# SUPRA-THRESHOLD PERCEPTUAL IMAGE CODING

*Thrasyvoulos N. Pappas* [1] *Thomas A. Michel* [2] *Raynard O. Hinds* [3]

[1]Bell Laboratories, Lucent Technologies, Murray Hill, NJ 07974, USA
Email: pappas@research.bell-labs.com

[2]Bell Laboratories, Lucent Technologies, Murray Hill, NJ 07974, USA
Email: jazz@research.bell-labs.com

[3]EECS Department, MIT, Cambridge, MA 02139, USA
Email: rohinds@mit.edu

## ABSTRACT

We investigate different algorithms and performance criteria for supra-threshold image compression. The algorithms include JPEG and perceptual JPEG, the Safranek-Johnston perceptual subband image coder (PIC), and the Said-Pearlman algorithm which is based on Shapiro's embedded zerotree wavelet algorithm. We also consider a number of performance criteria. These include mean-squared error, Watson's perceptual metric, a metric based on the PIC coder, as well as an eye-filter weighted mean-squared-error metric. Our experiments indicate that the PIC metric provides the best correlation with subjective evaluations. The metric predicts that at very low bit rates the Said-Pearlman algorithm and the $8 \times 8$ subband PIC coder perform the best, while at high bit rates the $4 \times 4$ subband PIC coder dominates.

## 1. INTRODUCTION

We investigated different algorithms and performance criteria for supra-threshold image compression, i.e. when the coding distortions exceed the threshold of visibility. Currently, most perceptual models provide the maximum amount of distortion that can be introduced to the signal without resulting in any perceived distortion. This is usually referred to as the just noticeable amount of distortion. However, in order to achieve very low compression rates, a certain amount of perceived distortion is unavoidable. Hence there is a need to derive quantitative (objective) measures of perceived distortion. In general, it is easier to obtain models for the perceptually transparent case; it is much more difficult to quantify perceived distortion.

We consider a number of algorithms over a wide range of bit rates, including very low rates where compression artifacts are clearly visible. The algorithms include JPEG and perceptual JPEG [1], the Safranek-Johnston perceptual subband image coder (PIC) [2]), and the Said-Pearlman algorithm [3] which is based on Shapiro's embedded zerotree wavelet (EZW) algorithm [4]. We also consider a number of modifications of the PIC coder to better adapt it to the low bit rates under consideration. While most of these algorithms were originally intended for threshold and near-threshold image compression, they provide an important starting point for deriving supra-threshold compression schemes. EZW is the only algorithm that was designed for low bit rates.

We also consider a number of performance criteria. These include the simple mean-squared error (MSE), Watson's perceptual metric [5], a metric based on the Safranek-Johnston PIC coder, as well as an eye-filter weighted MSE metric. Our experiments show that at low bit rates the coders generate a wide variety of distortions that are hard to quantify. As expected, the MSE metric does not correlate with subjective evaluations except at very low bit rates (where all metrics perform well), while Watson's metric is strongly biased towards JPEG. We found that the PIC metric provides the best correlation with subjective evaluations. The metric predicts that at very low bit rates (0.05 *to* 0.1 bits/pixel) the Said-Pearlman algorithm and the $8 \times 8$ subband PIC coder perform the best, while at high bit rates (0.4 bits/pixel and higher), when the distortions are not as severe but still noticeable, the $4 \times 4$ subband PIC coder dominates. However, the relative performance of the coders depends on the image characteristics and none of the metrics we consider can predict subjective quality consistently.

## 2. CODERS

The baseline JPEG algorithm uses $8 \times 8$ DCT's and a quantization table that is fixed for all DCT blocks. The perceptual JPEG [1] algorithm uses perceptual criteria to determine which DCT coefficients are perceptually insignificant and sets them to zero before the quantization step. A block diagram of the perceptual JPEG coder is shown in Fig. 1. The only difference from the standard JPEG algorithm is the pre-quantization block. Safranek showed that, at rates close to the perceptual threshold, the perceptual JPEG [1] algorithm offers greater compression without loss in subjective image quality. At lower rates, however, we were not able to obtain any significant compression gains; thus in this paper we use baseline JPEG.

The Safranek-Johnston PIC coder [2] uses an empirically derived perceptual masking model. A block diagram of the PIC coder is shown in Fig. 2. It uses GQMF analysis/synthesis filters. The base band is coded with DPCM while all other subbands are coded with PCM. All subbands use uniform quantizers with sophisticated entropy coding. The perceptual model specifies the amount of noise that can be added to each subband of a given image so that the difference between the output image and the original is just noticeable. To achieve higher compression ratios, the perceptual thresholds must be adjusted so that they result in the least objectionable distortion. Our experiments indicate that simple scaling of the thresholds provides the best performance. We also found that the texture and local brightness components of the perceptual model were *not* effective at any compression level; thus, we used only the "base sensitivity" for each subband. We tried both $4 \times 4$ (original PIC coder) and $8 \times 8$ subband decompositions. We also experimented with adding dither to the baseband but found that it does not offer any significant gains in the quality vs. compression efficiency domain.

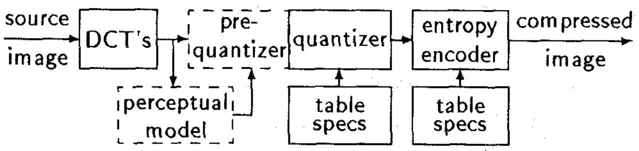



Figure 1. Perceptual JPEG Coder

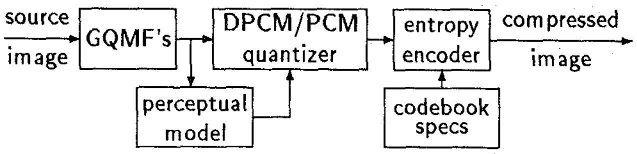



Figure 2. Perceptual Subband Image Coder (PIC)

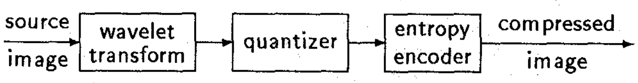



Figure 3. Embedded Zerotree Wavelet (EZW) Coder

The Said-Pearlman algorithm [3] is based on Shapiro's EZW algorithm [4]. It uses a hierarchical subband decomposition to efficiently encode magnitude and ordering information across different scales. Unlike the other two coders, this algorithm requires no training or prestored tables. It is designed to minimize the MSE metric. Thus, it does not use any explicit perceptual models.

## 3. PERFORMANCE AND METRICS

We considered a number of performance criteria and used them to evaluate the performance of the various algorithms over a wide range of bit rates ranging from extremely low (i.e. at the point where some of the coders completely break down) to very high where coding distortions are barely visible. We also conducted informal subjective evaluations of the performance of the coders using a variety of test images.

At low bit rates, the different coders generate different types of artifacts. The JPEG coder tends to generate a lot of blocking artifacts and artificial contours. The PIC coders tend to generate horizontal and vertical ringing, while diagonal lines have aliasing artifacts; they also generate midfrequency noise. The Said-Pearlman algorithm creates smudging in areas of texture and along diagonal edges; it also results in a lot of blurring, which is especially noticeable in large areas of texture. Thus, the performance of each algorithm depends strongly on image content. Fig. 4 shows a comparison of the different coders at 0.10 bits/pixel for the "Bank" image.

Our subjective evaluations indicate that at very low bit rates the Said-Pearlman coder and the $8 \times 8$ PIC coder perform the best. As the rate increases, the PIC coders generally improve faster than the other two coders and at high bit rates the $4 \times 4$ PIC coder usually dominates.

The simplest and most popular error metric is the MSE. Our subjective evaluations show that the metric predictions are wrong except at the lowest rates when the differences between the coders are striking. Fig. 5 shows that this metric is strongly biased towards the Said-Pearlman algorithm. This is not surprising since the algorithm is designed to minimize this metric. MSE also incorrectly predicts that JPEG performs better than the PIC coders at mid to high rates.

We also tried a weighted MSE metric, where the error is filtered by an FIR eye filter designed for a viewing distance of approximately 6 image heights. Fig. 6 shows that this metric is also biased towards the Said-Pearlman algorithm. Although better than MSE, this metric is still not consistent with our subjective evaluations.

Watson's perceptual metric [5] is defined as follows

$$\left\{ \frac{1}{N} \sum_{i,k} \left[ \frac{b_k(i) - \hat{b}_k(i)}{t_k(i)} \right]^2 \right\}^{\frac{1}{2}} \tag{1}$$

where $b_k(i)$ and $\hat{b}_k(i)$ are the $i$-th element of the $k$-th DCT block of the original and coded image, respectively, and $t_k(i)$ is the corresponding perceptual threshold. Our subjective evaluations show that the metric predictions are incorrect for many different images and over many coding rates. Fig. 7 shows that this metric is strongly biased towards the JPEG algorithm. This is not surprising since both the metric and JPEG are based on DCTs.

Finally, we used a metric based on the PIC coder

$$\left\{ \frac{1}{N} \sum_{i,k} \left[ \frac{s_k(i) - \hat{s}_k(i)}{\tau_k(i)} \right]^2 \right\}^{\frac{1}{2}} \tag{2}$$

where $s_k(i)$ and $\hat{s}_k(i)$ are the $i$-th element of the $k$-th subband of the original and coded image, respectively, and $\tau_k(i)$ is the corresponding perceptual threshold. We found that the PIC metric provides the best correlation with subjective evaluations. Fig. 8 shows the predictions of the PIC metric implemented using a $4 \times 4$ subband decomposition. It successfully predicts that at very low bit rates the Said-Pearlman algorithm and the $8 \times 8$ PIC coder perform significantly better than the other two coders It also predicts that Said-Pearlman performs slightly better than the $8 \times 8$ PIC coder; the difference is greater for images without a lot of texture and sharp edges. Finally, this metric predicts that at high bit rates the $4 \times 4$ PIC coder has the best performance. In fact, only the PIC metric predicts the relative performance of the $4 \times 4$ and $8 \times 8$ PIC coders consistently. In general, the PIC metric predicts the trends in relative coder performance but does not predict the exact position of the cross-over points.

The relative performance of the coders depends not only on the image characteristics but also on personal preferences. Thus it is very difficult for any metric to predict subjective quality consistently. One direction for improvement is to develop criteria that include more elaborate texture masking models.

## REFERENCES

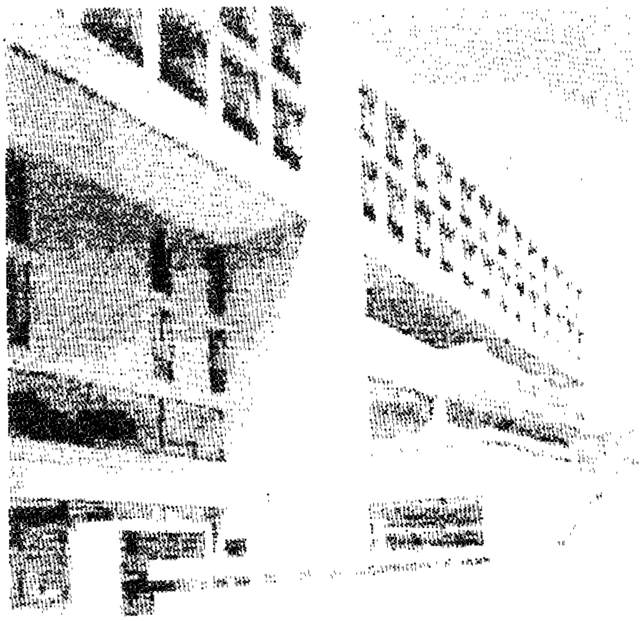
[1] R.J. Safranek, "A JPEG compliant encoder utilizing perceptually based quantization," *Proc. SPIE, vol. 2179, Human Vision, Visual Proc., and Digital Display V*, p. 117–126, Feb. 1994.

[2] R.J. Safranek, J.D. Johnston, "A perceptually tuned subband image coder with image dependent quantization and post-quantization," *ICASSP-89*, p. 1945–1948.

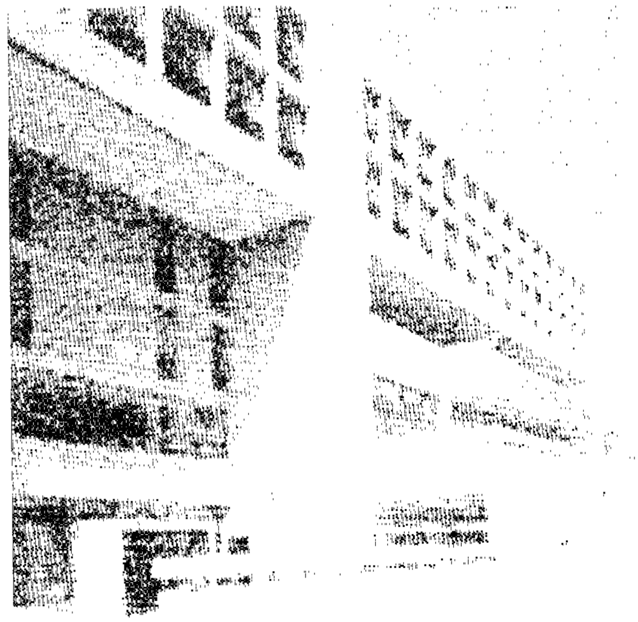
[3] A. Said and W.A. Pearlman, "A new fast and efficient image codec based on set partitioning in hierarchical trees," submitted to *IEEE Tr. Cir. Sys. Video Tech.*

[4] J.M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Tr. SP*, vol. 41, p. 3445–3462, Dec. 1993.

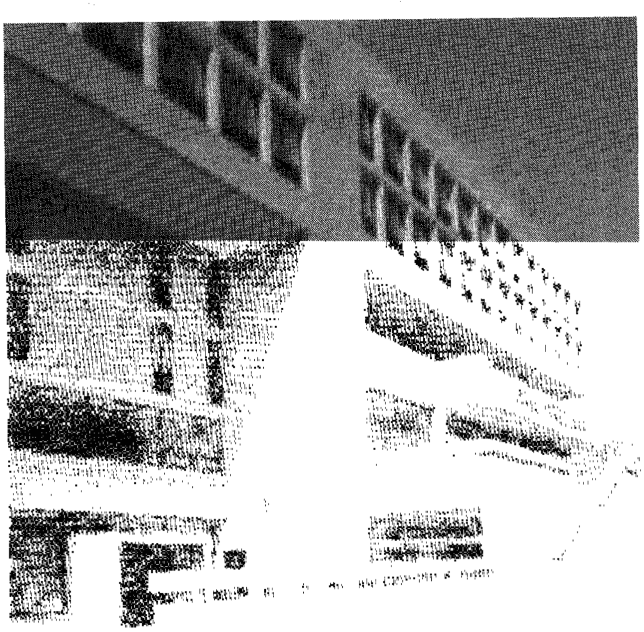
[5] A.B. Watson, "DCT quantization matrices visually optimized for individual images," *Proc. SPIE, vol. 1913, Human Vision, Visual Proc., and Digital Display IV*, p. 202–216, Feb. 1993.



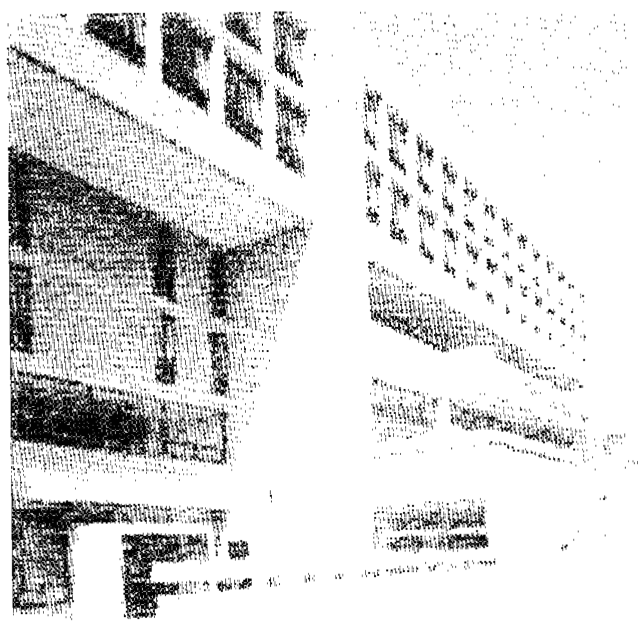
(a) JPEG Coder



(b) Modified PIC Coder ($4 \times 4$)



(c) Said-Pearlman Coder



(d) Modified PIC Coder ($8 \times 8$)

Figure 4. Coders at 0.10 bits/pixel ("Bank" image)

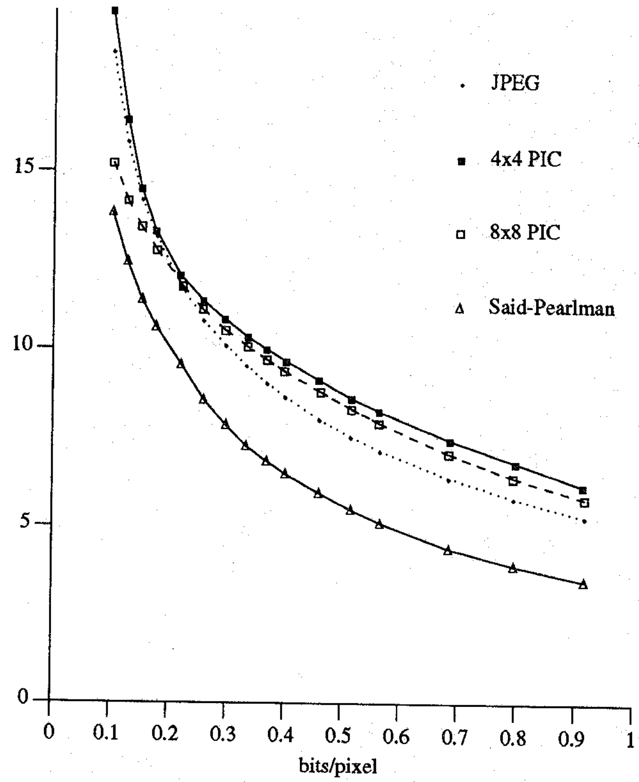


Figure 5. Mean-Squared Error Metric for "Bank"

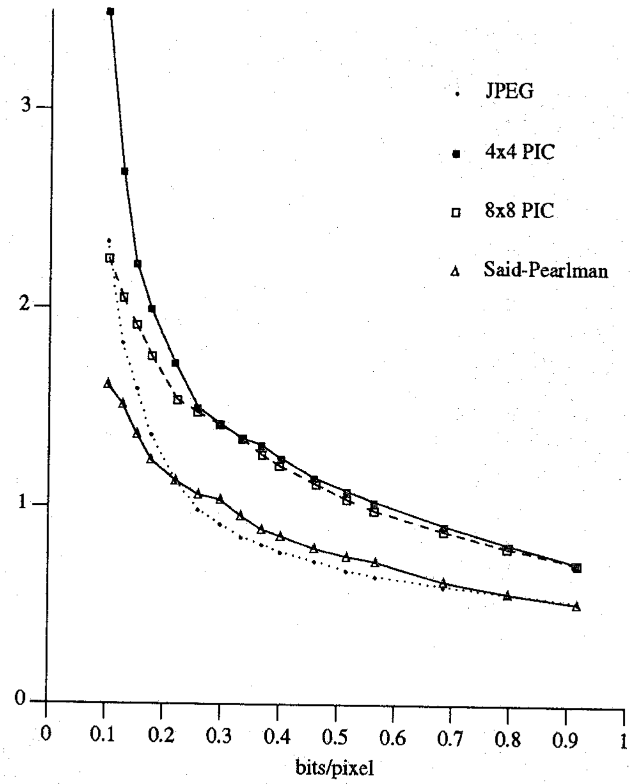


Figure 7. Watson's Perceptual Error Metric for "Bank"

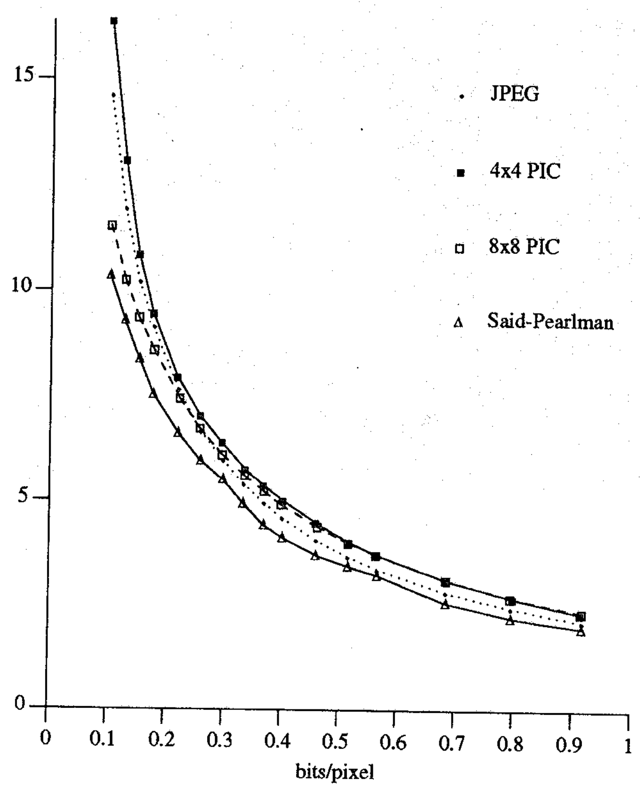


Figure 6. Weighted Mean-Squared Error Metric for "Bank"

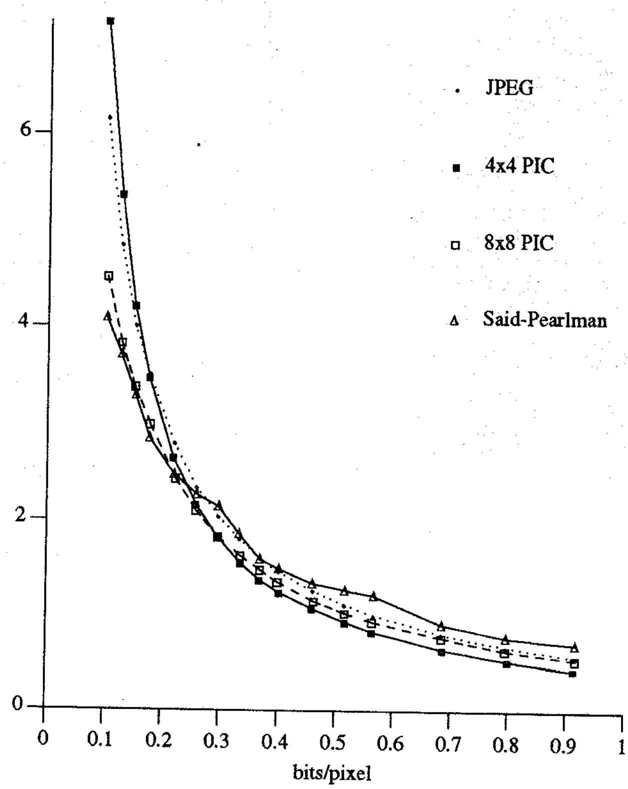


Figure 8. Perceptual (PIC) Error Metric for "Bank"