AUDIOVISUAL COMMUNICATIONS

Laboratory Session on the Discrete Cosine Transform (DCT)

Fernando Pereira

The objective of this lab session about the DCT is to get the students familiar with the way the DCT works with the target to compress images. For that purpose, the application "Comunicação de Imagem"[1] will be used which includes among others a module showing the various operations involved in DCT image compression of an 8×8 image block.

1. THE DCT TRANSFORM

The DCT is, with no doubt, the most popular transform in the field of image and video compression since it is used by the JPEG, H.261, H.263, MPEG-1 Video, MPEG-2 Video and MPEG-4 Visual and H.264/AVC coding standards, due to its high compression performance.

The DCT is computed as the product of two matrices, notably the image block (luminance or chrominances) and the DCT base functions (see Figure 1) according to the following expressions:

$$F(u,v) = 1/4\ C(u)\ C(v) \sum_{x=0}^{7}\sum_{y=0}^{7} f(x,y)\ \cos\left[\pi u(2x+1)/16\right]\cos\left[\pi v(2y+1)/16\right]$$

with u = 0, 1, 2, ..., 7

v = 0, 1, 2, ..., 7

$$C(u),\ C(v) = \begin{cases} 1/\sqrt{2} & u,v=0 \\ 1 & otherwise \end{cases}$$

Moreover, the inverse DCT is computed as:

$$f(x,y) = 1/4 \sum_{u=0}^{7}\sum_{v=0}^{7} C(u)C(v)F(u,v)\cos[\pi u(2x+1)/16]\cos[\pi v(2y+1)/16]$$

with x = 0, 1, 2, ..., 7

y = 0, 1, 2, ..., 7

---

[1] The application "Comunicação de Imagem" has been developed by Pedro Fernandes in the context of his 'Trabalho Final de Curso'. I would like to express here my appreciation for his work considering the possibilities this application has opened for the lab sessions of the Audiovisual Communications course.

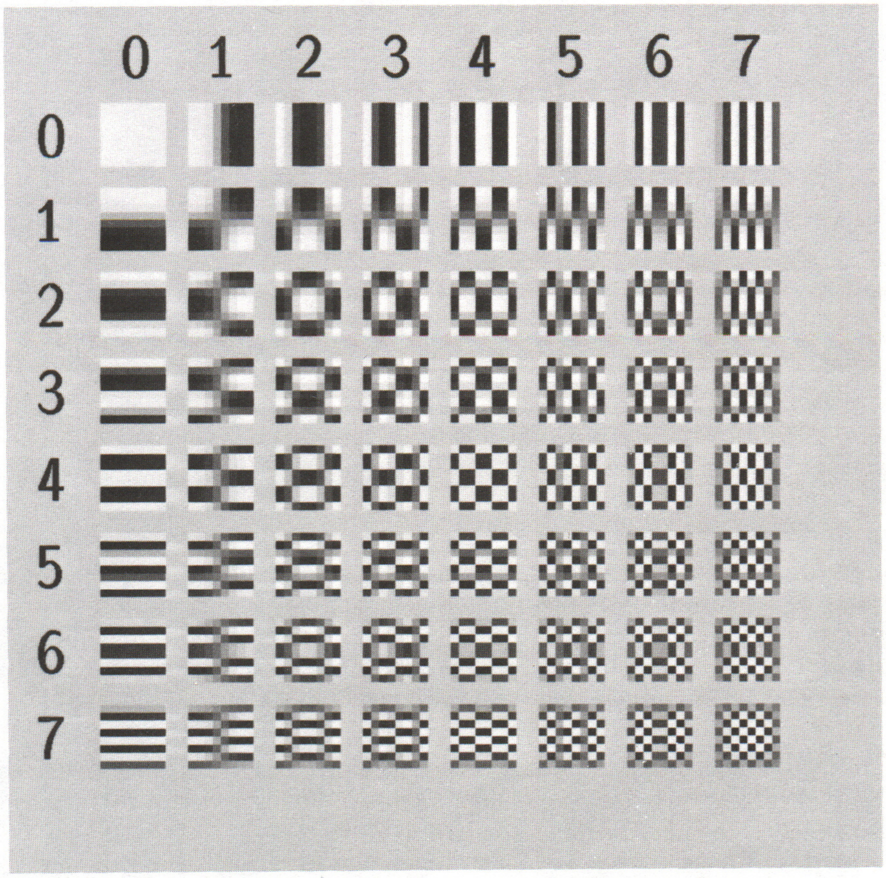

Figure 1 – Visual representation of the DCT base functions.

2. DESCRIPTION OF THE MATRICES

The DCT module uses several matrices which are described in the following:

2.1 Coding and Decoding Matrices

- * ***Original Image –*** Matrix with the 8×8 image block over which the DCT is applied this means spatial block with 64 PCM samples; each luminance sample has a value between 16 and 235, if the ITU-R 601 recommendation is followed, and between 0 and 255, otherwise.
- * ***DCT Coefficients –*** Matrix with the 8×8 DCT coefficients block; although the DCT coefficients are real values, they are represented with unit rounding (from -2048 to 2048).
- * ***Selected DCT Coefficients –*** Matrix with the DCT coefficients which have fulfilled the selection criteria defined by the 'Selection Filter' matrix (see below). The selection criteria are not normative and thus each encoder developer may adopt different solutions.
- * ***Quantized DCT Coefficients –*** Matrix with the quantized DCT coefficients using the quantization steps defined in the 'Quantization Step' matrix. Clipping may happen when the maximum quantization level is lower than the quantization level determined by the simple

division of the DCT coefficient by the corresponding quantization step; in recommendation H.261, this clipping may happen since the maximum quantization level is lower than the maximum quantization level that may be needed when the quantization step is very low. An example is the quantization of a DCT coefficient with value 1660 with quantization step: if the maximum quantization level available is 512, as in the H.261 recommendation, there will be clipping since the 'correct' should be 830. When clipping happens for a DCT coefficient, the corresponding cell in appear in red.

The quantization characteristic used in the application (but not normative) is:

Quantization_level = Int (DCT value / Quantization_step)

Remind that the quantization characteristic is not normative but uniform quantization as above is frequently used. The quantization process implies introducing an error that cannot be recovered, making the coding process lossy.

* ***Restored DCT Coefficients*** **–** Matrix with the DCT coefficient values recovered at the decoding based on the quantization levels and quantization steps used. The so-called 'inverse quantization' uses the received quantization levels and the corresponding quantization steps to recover the decoded DCT coefficient value, e.g. as in recommendation H.261:

  Dequantized_DCT_value = Quantization_step × (Quantization_level + 1/2)

  The inverse quantization process is typically normative which means the dequantized values for each quantization level and quantization step are fixed and well known in advance by the encoder, which may used this fact to simulate any type of quantization characteristic.

* ***Inverse DCT Transform*** **–** Matrix with the 8×8 samples resulting from the inverse DCT applied to the dequantized DCT coefficients. These sample values must be close to the 'Original image' values and the closer, the lower are the quantization steps.

* ***Differences*** **–** Matrix with the spatial luminance differences between the original encoded image block and the decoded image block resulting from the inverse DCT. These differences should be smaller for lower quantization steps. The application shows some objective quality metrics, notably the Root Mean Squared Error (RMS), which is the square root of the Mean Squared Error (MSE), the Signal to Noise Ratio (SNR) and the Peak SNR (PSNR):

$$SNR = 10\log_{10} \sum_{i=1}^{M}\sum_{j=1}^{N} x^2{}_{ij} \Big/ \sum_{i=1}^{M}\sum_{j=1}^{N} (x_{ij} - y_{ij})^2 \quad (dB)$$

where x is the original sample value at position (i,j) and y is the corresponding decoded value. To obtain the PSNR (and not the SNR), the SNR must be computed using the peak value for the signal, this means 255 for samples with 8 bits:

$$SNR = 10\log_{10} 255^2 \Big/ {}^{1}\!\!\diagup_{MN} \sum_{i=1}^{M}\sum_{j=1}^{N} (x_{ij} - y_{ij})^2 \quad (dB)$$

with

$$MSE = {}^{1}\!\!\diagup_{MN} \sum_{i=1}^{M}\sum_{j=1}^{N} (x_{ij} - y_{ij})^2 \quad (dB)$$

2.2 Control Matrices

A. ***Selection Filter*** – This filter uses two types of information, and thus two matrices, to decide how to filter the DCT coefficients:

* ***Selection*** – This matrix establishes the basic selection criteria for the coefficients selections according to:

☑ - The corresponding DCT coefficient is always selected for the next processing phase, the quantization, independently of its value (green cells).

☒ - The corresponding DCT coefficient never goes to the quantization phase, independently of its value (red cells).

🖩 - The corresponding DCT coefficient goes to the quantization phase if its amplitude is equal or higher than the corresponding threshold specific in the Filter Values matrix (white cells).

* ***Filter Values*** – This matrix defines the selection threshold for the DCT coefficients with 🖩 in the Selection matrix; this means that from the DCT coefficients with a 🖩 , only those which amplitude is equal or higher than this threshold will go the quantization phase.

B. ***Quantization Step*** – Matrix with the quantization step to be used for each DCT coefficient. The DCT coefficient (upper left corner) is usually quantized with a quantization characteristic without dead-zone, contrary to the other coefficients. Generally, all quantization steps can go from 1 to 255. In the H.261 mode, the DC coefficient is mandatory quantized with step 8, while the AC coefficients must use integer and even quantization step values, from 2 to 62. While the JPEG standard uses different quantization steps for the various DCT coefficients, depending on the human visual sensibility to each spatial frequency (based on the Lohscheller values), recommendation H.261 uses the same quantization step for all AC coefficients since prediction residues (noise) are coded. The DCT coefficients dequantization is made as recommended for recommendation H.261.

3. INTERACTION WITH THE APPLICATION

3.1 Description of the Switching Buttons

There are some switching buttons available in the bar at the button of the screen which allow to select same options, notably:

* **H.261 mode** – This button allows switching between the general mode and the H.261 mode. The difference between these two modes is related to the values allowed for the luminance (0-255 versus 16-235), for the quantization steps and for the quantized DCT coefficients.

* **Visualization of luminance values** – This button allows visualizing or not the luminance values on top of the corresponding matrices, notably in the Original Image, Inverse DCT Transform and Differences matrices.

- **Visualization of the DCT coefficients values** – This button allows visualizing or not the DCT coefficients values on top of the corresponding matrices, notably in the DCT Coefficients, Selected DCT Coefficients and Restored DCT Coefficients matrices.
- **Complementary image** – This button allow changing the Original Image to its complementary luminance image, this means ‘darks’ become ‘lights’ and vice-versa.

3.2 Changing the Matrices

There is the possibility to change the Original Image, DCT Coefficients, Selection Filter and Quantization Step matrices. To do that, the user must press twice of the mouse left button which opens another working window with the values of the matrix to be changed. The changes may be made in the following way:

- **Change a value in the matrix** – Just change the value in the corresponding cell in the open working window.
- **Applying 'value'** ***from*** **a certain DCT coefficient A to another coefficient B in the zig-zag order** – The value *value* is set for all positions in the zig-zag scanning order from position A to position B; counting starts with 0 (before closing the window, press *zig-zag fill*).
- **Random values –** The selected matrix is filled with random values having as maximum limit the value *value* (before closing the window, press *random*).
- **Lohscheller values –** The quantization step matrix is filled with half the Lohscheller values for the luminance; the Lohscheller values express the human visual sensitivity to the vertical and horizontal spatial frequency variations associated to each DCT coefficient.

3.3 Image Block Examples

The application includes 20 image block examples. The image blocks and associated data can be visualized using the ↑ e ↓ buttons; the image block number is shown on the button of the screen:

Example 1) single black pixel at the upper left corner of a fully white block
Example 2) line of black pixels on the top of a fully white block
Example 3) column of black pixels at the left of a fully white block
Example 4) line and column of black pixels on the top and at the left of a fully white block
Example 5) black and white alternating lines
Example 6) black and white alternating columns
Example 7) alternate black with mixed black and white lines and columns
Example 8) black central diagonal, downwards from left to right, in a fully white block
Example 9) black central diagonal, upwards from left to right, in a fully white block
Example 10) black cross made with central diagonals with a fully white block
Example 11) black cross as example 10 but without the black central pixels
Example 12) grey half image block with back diagonal in the other lighter grey half
Example 13) four black central pixels in a fully white block
Example 14) uniform grey block (128)
Example 15) sinusoidal horizontal light-dark transition block
Example 16) sinusoidal horizontal light-dark-light transition block
Example 17) fast sinusoidal horizontal transitions

Example 18) fast sinusoidal vertical transitions
Example 19) vertical and horizontal light-dark-light transitions
Example 20) vertical, horizontal and diagonal light-dark-light transitions

Audiovisual Communications Course Laboratories

Laboratory Session on the Discrete Cosine Transform (DCT)

Session _____

Nº _________ **Name** ____

Nº _________ **Name** ____

Nº _________ **Name** ____

Try to address the following questions in a way speedy way but always fully understand what you are doing; if you don’t understand, ask for help. The most important is to learn …

1. **DCT Coefficients versus Image Blocks:** For four more intuitive cases, from those included in the 20 image block examples, analyze and comment the set of non-null DCT coefficients obtained.

2. **Subjective and Objective Quality versus Quantization Step:** For the image block example number 13, analyze and comment the subjective (by visual inspection) and objective (SNR and PSNR) quality variations when the quantization steps are:

   a) the Lohscheller thresholds

   b) a constant quantization step equal to 8, 16, 32 and 64 (e.g. indicate in the application that you want a certain quantization step for all DCT coefficients from 0 to 63 in the zig-zag scanning order)

3. **Subjective and Objective Quality versus the Selection Threshold:** For the image block example number 13, analyze and comment the spatial and frequency domains impact of filtering the DCT coefficients with a 'selection filter' which:

   a) selects only the first six DCT coefficients (0 to 5 in the zig-zag scanning) which means that after the sixth DCT coefficient no coefficient is used, independently of their value

   b) selects the first 15, 30, 45 and finally all DCT coefficients ('killing' the other coefficients)

   c) selects only the DCT coefficients which are above a threshold of 200, 100, 50 and 25, with the exception of the DC coefficient which is always selected

4. **New ‘Selection Filter’:** Suggest a new ‘selection filter’, which is independent of the image block (this means it should be acceptable for all image blocks), which looks reasonable from the perspective of maximizing the subjective impact when coding with a certain amount of bits; explain the reasoning behind the design of your filter.

5. **DCT Coefficients at Decoder:** Explain the difference between the DCT coefficients computed at the encoder (‘DCT Coefficients’) and the DCT coefficients recovered at the decoder (‘Restored DCT Coefficients’).

6. **Quantized DCT Coefficients:** Identify which elements or operations determine the set of non-null DCT coefficients obtained on the ‘Quantized DCT coefficients’ matrix.

7. **Complementary Block DCT:** Analyzing the 20 image block examples and their complementary image blocks, this means the blocks where the luminance is computed as (255-Y), comment on what changes in terms of the DCT coefficients when the complementary image block is coded and not the original image block (use the chess board button to get the complementary image block)

8. **New Image Block:** Edit the ‘Original Image’ block and draw a not too simple letter, e.g., an R, as it could happen in a facsimile bilevel image. Analyze and comment the DCT transform performance to code this type of content.

9. **New DCT Coefficients Block:** Edit the ‘DCT Coefficients’ block and give 3 DCT block examples which image block corresponds to an intuitively expectable results, e.g., if only the DC coefficient is used than an uniform image block has to result.