Comparison of the First Order Wiener Kernel Transform of JPEG Compression with ISO 12233 SFR Measurements

R.B. Jenkin, M.A. Richardson and I.C. Luckraft
Cranfield University, Royal Military College of Science
Shrivenham, United Kingdom

Abstract

The Volterra and Wiener theories of non-linear systems provide techniques with which to evaluate the behaviour of non-linear systems. It may be shown that the First Order Wiener Kernel Transform may be thought of as being equivalent to the linear portion of the MTF of a system.

JPEG compression is such a non-linear system and as such measurement of its MTF has provided significant challenges. The First Order Wiener Kernel Transform of JPEG Compression (Version 6b) is determined under a variety of conditions and compared with SFR curves previously produced using ISO 12233. It is shown that there is agreement between the measurement methods. It is suggested that evaluation of the Wiener Kernel transform provides results that more closely represent the large-scale pictorial frequency response of JPEG compression due to the integration of intra and inter sub-image block effects. Further, that the technique represents an advantage over those previously considered.

Introduction

The measurement of the frequency response of an imaging system requires that it should be linear, spatially invariant and homogeneous.[1] If these conditions are not met then calculation of frequency based performance measures, such as the Modulation Transfer Function (MTF) or Noise Power Spectra (NPS), become prone to errors. Numerous pieces of work are documented regarding improvements that may be applied to the calculation of MTF in various systems.[2]

Of particular interest is the behaviour caused by non-linear components. These generally render the MTF 'scene dependant'. i.e. MTF varies with respect to the type of test target used and also the optical contrast of the test signal. The effects of non-linear sharpening on MTF have previously been simulated by the author.[3] Evaluation of MTF using edges of various contrast followed by basic interpolation of results was suggested to mitigate these difficulties.[3]

Also detailed was the application of Wiener Kernel measurement to extract a linear portion of frequency response from non-linear imaging systems. Results from computer based simulation showed that estimation of the first order Wiener kernel transform confidently extracted the linear portion of system response in the presence of non-linear agents. Further, that the first order Weiner Kernel was equivalent to the MTF of the linear agents.[3]

Evaluation of the MTF of JPEG compression has remained a challenge.[3,4] The success of the lossy compression system relies on the reduction of information in the image by discarding data relating to chrominance and high spatial frequencies.[4] The severity depends on the 'quality factor' specified and dictates the quantization tables used to perform the cull. The technique is therefore highly non-linear.[4]

Previous work by Ford et al[4-6] attempted to evaluate the MTF of JPEG compression in order to incorporate it into calculations of image quality metrics. Quality factor, however, was found to be a better overall indication of image quality than Perceived Information Capacity (PIC) or Square Root Integral with Noise (SQRIn).[4] A significant proportion of this result was attributed to the difficultly Ford encountered calculating the MTF of the compression system.[4]

The use of ISO 12233 Spatial Frequency Response[7,8] to evaluate JPEG offers significant advantages and has been previously documented by the author.[9] The potential shown by evaluation of the first order Wiener Kernel in previous simulations warrants its application to JPEG compression and comparison to results obtained using ISO 12233. Motivation for this work is to produce MTF curves that represent the overall pictorial effect of JPEG which are not subject to large scale corruption by non-linear processes.

ISO 12233, Volterra and Wiener Kernels

Numerous publications detail the theory used as the basis for ISO 12233 and it is not appropriate to reproduce it here, for example Refs. [2, 3, and 10].

The Volterra and Wiener theories of non-linear systems afford sophisticated descriptions of systems behaviour and as such a full account of their derivation is impossible here. An extensive work is published by Schetzen.[11] A useful summary and examples are provided by Burns.[12] A brief overview is given below.

The Volterra description of non-linear system behaviour is a generalized functional series.[12] The terms of the series are n-dimensional convolution integrals based on n-dimensional Volterra kernels.[11,12] For a one-dimensional stationary system:

$$
\begin{aligned}
r(x) = h_0 &+ \int_{-\infty}^{\infty} h_1(\tau) s(x-\tau)\delta\tau \\
&+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_2(\tau_1,\tau_2) s(x-\tau_1) s(x-\tau_2)\delta\tau_1\delta\tau_2 \\
&+ \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} h_n(\tau_1,\ldots,\tau_n) s(x-\tau_1)\ldots s(x-\tau_n)\delta\tau_1\ldots\delta\tau_n
\end{aligned} \tag{1}
$$

where *r(x)* is output, *s(x)* input, h_n the set of Volterra kernels and τ_n offset variables.[12]

The Volterra functionals rely on being able to describe system non-linearity as a power or polynomial series.[12] Increased system non-linearity is modelled by selecting significant terms of the series until *r(x)* describes behaviour to the desired degree of accuracy.[12]

In the same manner that a linear imaging system may be specified by either its LSF or OTF, a non-linear system may be specified by its set of Volterra kernels or their transforms. The transform for kernel h_n is given by:

$$H_n(\omega_1,\ldots,\omega_n) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} h_n(\tau_1,\ldots,\tau_n) e^{-2\pi i(\omega_1\tau_1+\ldots+\omega_1\tau_1)}\delta\tau_1\ldots\delta\tau_n \tag{2}$$

where H_n represents the n-dimensional Fourier transform of h_n and ω_n represents spatial frequency. Given a LNL system, after some mathematical working, it may be shown that its Volterra kernels are given by combinations of the linear component kernels.[12] The exact combination is determined by the order of the non-linearity.

Wiener developed the Volterra approach to consider non-linear system response to a white noise signal.[11,12] The Wiener expansion is given by:

$$
\begin{aligned}
r(x) = k_0 &+ \int_{-\infty}^{\infty} k_1(\tau) s(x-\tau)\delta\tau \\
&+ \int_{-\infty}^{\infty} k_2(\tau_1,\tau_2) s(x-\tau_1) s(x-\tau_2)\delta\tau_1\delta\tau_2 \\
&- C\int_{-\infty}^{\infty} k_2(\tau,\tau)\delta\tau \\
&+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} k_3(\tau_1,\tau_2,\tau_3) s(x-\tau_1) s(x-\tau_2) s(x-\tau_3)\delta\tau_1\delta\tau_2\delta\tau_3 \\
&- 3C\int_{-\infty}^{\infty} k_3(\tau,\tau_1,\tau_1) s(x-\tau)\delta\tau\delta\tau + \ldots
\end{aligned} \tag{3}
$$

where C is the spectral power of the input signal and k_n the set of Wiener kernels.[12] The expression is summarized by Burns as[12]:

$$r(x) = \sum_{n=0}^{\infty} G_n\left[k_n, s(x)\right] \tag{4}$$

where G_n is the set of Wiener functionals.[12] The Wiener functionals are orthogonal and thus the output of a system may be represented by:

$$r(x) = \lim_{n\to\infty}\left[G_0[k_0,s] + G_1[k_1,s] + \ldots + G_n[k_n,s]\right] \tag{5}$$

The benefit of this representation is that G_n operates only with kernels of order *n*. Therefore, Burns explains, evaluating additional terms of G_n does not change previously determined terms.[12] Thus, successive Wiener kernels may be determined until a good approximation of system behaviour is reached.[12] Differing non-linearities will produce terms of differing orders in the Wiener series. The first order term of the Wiener series may be thought of as representing the linear component of the system.[12] In a similar manner, the Wiener Kernel Transforms represent the Wiener Kernels in frequency space.

Estimation of the First Order Kernel Transform

For a white noise input, the results of an extensive derivation show that the first order Wiener kernel transform of a non-linear system may be estimated as[11,12]:

$$K_1(\omega) = \frac{\varepsilon\left[S^*(\omega)R(\omega)\right]}{C} \tag{6}$$

The Fourier transforms of the input and output of the system are denoted *S(x)* and *R(x)* respectively. The use of the complex conjugate is indicated by * and the ensemble average by ε.

Error in Transform Estimation

Burns concludes from empirical measurement that the standard deviation, σ_S, of single first order Wiener Kernel transform measurements is approximately the same as the kernel transform magnitude.[12] Taking *M* measurements, the standard deviation of the averaged result, σ_A, is reduced to:

$$\sigma_A = \frac{\sigma_S}{\sqrt{M}} \tag{7}$$

Experimental Method

The generation of SFR curves using ISO 12233 for JPEG compression by the author is published in detail in Ref. [9].

To estimate the first order Wiener Kernel of the compression system, various white noise targets of 1024x1024 pixels were generated using MATLAB.[13] Single channel monochrome targets were produced with noise of mean pixel value 128 and varying contrast. The range of minimum to maximum values of the noise varied between 32 and 224 pixel values. In addition colour targets were prepared in a similar manner by generating noise for red, green and blue channels. The generated images were recorded as 8 bit Targa files with no compression.

Targets were compressed using version 6b of the Independent JPEG Groups' implementation of the JPEG standard at various quality factors.[14] Monochrome targets were compressed using the monochrome setting for the compression standard.

Evaluation of the Wiener Kernel estimates was performed using MATLAB on a PC. Equation 6 was implemented and the original uncompressed files were compared to those after decompression. The first order Wiener Kernel is estimated for each row of the image, the results are then averaged to produce a single kernel for the image. The kernel estimates are then scaled to ensure the DC component has a value of one for direct comparison to the MTFs calculated using the ISO 12233 standard.

ISO 12233 strictly produces SFR curves as no attempt is made to correct for the frequency content of the target. The SFR curves used here have been corrected to account for the frequency content of the edges used in the analysis. Full details are given in Ref. [10].

As for previous work, the tone reproduction of the compression system was noted as varying with respect to the quality factor of the compression used.[10] The average gradient, however, is close to unity and thus the tone reproduction was again estimated as linear in both the monochrome and colour case.[10]

Results and Discussion

Figures 1, 2 and 3 show previous attempts to measure the MTF of JPEG compression with respect to quality factor. Figures 1 and 2 shown the results of Ford using traditional sine wave and edge techniques.[4] Figure 3 is the result of measurement using ISO 12233 with a monochrome edge of 64 pixel value transition (95-159) angled at 5°. Figure 4 shows the estimates of the first order Wiener Kernel transform using monochrome white noise of 64 pixel value range (95-159).

The differences between the results shown in Figures 1, 2 and 3 have previously been detailed.[4,10] Measurement with sine waves, Figure 1, produces optimistic MTF curves as the test signal power is contained in few of the DCT coefficients because spatial frequencies are presented individually. Thus, they are not subject to the effects of quantization.[4,10] The traditional edge technique used by Ford, Figure 2, is affected by the effects of phase and aliasing.[4] Also, because of the presence of multiple spatial frequencies, it is much more prone to the effects of the quantization stage.[4] Ford demonstrated that the position of the edge relative to sub-image blocks additionally affected results.[4]

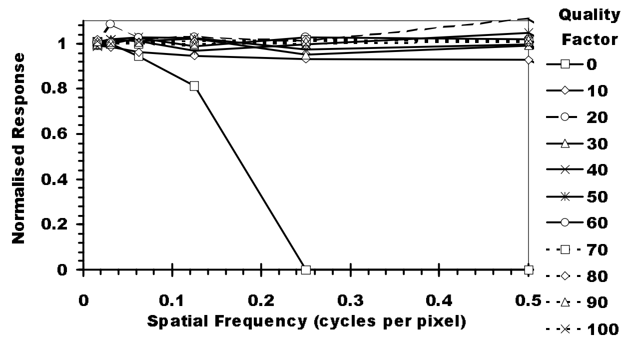

Figure 1. Ford's sine wave measurement with respect to quality factor. Reproduced from Ref. [4].

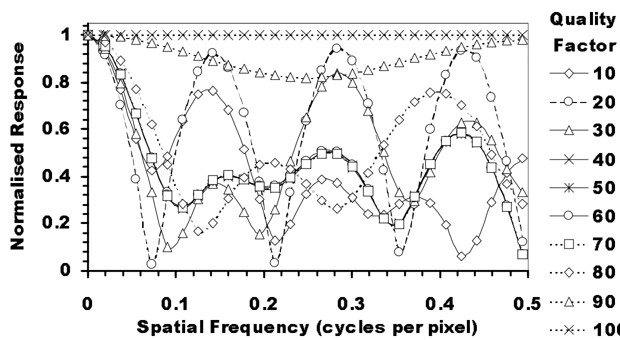

Figure 2. Ford's MTF results derived using the traditional edge technique. Diagram reproduced from Ref. [4].

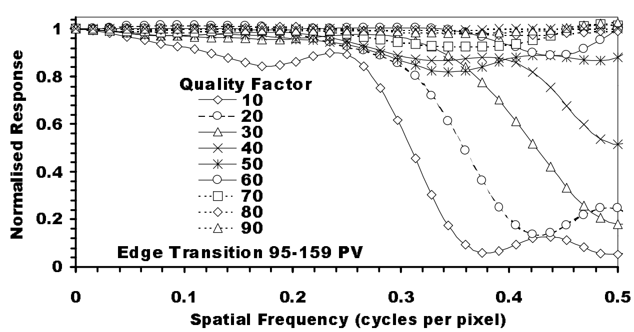

Figure 3. Measured MTF derived using ISO 12233 with respect to quality factor. Reproduced from Ref. [10].

The author's previous work showed that the integration of effects over a region of the image, afforded by the use of ISO 12233, provides results that vary in a more intuitive manner, Figure 3.[10] Band-limitation of the test signal appeared to reduce the effects of aliasing and the MTF generally decreased with increasing spatial frequency. Finally, reducing the quality factor resulted in a more rapid fall of the response as frequency increases.[10] These features were not apparent in curves produced using either the sine wave or traditional edge techniques. It was therefore argued that the results produced using ISO 12233 were superior as they represented the response of the system to multiple spatial frequencies and, due to integration, partially mitigated non-linear effects.

Figure 5 shows, however, that the results produced using ISO 12233 are affected by the size of the region of interest used. Further, it may also be suggested that some interaction between sub-image blocks and the sloping edge may be apparent as the curves do not monotonically decrease as shown in Figure 3. The integration of effects offered by ISO 12233 is effectively constrained by the size of the test image used and type of target.

In comparison, estimation of the first order Wiener Kernel transform affords large scale integration of the image, Figure 4. The family of curves are monotonically decreasing and, in agreement with ISO 12233, MTF generally improves with increasing quality factor. Using Equation 7, the standard deviation for each point in the graph is estimated as 3.1%. Thus, 95% of values will be described using $\pm 2\sigma_A$ or ±6.25% of the kernel transform magnitude.

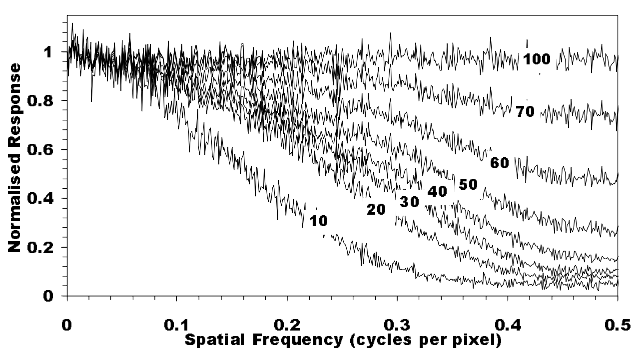

Figure 4. First order kernel transforms with respect to quality factor for monochrome white noise of range 95-159 pixel values.

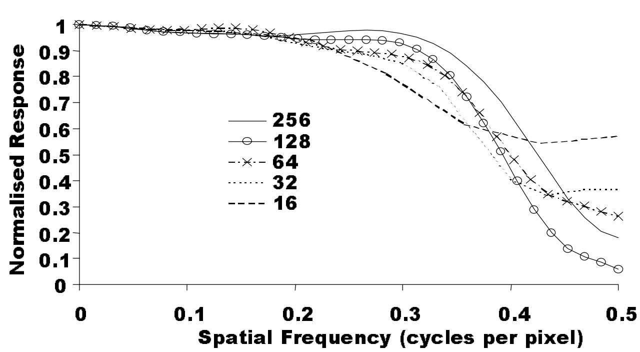

Figure 5. Variation in MTF with respect to the height of the region-of-interest selected (16-256 pixels) for an edge of transition 95 to 159 pixel values compressed using a quality factor of 30 and measured using ISO 12233. Reproduced from Ref. [10].

Figure 6 shows direct comparison of the results produced using ISO 12233 and the first order Wiener Kernel Transform. It may be seen that the results of the first order kernel transform are significantly lower than those produced using ISO 12233. It may be suggested that the division of the image by JPEG into sub-image blocks aids the reproduction of edges at the boundaries of these partitions. Though the edge used in ISO 12233 is sloping this advantage will be incorporated into the results. The random structure of the white noise target does not permit this to occur. For the results of any frequency analysis of JPEG compression to have maximum value in a particular field it is clear that statistically matching the structure and frequency content of the test target with the intended subject will produce optimum results. In the absence of this ability, it may be argued that the use of any evaluation methodology that provokes a bias in the results should be avoided.

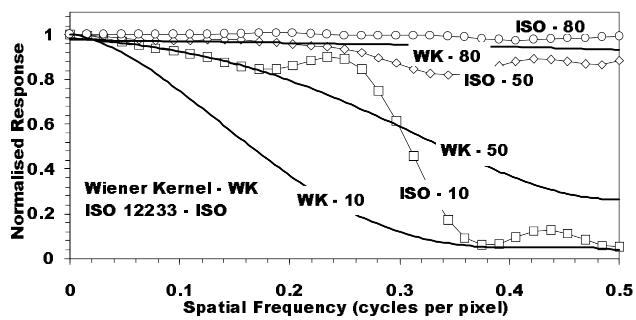

Figure 6. Comparison curve fits of the first order Wiener Kernel Transforms with results produced using ISO 12233 with respect to quality factor. Some curves are omitted for clarity.

Figures 7 and 8 show the measured first order kernel transform for white noise of varying contrast at quality factors of 90 and 30. In agreement with results produced using ISO 12233, in previous work,[10] it is seen that, at a high quality factor, the measured result does not vary with target contrast. At a low quality factor, however, the target contrast influences the result obtained. Poor signal strength leads to low MTF. This may be explained by the loss of small DCT coefficients at the quantization stage of the compression. The first order Wiener Kernel transforms are again significantly lower than those produced using ISO 12233 at a quality factor of 30. The previous explanation is suggested to account for this result.

Figures 9, 10 and 11 show the first order Weiner Kernel transforms for the red, green and blue channels respectively at various quality factors. As is shown in the previous work measured using ISO 12233, the frequency response of any single colour channel is lower than that for a monochrome signal. This is expected due to the conversion of the initial R, G, B, colour space to Y, Cb, Cr, and the subsequent sub-sampling of the chrominance information. The calculation of the colour space conversion distributes the test signal power throughout the Y, Cb and Cr channels resulting in reduced contrast. The sub-sampling reduces this response further. It may be seen that the green channel displays the best response, followed by the red then blue. This mimics the contribution made by each channel to the calculation of the luminance signal, below[10]:

$$Y=0.2989R+0.5866G+0.1145B \tag{12}$$

It is also evident that the noise displayed in the results also follows this pattern with least being exhibited in the green channel.

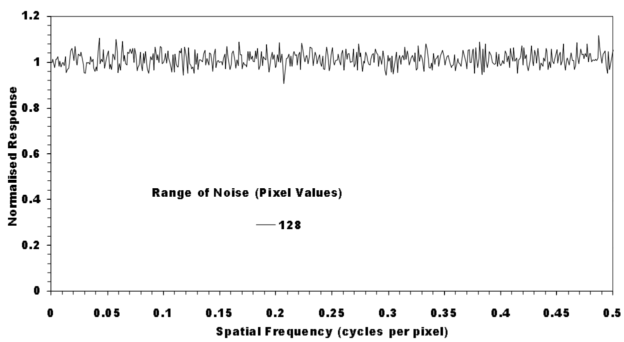

Figure 7. First order kernel transforms evaluated using monochrome white noise of varying contrast and compressed using a quality factor of 90. The mean value of the noise is 128.

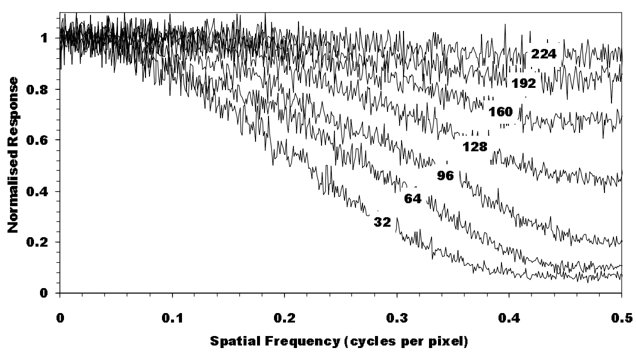

Figure 8. As above using a quality factor of 30.

As suggested in previous work, the variation of the results with respect to contrast of the test signal and colour channel demonstrate the highly complex nature of the compression system.[10] As such this should not be considered a deficit in the measurement methodology. The use of the first order Wiener Kernel transform has enabled the integration of results in a manner to reduce the effects of local non-linearities. In addition, the use of white noise appears to have reduced the interaction between sub-image blocks and the edge of the test target used in ISO 12233. The result is credible frequency response curves that may reasonably be employed in further calculations. Matching of the experimental conditions under which the compression system is measured to those in which it is used will further increase the value of the results as previously mentioned. In addition, development of methodologies to evaluate the departure of the imaging system from the general description offered by the curves would prove advantageous. Agents whose behaviour changes depending on the locale in which they operate are becoming increasingly prevalent in imaging systems as processing techniques develop in sophistication and it may be reasonably argued that the need for this approach can only increase in the future.

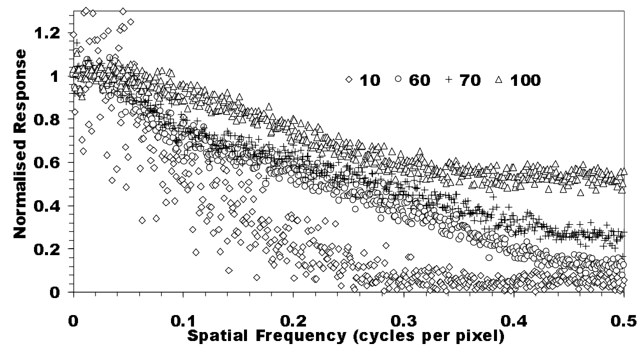

Figure 9. First order kernel transform for the red channel of image measured using white noise of mean 128 and 64 pixel value range. Curves have been omitted for clarity.

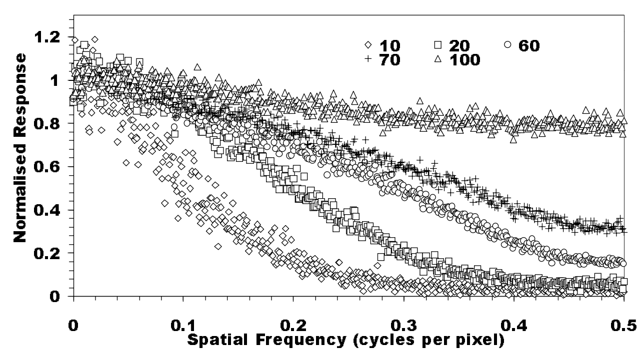

Figure 10. As above for the green channel.

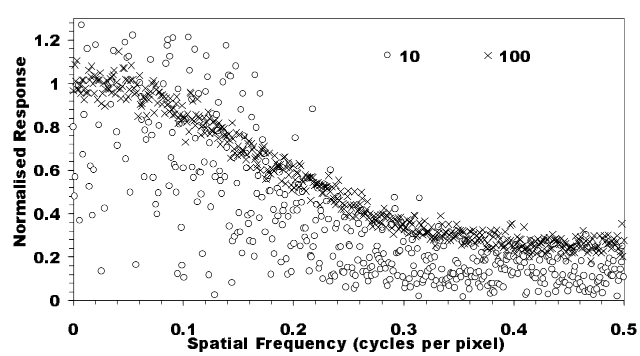

Figure 11. As above for the blue channel.

Conclusion

The first order Wiener kernel transform of JPEG compression has been evaluated under a variety of conditions and compared with results produced using ISO 12233. The technique has been shown to offer advantages because of its ability to integrate local effects to produce generalized curves. In particular, for JPEG compression, any possible interaction between test target structure and sub-image blocks is lost due to the use of white noise. This is believed to reduce bias in the results. The first order kernel transforms agree with the general trends exhibited by results produced using ISO 12233. Frequency response increases with quality factor. Chrominance based response varies with colour channel and is affected greatly by sub-sampling and quantization.

References

1. E H Linfoot, Fourier Methods in Optical Image Evaluation, Focal Press, London, 1964.
2. J C Dainty and R Shaw, Image Science, Academic Press, London, 1974.
3. R B Jenkin, R E Jacobson and N R Axford, In Proceedings, 54th Imaging Science and Technology PICS Conference, 21-23 April 2001, Montreal, Canada.
4. Ford A M, PhD Thesis, University of Westminster, UK. (1997)
5. R E Jacobson, A M Ford and G G Attridge, In Proceedings, Human vision and Electronic Imaging II, SPIE (1997)
6. A M Ford, In Colour Imaging Vision and Technology, Editors L W MacDonald and R Luo, John Wiley and Sons LTD, UK. (1999).
7. ISO/FDIS12233:1999(E). International Organization for Standardization, New York (1999)
8. D Williams, IS&T PICS 51st Conference Proceedings, 133 (1998)
9. R B Jenkin, R E Jacobson and K MacLennan-Brown, In Proceedings, 53rd Imaging Science and Technology PICS Conference, 26-29 March 2000, Portland, Oregon, USA.
10. S E Reichenbach, S K Park and Narayanswamy, Optical Engineering, 30(2), 170-177 (1991).
11. M Schetzen, The Volterra and Weiner Theories of Non-Linear Systems, Wiley, New York, 1980.
12. P Burns, Measurement of Non-Linear Systems, Masters Project, Clarkson University, 1977.
13. MATLAB Computation, Visualization, Programming, Version 6.12, The Mathsworks Inc., Natick, MA, USA.
14. Independent JPEG Group, JPEG Implementation Version 6b, http://www.ijg.org/.

Biography

Robin Jenkin received his Ph.D from University of Westminster (2001). He also holds M.Res Computer Vision and Image Processing (1996) from University College London and a BSc(Hons) Photographic and Electronic Imaging Sciences (1995) again from University of Westminster. Robin lectures in the field of electro-optics at Cranfield University. His current research interests include development of image quality measures for non-linear systems and their relationship to perceived image quality. Robin is also an active member of the Image Science Group of the Royal Photographic Society, UK.