Visibility of DCT basis functions:

Effects of display resolution.

Andrew B. Watson   Joshua A. Solomon   Albert J. Ahumada, Jr.

NASA Ames Research Center
Moffett Field, CA 94035-1000
beau@vision.arc.nasa.gov

Introduction

The JPEG, MPEG, and CCITT H.261 image compression standards, and several proposed HDTV schemes, employ the Discrete Cosine Transform (DCT) as a basic mechanism [1, 2]. Typically the 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). The QM is not defined by the standards, but is supplied by the user and stored or transmitted with the compressed images.

The principle that should guide the design of a QM is that it provide optimum visual quality for a given bit rate. QM design thus depends upon the visibility of quantization errors at the various DCT frequencies. In recent papers, Peterson *et al.* [3, 4] have provided measurements of threshold amplitudes for DCT basis functions at one viewing distance and several mean luminances. Ahumada and Peterson [5] have devised a model that generalizes these measurements to other luminances and viewing distances, and Peterson et al. [6] have 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 [7] has shown how this model may be used to optimize the quantization matrix for an individual image.

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.

Plausible Pixel Sizes

Visual resolution of the display (in pixels/degree of visual angle) is determined by display resolution (in pixels/cm) and viewing distance (in cm), according to the formula

$$(\text{pixels/degree}) = (\text{pixels/cm}) / \cot^{-1}[\text{distance}]$$

In the viewing situations for which block-DCT compression is contemplated, there are limits to the practical range of visual resolutions. At the high end, display resolution will be wasted on spatial frequencies which are not visible to the human eye. The limit of human spatial resolution is about 60 cycles/degree. Nyquist sampling of this frequency would require 120 pixels/degree. This corresponds to 300 dpi printing viewed at a distance of about 23 inches. At the low end, the pixel raster becomes visible. In these experiments, we have examined three viewing distances, 16, 32, and 64 pixels/degree, that span a large part of the range of useful viewing distances.

Methods

Detection thresholds for single basis functions were measured by a two-alternative, forced-choice method. Each trial consisted of two time intervals of 0.5 second, within one of which the stimulus appeared. 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 (basis functions were 1/2, 1/4, and 1/8 degree in width). We describe these three viewing distances as yielding effective visual resolutions of 16, 32, and 64 (magnified) pixels/degree.

During presentation, the luminance contrast of the stimulus was a Gaussian function of time, with a duration of 32 frames (0.53 sec) between $e^{-\pi}$ points. The peak contrast on each trial was determined by an adaptive QUEST procedure [8], which converged to the contrast yielding 82% correct. After completion of 64 trials, thresholds were estimated by fitting a Weibull psychometric function [9]. Thresholds are expressed as contrast (peak luminance, less mean luminance, divided by mean luminance), converted to decibel sensitivities ($-20 \log_{10}$[threshold])

To reduce the burden of data collection, 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 constrain our model sufficiently.
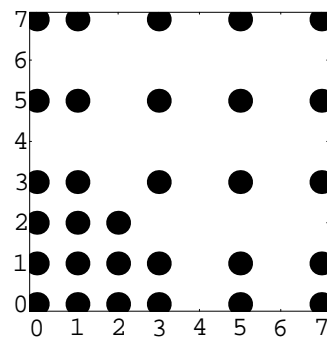

Figure 1. Subset of DCT frequencies used in the experiment.

To date, two data sets have been collected at the low resolution, five at the middle resolution, and one at the highest resolution, as shown in Table 1.

3

resolution (pixels/degree)	observer				
	abw	mjy	aig	sj	jas
16	0	30	0	30	0
32	7	30	60	30	30
64	0	30	0	2	0

Table 1. Thresholds collected for each observer and viewing distance.

Model of DCT Contrast Sensitivity

The model of DCT contrast sensitivity that we consider here is essentially that described by Peterson et al. [6] 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 *s0* (peak sensitivity), *f0* (peak DCT frequency at high luminances), and *k0* (inverse of the *latus rectum* of the parabola), and *r* (the orientation effect).

Results

Figures 2, 3, and 4 show decibel contrast sensitivities for the three viewing distances, along with curves showing the predictions of the best fitting version of the model. Within each figure, the three panels show data for horizontal frequencies {u, 0}, vertical frequencies {0, v}, 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. These plots, and the fits, do not include the thresholds at {0,0} (DC), which are reserved for a separate discussion. The data at 64 pixels/degree also omit 3 thresholds at very high frequencies which we suspect to be artifactual.
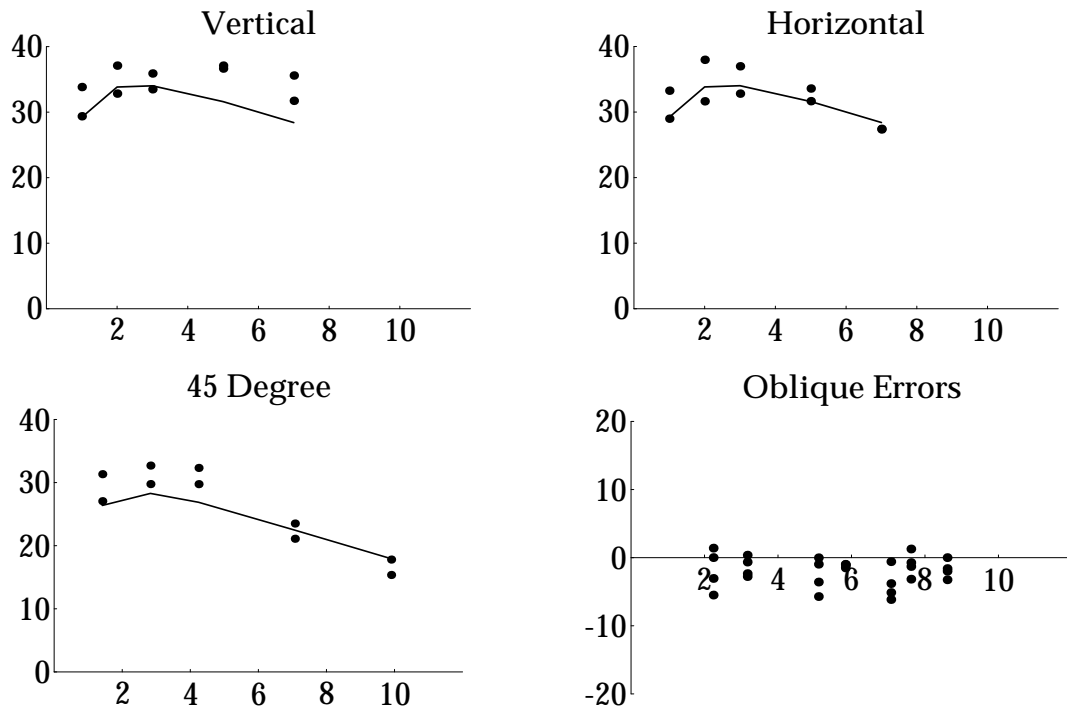
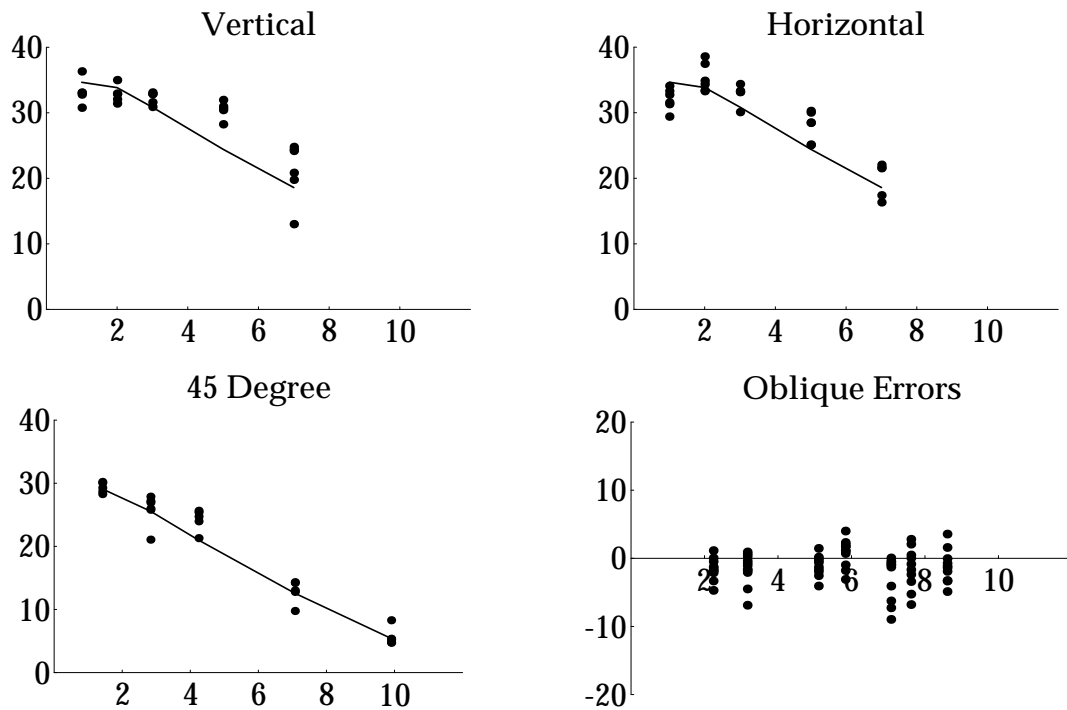Figure 2. DCT basis function sensitivities at 16 pixels/degree.

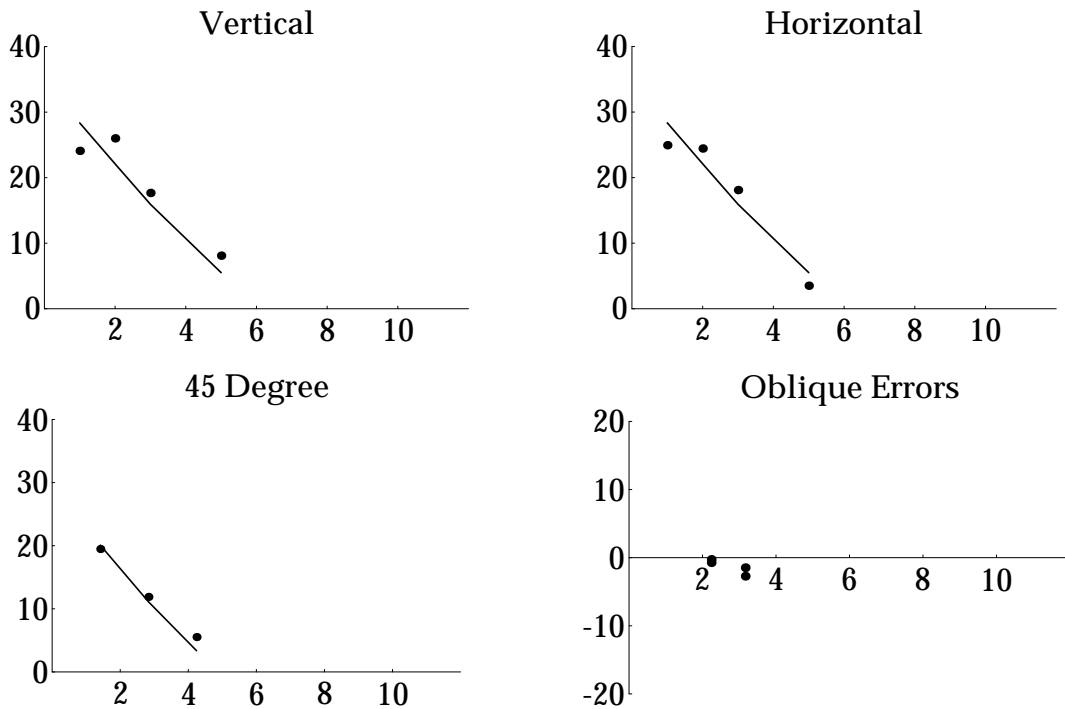Figure 3. DCT basis function sensitivities at 32 pixels/degree.

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

The fits are reasonable, though there appear to be some 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
s0	51.1	56.17	29.84
f0	3.68		
k0	1.728		
r	0.5115		

Table 2. Estimated model parameters.

The parameters *f0, k0, and r* (related to peak frequency, bandwidth, and orientation effects) are equated for all resolutions, while a separate value of *s0* (peak contrast sensitivity) is estimated for each of the three resolutions. The behavior of this parameter is worth considering. 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

6

sensitivity increased linearly with area (as it should for very small targets [10, 11, 12]) we would expect an increase of a factor of 4. If sensitivity increased due only to spatial probability summation [13, 14], 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 *s0* actual 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.

DC Sensitivities

Figure 5 shows the sensitivities for DC basis functions at the three visual resolutions. Ahumada et al. [5, 6]proposed as a working hypothesis that DC sensitivity is given by the peak sensitivity s0. 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.
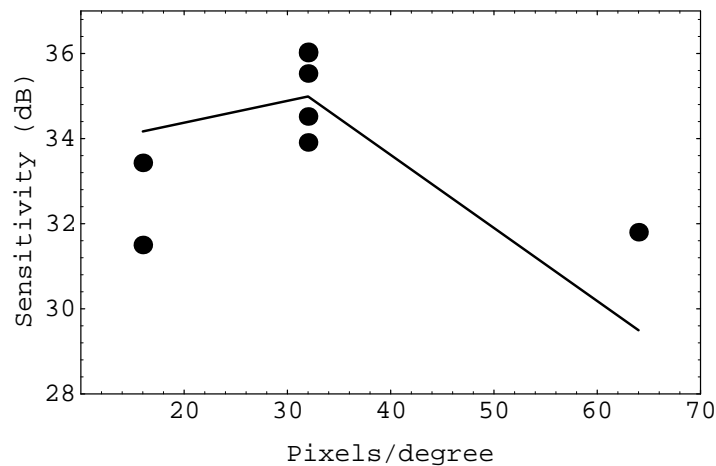

Figure 5. DC basis function sensitivities as a function of display visual resolution.

Discussion

We have examined the variation in visibility of single DCT basis functions as a function of display visual resolution. We have shown that the existing model [5, 6] accommodates resolutions of 16, 32, and 64 pixels/degree, provided that one parameter, the peak sensitivity s0, is allowed to vary. Variations in this

parameter are to some extent consistent with spatial summation, although sensitivity is lower at the lowest resolution than summation would predict.

Practical DCT quantization matrices must take into account both the visibility of single basis functions, and the spatial pooling of artifacts from block to block. Elsewhere we have shown that to a first approximation this pooling is consistent with probability summation[15]. If we consider two images of equivalent size in degrees, but visual resolutions differing by a factor of two, then the sensitivity to individual artifacts would be lower by $4^{1/4}$ in the higher resolution image due to the smaller block size in degrees, but higher by $4^{1/4}$ in the same image due to the greater number of blocks. Thus the same matrix should be used with both. The point of this illustration is that the overall gain of the best quantization matrix must take into account both display resolution and image size.

Acknowledgements

References

1. Pennebaker, W. B. and J. L. Mitchell. "JPEG Still image data compression standard." 1993 Van Nostrand Reinhold. New York.

2. Wallace, G. The JPEG still picture compression standard. Communications of the ACM. 34(4): 30-44, 1991.

3. 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.

4. 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.

5. 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.

6. 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.

7. 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.

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

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

10. 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 .

11. 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.

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

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

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

15. Peterson, H. A., A. J. Ahumada Jr. and A. B. Watson. The Visibility of DCT Quantization Noise. SID Digest of Technical Papers. XXIV: 942-945, 1993.