8 * (350 000 mm / 50 mm) = 38 mm.

So, circa 4 cm is the very best we can do in the utopic situation that the sensor is not limiting the spatial resolution of the image, while we also assume a static object and camera, an aberration-free lens, the object impeccably in focus, and no influence of the atmosphere. Now, let us put the same lens on a camera that captures thermal radiation at 10 500 nm or 10.5 µm:

λ = 10 500 nm (0.010500 mm)
f-number = 8
s = 350 m (350 000 mm)
f′ = 50 mm
Δx_{min} = 1.22 * 0.010500 mm * 8 * (350 000 mm / 50 mm) = 717 mm.

This shows why thermal imagery always features a larger spatial resolution. The only way to decrease the spatial resolution of that kind of imagery is by increasing the aperture (to *f*/1.4 for instance), flying lower (e.g. 200 m) or using a longer focal length lens (e.g. 150 mm).

λ = 10 500 nm (0.010500 mm)	λ = 10 500 nm (0.010500 mm)	λ = 10 500 nm (0.010500 mm)
f-number = 1.4	*f*-number = 8	*f*-number = 8
s = 350 m (350 000 mm)	s = 200 m (200 000 mm)	s = 350 m (350 000 mm)
f' = 50 mm	f' = 50 mm	f' = 150 mm
Δx_{min} = 126 mm	Δx_{min} = 410 mm	Δx_{min} = 239 mm

Finally, let us consider the physical size D of the aperture that a satellite-based imager would need in order to deliver a standard colour photograph with a spatial resolution of 10 cm. For this, we assume an orbit altitude of 450 km and need to use equation <18>:

$$\lambda = 550 \text{ nm } (0.000550 \text{ mm})$$
$$s = 450 \text{ km } (450\,000\,000 \text{ mm})$$
$$\Delta x_{min} = 10 \text{ cm } (100 \text{ mm})$$
$$D = 1.22 * 0.000550 \text{ mm} * 450\,000\,000 \text{ mm} / 100 \text{ mm} = 3020 \text{ mm.}$$

To minimise the length of the lens, the engineers could opt for an aperture of *f*/1.4 (because any larger *f*-number would yield a longer lens). Nevertheless, to reach an *f*/1.4 aperture with a lens that has a physical opening of circa 3 m, the lens still should have a focal length of 4.2 m.

$$\text{focal length } f' = \text{diameter aperture } D * f\text{-number} = 3020 \text{ mm} * 1.4 = 4228 \text{ mm}$$

Such a lens is not only massive; it is also very costly, exceptionally complicated to construct, and the satellite requirements for its deployment not trivial. For example: the largest high-quality optical lens known for civilian purposes has a diameter of 'only' 1.57 m. This lens belongs to the camera system of the Large Synoptic Survey Telescope. Such telescopes collect electromagnetic radiation mainly by reflection on large mirrors, after which a smaller lens can be used to project the collected radiation on the imaging sensor. Modern satellite-mounted imagers also rely on large reflecting telescopes instead of conventional camera-like lenses, because large mirrors are easier to construct. Let us illustrate their potential with one last example.

The camera of the classified American KH-11 KENNEN reconnaissance satellite features a telescope with an alleged 2.4 m-diameter mirror, watching Earth from an orbital altitude of circa 250 km. Equation <16> helps us to compute the angular limit of spatial resolution, while equation <3> can be used afterwards to express the spatial resolution of the resulting image:

$$\lambda = 550 \text{ nm } (0.000550 \text{ mm})$$
$$D = 2.4 \text{ m } (2400 \text{ mm})$$
$$s = 250 \text{ km } (250\,000\,000 \text{ mm})$$
$$\Delta\theta_{min} = 1.22 * 0.000550 \text{ mm} / 2400 = 0.0002796 \text{ mrad}$$
$$\Delta x_{min} = \Delta\theta_{min} * s = 0.0002796 \text{ mrad} * 250\,000\,000 \text{ mm} = 70 \text{ mm.}$$

A theoretical spatial resolution of 7 cm is awe-inspiring indeed, but consider the hardware that is needed for this. All these examples indicate why it is so difficult to get a spatial resolution from space that falls in the airborne range. Diffraction merely is working against you, and only a large aperture allows getting around it. On top of that, the atmosphere and speed of the satellite are not exactly in favour of spaceborne imaging either. In short: capturing images from space is an entirely different ballgame from acquiring them within the Earth's atmosphere, dramatically influencing the hard- and software designs. That is why reading a newspaper or spotting pottery shards from space currently still belong to science-fiction stories, even for orbital espionage technology.

4.3 Achieving a balance between two limits

From equation <10> in part 1, we know that the GSD should be at least three times as small as the smallest spatial distance we want to resolve in the image. For the spy satellite setup, this would mean a GSD of 2.3 cm or smaller. Using equation <10> as the connection, we now can combine the geometrical view (i.e. equation <5>) with diffraction theory (i.e. equation <18>):

$$3\frac{sp}{f'} = 1.22\frac{\lambda s}{D}$$ <19>

Figure 12 graphically translates equation <19>. Both clearly express the most fundamental ways to reduce the spatial resolution of an image (i.e. increase image details), either by:

- improving the limit imposed by detector sampling (left side of the equation), achieved by:
  - getting closer to the subject (*s*);
  - using a detector with a smaller photosite pitch (*p*);
  - applying a longer focal length lens (*f'*);
- improving the limit imposed by optical diffraction (right side of the equation), achieved by:
  - capturing electromagnetic radiation with a shorter wavelength (λ);
  - getting closer to the subject (*s*);
  - using a bigger lens aperture (*D*);

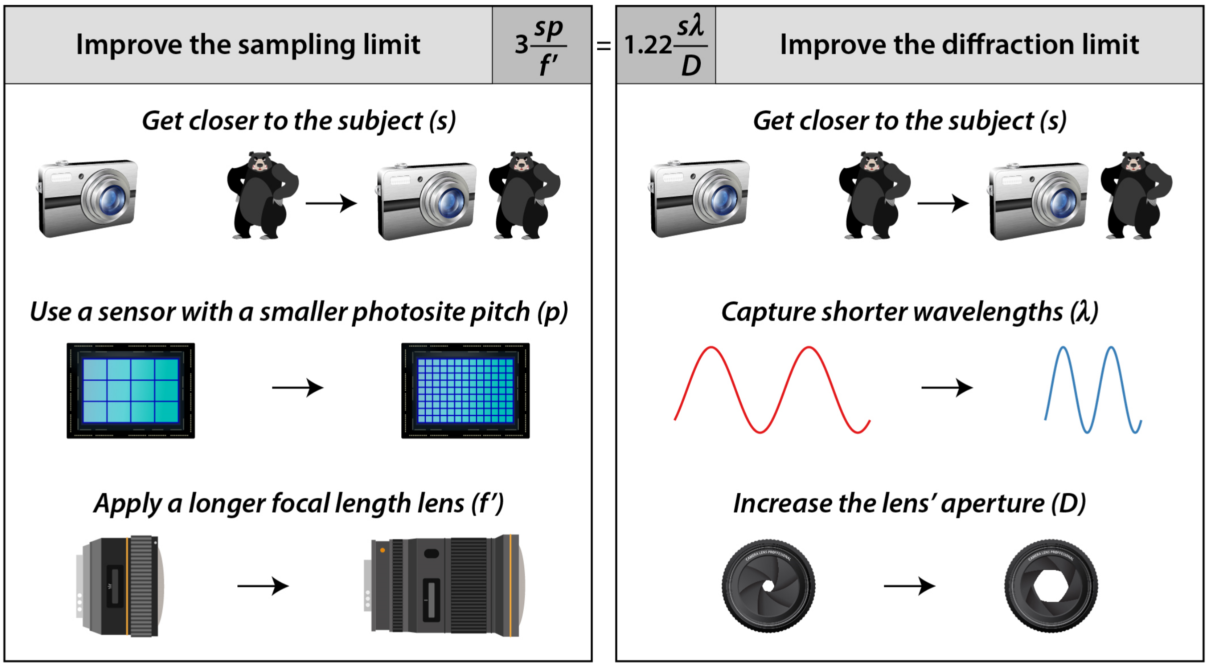


Figure 12 Achieving a smaller spatial resolution is possible via three geometry-related parameters (i.e. those defining the detector sampling limits) and three variables that determine the extent of optical diffraction.

The spatial resolving power limits imposed by the image sensor and optics thus need to be balanced for best spatial resolution performance. With equation <19>, we can now calculate the detector pitch *p* that an imager on the above-mentioned civilian satellite needs to yield this 10 cm spatial resolution:

s = 450 km (450 000 000 mm) → can be left out, because it occurs on both sides of equation <19>
f' = 4228 mm
λ = 550 nm (0.000550 mm)
D = 3020 mm
p = (1.22 * 0.000550 mm * 4228 mm) / (3 * 3020 mm) = 0.00031 mm = 0.31 µm.

Such a 0.31 µm detector pitch is – at least known from civilian systems – currently unrealistic. The image sensor carried by the WorldView-2 satellite features an 8 µm detector pitch. State-of-the-art image sensors used in very high megapixel phone cameras sport detector pitches of around 0.8 µm, but these suffer from poor performance in low-illumination conditions (of which satellite imaging is undoubtedly a good example). So assuming a three times higher and more realistic 1.0 µm detector pitch, equation <19> and Figure 12 illustrate that the same image spatial resolution can be achieved with a flying height that is circa three times lower (i.e. about 140 km). Since this alteration also influences the optical right-side of equation <19>, it has the positive effect that the aperture diameter can be reduced to about 940 mm.

4.4 Detector pitch-related diffraction threshold

In the first part of this entry on spatial resolution, we argued that digitising a line pair (or any feature of a given size) with three pixels seems more appropriate than the two pixels per line pair given by the Nyquist–Shannon sampling theorem (this is also reflected in equation <19>). Since we know from the Rayleigh criterion that Δx_{min} equals the radius of the Airy disc, it follows that the best 'match' between sensor characteristics and the optics is achieved when the detector pitch p equals one-third of the Airy disc radius. In other words: three pixels suffice to resolve the Airy disc radius entirely. The reasoning then goes that above that threshold, the optics diffraction limitation kicks in. So digitising the radius of the PSF's central part with more than three pixels does not improve the final spatial resolution of the image because the detector only resolves diffraction blur at that point.

However, photographic tests show that the visual impact of optical diffraction starts already earlier, somewhere between two and three times the detector pitch. Because it is impossible to give one definite number for all cameras and lenses, this three-pixel threshold is often redefined as twice Rayleigh's k factor of 1.22:

$$r_{\text{Airy\_max}} = 2.44p \text{ [µm]} \qquad \text{<20>}$$

Criterion <20> quantifies the **detector pitch diffraction threshold**. At that threshold, the optics' diffraction and detector sampling rate are <u>visually</u> in balance. Figure 9 illustrates this limit for a 4 µm detector pitch (commonly found amongst digital photographic cameras):

4 µm detector pitch diffraction threshold = 2.44 * 4 µm = 9.8 µm.

Above that value, optical diffraction *likely* deteriorates the image details, because the PSFs are simply too big regarding the detector pitch. Replacing and reshuffling some terms in equation <20> allows expressing the maximum *f*-number that can be used given this threshold:

$$1.22\lambda f_{\text{number\_max}} = 2.44p \rightarrow f_{\text{number\_max}} = \frac{2.44p}{1.22\lambda} = \frac{2p}{\lambda} \qquad \text{<21>}$$

This *f*-number is denoted the **diffraction-limited aperture**. When dealing with visible radiation, 700 nm should be used for λ since this is the upper sensitivity of any standard digital photo camera and diffraction is biggest at this wavelength. This indicates that optics-induced blur starts to dominate around and after *f*/11 in the visible range for most modern still cameras (see also Figure 9):

$f_{\text{number\_max}}$ = 2 * 4 µm / 0.7 µm = 11.4.

Aperture *f*/11 is, therefore, the lens opening at which a camera with a 4 µm detector pitch becomes diffraction-limited. At smaller apertures like *f*/16, *f*/22 or even *f*/32, the imaging sensor simply out-resolves the camera optics, hence limiting the system's total spatial resolving power and yielding a fuzzier image (Figure 13B). Note that the amount of optical diffraction is always determined by the optics; the size of the detector pitch just determines when and how well this limitation will be spatially resolved. For those reasons, ‘pixel peepers’ use a factor of 2 instead of 2.44 in Equation <20>, which in turn yields a factor of 1.64 in Equation <21>.

Equation <21> also teaches us that this diffraction threshold is lower for sensors with a smaller detector pitch. Smartphones commonly feature a 1.4 µm image sensor pitch, which means that optics-induced blur will deteriorate image details for apertures smaller than *f*/4 (see Figure 9):

$$f_{number\_max} = 2 * 1.4\ \mu m / 0.7\ \mu m = 4.$$

[Although an in-depth treatment is out of scope here, note that smaller apertures increase the depth of field (see Figure 13A). This relationship is a double-edge sword with significant consequences. Balancing diffraction versus depth of field is, for instance, vital when determining the ideal aperture for image-based three-dimensional modelling.]

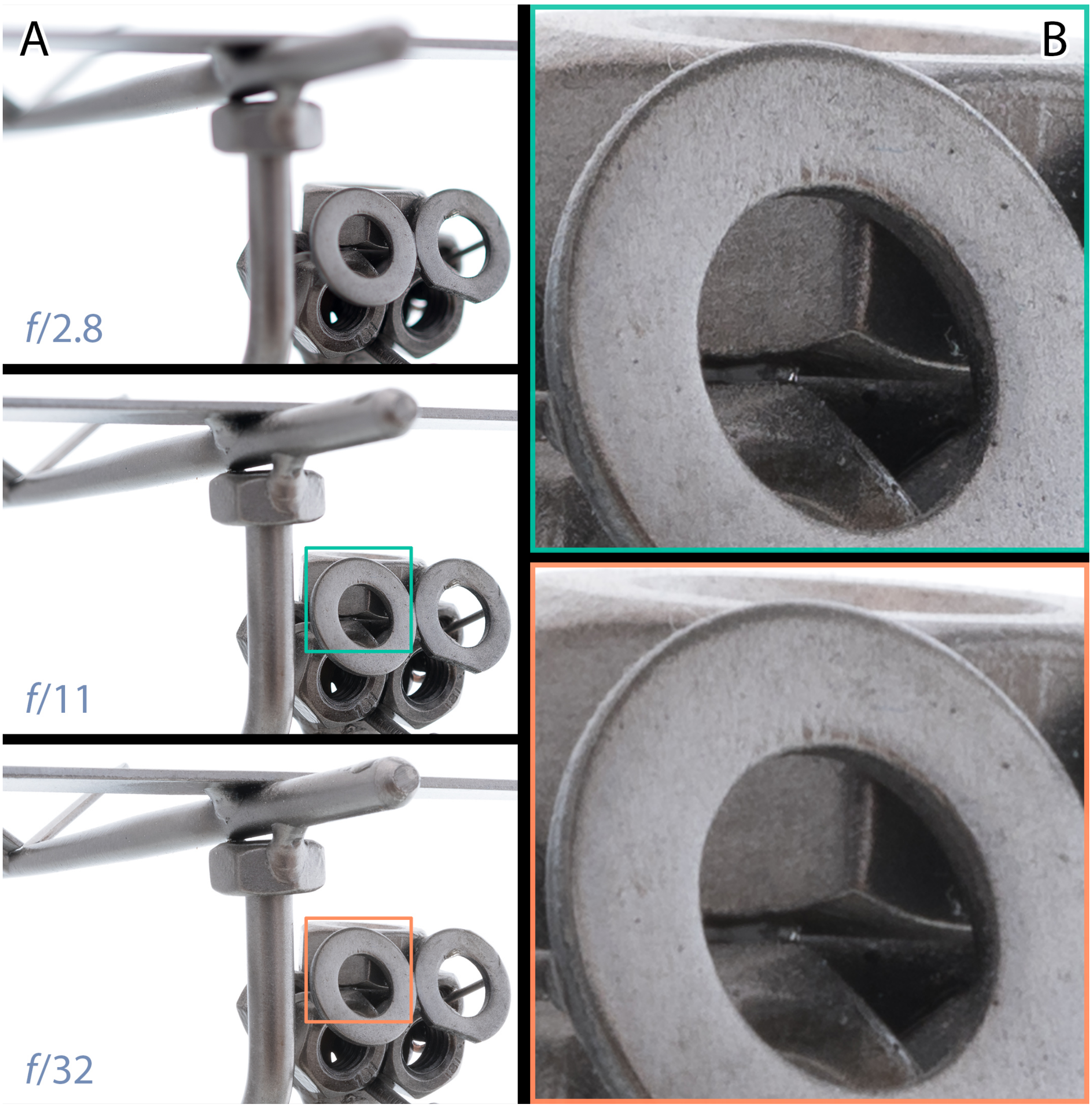


Figure 13 The insets of column A illustrate the influence of aperture on depth of field (check the wings and tail of the glider held by the figurine), while the insets of column B reveal how diffraction of small apertures blurs image details.

The reverse of equation <20> is also true: two points may not be resolvable when the detector pitch is too big. For example: there are simply insufficient samples generated by 4 µm photosites to capture the radii of *f*/2.8 Airy discs. Although this issue can be remedied with smaller photosites, it will lower the detector pitch diffraction threshold in turn. That is why Figure 12 treated both sides of equation <19> as imposed 'limits', and a well-designed imaging system must strike a balance between them.

Here, we also enter some fascinating territory. If we would like to quantify the limitation imposed by the detector sampling, should we:

- merely use the inverse of equation <20>, which states the condition when optical diffraction is expected to worsen the image details visually:

$$p_{\text{ideal}} = \frac{1}{2.44} r_{\text{Airy}} = 0.41 r_{\text{Airy}} \ [\mu\text{m}],$$ <22>

- or use the inverse of the number of pixels that are needed to resolve a feature (i.e. three, yielding 1/3 or 0.33), in agreement with equations <10> and <19>:

$$p_{\text{ideal}} = \frac{1}{3} r_{\text{Airy}} = 0.33 r_{\text{Airy}} \ [\mu\text{m}]$$ <23>

The difference between the two highlights one of the major difficulties when assessing and talking about spatial resolving power: detecting is not recognising and vice versa. We can typically spatially detect before we can spatially resolve, and we usually must spatially resolve an object before we can identify it. Remember Figure 6 of part 1: a circular object of 50 cm diameter could be detected using a GSD of 25 cm (i.e. two pixels per given object), but a GSD of 16.7 cm or even 12.5 cm (i.e. three or four pixels per object) were needed to more or less resolve and identify it.

This is a fundamental distinction that often gets lost in the remote sensing and photographic literature, as is the fact that all quantifications up to now are theoretical and simplified! The real world is much more complicated; images include noise, intervening atmosphere and motion blur, while the final spatial image resolution also depends on the form and contrast of the scene objects. Some of those factors will be considered in the third and last entry on spatial resolution. That part will delve into the world of spatial frequencies, bar patterns and modulations, and illustrate how these concepts can characterise imaging systems in ways none of the previous equations was able to do.

That being said, there is still one important message I would like you to take home: if there is no need to apply tiny apertures, avoid them. For most aerial imaging situations with full format or smaller sensors, this means sticking to apertures equal or bigger than *f*/8 (i.e. *f*-numbers ≤ 8), simply because the image will otherwise suffer from unnecessary softening. On top of that, *f*/8 (together with *f*/5.6) is the aperture where most lenses optically perform at their best (their so-called 'sweet spot'). Finally, smaller apertures also reduce the shutter speed, making it easier for aircraft-induced motion blur to sneak in the image.