1.2 JPEG †

The JPEG file format has emerged as a near universal image standard employed by nearly all commercial digital cameras. Given a three channel color image (RGB), compression proceeds as follows. An image is first transformed from RGB into luminance/chrominance space (YCbCr). The two chrominance channels (CbCr) are typically subsampled by a factor of two relative to the luminance channel (Y). Each channel is then partitioned into 8×8 pixel blocks. These values are converted from unsigned to signed integers (e.g., from $[0, 255]$ to $[-128, 127]$). Each block, $f_c(\cdot)$, is converted to frequency space, $F_c(\cdot)$, using a two-dimensional discrete cosine transform (DCT):

$$F_c(\omega_k, \omega_l) = \sum_{m=0}^{7} \sum_{n=0}^{7} f_c(m, n) \cos(\omega_k m) \cos(\omega_l n), \qquad (1.19)$$

where $\omega_k = 2\pi k/8$, $\omega_l = 2\pi l/8$, $f_c(\cdot)$ is the underlying pixel values, and c denotes a specific image channel. Note that this representation is simply the Fourier series where only the symmetric cosine basis functions are employed.

Example 1.3 The Fourier transform assumes that the underlying signal or image is periodic. What additional assumption does the DCT make? Show how this assumption leads to a basis of only cosine terms in Equation (1.19).

Depending on the specific frequency ω_k, ω_l and channel c, each DCT coefficient, $F_c(\cdot)$, is quantized by an amount $q_c(\cdot)$:

$$\hat{F}_c(\omega_k, \omega_l) \quad = \quad \text{round}\left(\frac{F_c(\omega_k, \omega_l)}{q_c(\omega_k, \omega_l)}\right). \qquad (1.20)$$

This stage is the primary source of data reduction and information loss.

With some variations, the above sequence of steps is employed by all JPEG encoders. The primary source of variation in JPEG encoders is the choice of quantization values $q_c(\cdot)$, Equation (1.20). The quantization is specified as a set of three 8×8 tables associated with each frequency and image channel (YCbCr). For low compression rates, the values in these tables tend towards 1, and increase for higher compression rates. The quantization for the luminance channel is typically less than for the two chrominance

channels, and the quantization for the lower frequency components is typically less than for the higher frequencies.

After quantization, the DCT coefficients are subjected to entropy encoding (typically Huffman coding). Huffman coding is a variable-length encoding scheme that encodes frequently occurring values with shorter codes, and less frequently occurring values with longer codes. This lossless compression scheme exploits the fact that the quantization of DCT coefficients yields many zero coefficients, which can in turn be efficiently encoded. Motivated by the fact that the statistics of the DC and AC DCT coefficients are different the JPEG standard allows for different Huffman codes for the DC and AC coefficients (the DC coefficient refers to $\omega_k = \omega_l = 0$, and the AC coefficients refer to all other frequencies). This entropy encoding is applied separately to each YCbCr channel, employing separate codes for each channel.