EL-GY 6123 Image and Video Processing, Fall 2018
Programming Assignment 7 (Due 5/16/2018)

1. Write a code that implements a basic form of the block-based hybrid video coder for coding a P-frame using a fixed block size of 8x8. For simplicity, consider intra-prediction using only the first 3 intra prediction modes shown below over 8x8 blocks, and inter-prediction using integer accuracy EBMA, with a specified search range, e.g. +/-24. You program should do the following for each block: i) find the best intra-prediction mode and the corresponding error block and its variance; ii) find the best MV for inter-prediction and the corresponding error block and its variance; iii) Choose the prediction block whose prediction error has the smaller variance; iv) Perform 8x8 DCT on the prediction error block of the chosen mode; v) Quantize all the DCT coefficients with the same quantization step-size (QS) q; vi) Count how many non-zero coefficients you have after quantization, vii) Reconstruct the error block by performing inverse DCT on quantized DCT coefficients; viii) Reconstruct the original block by adding the reconstructed error block into the best prediction block. Instead of developing a real entropy coder, we will use the total number of non-zero DCT coefficients as an estimate of the bit rate and ignore the bits needed to code the side information (mode info, motion vector, etc.). Your program should determine the PSNR of the reconstructed image (compared to the original image) and the total number of non-zero quantized DCT coefficients K, for a given quantization step-size q.

   Select two frames of a video, and apply your program to the second frame, by using the original of the first frame as the inter-prediction reference frame. Apply your program using several different q with a large range (e.g. 1 to 16 in steps of 1) and determine the corresponding PSNR and K for each q. Plot PSNR vs. K as your approximate PSNR vs rate curve. Note: if your video has very little motion, you should select two frames that are several frames apart. Otherwise, the inter-prediction errors can be all nearly zero.

   You can use the function cv2.dct(), to perform 2D DCT. You should write your own function for EBMA, quantization and counting the number of non-zeros, and for calculating MSE between two images and for converting MSE to PSNR.

   You can download the following test video foreman.avi) from NYU n course website

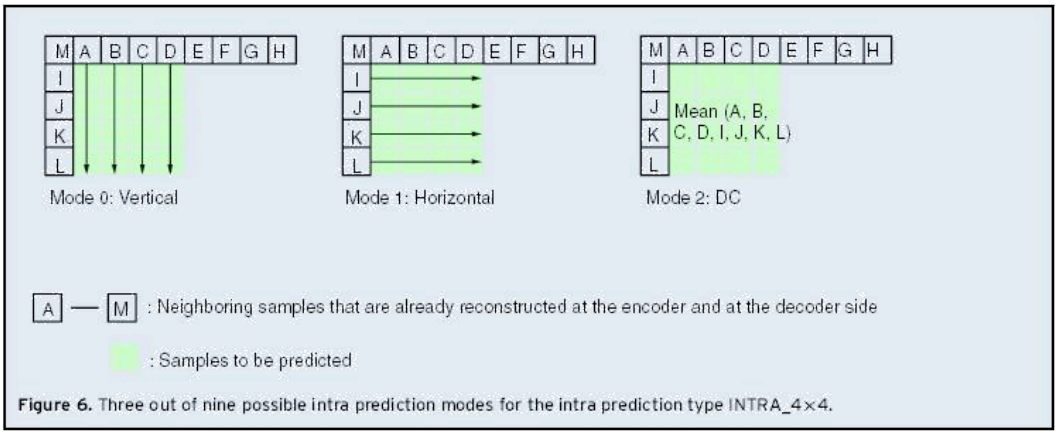

Figure 6. Three out of nine possible intra prediction modes for the intra prediction type INTRA_4×4.

3 intra-prediction modes illustrated at the 4x4 level. Please extend to 8x8 level. For the DC mode over 8x8 blocks, use the average of the 8 pixels on top and 8 pixels to the left, and the top-left pixel (M).

2. (**Bonus, extra 5 pt for the computer assignment**) Develop a code for coding a sequence of frames, with a GOP structure of IBPBPBP … I. For I-frame, you should use intra prediction only. For P-frame, you choose between intra prediction and inter prediction from the previous frame only. For B-frame, you choose between intra prediction, inter prediction frame the previous frame, and inter prediction from the following frame. In each case, use the prediction that leads to the least error variance. Record the average PSNR and number of non-zero DCT coefficients K for each frame type (I, P, and B) for the same q. Also record the average PSNR and K over all frames. Repeat for different q. Plot PSNR vs. K curves for different frame types as separate curves. Also plot average PSNR vs. K over all frames. You should observe that to achieve the same PSNR, I frame will require the highest K, followed by P, and then by B.

   Please note that you should use the decoded frame (with quantization error) as the reference frame for inter prediction. That is, if you want to code frames F1, F2, F3, F4, F5, using IBPBI mode, you should first code F1 as an I-frame, save reconstructed F1 frame; then code F3 as a P-frame, using reconstructed F1 (from quantized DCT coefficients) as the reference frame for inter prediction; then code F2 as a B-frame, using reconstructed F1 and F3 as reference frames.