A SINGLE HEISENBERG-GABOR BASED FIGURE-OF-MERIT BASED ON THE MODULATION TRANSFER FUNCTION OF DIGITAL IMAGING SYSTEMS

Corey Manders and Steve Mann

University of Toronto
Dept. of Electrical and Computer Engineering
10 King's College Rd.
Toronto, Canada

ABSTRACT

We propose a single figure-of-merit measure of resolution of a digital imaging system based on the work of Gabor in communication theory. Gabor's work was largely inspired by Heisenberg's developments in quantum theory, most notably his uncertainty theorem of quantum mechanics. Gabor's results look simultaneously at the frequency and spatial domain of a signal, making it ideal for the measure of the modulation transfer function and point-spread function of an imaging system. As opposed to the crude "megapixel" measure which is bantered about in the marketplace, we suggest a figure-of-merit which more accurately represents the resolution of the system. Given that the resolution measure we propose is condensed into a single number rather than a function such as the modulation transfer function or the point spread function, it is our intent to propose this scientific evaluation as a means for typical consumers to fairly judge the resolution of a camera. Finally, we use this measure to compare common digital SLR cameras with varying lenses.

1. INTRODUCTION: DESCRIBING A CAMERA'S RESOLUTION

An unfortunate development for typical consumers buying digital cameras is the evaluation of a particular camera by its measure of megapixels. Often consumers will prefer one camera over another simply because the pixel count is higher. Most engineers and scientists realize this measure is not entirely appropriate as a descriptor. Among the host of reasons why this is true: megapixel counts do not describe tonal response, they do not take into account the complete imaging system (for example the point spread function of the lens system), and often interpolated pixels are included in the count. Alternatively, engineers, scientists, and learned users of cameras prefer the data offered by the modulation transfer function [1]. The calculation of the modulation transfer function will be concisely explained in the following section. Those familiar with the modulation transfer function will undoubtedly understand that the function is a far superior measure than megapixels. However, the downside of a graph of the function is that it is not intuitive. An observer of a plot of the modulation transfer function of a camera must approximate at which frequency the function begins to decline, how long the function takes to decline, and the overall rate of decline of the function. General consumers of cameras are bolstered by the simplicity of megapixel counts in a society where bigger means better. Furthermore, modulation transfer functions are often plotted logarithmically, adding to the complexity of reading the data.

It should be noted that in the previous era of film based cameras, many professional grade films came with plots of the tonal response of the film in the form of a set of response curves. Though these plots were obviously only describing the response of the film and not the imaging system, these plots were most often thrown away. The common reason for throwing the information away could easily be speculated; the information is presented in a format too complicated for many users, and secondly, even with the information, the user was not able to accurately use the information in any meaningful way (with a few exceptions). We propose a system of measuring a cameras response by means of a single figure-of-merit of the modulation transfer function. This system results in a single number which is appropriate as a measure of both the tonal and spatial resolution of an imaging system.

After describing in detail our method of finding the figure-of-merit of a given imaging system, we measure four imaging systems. Two Nikon digital SLR cameras, the D70 and the D2H, with two lenses, an AF-S Nikkor 18-70mm 1:3.5-4.5G ED DX lens and an AF Nikkor 70-300mm 1:4-5.6 D ED lens. Each system was tested at the most open fstop (3.5 for the 18-70mm lens 4 for the 70-300mm lens). Similarly, the focal length of the two lenses were set to their shortest setting, 18mm and 70mm respectively. It should also be noted that the two cameras use different imaging technology. The lower priced Nikon D70 camera uses a 6.1 megapixel CCD sensor whereas the professional grade D2H uses a 4.1 megapixel "LBCast" JFET sensor developed by Nikon.

2. THE MODULATION TRANSFER FUNCTION

The sharpness of a photographic imaging system or of a component of the system is characterized by a parameter called modulation transfer function (MTF). This function is also termed the spatial frequency response and is relatively easy to calculate using a test chart. Several test charts are available, such as the USAF 1951 test chart. However, Norman Koren has developed a test chart which is much easier to use and is free to download from his website (www.normankoren.com). The basic pattern of the chart is shown in figure 1. It is also common to see this test pattern as

Fig. 1: The basic MTF test pattern, a sine wave of increasing spatial frequency.

black and white line pairs as opposed to the sine pattern shown in figure 1 giving rise to the scale lp/mm (line pairs per millimeter). However, given that the pattern used in this paper is the sine pattern, the measure cy/mm (cycles per millimeter) is more appropriate. One may easily recognize this pattern as a visual chirp, the sensor response of which is the basis of our analysis.

Most often, the test chart is positioned at a distance from the lens such that the lp/mm or cy/mm correspond to a millimeter of the sensor array. This implies that one knows the size of the sensor array. Even though this is relatively easy to find for most commercially available cameras, using this measure in this regard is not appropriate for the task for which we intend. A user of a digital camera or imaging system usually just wants to know the resolution of the imaging system, regardless of how large the physical sensor is. For this reason, we disregard the cycles per millimeter and rather consider cycles per image height and cycles per image width.

2.1. Calculating the MTF

Intuitively, the Modulation Transfer Function (MTF) may be described in the following manner: once the test chart is properly positioned and a test image is taken, we analyse the resulting images. At frequencies where the MTF of an imaging system is 100%, the pattern is unattenuated and retains full contrast. At the frequency where MTF is 50%, the contrast is half its original value. It should also be noted that all calculations are done on RAW image files. The non-linear range compression (described in detail in [2][3] and [4]), will modify the shape of the resulting curves. Using a linear response output, as is demonstrated in RAW image files, is essential in computing MTFs and resulting resolution parameters with consistency between cameras. It is well documented that the range compression applied to raw data internally in imaging system varies from camera manufacturer to camera manufacturer and indeed camera model to camera model. In the case where raw data is is not available from a given camera, the range compression must first be expanded to accurately measure the response. Such a technique is demonstrated in [3], where the resulting photoquantimetric values are light-linear, appropriate for the task. For added accuracy, the data in this paper was calculated from the uninterpolated Bayer pattern data.

Let V_b be the raw photoquantimetric value (pixel value from the raw data) which is the minimum value observed in the test image, likely observed in the low frequency region. Let V_w be the raw photoquantimetric value which is the maximum value observed in the test image, also likely to be observed in the low frequency region. Let V_{min} be the minimum luminance for a pattern near spatial frequency f. Correspondingly, let V_{max} be the maximum luminance for a pattern near spatial frequency f. The necessary definitions may then be stated as:

$$C(0) = \frac{V_w - V_b}{V_w + V_b} \quad \text{the low frequency (base) contrast}$$

$$C(f) = \frac{V_{max} - V_{min}}{V_{max} + V_{min}} \quad \text{is the contrast at spatial frequency} f$$

$$MTF(f) = \frac{C(f)}{C(0)} \tag{1}$$

$C(0)$ is the highest contrast found in the image of the test chart. The modulation transfer function at a frequency $f(MTF(f))$ is normalized by this contrast resulting in a measure in which the highest value is 1.0 (perfect modulation transfer) and 0.0 (no modulation transfer). At any horizontal co-ordinate, the test chart and resulting image represents a specific frequency f in cy/mm. For vertical resolution testing, the test chart is rotated 90 degrees, and consequently vertical co-ordinates represent a specific frequency.

2.2. Positioning the testchart

The test charts produced by Norman Koren are scaled such that the scales indicated on the test chart correspond to the test chart being 2.5mm or 5mm in length. Rather than millimeters, we consider cycles per image width (cy/iw). Then, if two contiguous lengths of the test chart are imaged, 10mm of the test image have been imaged in reference to the test chart scale. Thus, if the scale reads 2cy/mm this becomes 20cy/iw, 10cy/mm becomes 100cy/iw, etc. The fact that the scale is now in image widths allows the measure of resolution to be universal across sensors. To deal with different resolutions in image height, the test image is also imaged in a vertical position. A similar position is established for the camera to test chart distance, yielding a test pattern which may be examined in cy/ih (cycles per image height). When we conducted these tests on the cameras, we positioned the camera such that a single test chart was imaged so that the low frequency end started at the edge of the image and ended in the center of the image. This was done in both the vertical and horizontal cases. We chose to image the test chart such that the high frequency signals were largely in the centre of the frame to guard against any distortions (such as barrel distortion from the lens) which tend to affect the outer edges of the image more than the centre. Note that this will also result in higher resolution test results than if the image was taken with higher frequencies at the edge of the camera's field of view.

2.3. Collecting accurate data

With most cameras, a Bayer pattern is used in collecting red, green, and blue sensor values. Most commonly, these are alternating lines of red, green, red, green, etc. and green, blue, green, blue, etc. These values become interpolated in various manners to produce red, green, and blue pixel values at every pixel location, even though only one pixel colour exists at a specific location. Using programs such as David Coffin's free source program **dcraw** (available at http://www.cybercom.net/∼dcoffin/dcraw/), or else **neftoppm** (available at http://www.eyetap.org/∼corey/code.htm), the raw linear 12-bit uninterpolated Bayer pattern data may be collected. In the case of dcraw, the code must be modified so that no Bayer interpolation takes place, this will leave the raw sensor values in each of their original locations. The resulting 12-bit portable pixmap image (ppm) was then separated into three colour images (red, green, and blue) with NaN in the locations where there is no data.

Once the single ppm image was split into three colour images, the relevant data was cropped out of each colour channel's ppm. The resulting data was averaged down columns in the case of the horizontal test pattern, disregarding NaN values, and across rows in the case of the vertical test pattern, also disregarding NaN values. One should note that in the case of the green channel, the resolution will be higher than that of the blue and red channels due to the number of sensors being equal to the total number of red and blue sensors. From this data, the modulation transfer functions in the vertical and horizontal directions were computed.

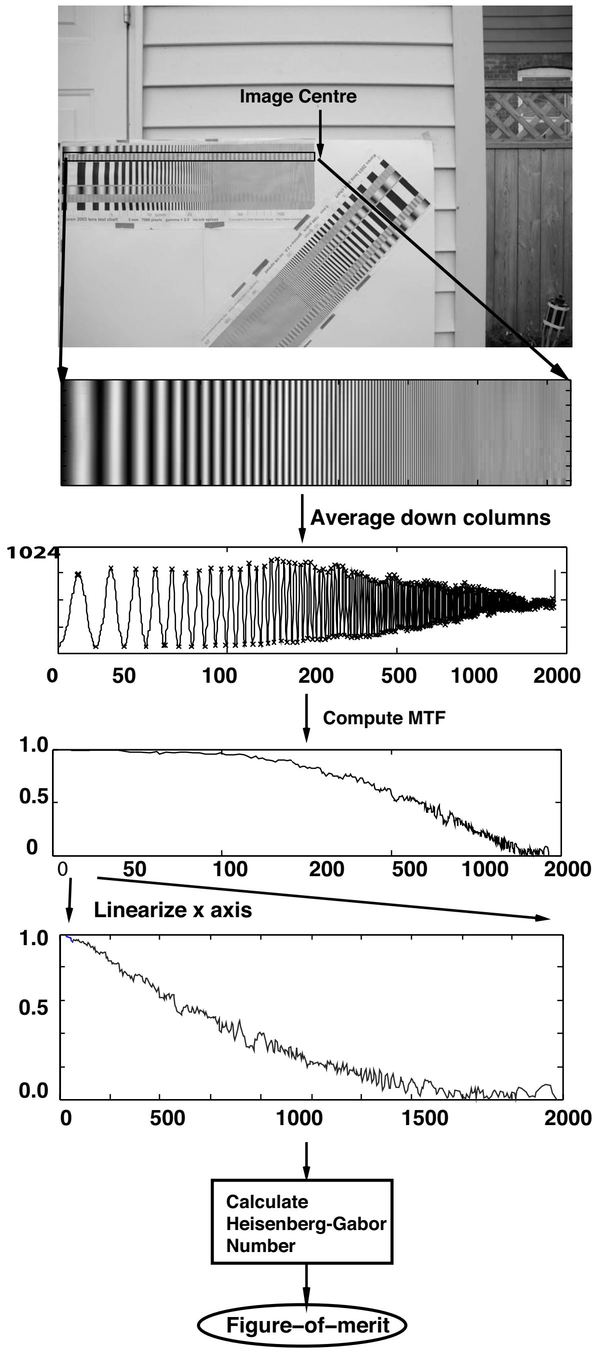

Fig. 2: A visual representation of the process used to find the Heisenberg-Gabor figure-of-merit. The initial image of the test pattern is taken such that one pattern begins at one edge of the image and ends in the centre. The Bayer pattern data is then extracted from the raw image file data and appropriately averaged excluding NaNs. From this data the MTF is calculated followed by linearizing the x-axis and applying equation 2. This yields a figure-of-merit for a single colour channel, which may then be linearly combined with other colour channels, producing a single figure-of-merit for a given imaging system.

3. PRODUCING A HEISENBERG-GABOR FIGURE-OF-MERIT

Using Heisenberg's uncertainty relation[5], Gabor proposes the concept of "effective frequency width" Δf and "the effective duration" Δt of a signal in his 1946 paper[6]. To measure the modulation transfer function (and possibly the corresponding point-spread function of the camera), we propose to use Gabor's Δf measure to quantify the resolution of a given camera. As the modulation transfer function may be viewed as a spatial frequency signal, we consider its effective frequency.

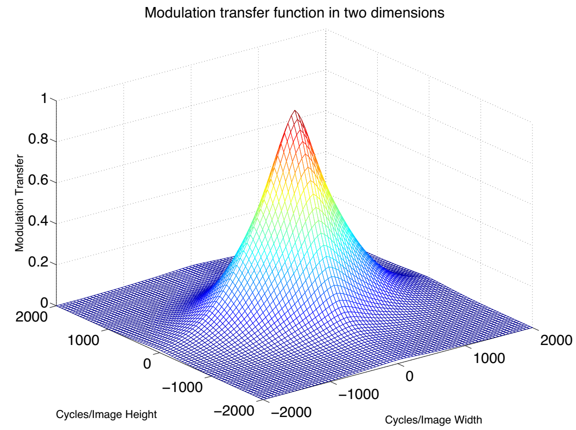

Fig. 3: The modulation transfer function plotted for both the vertical and horizontal test patterns. The Fourier transform of this measure is the point-spread function of the camera. As expected, the 2-D modulation transfer function is close to being Gaussian, and correspondingly, the point-spread function approaches being Gaussian.

3.1. Analytic Background

To find the values of Δf, the simplest method uses the first and second moments of the signal. Specifically we have,

$$\Delta f = \left[2\pi\overline{(f - \bar{f})^2}\right]^{\frac{1}{2}}. \quad (2)$$

Note that for ease of calculation, the use of the statistical identity $\overline{(f - \bar{f})^2} = \bar{f^2} + (\bar{f})^2$ is utilized. Given any signal $s(f)$ and its corresponding quadrature signal $\sigma(f)$ as in [6] we define a weight function

$$\psi^*\psi = \left[s(f)^2\right] + \left[\sigma(f)^2\right] \quad (3)$$

where the asterisk denotes the complex conjugate of the resulting analytic signal. The weight function is therefore the square of the absolute value of the signal. This can be considered the "power" of the signal and will be referred to by this name in what follows. Following the logic of Gabor, we do not consider the moments themselves, but rather the moments divided by M_0. For example, in our case we have:

$$\bar{f} = \frac{\int \psi^* f \psi df}{\int \psi^* \psi df} \qquad \bar{f^2} = \frac{\int \psi^* f^2 \psi dt}{\int \psi^* \psi df}. \quad (4)$$

Finally, we note the fact that the spatial frequency signal (the modulation transfer function), and the point spread function are related by a Fourier transform. This is what gives rise to the factor of 2π in the definition of Δt and Δf. Also, the point spread function may be found simply by taking the discrete Fourier transform of a symmetric version of the modulation transfer function. The symmetric modulation transfer function is produced by assuming the response of the imaging system will be identical for negative frequencies as positive frequencies, therefore enabling us to mirror the MTF around the y-axis.

3.2. The Δf measure of resolution

Using the information presented, one recognizes that in the spatial frequency domain, an imaging system should maximize its frequency response. Thus, we wish Δf to be as large as possible. On the contrary, the ideal point spread function is a delta function, thus we wish to minimize the measure of Δt. However, increasing Δf will under most circumstances decrease Δt. For this reason, we propose Δf as a reasonable measure of resolution in one axis. We must also remember that the measure may be taken in multiple directions and locations on the imaging system. We chose to measure the orthogonal vertical and horizontal directions, given the typical pixel layouts on imaging sensors. The horizontal measure, we choose to label Δf^{hor} and the vertical measure Δf^{vert}. Because we wish to maximize both the vertical and horizontal components of the Δf measure, a final measure of sensor resolution in one colour channel is proposed which is simply

$$\Delta f^{VH} = \Delta f^{\text{vert}} \times \Delta f^{\text{hor}} \tag{5}$$

3.3. Simultaneously evaluating colour channels

The previous work in this paper shows how to to derive a single figure-of-merit (Δf^{VH}) which may be applied to each of the colour channels taken from the uninterpolated Bayer pattern. One possibility is to only test and report the result for the green channel. This certainly makes sense from the perspective that the sensor array is populated more densely with green sensors, and the eye is most sensitive to light in the green range. Unfortunately, if the camera suffered from distortions in the red and blue channels, such a measure would be blind to this problem. In most digital cameras and imaging systems, the green channel will have a higher resolution which coincides with human perception. For this reason, we are suggesting that a valid measure of the three channels is to perform a YCbCr transformation on the values of the three channels, and take the Y component as a measure of the final sensor resolution. We term this measure Δf^{Y}, which may be calculated as:

$$\Delta f^{Y} = 0.299\Delta f^{VH}(\text{Red}) + 0.586\Delta f^{VH}(\text{Green}) + 0.114\Delta f^{VH}(\text{Blue}) \tag{6}$$

The base results for each colour channel using various camera bodies and lens combinations are given in table 1, and the resulting single number using equation 6 is shown in figure 2.

4. CONCLUSION

Given our method of combining colour channel resolution measures presented previously, the following final figures-of-merit are shown in table 2. The results are largely as one may predict. The Nikon D70 having a higher sample rate (it is a 6.1 megapixel camera as opposed to the 4 megapixel D2h), does indeed outperform the D2h in terms of the Heisenberg-Gabor rating. Also, changing the lens from the Nikkor 18-70mm to the Nikkor 70-300mm also reduces the performance. We expect this as the increased focal length should increase the width of the associated point-spread function, effectively convolving a larger blur kernel with the captured scene.

There are a few steps that may be taken to improve the accuracy of the measure. Specifically, aliasing artifacts are present in the data, creating unwanted results in the MTF calculation. This is particularly a problem when calculating the second moments of the MTF where small artifacts in high frequencies result in large perturbations from results which should be around 0. This could be overcome by averaging results where each image was taken with small pans (for the horizontal tests) or small tilts (for the vertical tests), and is the topic of ongoing research.

Imaging System	Horizontal Resolution	Vertical Resolution	Total Resolution
Green Channel			
D70,18-70mm lens	1678.55	1491.82	2.504×10^6
D70, 70-300mm lens	1883.33	1313.77	2.474×10^6
D2h, 18-70mm lens	1368.39	1316.65	1.802×10^6
D2h, 70-300mm lens	1432.25	1183.37	1.695×10^6
Blue Channel			
D70, 18-70mm lens	1628.89	1308.38	2.130×10^6
D70, 70-300mm lens	1336.92	1513.33	2.023×10^6
D2h, 18-70mm lens	1531.6	1776.53	2.721×10^6
D2h, 70-300mm lens	1386.02	1371.71	1.901×10^6
Red Channel			
D70, 18-70mm lens	1532.44	1307.57	2.004×10^6
D70, 70-300mm lens	1451.63	1473.52	2.138×10^6
D2h, 18-70mm lens	1329.64	1496.46	1.990×10^6
D2h, 70-300mm lens	1530.34	1312.43	2.009×10^6

Table 1: Heisenberg-Gabor (Δf) results for various combinations of Nikon D70 and Nikon D2h camera bodies with Nikkor 18-70mm and 70-300mm lenses.

Imaging System	**Figure-of-merit (Δf^{Y})**
Nikon D70, 18-70mm lens	2.309×10^6
Nikon D70, 70-300mm lens	2.280×10^6
Nikon D2h, 18-70mm lens	1.961×10^6
Nikon D2h, 70-300mm lens	1.811×10^6

Table 2: Resulting figures of merit for Nikon D70 and D2h camera bodies with Nikkor 18-70mm and 70-300mm lenses.

5. REFERENCES

[1] David Corley and Shirley Li, "Controlling image quality in a digital world," *Journal of the Society of Motion Picture and Television Engineers*, pp. 293–306, September 2004.

[2] F. M. Candocia, "A least squares approach for the joint domain and range registration of images," *IEEE ICASSP*, vol. IV, pp. 3237–3240, May 13-17 2002, avail. at http://iul.eng.fiu.edu/candocia/Publications/Publications.htm.

[3] C. Manders, C. Aimone, and S. Mann, "Camera response recovery from different illuminations of identical subject matter," in *Proceedings of the IEEE International Conference on Image Processing*, Singapore, Oct. 24-27 2004, pp. 2965–2968.

[4] Steve Mann, *Intelligent Image Processing*, John Wiley and Sons, November 2 2001, ISBN: 0-471-40637-6.

[5] W. Heisenberg, "Über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik," *Zeitschrift für Physik*, vol. 43, pp. 172–198, 1927, English translation: J. A. Wheeler and H. Zurek, Quantum Theory and Measurement Princeton Univ. Press, 1983, pp. 62-84.

[6] D. Gabor, "Theory of communication," *J. Inst. Elec. Eng.*, vol. Vol.93, no. 3, pp. 429–457, 1946.