Resolving some spatial resolution issues – Part 1: Between line pairs and sampling distances

Geert Verhoeven

One of the most common words in the remote sensing (or even general imaging) literature is ‘resolution’. Despite its abundant use and because the concept is often misjudged as uncomplicated, most modern literature relies on rather sloppy ‘resolution’ definitions that sometimes even contradict each other within the same text. In part, this confusion and misconception arises from the fact that technical as well as broader, application-specific explanations for resolution exist, both of them even relying on different ways to describe resolution characteristics. As a result, the term ‘resolution’ has been used for many years as a handy go-to term to cover many concepts: “this satellite produces images with a resolution of 30 m”; “there is an increasing number of high-resolution camera sensors on the market” or “the resolution of the human eye is coarser than an eagle’s eye”. Nowadays, one might wonder if resolution is a particular image characteristic, a property of the imaged scene or instead related to the imaging sensor or maybe the camera’s lens.

It is thus fair to say that the technical concept of resolution – or more specifically spatial resolution – and all its implications are commonly poorly understood, which leads to many popular, accepted but completely wrong statements. In the photographic literature, a widespread example is to refer to the total number of camera image pixels (i.e. the **pixel count**) as the image resolution of that specific digital camera. This is erroneous since the same 24-megapixel camera can capture a photograph of an Attic black-figure amphora as well as a complete submerged Greek temple. The resulting two photographs, although both are counting 24 megapixels, might reveal scene details of 0.01 cm and 2 cm respectively. In the remote sensing community, a prevalent misconception is that a satellite image with a 1 m resolution automatically means that we can recognise all objects in that image which have a width equal to or larger than 1 m.

In this two-part entry of our series, we will combine simple geometrical relationships (part 1) and fundamental laws of electromagnetic radiation (part 2) to shed some light on the term spatial resolution and explain its difference with the related concept of spatial resolving power. Similar to the previous two entries, this two-pieced text can only scratch the surface of this very complex topic. Notwithstanding, the aim is still to provide solid definitions and enough background knowledge to easily correct many of the “common knowledge” but ill-founded statements such as the ones mentioned above.

1 Basic definitions, concepts and units

1.1 Resolution versus resolving power

Not all images are created equally and, as such, not all have the same archaeological potential. We have seen in entry two that a digital image is a sampled (spatially, spectrally and temporally) and quantised (defined by the number of bits) representation of a real-world scene. The prime function of the imaging system is to resolve scene detail in any of these four dimensions. This text will deal with the spatial dimension and tackle the related concepts of **spatial resolution $R_{spatial}$** and **spatial resolving power $RP_{spatial}$**. A future entry will cover the three remaining dimensions of remote sensing products.

Any imaging system consists of a multitude of imaging hardware components aside from the necessary signal-processing algorithms and electronics. These components determine the imaging system’s

spatial resolving power $\boldsymbol{RP_{spatial}}$ (often shortened to **resolving power**), or its ability to separate the electromagnetic radiation that is reflected/emitted by neighbouring object points. Although there is no consensus throughout the literature on this, there are many good reasons to reserve the term spatial resolving power solely as an evaluator for (a component of) an imaging system (and thus not an image). As such, spatial resolving power quantifies the smallest detail that a lens, a film, a digital imaging sensor or a complete imaging system like a digital camera, telescope or microscope can spatially resolve.

Many authors interchange the term spatial resolving power with the term **spatial resolution**, although both terms are distinct. The **spatial resolution** $\boldsymbol{R_{spatial}}$ (also called the **limit of spatial resolution**, **minimum resolvable distance** or **limiting spatial resolution**) refers to a minimum distance Δx between two object points that are still resolved in the image. Although the term **resolution** is often used as shorthand for spatial resolution, it is always best to use the latter term, since resolution can also refer to the **measurement resolution**, indicating the smallest value change in a measurement of any instrument. For example: the measurement resolution of a distance measured by a ruler is often 1 mm and 1/100 of a second for a time interval determined by a digital stopwatch. [Note that measurement resolution is often wrongly denoted as “precision”. Also, mind that many authors still utilise the term spatial resolution to describe the imaging system. They should then at least use the term photographic or image spatial resolution when talking about the final product to avoid any confusion].

$$R_{spatial} = \frac{1}{RP_{spatial}} \text{ [mm]} \qquad \text{<1>}$$

$$RP_{spatial} = \frac{1}{R_{spatial}} \text{ [LP/mm]} \qquad \text{<2>}$$

Spatial resolution thus indicates the level of spatial detail observed in the image. Since it quantifies a minimum resolvable distance, it is a spatial domain metric expressed in µm (micrometre), mm, cm, m or inch. **Spatial resolving power** is merely the reciprocal of spatial resolution ($\mathbf{1/\Delta x}$). Consequently, a small spatial resolution translates into a high spatial resolving power and vice versa. Being a reciprocal value, spatial resolving power has units of mm^{-1}. However, a more practical unit is Line Pairs per millimetre (LP/mm, LP mm^{-1}) or Cycles per millimetre (C/mm, C mm^{-1}), in which every line pair or cycle consists of an adjacent black and white line.

Spatial resolving power is thus a spatial frequency metric, expressing “a certain amount of something per given unit”. One can now see how we are constantly dealing with different statements of spatial resolving powers in the digital world: the spatial resolving power of a printer is expressed in DPI (Dots Per Inch), of a scanner in SPI (Samples Per Inch), and in PPI (Pixels Per Inch) for a screen (Table 1).

Acronym/Initialism	Meaning	Application	Common values
LP/mm	Line Pairs per millimetre	Lens, sensor, camera system	60 LP/mm
DPI	Dots Per Inch	Printer	300 DPI
PPI	Pixels Per Inch	Screen	110 PPI
SPI	Samples Per Inch	Scanner	3200 SPI

Table 1 Common metrics of spatial resolving power.

1.2 The human eye

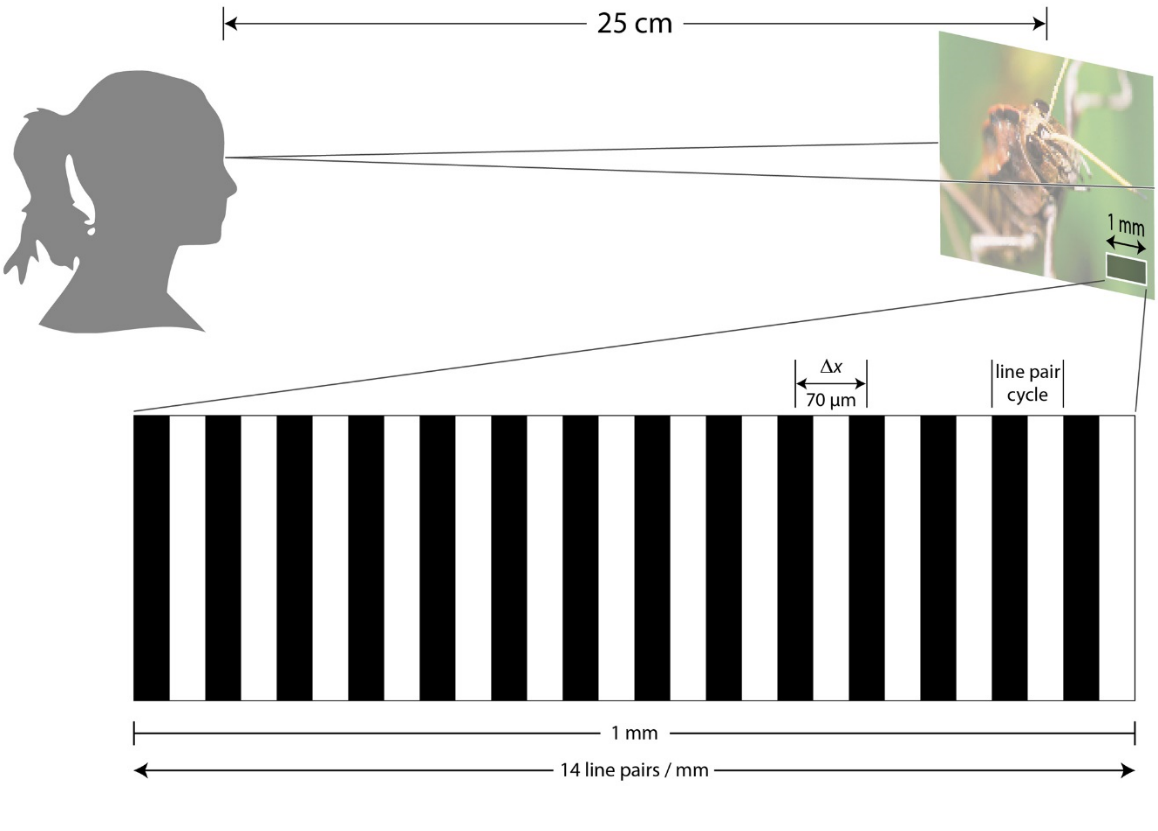
Let us resort to a very familiar optical instrument – the human eye – to illustrate this difference between spatial resolution and spatial resolving power. Although the exact value of the eyes’ resolving power depends on the testing method, a favoured test object is a grating: a chart with alternate light and dark stripes that are of equal width and parallel to each other (Figure 1). Every black or white stripe is called a line, and every black-white line combination forms a **line pair** or a **cycle**.

Under ideal circumstances and at a 25 cm reading distance, humans with excellent vision can spatially resolve about fourteen Line Pairs per mm (14 LP/mm). In other words, 14 LP/mm is the eye's spatial resolving power at a 25 cm reading distance (Figure 1). The width of every line pair quantifies then the spatial resolution of the resulting image on the retina. As mentioned before, it simply equals the reciprocal of the eyes' resolving power, yielding a distance Δx of 70 micrometres:

$$R_{\text{spatial}} = 1 / RP_{\text{spatial}} = 1 \text{ mm} / 14 \text{ line pairs} = 0.07 \text{ mm} / \text{line pair} = 0.07 \text{ mm}.$$

An often-encountered mistake throughout the literature is to quantify spatial resolution as the width of one line instead of the complete line pair. Since it is impossible to separate two black lines from each other, one must compute the linear separation Δx between the centres of the black lines with the white space in between. Because the spatial resolution metric provides a length, we can remove "line pair" from its units. [Note that some authors quantify spatial resolving power in lines per mm. Although this metric is numerically twice as high as the LP/mm statement, some authors consider them to be identical. It is thus best to stick to line pairs or cycles for quantifying spatial resolving power.]

For a human eye to tell two object points apart from 25 cm, they should thus not be spaced together more closely than 70 micrometres. Increasing the target distance to 5 m will increase the width of the lines that still can be resolved, also leading to a lower LP/mm value. To get around this, spatial resolution and spatial resolving power statements often make use of angles.



Spatial Resolving Power RP_{spatial} = 14 line pairs / mm

Spatial Resolution R_{spatial} = 1 mm / 14 line pairs = 0.07 mm / line pair = 0.07 mm

Figure 1 The ideal spatial resolving power of the human eye equals approximately 14 line pairs per mm.

1.3 Linear versus angular metrics

Besides the linear measures used so far, it is often beneficial to use angular statements. If two lines are drawn from your eye to the outer points of an object, the angle θ between these lines is termed the **angular size** of the object. From Figure 2, it is clear that the angular size θ of an object depends on its linear size *D* as well as the distance to the observer *s*. For small angles expressed in radians, this relationship can be formulated as follows:

$$\theta = \frac{D}{s} \text{ [rad]} \qquad \text{<3>}$$

To express the same angle in arcminutes, we add a radian-to-arcminutes conversion factor of 3438 (since there are 3438 arcminutes in one radian – consider Figure 2 for an overview of basic units).

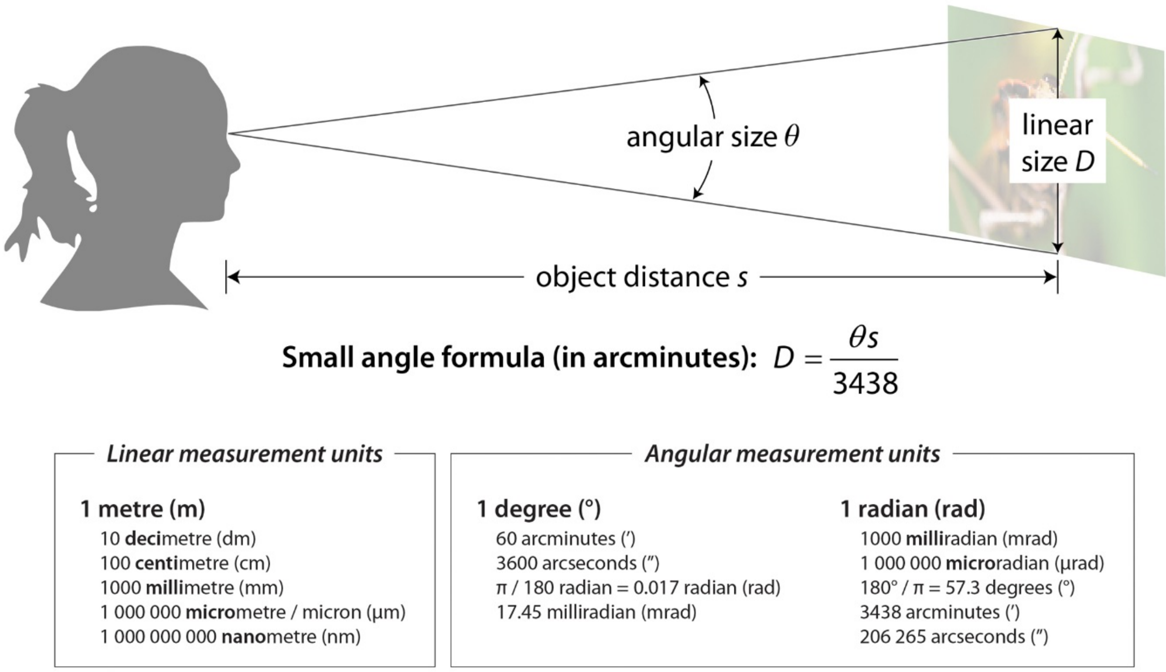
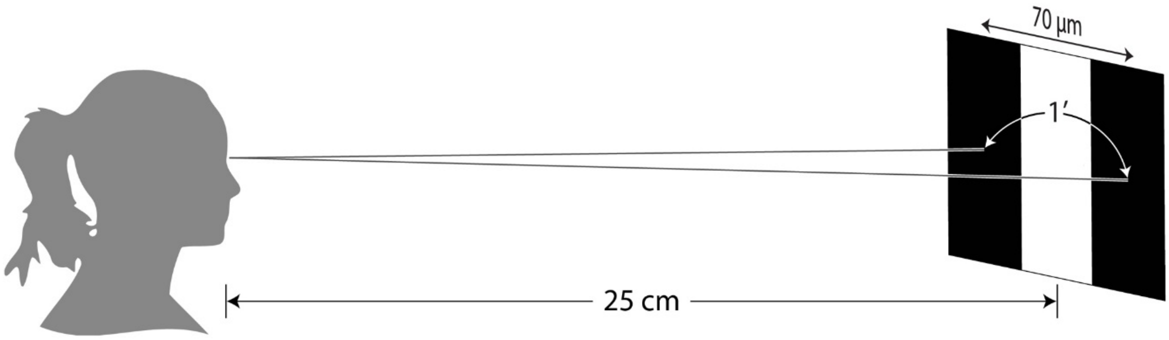


Figure 2 The angular size of an object equals the angle subtended by imaginary lines passing from the imaging system to the outer ends of that object. The figure also includes basic linear and angular measurement units.

Conversely, a given angular size can correspond to different actual linear sizes, all depending on the object distance *s*. So, although the minimum resolvable linear separation Δx depends on the distance from the eye (or any imaging sensor for that matter), its angular size does not. For that reason, it is beneficial to specify spatial resolution in terms of a **minimum resolvable angle** $\Delta\theta$, expressed in degrees, radians or milliradians. The **angular spatial resolving power** is again its reciprocal **$1/\Delta\theta$**, bearing units of line pairs per degree, cycles per (milli)radian or merely unitless reciprocal radians.

Using the human eye again as an example, its **visual acuity** as expressed by its **angular spatial resolving power** is sixty line pairs per degree (60 LP/°) or one cycle per arcminute (written as 1′ and equalling 1/60 of a degree or 60 arcseconds). Using the small-angle formula, we can compute that this angular spatial resolving power translates to a linear size of 70 µm at a viewing distance of 25 cm (Figure 3). So, when the **angular spatial resolution** of two small points or lines is less than one arcminute (in the retinal image or object space), we will not perceive them as individual entities. If their angular separation is larger, they can be distinguished by a healthy human eye. This is also the reason why we perceive out-of-focus objects as blurry. The size of the individual image points (of which this out-of-focus object is composed) all surpass 0.07 millimetres from a 25 cm viewing distance.



Angular Spatial Resolving Power $ARP_{spatial}$ = 1 line pair / arcminute

Angular Spatial Resolution $AR_{spatial}$ = 1 arcminute = 0.017°

Linear size D = 1 * 25 cm / 3438 = 0.007 cm

Figure 3 Under ideal conditions, the human eye has a visual acuity of one line pair per arcminute (or one cycle per 0.017°). This means that it is possible to separate two points that are 0.07 mm apart from 25 cm.

2 Basic geometrical relations

2.1 GSD and IFoV

To understand the factors that determine the spatial resolution of an image, it is beneficial to consider the problem from a simple geometrical point of view (Figure 4). From the previous entry on pixels, we know that an imaging sensor consists of a two-dimensional array of individual **photosites**, with the **photosite pitch *p*** equaling the distance from the centre of one photosite to the centre of an adjacent element.

The sensor is always a certain distance away from the **optical centre *O*** of the imaging system. For air- and spaceborne imaging, this distance equals the **focal length *f′*** of the lens. Focal length and photosite pitch are two variables that characterise the imaging system. Hence, they are part of the **image space**. However, the photosite pitch has a corresponding quantity outside the imaging system, in the so-called **object space**. The object-space conjugate of photosite pitch is the **Ground Sampling/Sampled/Sample Distance** or **GSD**. In non-remote sensing applications, one can resort to the more general term **Scene Sampling Distance** or **SSD**.

Figure 4 reveals that this GSD corresponds to a specific distance – measured on the surface of the imaged object – that results from projecting the photosite pitch in object space. Since it states the horizontally or vertically measured scene distance between two consecutive sample locations (and remember, every pixel is the result of one of those samples), GSD is one of the key factors that determine the final spatial resolution of an image. Even though the term is technically wrong (as explained in the previous entry), pixel size is often used as a synonym for GSD. The fact that some authors also use pixel size as a sensor characteristic, just further adds to the confusion and indicates why we should discard the term pixel size completely.

Besides GSD, there are four more relevant object space variables:

- the **footprint width** (***W***), equaling the dimension of the scene that is imaged at once by the imaging system in the absence of any motion. It is merely the object space conjugate of the width (horizontally and vertically) of the imaging sensor;

- the **Field of View** (**FoV**), the angular version of the footprint width and quantifying the angle in object space over which objects are recorded in the camera during the exposure;
- the **Instantaneous FoV** (**IFoV**) or the angle subtended by one photosite on the axis of the complete optical system. As such, it can be considered the angular version of the GSD. A synonym for IFoV is, therefore, **Detector Angular Subtense** of **DAS**;
- Figure 4 also depicts the **object distance *s***. In aerial or spaceborne imaging, this distance equals the flying height of the platform above the scene. It is the last variable we need to derive one of the most fundamental equations in all of photography, photogrammetry and remote sensing (Figure 5).

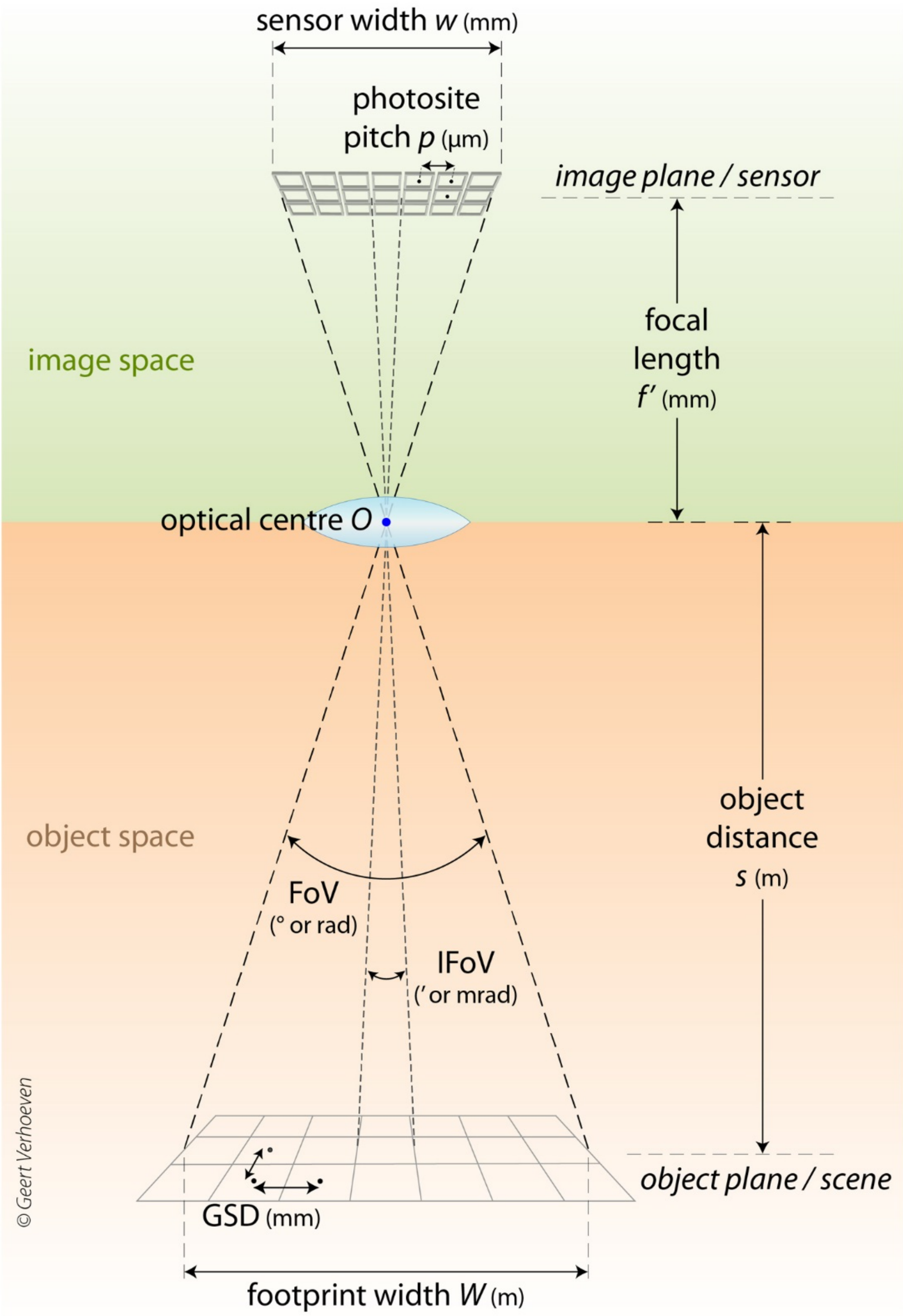


Figure 4 The basic geometrical properties of an air- or spaceborne imaging system.

2.2 Fundamental imaging equation

Using the relationships between the corresponding parts of similar triangles (i.e. triangles that have the same shape but different sizes), we can write down some very useful ratios: the photosite pitch *p* is proportional to the GSD in the same way as the sensor width *w* is related to the footprint *W* or as the focal length *f'* is relative to the object distance *s* (see figure 5). Mathematically, this becomes:

$$\frac{p}{GSD}=\frac{f'}{s}=\frac{w}{W} \qquad \text{<4>}$$

Any of those three ratios express the scale of the remotely sensed image. From here, it is trivial to express the GSD as a function of the object distance, photosite pitch and focal length:

$$GSD=\frac{sp}{f'} \text{ [m, cm, mm]} \qquad \text{<5>}$$

This straightforward formula says that the GSD is related to the detector pitch *p* by a scale factor or magnification *m*, which is the reciprocal of the image scale:

$$m=\frac{s}{f'}=\frac{W}{w} \qquad \text{<6>}$$

Although the detector pitch is a fixed sensor property, it can thus generate images with different GSDs by varying the distance between the camera and the scene (i.e. object distance *s*) and the focal length *f'* of the lens. Reducing the former or increasing the latter will lead to a smaller GSD and thus more details visible in the final image. Finally, one could also resort to a camera with reduced detector pitch *p*. [Although it might seem ideal from a spatial resolution point of view, note that a decreasing detector element spacing is not the ultimate solution since image noise will increase.]

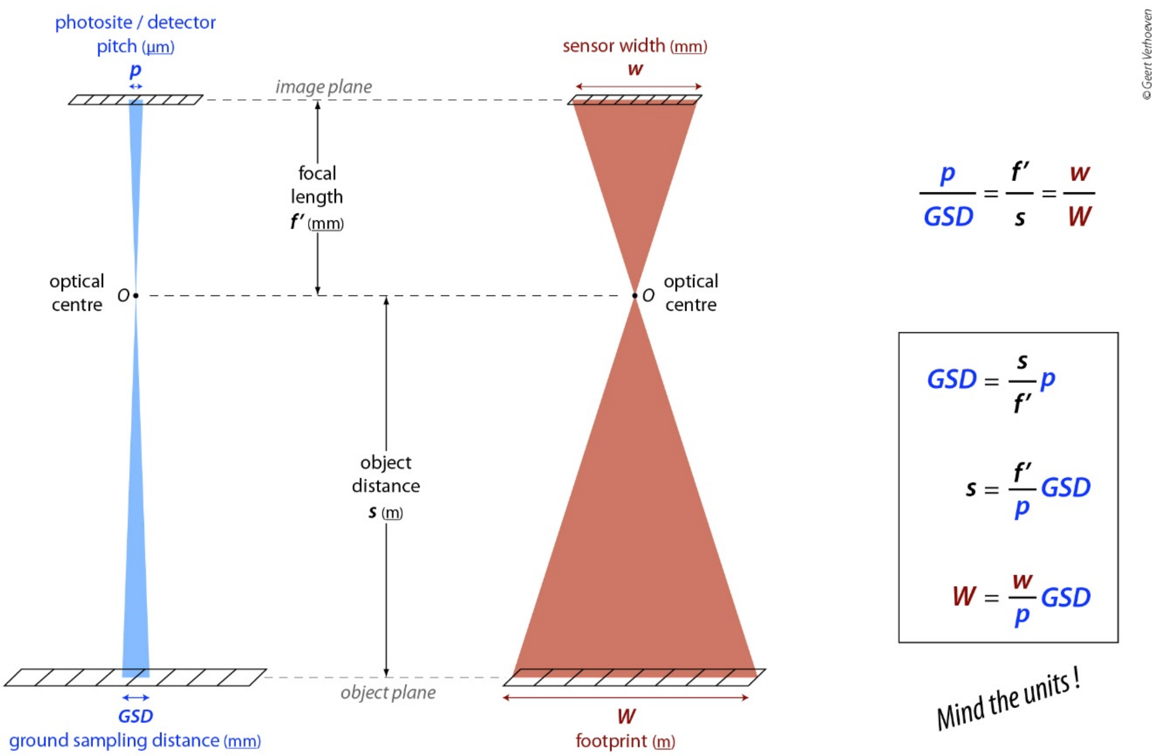


Figure 5 Deriving the purely geometry-based equation that is fundamental to all kinds of optical imaging.

To exemplify these relationships, consider an airborne Nikon D810 digital camera. Its imaging sensor measures 35.9 mm in length by 24 mm in width, and it creates images with 7360 pixels in length by 4912 pixels in width. The length of one photosite is thus 4.87 µm (i.e. 35.9 mm / 7360). Since the photosites are square, one can say that the photosite pitch *p* in both length and width equals 4.87 µm. Fitting this camera with a 50 mm lens to take aerial photographs from 350 m above a Celtic field system yields a GSD of 3.4 cm. The scale of this image is then 1/7000.

p = 35.9 mm / 7360 = 4.87 µm (0.00000487 m)
f' = 50 mm (0.05 m)
s = 350 m
GSD = 350 m * 0.00000487 m / 0.05 m = 0.034 m
Image scale = 0.05 m / 350 m = 1/7000
Scale factor or magnification *m* = 7000

Although FoV and IFoV can be considered the angular counterparts of respectively footprint width and GSD, we know from section 1.3 that these angular metrics can be stated without knowing the object distance. By equation <3>, it is possible to compute the IFoV in radians as follows:

$$IFoV = \frac{p}{f'} \text{ [rad]} \qquad \text{<7>}$$

For the previously mentioned Nikon D810 + 50 mm lens setup, this yields an IFoV of 20 arcseconds:

0.00000487 m / 0.05 m = 0.0974 mrad = 0.00558° = 20".

Since one can compute the GSD upon incorporating the object distance *s* in the formula for IFoV, GSD is also known as the **Ground-projected IFoV** or **GIFoV**. GIFoV (or GSD) is thus a linear property of an operational imaging system, obtained by projecting the detector's photosite width (or the detector pitch) on the ground.

2.3 From GSD to GRD (or the introduction of mister Nyquist)

It is essential to understand that parameters like GSD and GIFoV can only crudely approximate the smallest resolvable object in an image. Nevertheless, most of the remote sensing literature tacitly considers their value a direct expression of the size of the smallest distinguishable feature in an image. It is, for example, often assumed that an aerial photograph with a GSD of 10 cm allows distinguishing all vegetation marks that are 10 cm or larger.

However (and as covered in the previous contribution), creating an image requires a proper sampling of the scene's continuous spectral radiance in order to faithfully reconstruct the analogue signal from those digital samples. This is, however, only possible when the sampling follows the **Nyquist–Shannon sampling theorem**. Expressed by Harry Nyquist in 1928 and proven by Claude Shannon in 1949, this sampling theorem finds a very practical application in digital imaging. In this context, the sampling theorem can be stated as follows:

Assuming a perfect capturing system, an imaged scene can be reconstructed without artefacts if the original continuous spectral radiance signal did not contain frequencies at or above one-half of the sampling rate, the latter being determined by the sensor's detector pitch p.

In other words: a feature such as a line pair of size *D* will be imaged unambiguously when it is digitised by at least two pixels. The photosite pitch *p* should thus be equal to or smaller than *D*/2. For imaging sensors with square photosites, the **Nyquist frequency** (in LP/mm) can, therefore, be expressed as:

$$\text{Nyquist frequency} = \frac{1}{2p} \text{ [LP/mm]} \qquad \text{<8>}$$

where p is the photosite pitch in mm. The Nyquist frequency for the Nikon D810 imaging sensor with its photosite pitch p of 4.87 µm equals thus 102.7 LP/mm.

Nyquist frequency = 1 LP / (2 × 0.00487 mm) = 102.7 LP / mm

Digitising a line pair with exactly two image pixels would perfectly adhere to the sampling theorem and Nyquist frequency, but several tests have shown that one needs more than two pixels to digitise a line pair properly because the line pairs and the photosites are not necessarily in phase. Although six pixels per line pair would be ideal, image detail does not consist of perfect black and white stripes in normal photographic situations. Therefore, three pixels per line pair (or per feature width D) seems a good compromise. In practice, one can use the following equation to estimate the spatial resolving power of an imaging sensor based on its detector pitch p:

$$RP_{\text{spatial}} = \frac{1}{3p} \text{ [LP/mm]} \qquad \text{<9>}$$

Accordingly, the sensor inside the Nikon D810 has a theoretical resolving power of 68.5 LP/mm:

RP_{spatial} = 1 LP / (3 x 0.00487 mm) = 1 LP / 0.0146 mm = 68.5 LP/mm.

The reciprocal of this value provides the spatial resolution Δx of the image:

R_{spatial} = 1 mm / 68.5 = 0.0146 mm.

This spatial resolution value is, however, expressed in respect to the imaging sensor. To get the image's spatial resolution expressed in object space units – sometimes termed the **Ground Resolved Distance** (**GRD**) – we have to multiply the sensor-related value with the scale factor or the magnification m of that imaging setup. Computed as the ratio of the object distance s to the focal length f', the magnification is 7000 (i.e. 350 m / 0.05 m). As such, we get a GRD or spatial resolution of 102 mm.

0.0146 mm * 7000 = 102 mm

Since this value is exactly three times bigger than the GSD of that imaging setup (i.e. 34 mm – see above), one can compute the spatial resolution of the resulting image with the following formula:

$$R_{\text{spatial}}(\text{or GRD}) = 3 * GSD \text{ [m, cm, mm]} \qquad \text{<10>}$$

Thus, an aerial photograph with a GSD of 10 cm exhibits a theoretical spatial resolution that is much closer to 30 cm than to 10 cm. The same reasoning also counts the other way around: if you need to identify a feature of 50 cm in your aerial or spaceborne imagery, the GSD of these images should be around 16.7 cm or better (see Figure 6).

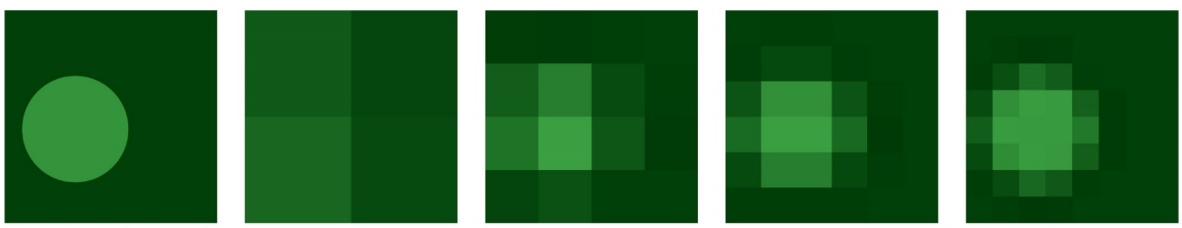


Figure 6 Sampling a scene with an object of 50 cm using different GSDs.

Despite incorporating the Nyquist sampling theorem in our coarse assessment of spatial resolving power and spatial image resolution, the quantifications mentioned above are still incorrect since they did not consider multiple other important variables:

- the part of the electromagnetic spectrum that is imaged;
- the properties of the lens system (such as aperture and aberrations);
- the noise level and bit depth of the sensor;
- the scene's geometry and contrast;
- the effect of neighbouring pixels;
- the amount and diffuseness of the illumination;
- the clarity of the atmosphere.

Although all of the above factors govern the final spatial image resolution, the first two are of the utmost importance, since they always put an absolute upper threshold on the amount of detail that can be resolved by the imaging system. We will explore this fact in the next contribution, which will represent the second part of this treatise on spatial resolution. More specifically, the second part of this text will explore the vital concepts of **point spread function** and **modulation transfer**, thereby effectively transitioning from the world of **geometrical optics** into the realm of **physical optics**.