Watson, A. B., Gale, A., Solomon, J. A. & Ahumada, A. J., Jr. (1994). Visibility of DCT Quantization Noise: Effects of Display Resolution Digest of Technical Papers, Society for Information Display, Annual meeting, San Jose, CA.

VISIBILITY OF DCT QUANTIZATION NOISE: EFFECTS OF DISPLAY RESOLUTION

Andrew B. Watson[1], Alan Gale[2], Joshua A. Solomon[1], Albert J. Ahumada, Jr.[1]
[1]NASA Ames Research Center , [2]San Jose State University

Abstract: The Discrete Cosine Transform is widely used in image compression. To better understand the visibility of DCT quantization errors and thereby design better perceptual quantizers, we measured visibility of DCT quantization noise as a function of display visual resolution in pixels/degree. Visibilities are consistent with a model incorporating effects of block size and spatial pooling.

INTRODUCTION

Background

In the JPEG, MPEG, and CCITT H.261 image compression standards, and several proposed HDTV schemes, a DCT is applied to 8 by 8 pixel blocks, followed by uniform quantization of the DCT coefficient matrix. The quantization bin-widths for the various coefficients are specified by a quantization matrix (QM).

QM design depends upon the visibility of quantization errors at the various DCT frequencies. In recent papers, Peterson *et al.* [1, 2] measured threshold amplitudes for DCT basis functions at one viewing distance and several mean luminances. Ahumada and Peterson [3] devised a model that generalizes these measurements to other luminances and viewing distances, and Peterson et al. [4] extended this model to deal with color images. From this model, a matrix can be computed which will insure that all quantization errors are below threshold. Watson [5] has shown how this model may also be used to optimize the quantization matrix for an individual image.

Objective

Visual resolution of the display (in pixels/degree of visual angle) may be expected to have a strong effect upon the visibility of DCT basis functions, and we therefore collected data to document this effect and to validate and enhance the model. We have examined three viewing distances that span a large part of the range: 16, 32, and 64 pixels/degree.

METHODS

Detection thresholds for single basis functions were measured by a two-alternative, forced-choice method. The stimulus was a single DCT basis function, added to the uniform gray background that remained throughout the experiment. Background luminance was 40 cd m^{-2}, and frame rate was 60 Hz. Observers viewed the display screen from distances of 48.7, 97.4, 194.8 cm. Display resolution was 37.65 pixels/cm. Images were magnified by two in each dimension, by pixel replication, to reduce monitor bandwidth limitations, resulting in magnified pixel sizes of 1/16, 1/32, and 1/64 of a degree, respectively at the three viewing distances. We describe these three viewing distances as yielding *effective* visual resolutions of 16, 32, and 64 (magnified) pixels/degree.

The contrast on each trial was determined by an adaptive QUEST procedure [6], which converged to the contrast yielding 82% correct. After completion of 64 trials, thresholds were estimated by fitting a Weibull psychometric function [7]. Thresholds are expressed as contrast (peak luminance, less mean luminance, divided by mean luminance), converted to decibel sensitivities ($-20 \log_{10}$[threshold])

We measured thresholds for only 30 of the possible 64 basis functions, as indicated in Fig. 1. We felt that thresholds would change sufficiently slowly as a function of DCT frequency that this sampling would be sufficient to constrain our model.

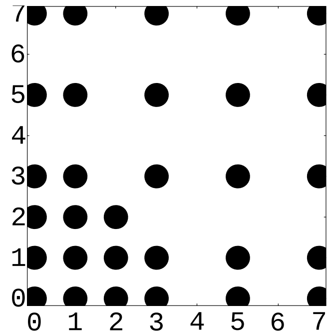

Figure 1. Subset of DCT frequencies used in the experiment.

MODEL OF DCT CONTRAST SENSITIVITY

The model of DCT contrast sensitivity that we consider here is essentially that described by Peterson et al. [4] In that model, log sensitivity versus log frequency is a parabola, whose peak value, peak location, and width vary with mean luminance. In addition, sensitivity at oblique frequencies ({u≠0,v≠0}) is reduced by a factor that is attributed to the orientation tuning of visual channels. The parameters of significance here are s_0 (peak sensitivity), f_0 (peak DCT frequency at high luminances), and k_0 (inverse of the *latus rectum* of the parabola), and r (the orientation effect).

RESULTS

Figures 2, 3, and 4 show decibel sensitivities for the three viewing distances. The data are accompanied by predictions of the best fitting version of the model. Within each figure, the three panels show data for vertical frequencies {0, v}, horizontal frequencies {u, 0}, 45 degree orientations {u, v=u}, and the remaining obliques {u>0, 0<v≠u}, all plotted against the radial frequency $f=\sqrt{u^2+v^2}$. In the case of the obliques, because there is no simple one-dimensional prediction to plot, we plot instead the deviations of the data from the model.

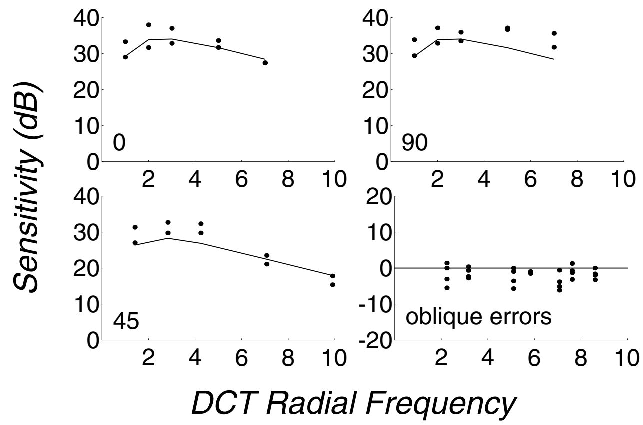

Figure 2. DCT basis function sensitivities at 16 pixels/degree.

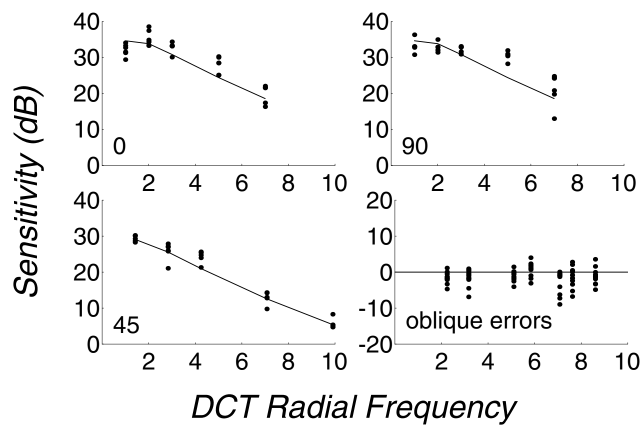

Figure 3. DCT basis function sensitivities at 32 pixels/degree.

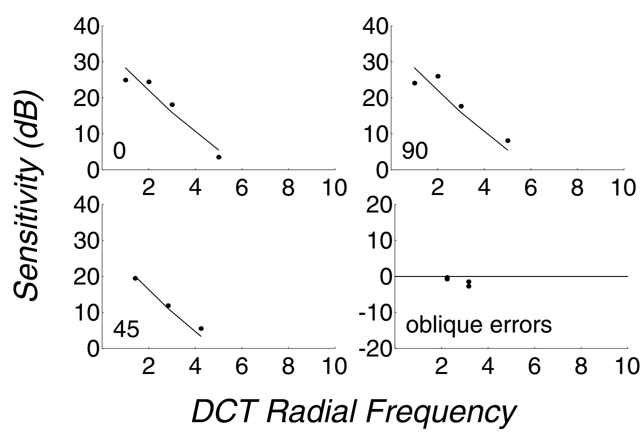

Figure 4. DCT basis function sensitivities at 64 pixels/degree.

The fits are reasonable, though there are some apparent systematic departures from the model. For reference, the RMS error of the raw data at the middle distance is 2.03 decibels, while the RMS error of the fit in Fig.s 2-4 is 2.94 decibels. The estimated parameters are shown in Table 2.

	pixels/degree		
	16	32	64
s_0	51.1	56.17	29.84
f_0	3.68		
k_0	1.728		
r	0.5115		

Table 2. Estimated model parameters.

The parameters f_0, k_0, and r (related to peak frequency, bandwidth, and orientation effects) are equated for all resolutions, while a separate value of s_0 (peak contrast sensitivity) is estimated for each of the three resolutions.

The parameter s_0 is plotted as a function of display visual resolution in Fig 5. Between 64 and 32 pixels/degree, it increases by a factor of 1.88. Between these two resolutions, the basis functions increase in size by a factor of two in each dimension. Thus if sensitivity increased linearly with area (as it should for very small targets [8, 9, 10]) we would expect an increase of a factor of 4. If sensitivity increased due only to spatial probability summation [11, 12], we would expect a factor of about $4^{1/4} = 1.414$. Thus the obtained effect is nearer to that expected of probability summation. At the closest viewing distance, despite a further magnification by 2, the parameter s_0 actually declines. While we would expect a smaller effect of size at the largest sizes, this decline is unexpected and may be due to 1) the relatively poor fit at this resolution, and 2) aspects of visual sensitivity which are not yet captured by the model.

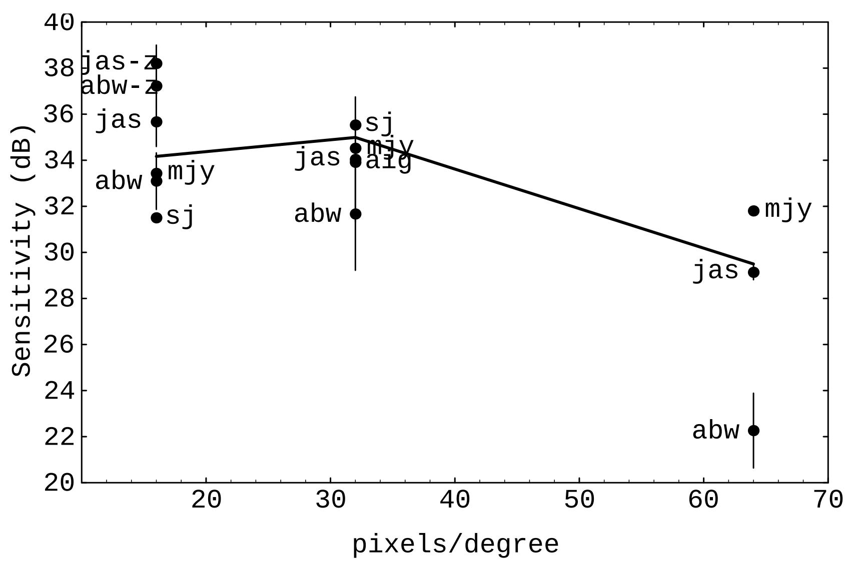

Figure 5. DC basis function sensitivities as a function of display visual resolution. Error bars of plus and minus one standard deviation are shown when multiple measurements were available. For clarity, points with error bars are labeled on the left, those without, on the right. The line indicates the parameter s_0 from Table 2.

DC Sensitivities

Figure 5 also shows the sensitivities for DC basis functions at the three visual resolutions. Ahumada et al. [3, 4]proposed as a working hypothesis that DC sensitivity is given by the peak sensitivity s_0. This prediction is given by the line drawn in Fig. 5. It captures some of the variation in the DC sensitivities, but further data will be needed to adequately test this model. The points in Fig. 5 at a resolution of 16 pixels/degree and labeled with the suffix "-z" were obtained by pixel-replication at the middle viewing distance, rather than use of the near distance. Their enhanced sensitivity suggests that viewing distance per se may have an effect, even when visual resolution is held constant. The substantial variability of DC thresholds at the highest resolution may be due to differences in accommodation between observers, perhaps as a function of age.

IMPACT

These results and the extended model should improve design of quantization matrices designed for different display visual resolutions.

ACKNOWLEDGMENTS

The authors would like to thank Shacolby Jackson and Mark Young for their help in this project. This work supported by NASA RTOP 199-06-12-31.

REFERENCES

1. Peterson, H. A. "DCT basis function visibility in RGB space." Society for Information Display Digest of Technical Papers. Morreale ed. 1992 Society for Information Display. Playa del Rey, CA.

2. Peterson, H. A., H. Peng, J. H. Morgan and W. B. Pennebaker. Quantization of color image components in the DCT domain. Human Vision, Visual Processing, and Digital Display. Proc. SPIE. **1453**: 210-222, 1991.

3. Ahumada, A. J., Jr. and H. A. Peterson. "Luminance-Model-Based DCT Quantization for Color Image Compression." Human Vision, Visual Processing, and Digital Display III. Rogowitz ed. 1992 Proceedings of the SPIE.

4. Peterson, H., A. Ahumada and A. Watson. "An Improved Detection Model for DCT Coefficient Quantization." Human Vision, Visual Processing, and Digital Display IV. Allebach ed. 1993 SPIE. Bellingham, WA.

5. Watson, A. B. "DCT quantization matrices visually optimized for individual images." Human Vision, Visual Processing, and Digital Display IV. Rogowitz ed. 1993 SPIE. Bellingham, WA.

6. Watson, A. B. and D. G. Pelli. QUEST: A Bayesian adaptive psychometric method. Perception and Psychophysics. **33**(2): 113-120, 1983.

7. Watson, A. B. Probability summation over time. Vision Research. **19**: 515-522, 1979.

8. Noorlander, C., M. J. G. Heuts and J. J. Koenderink. Influence of the target size on the detection threshold for luminance and chromaticity contrast. Journal of the Optical Society of America. **70**(9): 1116-1121, 1980 .

9. Graham, C. H., R. H. Brown and F. A. Mote. The relation of size of stimulus and intensity in the human eye: I. Intensity thresholds for white light. J. Exp. Psychol. **24**: 555-573, 1939.

10. Barlow, H. B. Temporal and spatial summation in human vision at different background intensities. Journal of Physiology. **141** : 337-350 , 1958 .

11. Graham, N., J. G. Robson and J. Nachmias. Grating summation in fovea and periphery. Vision Research. **18** : 815-825 , 1978 .

12. Robson, J. G. and N. Graham. Probability summation and regional variation in contrast sensitivity across the visual field. Vision Research. **21**: 409-418, 1981.