Effect of Gamma Correction in Discrete Imaging Systems upon Determination of MTF

R. Jenkin, R. E. Jacobson, N. R. Axford, and E. Allen
University of Westminster, Imaging Technology Research Group
Middlesex, United Kingdom

Abstract

Tonal manipulation is commonly implemented in discrete imaging systems in order to improve the subjective appeal of results or to emulate some other device. The use of integer mathematics in its application, however, causes loss of detected intensity levels. The reduction of information content that this represents introduces noise into MTF measurements of discrete systems.

A method to estimate a range in which the actual modulation exists from gamma corrected pixel values is presented. The models are compared with experimental data, the results discussed and conclusions drawn.

Gamma Correction

When the behaviour of an intended display system is known, the tone reproduction of the discrete acquisition device may be modified in order to achieve optimum objective tone reproduction for the combination.[1] Known as gamma correction, the modification takes the form:

$$V' = V^{\gamma_C} \tag{3}$$

where $\gamma_C = 1/\gamma_D$. V is the original signal, V' the modified signal, γ_C, the value of correction and γ_D the display gamma. The effect of gamma correction upon the transfer function of a system is to yield approximately linear tone reproduction as desired for optimum objective output.[2-5] Subjectively preferred tone reproduction often differs to the objective ideal and depends upon the display medium, viewing conditions and scene content.[2-5] For example, it is suggested that $\gamma = 1.6$ is optimum for viewing transparencies and $\gamma = 1.2$ for photographic prints under normal conditions.[5] In addition, the tone reproduction of the acquisition system will often not be linear and may be described using its own determined value of γ.

The tonal correction applied may be adjusted to take account of the above factors. The correction necessary may be calculated using:

$$\gamma_C = \frac{\gamma_T}{\gamma_A \gamma_D} \tag{4}$$

where γ_T is the target gamma and γ_A the value of the acquisition system.[6] For a system comprising of a digital camera, computer and CRT, the correction may be applied at any point before the physical display of the information.

It should be noted that a number of commercially produced electronic stills cameras employ *tonal shaping* rather than gamma correction in order to simulate the transfer characteristics of other devices, e.g. photographic emulsion. Often this approach is implemented as a *look-up table* (LUT). In such a case the gamma model has little or no basis to describe such tone reproduction and generally yields a poor mathematical fit. A more appropriate description in such case is the LUT itself. For the purposes of simplification the use of a gamma model assumed throughout this work.

Discrete Gamma Corrected Signals

Most discrete imaging devices produced presently include gamma correction or tonal shaping to condition signals for use with various output devices. Whilst this poses few difficulties for analogue systems, its application in digital systems produces a loss of intensity levels due to the limitation of integer calculations.[7,8] Output levels may not be used or alternatively have multiple mappings to input levels.[8,9] Evidence of this occurs as *stepping* in the systems tone reproduction curve, Figure 1.[6]

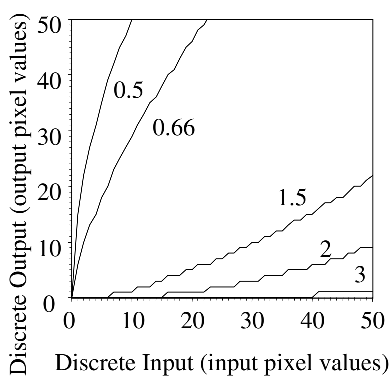

Figure 1. Stepping introduced to the gamma corrected signal due to the use of integer mathematics for the values of gamma correction shown.

A common method to mitigate the loss of pixel values is to acquire images at an increased bitdepth and sub-sample to the required output bitdepth after performing gamma correction.[7] Triantaphillidou, Jacobson and Ford have produced a computer routine to count remaining pixel values in gamma corrected discrete signals. The counting routine was compared with measurements of a commercially available 35mm film scanner and found to be accurate.[7]

Mathematically, concatenated gamma correction and sub-sampling of a digital signal may be described by:

$$G_Q = \left\lfloor \left(\frac{Q}{2^{d_Q}-1}\right)^{\gamma_C} \times \left(2^{d_G}-1\right) \right\rfloor \tag{5}$$

where d_Q is the bitdepth of the quantized input, d_G the bitdepth of the gamma corrected output, Q the input pixel value and G_Q the output pixel value associated with that input. The parentheses, $\lfloor \;\rfloor$, represent notation for a *floor* function. It is desired to examine the effect of gamma correction in isolation. The input signal, Q, therefore, is quantized as it has passed though the ADC before entering the gamma correction process. Thus, values of Q are restricted to positive integers between 0 and $2^{d_Q}-1$. Also d_Q and d_G are restricted to positive integers greater than or equal to one in order to describe physically realisable systems.

Linear Input Units

There exist many applications in the field of image science that require a linearized signal for mathematical or other manipulation. Any imaging system that does not exhibit a linearly proportional relationship of input with output cannot be considered linear. In these instances, the accepted practice is to linearize the output signal by applying the inverse of the systems tone reproduction curve to convert the signal into linear input units. A number of references document the procedure which is well known. An overview is given by Dainty.[10]

The approach works well for analogue systems, however, the described pixel value loss in discrete systems causes noise. Pixel values in the output are mapped to a number of levels in the input. It is impossible to correctly identify which input level gave rise to the output, Figure 2, and therefore only an *linear input range* may be calculated. Any measured output extrema which cannot be associated with a unique input value will cause ambiguity.

Mathematical Development

Calculation of Linear Input Range

The ability to calculate the extent of linear input range for a given output pixel value enables the calculation of limits within which the actual MTF will exist.

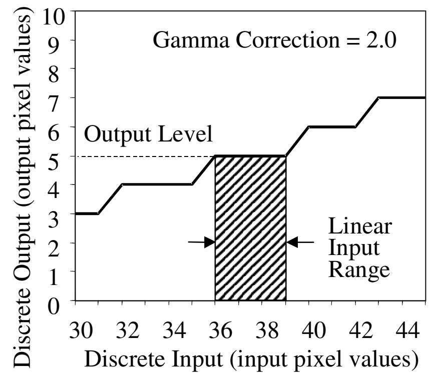

Figure 2. A magnified portion of a gamma corrected signal showing the ambiguity caused when linearizing a digital signal. The output value of 5 is mapped to input pixel values of 36 to 39.

For a given output pixel value, G_Q, the linear input range may be described as extending from Q_{MIN} to Q_{MAX}, Figure 2. When the input range is described in the above manner, the following will be true for a real system:

$$G_Q = \left\lfloor \left(\frac{Q_{MAX}}{2^{d_Q}-1}\right)^{\gamma_C} \times \left(2^{d_G}-1\right) \right\rfloor \tag{6}$$

As the minimum significant unit of the floor operation is one, the following is true:

$$\left[\left(\frac{Q_{MAX}}{2^{d_Q}-1}\right)^{\gamma_C} \times \left(2^{d_G}-1\right) \right] - G_Q < 1 \tag{7}$$

for $G_Q < 2^{d_Q}-1$ and therefore:

$$\left(\frac{Q_{MAX}}{2^{d_Q}-1}\right)^{\gamma_C} \times \left(2^{d_G}-1\right) < 1 + G_Q \tag{8}$$

As d_Q and d_G are limited to positive integers greater than or equal to one, it is true that $2^{d_G}-1 \geq 1$ and $2^{d_Q}-1 \geq 1$. Dividing Equation 8 by $2^{d_G}-1$ and raising to the power of $1/\gamma_C$ it is found:

$$\left(\frac{Q_{MAX}}{2^{d_Q}-1}\right) < \left(\frac{1+G_Q}{2^{d_G}-1}\right)^{\frac{1}{\gamma_C}} \tag{9}$$

Finally multiplying by $2^{d_Q}-1$:

$$Q_{MAX} < \left(\frac{1+G_Q}{2^{d_G}-1}\right)^{\frac{1}{\gamma_C}} \times \left(2^{d_Q}-1\right) \tag{10}$$

Q_{MAX} is the maximum value of the input range that will give the output value G_Q. The maximum value that Q_{MAX} may take is the RHS of the above equation. The above equation yields a *continuous* value of Q_{MAX} for a given value of G_Q. It should be remembered that gamma correction in the majority of digital acquisition systems operate on quantized signals, thus Q should be restricted to positive integer values. It is necessary, therefore to reduce Q_{MAX} to the nearest positive integer *below* the value calculated and thus it becomes:

$$Q_{MAX} = \left\lceil \left(\frac{1+G_Q}{2^{d_G}-1} \right)^{\frac{1}{\gamma_C}} \times \left(2^{d_Q}-1\right) \right\rceil - 1 \tag{11}$$

for $G_Q \le 2^{d_G}-1$. When $G_Q = 2^{d_G}-1$, $Q_{MAX} = 2^{d_Q}-1$. The parentheses, $\lceil\ \rceil$, represent a *ceiling* function and, as previous, $\lfloor\ \rfloor$ represent the use of a *floor* function in the following equations.

The lower value of the input range may be denoted Q_{MIN}. It may be calculated by first considering the mapping of Q to G_Q:

$$G_Q = \left\lfloor \left(\frac{Q_{MIN}}{2^{d_Q}-1} \right)^{\gamma_C} \times \left(2^{d_G}-1\right) \right\rfloor \tag{12}$$

The lowest input value for a given output will occur when:

$$\left\lfloor \left(\frac{Q_{MIN}}{2^{d_Q}-1} \right)^{\gamma_C} \times \left(2^{d_G}-1\right) \right\rfloor - G_Q = 0 \tag{13}$$

Therefore solving:

$$G_Q = \left(\frac{Q_{MIN}}{2^{d_Q}-1} \right)^{\gamma_C} \times \left(2^{d_G}-1\right) \tag{14}$$

Q_{MIN} may be found to be:

$$Q_{MIN} = \left(\frac{G_Q}{2^{d_G}-1} \right)^{\frac{1}{\gamma_C}} \times \left(2^{d_Q}-1\right) \tag{15}$$

The solution also needs to be restricted to positive integers. The calculated value above is less than or equal to the minimum value of the input range and thus the calculation of Q_{MIN} becomes:

$$Q_{MIN} = \left\lceil \left(\frac{G_Q}{2^{d_G}-1} \right)^{\frac{1}{\gamma_C}} \times \left(2^{d_Q}-1\right) \right\rceil \tag{16}$$

Q_{MAX} and Q_{MIN} may only be calculated for values of G_Q that exist in a gamma corrected output. Attempting to calculate the input range of an output value that does not exist will result in a breakdown of the formulae.

The numerical range of input values, α, associated with a given output value, G_Q, may be calculated:

$$\alpha = Q_{MAX} - Q_{MIN} \tag{17}$$

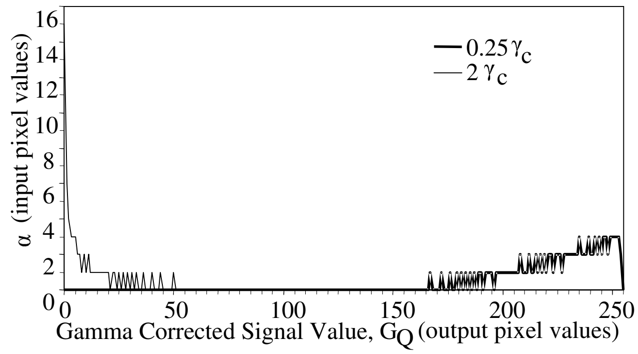

Figure 3. A graph of the numerical range of input values, α, versus gamma corrected signal values.

The *number* of input values, η, that will give rise to an output value is calculated using:

$$\eta = Q_{MAX} - Q_{MIN} + 1 \tag{18}$$

Plotting a graph of α against G_Q gives an indication of where the majority of uncertainty lies in the gamma corrected signal, Figure 3.

The figure illustrates that if the value of gamma correction used is above one, the majority of the noise lies in lower pixel values. Conversely, if the value of gamma correction is below one noise lies in the upper pixel values though to a lesser degree. Gamma correction of one effectively represents no modification of the signal and thus no noise is introduced.

Effect on MTF Determination

As the transfer function of the system is commonly used to convert the measured signal into linear input units, uncertainty in estimated input pixel values must be considered as an uncertainty in the calculated modulation transfer function. Establishing the variation in input values of maxima and minima will enable the calculation of a range in which the actual modulation will exist. Thus, the effect of gamma correction on MTF may be evaluated. For gamma corrected values of maxima and minima, M_G and N_G, of a sinusoidal signal, it is possible to calculate the range of input values that could have given rise to these using the equations above:

$$M_Q^{MAX} = \left\lceil \left(\frac{1+M_G}{2^{d_G}-1} \right)^{\frac{1}{\gamma_C}} \times \left(2^{d_Q}-1\right) \right\rceil - 1 \tag{19}$$

$$M_Q^{MIN} = \left\lceil \left(\frac{M_G}{2^{d_G}-1} \right)^{\frac{1}{\gamma_C}} \times \left(2^{d_Q}-1\right) \right\rceil \tag{20}$$

$$N_Q^{MAX} = \left\lceil \left(\frac{1+N_G}{2^{d_G}-1} \right)^{\frac{1}{\gamma_C}} \times \left(2^{d_Q}-1\right) \right\rceil - 1 \qquad (21)$$

$$N_Q^{MIN} = \left\lceil \left(\frac{N_G}{2^{d_G}-1} \right)^{\frac{1}{\gamma_C}} \times \left(2^{d_Q}-1\right) \right\rceil \qquad (22)$$

where M_Q^{MAX} and M_Q^{MIN} are the maximum and minimum input values resulting in M_G. The maximum and minimum input values resulting in N_G are N_Q^{MAX} and N_Q^{MIN}.

The range in which the actual modulation is contained may be calculated by evaluating maximum and minimum modulation possible, MTF_Q^{MAX} and MTF_Q^{MIN}:[11]

$$MTF_Q^{MAX} = \frac{M_Q^{MAX} - N_Q^{MIN}}{M_Q^{MAX} + N_Q^{MIN}} \qquad (23)$$

$$MTF_Q^{MIN} = \frac{M_Q^{MIN} - N_Q^{MAX}}{M_Q^{MIN} + N_Q^{MAX}} \qquad (24)$$

As Q_{MAX} and Q_{MIN} represent the extremes of input range that may result in the gamma corrected output, G_Q, MTF_Q^{MAX} and MTF_Q^{MIN} represent the extremes of input modulation the may result in the measured gamma corrected signal modulation. Due to the use of the ceiling function in Equations 19 and 22, these expressions cannot be simplified further. Given values of M_G, N_G, γ_C, d_Q and d_G, it is possible to calculate range of the input modulation within which the actual modulation will fall.

Experimental Simulation

As gamma correction for discrete systems is a mathematical process, its implementation in the modality is noiseless. Because of this, experimental testing of the technique may be performed using a simulation.

Arbitrary values were chosen to represent maxima and minima in the original signal, Q, at bitdepths of 12, 10 and 8 bits.

Equation 5 was used to apply various values of gamma correction and sub-sampling to create an output signal. The bitdepths chosen represent typical values for digital technology at present [12, page 33].

Using the gamma corrected maxima and minima, the range containing the input modulation was calculated with Equations 23 and 24. The actual modulation is then compared with the calculated range.

Results of the simulation are presented in Figures 4 to 9. All graphs are plotted with respect to quantized input signal amplitude as this is a primary variable which determines the severity of the effect of the gamma correction.

Figures 4 and 5 represent the effect typical of gamma correction performed for output to a CRT. Assuming linear tone reproduction of the acquisition device, $\gamma_A=1$, and a target gamma, γ_T, of unity, the gamma correction required, γ_C, is 0.45.[11] The input and output bitdepth is 8 bits.

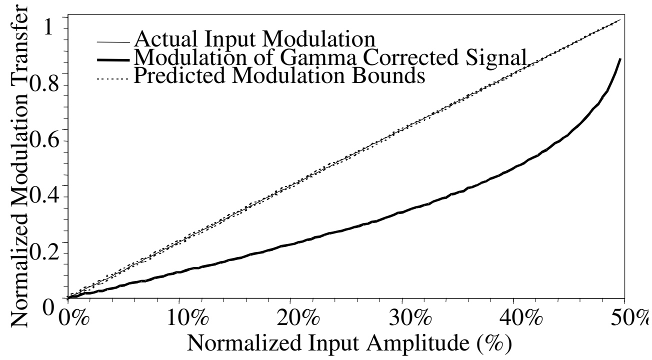

Figure 4. Predicted modulation range for a signal with gamma correction of 0.45 applied. The input and output bitdepths are both 8 bits.

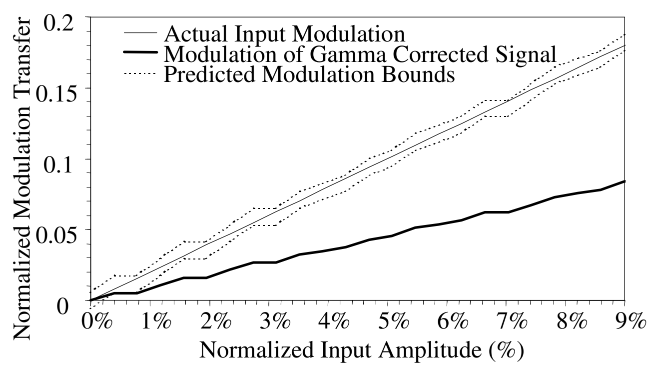

Figure 5. An enlarged portion of Figure 7.

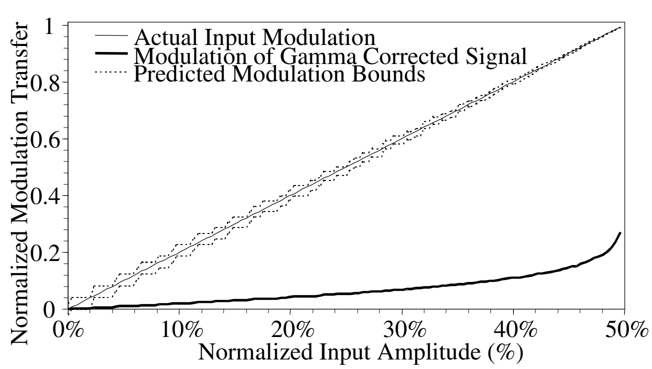

Figure 6.The predicted modulation range for a signal with gamma correction of 0.1 applied. The input and output bitdepth is 8 bits.

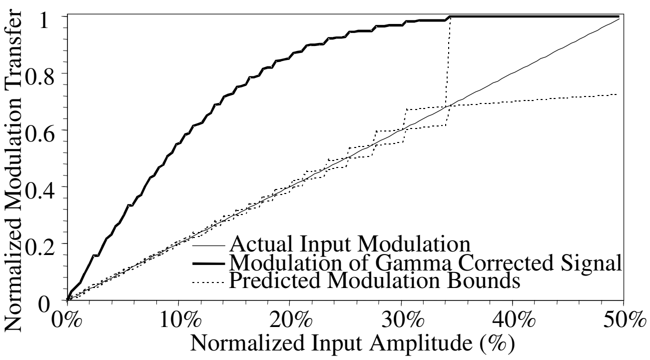

Figure 7. As for Figure 9 with gamma correction of 3 applied.

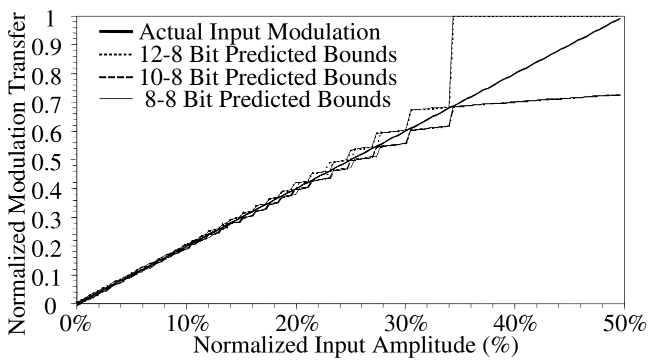

Figure 8. Predicted modulation range for a signal gamma corrected with a value of γ_c=3. The curves show the effect of changing input bitdepth. Both target and acquisition gamma is assumed to be unity. The output bitdepth is 8 bits.

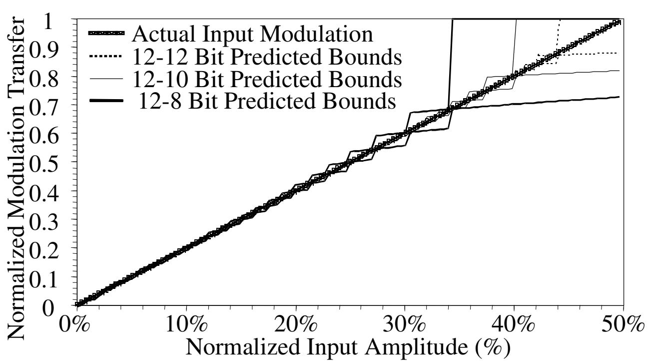

Figure 9. As for Figure 11, though changing the output bitdepth. The input bitdepth is 12 bits.

The figures initially demonstrate that gamma correction significantly affects the modulation of a signal. This is demonstrated by the difference between the curves for the actual and gamma corrected modulation. Transformation of exposures for discrete systems into linear input units is therefore vital.

Using values of the gamma corrected maxima and minima, Equations 23 and 24 successfully calculate a range which contains the actual input modulation. Figure 5 verifies that the input modulation is contained within the range at all points. The modulation range may be seen to be small in this instance and has a maximum value of •±0.01. Gamma correction is therefore not expected to significantly affect the determination of MTF for discrete systems that are gamma corrected for output to CRT, provided that the signal is transformed into linear input units.

The calculated modulation range is seen to be non-symmetric about the actual modulation, unlike that calculated for the effects of quantization in previous work.[11] It is therefore not possible to represent the effect of gamma correction as the input modulation plus and minus an error term as in Reference 13.

Figures 6 and 7 show the modulation range calculated for severe values of gamma correction, 0.1 and 3 respectively. The input and output bitdepths remain as 8 bits, also the target and acquisition values of gamma are unity.

The figures illustrate that as the value of the gamma correction is more extreme, the gamma corrected signal modulation deviates increasingly from the actual signal modulation. The more severe values of gamma correction cause more levels to be lost in the gamma corrected signal. This results in prediction of an increased range in which the actual modulation may lie.

The low value of gamma correction in Figure 6 results a greater predicted range for input with low modulation. The high value of gamma correction in Figure 7 results in increased predicted range for input signals with large modulation. This is unusual as sources of noise regularly affect signals with low modulation more severely than those with increased modulation. Large modulation normally results in increased signal to noise ratios for most processes.[10]

An explanation for this effect may be that gamma correction and the effect of integer mathematics is not an ergodic process. The results are calculated using values of the corrected signal and hence noise and signal components may not be separated.

Generally, the effect of high values of gamma correction may be seen to be more severe than that for low values. The predicted range is again shown to be non-symmetric in both figures. The actual input modulation continues to falls within the calculated range consistently.

Figures 8 and 9 demonstrate the effect of changing input and output bitdepths. It is clearly seen that changing the input bit depth of the signal has relatively little effect upon the predicted modulation range. Conversely, changing the output bitdepth has a more prominent effect. It may be seen that as the output bitdepth is increased, the predicted modulation range is smaller indicating that less errors will occur in measurements.

Conclusion

Gamma correction causes loss of available pixel values in discrete systems. This loss of values in turn causes error in MTF determination.

A technique has been shown that predicts the range of input modulation possible from gamma corrected maxima and minima. The limits have been shown to successfully contain the input value for a wide range of circumstances including severe gamma correction.

References

1. W.N. Sproson, Colour Science in Television and Displays Systems, Adam Hilger, Bristol, UK, 1983.
2. R.W.G Hunt, A model of colour vision for predicting colour appearance in various viewing conditions, Colour Reasearch and Applications, 12:297-314 (1987).
3. J.A.J. Roufs, The effect of gamma correction on perceived image quality, In IEEE International Display Conference, (1988).
4. G.C. Higgins, Image quality criteria, Journal of Applied Photographic Engineering, 3:53-60,1977.
5. C.J. Bartleson and E.J. Breneman, Brightness reproduction in the photographic process. Photographic Science and Engineering, 11:254-262, 197l.
6. A.M. Ford. Relationships between image quality and image compression. PhD Thesis, University of Westminster, UK, 1997.
7. S. Triantaphillidou, R.E. Jacobson, and A.M. Ford. Preferred tone reproductionof images on soft displays. In Exploring new tracks in imaging conference, Antwerp, Final Program and Proceedings, pg. 204-209, University of Antwerp, Belgium, 1998.
8. J.F. Blinn. Dirty pixels. IEEE Computer Graphics and Applications, 9:100-105, July 1989.
9. D.G. Pelli and L. Zhang. Acurate control of contrast on microcomputer displays, Vision Research, 31:1337-1350,1991.
10. J.C. Dainty, Methods of measuring the modulation transfer function of photographic emulsions. Optica acta, 18(11):795-813, 1971.
11. R. Jenkin, On the application of the modulation transfer function to discrete imaging systems, PhD Thesis, University of Westminster, UK, 2001.
12. R.C. Gonzalez and R.E. Woods, Digital imaging processing. Addison-Wesley Publishing Company, 1992.
13. R.B Fagard-Jenkin, R.E. Jacobson and J.R. Jarvis, The effect of quantization upon modulation transfer function determination. In S. Ohno, editor, Proceeding IS&T 1998 PICS Conference, Portland, Oregon, pages 135-140, Springfield, Virgina, USA, 1998.

Biography

Robin Jenkin received his BSc(Hons) in Photographic and Electronic Imaging Sciences from University of Westminster, UK, in 1995 and a research Masters in the field of computer vision from University College London in 1996.

Robin has completed a Ph.D in the field of discrete imaging system MTF evaluation at the University of Westminster. Robin works for Fuji Photo Film (UK) Ltd. and is a member of the Imaging Technology Research Group at University of Westminster. Current research interests include application of frequency response measures to non-linear imaging systems and the role of quality measures to predict performance of image processing algorithms.