# Cosine Transform Priors for Enhanced Decoding of Compressed Images.

Amos Storkey and Michael Allan

School of Informatics
University of Edinburgh
5 Forrest Hill, Edinburgh, EH1 2QL

**Abstract.** Image compression methods such as JPEG use quantisation of discrete cosine transform (DCT) coefficients of image blocks to produce lossy compression. During decoding, an inverse DCT of the quantised values is used to obtain the lossy image. These methods suffer from blocky effects from the region boundaries, and can produce poor representations of regions containing sharp edges. Such problems can be obvious artefacts in compressed images but also cause significant problems for many super-resolution algorithms. Prior information about the DCT coefficients of an image and the continuity between image blocks can be used to improve the decoding using the same compressed image information. This paper analyses empirical priors for DCT coefficients, and shows how they can be combined with block edge contiguity information to produce decoding methods which reduce the blockiness of images. We show that the use of DCT priors is generic can be useful in many other circumstances.

## 1 Introduction

A number of image compression methods, most notably including baseline JPEG (joint photographic experts group), use quantisation of the discrete cosine transform (DCT) coefficients in order to obtain a lossy compressed representation of an image. Put simply, baseline JPEG splits each image into 8x8 blocks and then performs a DCT on each image block. These are then quantised according to a preset quantisation schema which depends on the compression rate required. The quantised coefficients are then losslessly compressed and encoded to a bitstream, usually using Huffman codes. To decode the jpeg, the quantised coefficients are the obtained from the bitstream using the relevant lossless decompression. The quantised coefficients are then used directly in the inverse DCT to recreate the image.

The deficits of this scheme are that it can produce blocky artefacts [1] as each 8x8 block is treated independently, and that it can produce poor representation of regions with significant high frequency information.

In this paper we recognise the fact that the quantised DCT coefficients provide upper and lower bounds for the true coefficients. It is also possible to obtain empirical prior DCT coefficient distributions from the examination of many other 8x8 patches from a database of uncompressed images. Furthermore we can examine the pixel differences across block boundaries in the uncompressed image database and use that information as a prior measure for the blockiness effects of compressed images.

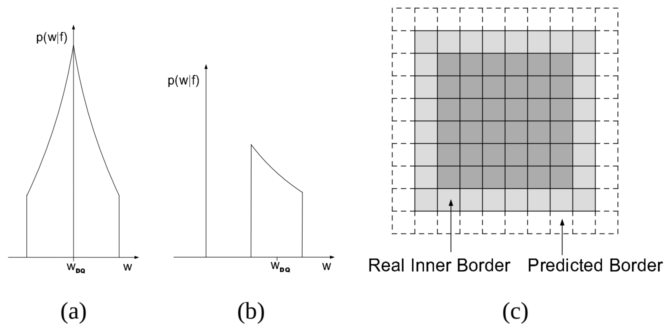



**Fig. 1.** Prior distribution of DCT coefficients given known quantised coefficients $\mathbf{f}$. (a) Zero-quantised coefficient. (b) Non-zero quantised coefficient. (c) Predicted border-pixels are given by inner border-pixels.

## 2 The Probabilistic Model

Consider first a single 8x8 image patch, and the set of pixels bordering each edge of that 8x8 patch. Let $\mathbf{w}$ denote the vector of DCT coefficients, $\mathbf{v}$ denote the set of quantised coefficients (and implicitly the constraints they give to the real coefficients). Let $\mathbf{y}$ denote the border-pixel intensities. Suppose we have some prior distribution over DCT coefficients $P(\mathbf{w})$ and are given the true border-pixel intensities $\mathbf{y}$, and quantised DCT coefficients. Then we wish to calculate the posterior distribution

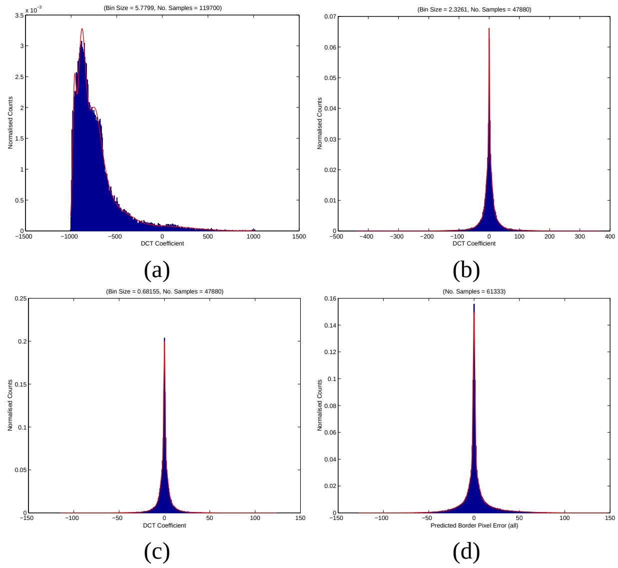
$$P(\mathbf{w}|\mathbf{y}, \mathbf{v}) \propto P(\mathbf{y}|\mathbf{w})P(\mathbf{w}|\mathbf{v}) \tag{1}$$

Here $P(\mathbf{w}|\mathbf{v})$ is the prior distribution $P(\mathbf{w})$ constrained to the region implied by the quantised coefficients $\mathbf{v}$, and renormalised. This is illustrated in Figure 1. $P(\mathbf{y}|\mathbf{w})$ is given by the model $P(\mathbf{y}|\mathbf{x})$ of the observed border pixels given the predicted border pixels $\mathbf{x}$ produced by the extrapolating the DCT basis functions 1 pixel over the boundary of the 8x8 region. This is simple to calculate using the basis function symmetry, and amounts to using the inner border pixels as the predicted border pixels (see figure 1).

## 3 Prior distributions

To use the above model two prior distributions are needed. First a prior distribution $P(\mathbf{w})$ for the DCT coefficients is required. Second we need the distribution of the true border pixels $P(\mathbf{y}|\mathbf{x})$ given the predicted values. Forms for both of these can be obtained empirically from a set of training images.

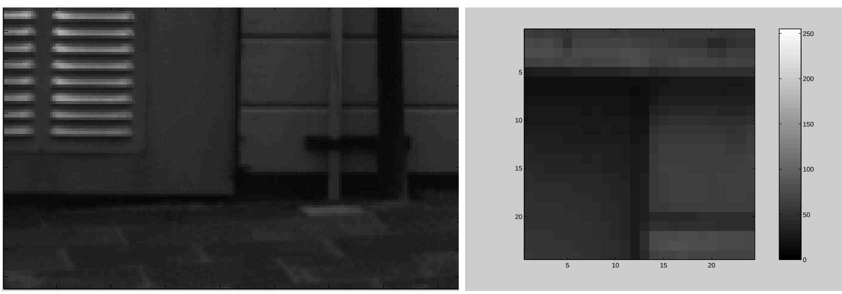
The source images used in this report can be found at http://www.hlab.phys.rug.nl/archive.html. The images were taken with a Kodak DCS420 digital camera. For details and a description of the calibration see the Methods section of [7]. The images contain many similar scenes of urban areas, landscapes and woodland areas. The linear intensity images have been used rather than the de-blurred set in order to reduce the possibility of artefact introduction. The camera used to capture the images does not deliver the the two outermost pixels along the edges and so it was necessary to remove them. To maintain an image that contains a whole number of 8x8 blocks, the eight outermost pixels along the edges were cropped. Also to reduce processing time the images were reduced in size by a factor of two along both the width and height using pixel averaging. The 12-bit images were intensity-scaled linearly to between 0 and 255 to comply with the JPEG standard. The first twenty images in the set were used as training images. The



(a) (b) (c) (d)

**Fig. 2.** Distribution of DCT coefficients: (a) The DC coefficient, (b) The (2,2) coefficient, (c) the (4,6) coefficient. (d) Distribution of differences between the predicted and true border pixels

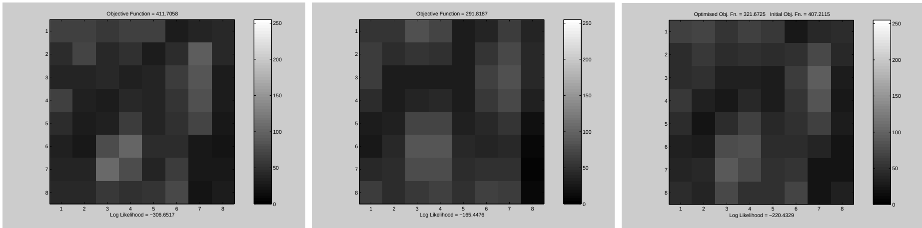
prior over discrete cosine transform coefficients was modelled as a factorised distribution $P(\mathbf{w}) = \prod_i P(w_i)$ where $w_i$ is the $i$th DCT coefficient. Then the prior for each coefficient was set using empirical values the training images. Histograms of the priors are given in Figure 2. Note that the lowest frequency coefficient (commonly called the DC coefficient) has a different structure from the other (AC) coefficients. The AC coefficients appear to have the same form, but have different distribution widths, where the higher frequency components are more tightly distributed around zero. The prior over



Original door scene — Original 3x3 block scene

**Fig. 3.** Two original scenes

the border pixels was also factorised into independent distributions for each pixel. The distribution of the difference between the predicted pixel value and the true pixel value was used as the model for $P(\mathbf{y}|\mathbf{w})$. The distribution obtained is illustrated in Figure 2d. More general forms for $P(\mathbf{y}|\mathbf{w})$ were tried, but they had little effect on the final outcome. Gaussian mixture models were fit to these histograms to provide a working functional representation of the distribution.



(a) (b) (c)

**Fig. 4.** Results of optimisation. (a) Original patch (b) Patch reconstructed from quantised DCT, (c) reconstructed from optimised DCT

## 4 Intermediate results

Although the form of model described is not the final model, as the true border pixel information is not available, it is instructive to see how it performs. Conjugate gradient optimisation was used to calculate the maximum a posteriori (MAP) DCT coefficients given the border pixel data and the quantised DCT coefficients. To test how well the original values can be recovered if a nearby local minima is found, the coefficients were initialised at the original values. For the image patch illustrated in figure 3b, and with a standard lossy decoding given by Figure 4a, we obtain Figure 4b using this approach. A representation of the true, quantised and optimised DCT coefficients are given in Figure 5. We can also initialise the values at the quantised coefficients. Figure 6 shows the use of this approach on a test image.

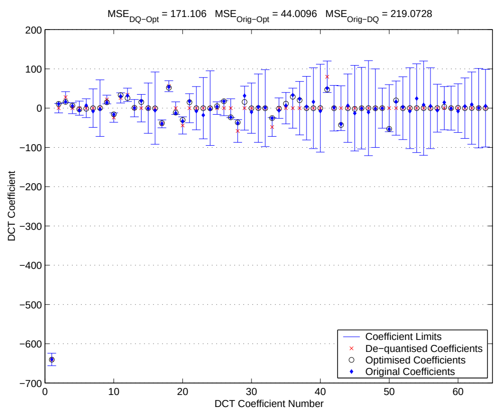



**Fig. 5.** Comparison of the different DCT coefficients.

Quantifiable assessment of image quality is notoriously hard and generally an unreliable measure of performance. However it is possible to use a perceptual error measure such as those of [4, 8, 5]. For example, with the image illustrated the perceptual error measure of [8] improves from $8.46$ to $8.39$. In general we find that for regions with fewer patches containing clear edges or linear features, most error measures (mean square error, signal to noise ratio, peak signal to noise ratio, perceptual error measure) improve, whereas there is loss in systems where edges occur. This is due to the fact that the current factorised prior does not contain edges and features as part of its model. For fuller comparisons with other techniques see [2].

Reconstructed using quantised coefficients — Reconstructed using optimised coefficients

**Fig. 6.** Results on door scene from 3

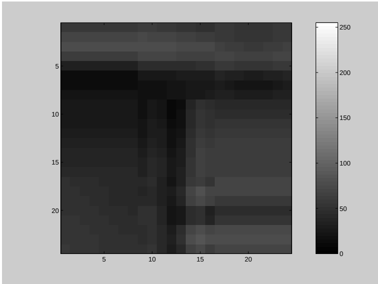
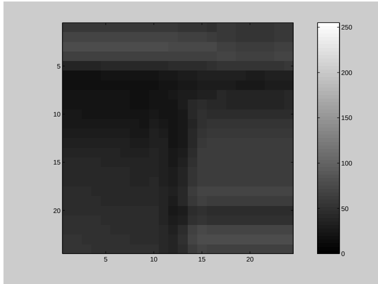
## 5 Full model

As the true pixel data is not available for the border pixels, the model just described is not directly useful. However it can be incorporated into a Markov Random Field model which can then be optimised using an iterative checkerboard approach. The simplest model is to use a model of the form

$$P(X) \propto \prod_i P(X_i) \prod_{(i,j)\in A,k} P(X_i^k - X_j^k) \tag{2}$$

where $X_i$ are the pixel values of the $i$th image patch and $X_i^s$ is the $k$th border pixel of the $i$th patch, and $A$ is the set of indices $(a, b)$ of adjacent image patches. Note that $P(X_i)$ and $P(X_i^k - X_j^k)$ are now clique potentials and not probability distributions. This model can be optimised using an iterative checkerboard scheme. First the coefficients of the 'black' coloured patches are fixed and $\prod_i P(X_i) \prod_{i\in W, j\in A(i),k} P(X_i^k - X_j^k)$ is optimised, for $W$ denoting the indices of the 'white' patches and $A(i)$ the adjacent patches to patch $i$. Then the white patches are fixed to their optimised values and the corresponding equation for the black patches is optimised. This process is repeated until a suitable convergence criterion is satisfied. Due to the symmetry of the $P(X_i^k - X_j^k)$, each step is guaranteed to increase the global $P(X)$. Note each $P(X_i)$ is implicitly given by the prior over DCT coefficients and the quantised coefficients. Again the optimisation was initialised at the known quantised values.

This approach was used on the patch illustrated in figure 7. The blocky artefacts of the quantised coefficients are significantly reduced.

Reconstructed using quantised coefficients. Reconstructed using optimised coefficients.

**Fig. 7.** Results of checkerboard iteration on 3x3 block scene

## 6 Continued work

Blockiness artefacts are only one of the problems of compressed images. More significant issues arise from within block edges. However the process of allowing freedom to choose the coefficients from within the range of values given by the quantised coefficients can be used in conjunction with any other information about the image patches. This might include more general patch priors than those given by factorised coefficients. We are currently working on a fuller quantitative analysis of these techniques.

One area of interest is that the artefacts of jpeg are often serious problems in super-resolution. [3] provides a good example of this. As this work places jpeg modelling within a probabilistic framework, we are working on ways to combine it directly with a number of super-resolution methods such as [6].


**Acknowledgements**

Amos Storkey thanks Microsoft Research, Cambridge for his fellowship.


## References


1. Albert J. Ahumada and Rensheng Horng. De-blocking dct compressed images. In *Human Vision, Visual Proc. and Digital Disp. V, SPIE Proc.*, volume 2179, pages 109–116, 1994.
2. M. Allan. Probabilistic methods for improving image decoding. Master's thesis, Division of Informatics, University of Edinburgh, 5 Forrest Hill, Edinburgh, EH1 2QL, UK, 2003.
3. W. T. Freeman, T. R. Jones, and E. C. Pasztor. Example-based super-resolution. *IEEE Computer Graphics and Applications*, 2002.
4. John Lazzaro and John Wawrzynek. Jpeg quality transcoding using neural networks trained with a perceptual error measure. In *Neural Computation*, volume 11, pages 267–296, 1999.
5. M. Miyahara, K. Kotani, and V.R. Algazi. Objective picture quality scale (pqs) for image coding. In *International Conference on Image Processing*, volume 3, 1995.
6. A. J. Storkey. Dynamic structure super-resolution. In S. Thrun S. Becker and K. Obermayer, editors, *Advances in Neural Information Processing Systems 15 (NIPS2002)*, pages 1295–1302. MIT Press, 2003.
7. J.H. van Hateren and A. van der Schaaf. Independent component filters of natural images compared with simple cells in primary visual cortex. In *R.Soc.Lon.B*, pages 359–366, 1998.
8. Andrew B. Watson. Dctune: A technique for visual optimisation of dct quantisation matrices for individual images. In *Society for Information Display Digest of Technical Papers*, volume XXIV, pages 946–949, 1993.