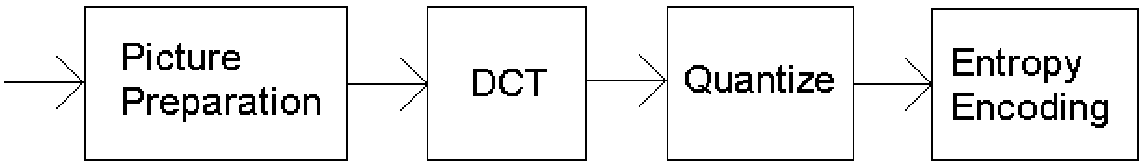
CSE 126/228F Lecture 5: April 15

JPEG (Part 2)

JPEG Stages



DCT

There are many transforms, most of which are very slow. This is important to consider since video demands real-time encoding and decoding. The JPEG committee took suggestions and empirically studied the use of several different transforms. Of the transforms studied, DCT (Discrete Cosine Transform) proved superior.

In JPEG, DCT operates on one block at a time. Because there are 64 elements in an 8x8 block, this is called the 64-element or 64-coefficient DCT. The DCT transform operates on this block in a left-to- right, top-to-bottom manner.

Today, we covered the DCT algorithm. In order to clear up any confusion from today's lecture, please refer to the book (section 6.5.2) for an explanation of the DCT algorithm. As mentioned today, the book's formula for DCT is incorrect, please use the one that follows:

Formula for FDCT:

$$S_{vu} = \frac{1}{4} C_u C_v \sum_{x=0}^{7} \sum_{y=0}^{7} S_{yx} \cos\left[\frac{(2x+1)u\pi}{16}\right] \cos\left[\frac{(2y+1)v\pi}{16}\right]$$

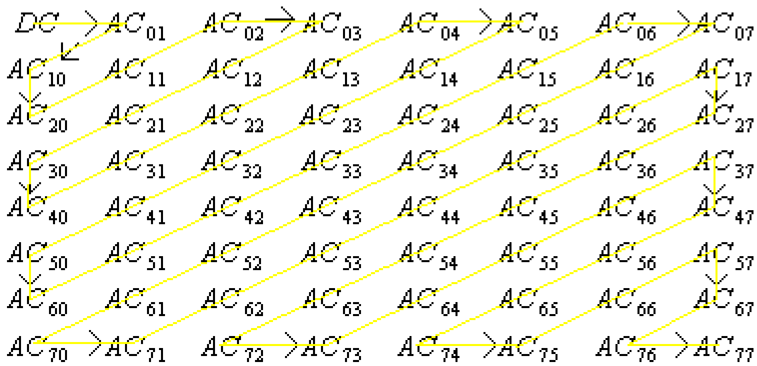
$$C_u, C_v = \frac{1}{\sqrt{2}}, u, v = 0$$

$$C_u, C_v = 0, otherwise$$

Notes about DCT:

The results of a 64-element DCT transform are 1 DC coefficient and 63 AC coefficients. The DC coefficient represents the average color of the 8x8 region. The 63 AC coefficients represent color change across the block. Low-numbered coefficients represent low-frequency color change, or gradual color change across the region. High-numbered coefficients represent high-frequency color change, or color which changes rapidly from one pixel to another within the block. These 64 results are written in a zig-zag order as follows, with the DC coefficient followed by AC coefficients of increasing frequency.

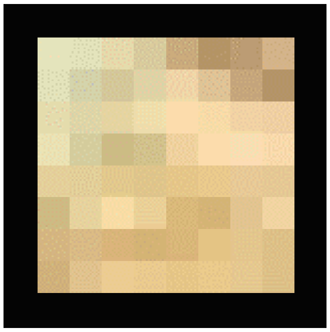
Zig-Zag sequencing:



Note that each diagonal line in this zig-zag sequence contains AC coefficients whose sum is constant. For example, the coefficients {30, 21, 12, 03} all add to 3.

Why is this ordering important? Well, if you think of a block of 8x8 pixels out of a coherent image, the pixels are likely to be very similar. If you run DCT on 64 pixels which are very similar, you will get a DC coefficient and some values for the low-frequency AC coefficients; the remaining coefficients will likely be at or near zero. Try to imagine creating an image out of pixels that wildly vary from their neighbors. The resulting image will more than likely not make much sense, it will just be a mess of dots. To give you an idea of how small an 8x8 region is, consider the following example:

The 8x8 region of pixels highlighted above looks like this (magnified 1600 times)



As you can see, this region does not deviate much from its average color. In addition, the change is slow and gradual across the block rather than sharp and abrupt from pixel to pixel.

This observation about images allows us to place a much greater importance on the DC and first few AC coefficients (beginning of zig-zag sequence) and it also allows us to assume there will be little or no values in the high-frequency AC coefficients (remainder of sequence).

Logically, if these values are of little importance we should be able to assign fewer bits to them in order to achieve greater compression. This naturally leads us to the stages of quantization and entropy encoding, which we will cover next time.