Programmed Automatic Exposure in Canon EOS Digital Cameras

Douglas A. Kerr, P.E.

Issue 1.1
July 4, 2005

ABSTRACT

Many modern cameras, both film and digital, offer (usually as their basic mode of operation) a "programmed automatic exposure" mode. In this mode the camera, after measuring the luminance of the scene, sets both aperture and shutter speed with no further intervention on the part of the photographer. This article discusses the details of this operation as found in the Canon EOS 10D, 20D, 300D (Digital Rebel), and 350D (Digital Rebel XT) digital single-lens reflex (SLR) cameras. It also discusses the related matters of *exposure compensation* (*exposure bias*) and *program shift*, tools that allow the photographer to "tweak" the programmed automatic exposure control mode to deal with the special needs of a particular shot.

CAVEAT

The information on which this article is based came in part from Canon publications or from other authentic sources, and from practical tests conducted by the author, but the description is based on the author's own interpretation of that information, combined with his general knowledge of the topics and principles involved. There is no guarantee that all the details presented exactly correspond to the actual working of the subject cameras.

INTRODUCTION

As with most modern cameras (simple or complex, film or digital), the Canon EOS series of digital single lens reflex (SLR) cameras offer a "programmed automatic exposure" exposure metering mode (mode "P", actually designated "program AE"). In this mode the camera, after measuring the luminance of the scene, sets both aperture and shutter speed with no further intervention on the part of the photographer. The algorithm followed by the camera in balancing shutter speed and aperture, so as to attain the overall exposure determined as appropriate by the metering system, is guided by a set of curves often called "program lines". We will shortly look at a graphical presentation of those curves and the algorithm through which they are used as it applies to the Canon EOS 10D, 20D, 300D (Digital Rebel), and 350D (Digital Rebel XT) cameras. The general principles doubtless apply to various other Canon EOS digital SLR cameras.

APEX NOTATION

in this article, we will often speak as if the camera used "APEX" notation internally for the various numerical quantities being processed. APEX—the Additive System of Photographic Exposure—is a system that uses base 2 logarithmic scales and units for various factors involved in exposure. Its original intent was to facilitate the manual computation of exposure using only addition and subtraction.

For each value, a change in one unit corresponds to a 2:1 change in the underlying factor—in other words, it works in “stops”.

The various APEX quantities all have “value” in their names, a signal that the APEX form of the factor is involved. The APEX values of interest in this article are:

- Aperture value (A_V). This represents the aperture, as an f/number. A_V 0 corresponds to f/1.0. A larger value represents a smaller aperture.
- Time value (T_V). This represents the exposure time, or shutter speed. TV 0 corresponds to 1.0 second. A larger value represents a shorter exposure time (faster shutter speed).
- Exposure value (E_V). This represents the joint effect of an aperture and an exposure time. It is defined as A_V+T_V. A larger value represents less exposure.
- Speed value (S_V). This represents the ISO “speed” (sensitivity) of the film or digital sensor system. S_V 5 corresponds to ISO 100. A larger value represents a higher sensitivity (speed).
- Brightness value (B_V). Thus represents the metered average luminance (“brightness”) of the scene (or equivalent), as determined in “reflected light” metering. A larger value represents greater luminance.

APEX abuses

There are several abuses of the APEX system that we often encounter in literature about the topic of this article.

<u>EV for scene brightness</u>. The APEX quantity E_V (exposure value) represents the joint effect on exposure of an aperture and an exposure time (shutter speed). There is an unfortunate practice of also using it to mean scene brightness. When this is done, it means ”the scene brightness that, for an ISO sensitivity of ISO 100, and for the exposure meter calibration factor used in the camera being discussed, would under the standard exposure equation call for an exposure of the E_V mentioned”. I discourage that usage. There is a perfectly good APEX quantity for scene brightness: B_V.

<u>A_V for f/number, T_V for seconds</u>. Often we find aperture, given as an f/number, labeled “A_V” (or “aperture value”)), or exposure time (shutter speed), given in seconds, labeled “T_V” (or “shutter speed value”). That is not correct usage. Those symbols or names should only be used for quantities expressed in APEX logarithmic units. (In fact Canon’s labeling of the aperture priority and shutter speed priority shooting modes as “A_V” and T_V” is a form of this error.)

SCENE METERING

The first step in the exposure control process is to determine the luminance of the scene. This can be done in many different ways. The most basic way is to essentially determine the average luminance of the scene (at least the average of luminance readings taken at many points across the scene).

The cameras discussed here do not offer that basic approach. Rather they offer two simple approaches and a complex one:

- Partial metering. Here, the luminance readings taken at a certain set of points on the scene, all confined to a relatively small circle in the center, are averaged.
- Center-weighted average metering. Here, the luminance readings at a number of points across the scene are averaged, but with the reading at the very center given more "weight" in determining the average (think of it as having "several votes"), the readings from the points around it having less weight, the ones around those having even less, and so forth.
- Evaluative metering. Here, the luminance readings from many points across the scene are evaluated with a complex "intelligent" algorithm that seeks (at the conclusion of the whole process we describe in this article) to provide the best overall exposure result regardless of the distribution of brightness of the scene.

In any case, the output of the metering stage of the process is a single number, which I will call the "equivalent average scene luminance". That means that the subsequent parts of the process treat this just as if it were the average scene luminance involved in the most primitive exposure control metering system.

The relative "calibrations" of the three schemes are often such that if the camera is regarding a uniform-luminance "scene" (a frame-filling "gray card", for example), we can expect very nearly the same exposure to be set regardless of which one of the three metering schemes is in use.

THE STANDARD EXPOSURE EQUATION

Exposure determination normally, in effect, follows a simple equation that tells what the exposure (in the sense of exposure value, E_V) should be for a given equivalent average scene luminance and a given ISO sensitivity.

Expressed in APEX form, that equation is:

$$A_V + T_V = B_V + S_V$$

which can of course be rewritten as:

$$E_V = B_V + S_V$$

Note that this does not seem to contain the *reflected light metering constant* by choice of which a manufacturer can implement his own view on the "best" exposure for a given scene luminance and ISO sensitivity. (It does show up, for example, in the "non-APEX" form of the equation.) That's because the scale used for Bv is not a fixed one, but is contingent on the value of *K* chosen by the manufacturer.

EXPOSURE DETERMINATION AND EXECUTION

Figure 1 allows us to follow the entire remaining scenario of exposure determination and execution. It consists of two sections.

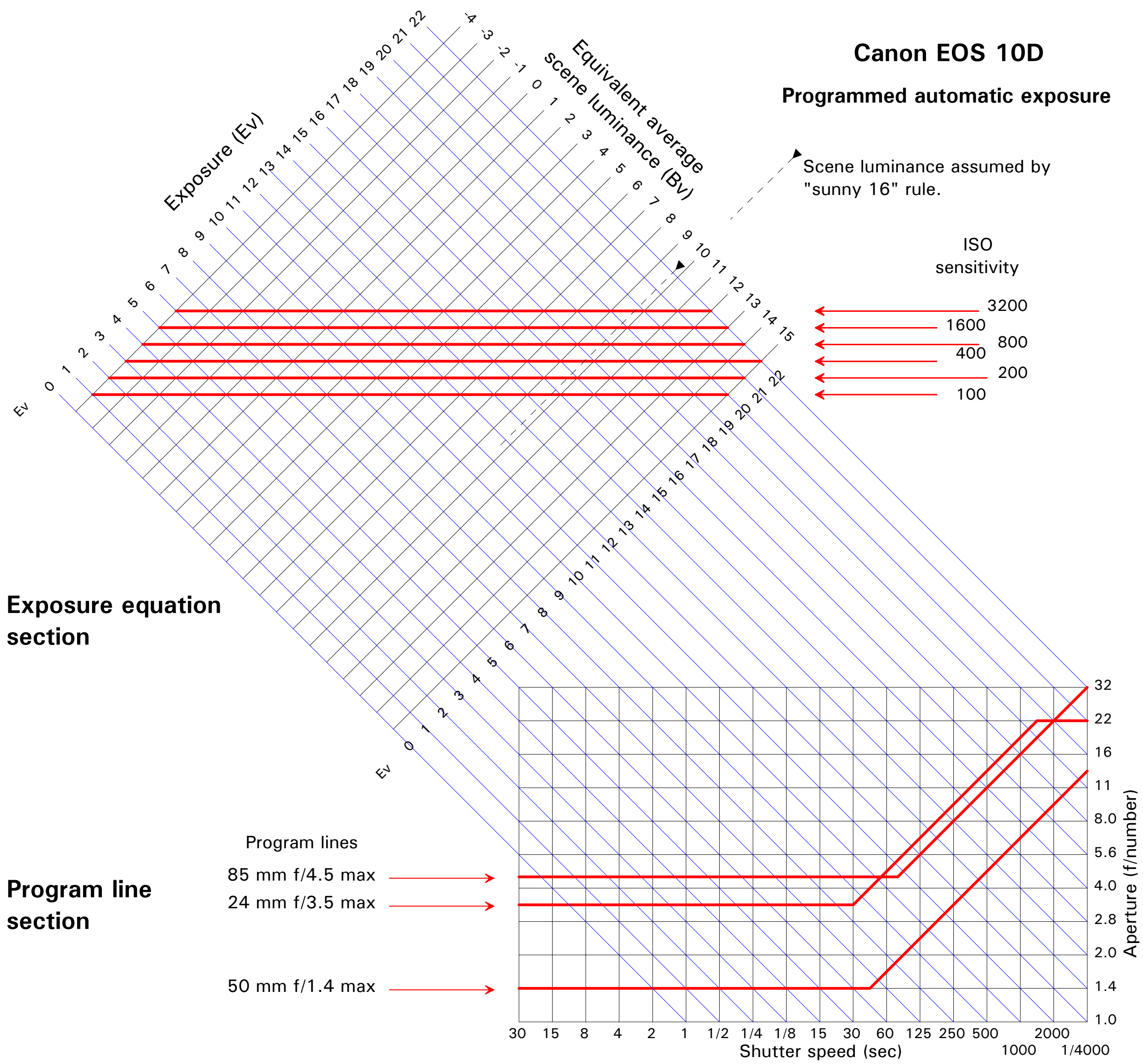

Figure 1. Programmed automatic exposure

Exposure equation section

The upper left part, labeled the "exposure equation" section, shows how the camera, having determined the equivalent average scene luminance (brightness) by metering, takes that result in combination with the ISO sensitivity in effect and determines the exposure, Ev, that seems desirable to properly generate the image. Recall that exposure, in this sense, means any combination of shutter speed and aperture with a certain joint effect.

This determination of the desirable exposure is, in effect, done by making a calculation guided by the standard exposure equation we discussed just above. The curves shown in that part of the figure implement that equation.

The calculation uses as its input the measured *effective average luminance* of the scene, shown here in terms of APEX brightness value, Bv. On our chart, we show a scale of Bv running from Bv -1 to Bv 15. This embraces the range of scene brightness over which the metering system in the cameras of interest is said to be usable.

By way of reference, a scene luminance of Bv 9.6 corresponds to the scene luminance ("under midday full sun at mid-latitudes") contemplated by the famous "sunny 16" rule for planning exposure outdoors in the absence of a meter. This value is shown on the Bv scale of the chart by a little black triangle.

The parameter of the calculation is the ISO sensitivity ("ISO speed") of the imaging system, which on these cameras can be set by the photographer over a range of ISO 100 to ISO 3200[1].

We reconstruct this stage of the exposure calculation this way. We start with the Bv value determined by the metering system. We fellow the corresponding black "Bv" line from that value downward and to the left until we hit the red line for the ISO sensitivity in effect. Then we follow the blue "Ev" line from that point downward to the right. The Ev value corresponding to that line is the Ev that the first stage of the calculations indicates as the exposure to be used for this shot.

Program line section

The second section of the chart, at the lower right, designated the "program line" section, represents the algorithm used by the camera in P mode in implementing the indicated Ev by choosing a particular shutter speed and a particular aperture, whose joint effect will correspond to that Ev. (This portion of the chart is adapted from the Canon manual for the EOS 10D.)

This graph has shutter speed (in seconds) as its *x* axis and aperture (as an f/number) as its *y* axis.

---

[1] In a "stock" EOS-300D. the ISO 3200 value is not available.

On this curve we see three red curves, called "program lines" [2]. They are illustrative of a large repertoire of such curves. Each is for a particular combination of the maximum available aperture of a lens and its focal length. Two of the lines shown, in fact, are for the same lens, a Canon EF-series 24-85 mm zoom lens. They correspond to the maximum aperture of the lens and its focal length at the two extremes of its "zoom" setting. The camera likely develops the curve to be used, on the spot, from knowledge of the maximum and minimum aperture and focal length of the lens in use, using an algorithm not shown on this chart (but which will be discussed later).

We follow the blue Ev line resulting from the first stage of the calculation until it hits the red program line for the lens in use. From that point, we read on the bottom scale the shutter speed to be used, and on the right-hand scale the aperture to be used. These are then set in the camera for use in the shot.

What are the principles upon which the development of the program line is based? From examination of the sample curves shown by Canon, it seems likely that they are these:

- Over the range where this is possible, to make a change in Ev, we "split the effort" equally between shutter speed and aperture. This dictates that the slope of the curve must be 1.
- The curve (unavoidably) flattens out at the maximum and minimum aperture available from the lens in use.
- The location of the curves are such that, for a given Ev, the shutter speed the curve calls for is greater for lenses of greater focal length. This is presumably because the longer the focal length, the faster the shutter speed that needs to be used to avoid an excess of motion blur from camera movement.

An example

We can follow the entire process for a specific illustrative case on figure 2.

We will assume the use of ISO 200 and an 85 mm f/4.5 lens, and that we have a metered equivalent average scene luminance of Bv 6. Following the black Bv 6 line to the ISO 200 curve, we see that the indicated exposure would be Ev 12. We follow the blue Ev 12 line to the program line section and its intersection with the red 85 mm f/4.5 program line. At this point, a shutter speed of 1/125 sec and an aperture of f/5.6 are indicated.

[2] "Program" here refers to the concept of "programmed" determination of shutter speed and aperture, the operation of the "P" shooting mode in these cameras.

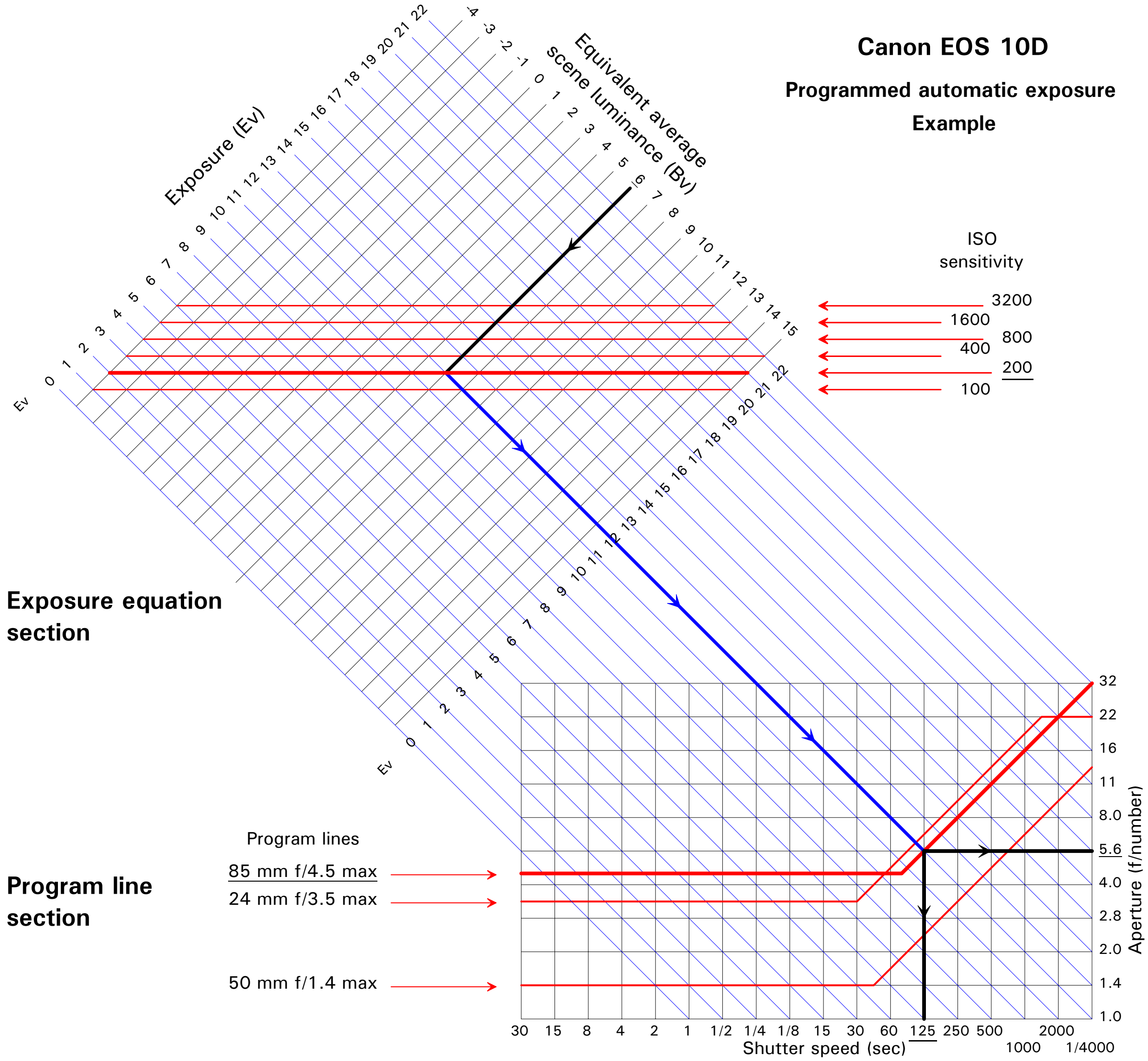

Figure 2. An example.

Granularity

The program line section is labeled with shutter speeds and apertures falling in a "full stop" series: adjacent values[3] nominally differ by a factor of 2 in their effect on exposure.[4] These cameras in fact can adopt shutter speeds and apertures falling between these values, but still only values from a preordained list. In some cases, the nominal increment between values in that list is 1/3 stop; in other cases, the

[3] The word "value" is used here in its general sense, not as an APEX code word.

[4] For shutter speed, there are small departures from a precise one-stop interval at various points to give "traditional" values. I do not know whether the camera treats 1/250 sec as exactly that or as 1/256 sec, the "theoretical" value for that step.

nominal increment is 1/2 stop.[5] In any case, the shutter speed and aperture determined by the program line are rounded to the nearest value in the series in effect. We will discuss some implications of this later.

INTERVENTION BY THE PHOTOGRAPHER

Exposure compensation

Often the photographer will wish to use the camera's metering system, in the programmed exposure mode, but will be aware of some situation of the scene that can be expected to cause the metering system to ordain too great, or too little, an exposure. Most cameras (including the ones of interest here) provide for a function called "exposure compensation"[6] (EC) to deal with this.

To employ this, the photographer sets an EC value (in Ev units). This tells the camera to use a greater, or lesser, exposure (in the direction and amount indicated by the EC setting) than that which the metering system would ordinarily utilize based on the metered scene luminance and ISO sensitivity in effect. Thus, if the photographer set an EC value of +1, then the camera will utilize an exposure one stop "hotter" than it would otherwise have used. (That means that it implements an Ev one unit less than indicated by the metering system.)

In Figure 3, we see, conceptually, how this happens. (The actual way in which this is done in the camera is not known to me.)

Here, in the exposure equation portion of the chart, we only show the equation line for ISO 100, but the normal line is accompanied by alternative lines for EC values of +1 and -1.

Note that if the EC has been set to +1, then for any given scene luminance, Bv, the Ev called for by the first stage calculation will be one unit less than with "no" EC setting (EC=0). If the EC has been set to -1, then for any given scene luminance, the Ev called for by the first stage calculation will be one unit greater than with EC=0.

The Ev output from the first stage is then turned into a shutter speed and aperture by the "program line" stage just as before.

The EC setting may be varied in steps of 1/3 or 1/2 stop, depending on the increment that applies to shutter speeds and apertures in the camera.

---

[5] For the 10D, the standard increment is 1/2 stop, but it may be changed to 1/3 stop. For the 20D, the standard increment is 1/3 stop, but It may be changed to 1/2 stop. For a stock 300D, the increment is 1/3 stop, and cannot be changed.

[6] This is known technically as "exposure bias".

Exposure compensation also applies to the working of the semi-metered exposure modes, aperture priority (Av) and shutter speed priority (Tv). It has no effect on the indication of the "little meter" in M (manual exposure) mode.

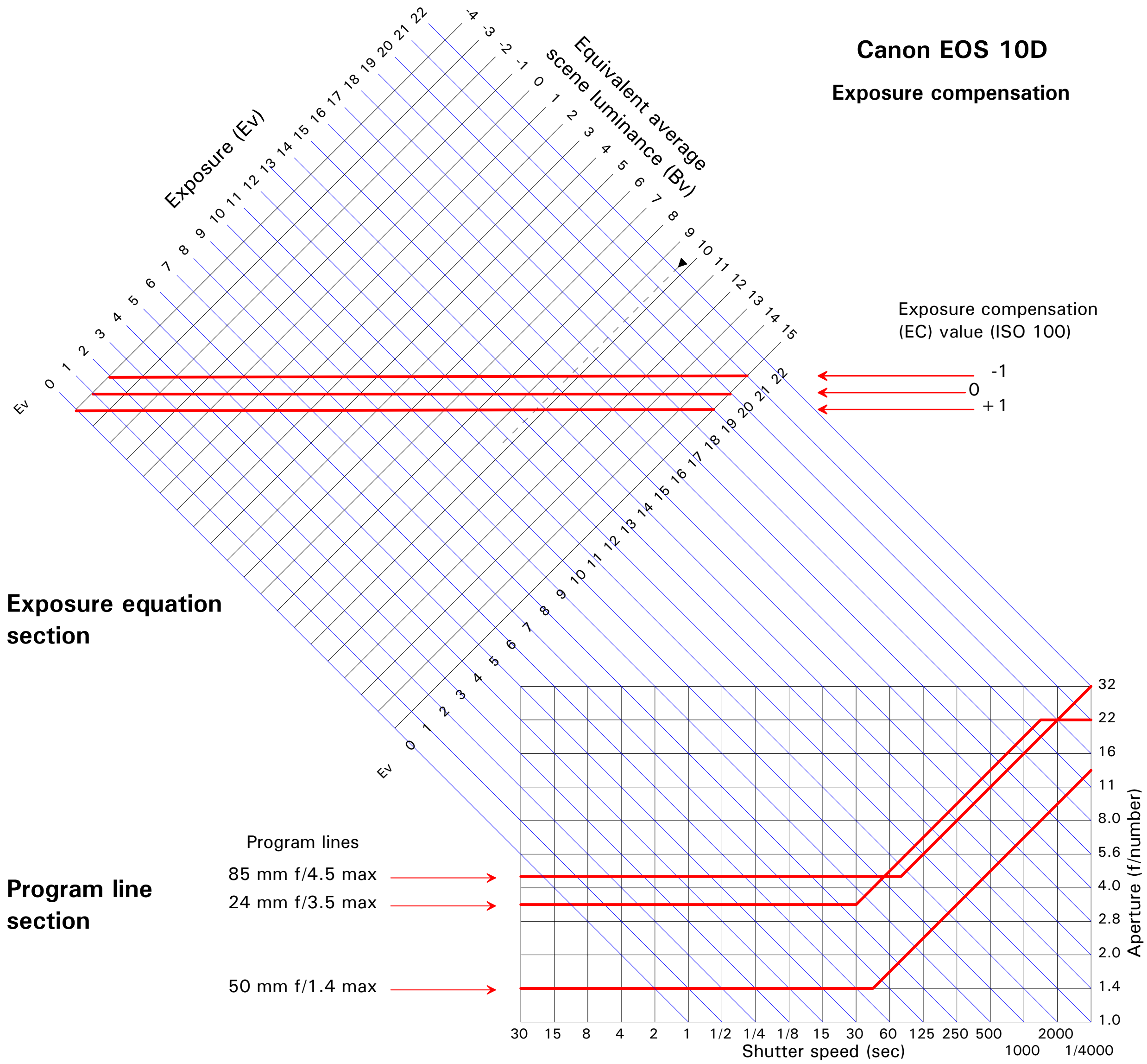

Figure 3. Exposure compensation (EC)

Program shift

Often the photographer will utilize the programmed automatic exposure mode and has no reason not to trust the system's choice of Ev, but takes issue with the balance of shutter speed and aperture called for by the system. Perhaps in view of the circumstances, he would rather use a faster shutter speed, and is willing to accept the larger aperture that would be needed to cooperate with that shutter speed to still yield the proper Ev.

Many cameras, including those of interest here, provide a function called "program shift" to deal with this need. This allows the photographer to shift the balance between shutter speed and aperture while still retaining the Ev called for by the first stage calculation. Figure 4 illustrates conceptually what this does.

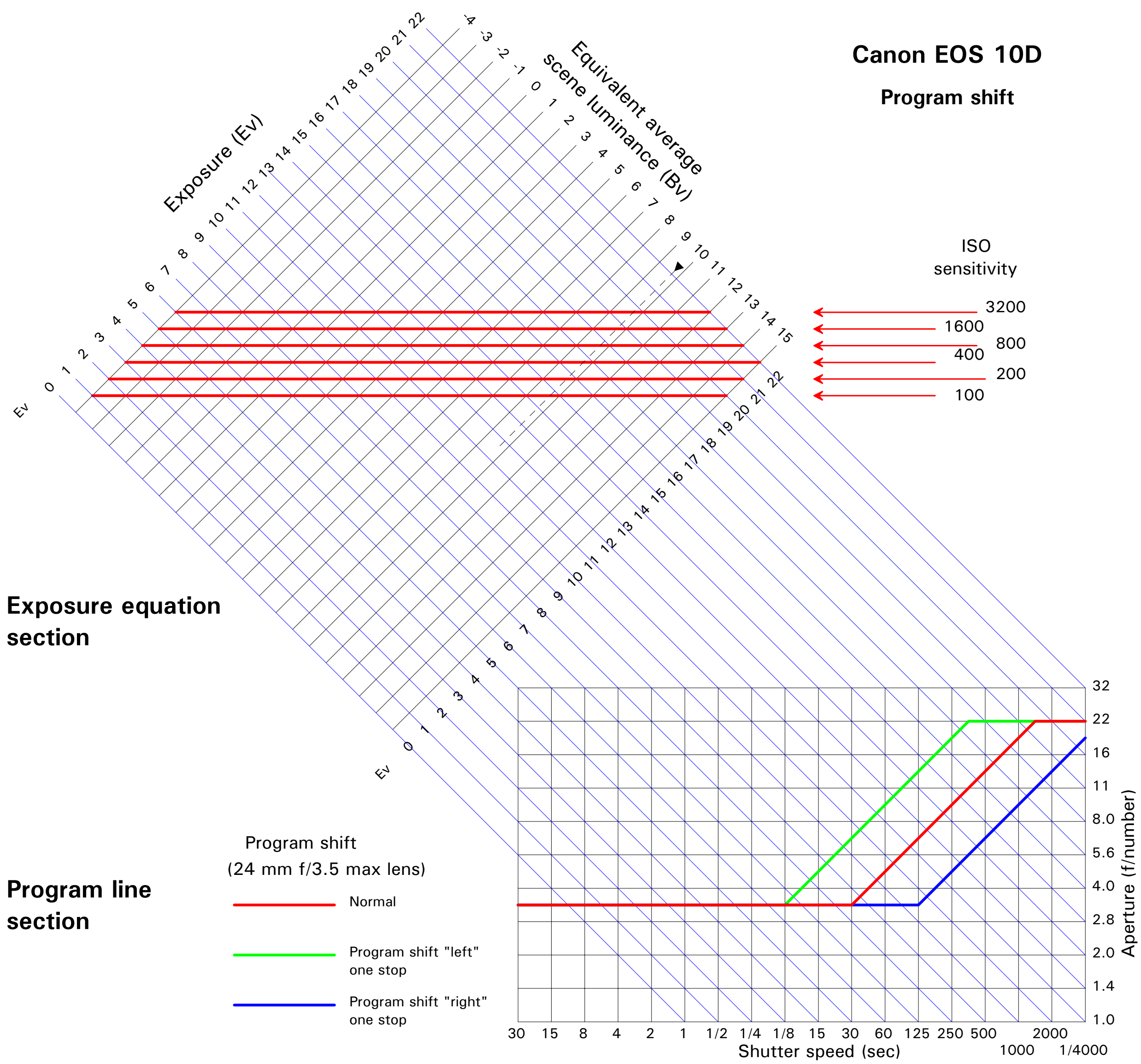

Figure 4. Program shift.

Here, in the program line section of the chart we see only the program line for one particular combination of focal length and maximum aperture (24 mm, f/3.5). But we see two variations of the "normal" curve. These correspond to a program shift of one stop to the "left" and one stop to the "right".[7] Note that the "one stop left"

[7] These are arbitrary names for the two directions, based on the shift of the curves on the chart. The names do correspond to the directions the command wheel on the cameras of interest must be turned to make that direction of shift.

curve would result, for any given Ev (sent down from the first stage), in a one-stop slower shutter speed but a one stop smaller aperture than if no EC were in effect.

The program shift works in 1/2 stop steps for any of the cameras of interest, regardless of whether the increment for shutter speed, aperture, and EC is 1/2 or 1/3 stop. An anomaly that can result from this when the other increments are 1/3 stop is discussed later.

METERING SYSTEM RANGE

For most sophisticated cameras, the manufacturer will state in the specifications a range over which the metering system works properly. This is typically in terms of a range of scene luminance, but is often (incorrectly) stated in terms of Ev, following the peculiar convention discussed (and denounced) above.

On the other hand, sometimes the limit on the range of the metering system is actually in terms of the range of Ev that the exposure equation stage of the process is able to output. In that case, the range of scene luminance for which the metering system is workable depends on the ISO sensitivity in effect.

In fact, in the case of the cameras of interest here, both kinds of limits are involved. It turns out that the "range" of the EOS 10D metering system (with an f/1.4 lens in place[8]) is:

- **Down** to a scene luminance of Bv -4 (which Canon describes as Ev 1, using the perverse convention mentioned above)

- **Up** to either (a) a scene luminance of Bv 15 (which Canon describes as Ev 20) or (b) a luminance which, in connection with the ISO sensitivity in effect, would lead to an exposure recommendation of Ev 22, whichever we hit first.

The two "Bv-based" limits are dependent on the maximum aperture of the lens in use, since the metering system sees the scene though that lens. I assume that for a smaller aperture than f/1.4, those two limits would increase accordingly. I assume that the "Ev 22" limit is not maximum aperture dependent.

Now, what happens in the case of luminances beyond that range? Is the output of the first stage just limited to the Ev output corresponding to the limit? Evidently not.

Based on practical observations in my studio, at the low end, the first stage is evidently able to give outputs that are as low an Ev as the camera could implement in turns of available shutter speed and aperture. So perhaps the limit means, "below this limit, the metering doesn't work accurately, and the calculated Ev might not be proper." I just don't know.

---

[8] The lens maximum aperture is involved since the luminance sensors see the scene through the lens, at its maximum aperture. Thus that aperture scales the sensitivity of the sensors to scene luminance.

At the high limit, I have no idea. I don't have a convenient way to test at a Bv of 5.3 stops brighter than "mid-latitude midday full sun". (My dermatologist wouldn't allow it, anyway.)

ANOMALIES

Program line anomalies

(Caveat: While the phenomenon reported in this section has actually been observed, the discussion of the mechanisms behind it is wholly conjectural on the part of the author.)

We noted before that the cameras of interest can only adopt shutter speeds and apertures from a preordained list, and that the nominal increment between successive available values is either 1/2 or 1/3 stop, depending on the camera and (where applicable) the increment choice made by the user.

Thus, if the EV line carrying the output of the first stage intersects the program line at a point where the shutter speed and/or aperture does not fall on one of the available values, the system presumably chooses the nearest available value(s) of each.

An additional possible complication is that I suspect that the output of the first stage of computation itself is quantized into values at a fixed increment before being fed to the second stage—maybe 1/2 stop, maybe 1/3 stop, maybe 1/6 stop. That is, the EV line cannot occur at just any arbitrary position.

Ideally, the way this would work would be that, as the scene luminance slowly decreases, first the shutter speed would get slower by one increment, and then later the aperture would increase by one increment, and so forth. But it doesn't seem to happen that way.

One possible reason is that the program line curves don't necessarily fall at locations that would bring this about. (If we take the curves as shown in the manual literally, they wouldn't.) And in fact to make them rigorously follow that scenario, different locations of the curves would be needed when 1/2 stop and 1/3 stop increments were in force. Canon may not have chosen to do that.

Another reason may be that the available values of shutter speed and aperture aren't exactly at the nominal spacings of 1/2 or 1/3 stop—they are at a series of customary "tidy" values. For example, from a shutter speed of 1/30 sec to 1/40 sec (consecutive values in the "1/3 stop" series) is 0.415 stop—not 0.333, which would correspond to 1/3 stop. If the camera actually uses the real available values of shutter speed and aperture—not the theoretical ones that would precisely follow the 1/3 stop or 1/2 stop series—this could also add to the unexpected behavior.

In any case, the overall result is that, if we gradually reduce the luminance of a test scene, we find that we may first encounter a shutter speed decrease of 1/3 stop, next an aperture increase of 1/3 stop, and then at the next critical point both a shutter speed decrease of 1/3 stop and an aperture increase of 1/3 stop.

Program shift anomaly

We mentioned above that the program shift function moves the program lines in increments corresponding to a change in the shutter speed and aperture of 1/2 stop, regardless of the increment in effect for shutter speed, aperture, and EC. an amount.

Thus, if an increment of 1/3 stop is in effect for those values, when program shift of one “click” is put into effect, its impact on shutter speed and aperture must be “rounded” so that the result is an available value. But the amount of program shift “ordered” still accumulates at 1/2 stop per click if the command dial is turned further.

As a result, if we turn the command dial to apply program shift, we may find that clicks alternately produce shutter speed and aperture changes of one “step” (1/3 stop) and two “steps” (2/3 stop).

In any event, the change in shutter speed and aperture are always consistent (resulting in a uniform Ev, except of course possibly for a small difference due to the fact that the shutter speeds and apertures are not always separated by precisely the nominal increment).

ACKNOWLEDGEMENT

Many thanks to Carla C. Kerr for her careful and insightful copy editing and proofreading of this daunting manuscript.

#