# JPEG COMPLIANT EFFICIENT PROGRESSIVE IMAGE CODING

*Jaehan In, Shahram Shirani, and Faouzi Kossentini*

Department of Electrical and Computer Engineering,
University of British Columbia,
Vancouver BC V6T 1Z4, Canada

## ABSTRACT

Among the different modes of operations allowed in the current JPEG standard, the sequential and progressive modes are the most widely used. While the sequential JPEG mode yields essentially the same level of compression performance for most encoder implementations, the performance of progressive JPEG depends highly upon the designed encoder structure. This is due to the flexibility the standard leaves open in designing progressive JPEG encoders. In this paper, a rate-distortion optimized JPEG compliant progressive encoder is presented that produces a sequence of bit scans, ordered in terms of decreasing importance. Our encoder outperforms a baseline sequential JPEG encoder in terms of compression, significantly at medium bit rates, and substantially at low and high bit rates. Moreover, unlike baseline JPEG encoders, ours can achieve precise rate/distortion control. Good rate-distortion performance at low bit rates and precise rate control, provided by our JPEG compliant progressive encoder, are two highly desired features currently sought for JPEG-2000.

## 1. INTRODUCTION

Images usually contain so much data that they need to be compressed prior to storage or transmission. JPEG (Joint Photographic Experts Group) is the current standard for compression and decompression of still, continuous tone, monochrome and color images. JPEG has four distinct modes of operation: sequential DCT-based, progressive DCT-based, lossless, and hierarchical [1].
Progressive image transmission using the progressive JPEG (PJPEG) mode is suitable for interactive image communication applications over bandwidth constrained channels. The least information necessary to represent an image is transmitted with as few bits as possible. Based on the receiver's request, the image is progressively reconstructed in several stages until the desired quality is achieved. While the sequential mode provides partial information of the full image, in the PJPEG mode of operation, one can see a rough approximation of the input image after the first scan has been decoded, and the image reproduction quality is then gradually improved as more scans are decoded. The PJPEG mode of operation is preferred in many applications [2]. In this paper, we present a new PJPEG compliant encoding algorithm that is rate-distortion (RD) optimized. The proposed algorithm produces an ordered sequence of scans that can be used to progressively encode/decode an image, achieving precise rate/distortion control and yielding relatively good compression performance at all bit rates, including the very low bit rates. In the next section, we provide a description of the PJPEG progressive mode of operation. The proposed PJPEG image encoding algorithm is discussed in Section 3. Experimental results and conclusions are presented in Sections 4 and 5.

This work was supported by MDSI, BC ASI and NSERC.

## 2. BACKGROUND: PROGRESSIVE JPEG

In the progressive JPEG (PJPEG) mode, each scan in the DCT transformed image corresponds to a few bits of one or more coefficients, and therefore represents a portion of the image being encoded/decoded. The DCT coefficients, each of which contributing differently to improving the image quality, can be divided into many groups or decomposed into bits as specified in the PJPEG mode, such that the reconstructed image quality increases as more groups of bits of the DCT coefficients are received and decoded. Two different modes are defined in PJPEG, spectral selection (SS) and successive approximation (SA). These two modes can be used independently or combined to provide full progression. In spectral selection, the DCT coefficients in each $8 \times 8$ block are segmented into frequency bands of variable lengths. Each frequency band is then encoded in one scan. In successive approximation, the precision of the DCT coefficients of the $8 \times 8$ blocks is reduced by dividing them by a power of two (or shifting right their binary representations) before encoding such coefficients. The received DCT coefficients are decoded and multiplied by the same power of two (or shifting left their binary representations) before the IDCT operation. The precision of the DCT coefficients is increased by one bit in each of the subsequent scans [3, 4].
PJPEG imposes a few constraints on the progressive encoder's structure. Such constraints, which must be satisfied by our proposed PJPEG design algorithm, are described below:

- The first scan to be encoded/decoded must contain only DC information.
- The DC and AC coefficients cannot be encoded into the same scan.
- The coefficient numbers in the same spectral band must be consecutive, i.e., their locations can be described by just two points: 1) start of band and 2) end of band.
- Only the first scan of any DCT coefficient can contain more than one bit, and the subsequent scans must represent only one precision bit at a time.
- Encoding/decoding of a bit of a coefficient requires that the proceeding bits of that coefficient be encoded/decoded in advance.

## 3. PROPOSED PJPEG CODING ALGORITHM

Understanding the relationship between the DCT coefficients and the reproduction quality can help us greatly design an efficient and effective PJPEG compliant encoder. However, ordering of $64$ DCT coefficients in an $8 \times 8$ block according to subjective importance is a very difficult task. Moreover, the bit rate usually supported by typical low-bandwidth channels is generally so low that only the first few more significant bits of a DCT coefficient can be encoded and transmitted. To address the above problems, all the bits of the 64 DCT coefficients in an $8 \times 8$ block are here assigned importance levels that involve the size (i.e., bit rate) of the resulting scan and the scan's contribution to reducing the reconstruction distortion, expressed in terms of the sum of squared errors. More specifically, each bit in a DCT coefficient is assigned a rate-distortion ratio, with numerator the reduction in reconstruction distortion, and denominator the scan's size in bits or bytes. The resulting value is then used to associate with each bit a priority level. Taking into account the general PJPEG encoder structure, the prioritized bits are then grouped subject to the specific PJPEG constraints, achieving high compression efficiency and low computational requirements.

### 3.1. Prioritization of Bits

Assuming 8-bit input precision, each DCT coefficient will have 11-bit precision, 10 bits for magnitude and one bit for sign. Excluding the sign bit, an $8 \times 8$ DCT coefficient block can be considered as a parallelepiped which contains $8 \times 8 \times 10$ or 640 bits. An image consists of a number of these parallelepipeds, whose bits should be encoded. The decoder assumes $8 \times 8 \times 10$ parallelepipeds of all "0"s and the "0"s are replaced by the actual values of the bits as they are decoded. Therefore, to quantify the effect of each bit on the quality of the decoded image, the distortion associated with each bit being considered "0" is calculated [1]. When the value of a bit is "0", its associated distortion will be equal to zero regardless of its position, as the decoder has already assumed all the bits to have a "0" value, and decoding a "0" bit will not reduce the reconstruction distortion. In the PJPEG coding mode, each bit of the parallelepiped should be transmitted along with the bits at the same position of all the parallelepipeds in an image. Thus, the overall distortion reduction value should be equal to the sum of the distortion reduction values associated with the bits at the same position in all the parallelepipeds in the image, that is, the overall distortion reduction $\Delta D_i$ associated with the $i$th bit is equal to

$$\Delta D_i = \sum_{j=1}^{N} (b_{i,j} 2^{p_i})^2, \tag{1}$$

where $i = 1, 2, ..., 640$, $N$ is the number of $8 \times 8$ blocks in the input image, $b_{i,j}$ is the value of the $i$th bit of the $j$th parallelepiped and $p_i$ is the position of the $i$th bit. Encoding of bits at different locations in the parallelepipeds also results in different scan sizes or bit rates. Thus, a rate-distortion optimized importance measure should involve the total distortion and bit rate values, and an appropriate importance measure $I_i$ is then given by

$$I_i = \frac{\Delta D_i}{\Delta R_i}, \tag{2}$$

where $I_i$, $\Delta D_i$ and $\Delta R_i$ are the priority value, distortion reduction and scan size (or bit rate increase) associated with the $i$th bit, respectively. In this manner, those bits that contribute the most distortion reduction while requiring the least number of bits (smallest scan size) are assigned the highest priority values. Based on the priority value defined in (2), all the $640$ bits in a $8 \times 8$ block can be sorted in terms of decreasing order of priority values. Unfortunately, however, the order obtained above cannot be used as-is if PJPEG compliance is to be maintained. PJPEG requires that, for a specific bit of a DCT coefficient to be encoded/decoded, all the more significant (in terms of bit position in the binary representation) bits of the subject coefficient be encoded/decoded first. The order of the bits obtained above seldom satisfies this constraint. To devise a coding method which is PJPEG compliant, both the priority values and the above PJPEG syntax constraint should be considered. Taking this constraint into account, the bits should be reordered. Figure 1 shows the first 20 bits with the highest priority values plus extra bits that should be encoded according to the JPEG constraints. Notice that the priority values associated with the first few bits are very large, as compared to those of the other bits. Such a feature is well-suited for low-bandwidth channels, such as those used in wireless applications. By transmitting the bits according to the proposed prioritization method, acceptable reproduction quality is quickly obtained after receiving the first $1-4$ bits.

### 3.2. Grouping of Bits

Encoding the bit planes according to the priority values of their corresponding bits and the above PJPEG constraint is still not "optimal", as such a method does not take into account the PJPEG header and encoding structures. In PJPEG, a header is assigned to each scan, carrying information such as the number of image components and bits in the scan. If each scan contains one single bit plane, the additional header information would be excessive, offsetting most of the compression gain that would have been obtained using such a method. Moreover, encoding/decoding $640$ bit planes is not necessarily the most compression efficient method in light of the specific coding methods allowed by PJPEG. Fortunately, PJPEG allows a group of consecutive most significant bits of a coefficient and/or bits at the same position of consecutive coefficients to be encoded in a single scan.

To determine which bits should be grouped together, we begin by grouping only consecutive bits in the same DC or AC coefficient. First, it is better (in terms of compression efficiency) to combine all the bits of the DC coefficient into one scan. This is partly because 1) the first group of bits is encoded as in baseline JPEG, and 2) all subsequent scans must represent only one bit, which is sent as-is for each $8 \times 8$ block. Another reason is that only a single header is needed when one scan represents all the bits of the DC coefficients. Figure 2 shows the PSNR versus rate for 4 cases: no grouping and groupings of 2, 3, and 4 most significant bits of the DC coefficient for the image LENA. Clearly, the PSNR increases (for the same bit rate) as more and more bits are grouped, but the rate of increase levels off after grouping $2-3$ bits. Naturally there is a trade-off between compression and progressiveness, and we therefore select the best balance between the two parameters.

Determining the "best" grouping method for the DC bits is rather simple since there are only $10$ possible different cases. Moreover, the same method can be applied independently to the AC coefficients in a straightforward way. However, combining successive approximation with spectral selection, grouping DCT bits can then

[1] Since the DCT is a unitary transform, the distortion in the DCT domain is theoretically equal to that in the spatial domain.

be performed in two directions, going from most to least significant bits and in the zig-zag direction from the first AC coefficient to the highest frequency AC coefficient. Obviously, the number of possible groupings is very large. In this work, the following method for grouping of the bits is used. First, the number of the most significant bits of each AC coefficient that should be grouped together is determined using the above described method. Whenever the groups containing the most significant bits of consecutive AC coefficients have the same number of bits, the bits of the coefficients are combined to form spectral bands. The rest of the bits of the coefficients which should be sent one bit precision at a time are grouped into spectral bands to the extent PJPEG allows. When two or more bits are grouped, their distortions are added. Using the rate of the grouped bits, the priority value representing each spectral band is recalculated. Each group is then encoded into a scan, and all the scans are transmitted according to their priority values in several stages. Finally the overall PJPEG transmission algorithm can be stated as follows:

1. Calculate the priority values of all the bits of all the DCT coefficients using (2).
2. Obtain the best grouping pattern, i.e., the number of the most significant bits combined to form the first group for each coefficient.
3. Combine the bits or group of bits as suggested above to form spectral bands.
4. Recalculate priority values as required for each scan.
5. Order and transmit all the scans according to their priority values in stages.

### 3.3. Rate Control

Rate control is the ability of coding an image up to a specified rate [6]. Our goal here is to seek the best possible image reproduction quality given the user-specified bit rate. As is well know, encoders that are baseline JPEG compliant cannot achieve automatic rate control, limiting their usefulness in many important applications. Rate control has thus become a key desired feature of the emerging JPEG-2000 standard.

Despite the various syntax constraints of PJPEG, precise rate control can still be achieved. In fact, we next propose a method that allows the above proposed PJPEG compliant encoder to maintain high reproduction quality subject to the constraint of a fixed bit rate. First, the bits of the DCT coefficients are ordered according to their priority values and the PJPEG syntax constraints, as explained in the previous section. If the scan being coded is the first DC scan or one of the first AC scans, and the size in bits of the encoded image is approaching the desired maximum number of bits, the encoding is interrupted and zeros are assigned to the rest of the bits needed for the subject scan. PJPEG has a very efficient method of encoding a sequence of zeros by using an EOB (End Of Band) marker. Provision is made so that we can accommodate the code bits representing the zeros and the required markers, such that the output bit rate never exceeds the desired one. In the later scans of the DC value, the bits are encoded as-is [1], and it is thus wasteful to assign zeros, as the bit rate would not be affected. Thus, if the total number of bits required to represent the scan would result in a bit rate that exceeds the desired one, the rate control routine in the encoding process moves to the next first AC scan to satisfy the required rate as described earlier.

## 4. EXPERIMENTAL RESULTS

In this section, our experimental results are presented. Figure 3 shows the PSNR for the PJPEG encoders based on the proposed algorithm and sequential JPEG at medium and high bit rates for the $512 \times 512$ image LENA transmitted in 14 stages. As it is evident, the performance of our algorithm outperforms that of sequential JPEG. Figure 4 shows the same information at low bit rates. As a matter of fact, below 0.2 bpp, the difference between sequential JPEG and our proposed algorithm is considerable. Some of the bit rates cannot be achieved by sequential JPEG. In the medium bit rate, however, the performance of the two algorithms are comparable. Note that sequential JPEG is designed for such medium bit rates.

The performance of the proposed constrained rate control scheme is shown in Table 1. The desired rates are selected to be 0.0625, 0.125, 0.25, 0.5 bpp. This table also shows the obtained bit rates. As it is evident from the table, using the above described rate control algorithm, the obtained bit rates are very close to the desired bit rates. The PSNR of the corresponding bit rates are qouted in Table 2.

## 5. CONCLUSIONS

The progressive mode of JPEG can perform well in many image transmission applications when designed efficiently due to the flexibility that the standard allows in designing PJPEG encoders. We developed a rate-distortion optimized PJPEG encoder that generates a sequence of ordered scans. To do this, the bits of the image in DCT domain are ordered based on their priority values. These bits are then grouped based on the design algorithm which provides a good balance between compression efficiency and progressiveness resolution. We also devised a PJPEG compliant rate-control algorithm by adding a simple control routine to the encoding process. Our encoder outperforms baseline JPEG substantially at low and high bit rates and achieves precise rate-control.

## 6. REFERENCES


[1] ISO/IEC, *Digital compression and coding of continuous tone still images: requirements and guidlines*. International Standard Organization, 1994.

[2] K. R. Rao and J. J. Hwang, *Techniques and Standards for Image, Video and Audio Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[3] W. Pennebaker and J. Mitchell, *JPEG Still Image Compression Standard*. New York: Van Nostrand Reinhold, 1993.

[4] G. K. Wallace, “The jpeg still picture compression standard,” *Communications of the ACM*, vol. 34, pp. 30–44, Apr. 1991.

[5] A. Said and W. A. Pearlman, “A new fast and efficient image codec based on set partitioning in hierarchical trees,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, pp. 243–250, June 1996.

[6] T. Chiang and Y. Q. Zhang, “A new rate control scheme using quadratic rate distortion model,” *IEEE Transactions on Circuit and Systems for Video Technology*, vol. 7, pp. 246–250, Feb. 1997.

| Desired bit rate (bpp) | Obtained bit rate (bpp) |
|---|---|
| 0.0625 | 0.0618 |
| 0.125 | 0.125 |
| 0.25 | 0.249 |
| 0.5 | 0.499 |

Table 1: Rate Control Results.

| Obtained bit rate (bpp) | Obtained PSNR |
|---|---|
| 0.0618 | 24.74 |
| 0.125 | 26.43 |
| 0.249 | 29.14 |
| 0.499 | 32.42 |

Table 2: PSNR for rate control.

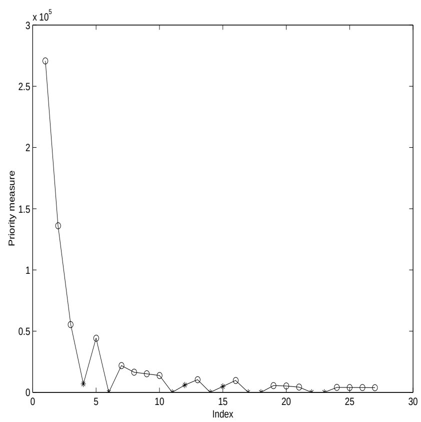


Figure 1: First 20 bits in descending order of priorities for the $512 \times 512 \times 8$ image LENA in DCT domain satisfying the PJPEG constraints

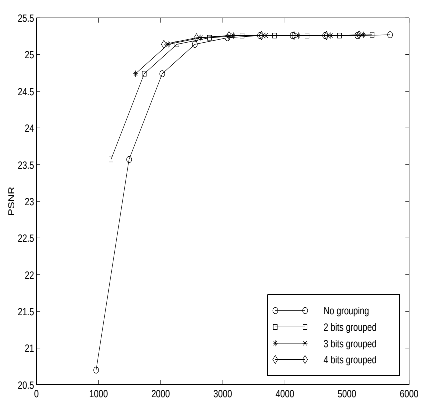


Figure 2: PSNR for the first 4 groupings of DC coefficient using successive approximation.

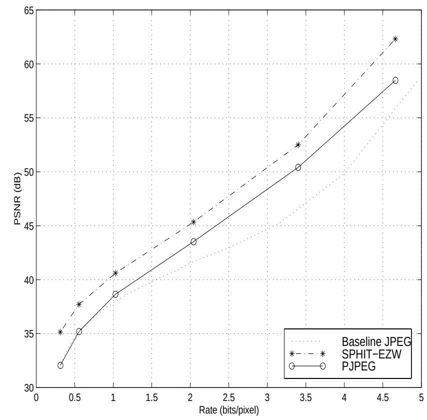


Figure 3: Comparison of the proposed algorithm with the sequential JPEG and SPHIT-EZW.

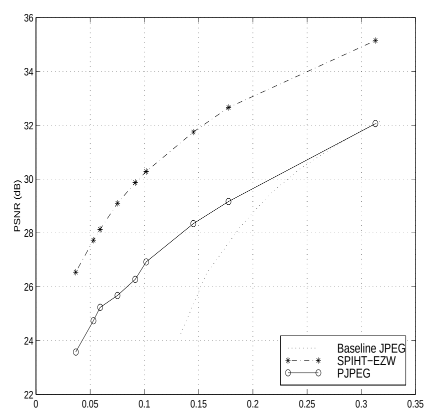


Figure 4: Comparison of the proposed algorithm with the sequential JPEG and SPHIT-EZW at low bit rates.