The Canon sRaw and mRaw Output Formats

Douglas A. Kerr

Issue 3
November 17, 2015

ABSTRACT AND INTRODUCTION

Many digital cameras (including many Canon models) offer an optional form of output file in which the data does not directly represent a "finished" image but rather is a more-or-less verbatim transcript of the digitized output of the collection of individual sensor photodetectors. This is referred to as the "raw" output. Recent middle- and upper-tier Canon dSLR cameras offer two alternatives to this output file type, described as the "sRaw" ("small raw") and "mRaw" ("medium raw") output files.

In this article I first give extensive background on various underlying topics. I next discuss the principles of the sensor arrangement that leads to the use of the raw output concept. Then I briefly review the "regular" raw output, what we do with it, and the advantages of operating in that mode. Next I describe the concept of the sRaw format, its important technical details, and why it is said to be beneficial. Finally, I introduce the mRaw format, and compare it with sRaw.

1. BACKGROUND

1.1 Color

Color is the property of light by which the human eye can distinguish different "kinds" of light. Color can be most understandably thought of in terms of two subordinate properties, *luminance* and *chromaticity*. Luminance is the property we can think of as telling us the brightness of the light of the color on interest (although rigorously, *brightness* means something slightly different).

Chromaticity is the property that describes the difference between "red" light and "green" light, and between "red" light and "pink" light. (I put those names in quotations because they in fact do not have exact meanings.)

Chromaticity is a two-dimensional property (in the mathematical sense), in that two numbers are required to specify it. Thus we recognize that, taking chromaticity together with luminance, three numbers are required to specify a color. That will be true for color-specifying approaches other than luminance-chromaticity.

And there are different approaches use, in two numbers, to specify chromaticity. Most of these can be defined by thinking of chromaticity as ranging over some type of two-dimensional plane.

Then, one approach to describing the chromaticity of a color is to give the *rectangular coordinates* of the point on that plane that corresponds to that chromaticity (they might be designated, in the familiar way, as x and y).

A second approach would be to describe the chromaticity of a color is to use *polar coordinates* (the radial distance from some origin to the chromaticity point of interest and the angle that described the direction of the point from the origin.

1.2 Chromaticity and chrominance

Related to chromaticity (and often confused with it because of the similarity of the names) is *chrominance*. Chrominance, like chromaticity, can be described in two ways, thinking of the property as ranging over a two-dimensional plane: with rectangular coordinates or polar coordinates.

Rather than trying to actually define chrominance directly, I will illustrate its nature in terms of its difference from chromaticity.

Imagine we have light of a certain color, which we think of as being described in terms of its luminance and its chromaticity. We now attenuate the light (such as by inserting a “neutral density filter” in the light’s path). The result is that the luminance of the color decreases, while the chromaticity (however described) **does not change**.

Now consider the original light but now think of its color in terms of luminance and chrominance. Again we attenuate the light. Now, the luminance has decreased but (if we use the rectangular coordinate approach to describing the chrominance), both coordinates of the chrominance **have also decreased** (in proportion to the decrease in luminance). If we used the polar coordinates approach to describing the chromaticity, we find that the “radial distance” (the “magnitude” of the chrominance) **has decreased** (in proportion to the decrease in luminance); the angle has not changed.

In both cases, the chrominance has, overall, decreased in proportion to the decrease in luminance.

1.3 Color spaces

In the use of the term we most encounter in conjunction with digital imaging, a color space is a defined structure by which a set of numerical values (usually three, sometimes more), spoken of as the *coordinates* of the color space, describe a color of light.

There are various approaches to the arrangement of a color space. It might describe the color in terms of its luminance and its chromaticity, with one coordinate for luminance and two for chromaticity (and, as discussed above, the latter might be by way of either rectangular or polar coordinates.

Or it might describe the color in terms of its luminance and its chrominance, with one coordinate for luminance and two for chrominance (and again, the latter might be by way of either rectangular or polar coordinates).

A different class of color spaces can be described as *tristimulus* color spaces. They describe a color in terms of the amounts of three kinds of light, of different chromaticity (called the "primaries" of the color space, that would need to be combined to produce light of the color being described (that is, light that would be recognized by a human viewer as having that color).[1]

An important subset of the tristimulus color spaces is the "RGB family", of which there are many specific variants. In all of them, the three primaries have chromaticities (varying among the different variants) that we can in any case speak of (imprecisely, of course) as being "red", "green", and "blue", and accordingly they are designated "R", "G", and "B".

1.4 The sRGB color space

A common color space used as the basis of the image files delivered by our digital cameras is the sRGB color space, a member of the RGB family (the designation in fact means "standard RGB", even though there are other standardized RGB-family color spaces). Here, three numerical values, designated R, G, and B, collectively spoken of as the sRGB *coordinates*, describe the color of each pixel. This color space is based on the concept that to make up the color of a certain pixel, we combine appropriate quantities of three kinds of light, with different (precisely-specified) chromaticities, known as the *sRGB primaries*.

The coordinate values R, G, and B, in fact tell the amounts of these three primaries to be combined to make up the color of interest, but in

---

[1] Purists reserve the term "tristimulus" for only a certain tristimulus color space, one used in scientific work to describe colors. It is called the "XYZ color space", and its primaries are invisible and thus imaginary. But purists may speak of it as the tristimulus color space.

a special way. They are actually a nonlinear transform of the values that actually describe the amounts of the three primaries.[2]

In this article, I will use r, g, and b to represent the actual values that describe "linearly" the amounts of the three primaries.

1.5 Trichromatic pixel sensors

It would be desirable in a digital still camera to have a sensor in which, at each pixel site of the delivered image, there was a photodetector that would report the color of the light landing there, by way of the coordinates of some defined color space (not necessarily, at this point, a recognized standard one).

One way to do this is with a *trichromatic pixel detector*. This has, at each pixel location, three photodetectors, each provided with a filter that shapes its spectral response. These filters would have their response peaks at wavelengths we can think of as broadly as "red", "blue", and "green". As a result, it is common to designate these three kinds of photodetectors as "R", "G", and "B", a usage that (as I will discuss later in a slightly different context) can lead to misunderstandings.

The outputs of these three photodetectors at a pixel location would, ideally, describe the color of the light there.[3]

This approach is actually used in some types of video camera, where there is a full set of one of these three types of photodetectors (one per pixel location) on each of three chips, all of which receive (in careful alignment) the image.

This approach is also implemented directly, with the three photodetectors essentially collocated at each pixel location, in some digital still cameras, notably the Sigma cameras that have the "Foveon" sensor system. Sony also has a patent on a structure that will accomplish this. And it is rumored at this writing that Canon will soon introduce cameras of this type.

---

[2] This nonlinear transformation is called, for historical reasons, "gamma precompensation".

[3] Alas, in reality, the three outputs do not consistently describe the color of the light, a matter that we will explore shortly

1.6 The color filter array (CFA) sensor arrangement

But, we don't have either of these arrangements on most digital still cameras. Rather, a clever compromise known as the *color filter array* (CFA) sensor is used.

Here, there is a photodetector at each pixel site. Each photodetector has one of three spectral responses, commonly identified as R, G, or B (because of the fact that their response peaks at wavelengths we can think of as broadly as "red", "blue", and "green"). They are arranged in a repetitive pattern. In the most common pattern (called the Bayer pattern, after its inventor, Bruce E. Bayer of Kodak), each block of four photodetectors (2 × 2) has one with an R filter, one with a B filter, and two with G filters.

The customary notation of R, G, and B, is used not only to identify the three types of photodetectors but their output values as well. The latter usage, in particular, can be misleading, as it suggests that these outputs correspond to the three coordinates of a color in the sRGB color space. They do not for two reasons:

- They are linear, not nonlinear (as are R, G, and B).
- They do not correspond (on a linear basis) to the amounts of the sRGB primaries needed to represent the color of interest.

Accordingly, to avoid these misunderstandings, I refer to these three types of photodetectors as D, E, and F, and their outputs as d, e, and f.

In order to determine the color of the light impinging on each pixel position, we need to have the d, e, and f values at that position.[4] But at each pixel position, we have only a **d**, an **e**, or an **f** value.

So before we can proceed to determine the color of the light impinging on every pixel position, we have to somehow fill in the missing values of d, e, and f. Simplistically, we could do this by interpolation between the values of that type in adjacent locations with that type of photodetector.

But the performance of such a simplistic system in actually allowing the determination of the color at each pixel location is not that good. So instead, we use a sophisticated algorithm that makes an "intelligent" estimate of the most likely value of each of the missing

---

[4] Actually, those values will not precisely describe the color of the light at that pixel location, but for the moment imagine that they can.

output values at a pixel location from examination of a number of the surrounding photodetector outputs (including those of other types). In any event, this process is usually referred to as “demosaicing”.

I note that the fact that there are twice as many “G” (to use the common notation) sensors as the other types recognizes the greater importance of the “G” outputs in determining the luminance of the color. This is important since the eye has a greater acuity with respect to the luminance aspect of light than the chromaticity aspect.

1.7 Metamerism

The human perception of chromaticity (and in fact, color is defined in terms of its human perception) depends on what is often called the spectrum of the light, which we can think of as a curve that tells us the relative amounts of power at different wavelengths in the light.

Now, there are an infinite number of spectrums that appear to the eye to have the same chromaticity (and thus, **do** have the same chromaticity). This matter is called *metamerism*. Kinds of light having different spectrums but the same chromaticity are called *metamers*.

1.8 Non-colorimetric sensors

As mentioned earlier, we would like for the values of the three outputs of our sensor at each pixel (direct or interpolated) to tell us the color of the light incident on the sensor. That means, among other things, that regardless of the spectrum of the incident light, so long as it has the same color, the sensor will report the same output values for it.

Such a sensor is said to be *colorimetric*, a word that essentially means “measures color”.

But it is not actually feasible to attain this in a practical sensor. Thus, when two samples of incident light have different spectrums but the same color (that is, they are metamers), the sensor may report different sets of output values. Or two samples of light with different colors may result in the sensor reporting the same set of output values.

This phenomenon is called *metameric* error. It cannot be avoided, or corrected for. However, in processing the outputs of a sensor, we take various steps to minimize its impact.

Because the three outputs of a sensor do not unambiguously describe the color of the incident light, they cannot rigorously be considered to be the coordinates of a color space (even a unique one for that sensor design). This is a major challenge in the processing of the sensor outputs into an image where we attempt to describe (for better or worse) the color of each pixel. Accordingly, I will speak of the sensor outputs as being the coordinates of a *pseudo-color space*.

1.9 The sYCC color space

It is often thought that in a JPEG file the colors of the pixels of the image are described in terms of coordinates of the sRGB color space. But they are not quite. Rather, they are described in terms of a transform of the sRGB color space called the sYCC color space.

“YCC” is short for “YCbCr”, the designation of a certain type of color space (whose coordinates are designated Y, Cb, and Cr). The designation “sYCC” means a certain standardized form of the YCbCr color space, just as “sRGB” means a certain standardized form of the RGB color space, and in fact the sTCC color space is derived from the sRGB color space.

The transform is mathematically very simple, and its inputs are the R, G, and B values of the color of interest.

The value Y is derived as a weighted sum of the values R, G, and B. Now if we were interested in the actual luminance of a color (which is also designated “Y”, although it is a different thing), we can calculate that as a weighted sum of r, g, and b (the “linear” coordinates of the sRGB color space). Thus we can perhaps understand how the sYCC value came to be designated “Y”.

But of course it is not the luminance of the color of interest. It is not even a non-linear form of the luminance of the color of interest. It is at best a sort of “pseudo-luminance”.

The coordinates Cb and Cr are the differences between the coordinates B and R and the pseudo-luminance, Y. Together, they describe a property that is something like the property of *chrominance*. But they together actually only describe a sort of pseudo-chrominance.

But despite the fact that Y is not the luminance of the color, and Cb and Cr do not describe its chrominance, Y, Cb, and Cr do rigorously describe a color. We can for example take the Y, Cb, and Cr values of a color in the sYCC color space and use the mathematical transform equations “backwards” to rigorously get the description of the color as the values R, G, and B under the sRGB color space.

1.10 Developing the JPG output file

The major stages in developing the JPG output file from the suite of sensor photodetector outputs are these:

- Demosaic the sensor photodetector outputs; that is, by intelligent interpolation, fill in the missing values so we have a d, e, and f value at each pixel location.[5] By the way, these values are usually held in a precision of 12 or 14 bits (depending on the camera model).

- At each pixel location, transform the d, e, f triple (which we can think of, with some reservations, as representing the color at that pixel location in the pseudo-color space of the sensor) into r, g, b, values that we consider to represent that color. There is considerable complication in doing this, but we need not be concerned with it here. These r, g, and b values are likely held in a precision of 12 or 14 bits. Recall that even if we do this in the best way possible, those r, g, b values to not rigorously describe the color of the light at that pixel position on the sensor

- At each pixel location, convert the linear values r, g, and b to their non-linear forms, as R, G, and B (there is a standard function for this defined for the sRGB color space). These may well be held to a precision of 12 or 14 bits.

- Convert this suite of R, G, B triples into sYCC form. The three components (Y, Cb, and Cr) for each pixel are at a precision of 8 bits.

- Discard a certain fraction of the CbCr pairs (the notion of “chrominance subsampling”).

- Pack the remaining data into a suitable format and compress it.

- Pack the resulting data into the appropriate format of a JPG file.

The discarding of a fraction of the Cb, Cr pairs, is desirable in that it reduces the amount of data to be recorded for the image. It is reasonable because the human eye is sensitive only to coarser changes in chrominance (which Cb, Cr sort-of describes) than to changes in luminance (which Y sort-of describes).

2. THE RAW OUTPUT FILE

While a JPG file can provide very useful images in a wide range of circumstances, in “difficult” situations, and when “very particular” results in image editing are desired, another approach can be

---

[5] If the photodetector uses a Bayer array, we actually have d, e1, e2, and f values. Usually the two kinds of e values are consolidated during the demosaicing operation, so we only have an interpolated d, e, and f value at each pixel location.

beneficial. Here, the collection of the digitized outputs of the entire array of photodetectors (all the d, e, and f values), described as “raw sensor data”, is delivered essentially verbatim in a special file format (called a raw[6] file).

Then, the remainder of the development process is performed not in the camera but rather in external software (sometimes called a “raw converter”). This has a number of advantages, including:

- If the exposure was not “ideal”, we can shift the mapping of photodetector output values onto image luminance values to "recover" from that. (One can of course only correct for a certain amount of exposure misadventure this way.)

- A demosaicing algorithm that “performs better” than that used by the camera can be employed.

- White balance color correction can be applied at this stage, when we may be in a good position to ascertain what “correction vector” gives us the most desirable result (perhaps even on a “season to taste” basis, or after measuring in the tentatively-developed image the chromaticity of an object in the scene we know to have “neutral” reflectivity).

- The finished image can be saved in a form that allows the image to be more precisely reconstructed for viewing or printing than with a JPEG file. For example, the image may be recorded in a TIFF file, in which we are free from the errors introduced by JPEG compression and can, if we wish, have all the coordinates expressed to 16-bit precision rather than the 8-bit precision of the JPEG file.

There are two prices paid for this improved capability:

- The raw data file delivered by the camera is substantially larger than a file carrying the developed image in JPEG form.

- There is another stage to be performed in the overall process of getting the image (one that is perhaps a bit complex for the user to execute).

Nevertheless, the use of the raw output, with “external development” of the finished image, is very popular among both professional and advanced amateur photographers. The “large file size” disadvantage

---

[6] This is often written as “RAW”, but that is inappropriate since it is neither an initialism nor an acronym. I use the form “raw” here, but I speak of a specific raw output file (used by Canon) as "Raw" to make it clear that this is a format name, not just a descriptive adjective.

has been greatly mitigated by the availability, at modest cost, of memory cards for cameras having quite large capacities, and the availability, at modest cost, of large-capacity hard drives for our computers.

And modern raw development software packages include many features that, in many “not too difficult” cases, can facilitate the user’s execution of the operation. Most of them even allow us to tell the software, “develop this one just as the camera would have” (the raw file generally includes all the parameters the camera **would have used** to internally develop the image).

3. CANON’S sRaw FORMAT

3.1 The history

Canon has for many years offered their Raw output format in most of their digital SLR cameras.

In 2007, in their EOS 1D Mark III digital SLR camera, Canon introduced a new output format option, the “sRaw” (“small raw”) format. It is now offered in all subsequent middle- and upper-tier Canon EOS dSLR’s (perhaps some models in other series as well).

A while later, a second similar format, now known as mRaw (“medium raw”) was introduced. But initially, this new format was known as “sRaw1”, and the original sRaw format then became known as “sRaw2”. But eventually these became known as “sRaw” and “mRaw”, respectively.[7]

I will begin with what was originally, and is now again, known as the sRaw format.

3.2 The sRaw format

Its salient attributes are:

- The delivered "raw" representation describes an image with exactly half the pixel dimensions (in each direction) of the camera’s "native" image (one-fourth the number of pixels). (By "native" image I mean one with approximately as many pixels as there are photodetectors in the sensor.)

---

[7] Remember, this is the same company that, having previously offered a landmark digital SLR called the EOS D30, a few years later brought out a wholly different one called the EOS 30D.

- The file size is on the order of half the file size of the "regular" Raw output file.

3.3 Generating the sRaw file

The basic procedure by which the sRaw file is generated, in the camera, is:

- We take each CFA cluster (four photodetectors with one "D", two "E", and one "F" type) and consider it to correspond to one pixel of our quarter-size delivered image (a "fat pixel").

- We take the average of the two “e” outputs, and use that as our “e” value. We transform each d, e, f triple into a form much like YCbCr (where d, e, and f are used as the transform inputs in place of R, G, and B). Not surprisingly, the resulting three values are usually labeled “Y”, “Cb”, and “Cr” (by parallel with the coordinates of a YCbCr color space proper), but to avoid misunderstanding, I will identify them as y, x, and z, respectively.

  This form differs from actual YCbCr encoding in that in YCbCr the three coordinates are derived from the nonlinear coordinates of the RGB color space, R, G, and B, whereas here they are derived from the three (linear) coordinates of the sensor “pseudo color space”, d, e, and f.

  We can. with a bit of imagination, consider y to be a (new kind of) pseudo-luminance of the pixel color, and x and z to give a (new kind of) pseudo-chrominance of the pixel color

- We record a "x, z" pair for only every other fat pixel, following the concept of "chromaticity subsampling" used in the sYCC scheme itself, at a ratio of 1:2. This specific chrominance subsampling pattern is formally designated “4:2:2” (don’t ask).[8]

- The resulting suite of data is compressed with a “reversible” (often called “lossless”) form of data compression (much like that used in JPEG) and saved in a file. This means that when the data is “decompressed”, the original data will be reconstructed, bit-for bit, without compression error. [The same is true of the regular Canon Raw data file.]

---

[8] We are able to do this because the human eye has lesser acuity for the *chrominance* aspect of an image than the *luminance* aspect. and y is sort-of luminance, and x, z describe sort-of chrominance..

3.4 The bottom line

Thus, in the sRAW format we retain the advantages of a raw output in allowing more precise manipulation of the data in post-processing, but substantially reduce the file size by:

- Supporting an image with only one fourth the pixel count.
- Exploiting the difference in human acuity for *"luminance"* vs. *"chrominance"* to reduce the amount of "chrominance" data to be included in the file.[9]

There is a penalty for the latter. Mechanically, the raw development software, to recover the full suite of raw values before further processing can be done, must interpolate between pairs of x,z pairs to reconstruct x and z values for pixels for which no x,z pair is recorded.

This interpolation does not actually retrieve for us the exact original x,z values for those pixels. So there is the potential of error in the retrieved raw data upon which the development of the image is based.

Note that, since for each pixel of the (smaller) image, there are D, E, and F ("E", "G", and "B") photodetectors, it would seem that we do not need to do "demosaicing" to get a d, e, and f output for each pixel in the development of the image. (But I am not certain of that.)

3.5 Is this a bargain?

The sRaw format was first introduced in the Canon EOS 1D Mark III digital SLR, a top-end digital camera. The "White Paper" produced by Canon USA for this camera mentions the availability of this new output format, and expresses its value this way:

> Photographers often require an image of lesser pixel dimensions than the maximum size for the camera (especially if the image is to be turned into deliverable prints of modest size). The use of the sRaw output file brings to such work the flexibility and control of image development given by use of the Raw file, but with a file size half the size [of the conventional image for the full-size image], thus easing storage space requirements.

---

[9] Of course, even in the conventional sYCC file, with inputs of R,, G, and B, the coordinates are not really (for Y) luminance and (for Cr and CB together) chromaticity, and in this case, where the input values are d, e, and f, the results are even further from representing luminance and chromaticity, but still the notion is somewhat valid. Sort of.

So, by being willing to have an image with half the pixel dimensions (one-fourth the pixel count) of that supported by the Raw format, we get a file that is half the size of the regular Raw file for the full-size image (and with slight jeopardy to the accuracy of the “pseudo-chrominance” values).

Wow! What a deal!

For later models, the Canon USA White Papers generally mention the availability of the sRaw output without any discussion of its advantages.

THE LESS-SMALL BROTHER—mRaw

Not too long after the introduction of the original sRaw format, in 2008, Canon introduced a second sRaw format in much the same vein, offering both as options in their later cameras. It gives a file that is "less small" than that of the original sRaw format.

In its first outings (in the EOS 50D and EOS 5D Mark II), it was called “sRaw1”, with the version we have already described then being called, in distinction, “sRaw2”. Eventually (starting with the EOS 7D), the new format became known as the mRaw format (“medium Raw”, I suppose), and the original format became again called “sRaw”.

I am less clear as to how this works than for sRaw.

The mRaw format differs from sRaw in these prominent ways:

- It represents an image having about one-half the pixel count of the image directly developable from a regular Raw file. (This image has roughly 70% the linear resolution of the “native” image.)

- The *chrominance* (x,z) is subsampled at a rate of 1:4 (that is, for every four pixels, there is only one set of x and z). This chrominance subsampling arrangement is formally designated “4:2:0” (don’t ask).

- The file size is approximately 3/4 the size of the Raw file for the corresponding full-size image, about 1.5 times the size of the sRaw file (for the “half-dimensions” image).

From some reverse engineering, I suspect that here, in the process of converting from photodetector outputs to the YCbCr-like form, there is a transformation from the "sensor" pseudo-color space to some other pseudo-color space (I have no idea exactly what). (It cannot actually be a true color space, since the d, e, f “color space” is not a true color space, as its values do not rigorously describe a color.)

As to the advantage of this format over other possibilities available to the photographer, I would hate to volunteer an opinion.

CONCLUSION

To this old telephone engineer, both sRaw and mRaw look like fairly clever solutions looking for a problem. The fact that sRaw manages to provide the raw data (almost) for a one-quarter pixel count image with half the file size of the regular raw representation for the full-size image would not, in my world, win a prize at an intermediate school science fair.

Still, if for a particular photographic task one needs only half the native resolution of the camera (one-fourth the pixel count), but wants to be able to process that image with all the subtlety that working with the raw data gives us, and save some storage space, then why not.

And, as always, I may be missing something.

Ah, the mind of Canon.

ACKNOWLEDGEMENTS

Much of the credit for “our” knowledge of the Canon sRaw (and mRaw) formats is due to Dave Coffin, the developer of *dcraw*, an open-source program that decodes raw files (of many types) and “develops” the images they describe. He has had to do this largely based on some very difficult “reverse engineering”.

Next in the food chain comes Laurent Clévy, who has “reverse engineered” the source code in *dcraw* for the Canon raw files (including for sRaw and mRaw), and has analyzed and described its implications. His input also included analysis of the various tags in the file structure using *exiftool*, developed by Phil Harvey, perhaps the dean of our metadata reverse engineers.

Joining the fray was Yang Gao, who, among other things, recognized the parallel between the encoding of the raw data in the sRaw format and that of the R, G, and B data in the YCbCr color model.

I came in late, and only had to figure out what it all means, and explain it.

Finally, thanks to my wife Carla C. Kerr for her keen and insightful copy editing of this difficult manuscript.

#