**Image Compression** is the practice of reducing the data storage size of an image file by rearranging the contents of the image into a more elegant format. This is done by 'condensing' the contents of the image using a data compression technique. Data compression methods generally fall into two categories: lossy and lossless. As implied by their names, these genres of compression algorithms are fundamentally different from the perspective of content reproduction.

**Lossy Compression**

Lossy algorithms are designed to discard information that is unnecessary for the approximate reproduction of a dataset. In the case of an image, a lossy compression algorithm has criterion that preserve components of the image matrix pertinent to visual recognition. These criterion are usually based on the nature of the human visual process.

An example of lossy image compression is the popular jpeg (joint photographic experts group) format. This format is built around an algorithm known as the Discrete Cosine Transform, or DCT. The DCT is a Fourier series of cosine waves that the jpeg algorithm uses to transform the contents of a two dimensional matrix into the frequency domain. This transform is called the spectral view of the image. The jpeg algorithm performs this operation on segments of the image defined by a grid of predetermined size. Commonly, the image is given a grid composed of 8*8 blocks of pixels. The two dimensional DCT is then applied to each block, independently. Since the human eye cannot easily detect high frequency signals in an image, the spectral blocks are truncated to contain only low frequency components. This truncation reduces the amount of information in storage since the 8*8 blocks are reduced dimensionally. However, high frequency information is lost during this operation. The image can then be stored. Upon reading the compressed data, the only additional piece of information required is the grid block size. The inverse DCT is thus performed on each block. Each pixel within the block is subsequently rounded to the nearest integer value.

Common issues encountered with jpeg compression include aliasing of periodic patterns and difficulty reproducing sharp changes in luma (i.e. sharp edges). The first issue is caused by the DCT being a waveform fit which generates coefficients for a sum series of cosines. By virtue of the fact that these coefficients are being truncated, data regarding repetition of particular image features may be unable to regenerate at the proper period. The second issue is related to the challenge of 'sufficiency' related to terms in a Fourier series and the approximation of a step function. Sharp edges in an image appear, from the

algorithm's perspective, as step functions instead of gradients. This issue is known in signal processing as the Gibbs phenomenon.

**Learn the theory behind jpeg:**
Example C Code for a 1-D DCT

**Lossless Compression**
Lossless algorithms preserve the exact content of the image data such that all pixel values remain identical and no approximations are made. These algorithms, unlike lossy compression techniques, do not rely on criterion based on visual interpretation. Instead, they rely on explicit relationships or mathematical assumptions made about the image as a discrete matrix of integer numbers.

An example of lossless compression is entropy encoding. Variations of entropy encoding find application in popular compression techniques such as zip files. Entropy encoding works by defining a library of symbols that represent strings repeated within the dataset. For example, if the string “12341234123491234” were compressed, it could be represented as a 'library' and the compressed string, “#=1234, ##9#”. By preserving the library and corresponding compressed string, the compressed data can be reassembled in an unaltered state.

Another form of lossless image compression is supported by the struggling* jpeg2000 image format. However, since jpeg2000 supports both lossless and lossy compressions based on discrete wavelet decomposition, it will not be discussed in this brief.
*http://en.wikipedia.org/wiki/JPEG_2000#Disadvantages_.2F_Legal_issues