Generalized Digital Pulse Code Modulation and its comparison with other compression schemes

Tarun Kumar Bhardwaj

ELECTRONICS AND COMMUNICATION DEPT, NETAJI SUBASH INSTITUTE OF TECHNOLOGY
DWARKA, DELHI 110053

Abstract: Image Compression is a process of removing redundant pixels from an image. There are various Image Compression Techniques available. Predictive Coding is one of the basic Image Compression Techniques. In Predictive Coding Pulse-code modulation (PCM) is a basic technique for image compression. In case of PCM the rate of the bit stream is simply reduced by removing a fixed number of least significant bits from each codeword so PCM coding technique is extremely simple but it has a poor coding efficiency. Another Predictive Coding technique is known as the differential pulse code modulation (DPCM). A DPCM removes pixel correlation and requantizes the residual pixel values for storage or transmission. The residual image has a variance much smaller than that of the original image. In this paper, we present a generalized DPCM (GDPCM) algorithm which is a Lossy Compression Scheme. This GDPCM technique provides better SNR with Low Complexity measure of complexity in MFLOPS. GDPCM has better SNR and coding efficiency than PCM and DPCM and it can adapt to the computational resources that are available in each application.

Keywords: predictive coding, JPEG, DPCM and complexity

1. INTRODUCTION

Images and videos are moved around the World Wide Web by millions of users almost in a nonstop fashion, and then, there is television (TV) transmission round the clock. This process of reducing the image and video data so that it fits into the available limited bandwidth or storage space is termed data compression. Data compression refers to the process of reducing the digital source data to a desired level and bandwidth compression refers to the process of reducing the analog bandwidth of the analog source. Today, most signals of interest (e.g., voice, audio, image, video) are digitally acquired (digitized) using A/D converters. A/D converters perform pulse-code modulation (PCM) with uniform quantization and fixed-length binary coding. This type of coding has several benefits such as it is the simplest source coding technique and it is trivial to locate (and decode) any codeword in the PCM bit stream (random access) and also when embedded quantization is used, the rate of the PCM bit stream can be easily reduced by discarding a fixed number of bits in each codeword (scalability). Despite all these advantages, PCM is not usually used in the storing or transmission of signals due to its poor coding efficiency. In general for images this Data compression is coined as Image Compression. There are various Image Compression Techniques available which can be broadly classified in to Lossless and Lossy Coding. In this Paper we deal with the Lossy Compression Schemes. Lossy Compression Techniques can be classified in to Predictive Coding, Transform Coding and Wavelet Coding. In lossless Image Compression Techniques Huffman Coding and Arithmetic Coding are used. One of the most commonly used Predictive Coding Technique is known as the differential pulse code modulation (DPCM) [1]. If the pixel correlation along only one dimension (row or column) is removed, then the DPCM is 1D DPCM. If the correlations along both dimensions are removed, then the resulting DPCM is known as 2D DPCM. A DPCM removes pixel correlation and requantizes the redundant pixel values for storage or transmission. The redundant image has a variance much smaller than that of the original image. The quantizer is fixed no matter how the decorrelated pixel values are interrelated. A variation on the theme is to use quantizer that adapt to changing input statistics, and therefore, the corresponding DPCM is called an adaptive DPCM. Due to limited bit width of the quantizer for the redundant image, edges are not preserved well in the DPCM. It also exhibits occasional streaks across the image when channel error occurs. In GDPCM these errors are eliminated. In case of GDPCM the input to the quantizer is adapted to always produce a unit variance differential signal. Basically, based on the previously quantized differential pixel, an estimate of the variance of the current difference pixel is obtained and it uses 2D linear prediction coding to produce the next predicted output.

2. COMPRESSION TECHNIQUES

![Various Compression Methods](image-url)

Figure 1: Various Compression Methods.
2.1.1 Predictive Coding

Image compression is achieved by removing pixel redundancies. For instance, pixels along a scan line are highly correlated. In a similar manner, pixels in a rectangular neighborhood are also highly correlated. A high correlation implies that a pixel is predictable based on the values of the pixels in its neighborhood. Therefore, one can transmit or store the pixel differences rather than the actual pixels themselves to achieve compression because the differential pixels are very nearly zero with a high probability and require less number of bits to code. This is the basis behind predictive coders. Predictive coding is used in Moving Picture Experts Group (MPEG) standards to code moving pictures.

a) DPCM

DPCM stands for Differential Pulse Code Modulation. In this case the important steps are sampling and Quantization. In this case quantization plays a key role we can use uniform as well as non uniform quantization. In case of non uniform quantization the quantization Error is low. While using DPCM prediction plays an important role it helps in predicting the value of next pixel. In this case the Signal to Noise Ratio value depends upon the prediction used. In case of 2-D prediction the value of Signal to Noise Ratio is high.

2.1.2. Transform Coding

In transform coding, a block of image pixels is linearly transformed into another block of transform coefficients of the same size as the pixel block with the hope that only a few of the transform coefficients will be significant and the rest may be discarded. A small amount of storage space is required to store only the significant transform coefficients, hence the compression take place. The original image can be obtained by performing the inverse transform of the reduced coefficient block.

a) JPEG

JPEG is an acronym for Joint Photographic Experts Group [4] and defines a set of still picture grayscale and color image compression algorithms and data transport format. In JPEG we have two compression algorithms, one is based on 2D DCT and the other is based on spatial prediction methods. The DCT-based algorithm is intended for compression quality ranging from very good to visually indistinguishable. JPEG defines four options or modes of operation, namely, sequential, progressive, hierarchical, and lossless coding. JPEG allows two entropy coders—Huffman and arithmetic coders [7]. The baseline system guarantees a reasonable level of functions in all decoders that use the DCT algorithms. It allows a restricted version of the sequential DCT algorithm for images with 8 bits per sample and Huffman coding.

2.1.3. Wavelet Coding

Compression of image data is achieved by removing the correlation among the pixels as in predictive coding or by compacting pixel energy in a few coefficients as in transform coding. This later principle is equally applicable to wavelet domain compression [8]. The main difference between transform coding and wavelet domain coding is that wavelet transform is typically applied to the whole image unlike the transform coding where the transform is applied to sub blocks. When we mention wavelet domain or wavelet transform, we mean discrete wavelet transform (DWT). In two-dimensional (2D) DWT, an image is transformed into four distinct sets of coefficients giving an approximation of the image at a lower resolution, and three details with orientations in the horizontal, vertical, and diagonal directions [8]. This is called one level of wavelet transform. In octave-band decomposition the approximation coefficients of the first level are again decomposed into four sets of coefficients to obtain a two-level 2D DWT.

a) JPEG 2000

Joint Photographic Experts Group In 2000 introduced another compression algorithm based on DWT and it became the JPEG2000 standard [9]. JPEG2000 is a recommendation that (1) specifies the decoding processes for decompressing compressed image data to reconstruct the image data, (2) specifies a code stream syntax containing information for interpreting the compressed image data, (3) specifies a file format, (4) provides guidance on the encoding processes for converting source image data to compressed image data, and (5) provides guidance on how to implement these processes in practice. We are concerned only with the basic compression algorithm used in JPEG2000 standard. The standard specifies both lossless and lossy compression algorithms. JPEG2000 uses 2D DWT as the compression vehicle.

2.2. GDPCM coding of images

GDPCM stands for Generalized Differential Pulse Coding Modulation. The basic idea behind this technique is that it gives better coding efficiency than PCM and has a low complexity. It uses a feature of DPCM. It also gives better coding efficiency than PCM and DPCM. It also removes the sensitivity issue of DPCM. GDPCM is the one in which the input to the quantizer is adapted to always produce a unit variance differential signal. Basically, based on the previously quantized differential signal, an estimate of the variance of the current difference pixel is obtained [2-3]. The quantizer input is then divided by the square root of the estimated variance so that the input difference pixel has a unit variance. Thus, the quantizer remains fixed while its input is adjusted from pixel to pixel to have a unit variance.

![Figure 2: Block Diagram of GDPCM](image-url)
In this equation which calculates the variance of the differential pixel at the quantizer input is estimated from the previous variance and the quantized differential pixel as
\[ \sigma_e^2(k) = (1-\gamma)\sigma_q^2(k-1) + \gamma \sigma_e^2(k-1) \]

Where \(0 \leq \gamma \leq 1\)

### 2.2.1. Design of a DPCM

There are two subsystems in the simple DPCM (1) the predictor [4-6] and (2) the quantizer. Thus, to design a DPCM for image compression, one must design the predictor and the quantizer. In most cases, linear prediction is used because it lends itself to analytical solution. While nonlinear prediction may prove to perform better, it is difficult to analyze and a closed form solution may not be feasible.

#### 2.2.2. Prediction

Consider a discrete sequence \(\{x[n], n = 0, 1 \ldots\}\), which may be pixels along a row of an image. If the pixels are highly correlated, then it makes good sense to estimate the value of the current pixel \(x[n]\) in terms of the previous pixels. That is to say that we can express the estimate or the predicted value of the current pixel as a weighted sum of the earlier pixels. Thus,

\[ x(n) = \sum_{m=1}^{p} a_m x(n-m) \]  

Equation describes a p-th order predictor because \(p\) previous pixels are used in estimating the current pixel value. Even though simple differencing is a degenerate case of linear prediction, it has no degree of freedom to optimize a cost function. Therefore, we will consider the general case where the predictor weights have to be determined to satisfy a chosen cost function.

For the first-order linear predictor, equation (1) reduces to

\[ x[n] = a x[n-1] \]  

In terms of the predicted pixel and the differential pixel, the actual pixel takes the form of

\[ x[n] = x[n] + e[n] \]  

Since the predicted pixel is some multiple of the previous pixel, we observe from equation (3) that the error or differential pixel and the previous pixel are orthogonal [4]. That is,

\[ E\{e[n] x[n-1]\} = 0 \]  

Equation (4) is formally written as \(e[n] \perp x[n-1]\). The prediction error is also called the innovation. Note that we are using the actual previous pixel in predicting the current pixel instead of using the previously reconstructed pixel \(x[n-1]\). This will simplify the analysis. The value of \(a\) is chosen so as to minimize a cost function. The usual cost function used in practice is the MSE between the actual and predicted pixels over the whole image. The MSE is written as

\[ MSE = E\{e^2[n]\} = E\{x[n] - x[n]\}^2 \]  

Here \(E\) denotes the expectation operation [5]

Thus, the prediction gain GP of the optimal first-order predictor is defined as the ratio of the input signal variance to that of the prediction error or differential signal

\[ SNRDPCM (dB) = SNRPCM (dB) + 10 \log GP \]  

A pixel in intensity images has correlation in both spatial dimensions. Therefore, it is meaningful to exploit the two dimensional (2D) correlation to achieve a higher compression. This may be accomplished by using a 2D linear predictor in the DPCM loop. Let the 2D first-order predicted pixel be expressed as

\[ x[m, n] = a1 x[m, n-1] + a2 x[m-1, n] + a3 x[m-1, n-1] + a4 x[m-1, n+1] \]  

2.3. Calculation of SNR

In case of GDPCM [9]

\[ SNR = 10*\log 10(\text{SigVar} / \text{std}(A1-A2)^2)) \]

\[ SNR = 10*\log 10(\text{SigVar}(\text{std}(A1[1:Height,1:Width])) - A2)^2)) \]

Here,

- \(A1\) = value of the current pixel.
- \(A2\) = value of the previous pixel.

In case of DPCM

\[ \text{SigVar} = (\text{std}(f[1:Height,1:Width]))^2 \]

\[ SNR = 10*\log 10(\text{SigVar}(\text{std}(f[1:Height,1:Width]) - y)^2) \]

Here,

- \(y\) = DPCM reconstructed image
- \(f\) = intensity image to be encoded.

### 2.4. Comparison in SNR

#### Table 1: Comparison in SNR

<table>
<thead>
<tr>
<th>S.No.</th>
<th>CODING SCHEME</th>
<th>BIT PER PIXEL</th>
<th>SNR IN dB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PCM</td>
<td>8</td>
<td>26.83</td>
</tr>
<tr>
<td>2</td>
<td>DPCM</td>
<td>4</td>
<td>23.87</td>
</tr>
<tr>
<td>3</td>
<td>GDPCM</td>
<td>3</td>
<td>29.52</td>
</tr>
<tr>
<td>4</td>
<td>JPEG</td>
<td>0.8</td>
<td>22.18</td>
</tr>
<tr>
<td>5</td>
<td>JPEG 2000</td>
<td>0.48</td>
<td>15.95</td>
</tr>
</tbody>
</table>

### 2.5. Computational Complexity

The complexity of the different versions of the different coding scheme discussed thus far are compared and listed in Table. The complexity is in terms of millions of floating point operations per second (MFLOPS). Only floating point multiplications are considered because floating point addition/subtraction is relatively much faster. The computational complexity increases with the size of the transform. In case of JPEG we use DCT which is a fast algorithm meaning that an N-point DCT

\[ \sum_{m=0}^{N-1} x(n) \cdot \sin \left( \frac{\pi m n}{N} \right) \]
can be computed in $N/2 \log_2 N$ operations; an operation here refers to multiplication. Thus, an $8 \times 8$ block DCT implemented in a row-column fashion requires 144 operations, while a $16\times16$ DCT needs 1024 operations, almost an order of magnitude increase in computations from that for $8\times8$ DCT.

2.6. Comparison in Complexity

Table 2: Comparison in complexity

<table>
<thead>
<tr>
<th>S.No.</th>
<th>CODING SCHEME</th>
<th>BPP</th>
<th>COMPLEXITY IN MFLOPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DPCM</td>
<td>4</td>
<td>7.86</td>
</tr>
<tr>
<td>2</td>
<td>GPCM</td>
<td>3</td>
<td>15.73</td>
</tr>
<tr>
<td>3</td>
<td>PCM</td>
<td>8</td>
<td>1.2</td>
</tr>
<tr>
<td>4</td>
<td>JPEG</td>
<td>0.8</td>
<td>60-100</td>
</tr>
<tr>
<td>5</td>
<td>JPEG 2000</td>
<td>0.48</td>
<td>60-100</td>
</tr>
</tbody>
</table>

3. RESULTS

- a) Comparison of uniform and non uniform quantizer
- b) Input image
- c) Unquantized differential image using GDPCM
- d) Quantized differential image using a 3-bit uniform quantizer.
- e) Reconstructed image with an SNR of 29.52 dB.
- f) Comparison in SNR of 1-D and 2-D DPCM
CONCLUSION
In this GDPCM, where the quantizer input is normalized by the factor that normalizes the quantizer input is computed recursively using a simple procedure. It is found that the GDPCM performs the best yielding an SNR of 26.80 dB using 1D prediction and 29.52 dB using 2D prediction, both at 3 bpp. The performance and complexity of the different Compression Schemes discussed thus far are compared and listed. The complexity is in terms of millions of floating point operations per second (MFLOPS). The best found to be the GDPCM, which yields better SNR at a low bit rate of 3 bpp with a complexity of around 10-16 MFLOPS. While in case of JPEG and JPEG 2000 the complexity is around 60-100 MFLOPS which is high as compared to GDPCM which gives a complexity of 10-16 MFLOPS. So GDPCM provide reasonable SNR with low Bit Rate and at low complexity.

FUTURE WORK
Using 2-D prediction in JPEG we can design a completely lossless compression Technique. This JPEG Based technique will provide coding efficiency and good Signal to Noise Ratio. The coding Complexity will increase. The numbers of operations will increases. In this case Optimum Quantizer can also be used to increase the Signal to noise ratio and coding Efficiency.

REFERENCES
[8] José Prades Nebot, Member, IEEE, Marleen Morbee, and Edward J. Delp, Fellow, IEEE “Generalized PCM coding of images” IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 21, NO. 8, AUGUST 2012