EXAMINING THE EFFECTS OF BASIS FUNCTION TRUNCATION IN THE DGT

Jeffrey A Bloom

NEC Research Institute
4 Independence Way
Princeton, NJ 08540

Todd R. Reed

University of California, Davis
Department of Elec. and Comp. Eng.
Davis, CA 95616

ABSTRACT

The Derivative of Gaussian Transform is a spatial/spatial frequency representation of an image that has been used in image and video compression algorithms. Both the forward and inverse transforms require $O(N^3)$ operations. In this paper, we consider truncating the length N basis functions to produce a sparse basis matrix and thus to reduce the computational load of the reconstruction. We examine the effects of truncation on basis function energy and evaluate the change in a number of basis quality measures as the basis function length is decreased.

1. INTRODUCTION

The Derivative of Gaussian Transform (DGT) [1] is a spatial/spatial frequency representation of an image. The basis functions are shifted, separable derivatives of a Gaussian and can be chosen to provide a complete representation of a discrete image. The DGT basis functions can be shown to have very good joint locality compared to Hermite functions and real Gabor functions making the DGT a good tool for use in image compression [2] and video compression [3] algorithms as well as in other tasks requiring local frequency processing. These functions have also been proposed as a model of the processing done in the primary visual pathway of the mammalian visual system [4]. This fact offers further motivation for their use in image processing applications.

As implemented in [1, 2, 3] for square images, both the decomposition and reconstruction are accomplished as the product of three $N \times N$ matrices as follows:

$$C = A^T F A \tag{1}$$
$$F = G^T C G \tag{2}$$

where F is the image, C is the coefficient array, G is a matrix whose columns are the 1D DGT basis functions, and A contains the dual basis defined as

$$A^T = (G^T G)^{-1} G^T. \tag{3}$$

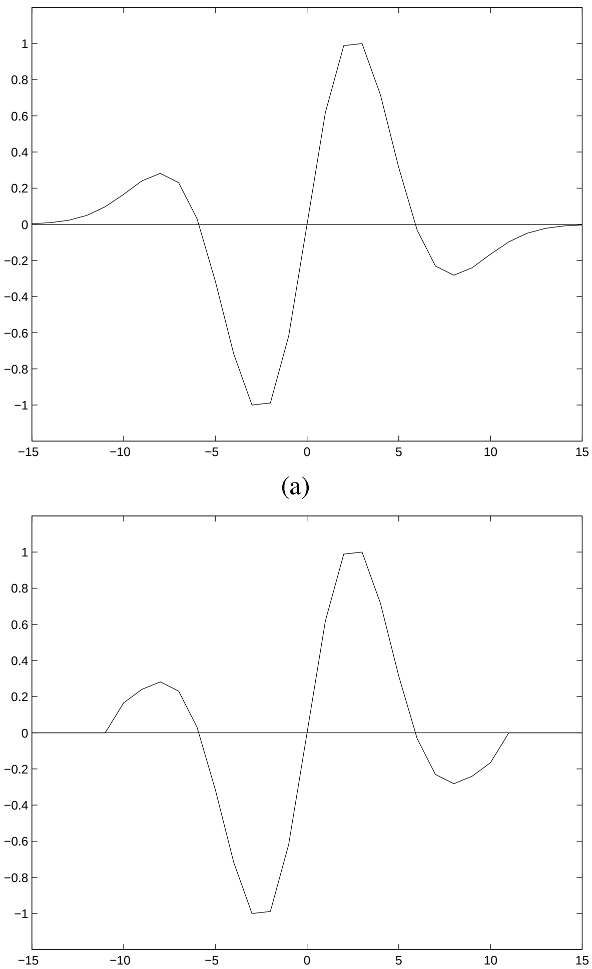
Each of these transforms in Equations 1 and 2 requires approximately $2N^3$ real multiply/add operations. If the basis functions of G have only L non-zero values, then the number of multiply/add operations in Equation 2 can be reduced. Consider the intermediate matrix, $I = G^T C$. Each element of I requires only L multiply/add operations. Since I is $N \times N$, its calculation requires LN^2 multiply/add operations and calculation of F requires a total of $2LN^2$ multiply/add operations. This represents a reduction of N/L multiply/add operations compared to the case where G is fully populated.

The 1D basis functions that make up the columns of the G matrix can be interpreted as Hermite polynomials in a Gaussian envelope, typically with standard deviation < 5 pixels. When N is much larger ($N = 256$ is used in [2] and [3]), the magnitudes of the basis function samples can quickly decay to negligible values. By truncating each of the basis functions some distance from its origin, the amount of reconstruction computation can be reduced with minimal effects on the resulting basis and quality of the reconstructed image. This truncation can be interpreted as windowing with a rectangular window. We will refer to the region inside the window as the region of support.

The basis functions are truncated, prior to shifting, to be length $L = 2t + 1$, where the support region radius, t, is less than $N/2$. The function magnitude is set to zero at all points that are outside the open interval $(-t, +t)$. An example is shown in Figure 1 where the third derivative from a sample basis is shown out to ± 15 pixels without truncation on the top and with truncation to ± 10 pixels ($t = 10$) below. In this paper, we consider only the case that the same support region is applied to all of the basis functions in the set.

2. EFFECTS ON ENERGY

The reduced support basis functions have less energy than their full-size counterparts. The reduction in energy can serve as a measure of the amount of support reduction. It can be shown [5] that the energy in a continuous, reduced



(a) (b)

Fig. 1. Third derivative, $\sigma = 3.4$: (a) without truncation and (b) truncated to ± 10 pixels.

support n^{th} order DGT basis function can be described by

$$E_{n,t} = A_n^2 \left(\frac{1}{\sigma\sqrt{2}}\right)^{2n-1} \int_{-t/\sigma\sqrt{2}}^{t/\sigma\sqrt{2}} H_n^2(x) e^{-2x^2} dx \quad (4)$$

where H_n is a Hermite polynomial of order n, σ is the standard deviation of the Gaussian, A_n is a scale factor used to normalize the basis functions, and t is the support region radius. The A_n necessary such that all the infinite support basis functions have a constant energy, E, can be described by

$$A_n = \sqrt{\frac{E 2^n \sigma^{2n-1}}{\sqrt{\pi}(2n-1)!!}} \quad (5)$$

where the symbol $m!!$, defined for odd integers, is used to specify the product

$$m!! = \prod_{j=0}^{(m-1)/2} (m-2j). \quad (6)$$

The ratio of the energy in a reduced support DGT basis function to that of the infinite support version of the same function is then given by

$$\frac{E_{n,t}}{E} = \frac{\sqrt{2}}{\sqrt{\pi}\,(2n-1)!!} \int_{-t/\sigma\sqrt{2}}^{t/\sigma\sqrt{2}} H_n^2(x) e^{-2x^2} dx. \quad (7)$$

Unfortunately, this expression has an integral that does not have a closed form solution, thus the data presented in this paper is found empirically, by measuring the energies in both the full-size and reduced support discrete functions.

3. EXERIMENTAL RESULTS

The uniform DGT basis used in this study is composed of eight Gaussian derivatives with derivative orders 0, 3, 10, 21, 36, 53, 74, and 99, all derivatives of a Gaussian with standard deviation of 3.4 pixels and length 256 pixels. This basis was chosen because it performs well as measured by the deleted band impulse, step, and constant responses. These quality measures, used to predict the amount of DC leakage and edge artifacts introduced by quantization of the coefficients, are based on work by Kronander [6] in the filterbank setting and by Villasenor *et al* [7] for wavelets.

As the region of support is decreased, the percent of retained energy, the conditioning of the basis matrix, and the four deleted band quality measures are evaluated. The percentage change in energies of the eight basis functions are plotted in Figure 2 for support regions less than ± 25 pixels.

One consideration in designing a DGT basis is that the basis matrix (G in Equation 2) be well conditioned. Poor conditioning on the basis matrix results in instability in the dual basis and thus in the decomposition. Figure 3 shows the condition number associated with the basis matrix as the region of support is decreased.

The four deleted band measures (DC ripple, DCR; step overshoot, SOS; impulse width, IW; and impulse side lobe strength, SLS) are shown in Figure 4, again as percentage changes with respect to the deleted band measures of the full-support basis.

We find that for support window radii greater than 11 pixels, the energy of the highest frequency basis function changes by less than 0.2% and that of the other 7 functions changes by less than 0.1%. This verifies the assumption that prior to application of the window, the basis function magnitudes were essentially zero outside the region of support. The basis matrix remains well conditioned all the way

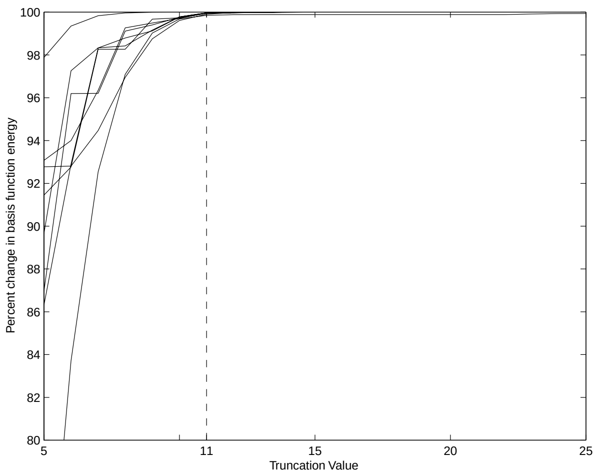


Fig. 2. Effects on basis function energy of limiting the region of support.

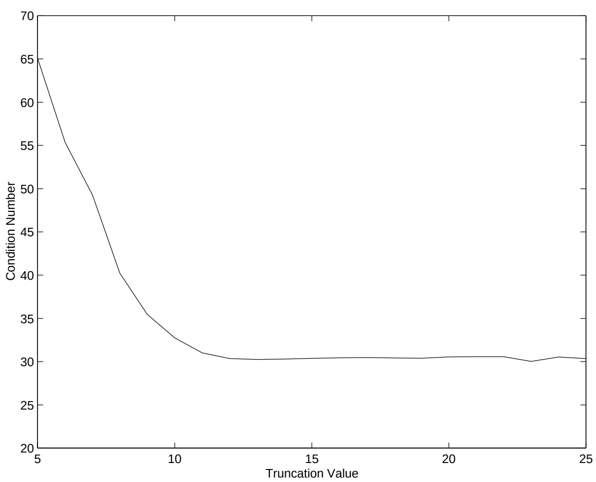


Fig. 3. Effects on condition number of limiting the region of support.

down to $t = 5$ and thus the stability of the decomposition will not be compromised by the limited support. Table 1 shows the truncation values above which the four deleted-band responses deviate from the full support function values by less than 1%.

4. IMAGE COMPRESSION

We now examine the effects of function truncation in an image compression example. Here we consider two cases. In the first compression experiment the basis matrix is composed of truncated functions and the dual basis is generated from this modified basis. This results in a "matched pair" of decomposition and reconstruction matrices. We call this case symmetric analysis/synthesis bases. As long as the ba-

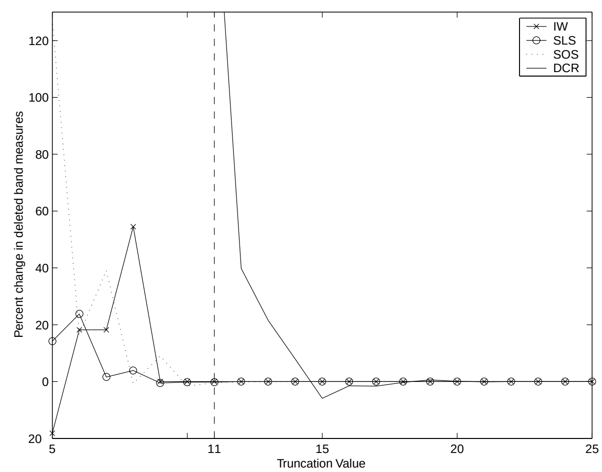


Fig. 4. Effects on deleted band measures of limiting the region of support.

Metric	Support Region Radius
DC Ripple	18
Step Overshoot	11
Impulse Width	9
Impulse Side Lobe Strength	9

Table 1. Minimum support radius providing $< 1\%$ change in deleted band metrics

sis matrix is well conditioned, the reduced support transform is complete and can losslessly represent any image. The variance of the coefficient distribution, however, can change and can be used as an approximation of the entropy in the coefficients. Increases in coefficient variance indicate a decrease in potential compression.

For the 256×256 *cammera man* image, the variance of the DGT coefficients was measured for regions of support from 127 to 5. This variance, plotted in Figure 5, changes slightly as the basis function support is decreased, but it stays within 1% of its full-length value until the support window falls below ± 11 pixels at which point it begins to rise steeply.

A typical rate-scalable compression scheme allows for the delivery of a coded source through a variety of channels of differing bandwidth. In this second compression experiment, we consider the reconstruction of a coded source by a variety of devices with varying computational budgets. The source image is decomposed using the dual basis generated from the full-length functions. A truncated basis set is then used to reconstruct the image from the resulting coefficients. We call this case asymmetric analysis/synthesis bases. As the basis function's region of support is decreased, the error in the reconstruction increases. The reconstruction error, shown in Figure 6(b), is found to be relatively constant and

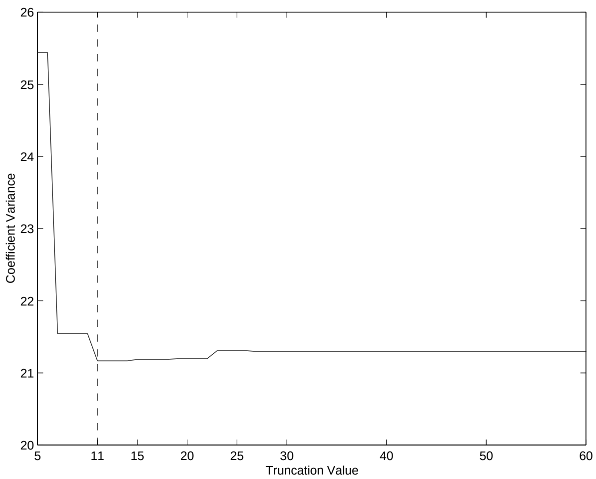


Fig. 5. Effect on coefficient variance of symmetric analysis/synthesis truncation

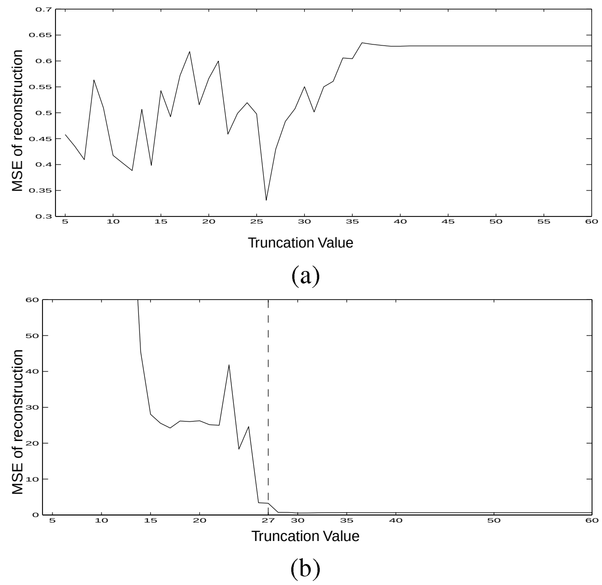


Fig. 6. Effect on reconstruction error of (a) symmetric and (b) asymmetric analysis/synthesis truncation.

less than one for bases with large support. As the support windows drops below ± 27, the MSE begins to rise sharply. This is in contrast to the symmetric analysis/synthesis case shown in Figure 6(a), in which the decomposition was performed with the dual to the truncated synthesis basis. Here, MSE remains essentially constant all the way down to $t = 5$ pixels. The scale of the MSE plot of Figure 6(a) has been expanded to show the slight fluctuation. Notice that while the maximum variation in MSE is 47% of it's nominal value of 0.63, the maximum absolute decrease in MSE is less than 0.3 which, with respect to the errors shown in Figure 6(b), is negligible.

5. CONCLUSIONS

The computational load of the DGT reconstruction can be significantly reduced without measurable impact on the basis function energy, the conditioning of the basis matrix, the deleted-band constant, impulse, or step responses, or the performance in an image compression application. Experiments show that the basis used in this study can be truncated at a distance of ± 18 pixels from the function origin with less than 1% change in DC ripple (the most sensitive of the deleted band quality measures). Using this reduced-support basis requires 7 times fewer multiply/add operations (2×37 multiply/add operations for each element F rather than 2×256). The limiting metric is the deleted-band constant response implying that further reduction in function support will begin to introduce additional ripple in the reconstruction when representation of local DC is required. Local DC is less of a problem in video where the image being represented is a motion compensated difference image. In this case, it may be possible to reduce the region of support to ± 11 pixels, which reduces the computational load by a factor of approximately 11 compared to that required by the full-size function representation.

6. REFERENCES

[1] J. A. Bloom and T. R. Reed, "A Gaussian derivative-based transform," *IEEE Trans. Image Proc.*, vol. 5, pp. 551–553, 1996.

[2] J. A. Bloom and T. R. Reed, "On the compression of still images using the Derivative of Gaussian transform," *Proc. ICIP*, vol. 3, pp. 433–437, 1998.

[3] J. A. Bloom and T. R. Reed, "On the compression of video using the Derivative of Gaussian transform," *Proc. Thirty-second Asilomar Conf. on Signals, Systems, and Computers*, pp. 865–869, 1998.

[4] R. A. Young, "Oh say, can you see? the physiology of vision," *Proc. SPIE*, vol. 1453 Human Vision, Visual Processing, and Digital Display II, pp. 92–123, 1991.

[5] J. A. Bloom, *The Derivative of Gaussian Transform*, Ph.D. thesis, University of California, Davis, 1999.

[6] T. Kronander, "New criteria for optimazation of QMF banks to be used in an image coding system," *Proc. ISCAS*, pp. 1354–1357, 1989.

[7] J. D. Villasenor, B. Belzer, and J. Liao, "Wavelet filter evaluation for image compression," *IEEE Trans. Image Proc.*, vol. 4, pp. 1053–1060, 1995.