**SIV864: Assignment 2**

**Shantanu Chaudhary (2010CS50295)**

Part A

* ***Report On DCT:***
  + **Energy Compaction in DCT:** The compaction performed by any transformation is measured by the ability of the tranform to compress the input data (usually in spacial domain) into as few coefficients as possible (in frequency domain). The coefficients with high amplitude relate to more energy in the frame which in turn gives a measure of more detail in the picture. Thus, the coefficients with large amplitude contribute more to the image details as compared to the coefficients with less amplitude. Now, by the process of quantisation, one can discard low amplitude coefficients without loss of significant details. This in turn doesn't introduce visual distortion in the reconstructed image. The coeffients are also related to the correlation in the image. DCT exploits the correlation in compressing image data. If the pixels in the input image are highly correlated, then the details of the image (in form of pixels) can be tranformed by DCT into less number of coefficients. Hence, more is the correlation, greater will be the compaction. If the correlation is more, then the compression through DCT would give near optimal results.
  + **DCT with respect to KLT:**
    - DCT gives optimal results for images in which correlation tends to unity, whereas KLT is optimal in energy compaction as it places as much energy as possible in as few coefficients as possible.
    - The KLT kernel is not separable as compared to DCT. This makes KLT more resource expensive as compared to DCT due to the inherent full matrix multiplication that needs to be performed.
    - The above difference also makes KLT data dependent and requires faster computation tranforms. Hence, the complexity of KLT is higher than DCT.
  + **DCT with respect to DFT:**
    - DFT is linear, separable (like DCT) and symmetric.
    - DFT is a complex transform (constaining imaginary i term). This means that it demands that both image magnitude and phase information would have to be encoded.
    - As compared to DCT, in case of DFT, its implicit periodicity gives rise to boundary discontinuities. And after quantization, Gibbs Phenomenon causes these boundary points to take incorrect values.
  + **Explanation of degradation of quality of image in Figure 15:** The removal of high-frequency coefficients results in removal of certain frequencies that were originally present in the sine wave. After losing certain frequencies it is not possible to achieve perfect reconstruction of the image.
* ***Salient Points (JPEG):***
  + Color image compression can be approximately regarded as compression of multiple grayscale images, which are either compressed entirely one at a time, or are compressed by alternately interleaving 8x8 sample blocks from each in turn.
  + For DCT, progressive-mode codecs, an image buffer exists prior to the entropy coding step, so that an image can be stored and then parceled out in multiple scans with successively improving quality.
  + Some simple intuition for DCT-based compression can be obtained by viewing the FDCT as a harmonic analyzer and the IDCT as a harmonic synthesizer.
  + The coefficient with zero frequency in both dimensions is called the “DC coefficient” and the remaining 63 coefficients are called the “AC coefficients” (Block Size is 8x8).
  + It has been observed that a fundamental property of DCT is that the FDCT and IDCT equations contain transcendental functions. Consequently, *no physical implementation can compute them with perfect accuracy*.
  + In terms of hardware implementation, hardware with different precision calculations will have different results.
  + For each DCT-based mode of operation, the JPEG proposal specifies separate codecs for images with 8-bit and 12-bit (per component) source image samples. That means imagewith 5-bit source need to be padded appropriately.
  + Quantization is defined as the division of each DCT coefficient by its corresponding quantizer step size, followed by rounding to the nearest integer.
  + Dequantization is the inverse function, which in this case means simply that the normalization is removed by multiplying by the step size, which returns the result to a representation appropriate for input to the IDCT.
  + The Zig-Zag ordering helps to facilitate entropy coding by placing low-frequency coefficients (which are more likely to be nonzero) before high-frequency coefficients.
  + Entropy Coding is done in JPEG by two methods: Huffman Coding, Arithmetic Coding.
  + The JPEG proposal specifies no required Huffman tables.
  + For color images with moderately complex scenes, all DCT-based modes of operation typically produce certain levels of picture quality for the certain ranges of compression. These levels are offer a guideline as to how quality and compression can vary significantly according to source image characteristics and scene content. (The units “bits/pixel” here mean the total number of bits in the compressed image - including the chrominance components - divided by the number of samples in the luminance component.)
    - 0.25-0.5 (moderate quality)
    - 0.5-0.75
    - 0.75-1.5
    - 1.5-2.0 (inditinguishable from original image)
  + For the lossless mode of operation, two different codecs are specified - one for each entropy coding method. The encoders can use any source image precision from 2 to 16 bits/sample, and can use any of the predictors except selection-value 0. The decoders must handle any of the sample precisions and any of the predictors.
  + JPEG does not specify or encode any information on pixel aspect ratio, color space, or image acquisition characteristics.
  + For JPEG in color images: Each component (colors) consists of a rectangular array of samples. A sample is defined to be an unsigned integer with precision P bits, with any value in the range [0, 2P -1]. All samples of all components within the same source image must have the same precision P which can be 8 or 12 for DCT-based codecs, and 2 to 16 for predictive codecs.
  + Let us say that that the sample dimernsions are xi and yi. Now to accomodate formats in which some image components are sampled at different rates, the components can have different dimensions. Let Hi and Vi be the relative horizontal and vertical sampling factors respectively for each component. Let X and Y be overall image dimensions. X and Y are defined as the maximum xi and yi for all components of the image. H and V are allowed only integral values (1-4). Thus, the encoded parameters are X, Y, Hi, Vi for each component of the image.
  + At the decoder side, the dimensions xi and yi are constructed using the following formula:

xi  = Ceiling ( X\*(Hi/Hmax))

yi = Ceiling ( Y\*(Vi/Vmax))

* + It also important to identify the encoding order and interleaving when the data is compressed and sub sequently decompressed. For displaying multiple-component images, many applications pipeline the process of decompression inorder to enable parallel processing. This process is complemented if the components are interleaved together within the compressed data stream.
  + JPEG defines Minimum Coded Unit (MCU) as the smallest group of interleaved units (A data unit is a sample in predictive coding and an 8x8 block of samples in DCT-based coding). When a unit is partitioned into horizontal and vertical segments then there can be various permutations and combinations for sequencing the coefficients (refer paper). Thus, interleaved data is an ordered sequence of MCUs, and the number of data units contained in an MCU is determined by the number of components interleaved and their relative sampling factors.
  + The JPEG proposal aloows some components to be interleaved and some to be non interleaved within the same compressed image.
  + In terms of specifications for tables, JPEG decoders can store upto 4 different decoders, and upto 4 different entropy coding tables.
  + Baseline sequential coding is for images with 8-bit samples and uses Huffman coding only. It also differs from the other sequential DCT codecs in that its decoder can store only two sets of Huffman tables (one AC table and DC table per set).
  + VLI Codes and VLC Codes (VLI: Valriable Length Integer, VLC: Variable Length Coding): Huffman codes (VLCs) must be specified externally as an input to JPEG encoders. The JPEG proposal includes an example set of Huffman tables in its information annex, but because they are application-specific. the VLI codes are far more numerous, can be computed rather than stored, and have not been shown to be appreciably more efficient when implemented as Huffman codes.
  + While discussing DCT-Progressive Mode, the paper discusses two methods by which a block of quantised DCT coefficients may be partially encoded. First is the process of “Spectral Selection” in which a specific band of coefficients from the zig zag sequence are encoded within a given scan. Second is the process of “Successive Approximation”, wherein the coefficients within the current band need not be encoded with full precision. Upon a coefficient’s first encoding, the N most significant bits can be encoded first, where N is specifiable. In subsequent scans, the less significant bits can then be encoded.
  + Hierarchical encoding is useful in applications in which a very high resolution image must be accessed by a lower-resolution display.

Part B

In my experiment, I have used th following mask with DCT as well as DFT.

Mask = [

1 1 1 1 0 0 0 0

1 1 1 0 0 0 0 0

1 1 0 0 0 0 0 0

1 0 0 0 0 0 0 0

0 0 0 0 0 0 0 0

0 0 0 0 0 0 0 0

0 0 0 0 0 0 0 0

0 0 0 0 0 0 0 0];

In other words, only 10 coefficients out of the total 64 are retained, rest are discarded. The program was run for both DCT and FFT. The input image was a colored image (lena). Next page shows input image and its gray color variant.

Input Image:





The output images after reconstruction are shows below:



The top image corresponds to reconstruction with DFT transforms and the lower corresponds to reconstruction with DCT transforms. We observe that inspite of 85% coefficient loss, DCT gives far more sharp results as compared to DFT.

**Mean Square Error:**

With DFT: 132.3715

With DCT: 39.6749