Digital Cameras – How many mega-pixels is optimum?

David Bartlett, 9th May 2005

Abstract

All too often we hear the first parameter by which a digital camera is specified is the number of mega-pixels in the sensor, and yet this is probably not the most important factor governing performance and will indeed become less important as sensor sizes grow.

So why this fixation with mega-pixels? Because the perception is that it governs the maximum size of the resulting usable image. This perception is not entirely true – there are other important factors that come into play.

In this article I look at the physics of image formation and assess what is important from the discerning photographer's point of view. This includes a review of lens performance, the operation of the sensor and ends with some thoughts on how many mega-pixels make technical sense.

How the camera works

In essence a camera is a box with a lens that projects an image of the subject onto a plane consisting of a material or device that is able to capture the light. Traditionally the image has been captured using film, but in a digital camera an electronic sensor is placed at the point where the image is formed. This sensor converts the image into electrical signals which are processed and stored on a memory card when the shutter is pressed. There are several important elements needed to capture a quality image:

- The lens. How sharp is it? How good is the image formed?
- The sensor physical size. How big should it be?
- Sensor performance, dynamic range and noise.
- Other factors such as post processing of the image, anti-aliasing filter, colour artefacts, white balance etc.

The last two are equally important in the context of capturing a quality image, but this article will consider the first two in more detail since these address the subject directly: "How many mega-pixels is optimum?"

The Lens

How lenses work

The lens is one of the most important parts of the camera – after all it is responsible to projecting the image of the subject onto the sensor.

It comprises one or more (generally several) glass elements which collect the light rays from the subject and refract them to form the image. There are many different makes, models and types of lenses but they can all be characterised by two physical parameters: focal length and aperture.

Focal Length

The focal length represents the distance between an ideal single refracting lens and the image it forms at the focal plane. The longer the focal length the bigger the image; or looking at it the other way around the narrower the angle of view for an image of fixed size. Zoom lenses allow the user to vary the focal length over a range.

The focal length of the lens works with the size of the image sensor to define the angle of view. For a 35mm camera a "standard" lens is generally considered to be about 50 mm focal length. For a 6x7 medium format camera it's about 90 mm, for an APS sized camera around 35 mm and for digital cameras with smaller sensors as short as 4 mm could be considered a "standard" lens.

Aperture

The aperture is the ratio between the focal length and the diameter of an ideal lens. Smaller indicates bigger diameter and a brighter image. The aperture represents the ability of the lens to collect light from the subject. Most lenses incorporate an adjustable iris (aperture) which allows the photographer to adjust the amount of light that the lens collects. The aperture affects the depth of field of the lens and so it may also be used in an artistic sense.

A professional lens for a 35mm camera would typically cover the range of apertures from *f2.8* to *f22*. A consumer lens would typically cover the range *f4.5* to *f16*. Compact digital cameras typically have a smaller range of apertures, maybe *f2* to *f8*.

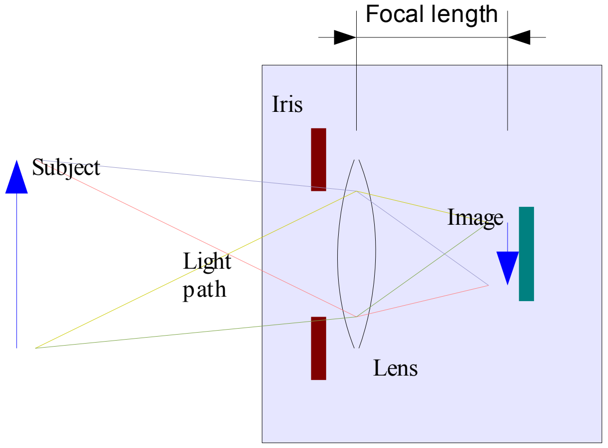

Figure 1: Optical path and image formation

Measuring lens sharpness

There are two widely used measures for lens sharpness – or resolving power:

- MTF or Modulation Transfer Function
- Limiting resolution based the Rayleigh principle

Both are measured (typically) in lines per millimetre, (strictly line-pairs per millimetre, or line cycles per millimetre) where the lines are pairs of black and white contrasting lines and the resolution is the spacing of the lines measured at the sensor plane.

Note that lens sharpness varies with aperture, position within the image and, for zooms, focal length too.

MTF

The MTF response is like the frequency response of an electronic filter but it describes the accuracy with which the lens reproduces detailed contrast within the subject.

In an MTF chart the response curve of the lens is plotted as contrast percentage. Curves are generally produced for different apertures and different line densities – measured in “lines per millimetre”. The MTF response represents the percentage of contrast between the line pairs that the lens has reproduced in the image. The 50% point represents where the lens has reproduced the line pairs with half as much contrast as possible. The lower the value the less contrast in the image.

The 50% point is equivalent to the half power point or -3 dB power point in an electronic filter response. The lens has a complex two-dimensional response, so it is not possible to attach a single MTF number to it.

Limiting resolution

Although the 50% point represents an important quantitative measurement of lens performance, the eye is able to discern detail in the image even when the contrast has fallen to as little as 5%. A measure sometimes used to indicate the maximum resolving power is based on the Rayleigh criterion, which represents a contrast of around 9%. For lower contrast than this there is little useful information in the image.

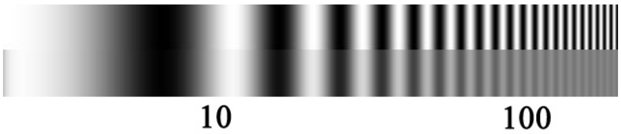

Figure 2: Source and image contrast compared

Figure 1 illustrates how the image contrast falls off as line density increases. The above figure has a 50% MTF around 40 lines per millimetre and a 9% MTF at approximately 150 lines per millimetre. It is representative of the resolving power of a professional slide film emulsion.

Image Formation, and limiting resolution

When light passes close to an object it is diffracted (bent), and therefore some of the light passing through the iris of the lens is diffracted and the resulting image is slightly blurred. The blurring of the image and the pattern formed by the diffracted light depends on the size of the objective lens, the focal length of the lens and the wavelength of the light.

For a circular aperture it takes the form of a central spot surrounded by a series of concentric circles. This is known as Airy's disc after the 19th century Astronomer Royal Sir George Airy who first solved the rather complicated wave equations and showed that the angle subtended between the centre of the image and the concentric dark and light rings surrounding it depended only on the diameter of the objective lens.

Since the lens aperture, A, is defined as the ratio between the diameter of the objective lens and the focal length, Airy's results can be simplified and approximated as follows:

$$r = 1.22 \cdot 5.6 \cdot 10^{-7} \cdot A \simeq 6.8 \cdot 10^{-7} \cdot A$$

where r is the minimum linear distance at the focal plane between two objects such that they can just be distinguished from one another. This is the Rayleigh criterion.

Note that the larger the aperture, the greater the minimum distance between two resolvable objects. Allowing a margin of error this leads to the practical approximation that the resolving power of the lens is at best a little less than A µm where A is the lens aperture. Put another way the resolving power of the lens is at best approximately *1400/A* lines per millimetre. Thus at *f14* a lens can resolve at best approximately 100 lines per millimetre.

By resolve here we mean that there is just sufficient contrast between the line pairs (9%) that the eye can still see distinguish them apart. The detail will appear very flat and almost completely washed out, as illustrated in Figure 2 – but the image does contain information that the eye can interpret.

Practical Lenses

Real lenses are designed to a compromise which means that stopped down at small apertures the resolving power is often aperture limited, whilst wide open the resolution is usually limited by the quality of the optical design. The best performance is usually achieved at around 2 stops below maximum aperture.

Since lens performance varies across the image frame from centre to edge and also for lines running radially as opposed to tangentially (equidistant from the centre), manufacturers normally produce MTF charts that show lens performance across the frame for two apertures (generally wide open and *f8*) and at two line densities (normally 10 and 30 lines per millimetre).

In order to illustrate the performance in a different way Figure 3 below shows MTF performance plotted against lines per millimetre for three different apertures for a fictitious lens. This chart shows how the optimum aperture (usually around *f*8) has good sharpness and contrast. Wide open the contrast is lower but performance does not fall off as quickly at high resolution. When stopped down the lens resolution is aperture limited and its performance falls off quickly with increasing resolution.

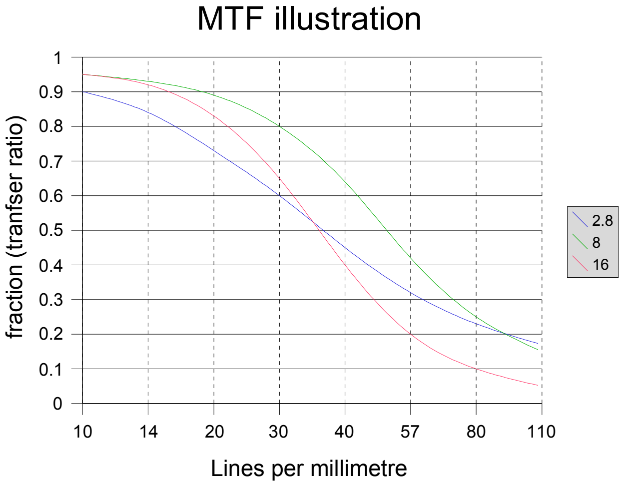

Figure 3: MTF against lines-per-millimetre

Another way of looking at lens performance is to plot the MTF against aperture. Figure 4 shows how the MTF might vary for a typical (fictitious) lens. The curves are plotted for different line densities (10, 20, 40 and 80 lines per millimetre) showing MTF against aperture.

Practical lenses typically show the inverted bathtub curve for resolving power with the optimum performance being around 2 stops down from maximum and performance falls off on either side of this peak. Values above 0.5 are considered “sharp” and values above 0.05 may be considered “resolvable”.

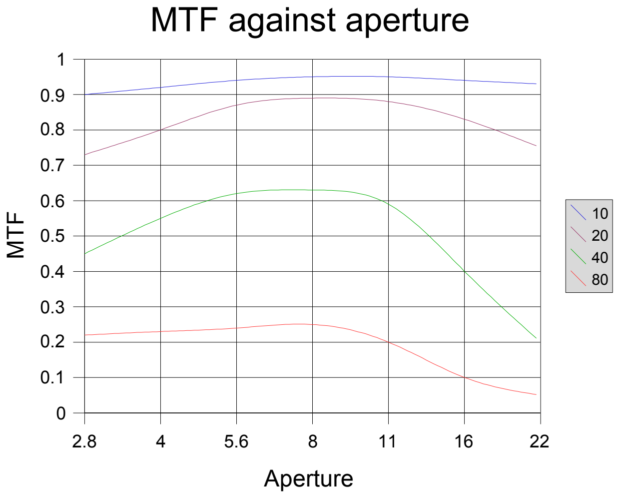

Figure 4: MTF at various resolutions

Another aspect of lens design is the “image circle”. This is the size of the image the lens is designed to create. The objective is to achieve linear even brightness across the image whilst preserving sharpness and contrast – all within a budget. This involves a number of design compromises, but the larger the required image circle the harder it is to achieve consistently high performance. Thus lenses designed for a smaller image circle usually exhibit better sharpness and contrast, and wider apertures can often be achieved. This is especially true with budget consumer lenses where cost is a major issue.

However, the aperture effect on sharpness always remains as a fundamental limitation.

The Sensor

The sensor in a digital camera should be matched to the lenses used, and there is little point in using a sensor with a higher resolution than that of the image formed by the lens.

Overview

The sensor is a flat rectangular electronic device that converts the incident light into an electrical signal which is digitised, processed and stored on a memory device in the camera. It consists of a two-dimensional array of photodiodes. Conventionally each pixel in a digital image comprises the three colour components red, green and blue (RGB), which together render the perceived colour for that pixel. This means that a 6 Mpel (mega-pixel) image on the computer comprises 6 million pixels each comprising red, green and blue colour components.

The sensor in the digital camera needs to measure the red, green and blue components separately, therefore each photodiode in the sensor array measures only one of the three colours. It has become conventional to call each photodiode a pixel, even though it measures only one of the RGB colour components. Thus a 6 Mpel camera sensor comprises 6 million photodiodes of which some are red, some green and some blue. In processing the measurements from the sensor array an image consisting of 6 million pixels in which each one has all three components present is done by combining and interpolating the measurements from adjacent groups of photodiodes. This process effectively converts 6 million measurements into 18 million values.

There are three main architectures for laying out the photodiodes on the sensor and distributing the red, green and blue measurements amongst them:

- The Bayer grid (used by most manufacturers)
- An octagonal arrangement (Fuji)
- The “stacked” foveon sensor used by Sigma.

Each has advantages and disadvantages.

The Bayer Grid

In this arrangement the photodiodes are laid out in a rectangular grid with two green photodiodes for each red and blue photodiode. Any group of four photodiodes therefore consist of two green, one blue and one red. In this way each set of four photodiodes can be used to generate a full RGB pixel. A Bayer grid with 6 million photodiodes comprises 3 million green, 1.5 million red and 1.5 million blue photodiodes. The grid is usually made slightly larger than the final required picture size since a few photodiodes around the edge cannot be fully used.

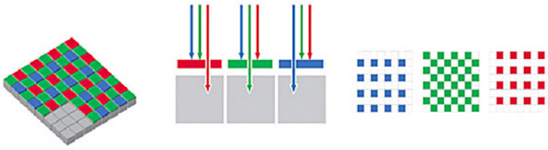

Figure 5: Bayer grid illustration

Fuji SuperCCD sensor array

An octagonal array of red, green and blue photodiodes is laid out. The pattern is similar to a Bayer grid turned through 45 degrees. Each photodiode is octagonal in shape and covers a larger area than a conventional square photodiode. This gives higher effective horizontal and vertical resolution and makes the sensor less sensitive to horizontal and vertical lines which appear cleaner to the eye. In the latest generation sensor each measurement site comprises a pair of photodiodes, a small one and a large one which together yield greater sensitivity and dynamic range.

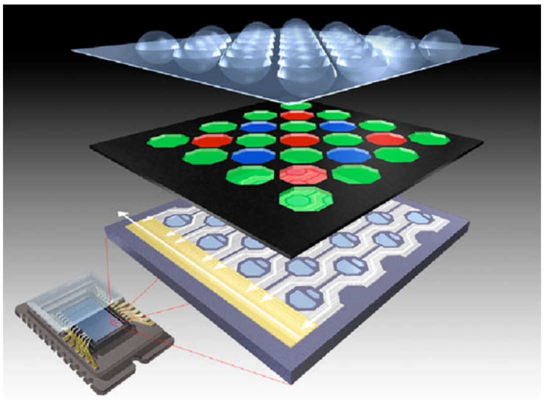

Figure 6: Fuji SuperCCD sensor construction (courtesy Fuji)

Sigma's Foveon sensor

The Foveon sensor stacks the red, green and blue photodiodes on top of one another, and therefore each measurement site on the sensor produces a full colour RGB pixel. Whilst theoretically ideal this poses a marketing challenge for Sigma. Some experts believe that the 3 Mpel Foveon sensor gives superior performance to conventional 6 Mpel Bayer grid sensors, but it does not quite match the performance of the 12 Mpel Bayer sensors. This is not really surprising since it has 9 million photodiodes, which places it between the two and yet it seems somewhat misleading to refer to it as a 9 Mpel sensor.

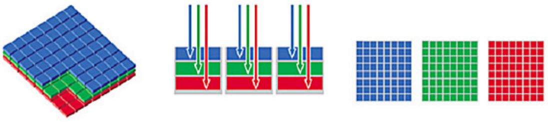

Figure 7: Sigma Foveon X3 sensor array

Sensor sizes and densities

In the days of celluloid a relatively small number of film sizes were adopted by the industry – for obvious reasons. In the digital age, however, every manufacturer can choose any sensor size they like since there are no compatibility issues to deal with.

As a result a large number of different digital sensor sizes have emerged. Some of the more common ones are as follows:

Sensor	length (mm)	width (mm)	area (mm^2)	Cameras
“645”	56.0	41.5	2324	
Full frame 35mm	36.0	24.0	864	Kodak DCS14n, Canon 1Ds, Contax N
Canon	28.7	17.8	511	Canon 1D
APS-C	25.1	16.7	419	
Nikon	23.7	15.6	370	Nikon D100 D70 D2X, Fuji S2 S3, Sigma, Pentax
Canon	22.7	15.1	343	EOS D30 D60 10D
Four-thirds	18.0	13.5	243	Olympus
2/3”	8.8	6.6	58	Sony DSC-F717, Minolta 7Hi
1/1.8”	7.2	5.3	38	Coolpix 4500, 995 and many others
1/2.7”	5.3	4.0	21	Minolta DiMAGE Xi and others

Table 1: Some of the more common sensor sizes

There are many other sensor sizes including bigger and smaller.

Matching the lens to the sensor

Based on the lens analysis presented earlier and assuming that one designs for the optimum aperture of a professional quality lens the maximum sensor resolution needed is about 125 lines per millimetre. Since each line consist of white and black and applying the principle of Nyquist, this implies a density of 250 pixels per mm, or 62.5 kpel per mm^2. There is little point in having more pixels than this since the image does not contain much more information for them to capture.

In practice the smaller the sensor and required image circle, the better the lens can be made and the wider the available operating aperture. Thus cameras using smaller sensors and lenses optimised for the smaller image circle may achieve optimum performance at a wider aperture and therefore may have better raw resolution. It is estimated that the smallest sensors might achieve 180 lines per millimetre based on an optimum aperture of *f4*.

This sets the reasonable upper bounds for different sensor sizes as follows:

Sensor	**true (Mpel)**	**bayer (Mpel)**
Full frame 35mm	50	150-200
“APS”	24	72-96
Four-thirds	16	48-64
1/1.8”	5	15-20
1/2.7”	3	9-12

Table 2: Upper bound for sensor size

The figures in the “true” column assume that there are three pixels (RGB) at each location. Splitting them and laying them out on a Bayer grid is a manufacturing choice and will not change the final achievable image resolution significantly.

The ultimate performance of the camera will only be achieved using the best quality lenses at their optimum settings.

Other Factors

As sensor densities increase, the area of each pixel reduces, sensitivity falls and noise increases.

Typical consumer cameras produce images in JPEG format in which each pixel is represented as RGB with 8 bit per colour depth.

Most professional cameras today produce images with resolutions of 12 bits per pixel (per colour). This gives additional headroom and better control of noise compared with consumer cameras. However, 12 bits per pixel is still below the density range available from professional slide films, and is often not enough to achieve the desired control of highlights and shadows. Consequently there is pressure to push this up towards 14 bits per pixel, and almost certainly in the future 16 bits per pixel.

Greater dynamic range requires lower noise but smaller photodiodes contribute more noise, so these two factors are working against one another. More progress will be needed before professional optics become the limiting factor.

To compound the problem larger images require more storage in the camera, and much faster processing in order to capture and store them. These factors need to be balanced against the professional requirements for minimal shutter lag and high frame rates for shooting sequences of images. A 200 Mpel sensor at 16 bits per pixel requires around 400 MB for each RAW image.

Depth of Field

Depth of field is the ability to throw the foreground or background out of focus. It is often used for artistic value, or to separate the image subject from the surroundings.

Whilst depth of field does not have a direct bearing on sensor size or number of pixels, the size of the sensor has a significant effect on the depth of field available.

Longer focal length lenses exhibit less depth of field for a given aperture. It can be shown that the relationship approximates a square law for distant objects.

A sensor half the size requires a lens of half the focal length to achieve the same field of view. Therefore as a rule of thumb the camera with the larger sensor and longer lens has half the depth of field of the smaller.

This is evident in small sensor consumer cameras which exhibit almost no control of depth of field and fixed focus can often suffice even at large lens apertures.

Thus a photographer who wishes to exploit depth of field of focus will require a camera with a larger sensor size. This is an artistic argument for using larger physical sensor sizes.

Conclusions

This article has shown that using current imaging methods based on the optics of refraction, there is a point beyond which increasing sensor size follows the law of ever diminishing returns.

Consumer cameras using smaller sensors are already approaching the limits of the optical resolution available. Allowing for the fact that they use matched lenses optimised for sensor size it seems that the maximum image size (un-interpolated) realistically obtainable from a 1/1.8” sensor is around 5 Mpel (true), which equates to 20 Mpel based on the traditional way of counting pixels in the camera. Images of this size are more than enough for most consumer needs and apart from the “mine is bigger than yours” factor, there seems little value is pushing sensor sizes of consumer cameras much beyond this.

Larger higher resolution images are best obtained by professional cameras using larger sensors and the highest quality professional lenses.

Current professional practice requires image sizes of 16 Mpel or more. In order to achieve these without image interpolation cameras using sensors no smaller than the “Four-thirds” system will be required.

What the article has not looked at is the scope for increasing imaging performance using methods other than better optics. Such methods may be based on image processing, or even completely new approaches to optics. We can only look forward to the creative genius of scientists as they try to find ways to push forward the boundaries of science.

References

1. “Fundamentals of Optics”, Jenkins, White, published by McGraw-Hill Kogakusha Ltd.
2. “CMOS Image Sensors”, Olympus, http://www.mic-d.com/curriculum/imageprocessing/cmoschips.html
3. “Foveon X3 Pixel Page”, http://www.x3f.info/technotes/x3pixel/pixelpage.html
4. “The Luminous landscape – tutorials”, Michael Reichmann, http://luminous-landscape.com/tutorials/
5. “Intro to resolution and MTF curves”, Norman Coren http://www.normankoren.com/Tutorials/MTF.html
6. “Lens Tutorial”, David Jacobson, http://www.photo.net/photo/optics/lensTutorial

About the Author

David is a Chartered Engineer (Electronics) with many years experience in the Electronics, Communications and IT industries. His specialist fields are Digital Signal Processing and Software Engineering, although he enjoys using his writing and photography skills in his work.

He combines his technology background with a keen knowledge and love of the natural world and the environment.

David runs his own business, Pyxidium Ltd., providing interim management skills to small organisations. He specialises in the fields of Mobile Wireless and Digital Imaging. As a freelance photographer he undertakes assignments and shoots for stock using a modern digital SLR.