Determination of the MTF of JPEG Compression Using the ISO 12233 Spatial Frequency Response Plug-in.

R. B. Jenkin, R. E. Jacobson and K. MacLennan-Brown
Imaging Technology Research Group, University of Westminster, Middlesex. HA1 3TP. United Kingdom.

Abstract

The Modulation Transfer Function (MTF) of the Joint Photographic Experts Group (JPEG) compression (Version 6b) [1] was evaluated using the ISO 12233 Spatial Frequency Response (SFR) Plug-in [2,3]. Results were obtained using Gaussian edges with respect to quality factor, target contrast and image channel perturbation. The effect of selection of the region-of-interest was also examined. Results were compared to those previously obtained using sine wave and traditional edge techniques [4].

This work shows that ISO 12233 represents an advantage over previous techniques. A number of issues associated with the extreme non-linearity of JPEG compression are surmounted. This advantage, however, is not of sufficient significance to justify inclusion of results in applications such as quality metrics [5].

The findings show that as quality factor is reduced the MTF of the compression is generally lowered. An interaction between target contrast and quality factor is shown to exist. As quality factor is reduced, variation in the measured MTF due to target contrast increases. Selection of the region-of-interest is shown to affect results.

Introduction

JPEG compression has become a widely recognized standard for lossy encoding [6]. Based primarily on the discrete cosine transform (DCT) its success lies in the useful compression ratios (1:10 to 1:25) that may be achieved [4]. The technique relies on reduction of information in the image, followed by entropy encoding of remaining data. Thus decompressed images differ from originals [4]. This is perceived as a change in quality of the image [4].

JPEG compression is included in many contemporary imaging chains, from digital cameras to the internet. In order to assess the overall quality of these chains some assessment of the quality of JPEG compression is needed.

Previous work by Ford et al. [4,7,8] compared the application of metrics such as Perceived Information Capacity (PIC) [9] and Square Root Integral with Noise (SQRIn) [10] to quality factor as a measure of quality. Quality factor was shown in that work to be a superior indication of overall quality [4]. The consideration for this is involved and references [4,7,8] provide useful information. A consistent and considerable problem, however, was evaluation of the MTF of the compression system [4]. Contrary to Fourier theory, evaluation by different techniques is not equivalent [4]. Compounding these problems is the non-stationary and non-isotropic nature of JPEG, demonstrated in reference [4].

The work of Reichenbach et al. has provided a partial solution concerning the non-stationary nature of digital systems by development of the sloping edge technique [11]. Adopted by Photographic and Imaging Manufacturers Association (PIMA) it has been developed into ISO 12233 and produced as a Adobe PhotoShop plug-in and also implemented in Matlab [3]. Whilst it is not appropriate to detail the exact method in this paper, it is given considerable coverage in literature, also for a number of applications [12-16].

A consequence of producing super-sampled edges is that the resultant SFR is effectively a mean response calculated over the region-of-interest selected. Initially avoiding the debate as to the validity of results, it is this that motivated an application to JPEG compression.

Experimental Method

The SFR plug-in requires that a sloping edge is used as input. A Gaussian function of an appropriate width is a good approximation to the Point Spread Function (PSF) of most imaging systems. This premise is used to produce various colour and monochrome targets.

Using appropriate software, an image of a step-edge (800 x 800 pixels) of the required magnitude is produced. This is rotated and Gaussian convolution applied. The result is cropped to the desired size (64 x 256 pixels) and stored in a lossless format. In order to keep this process consistent the edge manufacture, rotation and selection is centered in the image.

Consideration must be given to the effects of aliasing in the results. The effects will depend upon the energy of the test target spatial frequencies which lie above the Nyquist frequency of the system [17]. This may be engineered satisfactorily by adjusting the width of the Gaussian convolution kernel used. Figure 1 shows edges with a 5° slope that has been convoluted with kernels of varying width, σ. The edge transition is from a pixel value of 95 to 159.

Figure 2 shows the SFR generated by each edge. Though there is little visual difference, the graph illustrates that a Gaussian function with σ = 0.2 is most appropriate. There is a relatively small signal content above the Nyquist frequency and reasonable content below. If the spatial frequency content of the test signal is too small, as for σ = 0.5, the signal-to-noise of the results will decrease. An estimate of the potential for aliasing as defined by Kriss was calculated for the chosen target as 0.2 assuming that there is no appreciable signal above 2 cycles per pixel [17].

Using the above, test targets were produced to investigate a number of areas including angle of edge, selection of edge, test target contrast, compression ratio and colour channel. Exact details of targets may be found with the presented results.

Once prepared, targets were compressed using version 6b of the Independent JPEG Groups implementation of the JPEG standard [1] at various quality factors.

Measurement of the SFR of the compressed images was performed using the Matlab implementation of the standard on an IBM compatible personal computer.

The transfer function of JPEG compression varies with respect to the quality factor [4] and is primarily dependent on quantization tables used. However, it is also noted that average gradients of transfer functions are close to unity [4] within quantization limits. For simplicity the Opto-Electronic Conversion Function (OECF) used was a simple linear function in both monochrome and multi-channel cases.

ISO 12233 is designed to produce the SFR of a system. The response is so titled as it makes no attempt to correct the result for the frequency content of test targets used [3]. Assuming that, as for other systems, results may be cascaded, MTFs may be produced by dividing the output by the SFR of the input used.

Results and Discussion

Figures 3, 4, and 5 show measured MTF with respect to quality factor. Results shown in Figure 3 are those produced using ISO 12233 with a monochrome edge (95-159 pixel value) angled at 5°. Figures 4 and 5 are those of Ford using traditional sine wave and edge techniques.

It is clear from the figures that the results derived using the three methods are completely different. An attempt to reason for differences between the traditional sine wave and edge techniques has been provided by Ford [4,8]. This is summarized as follows. Representation of single spatial frequencies, i.e. sine waves, using the DCT places the majority of power in a minority of coefficients.

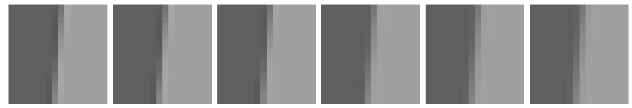

Figure 1. Sections of monochrome edges convoluted with Gaussian kernels of width (left to right) σ = 0, 0.1, 0.2, 0.3, 0.4 and 0.5 pixels. Complete size of originals 64 x 256 pixels.

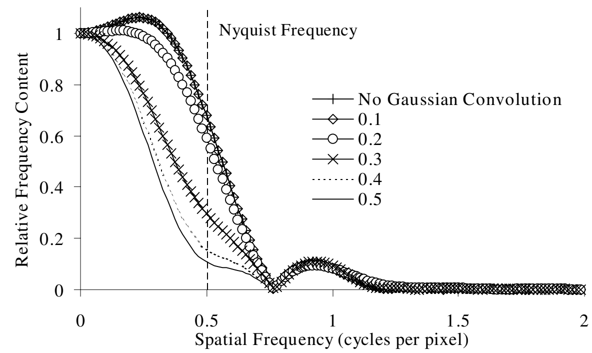

Figure 2. Relative spatial frequency content of the above edges

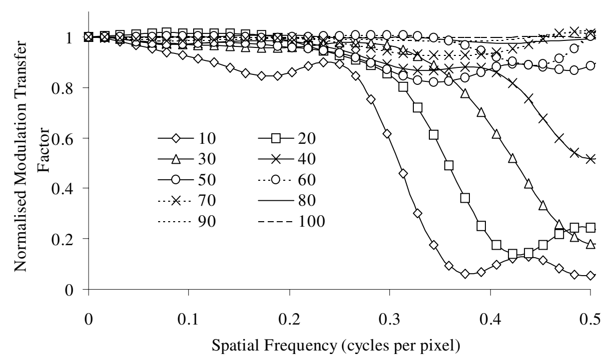

Figure 3. Measured MTF with respect to quality factor(10-100) for monochrome edges with a transition from 95 to 159 pixel values.

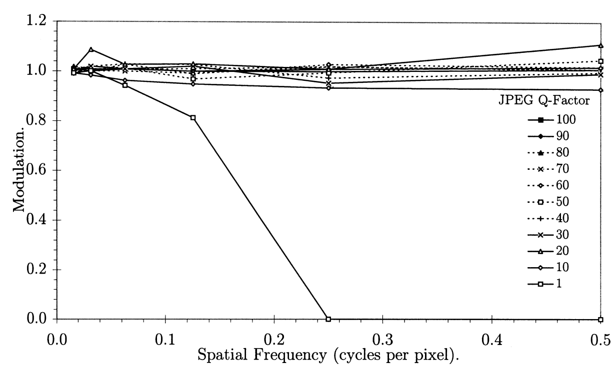

Figure 4. Ford's measurement of JPEG using a sine wave technique with respect to quality factor. Reproduced from reference [4].

Quantization scales the resultant coefficients, however, has little effect because of the above power distribution [4,8]. The traditional edge technique suffers from the accumulative effects of phase, aliasing and quantization [4,8]. Further due to the non-stationary nature of the compression these effects change with respect to position of the edge within the sub-image block as demonstrated in reference [4].

It may be argued that the results derived using edge techniques more closely resemble the pictorial effects of JPEG. The algorithm would never normally be applied to images containing anything other than numerous spatial frequencies. Figure 3, shows that results produced using ISO 12233 better correspond to Ford's edge results. Due to the construction of the super-sampled edge, effects generated because of the location of the edge within the sub-image block are mitigated. Effectively, a mean MTF is produced. This may be conceptually interpreted as the mean effect of the algorithm across an image. For critical applications, however, the deviation possible from this mean has to be considered.

It may be seen that the curves produced using ISO 12233 are generally higher than Ford's edge data. This is partially accounted for because the traditional technique only accounts for intra sub-image block effects. Inter-block effects occur when the edge position coincides with block edges. This coincidence aids edge reproduction, increasing local MTF. The super-sampled edge incorporates this into results unlike the traditional technique and is more representative of the overall pictorial effect.

Figure 6 demonstrates the effect of the sub-image blocks upon evaluation of the MTF. It may be seen that as the height of the region-of-interest selected is reduced so the MTF varies significantly.

The design of the test target which has limited frequency content above the Nyquist may also affect results. Ford's edge was not frequency limited and thus prone to aliasing.

Apart from the above differences the results provided by both ISO 12233 and Ford's previous work agree in that as quality factor increases, the general response increases. The variation in response is more regular for those produced with ISO 12233.

JPEG compression is an incredibly non-linear process. It would be naive not to investigate any variation with respect to other parameters in the system. Test target contrast is of significance. Figures 7 and 8 show edge MTFs derived for edge of varying contrast and two quality factors.

The figures clearly show that there is interaction between contrast of the test target used and quality factor. Whilst contrast has little effect using a high quality factor at low quality significant variation is seen. Further, it is seen that as contrast of the test target is reduced response is diminished. This may be explained because the effects of quantization on small signals will be relatively higher.

The variation in this additional dimension questions the usefulness of the results as variation in local MTF is compounded further. In order to employ results in quality metrics it is suggested that mean edge magnitude should be calculated to assess the mean MTF of the compression. The solution is far from ideal but allows some estimation of the magnitude of effects in systems that incorporate JPEG compression.

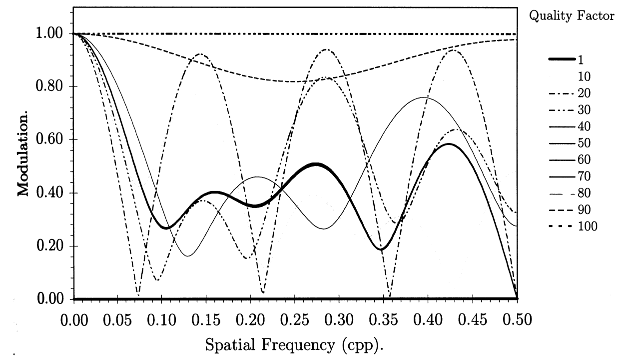

Figure 5. Ford's MTF results derived using the traditional edge technique. Diagram reproduced from reference [4].

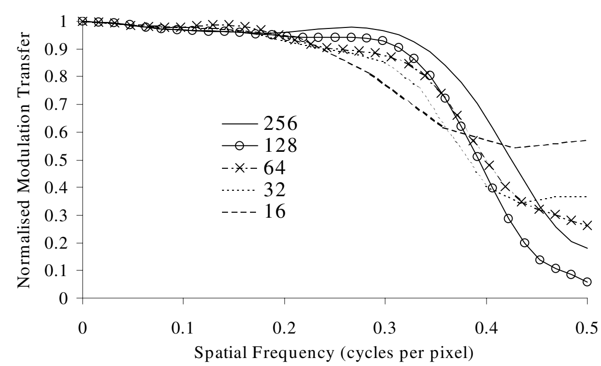

Figure 6. Variation in MTF with respect to the height of the region-of-interest selected (16-256 pixels) for an edge of transition 95 to 159 pixel values compressed using a quality factor of 30.

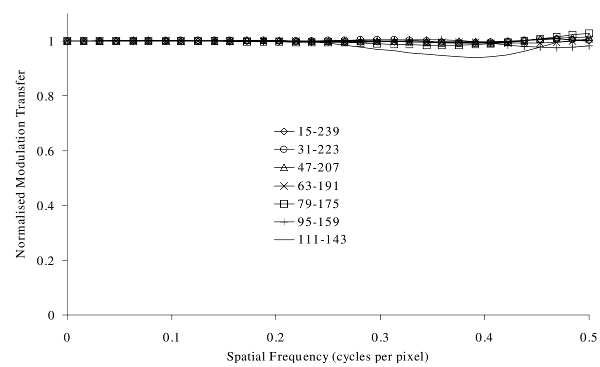

Figure 7. Responses using monochrome edges of varying magnitude compressed with a quality factor of 90.

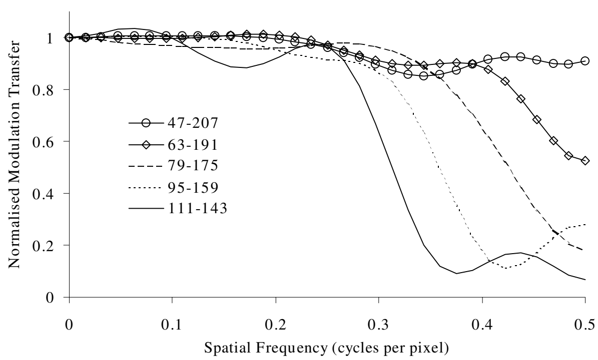

Figure 8. Responses using monochrome edges of varying magnitude compressed with a quality factor of 30.

In order to evaluate the effect of colour, edges of 64 pixel value magnitude (transition from 95 to 159) were created in each of the red, green and blue channels. The remaining channels were kept constant at a value of 128. Figures 9 and 10 show the resultant luminance based MTF using weighting coefficients of 0.3, 0.6 and 0.1 for the red, green and blue channels respectively.

The figures clearly belie the complexity of performing compression in a chrominance-luminance based colour space. Figure 9 shows that edge transitions in the red and green colour channels do not affect overall frequency response greatly. The response is similar to that generated for a monochrome edge of the same magnitude and quality factor. If a edge is generated in the blue channel, however, the response is significantly lower. This is surprising as the blue channel contributes least to the weighted average for calculation of luminance and the result is counter-intuitive. Why should a perturbation in a single channel of a colour image give a lower overall response than that for a monochrome image? It appears that this may be explained by examining the conversion from RGB to YCbCr primaries, below [4]:

$$Y=0.2989R+0.5866+0.1145B$$

$$Cb=-0.1687R-0.3312G+0.5000B$$

$$Cr=0.5000R-0.4183G-0.0816B \qquad (1)$$

It is seen that an edge in the blue channel will affect the Y, Cb and Cr channels. Figures 5 and 6 also show that as the contrast of a monochrome edge reduces so response is reduced. Aside from sub-sampling of the chrominance channels and differing quantization tables, the procedure to compress the Y, Cb and Cr channels is the same as for a monochrome signal. It is therefore apparent that a perturbation in the blue channel is converted into three low contrast perturbations in the Y, Cb and Cr channels. As low contrast signals are reproduced poorly the overall response could easily fall below that for an edge in a monochrome image.

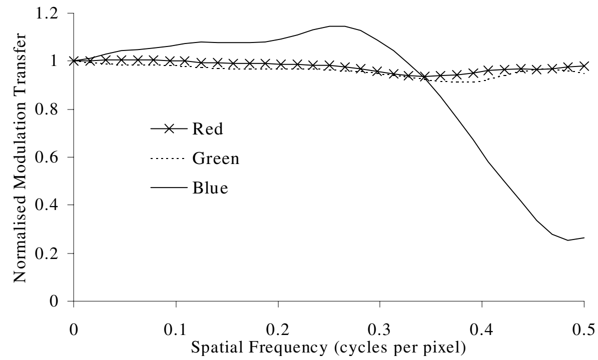

Figure 9. Calculated overall MTF for edge transitions in each of the red, green and blue colour channels compression using a quality factor of 90.

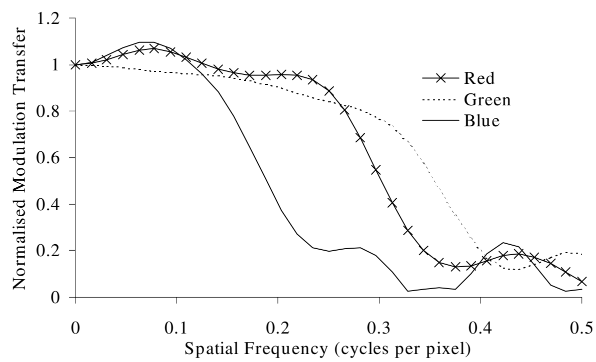

Figure 10. Calculated overall MTF for edge transitions in each of the red, green and blue colour channels compression using a quality factor of 30.

Figure 8, shows that, as expected using a quality factor of 30 results in a lower MTF. The findings show that as for a quality factor of 90, the MTF of the system when an edge is generated in the red and green channels is close to that for the monochrome case. Again for the blue channel the response is reduced. This result causes the MTF of JPEG compression to change with the colour of the edge used to evaluate it.

As in reference [4] JPEG compression has been shown to be highly scene dependent. The results generated using ISO 12233 however appear to be more consistent and closer to the pictorial effects of the compression. Ford's implementation of PIC and SQRIn showed that quality factor was a better quality metric than either. This may be due to the MTFs used to calculate this and re-calculation given these new findings may change this.

Consideration should be given to using the quantization tables themselves in order to evaluate the frequency response of the system when compared to a signal. Work is continuing in this area.

Conclusion

The measurement of JPEG compression using ISO 12233 represents an advantage over previous techniques used. The majority of this advantage relies on the generation of the super-sampled edge averaging the non-linear effects of the process. The results generated are closer to the overall pictorial effects of the compression.

JPEG compression, however, is highly scene dependent. Measured MTF is shown to vary with target contrast, colour and position. Because of the amount of averaging needed to reduce these results to a single mean MTF curve their use in image quality metrics may reasonable be questioned, but investigation of this is certainly warranted.

Acknowledgements

Many thanks to Dr. Adrian Ford upon whose original work this extension is based.

References

1. Independent JPEG Group, JPEG Implementation Version 6b, http://www.ijg.org/.
2. ISO/FDIS12233:1999(E). International Organization for Standardization, New York (1999).
3. Williams D., IS&T PICS 51st Conference Proceedings, 133 (1998).
4. Ford A. M., PhD Thesis, University of Westminster, UK. (1997).
5. Jacobson R. E., Journal of Photographic Science, 43, 7-16 (1995).
6. Pennebaker W. B. and Mitchell J. L., JPEG Still Image Data Compression Standard, Van Nostrand Reinhold, New York (1993).
7. Jacobson R. E., Ford A. M., and Attridge G. G., In Proceedings, Human vision and Electronic Imaging II, SPIE (1997).
8. Ford A. M., In Colour Imaging Vision and Technology, Editors MacDonald L. W. and Luo R., John Wiley and Sons LTD, UK. (1999).
9. Töpfer K. and Jacobson R. E., Journal of Information Recording Materials, 21, 5-27 (1993).
10. Barten P. G. J., In Human Vision, Visual Processing and Digital Display II, Editors Rogowitz B. E., Brill M. H. and Allebach J. P., SPIE (1991).
11. Reichenbach S. E., Park S. K. and Narayanswamy, Optical Engineering, 30(2), 170-177 (1991).
12. Okano Y., IS&T PICS 51st Conference Proceedings, 74 (1998).
13. Triantaphillidou S., Jacobson R. E. and Fagard-Jenkin R. B., IS&T 52nd PICS Conference Proceedings, 231 (1999).
14. Triantaphillidou S., and Jacobson R. E., IS&T 53rd PICS Conference Proceedings, p. 139 (2000).
15. Fagard-Jenkin R. B., Jacobson R. E. and Axford N., IS&T 52nd PICS Conference Proceedings, 225 (1999).
16. Burns P. and Williams D., IS&T 52nd PICS Conference Proceedings, 51 (1999).
17. Kriss M. A., IS&T 51st PICS Conference Proceedings, 247 (1998).

Biography

Robin Jenkin received his BSc(Hons) Photographic and Electronic Imaging Sciences degree from the University of Westminster in 1995. His masters degree in the field of computer vision and image processing was awarded by University College London in 1996. Now as a lecturer in image science at the University of Westminster Robin has just submitted his PhD thesis for examination in the field of evaluating digital systems frequency response. Robin's research interests lie in the field of quality determination and modeling.