Department of Electronics and Electrical Communication Engineering

Indian Institute of Technology Kharagpur

Image and Video Processing Laboratory
(EC69211)

Mini Project : JPEG Compression

Group - 1

Ashutosh Naik (20EC39049)
Arman Atibudhi (20EC39048)

Introduction

JPEG (Joint Photographic Experts Group) is a widely used image compression standard that significantly reduces the file size of digital images while maintaining an acceptable level of visual quality.

Steps involved in the JPEG compression pipeline:

Encoding

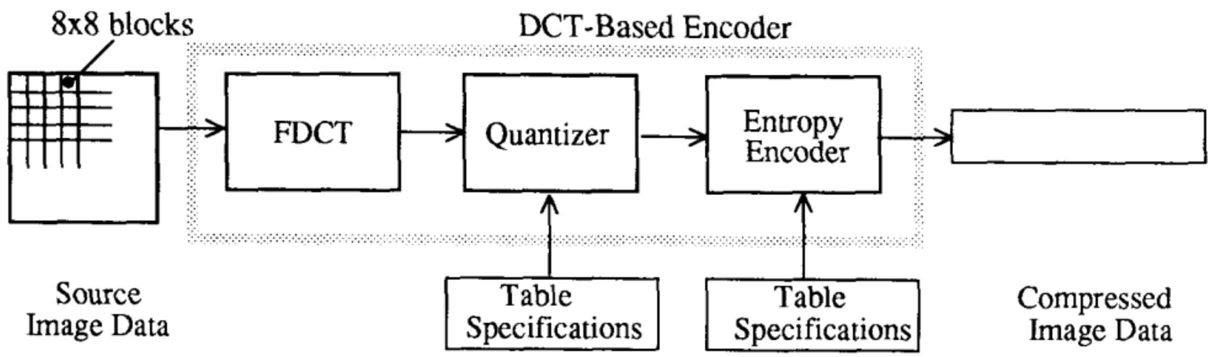


Figure 1. DCT-Based Encoder Processing Steps

1. *Colour Space Conversion*

JPEG typically starts with a colour space conversion. Most images are initially in the RGB (Red, Green, Blue) colour space, but JPEG uses the YCbCr colour space. The transformation involves separating luminance (Y) and chrominance (Cb and Cr) components. The human visual system is more sensitive to changes in brightness than colour, making YCbCr an appropriate choice for compression.

2. *Subsampling*

To further reduce redundancy, chrominance components (Cb and Cr) are subsampled. We used the subsampling scheme 4:2:0, where chrominance is horizontally and vertically subsampled by a factor of 2. This reduces the amount of data to be encoded while maintaining acceptable image quality.

3. *Block Division*

The image is divided into 8x8 pixel blocks. This block-based approach allows for localised compression and is a key aspect of the Discrete Cosine Transform (DCT).

4. *Discrete Cosine Transform (DCT)*

DCT is applied independently to each block in the image. DCT transforms spatial information into frequency information. High-frequency components, which represent fine details, are concentrated in the upper-left part of the DCT coefficient matrix. The coefficients are quantized to reduce precision, and quantization tables are employed to determine the step size.

$$F(u,v) = \frac{1}{4} C(u)C(v)\Big[\sum_{x=0}^{7}\sum_{y=0}^{7} f(x,y) *$$

$$\cos\frac{(2x+1)u\pi}{16} \cos\frac{(2x+1)v\pi}{16}\Big]$$

$$f(x,y) = \frac{1}{4}\Big[\sum_{u=0}^{7}\sum_{v=0}^{7} C(u)C(v)F(u,v)$$

$$\cos\frac{(2x+1)u\pi}{16} \cos\frac{(2x+1)v\pi}{16}\Big]$$

where: $C(u),\ C(v) = 1/\sqrt{2}$ for $u,v = 0$;

$C(u),\ C(v) = 1$ otherwise.

5. *Quantization*

Quantization reduces the precision of the DCT coefficients by dividing them by a quantization matrix. This step introduces some loss of information. Quantization tables determine the amount of compression applied to each frequency component.

$$F^{Q}(u,v) = \textit{Integer Round}\ \left(\frac{F(u,v)}{Q(u,v)}\right)$$

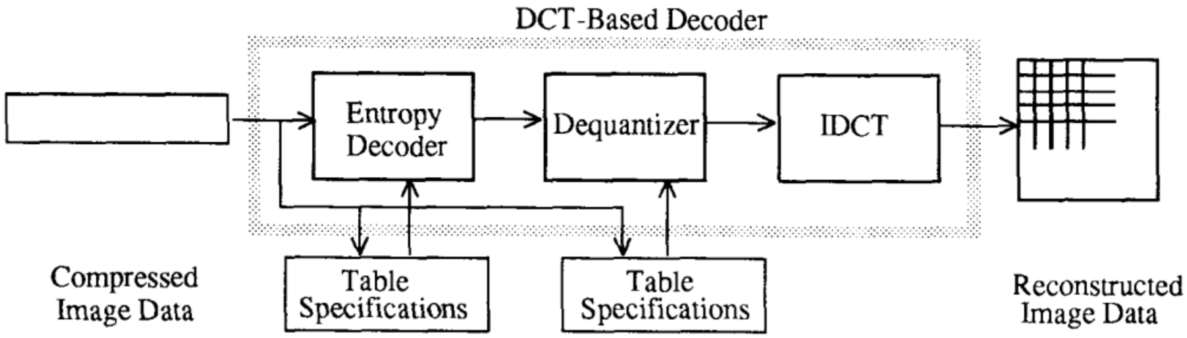
6. *Entropy Coding*

Entropy coding, specifically Huffman coding, is used to represent the quantized coefficients efficiently. Huffman coding assigns shorter codes to more frequent values, resulting in variable-length codes for different symbols. The Huffman tables are predefined and optimised based on the statistical properties of the quantized DCT coefficients.

7. *Bitstream Formation*

The final step involves assembling all compressed data into a bitstream. This bitstream includes information about the image size, colour space, quantization tables, and the compressed DCT coefficients. The resulting JPEG file is a compact representation of the original image.

Decoding



Steps involved in JPEG decoding

1. Bitstream Parsing

This involves reading the file header to extract information such as image dimensions, colour space, quantization tables, and Huffman tables. The parsed information is essential for subsequent decoding stages.

2. Entropy Decoding

Entropy decoding is the inverse of entropy coding. Huffman decoding is applied to the compressed bitstream to reconstruct the quantized DCT coefficients. Huffman tables, extracted from the bitstream during parsing, are used to decode variable-length codes back into symbol values.

3. Dequantization

Dequantization is the process of reversing the quantization applied during encoding. The quantized DCT coefficients are multiplied by the corresponding values in the quantization tables. This step reintroduces some of the information lost during quantization but does not fully recover the original coefficients.

$$F^{Q'}(u,v) = F^{Q}(u,v) * Q(u,v)$$

4. Inverse Discrete Cosine Transform (IDCT)

The Inverse Discrete Cosine Transform is applied to each 8x8 block, reconstructing the spatial information from the frequency domain. The IDCT process involves transforming the DCT coefficients back into pixel values for each block.

5. Upsampling

If chrominance subsampling was applied during encoding (e.g. 4:2:0), upsampling is performed to restore the full chrominance resolution. The inverse of the subsampling process is applied to reconstruct the Cb and Cr components.

6. Colour Space Conversion

The reconstructed luminance (Y) and chrominance (Cb and Cr) components are converted back to the RGB colour space. This step is necessary to obtain the final colour representation of the image.

7. Image Reconstruction

The reconstructed RGB components are combined to form the final decoded image. The image is post-processed to adjust colour balance, contrast, and brightness to enhance visual quality.

Conclusion

The JPEG compression pipeline is a sophisticated process that balances the trade-off between image quality and file size. By leveraging colour space conversion, subsampling, block division, DCT, quantization, entropy coding, and bitstream formation, JPEG achieves high compression ratios while preserving visual fidelity.