

Optical Engineering

OpticalEngineering.SPIEDigitalLibrary.org

Equivalence theory for cross-format photographic image quality comparisons

D. Andrew Rowlands

Equivalence theory for cross-format photographic image quality comparisons

D. Andrew Rowlands*

University of Cambridge, Department of Computer Science and Technology, Cambridge, United Kingdom

Abstract. Equivalence theory is commonly used by photographers in order to obtain equivalent photographs from cameras based on different sensor formats. A complete mathematical proof of equivalence theory that is valid for compound photographic lenses with any chosen pupil magnification is presented. An expression for a correction to the conventional equivalence formulae is derived, which is proven to be formally required whenever focus is set closer than infinity, although numerically significant only for macro photography. Additionally, this paper argues that equivalence theory should be used in practice to appropriately perform cross-format image quality comparisons and demonstrates how various common metrics should be evaluated. © 2018 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.OE.57.11.110801]

Keywords: photography; Gaussian optics; cameras; noise in imaging systems; image quality assessment.

Paper 181094T received Jul. 30, 2018; accepted for publication Oct. 18, 2018; published online Nov. 16, 2018.

1 Introduction

Many photographers are familiar with the term “35-mm full-frame equivalent focal length.” For a given lens focal length on a given camera format, this is the value of the focal length on the 35-mm full-frame format that would yield the same angular field-of-view (AFoV) when taking a photo from the same position (i.e., the same perspective). For example, a 16-mm focal length on the APS-C format has a 35-mm full-frame equivalent focal length of 24 mm. This concept is useful for photographers who are now using a smaller camera format but are historically familiar with the way in which 35-mm full-frame format focal lengths relate to the expected AFoV.

The above equivalence concept can be extended further. Along with the same AFoV from equivalent focal lengths, it is also possible to determine “equivalent f -numbers” and “equivalent ISO settings,” which together yield the same depth-of-field (DoF) and same shutter speed (exposure duration) on different formats.¹⁻⁴ This leads to the concept of “equivalent photos,” which is important for three main reasons:

1. Nowadays many photographers use several cameras based on different format sizes and need a framework for translating camera settings from one format to another.

For example, consider a focal length $f = 24$ mm, f -number $N = 2.8$, and ISO setting $S = 800$ on a full-frame camera. According to equivalence theory, in order to obtain an equivalent photo using an APS-C camera, the equivalent settings are $f = 16$ mm, $N = 1.8$, and $S = 342$.

2. Equivalent photos are produced using the same amount of light. However, a larger format has the capability of collecting more light than a smaller

format. This offers additional photographic capability such as a shallower possible DoF (for the same perspective, AFoV and shutter speed). When this additional capability is utilized, the smaller format will be unable to produce an equivalent photo. Therefore, equivalence theory can be used to determine the maximum photographic capability of a smaller format in relation to a larger one.

As an extreme example, consider a mobile phone with a tiny 1/2.5 in. sensor and 5-mm focal length. If the lowest available f -number is $N = 1.4$ and the lowest available ISO setting is $S = 50$, the equivalent settings on a full-frame camera are $f = 30$ mm, $N = 8.4$, and $S = 1818$. According to equivalence theory, the shallowest DoF that the mobile phone is capable of producing corresponds to using $N = 8.4$ on the full-frame camera. Furthermore, if ISO settings lower than $S = 1818$ are used on the full-frame camera, the mobile phone will be unable to provide a sufficiently long exposure duration (shutter speed) to produce an equivalent photo. The equivalent settings on an APS-C camera are $f = 20$ mm, $N = 5.5$, and $S = 777$.

Consequently, equivalence theory can be used to help choose an appropriate camera system for a given task. For example, a travel photographer may opt to choose a lighter and smaller overall camera system based on the 1 in. or micro four thirds formats if the additional photographic capability of a larger format such as APS-C or 35-mm full-frame will rarely be utilized in practice.

3. Although image quality (IQ) metrics are typically calculated using the same exposure settings irrespective of the sensor format, such practice is inappropriate because the same exposure settings on different formats do not yield photos with the same appearance characteristics, i.e., equivalent photos. This paper argues that IQ metrics should instead be compared using equivalent camera settings so that each format

*Address all correspondence to: D. Andrew Rowlands, E-mail: andrew.rowlands@hotmail.com

produces equivalent photos and receives the same amount of light. Fundamental aspects of IQ such as total image noise will then be of the same order of magnitude (i.e., roughly the same); remaining differences in IQ will only depend upon underlying differences in the lens and camera technology used by the camera models being compared as when comparing different models based on the same format. In principle, the larger format can only provide a significant step-up in IQ when utilizing its extra photographic capability, i.e., when equivalent settings do not exist on the smaller format.

Using image noise as an illustration, consider the above example of the full-frame camera compared to the mobile phone. Equivalence theory tells us that if ISO settings lower than $S = 1818$ are used on the full-frame camera, then the mobile phone will be incapable of producing photos with noise of the same order of magnitude as the full-frame camera. The equivalent ISO setting on an APS-C camera is $S = 777$.

In order to formalize the above concepts, the remainder of the introduction discusses traditional exposure strategy (Sec. 1.1), equivalent photos (Sec. 1.2), and cross-format IQ comparisons (Sec. 1.3). Subsequently, Sec. 2 provides a mathematical proof of equivalence theory and discusses each of the properties of equivalent photos. Section 3 gives some numerical examples. Section 4 discusses IQ in relation to equivalence theory and demonstrates how various fundamental IQ metrics should be evaluated. Section 5 provides further discussion as to how equivalence theory can explain the IQ differences between mobile phone cameras and larger format cameras. Finally, conclusions are made in Sec. 6.

1.1 Traditional Exposure Strategy

Traditional exposure strategy is designed to be independent of the sensor format used by the camera. For a typical photographic scene metered using average photometry, the same combination of f -number N and ISO setting S used on any format will provide the following:

- the same average photometric exposure $\langle H \rangle$ at the sensor plane;
- the same exposure duration or “shutter speed” t ;
- an output image at the standard lightness.

Average photometry assumes the use of a hand-held light meter or a simple in-camera metering mode. As discussed in Ref. 5, a typical scene is indirectly assumed to have an average luminance that is 18% of the maximum. Provided the ISO setting is defined using the standard output sensitivity (SOS) method,^{6,7} the average luminance for a typical scene metered using average photometry will be mapped to a standard lightness in the output JPEG file, specifically 50% lightness or middle gray. This is defined as a digital output level (DOL) of 118 for an 8-bit output JPEG image encoded using the sRGB color space.

The variables N , S , and t are often collectively referred to as the “camera exposure.” Although the camera exposure is format independent, there are various aspects of the resulting image appearance that are not. When used on different

formats with the lens focused at the same object-plane distance, the same camera exposure along with the same lens focal length f will lead to an image with the following properties:

- the same perspective and
- the same magnification.

However, the following aspects of the image appearance will be different:

- framing or AFoV
- DoF.

This is because the AFoV is dependent upon both f and the dimensions of the imaging sensor, and the DoF is dependent upon the circle of confusion (CoC) diameter, which is itself format dependent.⁸

Two further image characteristics will also be fundamentally different, both of which relate to IQ:

- total image noise and
- diffraction softening.

This is because, first, the sensor area of the larger format will collect a greater amount of light compared to the smaller format for the same average photometric exposure $\langle H \rangle$. Since photon shot noise obeys Poisson statistics and scales as the square root of the total amount of light collected, it follows that the larger format will produce a less noisy image since the overall signal-to-noise ratio (SNR) will be higher. Second, the image captured by the smaller format will suffer greater diffraction softening than the image captured by the larger format because the lens entrance pupil diameter will be smaller.

In summary, the use of the same camera exposure on different formats will not yield images that have the same appearance characteristics. One important consequence is that it is inappropriate to perform cross-format IQ comparisons using the same camera exposure on different formats.

1.2 Equivalent Photos

Discussions of equivalence theory have appeared in books^{3,5} and various online articles.^{1,2,4} Equivalent photos are defined as photos taken using cameras based on different sensor formats that have the following characteristics:^{3,4}

- (1) the same perspective;
- (2) the same framing (or AFoV);
- (3) the same display dimensions;
- (4) the same DoF;
- (5) the same shutter speed; and
- (6) the same lightness.

It is crucial to point out that these are all image attributes that depend only upon sensor format and are independent from the underlying camera and lens technology.^{3,4} In other words, equivalent photos are not expected to be identical. A detailed discussion of the six attributes of equivalent photos will be given in Sec. 2.

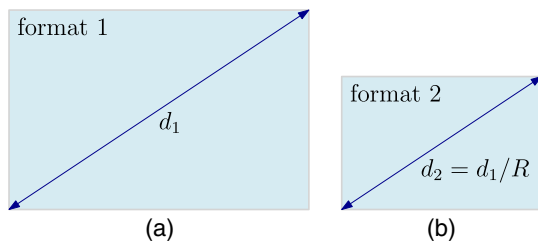


Fig. 1 The equivalence ratio R is defined as $R = d_1/d_2$. (a) Format 1 and (b) format 2.

In order to define how to take equivalent photos, of fundamental importance is the equivalence ratio R , which is the ratio between the lengths of the sensor diagonals:⁴

$$R = \frac{d_1}{d_2}.$$

Here, d_1 is the diagonal length of the larger sensor, and d_2 is the diagonal length of the smaller sensor, as illustrated in Fig. 1. For the special case that the larger format is 35-mm full-frame, the equivalence ratio is simply the traditional focal-length multiplier, informally referred to as “crop factor.” Note that precise equivalence cannot be achieved if the sensor formats have different aspect ratios.

Equivalent photos can be produced by different formats provided an “equivalent” combination of focal length f , f -number N , and ISO setting S is used on each respective format. Provided the object plane upon which focus is set is positioned beyond macro object distances:

- equivalent focal lengths are related through R ;
- equivalent f -numbers are related through R ; and
- equivalent ISO settings are related through R^2 .

Again using subscripts to denote formats 1 and 2, where 1 is the larger format, the relevant equations are as follows:

$$f_2 = \frac{f_1}{R}, \quad (1)$$

$$N_2 = \frac{N_1}{R}, \quad (2)$$

$$S_2 = \frac{S_1}{R^2}. \quad (3)$$

It will be shown in Sec. 2.2 that R must formally be replaced by the “working” equivalence ratio R_w , when focus is set on an object plane positioned closer than infinity.⁵ The replacement is required only for equivalent focal lengths and equivalent f -numbers and is not required for equivalent ISO settings. However, R can be used in general photographic situations because the practical significance of the replacement turns out to be negligible beyond macro object-plane distances.

As an example, let camera system 1 be based on the 35-mm full-frame format. In this case, Fig. 2 lists the corresponding value of R for a selection of smaller formats. Using the above equations, two examples of combinations of f , N , t , and S that result in an equivalent photo are listed in

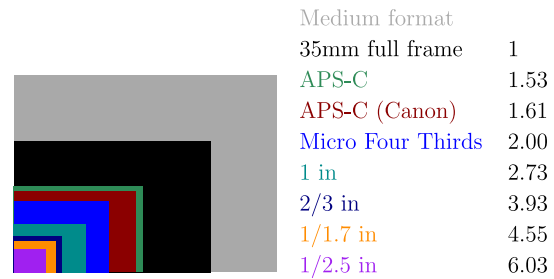


Fig. 2 The relative sizes of common sensor formats. Arbitrarily taking format 1 to be 35-mm full frame, the corresponding value of the equivalence ratio R is shown for each smaller format 2. The 1/1.7 and 1/2.5 in. sensor formats are commonly used in compact and mobile phone cameras.

Table 1 Example 1: combinations of f , N , t , and S , which produce an equivalent photo.

Format	f (mm)	N	t (s)	S (ISO)
35-mm full frame	24	2.8	1/100	800
APS-C	16	1.8	1/100	342
Micro four thirds	12	1.4	1/100	200
1 in.	9	1.0	1/100	108
2/3 in.	6	0.7	1/100	52
1/1.7 in.	5	0.6	1/100	39
1/2.5 in.	4	0.5	1/100	22

Table 2 Example 2: combinations of f , N , t , and S , which produce an equivalent photo.

Format	f (mm)	N	t (s)	S (ISO)
35-mm full frame	200	8	1/200	3200
APS-C	131	5.2	1/200	1368
Micro four thirds	100	4	1/200	800
1 in.	73	2.9	1/200	430
2/3 in.	51	2	1/200	207
1/1.7 in.	44	1.8	1/200	154
1/2.5 in.	33	1.3	1/200	88

Tables 1 and 2. The focal lengths and ISO settings have been rounded to the nearest integer, and the f -numbers have been rounded to one decimal place. The shutter speed t has been chosen arbitrarily in these examples since the required shutter speed depends upon the nature of the scene luminance distribution.

A real-world online demonstration that equivalent photos have the same DoF can be found in Ref. 1.

1.3 Cross-Format IQ Comparisons

In Sec. 1.1, it was pointed out that the same camera exposure used on different formats leads to images with different levels of noise and diffraction softening, the advantage belonging to the larger format. By contrast, equivalent photos have the following properties:

- total image noise of the same order of magnitude and
- same level of diffraction softening

As proven in Sec. 2, these properties arise from the fact that equivalent photos are produced by using the same lens entrance pupil diameter on each format instead of the same camera exposure.

The entrance pupil defines the flux entering the camera system and is typically the virtual image of the aperture stop seen through the front of the lens.⁹ Within Gaussian optics, the entrance pupil diameter D is given by $D = f/N$, where f is the front (anterior) effective focal length.⁵ Use of the same entrance pupil diameter automatically corresponds with the same level of diffraction softening, as discussed further in Sec. 4.6. Since equivalent photos are produced by using the same exposure duration (shutter speed) on each format, the total light received by each format will also be the same. It follows that equivalent photos will have total image noise of the same order of magnitude because generally the largest contribution to total image noise is photon shot noise, and this is proportional to the amount of light used to form the image.

Nevertheless, real world IQ differences (including total image noise) will inevitably occur in practice even when equivalent photos are taken. These will arise due to differences in the underlying camera and lens technology, such as:

- sensor quantum efficiency;
- read noise;
- sensor pixel count;
- lens aberrations;
- JPEG tone curve; and
- image processing.

In other words, since the total light received by each format is the same when equivalent photos are taken, it is factors such as those above that explain real-world cross-format IQ differences rather than format size. These factors will be discussed further in Sec. 4.

Although real-world IQ differences could favor any of the cameras being compared when equivalent photos are taken, the advantage of a larger format is that it offers extra photographic capability over a smaller format. This extra capability corresponds to situations in which the required equivalent camera exposure does not exist on the smaller format due to limitations in available f -number or ISO setting. Consequently, the smaller format is unable to provide the entrance pupil diameter and shutter speed required to match the larger format and produce an equivalent photo. For a given scene luminance distribution, the range of common entrance pupil diameters and shutter speeds available to two different formats can be termed the “equivalence overlap” between them.⁵ The equivalence overlap in turn

defines the range of available equivalent camera exposures. The larger the disparity in size between two formats, the smaller the equivalence overlap.

When the extra photographic capability offered by the larger format is utilized, the resulting photo will be produced using a greater amount of light than that achievable using the smaller format, and this in turn leads to appearance characteristics, such as shallow DoF or long-exposure motion blur, which cannot be achieved using the smaller format. Significantly, this potentially offers an IQ advantage in terms of SNR and resolving power (RP).

Whether or not a photographer is able to make use of the extra photographic capability offered by a larger format is an important factor to consider when choosing an appropriate camera system. The situations where the extra photographic capability can be utilized generally fall into two categories:

- (1) the entrance pupil diameter is set larger than that achievable using the smaller format
- (2) the exposure duration is set longer than that achievable using the smaller format.

As an example of the first type, consider a scenario in which the larger format is being used to take an action photo using a low f -number to isolate the subject from the background and provide a short exposure duration (fast shutter speed) in order to freeze the appearance of the moving subject. If an equivalent photo is attempted using the smaller format but an equivalent f -number does not exist, the smaller format will produce a photo with a deeper DoF and will be forced to underexpose in order to match the shutter speed used on the larger format. In this case, the extra exposure utilized by the larger format will, in principle, lead to a higher SNR. Furthermore, the image produced by the larger format will suffer less diffraction softening because the lens entrance pupil diameter will be larger.

As an example of the second type, consider the scenario where the larger format is being used to take a landscape photo with the camera set at the base ISO setting. If an equivalent photo is attempted using the smaller format but an equivalent ISO setting does not exist, the smaller format will not be able to produce a photo with a sufficiently long exposure duration (slow shutter speed) without overexposing the image. In other words, the smaller format will be unable to provide sufficient long-exposure motion blur. In this case, the extra exposure utilized by the larger format will again in principle lead to a higher SNR.

In summary, cross-format IQ comparisons should be carried out using equivalent camera settings over the equivalence overlap between the formats being compared. In this case, IQ for both formats is expected to be of the same order of magnitude. However, the extra photographic capability offered by the larger format should also be demonstrated beyond the equivalence overlap. In this regime, the larger format potentially offers higher IQ in terms of RP and SNR.

2 Formalism

Equivalence theory has previously been justified using approximate proofs that assume focus is set at infinity and use simplified formulae for the AFoV and DoF. Consequently, such approximate proofs fail to take into

account the fact that the relationship between equivalent focal lengths and between equivalent f -numbers is actually dependent on the distance to the object plane upon which focus is set.

A rigorous proof of equivalence theory has recently been given in Ref. 5. The proof is valid for compound photographic lenses with any chosen pupil magnification and with focus set at any chosen object-plane distance. The proof yields correction terms to the infinity-focus equivalence formulae. It will be shown in Sec. 3 that these correction terms become important at high magnifications. The details of the proof are expanded upon in the present section and are organized according to the following structure:

- Section 2.1 explains the condition for producing an image with the same perspective from different formats, specifically the requirement that the object-plane distance be the same when measured from the entrance pupil.
- Section 2.2 derives the condition for producing an image with the same framing from different formats, specifically a formula relating the equivalent focal lengths required. It is shown that when focus is set closer than infinity, the “working” equivalence ratio R_w formally replaces R . Section 2.2.1 derives practical expressions for R_w .
- Section 2.3 introduces the important concept of the CoC and discusses the requirement that equivalent photos be viewed at the same display dimensions.
- Section 2.4 derives the condition for producing an image with the same DoF (at the same perspective and framing) from different camera formats, specifically a formula relating the equivalent f -numbers required. It is shown that the equivalent f -numbers are related by the working equivalence ratio R_w . It is also proven that the same entrance pupil diameter is required on each format.
- Section 2.5 discusses the requirement that equivalent photos from different formats must be produced using the same exposure duration or shutter speed.
- Section 2.6 derives the condition for producing an output image with the same lightness from different camera formats, specifically a formula relating the equivalent ISO settings required. It is shown that these are always related by R rather than R_w because ISO settings are by construction independent of the distance to the object plane. It is proven that the “working” f -numbers are related by R , which is consistent with the fact that equivalent photos are produced using the same total amount of light.
- Section 2.7 summarizes the equivalence equations.

2.1 Same Perspective

For cameras based on different formats focused on a specified object plane, the same perspective requires the same object-plane distance measured from the lens entrance pupil of each camera.⁹

For generality, consider a compound photographic lens with any valid pupil magnification m_p . The pupil magnification is defined as follows:

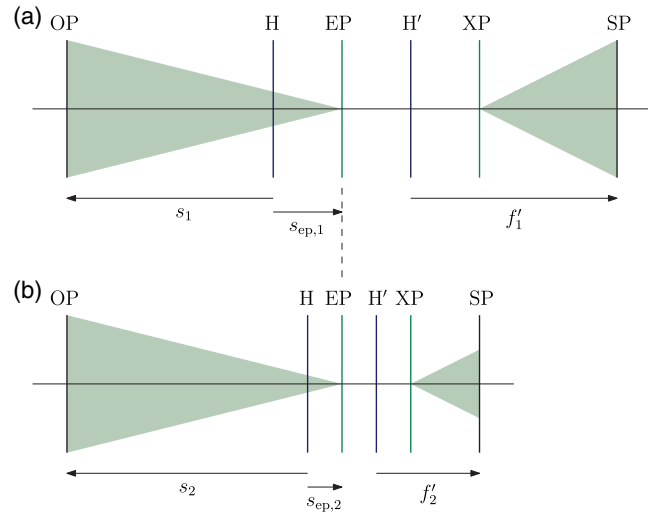


Fig. 3 For a general pupil magnification, (a) the vector distance $s_1 - s_{ep,1}$ for the larger format and (b) vector distance $s_2 - s_{ep,2}$ for the smaller format must be equal in order to achieve the same perspective and framing. In the diagram, OP is the object plane, H and H' are the first and second principal planes of the compound lens, EP is the entrance pupil, XP is the exit pupil, SP is the sensor plane, and f' is the rear (posterior) effective focal length. The principal planes and pupils are not required to be in the order shown. Figure reproduced from Ref. 5 courtesy of IOP Publishing Ltd.

$$m_p = \frac{D_{xp}}{D}, \quad (4)$$

where D_{xp} and D are the diameters of the exit and entrance pupils, respectively. When m_p differs from unity, the entrance and exit pupils will be located away from the principal planes of the compound lens. Precise equivalence between different formats is possible with focus set at any chosen object-plane distance only if the lens designs have the same symmetry and therefore the same pupil magnification. Such a scenario is illustrated graphically in Fig. 3. The sign convention has been adopted such that distances in front of H are positive and distances behind H are negative.

Let s_1 denote the distance from the first principal plane H to the object plane for format 1, and let $s_{ep,1}$ denote the distance from H to the entrance pupil of format 1. Analogously, let s_2 denote the distance from H to the object plane for format 2, and let $s_{ep,2}$ denote the distance from H to the entrance pupil of format 2. Since the total distance from the entrance pupil to the object plane must be the same for both cameras in order that the perspective (and framing) be the same, the following condition must hold:⁵

$$s_1 - s_{ep,1} = s_2 - s_{ep,2}. \quad (5)$$

This result will be utilized later in the proof.

2.2 Same Framing: Equivalent Focal Lengths

Recall the standard expression for the AFoV:

$$\alpha = 2 \tan^{-1} \left(\frac{d}{2bf} \right). \quad (6)$$

For completeness, a derivation of this formula is given in Sec. 7 as an appendix. Note that the apex of the AFoV is

situated at the lens entrance pupil.⁹ Here, f is the front (anterior) effective focal length,⁵ and d is the sensor length yielding the corresponding AFoV α measured in either the horizontal, vertical, or diagonal direction. The quantity b is the so-called “bellows factor,” which depends upon both $|m|$ and m_p :

$$b = 1 + \frac{|m|}{m_p}. \quad (7)$$

A useful practical expression for $|m|$ can be derived (see Sec. 7) as follows:

$$|m| = \frac{f}{s - f}. \quad (8)$$

At infinity focus, $s \rightarrow \infty$ and so $|m| \rightarrow 0$ and $b \rightarrow 1$. At closer focus distances (i.e., as the object plane upon which focus is set is brought forward from infinity) and assuming a traditional-focusing lens, the value of b gradually increases from unity. For a fixed framing, the AFoV therefore becomes smaller. Consequently, the object appears to be larger than expected, particularly at close-focusing distances.

The proof of equivalence given in this section is based upon a traditional-focusing lens. Such lenses achieve focus by movement of the whole lens barrel and not by altering their focal length. On the other hand, lenses that utilize front-cell or internal focusing may alter their focal length depending on the object-plane distance, and this can result in very different AFoV behavior as a function of object-plane distance to that described above. Nevertheless, the proof of equivalence remains valid for such lenses provided the “new” value for the focal length f is used in the bellows factor and AFoV formulae after focus has been set. In all cases, s is the object-plane distance measured from the first principal plane after focus has been set. A more detailed discussion of focus breathing is given in Sec. 8 as an appendix.

In order to derive the formula relating the equivalent focal lengths required on different formats in order to achieve the same framing or AFoV, consider format 1 with a sensor diagonal d and lens with front effective focal length f_1 focused at an object-plane distance s_1 measured from the first principal plane. Assuming a traditional-focusing lens, according to Eq. (6), the AFoV and bellows factor are as follows:

$$\alpha_1 = 2 \tan^{-1} \frac{d}{2b_1 f_1}, \quad b_1 = 1 + \frac{|m_1|}{m_p}. \quad (9)$$

Now, consider format 2 with a smaller sensor diagonal d/R , where R is the equivalence ratio introduced in Sec. 1.2. Assume the lens has front effective focal length f_2 and is focused on the same object plane positioned a distance s_2 from the first principal plane. Again, assuming a traditional-focusing lens, the AFoV and bellows factor in this case are as follows:

$$\alpha_2 = 2 \tan^{-1} \frac{d}{2Rb_2 f_2}, \quad b_2 = 1 + \frac{|m_2|}{m_p}. \quad (10)$$

The requirement that the two systems have the same AFoV demands that $\alpha_1 = \alpha_2$ and therefore

$$b_1 f_1 = R b_2 f_2. \quad (11)$$

Rearranging yields the following condition for equivalent focal lengths corresponding to the same AFoV on each format:

$$f_2 = \frac{f_1}{R_w}. \quad (12)$$

The “working” equivalence ratio⁵ denoted by R_w has been defined by

$$R_w = \left(\frac{b_2}{b_1} \right) R. \quad (13)$$

It is important to realize that $b_1 \neq b_2$ when focus is set on an object plane positioned closer than infinity. This is because the total refractive power Φ of a photographic lens is defined as the reciprocal of the effective focal length, $\Phi = 1/f_E$, where $f_E = f/n$ and $n \approx 1$ when the object-space medium is air. When equivalent photos are taken using cameras based on different formats, the equivalent front effective focal lengths f_1 and f_2 are not identical and therefore do not correspond to the same refractive power. Consequently, the systems have different bellows factors at the same perspective or object distance $s - s_{ep}$. This is evident from Eq. (8) when considering the examples of equivalent focal lengths given in Sec. 1.2 earlier.

At infinity focus, $b_1 \rightarrow 1$ and $b_2 \rightarrow 1$, and so, $R_w \rightarrow R$. Practical expressions for R_w are derived below.

2.2.1 Working equivalence ratio: practical expressions

Practical expressions for the working equivalence ratio R_w require rewriting Eq. (13) in terms of object distances and focal lengths rather than bellows factors.

Recall from Sec. 2.1 and Fig. 3 that $s_{ep,1}$ and $s_{ep,2}$ are the separations between the first principal plane and entrance pupil for each respective format. From Eq. (64) of Sec. 7, it follows that

$$s_{ep,1} = \left(1 - \frac{1}{m_p} \right) f_1, \quad s_{ep,2} = \left(1 - \frac{1}{m_p} \right) f_2. \quad (14)$$

Now, by utilizing Eq. (8), the bellows factors can be expressed in the following way:

$$b_1 = \frac{s_1 - s_{ep,1}}{s_1 - f_1}, \quad b_2 = \frac{s_2 - s_{ep,2}}{s_2 - f_2}. \quad (15)$$

Since the first requirement of equivalent photos is the same perspective, the result from Sec. 2.1 defined by Eq. (5) can be utilized, $s_1 - s_{ep,1} = s_2 - s_{ep,2}$. Combining this with Eq. (15) and substituting into Eq. (13) leads to a more explicit expression for R_w :

$$R_w = \left(\frac{s_1 - f_1}{s_2 - f_2} \right) R. \quad (16)$$

If the lens focal length of the larger format is known (format 1) and the equivalent lens focal length of the smaller format is required (format 2), then the format 2 terms s_2 and f_2 need to be eliminated from Eq. (16). With the

help of Eq. (5), algebraic manipulation leads to the following result:

$$R_w = \left(1 - \frac{m_{c,1}}{p_{c,1}}\right)R. \quad (17)$$

The correction $m_{c,1}$ arises due to the differing system magnifications, and the correction $p_{c,1}$ arises for a nonunity pupil magnification. These corrections are defined by

$$m_{c,1} = \left(\frac{R-1}{R}\right)\frac{f_1}{s_1}, \quad p_{c,1} = m_p + (1 - m_p)\frac{f_1}{s_1}. \quad (18)$$

Alternatively, if the lens focal length of the smaller format is known (format 2) and the equivalent lens focal length of the larger format is required (format 1), then the format 1 terms s_1 and f_1 need to be eliminated from Eq. (16). Again with the help of Eq. (5), algebraic manipulation leads to the following result:

$$R_w = \frac{R}{1 + \left(\frac{m_{c,2}}{p_{c,2}}\right)R}. \quad (19)$$

In analogy with above, the correction $m_{c,2}$ arises due to the differing system magnifications, and the correction $p_{c,2}$ arises for a nonunity pupil magnification. These are defined by

$$m_{c,2} = \left(\frac{R-1}{R}\right)\frac{f_2}{s_2}, \quad p_{c,2} = m_p + (1 - m_p)\frac{f_2}{s_2}. \quad (20)$$

For the purpose of determining equivalent focal lengths, which yield the same AFoV on different formats, the working equivalence ratio R_w provides a correction to R when focus is set on an object plane positioned closer than infinity. The correction vanishes when $R = 1$ or when $s_1, s_2 \rightarrow \infty$, in which case, R_w reduces to R .

For the special case of a symmetric lens design with $m_p = 1$, the separation terms defined by Eq. (14) vanish and the terms $p_{c,1}$ and $p_{c,2}$ are both unity. In this case, the object distances measured from the first principal plane will be equal for each format when equivalent photos are taken, $s_1 = s_2 = s$.

2.3 Same Display Dimensions

When viewing a photo, the level of detail resolved by an observer affects the perception of properties such as DoF. If a photo is viewed at a specified distance by an observer whose visual system has a known RP, the details resolved will depend upon the size of the photo, i.e., the enlargement factor from the optical image captured by the imaging sensor. Therefore, a fundamental requirement of equivalence theory is that equivalent photos be viewed at the same distance and same display dimensions.

Recall that in order to quantify observer RP in photography, the ability of the eye to resolve detail is defined using a pattern of line pairs.¹⁰ Each line pair consists of a vertical black stripe and vertical white stripe of equal width. As the width of the lines decreases, the stripes eventually become indistinguishable from a gray block. The least resolvable separation (LRS) measured in mm per line pair

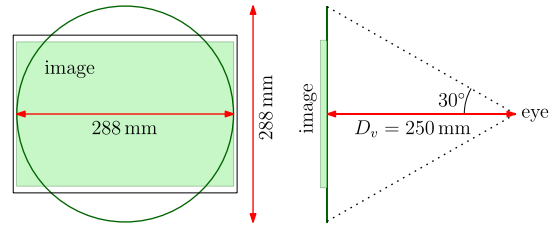


Fig. 4 A viewing circle of ~60 deg diameter defines the limits of near peripheral vision. This is shown in relation to a 3 × 2 image (green) printed on A4 (210 × 297 mm) paper (black border) viewed at the least distance of distinct vision, D_v . The cone of vision roughly forms a circle of diameter $D_v \tan 30^\circ = 288$ mm. Figure reproduced from Ref. 5 courtesy of IOP Publishing Ltd.

is the minimum distance between the centres of neighboring white stripes or neighboring black stripes when the pattern can still just be resolved by the eye. Observer RP is the reciprocal of the LRS and is measured in line pairs per mm (lp/mm). Observer RP depends upon the viewing distance. At the least distance of distinct vision D_v , which is defined as 10 in. or 25 cm, a value of around 5 lp/mm is often assumed.¹¹ However, RP varies considerably depending on the ambient conditions and the visual acuity of the individual.

The limits of near peripheral vision are defined by the 60-deg cone of vision, as illustrated in Fig. 4. If it is assumed that the width of the viewed photo just fits within the limits of near peripheral vision when viewed from D_v , then the enlargement factor x from a 35-mm full-frame sensor will be 8. In this case, the observer RP projected down to the sensor dimensions becomes 40 lp/mm. Mathematically,

$$RP(\text{on sensor}) = RP(\text{print viewed at } D_v) \times x, \quad (21)$$

where x is the enlargement and “print” refers to an image either printed or viewed on a display. The value RP (on sensor) does not refer to the actual capability or RP of the camera system itself. Nevertheless, it is clear that a 35-mm full-frame format camera system does not need to resolve a line pair of width less than the value $RP(\text{on sensor}) = 40$ lp/mm. This detail cannot be resolved by the observer of the output image when it has been enlarged by a factor of 8 and viewed at the least distance of distinct vision, D_v . These are typical viewing conditions assumed by the camera or lens manufacturer.¹¹

An equivalent viewpoint is that the value RP (on sensor) affords a certain amount of defocus blur that is undetectable to the observer to be present in the print. The allowed defocus blur can be treated rigorously by calculating the defocus point-spread function (PSF) using wave optics. However, for simple photographic calculations, the defocus blur is instead treated in a purely geometrical manner by assuming that the blur will be uniform over a circle that approximates the shape of the lens aperture. (A Gaussian function would be a more realistic model of the true geometric PSF.⁹) The size of the circle restricts the blur radius, and convolving this blur circle with the optical image formed on the sensor yields a blurred optical image. The circle on the sensor that affords the largest amount of undetectable blur in the print when it is viewed from a specified distance and at specified display dimensions, is known as the acceptable CoC, or simply the CoC.⁸

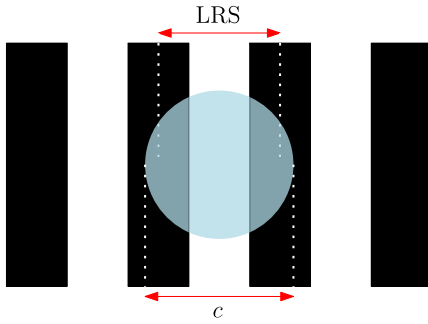


Fig. 5 The required CoC diameter on the sensor is slightly wider than the LRS between neighboring like stripes. A convolution of the CoC with the line pattern renders the stripes unresolvable to an observer of the print under the specified viewing conditions. In other words, the CoC describes the amount of defocus blur that can be tolerated before details in the print begin to appear “out of focus” to the observer. Figure reproduced from Ref. 5 courtesy of IOP Publishing Ltd.

By treating the CoC as a geometrical PSF and then calculating the cut-off frequency, the relationship between the value $RP(\text{sensor dimensions})$ defined by Eq. (21) and the corresponding CoC diameter c can be shown as follows:⁵

$$c = \frac{1.22}{RP(\text{on sensor})} = \frac{1.22}{RP(\text{print viewed at } D_v) \times x}. \quad (22)$$

This is illustrated graphically in Fig. 5. If $RP(\text{on sensor}) = 40 \text{ lp/mm}$ for a 35-mm full-frame camera, then the corresponding CoC diameter will be $c = 0.030 \text{ mm}$. Smaller or larger sensors require a smaller or larger diameter c , respectively, because the enlargement factor in Eq. (21) will change accordingly. To see this, consider format 1 with a sensor diagonal d and format 2 with a sensor diagonal d/R , where R is the equivalence ratio. Provided equivalent photos obtained from these formats are viewed from the same distance and at the same display dimensions, the enlargement factor appearing in Eq. (21) will be a factor R larger for format 2. Now, combining Eqs. (21) and (22) for both formats yields the following important result:

$$c_2 = \frac{c_1}{R}, \quad (23)$$

where c_1 and c_2 are the CoC diameters for formats 1 and 2, respectively. Table 3 lists example CoC diameters assumed by camera manufacturers for various formats.

It should be noted that a photographer can define a custom CoC based upon a known viewing distance L and enlargement x rather than those assumed by the manufacturer. The custom CoC diameter is obtained as follows:

$$\text{custom CoC} = c \times \frac{L}{D_v}, \quad (24)$$

where c is defined by Eq. (22). The value $RP(\text{print viewed at } D_v) = 5 \text{ lp/mm}$ can also be adjusted according to the viewing conditions and the visual acuity of the observer, if known.

Table 3 Example CoC diameters for various sensor formats.

Sensor format	Sensor dimensions (mm)	CoC (mm)
35-mm full frame	36.0 × 24.0	0.030
APS-C	23.6 × 15.6	0.020
APS-C (Canon)	22.3 × 14.9	0.019
Micro four thirds	17.3 × 13.0	0.015
1 in.	13.2 × 8.8	0.011
2/3 in.	8.8 × 6.6	0.008
1/1.7 in.	7.6 × 5.7	0.006
1/2.5 in.	5.8 × 4.3	0.005

2.4 Same Depth-of-Field: Equivalent f -Numbers

Recall the standard DoF equations:

$$\text{near DoF} = \frac{c(s - s_{ep})}{|m|D + c}, \quad (25)$$

$$\text{far DoF} = \frac{c(s - s_{ep})}{|m|D - c}, \quad (26)$$

$$\text{total DoF} = \frac{2|m|Dc(s - s_{ep})}{m^2D^2 - c^2}. \quad (27)$$

For completeness, a derivation of these formulae is given in Sec. 9 as an appendix. In the above, c is the CoC diameter described in the previous section, m is the magnification, $D = f/N$ is the entrance pupil diameter, and $s_{ep} = (1 - m_p^{-1})f$ is the separation between the first principal plane and the entrance pupil defined by Eq. (64) of Sec. 9 and illustrated in Fig. 3. This means that $s - s_{ep}$ is the object-plane distance measured from the entrance pupil rather than the first principal plane. Equations (26) and (27) can be applied for object-plane distances $s - s_{ep} < \mathcal{H}$, where \mathcal{H} is the hyperfocal distance defined by Eq. (76) of Sec. 9. At \mathcal{H} and beyond, the far DoF and total DoF are both infinite.

Now, consider format 1 with sensor diagonal d , front effective focal length f_1 , entrance pupil diameter D_1 , CoC diameter c_1 , and consider an object plane positioned a distance $s_1 - s_{ep,1}$ away from the entrance pupil. The total DoF is given by

$$\text{DoF}_1 = \frac{2|m_1|D_1c_1(s_1 - s_{ep,1})}{m_1^2D_1^2 - c_1^2}, \quad |m_1| = \frac{f_1}{s_1 - f_1}. \quad (28)$$

Now, consider format 2 with a smaller sensor diagonal d/R , front effective focal length f_2 , entrance pupil diameter D_2 , CoC diameter c_2 , and consider the same object plane positioned a distance $s_2 - s_{ep,2}$ from the entrance pupil. The total DoF is given by

$$\text{DoF}_2 = \frac{2|m_2|D_2c_2(s_2 - s_{\text{ep},2})}{m_2^2D_2^2 - c_2^2}, \quad |m_2| = \frac{f_2}{s_2 - f_2}. \quad (29)$$

However, according to Eq. (5) of Sec. 2.1, Eq. (12) of Sec. 2.2, Eq. (16) of Sec. 2.2.1, and Eq. (23) of Sec. 2.3, respectively, the following relationships hold when equivalent photos are taken:

$$\begin{aligned} s_2 - s_{\text{ep},2} &= s_1 - s_{\text{ep},1}, \\ f_2 &= \frac{f_1}{R_w}, \\ s_2 - f_2 &= (s_1 - f_1) \frac{R}{R_w}, \\ c_2 &= \frac{c_1}{R}. \end{aligned} \quad (30)$$

Combining the second and third equations above also shows that $|m_2| = |m_1|/R$. Substituting these relations into Eq. (29) yields the following result:

$$\text{DoF}_2 = \frac{2|m_1|D_2c_1(s_1 - s_{\text{ep},1})}{m_1^2D_2^2 - c_1^2}. \quad (31)$$

By comparison with Eq. (28), it can be concluded that $\text{DoF}_2 = \text{DoF}_1$ provided the following condition is satisfied:

$$D_2 = D_1. \quad (32)$$

In other words, a necessary condition for producing equivalent photos is the use of the same entrance pupil diameter on each format. As discussed in Sec. 1.3, in photographic situations where the smaller format cannot provide an entrance pupil diameter that matches that of the larger format, the smaller format will be unable to produce an equivalent photo. Since $N_1 = f_1/D_1$, $N_2 = f_2/D_2$, and $f_2 = f_1/R_w$, it now follows that

$$N_2 = \frac{N_1}{R_w}. \quad (33)$$

When focus is set on an object plane positioned closer than infinity, the above result reveals that the equivalence ratio R must formally be replaced by the working equivalence ratio R_w when relating equivalent f -numbers as well as equivalent focal lengths.

2.5 Same Shutter Speed

Equivalent photos taken by cameras based on different formats must be produced in the presence of the same amount of subject motion blur. This is defined as blur that occurs due to objects moving in the scene during the camera exposure. Since equivalent photos have the same perspective and framing, it follows that equivalent photos must be taken using the same exposure duration (shutter speed):

$$t_1 = t_2. \quad (34)$$

This requirement does not specify an appropriate shutter speed but merely states that it must be the same for each camera format. The required shutter speed depends upon the nature of the scene luminance distribution and the exposure strategy of the photographer.

2.6 Same Lightness: Equivalent ISO Settings

Even though equivalent photos are taken using the same shutter speed, the resulting lightness of the images will not be the same unless equivalent ISO settings are used rather than the same ISO settings. This is because the ISO setting determines the sensitivity of the camera digital output to incident photometric exposure, and different formats receive different levels of photometric exposure when equivalent photos are taken. The ISO equivalence relationship is derived in this section.

First, recall from Eq. (32) that the entrance pupil diameters on each format are the same when equivalent photos are taken. Since the shutter speeds are also required to be the same, it follows that equivalent photos are produced using the same total amount of light, as previously discussed in Sec. 1.3. More specifically, the total luminous energy Q incident at the sensor plane of both formats will be the same:

$$Q_1 = Q_2. \quad (35)$$

In order to rigorously derive this result, consider the photometric exposure at an infinitesimal area element on the sensor plane. This is defined by the well-known camera equation:

$$H = \frac{\pi}{4} LT \frac{t}{N_w^2} \cos^4 \varphi, \quad (36)$$

where L is the scene luminance at the corresponding scene area element. The cosine fourth rolloff factor describes the natural reduction in illuminance at the sensor plane that occurs away from the optical axis. It depends upon the angle subtended by the entrance pupil from the scene area element being considered. The factor T is the lens transmission factor, and t is the shutter speed. The working f -number N_w depends on the object-plane distance and can be expressed in the form $N_w = bN$, where b is the bellows factor defined by Eq. (7). At infinity focus, $b \rightarrow 1$ and the working f -number reduces to the f -number, $N_w \rightarrow N$.

Now, consider the larger format labelled format 1 and the smaller format labelled format 2. From Eq. (36), the photometric exposure H_1 at an infinitesimal area element dA_1 on the larger format sensor with focus set on a specified object plane is given by

$$H_1 = \frac{\pi}{4} LT \frac{t}{N_{w,1}^2} \cos^4 \varphi. \quad (37)$$

Analogously, the exposure H_2 at an infinitesimal area element dA_2 on the smaller format sensor with focus set on the same object plane is given by

$$H_2 = \frac{\pi}{4} LT \frac{t}{N_{w,2}^2} \cos^4 \varphi. \quad (38)$$

The working f -numbers for these systems are defined by

$$N_{w,1} = b_1N_1, \quad N_{w,2} = b_2N_2. \quad (39)$$

Substituting these into Eq. (33) and then utilizing Eq. (13) yields

$$N_{w,2} = \frac{N_{w,1}}{R}. \quad (40)$$

This result shows that the “working” f -numbers are always directly related through the equivalence ratio R when equivalent photos are taken with focus set at any chosen object-plane distance.

As illustrated in Fig. 3, the luminance and cosine fourth terms appearing in Eqs. (37) and (38) will be the same for both formats at the same perspective and framing (AFoV). It must also be assumed that the lens transmission factors are the same. (Lens transmission factors depend upon the underlying lens technology; the ISO 12232 standard assumes a standard value $T = 0.9$ when ISO settings are measured, along with $\varphi = 10$ deg and an infinite object distance.) Since the shutter speeds are also the same, combining Eqs. (37), (38), and (40) yields the following relationship:

$$H_2 = R^2 H_1. \quad (41)$$

The exposure at an infinitesimal area element on the smaller sensor is therefore a factor R^2 greater than the exposure at an infinitesimal area element on the larger sensor when equivalent photos are taken. The total luminous energy Q_1 and Q_2 incident at the sensor plane of the larger format and smaller format, respectively, during the camera exposure are given by integrating the photometric exposure over the corresponding sensor areas:

$$Q_1 = \int H_1 dA_1, \quad Q_2 = \int H_2 dA_2. \quad (42)$$

However, the area A_2 of the smaller format sensor is a factor R^2 smaller than the larger sensor format area A_1 , as shown in Fig. 6. Therefore,

$$A_2 = \frac{A_1}{R^2}, \quad dA_2 = \frac{dA_1}{R^2}. \quad (43)$$

Substituting Eqs. (41) and (43) into Eq. (42) shows that the total luminous energy incident at the sensor plane of both camera formats is equal when equivalent photos are taken, thus proving Eq. (35).

Now, consider the arithmetic average photometric exposures $\langle H_1 \rangle$ and $\langle H_2 \rangle$ for both camera formats. These are defined as follows:

$$\langle H_1 \rangle = Q_1/A_1, \quad \langle H_2 \rangle = Q_2/A_2. \quad (44)$$

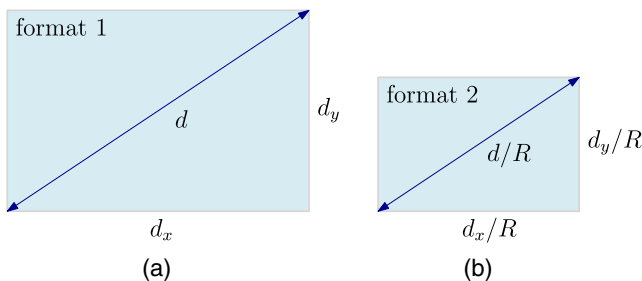


Fig. 6 (a) The area of sensor format 1 is $d_x d_y$, and (b) the area of sensor format 2 is $(d_x/R)(d_y/R)$. Format 2 is therefore a factor R^2 smaller than format 1.

Utilizing Eqs. (35) and (43) yields

$$\langle H_2 \rangle = R^2 \langle H_1 \rangle. \quad (45)$$

In traditional exposure strategy, the product of the arithmetic average exposure with the exposure index (ISO setting) S defines a photographic constant P , which is independent of sensor format. (The ISO 12232 standard⁷ uses $P = 10$, which indirectly implies an average scene luminance of 18% for a typical photographic scene.⁵) This means that

$$\langle H_1 \rangle S_1 = \langle H_2 \rangle S_2 = P. \quad (46)$$

Now, combining Eqs. (45) and (46) yields the final result:

$$S_2 = \frac{S_1}{R^2}. \quad (47)$$

The required ISO setting on the smaller format is therefore a factor R^2 lower than the required ISO setting on the larger format when equivalent photos are taken. Equation (47) holds when focus is set at any chosen object-plane distance.

Brightness is commonly used to describe the nonlinear perceptual response of the human visual system to luminance. Lightness can be thought of as brightness defined relative to a reference white. While brightness as a descriptor ranges from “dim” to “bright,” lightness ranges from “dark” to “light.” The lightness function L^* defined by the CIE is illustrated in Fig. 7 and is specified by the following formula:

$$L^* = 116f(Y) - 16, \quad (48)$$

where

$$f(Y) = \begin{cases} Y^{1/3} & \text{if } Y \geq \delta^3 \\ \frac{Y}{3\delta^2} + \frac{4}{29} & \text{otherwise} \end{cases}, \quad (49)$$

and $\delta = 6/29$. Here, $Y = L/L_n$ denotes relative luminance and so L_n is the luminance of the reference white. Lightness is therefore a nonlinear function of relative luminance, as shown in Fig. 7, and takes values between 0 and 100%. Note that 18% relative luminance corresponds to 50% lightness, which is commonly referred to as “middle gray.”

Since 2004, Japanese camera manufacturers have been required by CIPA to use the SOS method⁶ (or the related REI method) to determine camera ISO settings. The SOS method is based on a measurement of the photometric exposure required to map 18% relative luminance to a standard DOL in the output JPEG file. For an 8-bit JPEG file encoded using the sRGB color space, CIPA chose 118 as the standard DOL because this value corresponds with middle gray (50% lightness) on the standard gamma curve of the sRGB color space, as shown in Fig. 7. In other words, 18% relative luminance will always correspond with 50% lightness in the output JPEG file, irrespective of the shape of the JPEG tone curve used by the camera. Camera ISO settings will be discussed further in Sec. 4.

As already mentioned, the value of the photographic constant $P = 10$ corresponds to assuming that the average scene luminance will be 18% of the maximum for a typical photographic scene. Provided the scene is typical and traditional metering based on average photometry is used, the SOS

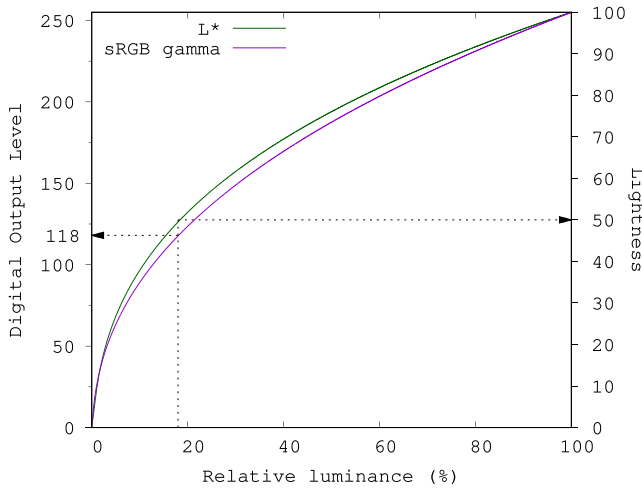


Fig. 7 18% relative luminance corresponds to 50% lightness (middle gray) on the CIE lightness curve used by the CIE LAB color space. The corresponding 8-bit DOL in the sRGB color space is 118. This DOL is 46% of the maximum DOL since the sRGB gamma curve is not identical to the CIE lightness curve.

method described above ensures that the average scene luminance will map to a standard lightness in the output JPEG file. In other words, the same camera exposures used on the same format will map the average scene luminance to the standard lightness. When equivalent photos are taken using different formats, the same applies to the equivalent camera exposures.

In summary:

- Provided equivalent focal lengths and equivalent f -numbers are related through the working equivalence ratio R_w , equivalent photos are produced using the same total amount of light.
- The ISO settings required on different camera formats will always be directly related through the square of the standard equivalence ratio R when equivalent photos are taken. This relationship is independent of the distance to the object plane upon which focus is set.
- When equivalent ISO settings are used to take equivalent photos, the average scene luminance for a typical photographic scene will map to the same standard lightness in the output JPEG files.

2.7 Equation Summary

2.7.1 Focus set at any chosen object-plane distance

Denoting the larger and smaller formats by subscripts 1 and 2, respectively, equivalent focal lengths and equivalent f -numbers are related by the “working” equivalence ratio R_w when equivalent photos are taken with focus set at any chosen object-plane distance:

$$f_2 = \frac{f_1}{R_w}, \quad N_2 = \frac{N_1}{R_w}.$$

On the other hand, equivalent ISO settings are always related through the standard equivalence ratio R :

$$S_2 = \frac{S_1}{R^2}.$$

The same holds for the “working” f -numbers:

$$N_{w,2} = \frac{N_{w,1}}{R}.$$

A general expression for R_w is as follows:

$$R_w = \left(\frac{b_2}{b_1} \right) R,$$

where b_2/b_1 is the ratio of the bellows factors when focusing at the specified object plane. Practical expressions for R_w have been given in Sec. 2.2.1. For example, if f_1 or N_1 corresponding to the larger format are known, then

$$R_w = \left(1 - \frac{m_{c,1}}{p_{c,1}} \right) R,$$

and

$$m_{c,1} = \left(\frac{R-1}{R} \right) \frac{f_1}{s_1}, \quad p_{c,1} = m_p + (1 - m_p) \frac{f_1}{s_1}.$$

The correction $m_{c,1}$ arises due to the difference between the equivalent system magnifications, and the correction $p_{c,1}$ arises for a nonunity pupil magnification.

2.7.2 Focus set at infinity

When focus is set at infinity, the working equivalence ratio $R_w \rightarrow R$ and the working f -number $N_w \rightarrow N$. The equivalence equations then reduce to the following three equations presented in Sec. 1.2 earlier:

$$f_2 = \frac{f_1}{R}, \quad N_2 = \frac{N_1}{R}, \quad S_2 = \frac{S_1}{R}.$$

In practice, the numerical difference between R_w and R turns out to be negligible beyond macro object-plane distances and so these equations at infinity focus can be used in place of the exact equations in ordinary photographic situations. This will be shown by example in Sec. 3.

2.7.3 Other equivalence equations

Several other equations of interest arise, which hold when focus is set at any chosen object-plane distance including infinity:

$$\begin{aligned} D_2 &= D_1, \\ Q_2 &= Q_1, \\ \langle H_2 \rangle &= R^2 \langle H_1 \rangle, \\ c_2 &= \frac{c_1}{R}, \\ |m_2| &= \frac{|m_1|}{R}, \end{aligned}$$

where D is the entrance pupil diameter, Q is the total luminous energy, $\langle H \rangle$ is the average sensor-plane photometric exposure, c is the CoC diameter, and $|m|$ is the magnification.

3 Numerical Examples

Assume that the focal length, f -number, and ISO setting are known on a 35-mm full-frame camera labelled format 1. The task is to find the equivalent focal lengths, f -numbers, and

ISO settings on a range of other formats. In the following examples, the use of a traditional-focusing lens with constant focal length is assumed. For internally focusing lenses that change their focal length upon focusing (see Sec. 8), knowledge of the “new” focal length upon setting focus would be required.

In the present case, the appropriate formula for the working equivalence ratio R_w is that given in Sec. 2.7.1. If it is assumed that the pupil magnification is unity for simplicity so that $p_{c,1} = 1$, then the object-plane distances measured from the entrance pupil of both camera formats will be the same, $s_1 = s_2 = s$. Rearranging the expression for R_w now yields

$$\frac{s}{f_1} = \frac{1 - (1/R)}{1 - (R_w/R)} \quad \text{for } m_p = 1. \quad (50)$$

Given the equivalence ratio R for format 2, this formula enables the ratio R_w/R to be plotted against s/f_1 , the ratio of object distance to format 1 focal length.

Figure 8 illustrates the results for a selection of format 2 sizes. Although larger magnifications are possible, the minimum value shown on the horizontal axis, $s/f_1 = 2$, corresponds to 1:1 reproduction ($|m| = 1$). It can be seen that the minimum value of R_w/R becomes smaller as the size difference between format 1 and 2 increases. Consequently, $R_w/R \rightarrow 1$ more slowly as $s/f_1 \rightarrow \infty$ for larger R . Significantly, the ratio R_w/R very quickly increases with s/f_1 for all format 2 sizes. Even for the small 1/2.5 in. format, R_w/R already reaches 0.9 when $s/f_1 \approx 8.33$, which still defines an object plane positioned very close to the lens entrance pupil.

As a numerical example, consider format 1 set at ISO 400 and fitted with a 100 mm macro lens set at $N = 4$. At infinity focus, $R_w \rightarrow R$ and the equivalent camera exposures listed in Table 4 for a selection of format 2 sizes produce equivalent photos. The shutter speed is arbitrary because it depends upon the scene luminance distribution, but nevertheless must be the same for all formats. Both R and R_w along with the f -number have been rounded to two decimal places, the focal lengths have been rounded to one decimal place, and the ISO settings have been rounded to the nearest integer.

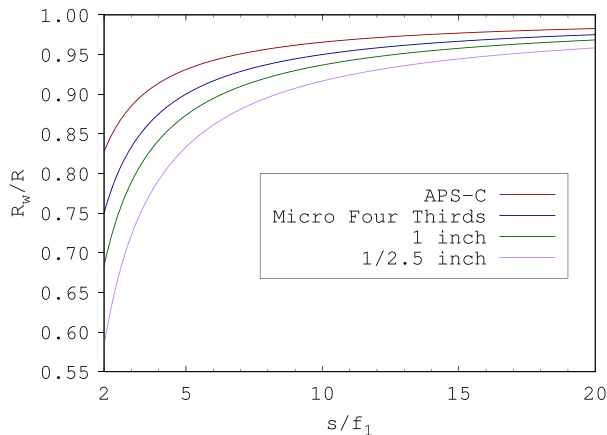


Fig. 8 R_w/R ratio as a function of object distance to focal length ratio s/f_1 for a selection of sensor formats when comparing with the 35-mm full-frame format.

Table 4 Example equivalent camera exposures when focused at infinity.

Format	R	R_w	$f(\text{mm})$	N	$S(\text{ISO})$
35-mm full frame	1	1	100	4	400
APS-C	1.53	1.53	65.4	2.62	171
APS-C (Canon)	1.61	1.61	62.0	2.48	154
Micro four thirds	2.00	2.00	50.0	2.00	100
1 in.	2.73	2.73	36.7	1.47	54
2/3 in.	3.93	3.93	25.4	1.02	26
1/1.7 in.	4.55	4.55	22.0	0.88	19
1/2.5 in.	6.03	6.03	16.6	0.66	11

3.1 Portrait Photography

Now, consider focus set on an object plane positioned 5 m from the entrance pupil of the 100 mm lens on format 1. In this case, the ratio $s/f_1 = 50$, which lies in the portrait photography regime, and the magnification is found to be $|m| \approx 0.02$ by using the formula $|m| = f_1/(s - f_1)$. The equivalent camera exposures listed in Table 5 for a selection of format 2 sizes produce equivalent photos. This data confirms the trend shown in Fig. 8; it is clear that there is negligible difference from the infinity focus results. Evidently, R_w is very well-approximated by R in the portrait regime and beyond toward infinity.

3.2 Macro Photography

Macro object distances are traditionally defined as 1:1 magnification ($|m| = 1$) or larger. Using the formula $|m| = f_1/(s - f_1)$, when $|m| = 1$, the ratio $s/f_1 = 2$. From Fig. 8, it is evident that the equivalence correction terms are important in this regime, particularly for very small formats. For example, the ratio $R_w/R \approx 0.58$ when $|m| = 1$ for the 1/2.5 in. format.

For the same example 100 mm lens on format 1, the equivalent camera exposures listed in Table 6 produce

Table 5 Example equivalent camera exposures at portrait object-plane distances.

Format	R	R_w	$f(\text{mm})$	N	$S(\text{ISO})$
35-mm full frame	1	1	100	4	400
APS-C	1.53	1.52	65.8	2.63	171
APS-C (Canon)	1.61	1.60	62.5	2.50	154
Micro four thirds	2.00	1.98	50.5	2.02	100
1 in.	2.73	2.69	37.1	1.49	54
2/3 in.	3.93	3.87	25.8	1.03	26
1/1.7 in.	4.55	4.48	22.3	0.89	19
1/2.5 in.	6.03	5.93	16.9	0.68	11

Table 6 Example equivalent camera exposures at macro object-plane distances.

Format	R	R_w	$f(\text{mm})$	N	$S(\text{ISO})$
35-mm full frame	1	1	100	4	400
APS-C	1.53	1.26	79.1	3.16	171
APS-C (Canon)	1.61	1.31	76.5	3.06	154
Micro four thirds	2.00	1.50	66.7	2.67	100
1 in.	2.73	1.86	53.7	2.15	54
2/3 in.	3.93	2.47	40.5	1.62	26
1/1.7 in.	4.55	2.78	36.0	1.44	19
1/2.5 in.	6.03	3.51	28.5	1.14	11

equivalent photos at 1:1 magnification. The numerical differences with the portrait and infinity focus cases are large. Significantly, larger equivalent focal lengths and equivalent f -numbers are required on smaller formats than would be expected using the traditional infinity-focus formulae. Although the pupil magnification m_p has been set to unity in the above example, a nonunity pupil magnification will also have a large effect in the macro regime and should be included in real-world calculations.

In conclusion, the equivalence ratio R_w is important for macro photography. However, R_w is well-approximated by the equivalence ratio R in general photographic situations, and the simplified equivalence equations can be used.

4 Equivalence and Image Quality

Section 1.3 argued that cross-format IQ comparisons should be carried out using equivalent camera settings over the equivalence overlap between the formats being compared. In this case, IQ for both formats is expected to be of the same order of magnitude, although real-world differences will be revealed due to differences in the underlying camera and lens technologies. When equivalent camera settings do not exist on the smaller format due to equivalent f -number or equivalent ISO setting limitations, the extra photographic capability of the larger format can be utilized and this potentially offers higher IQ in terms of RP and SNR.

Objective IQ metrics such as RP, SNR, and engineering dynamic range are useful for evaluating camera system capability. However, perceived IQ metrics, such as photographic dynamic range and sharpness, are often more useful in photography because they take into account the conditions under which the output image will be viewed.

The purpose of this section is to explain how to appropriately apply equivalence theory when performing cross-format IQ comparisons and, in particular, the appropriate units that should be used. Since IQ metrics are commonly evaluated as a function of ISO setting, this section begins with an explanation of how ISO settings are derived in modern photography. Subsequently, noise and SNR, dynamic range, RP, and sharpness are all discussed.

It should be noted that unless a specific object-plane distance is being considered, the simplified equivalence

equations at infinity focus can be used when performing cross-format IQ comparisons.

4.1 ISO Sensitivity Settings

Prior to 2004, camera manufacturers typically used the saturation-based method from ISO 12232⁷ for determining camera ISO settings. This method involved a measurement of the photometric exposure required to saturate the JPEG output for a known scene luminance. (For an 8-bit JPEG file, saturation corresponds to a DOL of 255.) However, a drawback of this method was that the actual shape of the JPEG tone curve below saturation was not taken into account. This meant that different camera models could produce images with middle gray (18% relative luminance or 50% lightness) mapped to a different DOL for the same exposure settings. In other words, 18% relative luminance did not necessarily map to 50% lightness in the output JPEG file, and so photographers were faced with the unsatisfactory situation that certain cameras would produce images that appeared darker or lighter than those from other cameras when using the same exposure settings.

In order to address the above issue, CIPA introduced the SOS method⁶ in 2004, and this was subsequently incorporated into the ISO 12232 standard. Since that time, Japanese camera manufacturers have been required by CIPA to use the SOS method (or the related REI method) to determine camera ISO settings.

As briefly described in Sec. 2.6 earlier, the SOS method is based upon a measurement of the photometric exposure required to map 18% relative luminance to a standard DOL in the output JPEG file. For an 8-bit JPEG file encoded using the sRGB color space, CIPA chose 118 as the standard DOL because this value corresponds with middle gray (50% lightness) on the standard gamma curve of the sRGB color space, as shown in Fig. 7 of Sec. 2.6. In other words, 18% relative luminance will always correspond with 50% lightness in the output JPEG file, irrespective of the shape of the JPEG tone curve used by the camera. If the camera manufacturer alters the shape of the JPEG tone curve, then the required exposure duration and hence measured ISO value will adjust accordingly in order to ensure that 18% relative luminance maps to the standard DOL. This can affect IQ. For example, camera manufacturers can sacrifice SNR for extra highlight headroom in the JPEG output.¹²

Once the camera ISO settings have been determined, it is instructive to consider their use as part of a practical exposure strategy for a real-world scene. As mentioned in Sec. 2.6 earlier, a “typical” photographic scene is assumed to have

Table 7 From top to bottom, cameras based on the 35-mm full frame, APS-C, micro four thirds, and 1 in. formats, respectively, all set at the same camera exposure.

Camera	$f(\text{mm})$	N	$t(\text{s})$	$S(\text{ISO})$
Canon 1D-X	85	2.0	1/1.3	200
Fujifilm X-A1	56	2.0	1/1.3	200
Panasonic GH4	42.5	2.0	1/1.3	200
Nikon 1 V3	32	2.0	1/1.3	200

an average relative luminance of 18%, i.e., an average scene luminance that is 18% of the maximum. (This follows from the value of the photographic constant $P = 10$.) For such a scene, the use of traditional metering based on average photometry will analogously ensure that the average scene luminance will map to 50% lightness in the output JPEG file. If the scene is not typical, then the photographer can either use exposure compensation to adjust the metered average luminance, use a “matrix” metering mode, or a combination of both.

4.1.1 “RAW” ISO settings

Although IQ comparisons based on JPEG output are useful for photographers who primarily use JPEG output from the camera, many photographers prefer to process the RAW data themselves using RAW conversion software that provides more control over the nature of the output image. Furthermore, processing the RAW data manually enables the maximum IQ capability of the camera to be fully extracted.

However, it should be evident from the above description of modern camera ISO settings that such settings correspond with the JPEG output only and are not valid with RAW data. One of the main reasons for this is that the ISO measurements include any mid-tone digital gain that may have been applied in order to alter the JPEG tone curve away from the standard gamma curve of the output color space. In order to provide fair RAW IQ comparisons, it would be necessary to define “RAW ISO values,” which directly correspond with the analog gain settings used to produce the linear RAW data. The SOS method cannot be applied to RAW data, however, the “saturation-based ISO speed” method that was used prior to 2004, and remains part of the ISO 12232 standard, could feasibly be applied to RAW data. In fact, this is precisely the method used by some online sources of camera IQ measurements, such as DxOMark[®].¹³ No correspondence should be expected between the ISO values used by such sources and the ISO settings labelled on the camera; such labelled settings are only for use with in-camera JPEG output. Much confusion arises from DxOMark’s use of the term “measured ISO” to label their ISO values. “Saturation-based RAW ISO” would be a more appropriate term that distinguishes between camera JPEG and measured RAW values.

Although IQ metrics are commonly evaluated as a function of ISO setting, it should be remembered that the ISO setting itself is not a performance metric. For example, a higher quantum efficiency is favorable and this increases ISO sensitivity, both in JPEG and RAW. On the other hand, a higher full-well capacity (FWC) per unit area is also favorable, but this decreases ISO sensitivity when measured using the saturation-based RAW method.

4.2 Noise

Although SNR is a useful quantifiable measure of IQ, it is also useful to gain a visual impression of the image noise to be expected at a given ISO setting for a typical photographic scene.

For photographers who primarily use JPEG output from the camera, camera reviewers may, for example, compare (or provide a comparison tool for) images obtained using the camera default JPEG settings. In this case, the quality of

the internal camera signal processing and noise filtering will be taken into account. For photographers that primarily process RAW output themselves, camera reviewers may instead provide images that have been processed from the RAW files using a standard workflow with noise filtering disabled.

In order to demonstrate equivalence theory using the visual impression of image noise, consider the cameras listed in Table 7 all set at the same camera exposure. These cameras are based on the 35-mm full frame, APS-C, micro four thirds, and 1 in. formats, respectively. Lenses with equivalent focal lengths have been used so that the AFoV is the same in all cases. However, the images are not equivalent because the total light collected and DoF are different in all cases.

The left column of Fig. 9 shows corresponding images reproduced from Ref. 1. Images (a)–(d) correspond to the Canon, Fujifilm, Panasonic, and Nikon cameras, respectively. The images were obtained by processing the RAW files using Adobe[®] Camera Raw with the Adobe[®]

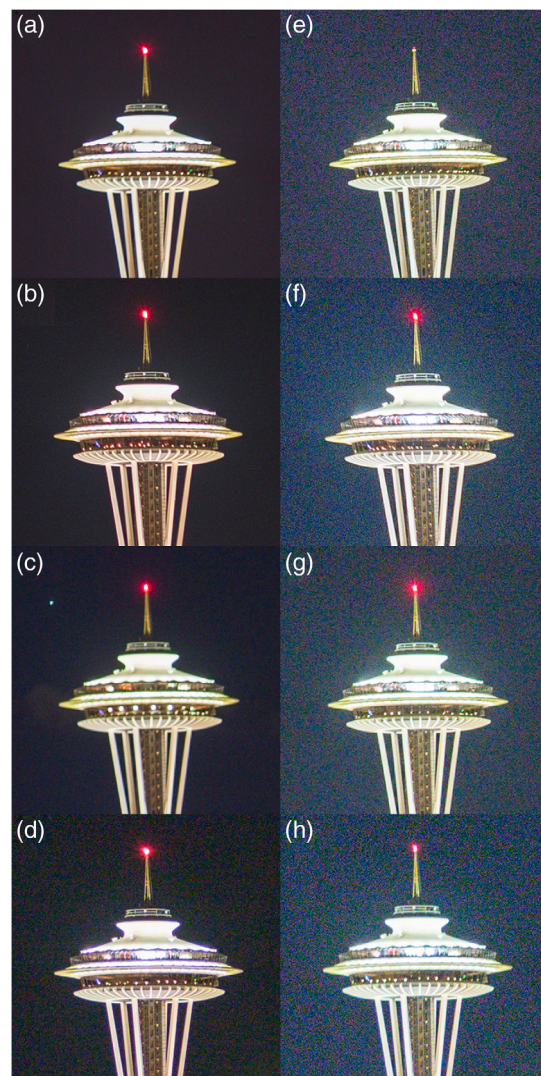


Fig. 9 (a)–(d) Noise appearance for cameras using the same camera exposure corresponding to Table 7. (e)–(h) Same as left column but using equivalent camera exposures corresponding to Table 8. The figure is best viewed on a computer monitor. Figure derived from Ref. 1 courtesy of www.dpreview.com.

Table 8 From top to bottom, cameras based on the 35-mm full frame, APS-C, micro four thirds, and 1 in. formats, respectively, all set at equivalent camera exposures. The nearest exposure variables available on the camera for equivalence to hold are listed, with the ideal equivalent values given in brackets.

Camera	f (mm)	N	t (s)	S (ISO)
Canon 1D-X	85	5.6	1/1.3	3200 (2975)
Fujifilm X-A1	56	3.6 (3.7)	1/1.3	1250 (1272)
Panasonic GH4	42.5	2.8	1/1.3	800 (744)
Nikon 1 V3	32 (31)	2.0 (2.1)	1/1.3	400

Standard color profile and noise filtering minimized. The images were subsequently downsampled to a common pixel count. Note that the Adobe Standard color profile aims to give a consistent contrast response across cameras but is calibrated to respect the manufacturer's JPEG ISO implementation. In other words, the images all appear equally as light as each other.

The noise clearly becomes more apparent as the format size decreases. This is to be expected because the largest contribution to the image noise is photon shot noise. As mentioned in Sec. 1.1 earlier, photon shot noise scales as the square root of the total amount of light collected (or more precisely, the number of photoelectrons generated) and so overall, SNR is higher as the format size increases.

Now, consider the same cameras all set at the "equivalent" camera exposures listed in Table 8. The right column of Fig. 9 again shows corresponding images reproduced from Ref. 1 using the same RAW workflow as above. Slight lightness adjustments have been made to compensate for the use of any nonideal equivalent settings. Since the images are equivalent, it is clear that the level of visual noise is now very similar for all cameras.

Remaining differences in noise are due to differences in the underlying camera and lens technology, for example:

- sensor quantum efficiency;
- read noise;
- sensor pixel count and image resampling;
- JPEG tone curve processing.

A higher quantum efficiency enables more photoelectrons to be produced for a given level of photometric exposure. This increases the sensitivity of the camera digital output to incident photometric exposure and raises the ISO value (both JPEG and RAW) corresponding to the selected analog gain setting. In other words, more light can be collected and a higher SNR can be achieved at a given ISO value compared to a camera with lower quantum efficiency. When comparing equivalent photos, the same applies to the equivalent ISO settings because in this case, the total incident light is the same rather than the level of incident photometric exposure.

The noise floor or read noise is another factor, which will become apparent in darker regions of the image. When producing equivalent photos, the value of the equivalent ISO setting is lower on a smaller format because a smaller format requires less gain to achieve the same DOL when equivalent

photos are taken, as proven in Sec. 2.6 earlier. In principle, this compensates for the format size difference in terms of read noise. In practice, there will be a dependence upon signal processing technology. There will also be a dependence upon sensor pixel count because the total read noise for an aggregate of smaller pixels is generally different to that of a larger pixel of equal area. More formally, it is actually differences in the read noise per percentage sensor area at equivalent ISO settings that matters when comparing equivalent photos. This is discussed further in Sec. 4.3 below, which covers SNR in more detail.

Sensor pixel count also plays another role, which primarily affects photon shot noise when comparing equivalent photos. Because equivalent photos are required to be viewed at the same display dimensions, different pixel counts require different amounts of resampling in order to match the required display resolution in pixels per inch. Downsampling reduces noise at the expense of resolution, whereas upsampling preserves any resolution advantage without reducing noise. As discussed in Sec. 4.3 below, sensor pixel count can be automatically taken into account when measuring SNR by using a metric such as SNR per percentage sensor area.

Since the camera ISO settings are based upon the JPEG output and are determined using the SOS method described in the previous section, camera manufacturers can alter the shape of the JPEG tone curve in order to position middle gray (DOL = 118 in the JPEG output) at any desired position on the sensor response curve. For example, shifting DOL = 118 to correspond with a lower RAW value can increase the highlight headroom in the JPEG output while maintaining the standard mid-tone lightness that defines the ISO setting.¹² However, this will lower the SNR because digital gain is effectively being applied to the mid-tone region. Camera manufacturers will choose to balance such trade-offs differently, and this will have an effect on the level of noise seen in the JPEG output.

Finally, the f -number and ISO setting used on the full-frame format in the example given in Table 8 are both relatively high. This means that there is a large equivalence overlap with the smaller formats. A major step-up in noise performance can be achieved using a larger format only when its extra photographic capability is utilized. As discussed in Sec. 1.3 earlier, this requires the use of a low ISO setting and/or f -number that do not have equivalents on the smaller format.

4.3 Signal-to-Noise Ratio

SNR is a more precise way of quantifying camera capability in terms of image noise. Typically, SNR is measured using RAW data and so "RAW" ISO settings should ideally be used. If the camera ISO settings are used instead, the results will favor cameras that place middle gray (DOL = 118 in the JPEG output) higher on the sensor response curve.

Photon transfer curves are plots of noise (or alternatively SNR) as a function of normalized exposure value (EV), and the curves can be plotted at various ISO settings. The normalized EV is simply the number of stops and can be calculated based on input-referred units (photoelectron count) or output-referred units (RAW values). The normalized EV based on input-referred units does not take ISO gain into account and so the speed value is not included. On the

other hand, output-referred units do take into account the ISO gain and so the speed value is included in the number of stops.

When performing cross-format comparisons using equivalence theory, the total light received by each format must be the same. It is therefore most convenient to use output-referred units when specifying the normalised EV (i.e., the base 2 logarithm of the RAW value specified as a digital number or analog-to-digital unit.) Cross-format comparisons can then be performed simply by comparing curves at equivalent ISO settings.

SNR itself is typically specified as a measure per sensor pixel (photosite). However, per photosite measurements are not particularly useful when comparing sensors with different photosite sizes, the reason being that SNR is dependent upon the area over which it is measured.¹⁴ For example, consider two sensors that are identical other than photosite count, one having four times as many photosites as the other. The sensor with the larger photosites will have a greater SNR per photosite because the light-gathering area per photosite is larger. However, this does not mean that the higher-resolution sensor is noisier; the same SNR per photosite could in principle be achieved using the higher-resolution sensor simply by binning every group of four photosites together, either on the sensor or by subsequent image resampling. In other words, a more appropriate measure of SNR when comparing sensors of the same format is SNR per percentage sensor area.¹⁴

When performing cross-format comparisons, SNR should be specified in a way that corresponds to a comparison of equivalent photos, i.e., the use of equivalent camera exposure settings. Since the display dimensions of equivalent photos must be the same, the different enlargement factors from the sensor areas to the viewed output image dimensions must be taken into account. It turns out that SNR per percentage sensor area is again the appropriate measure to use. In this case, it takes into account both sensor pixel size and format size. Although the measurement area scales in direct proportion with the format size, the level of photometric exposure is greater as the format size decreases because the equivalent ISO setting is lower for the same image lightness. In other words, the measure is based upon the same amount of incident light when comparing each format, at least over the equivalence overlap between them. Photon transfer curves can be conveniently compared by plotting SNR per percentage sensor area as a function of normalized EV based on RAW value, and then comparing the curves at equivalent ISO settings. The main advantage of a larger format is the extra photographic capability afforded by low ISO settings that have no equivalent on the smaller format.

Finally, it should be mentioned that FWC per percentage sensor area is an important sensor characteristic that can increase the maximum achievable SNR and in turn lead to greater dynamic range.

4.4 Dynamic Range

Camera reviewers may measure “highlight” and “shadow” dynamic range. These are per-pixel metrics that are only valid with JPEG output and are intended to give information about the nature of the JPEG tone curve used by the camera manufacturer. Highlight dynamic range is a measure of the number of exposure stops needed to increase middle gray

(DOL = 118) to saturation (DOL = 255). Analogously, shadow dynamic range is a measure of the number of exposure stops needed to increase DOL = 1 to DOL = 118.

However, it is the dynamic range in the RAW data that is more important in terms of camera capability. Dynamic range is defined in engineering as the ratio of the maximum signal to the minimum usable signal, with the latter defined as the signal corresponding to SNR = 1. This engineering definition of dynamic range is commonly used to quantify the dynamic range of digital cameras, with SNR typically calculated as a per sensor pixel (photosite) value using RAW data.

However, as discussed in the previous section, it is much more appropriate to use SNR per percentage sensor area when comparing digital cameras. This measure takes into account sensor pixel count when comparing the same formats, and both sensor pixel count and format size when performing cross-format comparisons using equivalent camera exposure settings. In order to illustrate the former case, consider the common misconception that larger photosites automatically provide higher dynamic range due to a higher FWC. This is not the case in practice because larger photosites also have a larger surface area for collecting light, and so the ratio between the maximum and minimum signal (photoelectron count) remains the same provided the same sensor technology is used. It is actually FWC per percentage sensor area which matters and this could favor either larger or smaller photosites depending on the technology used. On the other hand, read noise does have a dependence on sensor pixel count because the total read noise for an aggregate of smaller photosites is generally greater than that of a larger photosite of equal area, as already mentioned in Sec. 4.2. It follows that larger photosites may provide an SNR advantage at high ISO settings, where the contribution to the read noise upstream from the ISO gain amplifier dominates,¹⁴ and this can lead to a greater DR because of a lower “usable” signal.

In conclusion, SNR should ideally be normalized by percentage sensor area when calculating the dynamic range of digital cameras. Nevertheless, one issue that arises is which actual percentage value to choose as this will affect the absolute dynamic range values. A way forward is provided by the CoC described in Sec. 2.3. Since the CoC area scales in direct proportion with format size, SNR per CoC area is an equally valid measure that can be used instead of SNR per percentage sensor area. It can be calculated as follows:

$$\text{SNR per CoC} = \text{SNR per photosite} \times \sqrt{\frac{\pi(c/2)^2}{A}}, \quad (51)$$

where A is the photosite area and c is the CoC diameter. The significance of using the CoC area is that the corresponding area on the viewed output image is the minimum area over which an observer can actually resolve detail and observe dynamic range. This leads to a definition of perceived dynamic range referred to as photographic dynamic range (PDR).¹⁵ Since a typical observer of the output image would consider the signal corresponding to SNR = 1 to be too noisy to be useful in practice, a more appropriate definition can be used for the minimum usable signal. For example, Ref. 15 uses SNR = 20 after normalizing for

Table 9 Photographic dynamic range (PDR) measured in stops at the equivalent ISO settings listed in Table 8. Data sourced from Ref. 16.

Camera	Format	S(ISO)	PDR
Canon 5D mk4	35-mm full frame	2975	7.28
Canon 1D-X	35-mm full frame	2975	7.22
Fujifilm X-A1	APS-C	1272	7.29
Panasonic GH4	Micro four thirds	744	7.36
Nikon 1 V3	1 in.	400	6.86

the CoC area. It follows that in order to calculate the PDR, the lower PDR limit is the RAW level at which the SNR per photosite equals $20/\sqrt{\pi(c/2)^2/A}$, and the upper PDR limit is the value of the upper RAW clipping point. The PDR expressed in stops is then given by dividing the value of the upper limit by the lower limit and taking the base 2 logarithm.

Table 9 lists the PDR values for the same cameras and equivalent ISO settings used in the noise comparison of Sec. 4.2 with the addition of the full-frame Canon 5D mk4 camera. The PDR values are seen to be in very close agreement. Figure 10 shows that this equivalence behavior is maintained over the equivalence overlap between the cameras, at least for the Fujifilm, Panasonic, Nikon, and Sony cameras. At lower ISO settings, the extra photographic capability is evident as the format size increases. For example, the maximum PDR of the Panasonic GH4 occurs at ISO 100, however, the Nikon 1 V3 is unable to produce a PDR value of the same order of magnitude because its lowest ISO setting is 160. The equivalent ISO setting, ISO 54, does not exist and so, an equivalent photo cannot be produced. The camera is unable to use the same exposure duration as the Panasonic camera without overexposing the photo. Interestingly, the full-frame Canon 1D-X does not appear to take advantage of the extra photographic capability that its sensor size affords because the PDR curve levels off at

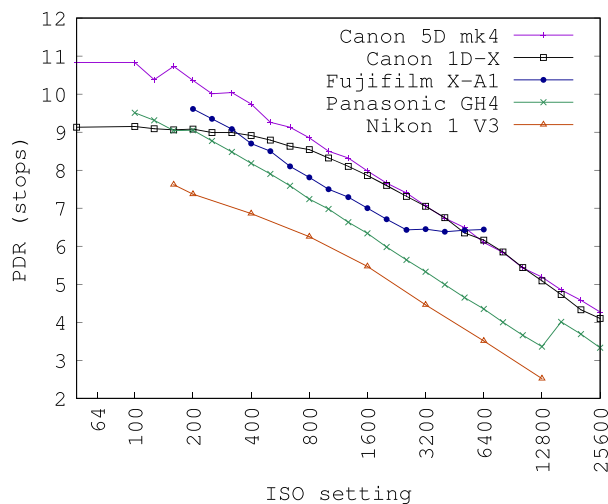


Fig. 10 Photographic dynamic range as a function of ISO setting for a selection of cameras. The data is not available at all ISO settings. Data sourced from Ref. 16.

low ISO settings. On the other hand, the more recent full-frame Canon 5D mk4 camera does take advantage of its extra photographic capability at low ISO settings. The long exposure duration available at these ISO settings enables the camera to achieve a higher SNR, which in turn leads to higher PDR. Furthermore, the Canon 5D mk4 has a very low base saturation-based RAW ISO setting. In a modern camera with high quantum efficiency, this is characteristic of a sensor with a high FWC per percentage sensor area.

4.5 Resolving Power

System RP describes the full capability of a camera and lens system to resolve detail. It is determined by the system MTF cut-off frequency. Many component contributions to the system MTF can be defined, and typically, an effective system cut-off frequency is used, which is defined by the spatial frequency at which the system MTF drops to a small percentage value, such as 10%. The system cut-off frequency is typically limited by either the sensor Nyquist frequency or the lens cut-off frequency, whichever is lower at the selected lens f -number. When the system RP is not limited by the lens, an optical low-pass filter may be required to lower the system cut-off frequency down to the sensor Nyquist frequency in order to prevent aliasing.

When performing cross-format comparisons, spatial frequencies μ at the sensor plane scale in proportion with the equivalence ratio R between the formats being compared, i.e.,

$$\mu_2 = R\mu_1, \quad (52)$$

where 2 denotes the smaller format and 1 denotes the larger format. This means that an RP of say 200 lp/mm on the 35-mm full-frame format corresponds with 306 lp/mm on APS-C ($R = 1.53$) and 400 lp/mm on micro four thirds ($R = 2$). To see this explicitly, recall that equivalent photos are viewed from the same distance and at the same display dimensions, and the smaller sensor diagonal is a factor R smaller than the larger sensor diagonal. Therefore, the equivalent image spatial frequencies on the smaller sensor plane must be a factor R larger than those on the larger sensor plane because the projected optical image needs an extra enlargement factor of R in order to match the dimensions of the viewed output image.

For example, consider the lens cut-off frequency for an ideal aberration-free diffraction-limited lens, i.e., the diffraction cut-off frequency. This is given by

$$\mu_c = \frac{1}{\lambda N}. \quad (53)$$

The diffraction cut-off frequency is reduced as the f -number increases. Since equivalent photos are produced using equivalent f -numbers according to $N_2 = N_1/R$, the diffraction cut-off frequency increases in proportion with R , which is consistent with the above analysis.

Fortunately, there is an alternative spatial frequency unit that is very convenient when performing cross-format comparisons, line pairs per picture height (lp/ph). This is related to lp/mm in the following way:

$$\text{lp/ph} = \text{lp/mm} \times \text{ph(mm)}. \quad (54)$$

In the present context, picture height refers to the short edge of the imaging sensor, which is proportional to the sensor diagonal. This means that the unit is format-independent and equivalent spatial frequencies are automatically accounted for provided the formats have the same aspect ratios. In particular, the sensor Nyquist frequency expressed in lp/ph units depends only on sensor pixel count (and fill-factor) and is independent of the actual pixel pitch.

It should be apparent that when performing cross-format system capability comparisons, the larger format can in principle gain a RP advantage when the system cut-off frequency is limited by the lens and an equivalent f -number on the smaller format does not exist. For example, when $N = 2.8$ on the 35-mm full-frame format, an equivalent f -number does not exist on the small 1/2.5 in. format for which $R = 6.02$. However, real-world comparisons are strongly affected by the nature of the lens aberrations, the presence/absence of a low-pass filter, pixel count, and other optical phenomena that can profoundly affect the effective system cut-off frequency.

On the 35-mm full-frame format, the lens cut-off frequency typically reaches its maximum at an f -number a stop or two higher than the lowest available because of geometric lens aberrations that dominate at the maximum entrance pupil diameter. As the format size decreases, there is less need to stop down to alleviate the effect of aberrations because the maximum achievable entrance pupil diameter is smaller.

4.6 Sharpness

The standard viewing conditions assumed by the camera or lens manufacturer typically involve an enlargement factor of 8 on the 35-mm full-frame format.¹¹ As discussed in Sec. 2.3, this means that a full-frame camera does not actually need to resolve spatial frequencies greater than 40 lp/mm on the sensor plane because further detail cannot be resolved by an observer of the output image under these viewing conditions. In other words, the RP of a camera system only becomes important when the output image is viewed under more extreme conditions (such as a greater enlargement or closer viewing distance) than the standard viewing conditions assumed by the camera or lens manufacturer.

Under the standard viewing conditions described above, the nature of the system MTF curves between 10 and 40 lp/mm on the full-frame format is much more significant in terms of IQ. When combined with the contrast sensitivity function of the human visual system, it is the system MTF values at these spatial frequencies that primarily determine perceived image sharpness. Several sharpness metrics exist, in particular, subjective quality factor¹⁷ has found application in photography.

From the discussion of spatial frequencies in the previous section, it follows that the nature of the MTF curves between $10R$ and $40R$ lp/mm (where R is the equivalence ratio) are most important for determining image sharpness under standard viewing conditions when performing cross-format capability comparisons between the full-frame format and a smaller format. For example, system MTF at 20 lp/mm on the 35-mm full-frame format should be compared with 30 lp/mm on APS-C and 40 lp/mm on micro four thirds. This means that the smaller format needs to match the system MTF of the larger format for all relevant equivalent spatial

frequencies in order to achieve the same perceived sharpness. Naturally, the component lens MTF curves should be based on equivalent focal lengths and equivalent f -numbers. In practice it is more convenient to use the line pairs per picture height (lp/ph) spatial frequency unit defined in the previous section.

An important contribution to perceived image sharpness is lens diffraction softening. This arises due to the fact that the diffraction cut-off frequency scales with f -number according to Eq. (53). Equivalent photos suffer the same level of diffraction softening. This is because blur due to diffraction softening will generally not become visible until the Airy disk diameter $d_{\text{Airy}} = 2.44\lambda N$ (corresponding to the diameter of the diffraction PSF) approaches the CoC diameter c . Since the equivalent f -numbers and CoC diameters scale in proportion with R according to Eqs. (2) and (23), the net result is that equivalent f -numbers on different formats lead to the same level of diffraction softening when the equivalent output images are viewed at the same distance and display dimensions. For the c value corresponding to standard viewing conditions, diffraction softening generally becomes noticeable at $N = 16$ on the full-frame format, $N = 10.5$ on the APS-C format, and $N = 8$ on micro four thirds.

Sensor pixel count is another important contribution to perceived image sharpness. Although a higher sensor pixel count only has a negligible effect on the system RP when the system is limited by the lens, a higher sensor pixel count can improve perceived images sharpness irrespective of the nature of the limiting component contribution to the system MTF. This can be seen by considering the detector-aperture contribution to the system MTF:

$$\text{MTF}_{\text{det-ap}}(\mu_x, \mu_y) = |\text{sinc}(d_x\mu_x, d_y\mu_y)|, \quad (55)$$

where d_x, d_y are the dimensions of the photosite detection areas in the horizontal and vertical directions, and μ_x, μ_y are the horizontal and vertical spatial frequencies on the sensor plane. The detection areas d_x, d_y are related to the pixel pitches p_x, p_y via the fill-factor. Recall that the detector cut-off frequency is defined as the spatial frequency at which the detector-aperture MTF first drops to zero, and for a 100% fill-factor the sensor Nyquist frequency in a given direction is half the detector cut-off frequency in that direction. The important point is that increasing the sensor pixel count increases the sensor Nyquist frequency, and this in turn increases the detector-aperture MTF at the important (lower) spatial frequencies that are most relevant to standard viewing conditions. This occurs even when the higher sensor Nyquist frequency cannot increase the system RP. However, perceived sharpness will not necessarily be improved over another camera system with a lower pixel count because that system may have higher system MTF curves at the important frequencies. For example, the presence of an optical low-pass filter will reduce the system MTF at the important spatial frequencies, even when the system RP is limited by the lens f -number. As sensor pixel count increases, there is less need for an optical low-pass filter and it is becoming common to see cameras without a low-pass filter fitted in order to improve perceived sharpness. However, it may be possible to observe aliasing at low f -numbers in some cameras that do not have a low-pass filter.

5 Equivalence and Mobile Phone Cameras

Equivalence theory can be readily applied to translate camera exposure settings between a traditional mobile phone camera and a larger-format camera provided the equivalence ratio R is known. This is particularly important because mobile phone manufacturers typically specify the lowest available lens f -number without specifying the sensor size. Along with the base ISO setting, knowledge of the lowest available f -number and sensor size are both necessary for determining photographic capability.

Similarly, equivalent camera exposure settings should be used as a framework for comparing IQ between different mobile phone cameras and between mobile phone and larger-format cameras. However, mobile phone manufacturers rarely use the equivalence framework correctly and this can lead to misleading claims in terms of IQ.

Nevertheless, there are photographic situations in which the IQ produced by mobile phone cameras can match that of larger-format cameras. This can occur when equivalent photos are taken, i.e., equivalent camera exposure settings are used on the mobile phone and the larger-format camera. Under these conditions, the entrance pupil diameter on both formats will be the same and the total light received will be the same. However, the equivalence overlap will typically be small due to the large size difference between the mobile phone sensor and larger-format sensor.

Mobile phone cameras generally use small sensors in order to reduce the size of the lens. Although 10-bit analog-to-digital converters are typically used in order to reduce costs, note that imaging sensor manufacturers are able to achieve similar quantum efficiency and read noise per percentage sensor area across a wide range of sensor sizes. Reference 18 provides a useful discussion of the design strategy for mobile phone camera lenses. A fixed focal length is typically used in order to reduce size and cost. The maximum entrance pupil diameter is restricted due to the small physical size. Indeed, the lowest f -numbers on larger-format cameras do not have equivalents on mobile phones. This means that aberrations are not the limiting factor in terms of RP since these only dominate at large entrance pupil diameters. However, technological requirements are more demanding due to the increased sensitivity to misalignment.

Because of the small equivalence overlap, there will be many photographic situations in which the mobile phone is unable to take an equivalent photo. In such cases, the IQ of photos taken with the mobile phone will typically appear to deteriorate when viewed at larger display sizes, unlike photos produced from larger-format cameras.

This section begins with a discussion of why photos produced from any modern camera appear to have good enough IQ when viewed at small display sizes such as mobile phone screens, at least in terms of noise and perceived sharpness. Subsequently, the equivalence overlap between an example mobile phone and an example larger-format camera is described, which helps to predict how the photographic output is likely to compare in terms of IQ under various photographic conditions. IQ differences will become more apparent when the photos are viewed at larger display sizes. Finally, example features based on computational photography techniques that are available on recent smartphones are briefly discussed; equivalence theory is less applicable when such features are enabled.

5.1 Display Size

The fundamental reason that images appear to be less noisy when viewed at smaller display sizes relates to the fact that the CoC diameter increases as the display dimensions decrease according to Eq. (22). For example, consider a mobile phone with a 1/2.5 in. sensor and 5 in. screen in landscape orientation. At the least distance of distinct vision (25 cm), the CoC diameter will be doubled from the 0.005 mm value given in Table 3 since the enlargement factor from the sensor dimensions to the mobile phone screen dimensions will be approximately half that to the A4 paper size used to calculate Table 3. In this case, the CoC area on the sensor will be four times as large.

In Secs. 4.3 and 4.4, it was explained that SNR is dependent on the size of the sensor area over which it is measured. In the present context, the SNR per CoC on the sensor is a measure that corresponds with the minimum area on the viewed output image over which detail and noise can be perceived by a human observer. According to Eq. (51), the SNR per CoC will be doubled when viewing an image on the mobile phone compared to viewing at A4 paper size, and so perceived noise will be lower. Although the lower noise occurs at the expense of perceived scene resolution due to the limited RP of the human visual system, mobile phone camera JPEG engines typically apply sharpening optimized for mobile phone display sizes in order to improve perceived sharpness. Various other image processing algorithms including noise reduction will also be applied.

Since the CoC diameter is the limiting factor in determining perceived noise and resolution, downsampling an image to a lower pixel count such as the mobile phone screen pixel count will not have a noticeable effect on IQ under the viewing conditions described above. However, downsampling can improve noise when “zooming in” to the image, albeit at the expense of captured scene resolution. In this case, the CoC is temporarily reduced to a small size and the SNR associated with individual pixels becomes more important. To see that downsampling can improve the SNR associated with individual pixels, consider a crude downsampling filter that averages over every block of four pixels in order to downsample an image to 25% of its original pixel count while maintaining the same image dimensions. Although this process discards captured scene resolution, temporal noise adds in quadrature and so each “larger” pixel in the downsampled image is associated with an SNR of up to twice that of any of the individual pixels in the original image when only temporal noise is considered.¹⁴ This higher SNR is maintained when the image dimensions are subsequently reduced.

5.2 Equivalence Overlap Example

As an example of the equivalence overlap, consider the Google Pixel 2 smartphone, which has a 1/2.6 in. sensor ($\sim 5.5 \text{ mm} \times 4.1 \text{ mm}$), a base ISO setting $S = 50$, and a lens with focal length $f = 4.45 \text{ mm}$ and fixed f -number $N = 1.8$. Also, consider a camera based on the micro four thirds format ($17.3 \text{ mm} \times 13.0 \text{ mm}$), which has an equivalence ratio $R = 2$ with 35-mm full frame. The Google Pixel 2 and micro four thirds sensors both have a 4:3 aspect ratio and the equivalence ratio between them is $R \approx 3.15$. Therefore, the maximum photographic capability of the Google Pixel 2 corresponds to using $S \approx 500$ and $N \approx 5.6$

on the micro four thirds format, along with a lens focal length $f \approx 14$ mm in order to produce an equivalent photo.

In order to use equivalence theory as a framework for comparing the IQ of the above camera formats, it is instructive to group the camera exposures used on the micro four thirds format into the three possible scenarios listed below.

1. $S \geq 500$ and $N \geq 5.6$ on micro four thirds.

In this case, the mobile phone will be able to produce an equivalent photo. The IQ of the mobile phone and micro four thirds cameras will therefore be of the same order of magnitude, i.e., similar, and will differ only due to differences in the underlying sensor and lens technology, such as quantum efficiency, read noise per percentage sensor area, and lens aberrations. Perceived IQ at larger display sizes will also be of the same order of magnitude.

2. $S \geq 500$ and $N < 5.6$ on micro four thirds.

The mobile phone will be unable to produce an equivalent photo because an equivalent f -number does not exist; the DoF will be deeper and the required exposure duration (shutter speed) will be longer than that on the micro four thirds format. Although the system RP is dependent upon the lens aberrations, the micro four thirds format can potentially resolve more detail due to a higher diffraction cut-off frequency [see Eq. (53)], and this may become apparent at larger display sizes. Nevertheless, the total light received by both formats will be the same due to the longer exposure duration used on the mobile phone. Consequently, image noise will be of the same order of magnitude. Although thermally induced noise may be higher on the mobile phone, this contribution will typically be negligible at exposure durations of less than a second. However, the longer exposure duration required on the mobile phone could lead to camera shake in low-light conditions unless optical image stabilization is available. (The Google Pixel 2 does feature optical image stabilization.) Camera shake can greatly lower the system cut-off frequency and will become more apparent at larger display sizes.

3. $S < 500$ on micro four thirds.

The mobile phone will be unable to produce an equivalent photo because it will be unable to collect as much light as the micro four thirds format without overexposing the photo, irrespective of the camera exposure. Consequently, the photo from the micro four thirds camera is likely to be less noisy, and

this will become more apparent at larger display sizes. However, the mobile phone diffraction cut-off frequency can match that of the micro four thirds format, provided $N \geq 5.6$ is used on micro four thirds.

Comparison between any other mobile phone sensor format and larger camera sensor format can be achieved by replacing the S and N limits above with the appropriate values.

As discussed further in the section below, computational photography features available on modern smartphones such as the Google Pixel 2 must be disabled in order to apply equivalence theory in the manner above. With such features disabled, Fig. 11 shows a crude visual noise comparison between the Google Pixel 2 and the Olympus E-M1 camera, which is based on the micro four thirds format. Equivalent camera exposure settings were used; $S = 80$ and $N = 1.8$ on the Google Pixel 2 (left diagram), and $S = 800$ and $N = 5.6$ on the Olympus E-M1 (middle diagram). The same exposure duration $t = 1/2$ s was used in both cases along with equivalent focal lengths. In order to eliminate the different JPEG processing by the manufacturer from the comparison, the images were obtained from the RAW data by processing Adobe® DNG files using default settings and noise reduction minimized. The Google Pixel 2 image was subsequently upsampled to match the sensor pixel count of the Olympus E-M1 and crops from the images were then taken. Evidently, visual noise is of the same order of magnitude for both cameras when equivalent photos are taken.

The diagram on the right shows the same result for the Olympus E-M1 when the shutter speed was maintained at $t = 1/2$ s but the ISO setting was lowered to $S = 200$ and the f -number lowered to $N = 2.8$ in order to widen the aperture and let in two extra stops of light. Consequently, the image is less noisy. Equivalent settings do not exist on the Google Pixel 2, however, such limitations can be overcome by using computational photography features such as HDR+ discussed below.

5.3 Computational Photography

Due to the small equivalence overlap between a mobile phone camera sensor and a larger format sensor described above, the mobile phone camera is unable to receive as much light as a camera based on a larger format in most photographic situations. This places limitations on SNR, dynamic range, RP, and possible shallow DoF. Recent smartphones now employ computational photography techniques in an attempt to overcome such limitations, and this means that equivalence theory cannot always be applied.



Fig. 11 (Left) Google pixel 2; $S = 80$, $N = 1.8$, $t = 1/2$ s. (Middle) Olympus E-M1; $S = 800$, $N = 5.6$, $t = 1/2$ s. (Right) Olympus E-M1; $S = 200$, $N = 2.8$, $t = 1/2$ s.

Many computational photography techniques are based upon multiple camera exposures or even multiple cameras. For example, the high dynamic range (HDR) capture mode is sometimes the default mode on modern smartphones. While traditional HDR imaging works by combining frames from multiple camera exposures with differing exposure durations, alternative techniques have been developed, such as the HDR+ technology¹⁹ available on the Google Pixel/Nexus smartphones. This works by taking multiple (typically 10) underexposed frames, each using the same exposure duration. Underexposing each frame ensures that highlight data is preserved and the shorter exposure duration required by each frame minimizes camera shake. The frames are subsequently aligned using the lucky imaging technique and then averaged. Averaging frames averages the temporal noise and therefore increases SNR and dynamic range, a technique often used in scientific photography and astrophotography. The shadows are then lifted and the image adjusted, the result of which is a photo with improved IQ and an attractive appearance. Other computational photography techniques have been developed for smartphones that have dual cameras on the rear. For example, dual cameras can compute a stereo depth map that enables images with synthetic shallow DoF to be produced, mimicking the shallow DoF produced by larger-format cameras at wide apertures (low f -numbers). The portrait mode found on the Google Pixel 2 camera is able to achieve a similar effect by using only one camera; a convolutional neural network is used to compute a segmentation mask for the subject, and information from phase-detect autofocus is used to compute the depth map.

Equivalence theory cannot be easily applied when techniques such as those described above are employed, either in terms of translating camera exposure settings between different format sizes or when evaluating IQ. Although it is possible to disable such features (as done in the crude noise comparison of Fig. 11), doing so will not utilize the full capability of the smartphone in terms of photographic output. Instead, metrics such as those developed by DxOMark^{®20} can be useful when comparing the IQ between different smartphone cameras.

6 Conclusions

Although traditional exposure strategy is designed to be independent of camera sensor format, the same camera exposure used on different formats will not lead to images with the same appearance characteristics. Equivalence theory enables photographers to calculate the equivalent focal lengths, f -numbers, and ISO settings required on different camera formats in order to produce equivalent photos, i.e., photos with the same appearance characteristics including perspective, framing (AFoV), DoF, and shutter speed. The availability of equivalent camera exposure settings on different camera models establishes the equivalence overlap between them. Larger formats offer extra photographic capability beyond the equivalence overlap. Equivalence theory can therefore provide useful information about the photographic capability of a given camera format along with its suitability for a given application.

This paper has provided a complete mathematical proof of equivalence theory. The proof introduces a working equivalence ratio that is shown to be formally required whenever

the object plane upon which focus is set is brought forward from infinity. It has been demonstrated that the working equivalence ratio has a large numerical effect at macro object distances. It has also been proven that the same entrance pupil diameter is required on each format in order to produce equivalent photos and that equivalent photos are produced using the same amount of light. The extra photographic capability offered by a larger format such as shallower possible DoF or longer possible exposure duration corresponds to situations in which the smaller format is unable to provide a sufficiently large entrance pupil diameter or sufficiently low equivalent camera ISO sensitivity.

Since equivalent photos are produced using the same amount of light, equivalent photos have IQ of the same order of magnitude. The real-world IQ differences depend upon the underlying camera and lens technology rather than the amount of light. This paper has argued that equivalence theory should therefore be used as a framework for performing cross-format IQ comparisons and has demonstrated how equivalence theory can be used in practice to appropriately perform such comparisons. When utilized, the extra photographic capability offered by a larger format beyond the equivalence overlap with a smaller format can potentially provide higher IQ, at least in terms of SNR and RP.

7 Appendix A: Derivation of the AFoV Formula

For completeness, a derivation of the standard formula for the AFoV is presented here.

The geometry defining the framing or AFoV for a compound photographic lens is illustrated in Fig. 12. When focused at the object plane OP, the corresponding AFoV in the vertical direction has been denoted by α . The apex of the AFoV is located at the lens entrance pupil.⁹ Simple trigonometry yields an expression for α in terms of the object-space quantities h , s , and s_{ep} :

$$\tan \frac{\alpha}{2} = \frac{h}{s - s_{ep}}. \quad (56)$$

Utilizing the fact that the sensor height $d = 2h'$ and the magnification $|m| = h'/h$, the above expression may be written as follows:

$$\tan \frac{\alpha}{2} = \frac{d}{2|m|(s - s_{ep})}. \quad (57)$$

The AFoV in the horizontal or diagonal directions is straightforwardly obtained by replacing d , h , and h' with the appropriate lengths. The above expression is valid for a photographic lens provided the object-plane distance s and entrance pupil distance s_{ep} are both measured from the first principal plane of the compound lens assembly. The distance $s - s_{ep}$ describes the object-plane distance measured from the entrance pupil.

In order to obtain a more convenient expression for the AFoV, the distances s and s_{ep} can be eliminated in favor of the Gaussian magnification $|m|$ and pupil magnification m_p , respectively. This can be achieved in three steps:

1. In order to determine the expression for s , first consider the magnification expressed in terms of the

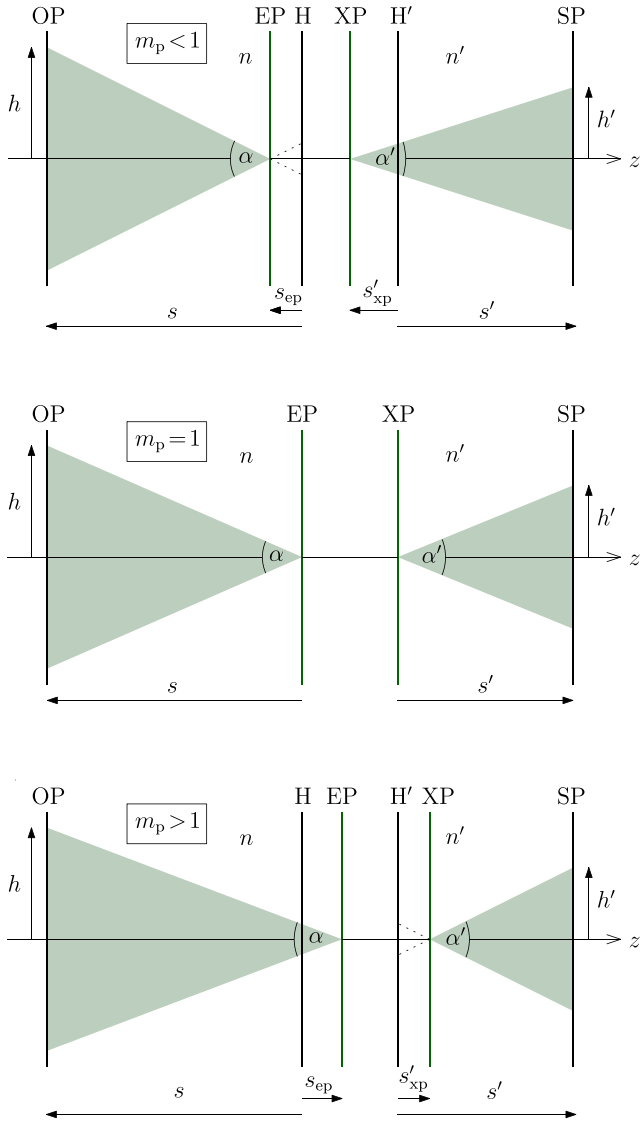


Fig. 12 Geometry defining the AFoV when the pupil magnification (a) $m_p < 1$, (b) $m_p = 1$, and (c) $m_p > 1$. The principal planes and pupils are not required to be in the order shown. Figure reproduced from Ref. 5 courtesy of IOP Publishing Ltd.

object and image distances measured from the first and second principal planes, respectively.⁹

$$|m| = \left(\frac{n}{n'}\right) \frac{s'}{s}. \quad (58)$$

Here, n and n' are the refractive indices of the object-space and image-space media, respectively. The distance s' can be eliminated by applying the Gaussian conjugate equation:

$$\frac{n}{s} + \frac{n'}{s'} = \frac{1}{f_E}. \quad (59)$$

The effective focal length f_E is related to the front and rear effective focal lengths f and f' according to the following expression.^{21,22}

$$\frac{n}{f} = \frac{n'}{f'} = \frac{1}{f_E}. \quad (60)$$

Combining the above equations yields the required expression for s :

$$s = \left\{1 + \frac{1}{|m|}\right\} f. \quad (61)$$

- In order to determine the expression for s_{ep} , it is convenient to eliminate the magnification variable $|m|$ by considering focus set at infinity. This scenario is illustrated in Fig. 13. The distance from the second principal plane to the exit pupil has been denoted by s'_{xp} . Utilizing the fact that the principal planes are planes of unit magnification, the geometry reveals that

$$s'_{xp} = (1 - m_p)f'. \quad (62)$$

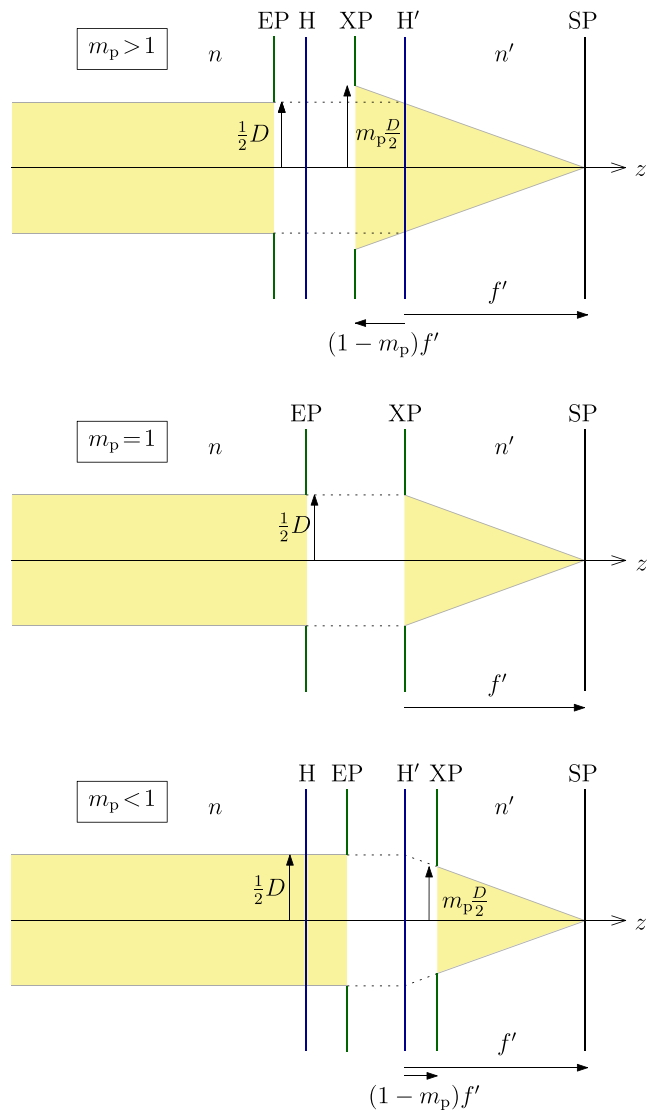


Fig. 13 Gaussian pupil and focal plane distances at infinity focus. (a) $m_p > 1$, (b) $m_p = 1$, and (c) $m_p < 1$. The pupils and principal planes are not required to be in the order shown. Figure reproduced from Ref. 5 courtesy of IOP Publishing Ltd.

The entrance pupil distance s_{ep} can now be obtained by applying the Gaussian conjugate equation to the pupil distances:

$$\frac{n}{s_{ep}} + \frac{n'}{s'_{xp}} = \frac{1}{f_E}. \quad (63)$$

Using the relationship between the focal lengths defined by Eq. (60) yields

$$s_{ep} = \left(1 - \frac{1}{m_p}\right)f. \quad (64)$$

3. Using the results of the previous two steps, the distance from the entrance pupil to the object plane can be written as follows:

$$s - s_{ep} = \left(\frac{1}{|m|} + \frac{1}{m_p}\right)f. \quad (65)$$

Substituting into Eq. (57) yields the standard AFoV formula defined by Eq. (6):

$$\alpha = 2 \tan^{-1}\left(\frac{d}{2bf}\right).$$

It should be noted that contrary to popular belief, it is the front (anterior) effective focal length f that should appear in this expression rather than the rear effective focal length f' or effective focal length f_E .⁵ However, $f = f' = f_E$ when the object-space and image-space refractive media are both air. The quantity b appearing above is the bellows factor defined by Eq. (7):

$$b = 1 + \frac{|m|}{m_p}.$$

Equation (8) defines a useful practical expression for $|m|$, which is obtained by rearranging Eq. (61):

$$|m| = \frac{f}{s - f}.$$

8 Appendix B: Focus Breathing

As illustrated in Fig. 14, a traditional-focusing lens achieves focus by movement of the whole lens barrel. In (a), the object plane (OP) is at infinity so that the rear focal plane passing through F' defines the image plane (IP), which coincides with the sensor plane (SP). In (b), the object plane has been brought forward so that the image plane moves behind the sensor plane a distance e . Here, the symbols l and l' have been used to denote the object and image distances measured from the principal planes H and H' , which satisfy the Gaussian conjugate equation. In (c), the whole lens barrel is brought forward in order to eliminate e . It follows that the required movement e satisfies

$$\frac{n_1}{l - e} + \frac{n_2}{f' + e} = \frac{1}{f_E}, \quad (66)$$

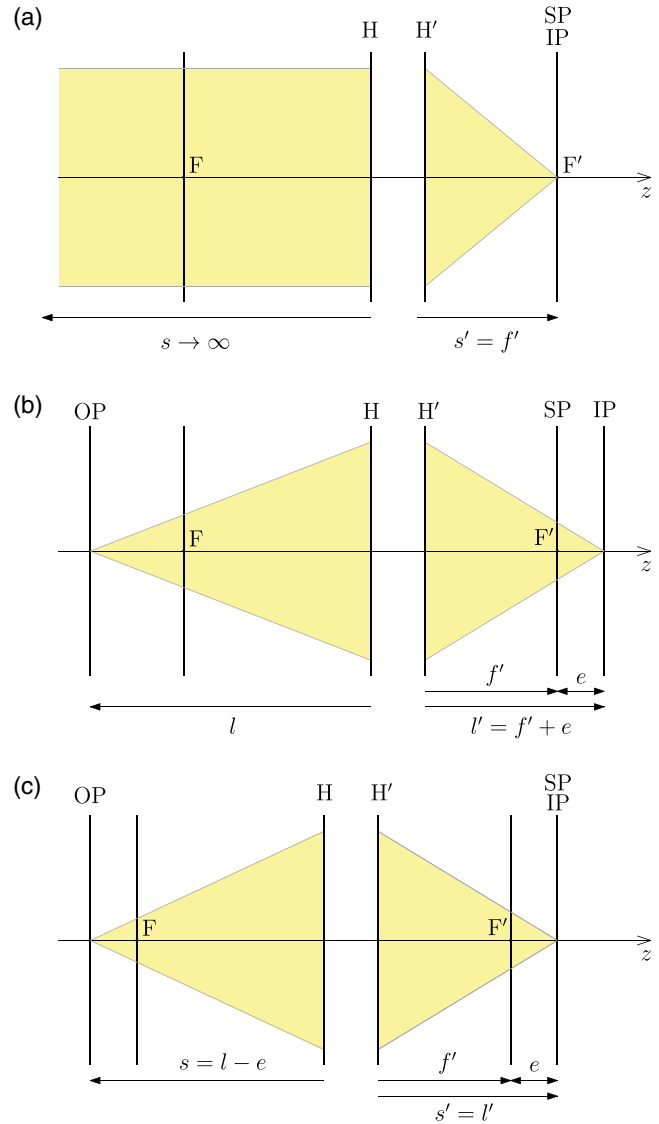


Fig. 14 Geometry of traditional-focusing movement. The figure is for illustration purposes only and is not to scale. (a) Focus set at infinity, (b) OP brought forward, (c) Focus set at OP.

where n_1 and n_2 are the object-space and image-space refractive indices, respectively, and f_E is the effective focal length. This leads to a quadratic equation for e . Subsequently, the familiar labels s and s' satisfying the Gaussian conjugate equation can be used to replace the distances $l - e$ and l' , respectively.

The important point to note from the above description is that the focal length of a traditional-focusing lens does not change when focus is set. However, the magnification $|m|$ increases at closer focus distances, i.e., as the object plane upon which focus is set is brought forward from infinity. This is evident from Eq. (8), $|m| = f/(s - f)$, where f is the front effective focal length and s is the distance from the first principal plane to the object plane after focus has been set; it is the latter distance that is reduced by the whole lens barrel being brought forward. Consequently, the bellows factor b defined by Eq. (7) increases, $b = 1 + (|m|/m_p)$. Therefore, the AFoV decreases according to Eq. (6), $\alpha = 2 \tan^{-1}(d/(2bf))$. Objects therefore appear

larger than expected. Although some photographers refer to any change in AFoV with object-plane distance as “focus breathing,” the term should not be used for traditional-focusing lenses because the increase in magnification and subsequent decrease in AFoV described above is a natural characteristic due to the physics.

However, many modern lens designs use front-cell focusing or internal focusing. The latter type of lens contains so-called “floating” elements that can be moved along the optical axis in order to alter the spacing between groups. Since the distance between refractive elements affects the total refractive power and therefore the effective focal length f_E and front effective focal length f , it follows that f necessarily changes when focus is set. Internal focus designs are particularly suited to fast autofocus applications since only a small movement of the floating element is required. Furthermore, the front and rear elements can be kept fixed so that the lens does not extend upon focusing.¹¹

The AFoV formula defined by Eq. (6) is independent of the type of focusing used. However, in the case of an internally focusing lens, the “new” value for f after focus has been set must be used. Typically, the focal length decreases at closer focus distances. This affects the AFoV value in two ways:

1. The AFoV is increased from its “traditional-focusing value” since f appears in the denominator of the AFoV formula.
2. The magnification $|m|$ and therefore the bellows factor b itself are reduced from their traditional-focusing values.

Both of these factors work to increase the AFoV compared to its traditional-focusing value. This phenomenon can be referred to as focus breathing.

The amount of focus breathing exhibited by an internally focusing lens depends very much upon the details of the lens design. Breathing is a design parameter that can be considered as important or unimportant when setting the design requirements. Some lenses are designed to minimize any change in AFoV with focus distance, while such changes are of minimal importance in the design of others.

For example, let the marked lens focal length (i.e., focal length at infinity focus) be labelled by f_A , and let the focal length and bellows factor after setting focus on an arbitrarily-chosen object plane be f_B and b_B , respectively. If the internally focusing lens is designed such that $b_B f_B = f_A$ upon setting focus at any object plane, the AFoV will always remain fixed and the “working” f -number (see Sec. 2.6) will not deviate from the marked lens f-number. In this case, focus breathing has been cleverly used to completely compensate for the reduction in AFoV that would have occurred due to the presence of the traditional bellows factor. However, if $b_B f_B < f_A$ upon setting focus, the traditional bellows factor will be over-compensated for and objects will appear to shrink. This effect generally becomes appreciable only when focusing at portrait object distances and closer. For example, some 70 to 200 mm zoom lenses on the full-frame format set at 200 mm reduce their focal lengths to around 160 to 170 mm at portrait object distances. Zoom lenses can also exhibit breathing properties that vary as a function of zoom.

In summary, the proof of equivalence is valid for a traditional-focusing lens. The proof is also valid for an internally focusing lens provided the “new” value for the focal length f is used in the bellows factor and AFoV formulae after focus has been set. In all cases, s is the object-plane distance measured from the first principal plane after focus has been set.

Although the effects of focus breathing are likely to be similar if the lens designs are similar, a general formula for the change in focal length that may occur when using an internally focusing lens does not exist because any change in focal length away from the value marked on the lens barrel depends very much upon the complexities of the specific lens design. Unfortunately, such information is not commonly reported by the manufacturers.

9 Appendix C: Derivation of the DoF Formulae

For completeness, a derivation of the standard DoF equations is presented here. The derivation follows that given in Ref. 8 but with the pupil magnification included.

Consider the geometry illustrated in Fig. 15. The object plane upon which focus is set is positioned a distance s away from the first principal plane H. The corresponding image plane coincides with the sensor plane at a distance s' from the second principal plane H'. Now, consider the upper diagram, where a point object is placed at the optical axis in front of the object plane at a distance s_n away from H. The rays from this point converge behind the sensor plane

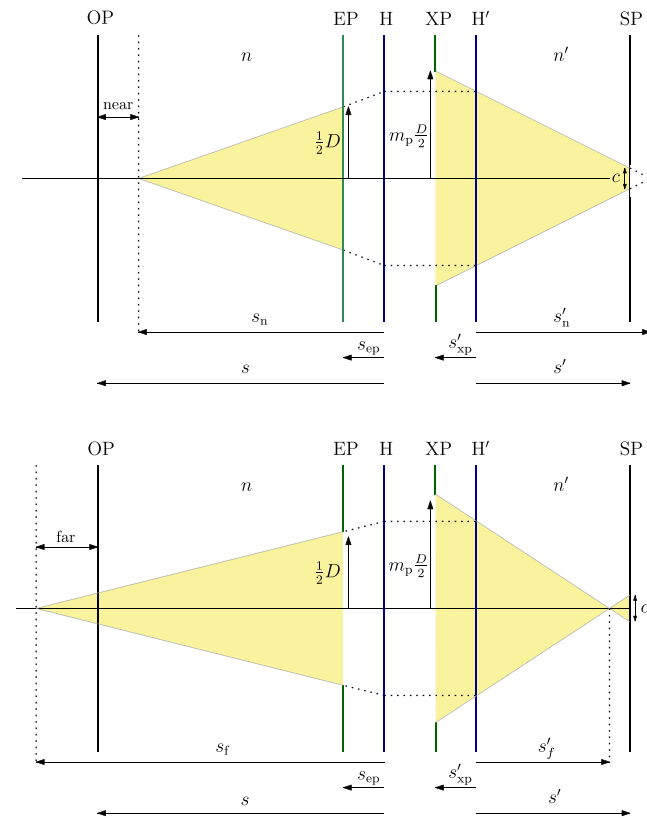


Fig. 15 Geometry for the DoF equations. The pupil magnification $m_p > 1$ in this example. The (a) near DoF and (b) far DoF boundaries are defined by the distances, where the blur spot diameter is equal to the prescribed acceptable CoC diameter c . The total DoF is given by the sum of the near and far DoF. Figure reproduced from Ref. 5 courtesy of IOP Publishing Ltd.

at a distance s'_n away from H' . At the sensor plane, the image of this point is a blur spot with diameter labelled c . In the lower diagram, rays from a point object positioned behind the object plane converge in front of the sensor plane. Such a point also appears as a blur spot at the sensor plane. At a particular object distance s_f measured from H , the blur spot diameter is again equal to c .

Provided the diameter c does not exceed a prescribed value, objects situated between s_n and s_f will remain acceptably sharp or in-focus at the sensor plane. This region defines the DoF. The blur spot with the prescribed diameter c is precisely the acceptable CoC described in Sec. 2.3.

A straightforward way to derive the DoF equations is to first project the CoC onto the object plane OP according to the system magnification.⁸ This yields a circle of diameter $c/|m|$ at the object plane. Graphical consideration of the upper and lower diagrams of Fig. 15 and the use of similar triangles then reveals two simple expressions for c :

$$\frac{c}{|m|} = \left(\frac{s - s_n}{s_n - s_{ep}} \right) D, \quad (67)$$

$$\frac{c}{|m|} = \left(\frac{s_f - s}{s_f - s_{ep}} \right) D. \quad (68)$$

The distance s_{ep} is defined by Eq. (64) of Sec. 8. The DoF boundaries s_n and s_f can be found by rearranging Eqs. (67) and (68):

$$s_n = \frac{|m|Ds + cs_{ep}}{|m|D + c}, \quad s_f = \frac{|m|Ds - cs_{ep}}{|m|D - c}. \quad (69)$$

The total DoF is the distance between the near and far boundaries, $s_f - s_n$. Algebraic manipulation leads to the following formula given as Eq. (27):

$$\text{total DoF} = \frac{2|m|Dc(s - s_{ep})}{m^2D^2 - c^2}.$$

Notice that the distance $s - s_{ep}$ is the object-plane distance measured from the entrance pupil rather than the first principal plane H . When the pupil magnification is unity, the term s_{ep} vanishes.

It is also useful to define the “near” DoF and the “far” DoF. These are the components of the total DoF measured in front of the object plane and behind the object plane, respectively. In other words, the near DoF is the distance $s - s_n$, and the far DoF is the distance $s_f - s$. Algebraic manipulation yields Eqs. (25) and (26):

$$\text{near DoF} = \frac{c(s - s_{ep})}{|m|D + c},$$

$$\text{far DoF} = \frac{c(s - s_{ep})}{|m|D - c}.$$

An alternative form of the DoF equations is obtained by introducing the quantity h defined by

$$h = \frac{Df}{c} = \frac{D^2}{c}. \quad (70)$$

Now, the DoF equations can be written in the form:

$$\text{near DoF} = \frac{(s - f)(s - s_{ep})}{h + (s - f)}, \quad (71)$$

$$\text{far DoF} = \frac{(s - f)(s - s_{ep})}{h - (s - f)}, \quad (72)$$

$$\text{total DoF} = \frac{2h(s - f)(s - s_{ep})}{h^2 - (s - f)^2}. \quad (73)$$

The ratio of the near DoF to the far DoF decreases from unity as the object-plane distance $s - s_{ep}$ increases. Eventually, a distance $s - s_{ep} = \mathcal{H}$ is reached at which the near to far DoF ratio reduces to zero because the far DoF extends to infinity. This value \mathcal{H} measured from the entrance pupil is known as the hyperfocal distance.⁸ The far DoF extends to infinity when the denominator of Eq. (72) is zero, in which case:

$$0 = h - (\mathcal{H} + s_{ep} - f), \quad (74)$$

and so

$$\mathcal{H} = h - s_{ep} + f. \quad (75)$$

Since $c = |m|D$ when the rear DoF extends to infinity, this can alternatively be expressed as follows:

$$\mathcal{H} = \left(\frac{D}{|m|} + \frac{D}{m_p} \right) N. \quad (76)$$

Substituting Eq. (75) into Eq. (71) shows that the corresponding near DoF is as follows:

$$\text{near DoF} = \frac{\mathcal{H}}{2}. \quad (77)$$

In other words, at the hyperfocal distance $s - s_{ep} = \mathcal{H}$, the far DoF extends to infinity and the near DoF extends to half the hyperfocal distance itself. According to Gaussian optics, focusing at the hyperfocal distance yields the maximum available DoF for a given combination of camera settings. This means that the far DoF and total DoF equations are valid for object-plane distances $s - s_{ep} < \mathcal{H}$. At \mathcal{H} and beyond, the far DoF and total DoF are both infinite.

Acknowledgments

The author would like to thank Professor R. Barry Johnson for many useful comments.

References

1. R. Butler, “What is equivalence and why should I care?,” 2014, <http://www.dpreview.com/articles/2666934640>.
2. R. Butler, “Equivalence is useful if you have any interest in light (and as a photographer, you probably should),” 2017, <http://www.dpreview.com/learn/2799100497>.
3. C. S. Johnson, “Science for the curious photographer: an introduction to the science of photography,” A K Peters Ltd., Natick, Massachusetts (2010).
4. J. James, “Essay on equivalence,” <http://www.josephjamesphotography.com/equivalence> (Retrieved on January 31, 2017).
5. D. A. Rowlands, *Physics of Digital Photography*, IOP Publishing, Bristol (2017).

6. CIPA DC-004, "Sensitivity of digital cameras," Camera & Imaging Products Association (2004).
7. ISO 12232:2006, "Photography - digital still cameras - determination of exposure index, ISO speed ratings, standard output sensitivity, and recommended exposure index" (2006).
8. R. Kingslake, *Optics in Photography*, Vol. **PM06**, SPIE Publications, Bellingham, Washington (1992).
9. R. Kingslake and R. B. Johnson, *Lens Design Fundamentals*, 2nd ed., Elsevier/Academic Press, Amsterdam, Boston (2010).
10. J. B. Williams, *Image Clarity: High-Resolution Photography*, Focal Press, Boston (1990).
11. S. F. Ray, *Applied Photographic Optics: Lenses and Optical Systems for Photography, Film, Video, Electronic and Digital Imaging*, 3rd ed., Focal Press (2002).
12. R. Butler, "Behind the scenes: extended highlights!," 2011, <http://www.dpreview.com/articles/2845734946>.
13. "DxOMark measurements for lenses and camera sensors," <http://www.dxomark.com/About/In-depth-measurements/Measurements/ISO-sensitivity>.
14. E. Martinec, "Noise, dynamic range and bit depth in digital SLRs," 2008, <http://theory.uchicago.edu/ejm/pix/20d/tests/noise>; <http://www.photonstophotos.net/Emil20Martinec/noise.html>.
15. W. J. Claff, "Sensor analysis primer - engineering and photographic dynamic range," 2009, http://www.photonstophotos.net/GeneralTopics/Sensors_&Raw/Sensor_Analysis_Primer/Engineering_and_Photographic_Dynamic_Range.htm.
16. <http://www.photonstophotos.net/Charts/PDR.htm>
17. E. M. Granger and K. N. Cupery, "An optical merit function (SQF), which correlates with subjective image judgments," *Photogr. Sci. Eng.* **16**, 221–230 (1973).
18. T. Steinich and V. Blahnik, "Optical design of camera optics for mobile phones," *Adv. Opt. Technol.* **1**(1–2), 51–58 (2012).
19. S. W. Hasinoff et al., "Burst photography for high dynamic range and low-light imaging on mobile cameras," in *SA '16: SIGGRAPH ASIA 2016 Technical Briefs*, New York, ACM (2016).
20. D. Cardinal, "How DxOMark Mobile scores smartphone cameras," 2018, <https://www.dxomark.com/dxomark-mobile-scores-smartphone-cameras>
21. J. E. Greivenkamp, *Field Guide to Geometrical Optics*, SPIE Field Guides, Vol. **FG01**, SPIE Publications, Bellingham, Washington (2004).
22. W. J. Smith, *Modern Optical Engineering*, 4th ed., McGraw-Hill Professional, New York (2007).

D. Andrew Rowlands received his PhD in theoretical condensed matter physics from the University of Warwick, United Kingdom, in 2004. He was awarded an EPSRC Research Fellowship at the University of Bristol, United Kingdom, and subsequently worked at Lawrence Livermore National Laboratory, US, and Tongji University, China. He recently authored the IOP book *Physics of Digital Photography* and is currently a research associate in computational photography at the University of Cambridge, United Kingdom.