A Comparison of Discrete Orthogonal Basis Functions for Image Compression

O. Hunt, and R. Mukundan

University of Canterbury, Dept. Computer Science & Software Engineering.
ojh16@student.canterbury.ac.nz

Abstract

In this report we analyse the image reconstruction accuracy when using different orthogonal basis functions as the kernel for a reversible image transform. In particular we examine the Discrete Cosine Transform(DCT), Discrete Tchebichef Transform(DTT), Haar Transform, and Walsh-Hadamard Transforms(WHT).
We have found that the DCT provides the greatest energy compactness properties for continuous tone images(such as photographs). For images demonstrating rapid gradient variations the Haar Transform performs significantly better than any of the other transform we have analysed, although its performance on continuous tone images is substantially worse than either the DCT or DTT. The WHT performs poorly on either image type.

Keywords: Discrete Orthogonal Functions, Discrete Tchebichef Transform, Image Reconstruction, Image Compression.

1 Introduction

Orthogonal moments have demonstrated many desirable properties in the field of image processing, especially in feature and object recognition, however they also demonstrate some data compaction properties. In this paper we present the results of a study of the data compaction properties of transforms of the form:

$$\tau(u,v) = \sum_{x}^{N-1}\sum_{y}^{N-1} f(x,y)g(x,y,u,v) \qquad (1)$$

Where u and v are coordinates in the transform domain, $f(x,y)$ is a function returning the intensity of an image at coordinates x and y, and $g(x,y,u,v)$ is the kernel function for the transform.

By using different moment sets to provide the kernel, g, it is possible to compare the energy compactness of the resultant data sets. In this study we have used the Discrete Cosine Transform(DCT), The Discrete Tchebichef Transform(DTT), and the Walsh-Hadamard Transform(WHT).

By analysing the Mean-Square-Error(MSE) of the reconstructed image as more components of the transformed data are used, we discovered that the DTT performs very similarly to the DCT, performing only marginally worse on photographic images, and on par when used on images demonstrating sharp variations in gradient(typically present in "vector-art"). While the WHT showed an improvement in it's data compaction properties when processing vector art, it's performance was significantly below that of the DCT and DTT.

2 The Transforms

The transforms used in this study all define their kernel function to be a product of a 2-dimensional function, $g'(i,j)$ such that $g(x,y,u,v) = g'(x,u)g'(y,v)$. In this section we describe each of the transforms, and the kernel, g', that they are based on. In figure 1 we have provided a graphical representation of an 8×8 kernel, as produced by each of the basis functions we have studied.

2.1 Discrete Cosine Transform

The DCT is one one the most well known transforms in image processing, and is used in many different fields, including compression(as the basis for major standards such as JPEG, and MPEG 1 and 2). The kernel for the DCT is derived from the orthonormal Tchebichef polynomials[1], resulting in the following definition of g':

$$g'(x,u) = \lambda(u)\cos\frac{\pi(2x+1)u}{2N} \qquad (2)$$

where

$$\lambda(u) = \begin{cases} \sqrt{\frac{1}{N}}, & u=0 \\ \sqrt{\frac{2}{N}}, & otherwise \end{cases} \qquad (3)$$

Which produces the kernel shown in figure 1(a).

2.2 Discrete Tchebichef Transform

The DTT is a relatively new transform, that uses the Tchebichef moments to provide a basis matrix. As with the DCT the DTT is derived from the orthonormal Tchebichef polynomials, which leads us to presume that it will exhibit similar energy compaction properties[2, 3]. The kernel of the DTT is defined as:

$$g'(x,u) = t_u(x) \tag{4}$$

Where $t_p(x)$ is the Tchebichef the p^{th} order set of the Tchebichef moments. These can be defined using the following function over the discrete range $0..N-1$:

$$t_p(x) = u! \sum_{k=0}^{u} -1^{u-k} \binom{N-1-k}{u-k} \binom{u+k}{u} \binom{x}{k}$$

More frequently we define them using the following recurrence relation:

$$\begin{aligned} t_0(x) &= \frac{1}{\sqrt{N}} \\ t_1(x) &= (2x+1-N)\sqrt{\frac{3}{N(N^2-1)}} \\ t_p(x) &= A_1 x + A_2) t_{p-1}(x) + A_3 t_{p-2}(x) \end{aligned}$$

where $p = 2..N-1$, and,

$$\begin{aligned} A_1 &= \frac{2}{p}\sqrt{\frac{4p^2-1}{N^2-p^2}}, \\ A_2 &= \frac{1-N}{p}\sqrt{\frac{4p^2-1}{N^2-p^2}}, \\ A_3 &= \frac{p-1}{p}\sqrt{\frac{2p+1}{2p-3}}\sqrt{\frac{N^2-(p-1)^2}{N^2-p^2}}. \end{aligned}$$

The kernel produced by the DTT is shown in figure 1(b).

2.3 The Haar Transform

The basis function for the Haar transform is unique among the functions we have studied as it defines what is referred to as a ‘wavelet’. ‘Wavelets’ are a class of functions where a ‘mother’ function is scaled and translated to produce the final set of components. The Haar transform derived from the simple piecewise function:

$$\Psi(x) = \begin{cases} 1, & 0.0 \leq x < 0.5 \\ -1, & 0.5 \leq x < 1.0 \\ 0, & otherwise \end{cases} \tag{5}$$

We then define the transform kernel itself through the following function:

$$\phi_{00}(x) = 1, \qquad x \in [0,1] \tag{6}$$

$$\phi_{pq}(x) = 2^{p/2}\Psi(2^p x - q + 1) \qquad x \in [0,1] \tag{7}$$

The final result of this transform is visible in figure 1(d).

2.4 Walsh-Hadamard Transform

The WHT is the simplest of the transforms we have studied, and is used primarily as a reference point, allowing us to compare the performance of the complex DCT and DTT kernels, to a simpler, more readily computed one. The Walsh-Hadamard is the general name given to either the Hadamard Transform, as the components of the kernel are merely a reordering the Walsh functions. The kernel is defined as:

$$g'(x,u) = \frac{1}{\sqrt{N}} \prod_{i=0}^{n-1} -1^{b_i(x) * b_{n-1-i}(u)} \tag{8}$$

Where $b_i(x)$ returns the i-th bit of x, and $N = 2^n, n \in I$. This leads to the kernel shown in figure **??**.

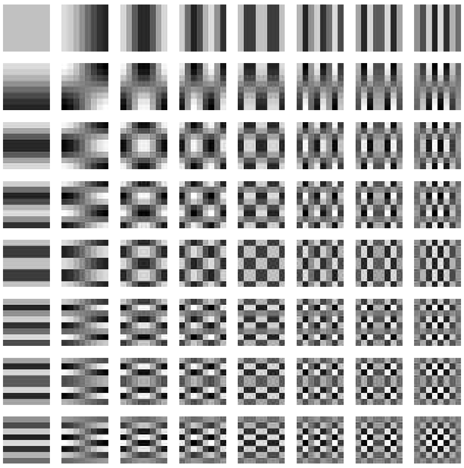
2.5 Notes on the DCT and DTT

The DCT and DTT are both derived from the set of discrete Tchebichef polynomials which leads to them sharing a number of similarities. We can see in section 4 that they share similar performance levels. They also provide kernels that share many similarities, the kernels for an 8x8 DCT and DTT are shown in figures 1(a) and 1(b).

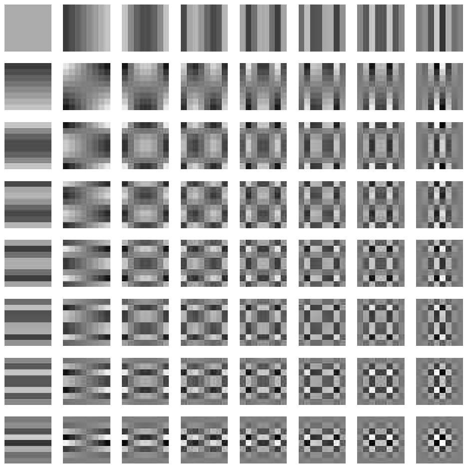
3 Experimental Method

Calculating the energy compactness of a transform can not be calculated directly, and instead is approximated by analysing the reconstruction error of the transform when the number of components used for reconstruction is restricted to less than the total number available. To measure this we need some measure to represent the reconstruction accuracy. For this we used standard Peak Signal to Noise Ratios(PSNR), calculated for the difference between the input and output images for each transform.

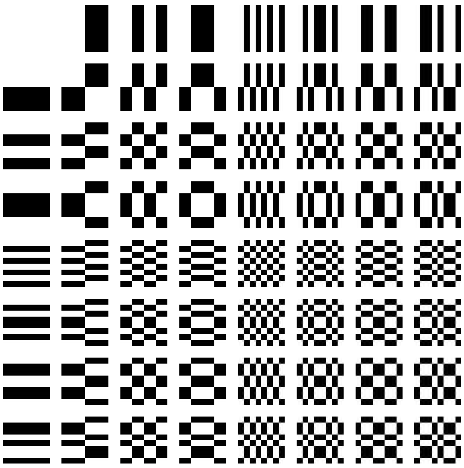
We must also choose the order in which to include components, as the low order components have a much greater effect on the reconstructed image the the high order components do. For this we use the standard ‘zig-zag’ pattern used by the JPEG



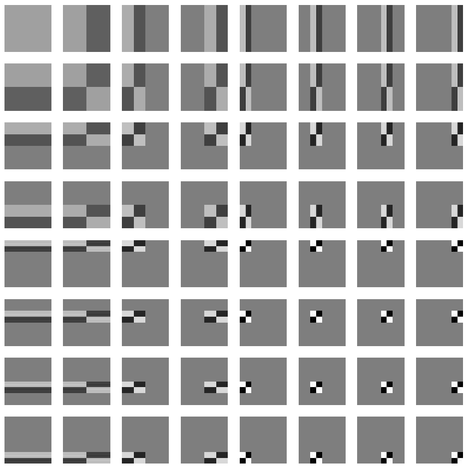
(a) Discrete Cosine Transform



(b) Discrete Tchebichef Transform



(c) Walsh-Hadamard Transform



(d) Haar Transform

Figure 1: Graphical representation of each of the basis functions we have studied.

standard(see Figure 2) which sorts the transform so the the lowest order components are processed first, and the highest order moments last.

Transforms such as as these are typically performed on separate blocks of the image sequentially, rather than over the entire image at once. In our experiments we performed the transforms on 16x16 pixel blocks, meaning a maximum of 256 components are needed to accurately reconstruct each block. All computation was performed using double precision floating point values.

4 Experimental Results

In this paper we include only the results for the common 'Lenna' image, and a simple image with rapid intensity variation, 'ruler', these are shown in figure 3. A more in depth analysis of the transforms we have described is currently being undertaken.

To compare the performance of each transform we measured the reconstruction error, as we increase the number of components used for the reconstruction of each block. The results are shown in figures 4 and 5. These results show that there is little difference between the DCT and DTT, and the WHT provides significantly worse energy compactness than any of the other transforms. The results also show that the Haar Transform is extremely biased towards images exhibiting rapid gradient changes. Figure 5 shows that both the DCT and DTT provide much poorer performance on the sharp changes of the 'Stripes' image, in this image the DTT outperforms the DCT, the WHT provides comparable performance

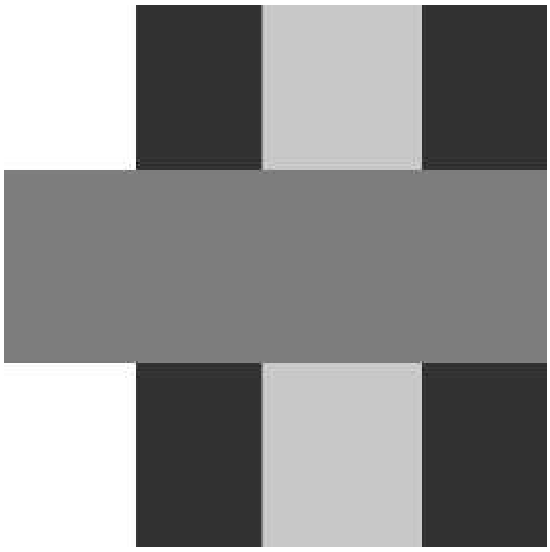


Figure 3: The ‘Lenna’ and ‘Stripes’ images.

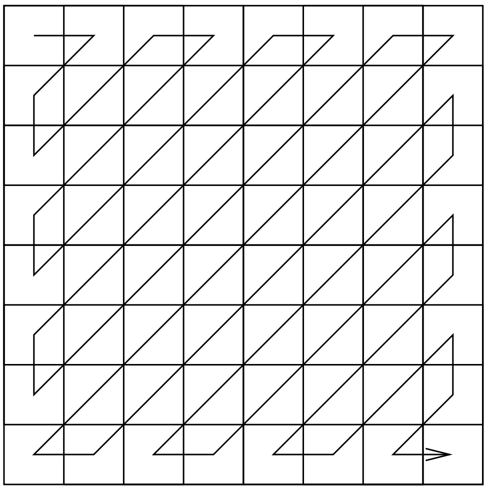


Figure 2: The ‘zig-zag’ pattern used to define the order of components in an 8 by 8 pixel transform.

till approximately half of the components have been used, at which point it starts to fall behind dramatically. This implies that the low order moments of the WHT are less dependant on the image gradients than those of either the DCT or DTT. In this case the Haar Transform significantly outperforms all the other transforms that we have studied.

As no quantisation is performed there is no error in the fully reconstructed image, so PSNR becomes infinite, in order to minimise the possibility of this artifact hiding useful information we have not included the effect of the last component being added to the reconstructed image.

5 Conclusion

Our results show that the type of image being compressed has a significant effect on the performance of the transforms being used. For images exhibiting rapid gradient variations the Haar Transform is clearly superior to all the other transforms, however it is less effective on continuous tone images. For such images either the DCT or DTT could be used, as they exhibit very similar levels of performance. For continuous tone images the DCT is slightly more effective than the DTT, although for images exhibiting rapid gradient changes the reverse is also true.

The performance of the WHT is significantly worse than any of the other transforms we have studied. In continuous tone images the level of performance is continually lower, for images exhibiting rapid gradient changes the performance level is erratic at best, and is still largely below that of the other transforms.

It should be noted that these results however do not guarantee any immediate compression, as it would still be necessary to quantise the output from the transforms, either through immediate culling of components as we have done, or through some other more complex mechanism such as the quantisations schemes used in the JPEG image compression standard[4].

References

[1] N. Ahmed, T. Neterajan, and K. R. Rao, "Discrete cosine transform," *IEEE Trans. on Computers*, vol. 23, pp. 90–93, 1974.

[2] R. Mukundan, "Improving image reconstruction accuracy using discrete orthonormal moments," in *CISST 03 International Conference*, pp. 287–293, 2003.

[3] R. Mukundan, S. H. Ong, and P. A. Lee, "Image analysis by tchebichef moments," *IEEE Transactions on Image Processing*, vol. 10, no. 9, pp. 1357–1364, 2001.

[4] G. K. Wallace, "The jpeg still picture compression standard," *Commun. ACM*, vol. 34, no. 4, pp. 30–44, 1991.

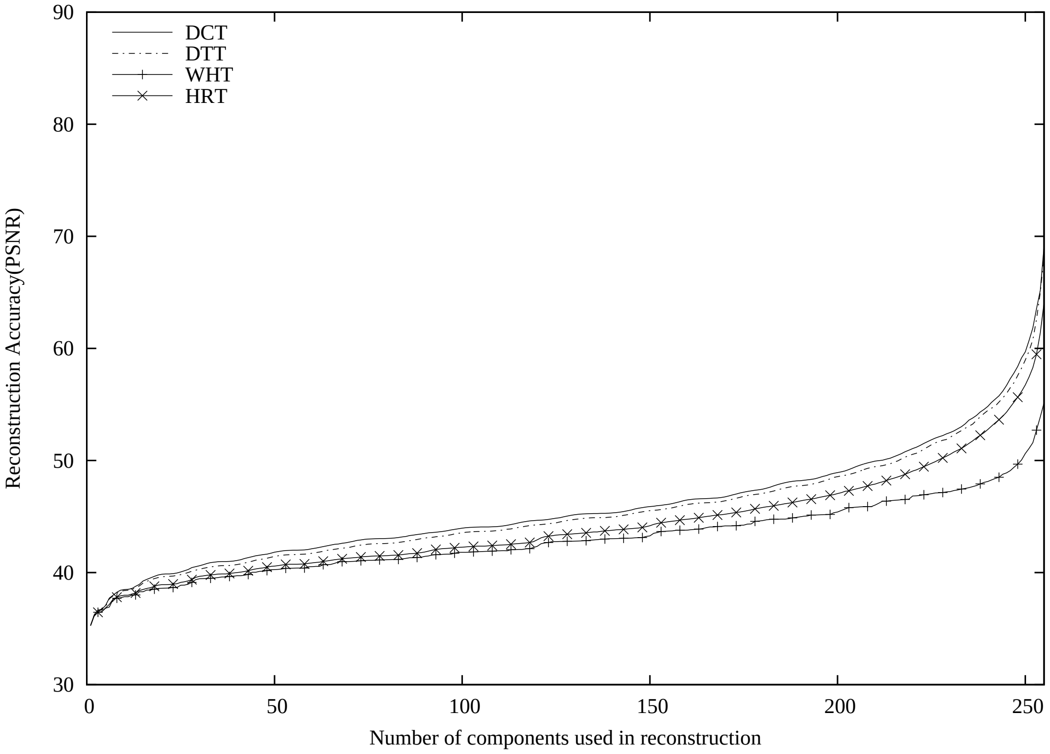


Figure 4: Reconstruction error for the ‘Lenna’ image.

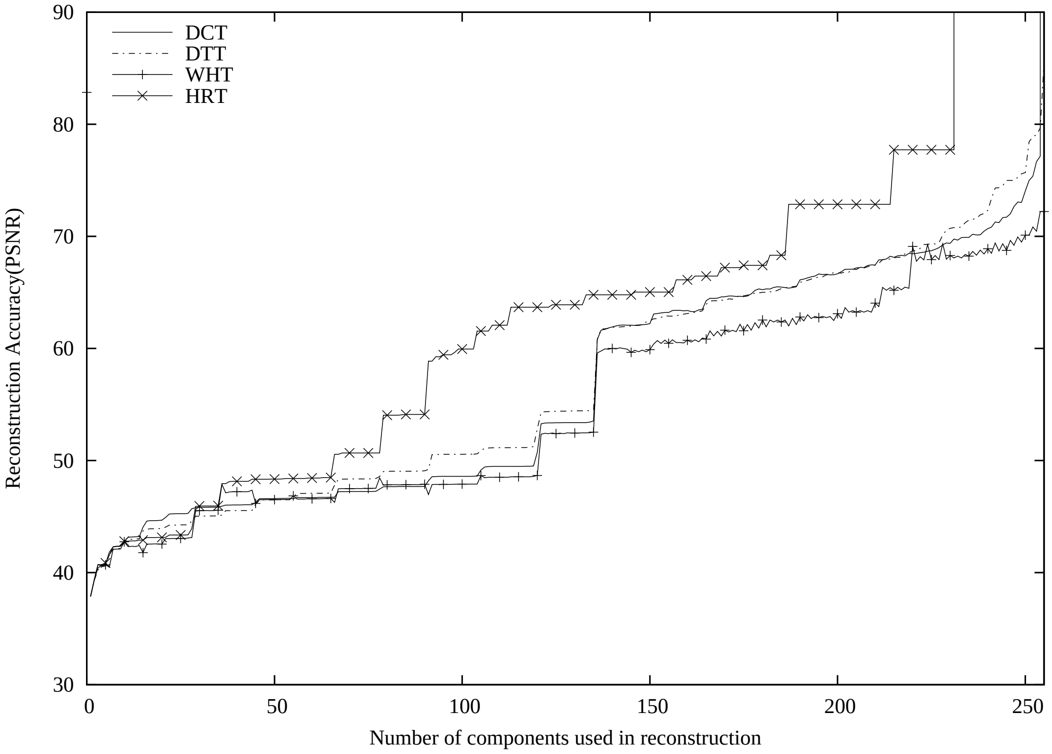


Figure 5: Reconstruction error for the ‘Stripes’ image.