Comparison of Priority Partition Methods for VBR MPEG

Brian DeCleene, Pramod Pancha, Magda El Zarki, Henrik Sorensen
Department of Electrical Engineering,
University of Pennsylvania
Philadelphia, PA 19104
Tel.: (215)898-9781

Abstract
Successful exploitation of prioritized ATM networks requires matching the characteristics of the coded source with the unique properties of the network. This is particularly true for video which is very bandwidth intensive. In this paper we describe and investigate the performance of three prioritization schemes in terms of the traffic characteristics, SNR, and visual quality. These schemes are further designed such that the high priority data stream is MPEG-1 compliant and only a single pass through the encoder is required. A technique for choosing parameter values which achieve a specified high priority traffic bandwidth is presented for two of these schemes. The technique is demonstrated to be remarkably accurate for different video sequences.

1.0 Introduction
Advances in optical transmission technology have led to the development of fast packet switched networks. In these networks, Asynchronous Transfer Mode (ATM) is expected to achieve dynamical allocation of bandwidth according to the needs of the service. This in turn allows the network to support a larger number of services through statistical multiplexing. A consequence of using ATM is that, during times of congestion, packets will be delayed, and occasionally dropped.

Rather than randomly dropping packets during network congestion, priority scheduling schemes may be used to alleviate the problem by delaying or dropping non-essential packets so that the overall service performance is only minimally degraded. This can take the form of detecting network degradation before the onset of congestion and then taking appropriate actions. Policing schemes, on the other hand, try to eliminate the conditions that cause congestion. In the first approach, prevention of congestion is the user's responsibility; while with policing schemes, the network assumes this responsibility.

While policing and congestion notification schemes try to avoid and alleviate the occurrence of congestion respectively, prioritization schemes specify how the network should behave if congestion occurs. Prioritization allows the coder to specify the relative importance of the various packets to the network so that only less important packets are lost. Thus, a service can ensure a minimum quality of service on a prioritized network.

For network transmission, the traffic characteristics of each priority stream must be able to conform to the requirements that the network control algorithms place on them. This is especially vital for the high priority (HP) stream since losses in this traffic component will result in an unacceptable degradation in the quality of service. It is therefore important for the HP traffic to stay within the bounds allowed by the network's admission control algorithm in order to guarantee delivery of these cells. For example, if a constant bandwidth virtual circuit is reserved for video transmission, then the optimal traffic characteristics would require that the HP traffic generated by the encoder be less than this reserved bandwidth. The ability to adjust the parameters of the prioritization mechanism to generate the desired traffic characteristics while maintaining good subjective visual quality is therefore the final goal for any priority partition scheme.

Besides producing particular bit rate characteristics, the priority mechanism must identify which components of the source are most crucial to the end-user. We shall call this ranking the "priority partition." Typically, since entropy coding obscures the important characteristics of the video data, prioritization must occur before variable length encoding. Further, if entropy coding is performed without considering the specifics of the priority partition, then there may be a loss of coding efficiency. Thus, the performance of a prioritization scheme relies considerably on the type of coding algorithm utilized.

A number of researchers have proposed two-layer priority partitions which first divide the video data into high and low priority (LP) information and then input this data into a pair of distinct video encoders. Similarly, layered approaches require a number of successive passes over each frame of data refining the picture[1]. These proposals

This work was supported in part by the National Science Foundation under grant NCR-9016167.

5d.4.1
have the advantage that the division between priority partitioning and encoding is distinct. However, in these approaches, each priority level may require its own encoder and decoder pair resulting in a system that is computationally complex.

An alternative approach is to exploit the characteristics of the encoder to partition the data. This approach is motivated from the observation that lossy encoding algorithms inherently attempt to confine losses to the less important features of the data. Hence, lossy encoders naturally “prioritize” the data. Using this approach, the overall complexity of the system is greatly reduced and prioritization can be performed using only one pass through the encoder.

In this paper we consider partitioning techniques which are suitable for implementation within encoders that are compatible with the MPEG1 algorithm. This class of encoders is primarily characterized by its utilization of transform coding based on the DCT2. In addition, motion compensated prediction techniques are utilized to remove temporal redundancies in the visual sequence. A detailed description of MPEG and the characteristics of the variable bit rate encoder utilized in this study are available in [2],[3], and [4].

The paper is organized as follows: The next section introduces and describes three priority partitions based on the frequency and energy content of each block. Section 3 compares the traffic characteristics of these priority partitions in order to gain insights into their impact on ATM networks. The visual performance of these schemes is described in Section 4. Finally in Section 5, we describe a method by which parameters for two of the partition schemes can be chosen in order to satisfy a bandwidth constraint on the HP stream.

### 2.0 Priority Partitions

In this section we define a number of video priority partitions. Naturally, since the MPEG algorithm does not specify how multiple priorities should be transmitted and reconstructed, these algorithms cannot be fully MPEG compliant. Consequently, our algorithms satisfy a modified condition of compliance. For all of the approaches considered here, the HP data stream can be decoded using any MPEG compliant decoder and will maintain a minimum acceptable quality of service.

The MPEG algorithm defines a number of distinct coding layers which encompass different stages of coding. The lowest two of these layers, macroblock and block, perform the majority of the true compression. Each macroblock is defined as a 16x16 region of the image which is subdivided into 8x8 blocks. If the macroblock is predictive or interpolated, then motion vectors describe the best estimate of the current macroblock from previous (and/or future) macroblocks. These motion vectors are always assigned to the HP data stream. The data not accounted for in the motion estimate is coded within each of the blocks. Each 8x8 block of the frame is initially transformed using a DCT. The result is quantized component-by-component. After quantization, the 8x8 array is arranged into a linear vector going from low frequencies to high frequencies.

We shall introduce the priority partitions at the block level. This is illustrated in Figure 1. By considering information from the previous stages of the block coding, the priority partition makes a decision upon which components should be designated HP and which should be designated LP. In order to minimize the mean squared error (MSE), the priority partition decision should be performed based on the coefficient values of the output of the DCT as it provides an accurate representation of a frequency component's importance. However, we use the quantized DCT coefficient values for our decision rules for two reasons. First, the bit rate is impacted by the quantized array and not the actual values generated by the DCT. Secondly, the quantizer matrix provides an additional bias based on subjective experiments towards the lower frequency DCT coefficients. Since our aim is to achieve the best quality-of-service (as opposed to minimal MSE) for a given bit rate and to simultaneously exploit the inherent “prioritization” within the encoder, our priority partition decision rules are usually based on quantized DCT values. Finally, since the MPEG algorithm uses DPCM coding for the DC coefficient, all of the priority partitions presented here include this DC value within the HP data stream to prevent accumulation of any DC error. Therefore, the partitions presented here only control the bits generated by the AC components of each block. Hence, in our prioritization schemes even if we assign everything to LP, there is still a minimal HP data stream consisting of the block DC components and the headers of all layers above the macroblock.

### 2.1 Frequency Truncation (FT)

This priority partition is based on the observation that low frequencies are more visually important than the high frequencies of an image. Therefore, we shall identify a certain number of low frequency components as HP and

---

1. Motion Pictures Expert Group
2. Discrete Cosine Transform
transmit the remaining information as LP. Although sophisticated systems such as subband filter banks can be implemented to identify these high and low frequency components [5][6], we shall rely on the block DCT used in the MPEG algorithm to analyze the frequency spectrum of the signal. As a result, the first \( \beta_{FT} \) AC coefficients within each encoded block are defined as HP data stream while the remaining set is sent LP. This is illustrated in Figure 2.

Of the priority partitions introduced in this paper, frequency truncation is the least computational intensive since the zigzag scanning of the DCT coefficients rearranges the elements of each block into a linear vector with the lowest frequency components at one end and the higher frequency components at the other. Thus, the HP implementation simply requires zeroing out all of the coefficient beyond \( \beta_{FT} \). The LP traffic is generated by zeroing out all the DCT coefficient values that are located at or below this threshold, \( \beta_{FT} \).

2.2 Minimum Distortion (MD)

Although the frequency truncation priority partition works well for still images, motion compensation and differential coding within the MPEG algorithm generate a large fraction of high frequency components with significant energy. We would therefore like to select the \( \beta_{MD} \) coefficients with the highest energy and transmit them as HP as shown in Figure 2. Since this yields the minimum MSE between the reconstructed data and the actual data for a given number of non-zero coefficients, this partition is optimal in the distortion sense.

Unfortunately, this simple interpretation of the minimum distortion algorithm is complicated when we use the quantized DCT values to select the largest elements. If the quantizer value is not constant for all elements of the array (e.g. when using the recommended quantizer matrix for intraframes), then we may not be actually selecting the “best” values. Despite the decreased optimality in terms of the MSE, using the quantized values provides better control of the HP bit rate and only modestly impacts the overall quality. All the results presented here attributed to the minimum distortion technique were calculated using the quantized DCT values.

Figure 2: Priority Partition schemes. The priority partition separates the zig-zag vector into two components. This is illustrated for the three techniques presented in this paper with \( \beta_{MD} \) D stands for the DC component.

Implementation of the minimum distortion technique is more computationally complex than frequency truncation. For a given array, the top \( \beta_{MD} \) values must be determined. This can be implemented using an array of size \( 2^N \) to track the peak values and their corresponding indexes using only a single pass. Clearly, for small thresholds, the complexity is small. However, as the threshold increases, the problem becomes one of sorting all of the quantized DCT values and the time to compute these values increases as well.

2.3 Energy Threshold (ET)

A different approach to achieving a priority partition, which does not involve selecting a fixed number of coefficients to send per block, is to transmit all coefficients which exceed a fixed energy threshold as HP. Clearly, this causes the number of coefficients transmitted per block to vary so that we define \( \beta_{ET} \) as the average number of non-zero coefficients transmitted per block. This technique is similar to progressive priority coding which has been demonstrated to achieve superior performance for truncating the data stream in a fixed bit-rate H.261 encoder [7]. In addition, the computational complexity of thresholding is comparable to truncation.

Like minimum distortion, elements can be selected on the basis of either the original DCT array or the quantized array. If the quantized values are used, then the selected components may not actually represent the components with the largest energies. More importantly, the dynamic range of values is much smaller after quantization. Therefore, in order to provide accurate control of the bit rate, energy thresholding, unlike the other techniques, uses the unquantized DCT coefficients to select terms to transmit as HP. This is shown by the dotted line in Figure 1.

An intuitive understanding of the behavior for each of these priority partitions can be obtained by considering the distribution of blocks with a given number of non-zero coefficients. This is shown in Figure 3. As we can see, both frequency truncation and minimum distortion ensure
that the maximum number of non-zero coefficients that are transmitted is at most equal to their corresponding thresholds. Since thresholding is defined in terms of the average number of transmitted non-zero coefficients, the range of the non-zero coefficients is much larger. A difference between the minimum distortion partition and the frequency truncation is that rate distortion maps all the blocks with more than $\beta_{MD}$ coefficients to exactly $\beta_{MD}$ while truncation distributes the blocks exceeding $\beta_{PT}$ throughout the allowed range. It is also interesting to note that the minimum distortion partition exactly follows the full transmission distribution for values less than $\beta_{MD}$. That is, any blocks with fewer than $\beta_{MD}$ non-zero coefficients are transmitted in their entirety unlike truncation and energy thresholding. In the next sections, we shall investigate the properties of these three prioritization schemes at a fixed average HP bit rate.

The empirical results that are presented here were gathered from our database of video sequences. Included in this set were 512x480 pixel movie segments of 200 frames from “Indiana Jones and the Temple of Doom” and “League of Our Own” operating at a frame rate of 24 frames/s as well as popular 352x240 pixel sequences such as “Table Tennis”, “Flower Garden”, and “Bike” with a frame rate of 30 frames/s. In addition, all sequences were coded with motion compensation. It should be noted that although longer sequences will result in higher confidence levels for parameter values, our purpose of the next two sections is to portray the general characteristics of each priority partition. We have therefore restricted our attention in the next two sections to the first 80 frames from the segment of “League of Our Own.”

It is important to note that although the $\beta$’s share common definitions, we do not require or expect that $\beta_{PT} = \beta_{MD} = \beta_{ET}$ will achieve the same HP bit rate. For example, to achieve an overall bit rate of approximately 3.6 Mbps for 80 frames of the movie “League of Our Own”, it was necessary to set $\beta_{PT}=17$, $\beta_{MD}=5$, and $\beta_{ET}=4.4$. The difference in the overall bit rate for equivalent $\beta$’s can be attributed to a couple of observations. First, as can be seen in Figure 3, the distribution of non-zero coefficients is very different for each of the schemes. It follows that they compress differently. Along the same lines, the MPEG coding algorithm not only entropy codes the non-zero values but also compresses the runs of zeros between the values. Consequently, just as the distribution of values is different for each of the priority partitions, so is the distribution of runs of zeros. The combination of these effects forces us to select distinct thresholds for each priority partition method in order to make relevant comparisons.

Finally, it should be noted that the compression efficiency of the prioritized coder will always be lower than that of a similar non-prioritized coder. We noted that when we encode the LP block data using the MPEG variable length code tables the loss in efficiency caused by prioritization is relatively constant for different thresholds for each scheme. This efficiency loss was approximately 10-15% for frequency truncation and 20% for minimum distortion and for energy threshold.

### 3.0 Packet Distributions

A priority partition is only successful if the network can maintain the integrity of the HP channel. This objective is only possible if the temporal characteristics of the generated packets are well behaved. From the view point of the network designer, the packet distribution and auto-correlation measures can help assess the expected performance of a given traffic source in an ATM environment. Therefore, in this section, we investigate the distribution of cells per slice and the autocorrelation function of both cells per slice and cells per frame.

The cells per slice statistics for the priority partition schemes at identical bit rates are shown in Table 1. Upon examining the average number of cells per slice, we can observe that the frequency partitioning algorithm results in a wider disparity between the average rates in intraframes and predictive frames than the other algorithms. This can be explained by the fact that encoding of intraframes is highly biased towards transmission of the low frequency content of the image through the use of a nonuniform quantizer matrix for the DCT coefficients. In predictive frames, a uniform quantizer matrix is used for all DCT coefficients and, as a result, this bias does not exist. Since the threshold and minimum distortion schemes utilize the magnitude of coefficient values as their selection criterion, they are less affected by the difference in frequency content of intraframes and predictive frames. We should also note that threshold partitioning results in the largest variance of slice sizes of all the partition schemes. The reason for this large variance is that the nature of the threshold

<table>
<thead>
<tr>
<th>Table 1. Cells/slice statistics at fixed HP bit rate. ($\beta_{PT} = 17$, $\beta_{ET} = 4.4$ and $\beta_{MD} = 5$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency (Intra / Pred.)</td>
</tr>
<tr>
<td>---------------------------</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Variance</td>
</tr>
<tr>
<td>Maximum</td>
</tr>
<tr>
<td>Minimum</td>
</tr>
</tbody>
</table>

3. We shall define a slice as a horizontal strip of macroblocks across a frame.
partition algorithm causes it to select a large number of non zero coefficients when energy in a block is large and evenly spread out while at the same time it will select few coefficients if the coefficient values fall below the threshold.

A potentially useful observation is that the minimum distortion algorithm has a significantly lower variance than the other two algorithms. Further, the maximum number of cells per slice is also substantially lower. As a result, traffic burstiness, which is defined as the peak to average rate ratio, is significantly lower for the this method (1.5 vs. 2.2 for threshold and 2.0 for frequency partitioning algorithms). This is a desirable quality in the traffic behavior since bursty traffic is usually subjected to larger losses in networks.

Figure 4 corroborates the observations on the cell generation process that were made from the statistics in Table 1. For example, it can be clearly be seen that the threshold partitions lead to the greatest spread in the cells per slice distribution as mentioned above. The rate distortion partition method appears to have a very attractive property in that its range of cells generated per slice is very narrow and that there is a sharp cutoff in the distribution; values less than 18 cells/slice and values greater than 32 cells/slice rarely occur. Since the observed range of cell generation rates is fairly small this makes it easier to support such a stream utilizing a quasi-constant bit rate channel.

Finally, the autocorrelation functions shown in Figure 5 depict the time behavior of the traffic generated by each partition method. In terms of the cells/slice, the correlation between slices in the same position in adjacent frames was found to be high for all schemes. On the other hand, the autocorrelation behavior of cells per frame when comparing the three schemes shows some interesting differences. We observe that the energy threshold and minimum distortion partitions tend to mask the burstiness effect of the intraframes when compared to the frequency truncation scheme. In other words, by choosing the priority partition

![Figure 4. Cells/Slice distribution for predictive frames for a fixed bit rate.](image)

![Figure 5. Autocorrelation of (a) Cells/frame (b) Cells/Slice based on coefficient values rather than by frequency truncation, the traffic characteristics](image)

4.0 Quality of Service

While the packet distributions are relevant to the network, the ultimate goal is to maintain the best minimum quality of service within the bandwidth constraints. Ideally, the priority partition should generate a low HP bit rate; yet maintain minimal distortion. Setting aside issues of computational complexity and packet distribution, the scheme that best achieves these minimum distortion goals is the most favorable. In this section, we will compare the three schemes in terms of their mean-square-error (MSE) as well as their subjective visual quality.

For a given bit rate, the corresponding average frame SNR is shown in Figure 6. Clearly as the bit rate for all of the priority partitions approach the rate required for full transmission, they converge to the same point. We also see that truncation and minimum distortion both result in comparable performance with minimum distortion performing slightly better for intraframes. The implication is that the block DCT efficiently compresses most of the energy into the low frequency components. Therefore, the truncation priority partition and the minimum distortion priority partition generally select the same set of coefficients per block.

Perhaps the most striking feature of Figure 6 is that thresholding generates the best video in terms of the SNR at a given bit rate. For example, for the first 80 frames of "League of Our Own" at the bit rate of 3.6 Mb/s used in the last section, the intraframe SNR for the HP data stream
using thresholding is around 32dB compared to 28dB for truncation and minimum distortion. This 4dB gain is a consequence of the fact that thresholding is not constrained with a hard limit upon the number of non-zero coefficients per block. Thus, thresholding concentrates more of its limited bandwidth where the block energy is significant.

We tested for visual quality by examining the quality of the reconstructed frames of the HP stream for each of these schemes on a frame by frame basis. A preliminary observation that we made was that the correlation between the results using MSE and that of visual quality was quite poor. At the same bit rate, the visual quality of the truncation was subjectively superior to the other priority partitions for most predictive frames. The quality of the reconstructed intraframes were comparable for all partitions. We attributed this to greater continuity of the frequency components between adjacent blocks with frequency truncation which resulted in less blocking in the reconstructed video. This can also be attributed to the fact that the MPEG algorithm compresses the truncated data stream more efficiently than the other priority partitions. Hence, a greater number of non-zero coefficients could be transmitted for the frequency truncation at the same bit rate. For example, at 3.6 Mb/s, an average of 6.7 non-zero intraframe coefficients/block were allocated to HP for truncation compared to 4.4 and 4.0 for thresholding and minimum distortion respectively. Another factor is that with thresholding and minimum distortion high frequency error can accumulate. Since truncation never transmits these coefficients, the high frequency error remains constant on average and we notice that the visual quality does not degrade as much with each successive predictive frame.

5.0 Bandwidth Allocation

In the previous sections, we studied the behavior of the proposed priority partitions at the same overall HP bit rate. Controlling the HP bit rate poses a difficult challenge given the non-stationarity of video. Yet, the ability to estimate and control the bandwidth required by the HP traffic is fundamental to the overall success. For the techniques outlined in the previous section, this task amounts to selecting the appropriate threshold such that the generated HP traffic satisfies certain network constraints. In this section, we propose a technique for selecting the appropriate threshold such that a given fraction of the overall bit rate generated from the AC components is designated as HP. We will designate this function as $F(\beta)$.

From the definitions, a couple of observations about $F(\beta)$ can be made. Clearly, $F(\beta)$ should be a strictly increasing function since increasing the threshold allows more non-zeros coefficients through. In addition, we note that when the threshold is zero, then all of the AC components are designated as LP (i.e. $F(0) = 0$). Also, as was observed in Figure 3, the maximum number of non-zero coefficients transmitted for frequency truncation and minimum distortion is constrained by the corresponding threshold. Therefore, if the last non-zero element of a given block occurs at index $n_{PT}$, then selecting $\beta_{PT} \geq n_{PT}$ will allocate all of the AC coefficient to HP. So, $F(\beta_{PT}) = 1$ when $\beta_{PT} \geq n_{PT}$. Analogously, when $n_{MD}$ represents the number of non-zero coefficients within the block, then the same condition holds true for minimum distortion.

A study of this function for a collection of video sequences has revealed that a similar relationship holds for frequency truncation and minimum distortion. An example is shown in Figure 7. In particular, if $n$ is defined as $n_{PT}$ or $n_{RD}$ according to the prioritization technique, then set

$$x = \min(m_n + \sigma_n, 63)$$

(EQ 1)

where $m_n$ and $\sigma_n$ are the mean and standard deviation of $n$ respectively. The fraction of excess bits allocated to HP is approximately,

$$F(\beta) \approx \begin{cases} 
\min(\beta / x, 1) & x \leq 8 \\
A_x [1 - \exp(B_x \beta)] & \text{otherwise}
\end{cases}$$

(EQ 2)

Note that both $A_x$, $B_x$ are a function of $x$ and given in Table 2 at the end of this document. This function satisfies all of the constraints described in the previous section and, as is evident from the figure, this approximation fits the experimental data well. Also, it reveals that bit rate is
more sensitive at low thresholds: a small error can mean a large deviation in the expected bit rate.

From this approximation, an algorithm for controlling the excess bit rate is easy to derive. For each picture, the mean and standard deviation of the appropriate \( n \) is evaluated during encoding. Separate values are kept for the intra macroblocks and the predictive macroblocks. This may be easily accomplished by maintaining a running sum of \( n \) and its square along with a counter of the number of blocks. At the end of the frame, the variable \( \chi \) is calculated and from this, the appropriate threshold is selected to maintain a given excess fraction of bits. By reevaluating \( \beta \) at the picture level, we are able to track the non-stationary nature of arbitrary video sequences.

Figure 8 illustrates the performance of this bandwidth allocation scheme for 100 frames of the video "Indiana Jones and the Temple of Doom" and "League of Our Own." As we can see, the minimal bit rate was 0.68 Mbit/s and 0.74 Mbit/s for Indiana and League respectively. This minimal quality consists of the DC component for each block along with the motion vectors. By using the threshold of the previous frame for the current frame, we were able to linearly scale the additional data between this minimum level up to the full video. One advantage of this approach is that it is scalable. For example, while the overall bit rate for both video sequences was different by about 4.5 Mbit/s, the relationship remained linear for both. Therefore, we can accurately control the quality and the bit rate of an arbitrary video sequence. It is also interesting to note that both of the videos had the same minimal quality bit rates. This demonstrates that by controlling the AC coefficients alone, the video bit rates can be accurately adjusted from a common level. We have also verified the validity of (EQ 1) and (EQ 2) on other video sequences such as "Table Tennis" and "Flower Garden." With all of these, the relationships remained valid.

Figure 8: Theoretical vs Experimental Bandwidth Usage. The observed proportion of excess bits over the minimal quality matches nicely. The top figure represents "Indiana Jones and the Temple of Doom." The bottom figure is for the movie "League of Our Own."

Although we controlled the threshold at the picture layer, the video can also be controlled at other levels as well. For example, very tight control of the bandwidth can be achieved by monitoring the threshold on a block-by-block level. In order to implement this, however, the variable \( x \) needs to be redefined as the mean of \( n \). Also since curves are not as correlated at the block layer as they are at the picture level, the threshold should be calculated first and then applied to the current block. Because of the additional cost in computational complexity, we deemed that picture level tracking was sufficient and did not interfere with the advantages of the various packet distributions significantly.

While it is possible to accurately control the bandwidth for both frequency truncation and minimum distortion, energy thresholding is much more sensitive to different types of video. As can be seen in Figure 7, energy thresholding increases more rapidly than both frequency truncation and minimum distortion and saturates at a much lower value. This produced a very small range of thresholds that actually controlled the performance of the
encoder. Furthermore, the definition of \( \beta_{ET} \) as the average number of non-zero coefficients transmitted makes it difficult to calculate until a full analysis of the frame is performed. Combining these problems, energy thresholding was very difficult to control.

Rather than controlling the bandwidth allocation as the coding occurs, all analysis was performed first and then the appropriate values fed to the encoder. For a given video source, a table was generated based on the empirical distribution of the coefficient energies within each unquantized DCT block. Recall that unlike the previous techniques, we are forced to use the unquantized data in order to assure a larger range of values. Based on this table, a threshold was selected to ensure that a given percentage of non-zero coefficients was transmitted. As expected, a threshold of zero assigns all of the coefficients to HP and hence, \( \beta_{ET} \) is 64 (the number of coefficients in an 8x8 block). On the other hand, if the threshold is sufficiently large, no components are assigned to the HP data stream. Once a threshold was selected, it was fixed for the entire coding process.

6.0 Summary

We investigated the performance of three priority partitioning schemes; frequency truncation, minimum distortion and energy threshold, for a variable bit rate MPEG compatible encoder. One constraint of using this particular coding algorithm is that the bit stream resulting from priority partitioning must still be MPEG compatible.

The source traffic characteristics for each of these three methods were also examined. In terms of the statistics of the cells generated per slice, the minimum distortion method appears to have a significant advantage over the other methods. The traffic characteristics in interframes and intraframes are very similar and the traffic burstiness is significantly lower than for the other schemes.

We discovered that in terms of minimum distortion performance at a fixed bit rate, the energy threshold method outperforms the frequency truncation and minimum distortion schemes. Yet, when the subjective visual quality of the three methods was compared, the quality of video generated by the frequency truncation scheme was visually more appealing than the other schemes for the medium bit rates while at high bit rates there was, as expected, virtually no difference in the three schemes.

Finally, a control technique for managing the average excess bits allocated to HP is demonstrated for the frequency truncation and minimum distortion schemes. It is shown that this technique works extremely well and is remarkably insensitive to the input sequence characteristics. The difficulties of tracking energy thresholding is discussed.

References


Table 2 : Parameters for \( F(y) \) of EQ 2

<table>
<thead>
<tr>
<th>x</th>
<th>( A_x )</th>
<th>( B_x )</th>
<th>x</th>
<th>( A_x )</th>
<th>( B_x )</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>1.0001</td>
<td>-0.2142</td>
<td>37</td>
<td>1.0195</td>
<td>-0.0773</td>
</tr>
<tr>
<td>10</td>
<td>1.0001</td>
<td>-0.2134</td>
<td>38</td>
<td>1.0216</td>
<td>-0.0747</td>
</tr>
<tr>
<td>11</td>
<td>1.0001</td>
<td>-0.2126</td>
<td>39</td>
<td>1.0239</td>
<td>-0.0722</td>
</tr>
<tr>
<td>12</td>
<td>1.0002</td>
<td>-0.1958</td>
<td>40</td>
<td>1.0265</td>
<td>-0.0697</td>
</tr>
<tr>
<td>13</td>
<td>1.0003</td>
<td>-0.1856</td>
<td>41</td>
<td>1.0294</td>
<td>-0.0673</td>
</tr>
<tr>
<td>14</td>
<td>1.0004</td>
<td>-0.1782</td>
<td>42</td>
<td>1.0326</td>
<td>-0.0649</td>
</tr>
<tr>
<td>15</td>
<td>1.0006</td>
<td>-0.1681</td>
<td>43</td>
<td>1.0362</td>
<td>-0.0625</td>
</tr>
<tr>
<td>16</td>
<td>1.0008</td>
<td>-0.1608</td>
<td>44</td>
<td>1.0404</td>
<td>-0.0600</td>
</tr>
<tr>
<td>17</td>
<td>1.0011</td>
<td>-0.1528</td>
<td>45</td>
<td>1.0451</td>
<td>-0.0576</td>
</tr>
<tr>
<td>18</td>
<td>1.0014</td>
<td>-0.1465</td>
<td>46</td>
<td>1.0506</td>
<td>-0.0552</td>
</tr>
<tr>
<td>19</td>
<td>1.0017</td>
<td>-0.1414</td>
<td>47</td>
<td>1.0569</td>
<td>-0.0527</td>
</tr>
<tr>
<td>20</td>
<td>1.0021</td>
<td>-0.1359</td>
<td>48</td>
<td>1.0644</td>
<td>-0.0502</td>
</tr>
<tr>
<td>21</td>
<td>1.0025</td>
<td>-0.1312</td>
<td>49</td>
<td>1.0734</td>
<td>-0.0476</td>
</tr>
<tr>
<td>22</td>
<td>1.0030</td>
<td>-0.1265</td>
<td>50</td>
<td>1.0843</td>
<td>-0.0450</td>
</tr>
<tr>
<td>23</td>
<td>1.0035</td>
<td>-0.1223</td>
<td>51</td>
<td>1.0977</td>
<td>-0.0423</td>
</tr>
<tr>
<td>24</td>
<td>1.0041</td>
<td>-0.1181</td>
<td>52</td>
<td>1.1146</td>
<td>-0.0394</td>
</tr>
<tr>
<td>25</td>
<td>1.0047</td>
<td>-0.1144</td>
<td>53</td>
<td>1.1363</td>
<td>-0.0365</td>
</tr>
<tr>
<td>26</td>
<td>1.0055</td>
<td>-0.1103</td>
<td>54</td>
<td>1.1653</td>
<td>-0.0333</td>
</tr>
<tr>
<td>27</td>
<td>1.0063</td>
<td>-0.1067</td>
<td>55</td>
<td>1.2052</td>
<td>-0.0299</td>
</tr>
<tr>
<td>28</td>
<td>1.0071</td>
<td>-0.1035</td>
<td>56</td>
<td>1.2631</td>
<td>-0.0263</td>
</tr>
<tr>
<td>29</td>
<td>1.0081</td>
<td>-0.1000</td>
<td>57</td>
<td>1.3530</td>
<td>-0.0224</td>
</tr>
<tr>
<td>30</td>
<td>1.0091</td>
<td>-0.0969</td>
<td>58</td>
<td>1.5081</td>
<td>-0.0180</td>
</tr>
<tr>
<td>31</td>
<td>1.0102</td>
<td>-0.0939</td>
<td>59</td>
<td>1.8270</td>
<td>-0.0130</td>
</tr>
<tr>
<td>32</td>
<td>1.0115</td>
<td>-0.0908</td>
<td>60</td>
<td>2.7872</td>
<td>-0.0072</td>
</tr>
<tr>
<td>33</td>
<td>1.0128</td>
<td>-0.0880</td>
<td>61</td>
<td>11.1110</td>
<td>-0.0015</td>
</tr>
<tr>
<td>34</td>
<td>1.0142</td>
<td>-0.0853</td>
<td>62</td>
<td>11.1110</td>
<td>-0.0015</td>
</tr>
<tr>
<td>35</td>
<td>1.0158</td>
<td>-0.0826</td>
<td>63</td>
<td>11.1110</td>
<td>-0.0015</td>
</tr>
<tr>
<td>36</td>
<td>1.0176</td>
<td>-0.0798</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>