ELEC 5306: Advanced Signal Processing: Video Compression

1. Objective

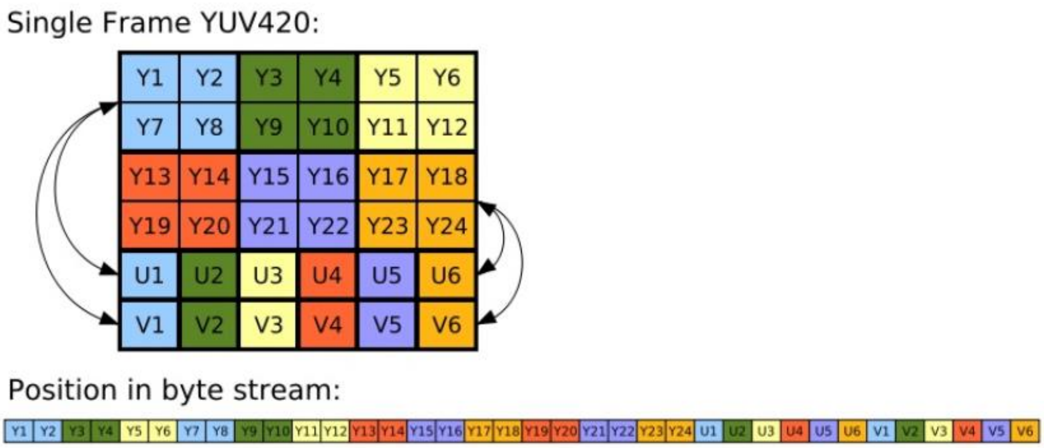
To read and show the frame from a YUV420 file;

To investigate the uniform quantizers;

2. Video Formation

YUV is a color space commonly used as part of a color image representation, which encodes a color frame taking human perception into account and allowing reduced bandwidth for chrominance components. Therefore, YUV formation typically enables transmission to be more efficiently masked by the human perception than using a direct RGB-representation.

YUV420 is a video format, which means that the Y, U, and V values are grouped together instead of interspersed. The reason is that by grouping the U and V values together, the video frame becomes much more compressible. In other words, when given an array of a frame in the YUV420p, all the Y values come first, followed by all the U values, followed finally by all the V values.



1: YUV420 example (4×6)

Fig. 1 shows an example of YUV420 format with the size 4×6 pixels. The U and V values correspond to each 2 by 2 block of the frame, meaning each U and V entry applies to four pixels. After the Y values, the next X/4 positions are the U values for each 2 by 2 block, and the next X/4 positions after that are the V values that also apply to each 2 by 2 block.

Input: a given YUV420 file called “Birds_part_420.yuv” (1024 x 768 pixels)

Output: Show the 5th frame in the matlab (it should be a color image!)

3. Uniform Quantization

In image and video processing, since the quantization compresses the continuum of analog signal to a finite number of discrete values, it must introduces some distortions in the displayed analog image. This distortion is known as quantization error and might shown as patches, especially in flat region, which is also known as contouring artifacts.

Commonly, we define a scalar quantizer $\mathrm{Q}(\cdot)$ to be a mapping of input values $\{\mathrm{d_k}; \mathrm{k} = 1, \ldots, \mathrm{N}\}$ to output $\{\mathrm{r_k}; \mathrm{k} = 1, \ldots, \mathrm{N}\}$. Thus, we have:

$$\mathrm{Q}(\mathrm{x}) = r_k \tag{1}$$

We assume that the quantizer output levels maintain the following relationship:

$$\mathrm{r_1} < \mathrm{r_2} < \cdots < \mathrm{r_N} \tag{2}$$

The number of bits required to denote any one of the output levels is:

$$\mathrm{B} = \lceil \log 2\mathrm{L} \rceil (\mathrm{bits}) \tag{3}$$

where $\lceil x \rceil$ is the nearest integer equal to or larger than x.

The optimal MSE quantizer design amounts to determining the input decision intervals $\{\mathrm{d_k}\}$ and corresponding reconstruction levels $\{\mathrm{r_k}\}$ for given L levels such that the MSE is minimized:

$$\mathrm{MSE} = \mathrm{E}[(\mathrm{x} - \mathrm{x_1})^2] = \sum_{j=1}^{L} \int_{d_j}^{d_{j+1}} (x - x_1)^2 p_x(x) dx \tag{4}$$

Minimization of equation (4) is obtained by satisfying the following two conditions:

$$\mathrm{d_j} = \frac{d_{j+1} + d_{j-1}}{2} \tag{5}$$

$$\mathrm{r_j} = \frac{\int_{d_j}^{d_{j+1}} x p_x(x) dx}{\int_{d_j}^{d_{j+1}} p_x(x) dx} \tag{6}$$

When the ***pdf*** (probability density function) of the analog signal is uniform, the decision intervals and output levels of the quantizer can be computed analytically as shown below. In this situation, the decision intervals are all equal as well as the intervals between the output levels. The quantizer is called a uniform quantizer. The uniform ***pdf*** is given by

$$\mathrm{p_x}\,(\mathrm{x}) = \frac{1}{x_{max} - x_{\min}} = \frac{1}{\mathrm{d_{L+1}} - \mathrm{d_1}} \tag{7}$$

Also, we can derive:

$$\mathrm{d_{j+1}} - \mathrm{d_j} = \mathrm{d_j} - \mathrm{d_{j-1}} = \Delta\,(2 \le \mathrm{j} \le \mathrm{L}) \tag{8}$$

Then, the (6) can be simplified as:

$$\mathrm{r_j} = \frac{\mathrm{d_{j+1}} + \mathrm{d_j}}{2} = \mathrm{d_j} + \frac{\Delta}{2} \tag{9}$$

From the above equation, we see that the interval between any two consecutive decision boundaries is the same and equal to the quantization step size Δ. Finally, we can find that the Δ of the uniform quantizer is related to the number of levels L or the number of bits B, as given by

$$\Delta = \frac{\mathrm{x_{max}} - \mathrm{x_{min}}}{\mathrm{L}} = \frac{\mathrm{d_{L+1}} - \mathrm{d_1}}{2^{\mathrm{B}}} \tag{10}$$

Finally, the design of the optimal uniform quantizer is carried out as follows:

1. To determine the input range L and step size Δ, where $d_1 = x_{min}; d_{L+1} = x_{max}$
2. To find the L output levels from (9);
3. To quantize a sample x according to this equation $Q(x) = r_j$, if $d_j \leq x \leq d_{j+1}$

Input: the **Y** component (8bpp) in the first experiment.

Output: Display both the original and requantized (dequantized) images, observe the differences, calculate the PSNR. For display purpose, set B to be 4 and 6 bits. For experimental plotting, vary B from 1 to 7 bits.

$$MSE = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [I(i,j) - R(i,j)]^2$$

$$PSNR = 10\log_{10}\left(\frac{Max_I^2}{MSE}\right)$$

where I(i; j) is the original data and R(i; j) is the reconstructed one; M and N is the width and height respectively; Max_I is the maximum value in I.