ORTHONORMAL WAVELETS WITH BALANCED UNCERTAINTY

D. M. Monro, B. E. Bassil and G. J. Dickson

School of Electronic and Electrical Engineering, University of Bath
Claverton Down, BA2 7AY, England

e-mail: D.M.Monro@bath.ac.uk garyd@ee.bath.ac.uk
Fax: +44 1225 826073
Internet: http://dmsun4.bath.ac.uk/

ABSTRACT

This paper addresses the question: 'What makes a good wavelet for image compression?', by considering objective and subjective measurements of quality. A new metric is proposed for the design of the Finite Impulse Response (FIR) filters used in the Discrete Wavelet Transform (DWT). The metric is the diagonal of the Heisenberg uncertainty rectangle, with time weighted by a factor k relative to frequency. Minimization of the metric balances the time and frequency spreads of the filter response. The metric can be computed directly from the filter coefficients, so it can be used to optimize wavelets for image compression without the cost of repeatedly compressing and decompressing images. A psychovisual evaluation carried out with 24 subjects demonstrates that orthonormal FIR filters designed this way give good subjective results with zerotree image compression.With suitably chosen k, both better subjective quality and lower RMS error are achived than with wavelets chosen for maximum regularity.

1. BACKGROUND

The theory of continuous and discrete wavelet transforms [1, 2] has inspired much basic and applied research in signal and image processing, as well as revitalizing the study of sub-band filtering [3, 4, 5]. The Discrete Wavelet Transform (DWT) is obtained by repeated filtering and sub-sampling into two bands with low- and high-pass Finite Impulse Response (FIR) filters, $\{h_{0,}\ldots, h_{L-1}\}$ and $\{g_{0,}\ldots, g_{L-1}\}$ respectively. The inverse process gives perfect reconstruction if the wavelet is orthonormal. This is easily shown to be the case [4] if the filter coefficients satisfy

$$\sum_n h_n\, h_{n+2j} = \delta_{j,0}$$

with $$g_n = (-1)^n\, h_{L-n-1} \qquad (1$$

Wavelets exist only if a regularity condition is satisfied,

$$\sum_n h_n = \sqrt{2} \qquad (2$$

2. HEISENBERG UNCERTAINTY

The lower limit on the time-frequency resolution that can be obtained with wavelet transforms is theoretically [6]

$$\Delta\omega\, \Delta t \geq 1/2 \qquad (3$$

This states that the time and frequency localization of information cannot simultaneously be arbitrarily small. It can be shown that the continuous, infinite time, Gabor wavelet [6] is the only wavelet which achieves this minimum. The lower the Heisenberg uncertainty, the better is the resolving power of the wavelet. The uncertainty is invaraint over the scales used in the DWT; as data is filtered and subsampled, the time resolution is successively halved and the frequency resolution is doubled, preserving their product $\Delta\omega\, \Delta t$.

We can compute the bandwidth $\Delta\omega$ and time dispersion Δt directly for a FIR filter from its coefficients:

$$\Delta\omega^2 = \frac{\pi^2}{3} + 4 \sum_{n=0}^{L-2} \sum_{m=n+1}^{L-1} \frac{(-1)^{m-n}}{P\,(m-n)^2}\, h_m\, h_n$$

where $P = \sum_{n=0}^{L-1} h_n^2 = 1$ in the orthonormal case.

$$\Delta t^2 = \sum_{n=0}^{L-1} (n - \overline{t}\,)^2\, h_n^2$$

where $$\overline{t} = \frac{\sum_n n\, h_n}{\sum_n h_n} \qquad (4$$

3. A PROPOSED QUALITY METRIC

Heisenberg uncertainty is not itself a useful figure of merit for wavelet design. Although the product $\Delta\omega\, \Delta t$ is the area of a rectangle in time/frequency space, it tells us nothing about its aspect ratio. Figure 1 illustrates how three different

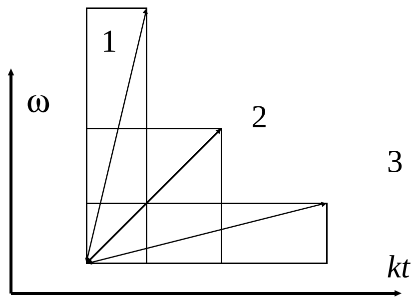


Figure 1. Uncertainty rectangles of three FIR filters with identical $\Delta\omega\,\Delta t$. Filter 2 minimizes the metric $M(k)$ for filters with the same product $\Delta\omega\,\Delta t$.

filters could have the same uncertainty while having vastly different time and frequency resolution. The metric we propose is a balanced uncertainty, which will favour an uncertainty rectangle with the shortest diagonal in a weighted time/frequency space. Consider the function

$$M(k) = \Delta\omega^2 + k^2\,\Delta t^2 \qquad (5$$

The parameter k specifies the relative importance attached to time and frequency resolution, and may be image dependent. To design a particular type of wavelet filter, we choose a value of k and minimize $M(k)$ directly from the FIR filter coefficients.

4. DESIGN EXAMPLES

Although the principle of balanced uncertainty extends to all types of wavelet, we demonstrate it here using orthonormal wavelets, because the number of free parameters is more manageable. In the orthonormal case there are $L/2 - 1$ free parameters in solving for the coefficients h_n in Equations 1 and 2, whereas in the biorthogonal case we would have to consider analysis and synthesis filters of different lengths.

The Haar wavelelet is the only orthonormal solution for L=2, for which $\Delta\omega\,\Delta t = 0.568$. Also, if we minimize Δt for any filter length, we again obtain the Haar wavelet, *i.e*. with two adjacent nonzero coefficients.

For L=4, there is one free parameter, and we can minimize $M(k)$ easily. Minimization with $k = 0$ produced the well-known Daubechies wavelet, OrthD4, which therefore minimizes $\Delta\omega$ without regard to Δt. We obtained a range of other wavelets B4*(k)* by minimizing $M(k)$ for $k^2 = 0, 0.1, 0.2, \ldots, 0.9$. For comparison, we also minimized $\Delta\omega\,\Delta t$, to get MinUncert4.

Similarly, for $L = 6$ there are two free parameters, and we can compare the corresponding filters OrthD6, B6*(k)* which minimizes $M(k)$ for $k^2 = 0, 0.1, 0.2, ..., 0.9$, and MinUncert6. OrthD6 is similar to, but not identical to B6(0).

Name	k^2	$\Delta\omega$	Δt	$\Delta\omega\,\Delta t$	e_{rms}
Haar2	∞	1.136	0.500	0.568	9.89
OrthD4	0	1.033	0.612	0.633	8.95
B4, $k^2 = 0.4$	0.4	1.035	0.597	0.618	8.92
MinUncert4	n.a.	1.103	0.506	0.559	9.56
OrthD6	≈0	0.988	0.650	0.641	8.88
B6, $k^2 = 0.4$	0.4	0.995	0.686	0.682	8.75
MinUncert6	n.a.	0.963	0.635	0.611	9.47

Table 1. Uncertainty parameters and e_{rms} on Gold Hill at 0.2 bpp for some orthonormal FIR wavelets.

Table 1 Gives the uncertainty metrics for the design examples along with the RMS errors obtained when used in zerotree compression of the standard test image Gold Hill to 0.2 bpp (40:1). The version of B6*(k)* given at $k = 0.4$ is the best psychovisual result as determined in the next Section, and we show also the corresponding B4.

5. PSYCHOVISUAL EVALUATION

We evaluated the wavelets subjectively, using our implementation of Shapiro's zerotree coder [7]. A total of 24 nonspecialist volunteers viewed 10 versions of the same image optimized for $k^2 = 0, 0.1, 0.2, \ldots, 0.9$ at the same compression using a particular length of orthonormal filter, and were asked to pick the one of best visual quality. Four images were displayed simultaneously on the screen of a workstation, initially from both ends of the range, with two other images spaced as equally as possible in k^2. After each selection the image farthest away in k^2 was discarded, and a new range was defined by the extremes of the three remaining images. The choice quickly narrowed to a final selection from four images adjacent in the range. The 8bpp Y (luminance) components of Gold Hill, Barbara2 and Boats were used, at 20:1, 40:1 and 60:1. With fourth order wavelets, the compressions were not of sufficient quality to give consistent results. With sixth order orthonormal wavelets, clear preferences emerged.

A one way analysis of variance using the F-test was carried out to determine if the results were consistent at each compression ratio, and the Barbara2 image was excluded at the 95% significance level. A similar analysis of the combined results from Gold Hill and Boats was used to see if the results were consistent across all three compression ratios, and all were accepted at the 95% level. The result over these 144 trials was mean $k^2 = 0.40$, standard deviation 0.24. The histogram of these combined results is shown in Figure 2. The Barbara2 image gave a similar result at 20:1, but lower values of k^2 at 40:1 and 60:1, both near 0.2.

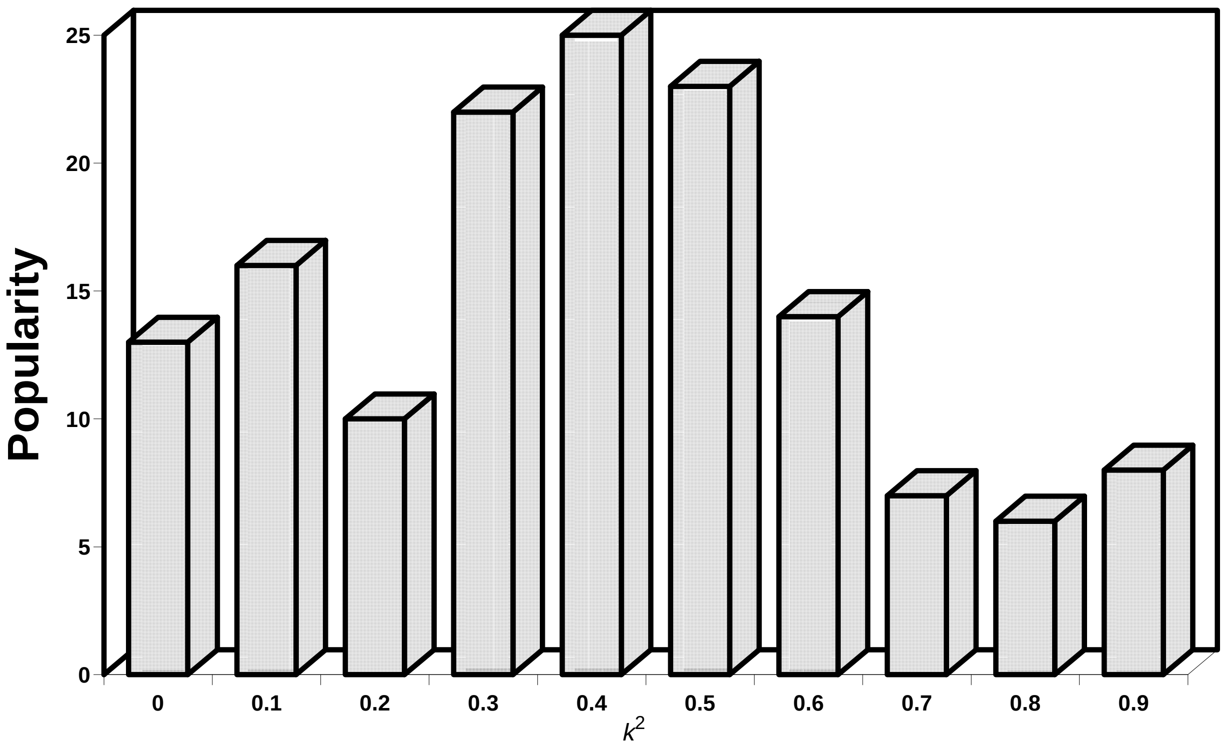


Figure 2. Histogram of 144 subjective preferences, by 24 subjects on 2 images at 3 compression ratios.

6. DISCUSSION AND CONCLUSIONS

The psychovisual evaluation found B6*(k)* with $k^2 = 0.40$ to give good visual results in the examples studied, although for special classes of images, we would expect other values of k to be optimum. For example the Barbara2 image has a large area of smooth skin and subjects preferred slightly larger time spreading at higher compression ratios.

Perhaps the most important comparison is with the Daubechies class of orthonormal wavelets, chosen for ‘maximum regularity’, which is based on placing the maximum number of z-plane zeros of the FIR filter at $z = -1$. This tends to give the filter a flat response and sharp cutoff in the frequency domain, at the expense of ringing of the filter. Our results show this to be a visually suboptimum design principle in selecting wavelets for image compression.

The unweighted uncertainty $\Delta\omega\Delta t$ does not predict degradation, either visually or in terms of MSE. However by computing the diagonal of the k-weighted uncertainty rectangle, $M(k)$, we obtain a useful quality metric for wavelet compression. Although the wavelets designed at $k^2 = 0.4$ have the lowest RMS errors of those shown in Table 1, they are not in general the smallest RMS errors over all k, and we have also obtained other wavelets giving lower RMS errors by annealing over the wavelet coefficients. It is also of practical significance that $M(k)$ can be computed directly from the FIR coefficients, so wavelets can be optimized without repeated cycles of compression and decompression.

Wavelets with a low or zero weighting k^2 minimize the bandwidth $\Delta\omega$ of the FIR filters at the expense of a wider impulse response Δt, which is visible as in ghosting and/or blurriness of the image. Those with a higher value of k^2 are sharper, but the blockiness of the zerotree structure appears. The Haar filter is the blocky extreme of the range, while the orthonormal Daubechies filters are at or near the blurry extreme. Because of the widespread use of the Daubechies filters, it is a significant finding that these are not in general close to the psychovisual optimum.

The wavelet B6*(k)* with $k^2 = 0.4$ gives the best psychovisual quality over a range of compression ratios for two standard test images. Figure 3 shows images for visual comparison of three 6 coefficient filters, in which B6*(k)* with $k^2 = 0.4$ clearly outperforms both OrthD6 and MinUncert6 in the Gold Hill image.

The coefficients of B6*(k)* with $k^2 = 0.4$ are:

$h_0 = 0.51065493$ $h_1 = 0.81006904$ $h_2 = 0.24732487$

$h_3 = -0.13503181$ $h_4 = -0.05087302$ $h_5 = 0.03206956$

7. REFERENCES

[1.] Daubechies, I., ‘Orthonormal bases of compactly supported wavelets’, Comm. Pure Appl. Math., Vol 41, pp 909-966, 1988.

[2] Mallat, S., ‘A theory for multiresolution signal decomposition: The wavelet representation’, IEEE Trans. Pattern Anal. Mach. Intel., Vol. 11, July 1989.

[3] Smith, M. J. T. and Barnwell, T. P., ‘A new filter bank theory for time-frequency representation’, IEEE Trans. Acoust., Speech and Signal Proc., Vol ASSP-35, No. 3, pp 314-327, March 1987.

[4] Antonini, M., Mathieu, P., and Daubechies, I., ‘Image coding using wavelet transform’, IEEE Trans. Image Proc., Vol.1, No. 2, pp 205-220, April 1992.

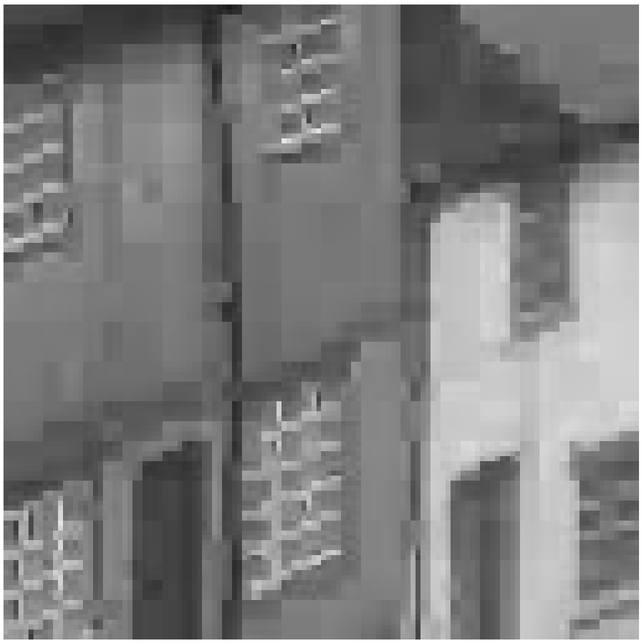
[5] Vetterli, M., and Herley, C., Wavelets and filter banks: Theory and design’, IEEE Trans. Signal Proc., Vol. 40, No. 9, pp.2207-2232, September 1992.

[6] Gabor, D. ‘Theory of Communication’, J. of the IEE, Vol. 93, pp. 429-457, 1946.

[7] Shapiro, J. M., ‘Embedded image coding using zerotrees of wavelet coefficients’, IEEE Trans. Signal Proc., Vol. 41, No. 12, pp. 3445-3462, December 1993.

(b) B6(0.4), Best Psychovisual, minimizes *M*(0.4), e_{rms}= 8.75

(a) OrthD6, maximum regularity, e_{rms}=8.97



(c) Minimum Uncertainty 6, e_{rms} = 9.42

Figure 2. Detail from the luminance (Y) CCITT test image Gold Hill after compression of the whole image at 40:1 (to 0.2 bpp).