Introduction to image compression

K J Somaiya College of Engineering

Somaiya Vidyavihar University

Dr Makarand G Kulkarni

Image Compression

Image can be sampled sufficiently finely so that a binary data stream can represent the original data to an extent that is satisfactory to the most discerning eye.

Since we can represent a picture by something between a thousand and a million bytes of data, we should be able to apply the techniques studied earlier directly to the task of compressing that data for storage and transmission.

Image Compression

1. High quality images are represented by very large data sets. A photographic quality image may require 40 to 100 million bits for representation. These large file sizes drive the need for extremely high compression ratios to make storage and transmission (particularly of movies) practical.

2. Applications that involve imagery seem to be inherently linked to immediate human consumption, and so need to be fast in execution on computers and in transmission. Television, movies, computer graphical user interfaces, and the World Wide Web are examples of applications in which imagery must be moved from storage or across some kind of distribution network very quickly for immediate human intake.

3. Imagery has the quality of higher redundancy than we can generally expect in arbitrary data. For example, a pair of adjacent horizontal lines in an image are nearly identical (typically), while, two adjacent lines in a book have essentially no commonality.

Image Compression

For the first two we will always want to apply highest level of compression technology available for the movement and storage of image data.

The third factor indicates that compression ratios will usually be quite high.

The third factor also says that some special compression techniques may be possible that will take advantage of the structure and properties of image data.

The close relationship between neighboring pixels in an image can be exploited to improve the compression ratios.

This is very important for the task of coding and decoding image data for real-time applications.

Image Compression

Another interesting point to note is that the human eye is very tolerant to approximation error in an image. Thus, it may be possible to compress the image data in a manner in which the less important information (to the human eye) can be dropped. That is, by trading off some of the quality of the image we might obtain a significantly reduced data size. This technique is called **lossy compression**, as opposed to the **lossless compression** techniques discussed earlier. This sentiment, however, can never be expressed with regards to, say, financial data or even textual data! Lossy compression can only be applied to data such as images and audio for which human beings will tolerate some loss of fidelity.

The JPEG Standard for **Lossless** Compression

Some compression can be achieved if we can predict the next pixel using the previous pixels. In this way we just have to transmit the prediction coefficients (or difference in the values) instead of the entire pixel. The predictive process that is used in the lossless JPEG coding schemes to form the innovations data is also variable. However, in this case, the variation is not based upon the user's choice, but rather, for any image on a line-by-line basis. The choice is made according to that prediction method that yields the best prediction overall for the entire line.

The JPEG Standard for **Lossless** Compression

There are eight prediction methods available in the JPEG coding standards. One of the eight (which is the no prediction option) is not used for the lossless coding option that we are examining here. The other seven may be divided into the following categories:

1. Predict the next pixel on the line as having the same value as the last one.
2. Predict the next pixel on the line as having the same value as the pixel in this position on the previous line (that is, above it).
3. Predict the next pixel on the line as having a value related to a combination of the previous, above and previous to the above pixel values. One such combination is simply the average of the other three.

The JPEG Standard for Lossless Compression

The differential encoding used in the JPEG standard consists of the differences between the actual image pixel values and the predicted values. As a result of the smoothness and general redundancy of most pictures, these differences consists of a series of relatively small positive and negative numbers that represent the small typical error in the prediction. Hence, the probabilities associated with these values are large for the small innovation values and quite small for large ones. This is exactly the kind of data stream that compresses well with an entropy code.

The JPEG Standard for **Lossy** Compression

The differential encoding used in the JPEG standard consists of the differences between the actual image pixel values and the predicted values. As a result of the smoothness and general redundancy of most pictures, these differences consists of a series of relatively small positive and negative numbers that represent the small typical error in the prediction. Hence, the probabilities associated with these values are large for the small innovation values and quite small for large ones. This is exactly the kind of data stream that compresses well with an entropy code.

The JPEG Standard for **Lossy** Compression

The JPEG standard includes a set of sophisticated lossy compression options which resulted from much experimentation by the creators of JPEG with regard to human acceptance of types of image distortion. The JPEG standard was the result of years of effort by the JPEG which was formed as a joint effort by two large, standing, standards organizations, the CCITT (The European telecommunications standards organization) and the ISO (International Standards Organization).

The JPEG lossy compression algorithm consists of an image simplification stage, which removes the image complexity at some loss of fidelity, followed by a lossless compression step based on predictive filtering and Huffman or Arithmetic coding.

The lossy image simplification step, which we will call the image reduction, is based on the exploitation of an operation known as the **Discrete Cosine Transform** (DCT), defined as follows.

$$Y(k,l) = \sum_{i=0}^{N-1}\sum_{j=0}^{M-1} 4y(i,j)\cos\left(\frac{\pi k}{2N}(2i+1)\right)\cos\left(\frac{\pi l}{2M}(2j+1)\right) \tag{1.88}$$

where the input image is N pixels by M pixels, $y(i, j)$ is the intensity of the pixel in row i and column j, $Y(k, l)$ is the DCT coefficient in row k and column l of the DCT matrix. All DCT multiplications are real. This lowers the number of required multiplications, as compared to the DFT

The JPEG Standard for Lossy Compression

In the JPEG image reduction process, the DCT is applied to 8 by 8 pixel blocks of the image. Hence, if the image is 256 by 256 pixels in size, we break it into 32 by 32 square blocks of 8 by 8 pixels and treat each one independently. The 64 pixel values in each block are transformed by the DCT into a new set of 64 values. These new 64 values, known also as the DCT coefficients, form a whole new way of representing an image. The DCT coefficients represent the spatial frequency of the image sub-block. The upper left corner of the DCT matrix has low frequency components and the lower right-corner the high frequency components (see Fig. 1.19). The top left coefficient is called the **DC coefficient**. Its value is proportional to the average value of the 8 by 8 block of pixels. The rest are called the **AC coefficients**.

DC coefficient

Low frequency coefficients

4.32	3.12	3.01	2.41
2.74	2.11	1.92	1.55
2.11	1.33	0.32	0.11
1.62	0.44	0.03	0.02

Higher frequency (AC coefficients)

Fig. 1.19 Typical discrete cosine transform (DCT) values for a 4 x 4 image block. In JPEG, 8 x 8 pixel blocks are used.

The JPEG Standard for **Lossy** Compression

So far we have not obtained any reduction simply by taking the DCT. However, due to the nature of most natural images, maximum energy (information) lies in low frequency as opposed to high frequency. We can represent the high frequency components coarsely, or drop them altogether, without strongly affecting the quality of the resulting image reconstruction. This leads to a lot of compression (lossy). The JPEG lossy compression algorithm does the following operations:

1. First the lowest weights are trimmed by setting them to zero.
2. The remaining weights are quantized (that is, rounded off to the nearest of some number of discrete code represented values), some more coarsely than others according to observed levels of sensitivity of viewers to these degradations.

The JPEG Standard for Lossy Compression

Now several lossless compression steps are applied to the weight data that results from the above DCT and quantization process, for all the image blocks. We observe that the DC coefficient, which represents the average image intensity, tends to vary slowly from one block of 8×8 pixels to the next. Hence, the prediction of this value from surrounding blocks works well. We just need to send one DC coefficient and the difference between the DC coefficients of successive blocks. These differences can also be source coded.

We next look at the AC coefficients. We first quantize them, which transforms most of the high frequency coefficients to zero. We then use a **zig-zag** coding as shown in Fig. 1.20. The purpose of the zig-zag coding is that we gradually move from the low frequency to high frequency, avoiding abrupt jumps in the values. Zig-zag coding will lead to long runs of 0's, which are ideal for RLE followed by Huffman or Arithmetic coding.

The JPEG Standard for **Lossy** Compression

4.32	3.12	3.01	2.41
2.74	2.11	1.92	1.55
2.11	1.33	0.32	0.11
1.62	0.44	0.03	0.02

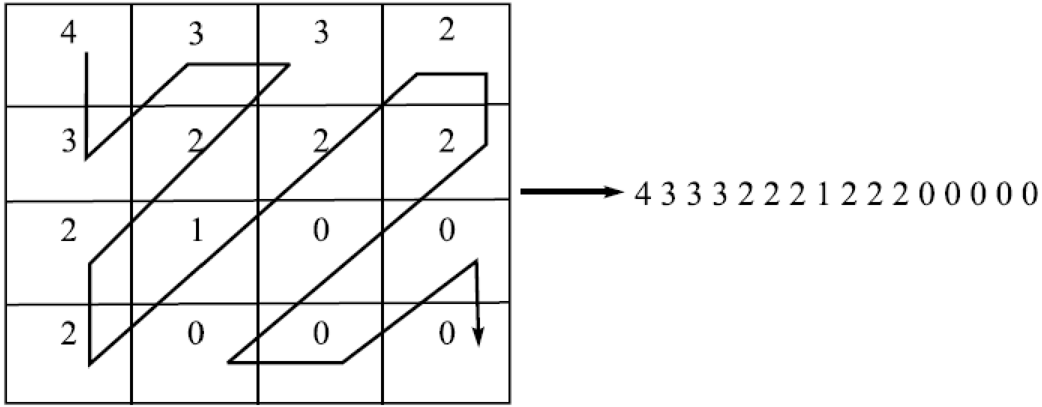

Fig. 1.20 An example of quantization followed by zig-zag coding.

The typically quoted performance for JPEG is that photographic quality images of natural scenes can be preserved with compression ratios of up to about 20:1 or 25:1. Usable quality (that is, for non critical purposes) can result for compression ratios in the range of 200:1 up to 230:1.