## Title
A Progressive Analysis Based Perceptually Transparent Coder for Still Images



## Authors
Algazi, Ralph
Ford, Gary
Estes, Robert R.
et al.

## Publication Date
1996

Peer reviewed

# A progressive analysis based perceptually transparent coder for still images

V. R. Algazi G. E. Ford R. R. Estes Jr. A. El-Fallah

Center for Image Processing and Integrated Computing (CIPIC)
University of California, Davis

## ABSTRACT

The encoding of images at high quality is important in a number of applications. We have developed an approach to coding that produces no visible degradation and that we denote as *perceptually transparent*. Such a technique achieves a modest compression, but still significantly higher than error free codes.

Maintaining image quality is not important in the early stages of a progressive scheme, when only a reduced resolution preview is needed. In this paper, we describe a new method for the progressive transmission of high quality still images, that efficiently uses the lower resolution images in the encoding process. Analysis based interpolation is used to estimate the higher resolution image, and reduces the incremental information transmitted at each step.

This methodology for high quality image compression is also aimed at obtaining a compressed image of higher perceived quality that the original.

**Keywords:** Image coding, data compression, high quality imaging, noise removal, perceptual coding.

## 1 INTRODUCTION

Most image coding techniques are directed at the efficient digital representation of original images of moderate quality. The quality of the original image serves as an implicit measure of the additional tolerable distortion that the coder may introduce. For very high quality and super high definition images, a more suitable goal is to allow coding errors, but only if they are imperceptible. We refer to such an approach as *perceptually transparent coding*.

In perceptually transparent coding, our primary goal is to control image quality. Secondarily, we wish to compress the image. Quantization errors are difficult to control in transform domain techniques because the they are distributed over the entire transform block and interact in a complex manner. As such, we restrict our attention to spatial domain processing, where we have sufficient control over the errors introduced.

In a differential quantization based approach, we use a good low frequency approximation of the image which allows us to coarsely, non uniformly quantize the remainder, exploiting properties of the human visual system, such that the quantization errors are invisible in the reconstructed output image. The resulting representation is then compressed using error-free codes. This approach has several beneficial characteristics: in addition to resulting in coders with substantially better performance than error free coding at a comparable quality, by considering first the changes that can be made to images that do not produce perceptible distortion, we identify image parameters or characteristics that are onerous for the encoder, but that are perceptually unimportant.

We further note that significant redundancy exists in images in the form of quantization noise introduced

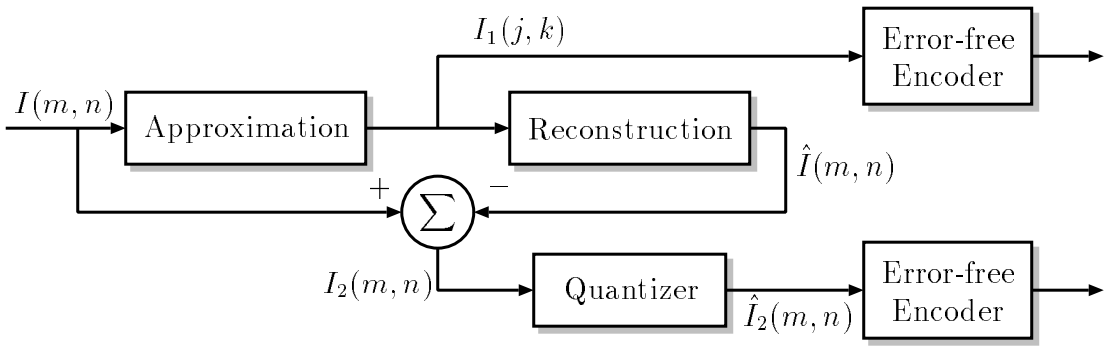


Figure 1: Differential Quantization

by the sampling process. As such, we preprocess the image using newly developed anisotropic diffusion based techniques that can remove this noise without introducing any perceptible artifacts. In many cases such processing can improve image quality, while simultaneously increasing the compressibility (often a conflicting goal.) Here, we limit the amount of preprocessing to maintain perceptual transparency, but expect the simplified images to be easier to predict and, therefore, more compressible.

Finally, we consider the error-free encoding of the non uniformly quantized remainder image. This image has characteristics similar to an edge map, and is high in structure. As such, we propose to use analysis based techniques to encode it, that is, the encoding process is driven by (an analysis of) the data. We consider, primarily, a pyramidal representation for the remainder image. In this progressive representation, higher resolution images are estimated from the lower resolution images using analysis based interpolation strategies. At each stage, only the error images must be transmitted.

## 2 DIFFERENTIAL QUANTIZATION[3,4]

There are three properties of human visual perception which can be used to achieve high image quality while reducing the information content, or bit rate. These are the nonlinear perception of luminance according to Weber's law, the very substantial decrease of contrast sensitivity for spatial-frequencies above 8 cycles/degree, and the visual masking of perturbations or errors by the activity in the image. Most common visual artifacts encountered in images encoded by current techniques occur in the vicinity of high contrast edges, or near the transition between image regions, and are caused by inadequate control of the spatial distribution of errors. Artifacts with a spatial structure, such as contouring, or the end of block effect in transform coders and vector quantization (VQ) are quite perceptible and highly objectionable. Differential quantization circumvents this problem by providing an excellent approximation in the flat portions of an image. The approximation is not as good near edges or in active areas, but in these regions, visual masking, which extends over several minutes of solid angle, allows for substantial errors to occur below the visual threshold of perception. A diagram of the differential quantization approach which exploits these visual properties is shown in Figure 1.

In this scheme, we compute a low frequency approximation to the original image, and then exploit the properties of this approximation by coarsely quantizing the difference between this approximation and the original image. That is, given an original $M \times N$ image, $I(m,n)$, we compute an approximation, $\hat{I}(m,n)$, which can be reconstructed from a smaller, subsampled image, $I_1(j,k)$.* We then coarsely quantize the difference $I_2(m,n) = I(m,n) - \hat{I}(m,n)$ to obtain the quantized remainder image, $\hat{I}_2(m,n)$, which we use as an alternate representation, $\{I_1, \hat{I}_2\}$, of the original image.

In this paper, we use a spline based low frequency approximation, computed from an $8 \times 8$ subsampled version

*In general, the approximation does not have to be in the form of an image, but typically, it is.

of the original, which, when coupled with a 45 level non uniform quantizer results in images which are perceptually indistinguishable from the original.

## 2.1 Spline based approximation

Consider the image to be a surface in 3 dimensional space. We sample the input, $M \times N$ image, $I(m,n)$, on a rectangular grid to obtain the subsampled, $J \times K$ image, $I_1(j,k)$, where $J = \lceil M/S \rceil$, $K = \lceil N/S \rceil$, and $S$ is the subsampling factor. Using bicubic spline patches, we then determine, from the $\{I_1(j,k)\}$, a smooth surface $\hat{I}(x,y)$ that interpolates these sampled values and is continuous at the patch boundaries.[1] The approximation, $\hat{I}(m,n)$, of the original image is obtained by subsampling $\hat{I}(x,y)$ at $x = m/S$ and $y = n/S$, where $S = 8$ is the subsampling factor we have chosen to exploit visual masking of errors in the active areas of the image.

Alternatives to the use of bicubic splines are linear splines on a rectangular or quincunx sampling grid. We have found them slightly inferior for our purpose. Another alternative is to use an FIR approximation to an ideal low pass filter both prior to sampling and for interpolation. We have found this approach unsatisfactory because of the serious visual artifacts caused by ideal low pass filtering and interpolation of images.[10] We realize that anti–aliasing of the input prior to subsampling will lead to a better low frequency approximation, but it also leads to an increase in the number of samples that must be encoded and usually to an increase in bit rate, so that it is not done. Noise removal preprocessing does, however, provide some anti-aliasing which we exploit.

The subsampled array $\{I_1(j,k)\}$, from which we compute the spline approximation, $\hat{I}(m,n)$ is represented with 8-bit accuracy and is not further encoded, i.e. the upper error free encoder of Figure 1 is not used. For an $8 \times 8$ subsampling grid, this adds 0.125 bits per pixel to the overall bit rate of the code.

## 2.2 Non uniform quantization and perceptual transparency

The decomposition discussed in the previous section results in a reduced entropy because the remainder image has a significantly lower variance than the original image. However, the number of quantization levels for the remainder remains high. In fact, the dynamic range of the remainder is almost 9 bits. The principal advantage of the differential quantization scheme, though, comes from our ability to exploit the characteristics of this remainder image.

### 2.2.1 Luminance and brightness

It is well known that humans do not distinguish 256 shades of gray. The use of 8 bit gray scale images is due to the fact that at low luminance levels, the just noticeable difference (JND) in luminance is approximately 1/256. Further, any deviation in the mapping from the quantized signal to the brightness of the display will result in perceptible contouring in the low frequency subareas of the image if fewer quantization levels are used.

The number of gray levels in our scheme is always greater than 256 because of the additive contributions of both the (continuous) spline approximation and the (quantized) remainder. Further, the spline approximation is best in the low frequency subareas of the images, making coarser quantization of the remainder now feasible.

### 2.2.2 Visual masking

It is also known that errors in images are substantially less visible in active portions of the image.[11] The phenomenon of visual masking by image activity is generally difficult to exploit in image coding since it requires some analysis of the image. Here, since the remainder is the difference between a smooth approximation and the original, large values in the remainder correspond to the most active portions of the image, for which visual masking will be significant. It has been determined that visual masking, at a viewing distance of 6 times picture height, will occur for a distance of up to six or seven pixels from a sharp transition. This suggests the use of a

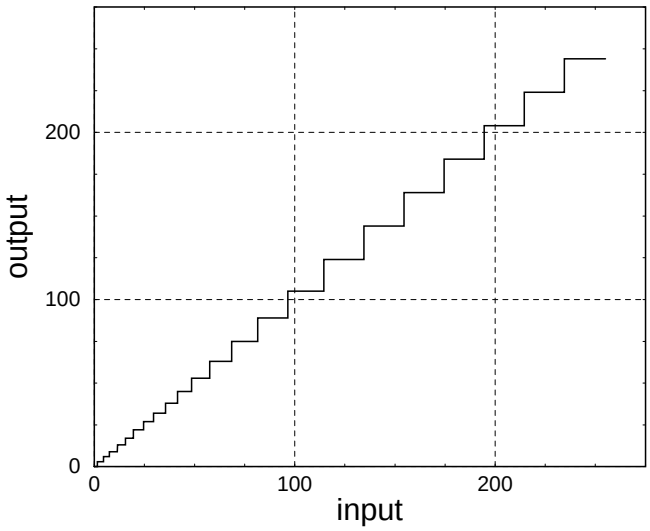


Figure 2: Non uniform quantizer. Note: only half of the symmetric quantizer is shown.

$8 \times 8$ subsampling grid, so that the maximum distance from the two dimensional grid of subsamples is less than six pixels.

### 2.2.3 Non uniform quantization

Based on the above considerations, we devised a non uniform quantization scheme for the remainder. Such a scheme provides fine quantization for the low frequency subareas, where the remainder is small, and allows for coarse quantization in the active areas in the image.

We designed a minimum mean square error quantizer which provides a desired non uniform characteristic based on the first order probability density function of the remainder. The number of quantization levels was progressively decreased until we reached the threshold of perception for the quantization error.

We find that, for all images, quantization of the remainder to approximately 5 bits, or 32 levels, is sufficient to ensure perceptual transparency. Further work, based on a statistical analysis, results in a *universal* non uniform quantizer which assures perceptual transparency for *all* images. This non uniform quantizer has 31 levels in the range of $-115$ to $+115$,[3] and uses a uniform quantizer with step size 20 for larger errors. The theoretical maximum number of levels needed by this quantizer (Figure 2) is then 45.

# 3 ADAPTIVE NOISE REDUCTION

Images commonly used for the evaluation of coding algorithms are quite noisy. To quantify this statement, we have analyzed several images, including images from the Super High Definition (SHD) image test set provided by Nippon Telegraph and Telephone (NTT).

The analysis was performed by locating in each image, a relatively flat region of sufficient extent to allow statistical analysis. These regions generally exhibit a slow trend or shading, which we removed by subtracting a $3 \times 3$ running average. The residual noise was then quantized to integers and analyzed — for all images, we found that the noise is additive with approximately constant variance. The results are shown in the first three columns of Table 1. Of the images in our test set (Figure 6), only 6 had large enough flat regions to perform the analysis.

| image | before | | | after | | |
|---|---|---|---|---|---|---|
| | MSE | PSNR | entropy | MSE | PSNR | entropy |
| bldg | 0.705 | 49.65 | 1.869 | 0.078 | 59.23 | 0.241 |
| daisys | 8.702 | 38.73 | 3.559 | 0.436 | 51.74 | 0.959 |
| lena | 3.386 | 42.83 | 2.904 | 0.070 | 59.71 | 0.300 |
| smile | 0.163 | 56.00 | 1.037 | 0.020 | 65.18 | 0.018 |
| wheel | 1.231 | 47.23 | 2.255 | 0.058 | 60.47 | 0.262 |
| wine | 1.307 | 46.97 | 2.262 | 0.098 | 58.22 | 0.441 |

Table 1: Noise characteristics before and after 10 iterations of CPF based preprocessing.

No gamma correction was done.

Observe that the noise variance and the corresponding entropy are quite high, even for the SHD images. For additive Gaussian noise, at 45 $dB$ PSNR, we predict an entropy of 2 bits/pixel.[6] The values in the table were measured experimentally and the corresponding histograms confirm Gaussian behavior.

Such noise has a large effect on the performance of coders at high quality levels.[6] We have found, however, that anisotropic diffusion based adaptive noise reduction techniques can substantially reduce this noise while maintaining the structured image details, important in the perception of image quality.[2,9]

### 3.1 Anisotropic diffusion

In adaptive noise reduction, an interactive data dependent filtering algorithm is used. It can be shown that filtering with a family of Gaussian filter kernels $G(x, y, t)$ with variance parameter $t$, i.e.

$$I(x, y, t) = I(x, y) * G(x, y, t), \tag{1}$$

is equivalent to the partial differential diffusion equation

$$I_t = c\nabla^2 I = c(I_{xx} + I_{yy}), \tag{2}$$

where the subscripts denote partial derivatives, and $\nabla^2$ is the Laplacian. In anisotropic diffusion, we allow the conduction coefficient, $C(x, y, t)$, to vary with respect to space and time, so that

$$I_t = C(x, y, t)\nabla^2 I + \nabla C \bullet \nabla I = \nabla \bullet [C(x, y, t)\nabla I], \tag{3}$$

where $\nabla$ represents the gradient operation and $\nabla\bullet$, the divergence. Typically, $C = g(\nabla I)$, where $g$ is a nonlinear function to be specified. In our previous work, we use adaptively scaled mean curvature diffusion (MCD)[7] by choosing

$$C = g(\nabla I) = \frac{1}{\sqrt{1 + A^2 \left|\nabla I\right|^2}}, \tag{4}$$

where $A$ is a scaling parameter (which is allowed to vary with time). As such, it can be shown[7] that the local rate of diffusion is equal to twice the mean curvature, $H$, of the image surface about each pixel.

This leads to a very effective, adaptive, iterative noise reduction technique. MCD preserves image structure, characterized by regions of consistently high gradients, and substantially reduces independent, random noise. It, however, also tends to round corners and other features characterized by higher order structure, such as edge intersections.

In more recent work,[8] this problem is dealt with directly, and a diffusion coefficient,

$$C = \frac{1}{|\nabla g|\sqrt{1 + \left[2H(|\nabla g| - 1)\right]^2}}, \tag{5}$$

is developed which preserves corner structures much better. Using this filter, denoted the corner preserving filter (CPF), more iterations are allowed (yielding more noise reduction) while still maintaining perceptual transparency.[8]

Ten CPF iterations results in more than 10 $dB$ of noise reduction in the flat portions of the image (as can been seen in Table 1), while introducing no perceptible changes — as long as the PSNR of the original image is at least 46 dB, which implies that the noise is not perceptible. However, the performance of the coder is substantially improved, as discussed below. Note that, for noisy images, adaptive noise removal may actually *improve* the image quality.

# 4 ERROR FREE ENCODING OF THE QUANTIZED REMAINDER

In the previous sections, we have justified the use of differential quantization using a non uniform quantizer as a means for generating perceptually transparent codes, and the use of anisotropic diffusion based noise removal, but we have not dealt with the issue of compressing the resulting, quantized remainder image. That is the topic of this section.

In the proposed scheme, the non uniform quantization is as coarse as possible while still maintaining perceptual transparency. Therefore, to avoid additional and generally uncontrollable image degradation that may result by further quantization, all subsequent coding of the quantized remainder is error free.

The quantized remainder image is just an 45 level gray scale image, and thus we can encode it with any lossless encoding method available for gray scale images. As luck, or fate, would have it, there are not many choices. Our first attempts used DPCM, but here we consider the use of progressive techniques based on hierarchical pyramids and analysis based interpolation strategies. Recall, that for the remainder of this section, we will be only be concerned with the error free coding of the non uniformly quantized remainder image.

Note that, in our situation, the prediction error is not guaranteed to fall on one of the allowed quantization levels and, even if the prediction is quantized similarly, the difference between the two non uniformly quantized values does not necessarily lie on an allowed quantization level. In short, an error free encoding strategy cannot be developed in such a fashion. We can however, enumerate the quantization levels and take differences between these indices to obtain an error-free encoding technique. Other techniques, such as error feedback,[13] can also be used. Furthermore, the sums in the encoder and decoder can be taken modulo 45 (the number of levels in the non uniform quantizer) without affecting the error free nature of the code. This allows us to represent the difference between indices with 45 levels, instead of 89, which leads to a coding gain.

## 4.1 Interpolation based pyramids

We now consider a progressive representation of the quantized remainder, and the reconstruction — by interpolation — of the highest resolution image from lower resolution subimages. The basic scheme is shown in Figure 3, where we show only 2 stages of the pyramid and haven't included the error free encoding that occurs between the encoder and decoder.

The hope is that analyzing the lower resolution images, which in some sense supply non causal information about the signal, will lead to better predictions than standard, non analysis based techniques, such as DPCM, resulting in smaller errors and a more compressible image representation. In previous work,[1] we used bilinear interpolation and a directional interpolation strategy as the basis for the image pyramid and obtained encouraging results. The results obtained using the directional interpolation were 9% better than those obtained using DPCM.[†]

[†] We will use DPCM to indicate our DPCM encoding technique for encoding the quantized remainder. We use JPEG-DPCM to refer to the standard error-free DPCM encoder used in the JPEG standard.

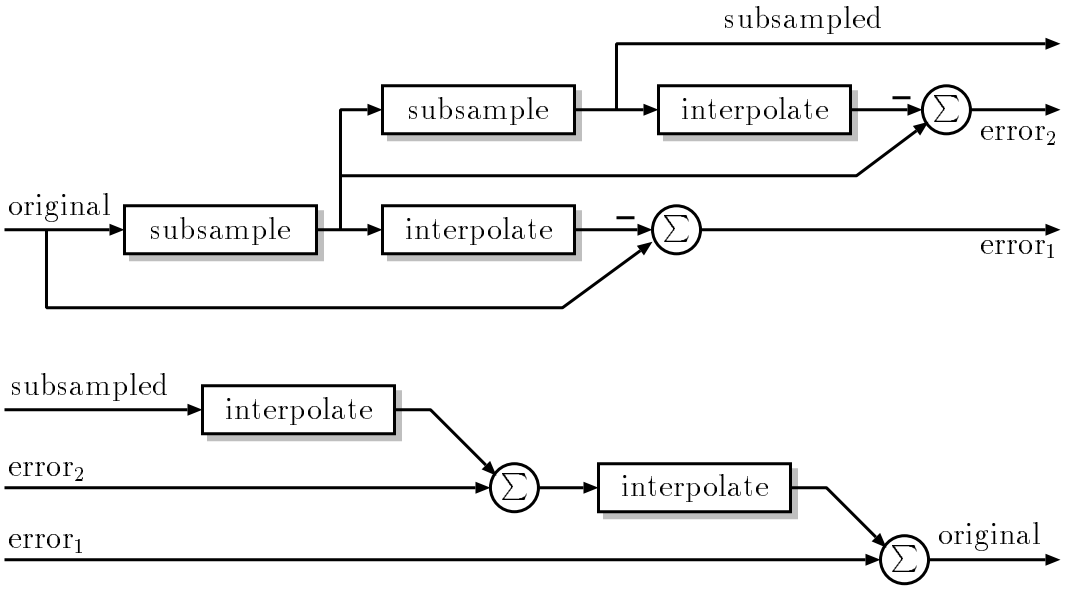


Figure 3: Two stage pyramidal code: encoder (top) and decoder (bottom).

The directional interpolation algorithm[1] used previously (Figure 4) computes Robert's gradient angle estimates from the low resolution images in the pyramid and then interpolates these to the high resolution grid. From the high resolution angle estimates a low frequency direction is determined that guides the interpolation algorithm. It is difficult to compute a robust high resolution gradient estimate and when the estimate is incorrect large prediction errors result. In that work, we used our confidence in the gradient estimate to blend the results with those obtained with a more conservative, bilinear interpolation. Recently, however, we have developed an analysis based anisotropic diffusion interpolation algorithm which, in a sense, performs directional filtering, yet does not use *high resolution* gradient estimates.

The diffusion based algorithm is quite simple. It simply consists of starting with an initial approximation of the high resolution image, such as a bilinear interpolated version, and then applying the $MCD$ algorithm while holding the subsampled pixel values fixed. By constraining the solution, we can obtain images which are good approximations of the original, hopefully in the sense of reduced entropy.

Typically, subsampling and interpolation are performed on a rectangular grid, in which case, the smallest symmetric subsampling yields a 4:1 reduction in the number of pixels in each stage of the pyramid. As an alternative, quincunx sampling only reduces the number of pixels by a factor of two at each stage, and we can cascade two stages of quincunx interpolation to get the equivalent of one stage of rectangular interpolation. In additive decomposition based coding, such as this, the quincunx technique has many advantages. Briefly, each pixel is encoded with respect to its four nearest neighbors at the *current* resolution, and the special cases associated with rectangular interpolation reduce to a single, simple case (Figure 5). See our previous work[1] for a more thorough discussion. Indeed, we showed in that work the advantages of using a quincunx pyramid. However, here we will restrict our attention to rectangular subsampling as we have not developed a suitable framework for quincunx interpolation using our new anisotropic diffusion based interpolation algorithm.

A very important clarification is that we are primarily interested in potential gains of using analysis based interpolation techniques, or analysis based techniques in general, for the encoding of the quantized remainder images. These images have properties similar to edge maps, with very obvious structural dependencies, and due to the noise removal preprocessing, are relatively noise-free. We feel that these are favorable conditions for an analysis based interpolation algorithm.

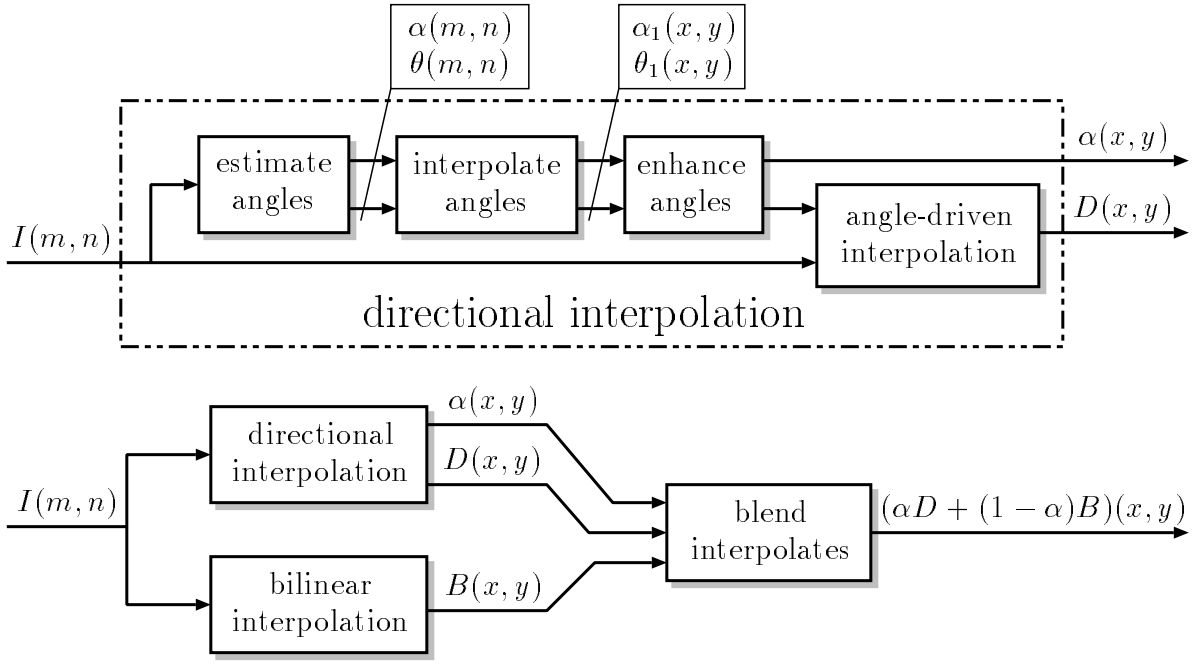


Figure 4: Overview of blended, directional interpolation strategy (top) and directional interpolation (bottom).

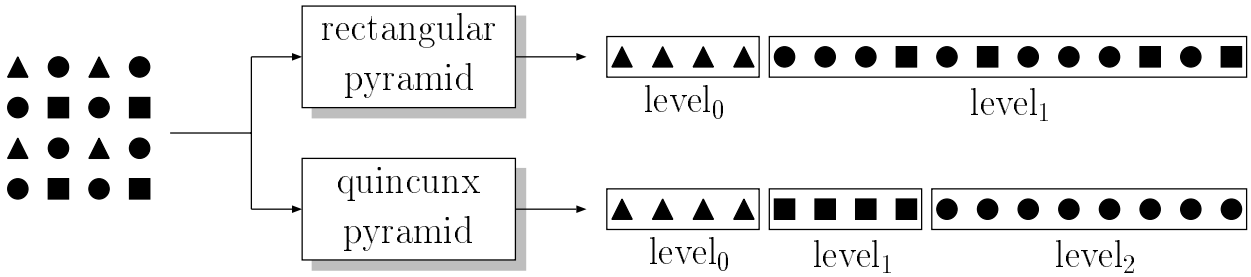


Figure 5: Quincunx sampling results in more symmetric neighborhoods, and better predictions than rectangular pyramids.

### 4.2 Error free coding

We have yet to discuss the actual encoding strategy used to encode the error images. Given the sometimes large percentage of zeros in these images, Huffman codes can be inefficient. Further, these zeros exhibit coherency, since the errors are localized in the active areas in the image. To exploit this coherency and make Huffman coding more efficient, we encode the error images using a hybrid strategy which first encodes the position of the non zero pixels using a binary image encoder (the QM-code) with a seven pixel predictor, after which we encode the non zero pixels with a Huffman code. This is a simple implementation of the color shrinking technique[5] for encoding sparse sources. This strategy results in an average 7% bit rate reduction for the analysis based pyramids (after noise preprocessing), but for the simpler images can be as high as 27%. The same numbers for the original (non preprocessed images) are 1.5% and 10%, respectively, illustrating the increasing importance of efficient coding as the the predictions become more accurate. In addition, the 4096 samples from the image transmitted by the differential quantization stage propagate through the the entire pyramid and, thus, the lowest resolution image in our pyramid consists entirely of zeros, and need not be encoded.

# 5 RESULTS

In this section, we summarize the results we have obtained with our perceptually transparent coding scheme, with and without adaptive noise removal. For each case, we have encoded the remainder using DPCM, as outlined in Section 4, and five interpolation based pyramids, as described in Section 4.1. The test set consists of the nine $512 \times 512$, 8 bit images shown in Figure 6. The results obtained are presented in Tables 2, 3 and 4.

In Tables 2 and 3, each pair of results consists of a Huffman encoded bit rate, followed by a color shrinking based result, as discussed in Section 4.2, in which the 0 entries are encoded as a binary activity mask using standard binary image encoding techniques. The first column is the output of the differential quantization stage, after non uniform quantization, and is the quantized remainder which we further encode in an error free fashion. The second column presents the results obtained by DPCM encoding the quantized remainder, while the last five columns present the results for 5 different pyramid based encoding techniques. Three iterations of MCD are applied to a bilinear approximation to obtain the diffusion based results. Quincunx results are not presented, in this case, since we haven't developed a suitable quincunx based interpolation algorithm yet. As a summary, some of these results are repeated in Table 4, and compared with JPEG-DPCM (for each image we used the kernel that gave the best results.) Note that such a scheme is not progressive and that we expect progressive schemes to be less efficient.

Comparing the pyramid based results, we see that the results for directional interpolation on a quincunx pyramid are clearly best, with the bilinear quincunx pyramid a close second. The rectangular pyramids are slightly worse, and are worse even than the corresponding DPCM results. These results seem to be in conflict with results we have previously presented,[1] but we are presenting different DPCM results here. In the prior work, we presented results obtained by Huffman encoding the entire DPCM predicted image. Here, we do not encode the 4096 pixels which are known from the $8 \times 8$ subsampled values, and use color shrinking to obtain a more effective representation. When color shrinking is applied to all techniques, we find that the previously presented gains, due to pyramid coding, are lost. However, it is still significant that we can compress the image as effectively using a hierarchical technique. Improved analysis based interpolation techniques may lead to better results than DPCM.

In comparing the analysis and non analysis based pyramid encoding techniques, we find that we do obtain moderately better results using analysis, with the quincunx pyramid giving a 2% bit rate reduction with respect to the bilinear quincunx result. Quincunx sampling results are 5-6% better than the corresponding rectangular results. Finally, the analysis based rectangular pyramids are slightly better than bilinear interpolation.

The anisotropic diffusion based results are slightly better than the results obtained with the previous directional interpolation algorithm (for rectangular pyramids), but since better results are obtained with quincunx pyramids, for which we do not have a diffusion based algorithm, its true merit is not known. It is however a simpler strategy, since no blending has been used. Examining the interpolated approximations using the diffusion based interpolation, we found that they are not nearly as good as we expected, most likely due to the effect of aliasing on the gradient computation. Although, this algorithm doesn't use high resolution gradient estimates, which we touted earlier as a possible advantage, it appears that it would benefit from such. Accurate gradient estimation seems to be very important for analysis based interpolation. Even with accurate gradient computations, we need to reexamine our interpolation rules, given this information. Alternatives have been presented,[12,14] which we plan to evaluate in future studies.

Note that with a good directional interpolation strategy, coarser quantization could be used while still maintaining perceptual transparency, since the errors would be more localized to edges and areas where masking is significant.

In Table 3, we show the effect of adaptive noise reduction. For the DPCM scheme, the noise reduction leads to a 14% decrease in bit rate, averaged over all images. For the quincunx pyramid it results in a 13% reduction. The gain of perceptually transparent coding, over JPEG-DPCM (Table 4) is slightly over 50%, i.e., we have compressed the images by an additional factor of two such that the error introduced are not visible.

| Image | quantized remainder | DPCM | bilinear | | directional | | diffusion |
|---|---|---|---|---|---|---|---|
| | | | rectangular | quincunx | rectangular | quincunx | rectangular |
| baboon | 4.931/4.924 | 4.255/4.262 | 4.378/4.379 | 4.273/4.271 | 4.386/4.387 | 4.246/4.251 | 4.376/4.376 |
| bldg | 4.632/4.459 | 2.659/2.603 | 2.942/2.868 | 2.726/2.671 | 3.014/2.946 | 2.614/2.575 | 2.905/2.834 |
| daisys | 4.541/4.518 | 3.350/3.336 | 3.489/3.464 | 3.360/3.352 | 3.476/3.460 | 3.308/3.304 | 3.472/3.448 |
| flowers | 3.867/3.698 | 1.938/1.935 | 2.143/2.134 | 1.987/1.960 | 2.085/2.074 | 1.966/1.939 | 2.072/2.053 |
| lena | 3.875/3.827 | 2.778/2.806 | 2.865/2.879 | 2.768/2.762 | 2.840/2.842 | 2.719/2.719 | 2.851/2.865 |
| lynda | 3.122/2.871 | 1.868/1.858 | 1.961/1.938 | 1.833/1.805 | 1.897/1.877 | 1.823/1.794 | 1.922/1.894 |
| smile | 2.438/2.098 | 1.595/1.439 | 1.717/1.552 | 1.606/1.450 | 1.672/1.526 | 1.598/1.442 | 1.692/1.526 |
| wheel | 3.624/3.413 | 2.035/2.037 | 2.249/2.226 | 2.114/2.103 | 2.178/2.164 | 2.058/2.049 | 2.205/2.184 |
| wine | 4.071/3.825 | 2.221/2.155 | 2.375/2.309 | 2.222/2.174 | 2.304/2.250 | 2.173/2.130 | 2.316/2.249 |
| average | 3.900/3.737 | 2.522/2.492 | 2.680/2.639 | 2.543/2.505 | 2.650/2.614 | 2.501/2.467 | 2.646/2.603 |

Table 2: Transparent coding on the original images.

| Image | quantized remainder | DPCM | bilinear | | directional | | diffusion |
|---|---|---|---|---|---|---|---|
| | | | rectangular | quincunx | rectangular | quincunx | rectangular |
| baboon | 4.901/4.862 | 4.163/4.122 | 4.297/4.254 | 4.171/4.136 | 4.298/4.257 | 4.139/4.114 | 4.290/4.242 |
| bldg | 4.528/4.111 | 2.475/2.310 | 2.766/2.573 | 2.563/2.405 | 2.835/2.653 | 2.450/2.310 | 2.729/2.535 |
| daisys | 4.484/4.358 | 3.122/3.038 | 3.298/3.206 | 3.137/3.079 | 3.284/3.195 | 3.081/3.026 | 3.265/3.174 |
| flowers | 3.832/3.589 | 1.869/1.827 | 2.096/2.045 | 1.917/1.854 | 2.026/1.969 | 1.898/1.831 | 2.019/1.953 |
| lena | 3.615/3.167 | 2.230/2.066 | 2.387/2.215 | 2.233/2.106 | 2.324/2.164 | 2.178/2.058 | 2.348/2.175 |
| lynda | 2.911/2.432 | 1.593/1.427 | 1.753/1.576 | 1.606/1.431 | 1.681/1.512 | 1.594/1.417 | 1.699/1.511 |
| smile | 2.154/1.561 | 1.413/1.008 | 1.526/1.127 | 1.417/1.044 | 1.477/1.095 | 1.411/1.036 | 1.493/1.089 |
| wheel | 3.320/2.661 | 1.774/1.453 | 1.963/1.628 | 1.822/1.545 | 1.896/1.575 | 1.771/1.496 | 1.904/1.567 |
| wine | 4.015/3.649 | 2.136/2.014 | 2.307/2.178 | 2.136/2.037 | 2.227/2.110 | 2.086/1.992 | 2.237/2.107 |
| average | 3.751/3.377 | 2.308/2.141 | 2.488/2.311 | 2.334/2.182 | 2.450/2.281 | 2.290/2.142 | 2.443/2.261 |

Table 3: Transparent coding on the noise reduced images (10 CPF iterations).

| Image | original | | | 10 CPF iterations | | |
|---|---|---|---|---|---|---|
| | JPEG | DPCM | directional | JPEG | DPCM | directional |
| baboon | 6.572 | 4.262 | 4.251 | 6.470 | 4.122 | 4.114 |
| bldg | 4.233 | 2.603 | 2.575 | 3.782 | 2.310 | 2.310 |
| daisys | 5.453 | 3.336 | 3.304 | 5.143 | 3.038 | 3.026 |
| flowers | 3.524 | 1.935 | 1.939 | 3.246 | 1.827 | 1.831 |
| lena | 4.704 | 2.806 | 2.719 | 3.777 | 2.066 | 2.058 |
| lynda | 3.655 | 1.858 | 1.794 | 2.694 | 1.427 | 1.417 |
| smile | 2.909 | 1.439 | 1.442 | 2.029 | 1.008 | 1.036 |
| wheel | 3.640 | 2.037 | 2.049 | 2.623 | 1.453 | 1.496 |
| wine | 3.977 | 2.155 | 2.130 | 3.700 | 2.014 | 1.992 |
| average | 4.296 | 2.492 | 2.467 | 3.718 | 2.141 | 2.142 |

Table 4: Advantages of preprocessing and perceptually transparent coding. Note: the directional results are for a quincunx pyramid.

Figure 6: Test images. In order, left to right and top to bottom: baboon, bldg, daisys, flowers, lena, lynda, smile, wheel, and wine. Of these, flowers, smile, wheel, and wine are from the SHD test set.

# 6 DISCUSSION

This paper makes two major contributions. The first one is conceptual. It is possible to process and represent images so as to improve their compressibility without loss of image quality. In particular, adaptive noise reduction leads to a substantial increase in compressibility with no visible change in the image. The second contribution is that a carefully designed progressive code, when used in conjunction with image analysis, results in a hierarchical code which is as efficient as good DPCM based techniques. We expect, with further effort, that they could be made better, in contrast to the slight loss of performance generally associated with progressive coding schemes.

Note that our approach to perceptually transparent coding, which controls image quality by first introducing imperceptible changes in the image, now requires increased attention to efficient error free coding schemes for quantized gray scale and color images.

# 7 ACKNOWLEDGMENTS

This research was supported in part by the UC MICRO program, Pacific Bell, Lockheed and Hewlett Packard.

# 8 REFERENCES

[1] V. R. Algazi, G. E. Ford, R. R. Estes, A. El-Fallah, and A. Najmi. Progressive perceptually transparent coder for very high quality images. In *Proceedings of the SPIE Symposium on Applications of Digital Image Processing XVII*, 1994.

[2] V. R. Algazi, G. E. Ford, M. Mow, and A. Najmi. Design of subband coders for high quality images, based on perceptual criteria. In *Proceeding of the SPIE Symposium on Applications of Digital Image Processing XVI*, volume 2028, pages 40–6, July 1993.

[3] V. R. Algazi, E. Maurincomme, and G. E. Ford. Error free and transparent coding of images using approximations by splines. In *Proceedings of the SPIE Symposium on Image Processing Algorithms and Techniques II*, volume 1452, pages 364–70, Boston, MA, November 1991.

[4] V. R. Algazi, T. R. Reed, G. E. Ford, E. Maurincomme, I. Hussain, and R. Potharlanka. Perceptually transparent coding of still images. *IEICE*, E75-B(5):340–8, March 1992.

[5] V. Ralph Algazi, Phillip L. Kelly, and Robert R. Estes. Compression of binary facsimile images by preprocessing and color shrinking. *IEEE Transactions on Communications*, 38(9):1592–8, September 1990.

[6] V. Ralph Algazi, T. R. Reed, Gary E. Ford, and R. R. Estes. Image analysis for adaptive noise reduction in super high definition image coding. In *Proceedings of the SPIE Symposium on Visual Communications and Image Processing*, volume 1818, Boston, MA, November 1992.

[7] A. I. El-Fallah and G. E. Ford. Mean curvature evolution and surface area scaling in image filtering. Submitted to the IEEE Transactions on Image Processing, June 1994.

[8] A. I. El-Fallah, G. E. Ford, V. R. Algazi, and R. R. Estes, Jr. The invariance of edges and corners under mean curvature diffusions of images. In *Proceedings of the SPIE, Image and Video Processing III*, San Jose, CA, February 1995.

[9] Adel El-Fallah and Gary E. Ford. Image filtering by gradient inverse inhomogeneous diffusion. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, volume V, pages V73–6, April 1993.

[10] T. A. Hentea and V. R. Algazi. Perceptual models and filtering of high–contrast achromatic images. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-14(2):230–45, 1984.

[11] A. Netravali and B. Haskell. *Digital Pictures*. Plenum Press, 1988. Chapter 4.

[12] John A. Robinson. Compression of super-high-definition multimedia images. In *Proceedings of the SPIE, Very High Resolution and Quality Imaging Conference*, volume 2663, San Jose, California, February 1996.

[13] K. M. Uz, M. Vetterli, and D. LeGall. Interpolative multiresolution coding of advanced television with compatible subchannels. *IEEE Transactions on CAS for Video Technology*, 1(1):88–99, March 1991. Special Issue on Signal Processing for Advanced Television.

[14] Xiaolin Wu, Nasir Memom, and Khalid Saywood. A context-based adaptive lossless/nearly-lossless coding scheme for continuous tone images. Submitted to ISO as a candidate algorithm for the next international standard for lossless compression of continuous-tone images, 1995.