1. JPEG Encoding

Assume the resolution of the image is 640x480. After the picture preparation stage in the JPEG encoding, how many Y, U, V blocks are there in the image? If the entire region of Y, U and V blocks must be encoded before the next region is encoded, what is the sequence of the input blocks into the DCT stage?

Sol. In JPEG encoding, the raw image is divided into 3 planes (Y,U and V) and these planes are encoded separately. The U and Y planes are first subsampled by a factor of four (both x and y directions) and the result region (a group of 4x4 pixels) is called a meta-region. Therefore, in an image, there are \((640*480*1/16)\) meta-regions in U and V planes. Then, 8x8 blocks of values in Y plane and 8x8 blocks of meta-regions in U and V planes will be sent into the remaining stage. Since the block has a size of 64 values/meta-region, in conclusion, there are totally \((640*480*1/64)\) blocks of Y and \((640*480*1/16*1/64)\) U and V frames.

Due to the subsampling of values in U and V planes with a factor of four, one block of U and Y will be sent together with 16 blocks of Y.

2. MPEG Frame Sequence

Give a frame sequence (in an order with respect to their presentation time) as shown below.

\[I_1, B_2, B_3, B_4, P_5, B_6, I_7, P_8, B_9, I_{10}, B_{11}, B_{12}, B_{13}, I_{14}, \ldots\]

1.1 What is the sequence of frames in an order with respect to their decoding time?

Sol. \[I_1, P_5, B_2, B_3, B_4, I_7, B_6, P_8, I_{10}, B_9, I_{14}, B_{11}, B_{12}, B_{13}, \ldots\]

1.2 Assume that the frame keeps repeating every after 13 frames and the first 13 frames are in the same GOP. Is the first GOP an independent GOP?

Sol. No, the frames \(B_{11}, B_{12}\) and \(B_{13}\) all depends on the frame \(I_{14}\), which is in the second GOP.

1.3 If the answer in the above question is ‘no’, how can you modify the type of a frame (just one) to make the first GOP an independent GOP?

Sol. Change the type of the fourteenth frame from \(B_{13}\) to \(P_{13}\) (or \(I_{13}\))
And we can write the sequence of frame in an order with respect to their decoding time as

I₁, P₅, B₂, B₃, B₄, I₇, B₆, P₈, I₁₀, B₉, P₁₃, B₁₁, B₁₂, I₁₄ …

1.4 If the P₅ is lost, what frames will be also lost?

Sol. Because the frames B₂, B₃, B₄ and B₆ all depend on the frame P₅, if P₅ is lost, these frames couldn’t be decoded.

3. MPEG 1

Suppose the resolution of each video frame is 320x240 and the refresh rate is 30 frame/s. We consider the video of the balloon that is moving straightly upward at 18 miles/hour speed, starting from the bottom in the first frame (moving vertically only). If the frame is showing 1/100 mile height track, which frame will show the picture of the balloon that reaches the top of the frame. Assume that at the frame number 0, the top of the balloon is at the bottom edge of the frame exactly.

Sol.

The balloon is moving upward at the 18 miles/hour speed. Relatively speaking, in one second, it will move vertically by 0.005 miles. Since the height track of the frame is 0.01 miles, it takes 2 second or 60 frames for the balloon to reach the top of the frame. At that time, the 60th frame will be shown.