5C1 Video Processing: Assignment II

Dian Zhuang

20 April 2023

---

1. Describe, explain and justify the algorithm you use to create the representations. Explain how you generate the PSNR that is used to compare representations and make decisions.

1.

- FFMPEG transcoding: I use the standard PSNR calculation method to calculate the PSNR values of videos at different bit rates generated by transcoding at different bit rates, which are then compared with the original video: First, transcode the given 3 videos, and then, compare each frame of the transcoded videos with the original video frame by frame, calculate the PSNR for each frame, sum them up and take the average to obtain the PSNR values at different bit rates.

  The transcoding command is: "ffmpeg -s 640x274 -pix_fmt yuv420p -i dancing.640x274.yuv -b:v 2048k -vcodec libx264 -pix_fmt yuv420p -r 24 -nal-hrd cbr test2048k.mp4"
  In this command, I used "-s 640x274" to indicate the resolution of dancing.640x274.yuv.
  By using "-pix_fmt yuv420p", I specified the storage format of each frame in dancing.640x274.yuv and the storage format of the encoded output.
  By using "-b:v 2048k", I set the output bitrate for transcoding.
  By using "vcodec libx264", I chose the x264 encoding scheme.
  By using "-r 24", I set the output frame rate to 24, which is consistent with the frame rate of "dancing_org.mp4".
  By using "-nal-hrd cbr", I set the use of CBR constant bitrate encoding.
  With this command, we can transcode dancing.640x274.yuv to output a video file named "test_640x274_2048k.mp4" for the subsequent PSNR calculation.

- PSNR Calculation: The calculation method for the PSNR of the two videos is as follows:
  - The transcoded video is denoted as v.mp4, and the original video is denoted as "dancing_org.mp4". The frames of the two videos are compared frame by frame, and the corresponding frame's PSNR is calculated. The algorithm for the calculation is described as follows:
    * Read a frame from each of the two videos, then use the bicubic algorithm to scale the resolution of the transcoded video frame to the same resolution as the original video. After that, calculate the Mean Squared Error (MSE) between the two frames: Assuming the original video has a resolution of m x n, let's denote a frame from the original video as I, and a corresponding frame from the transcoded video as K, with a resolution of m x n as well. The formula for calculating the Mean Squared Error (MSE) is as follows:

$$\text{MSE} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [(I(i,j) - K(i,j))^2] \tag{1}$$

    * Then the formula for defining PSN(db) is:

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

      In the formula, if the video format input is uint8, the max value is 255.
  - Add the PSNR of the two input videos frame by frame, then take the average to get the PSNR of the two videos. If you have calculated the PSNR for c frames, the average PSNR is:

$$\text{PSNR}_{\text{avg}} = \frac{1}{c} \sum_{i=0}^{c-1} \text{PSNR}_i \tag{3}$$

2. Show on a single R/D plot in Figure **??**, the RD curves for each of your representations. Place appropriate legends and label your axes appropriately. Show on the plot your estimate for the crossover bitrates i.e. the maximum bitrate where the quality of representations at 138p and 274p are greater than 274p and 548p respectively.

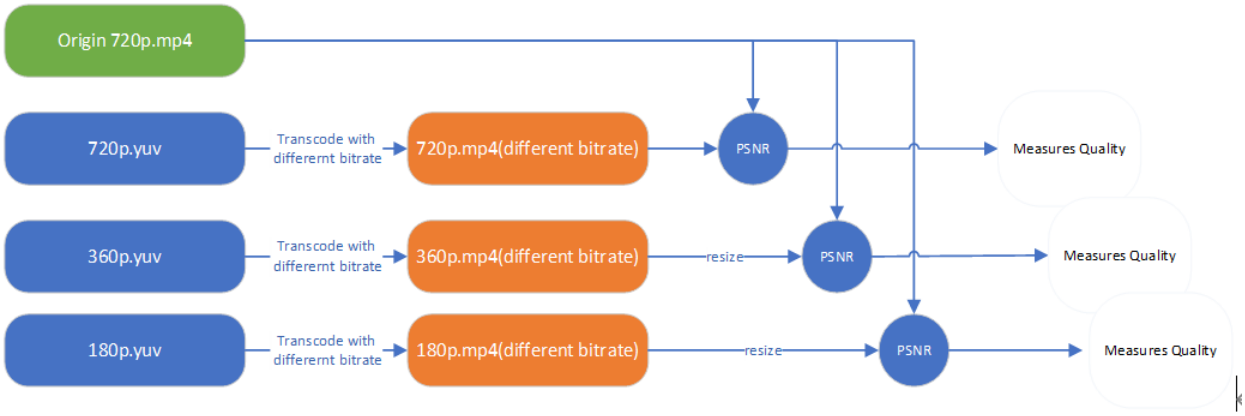

Figure 1: Algorithm Flow Diagram.

resolution	Bitrate (Kbps)						
720p	512	1024	2048	3072			
PSNR(db)	25.46	27.13	28.69	29.83			
360p	96	128	256	384	512	1024	2048
PSNR(db)	22.67	23.37	24.86	25.68	26.20	27.33	28.16
180p	64	96	128	256	512	1024	
PSNR(db)	22.19	22.72	23.02	23.54	23.82	23.96	

Figure 2: Dataset.

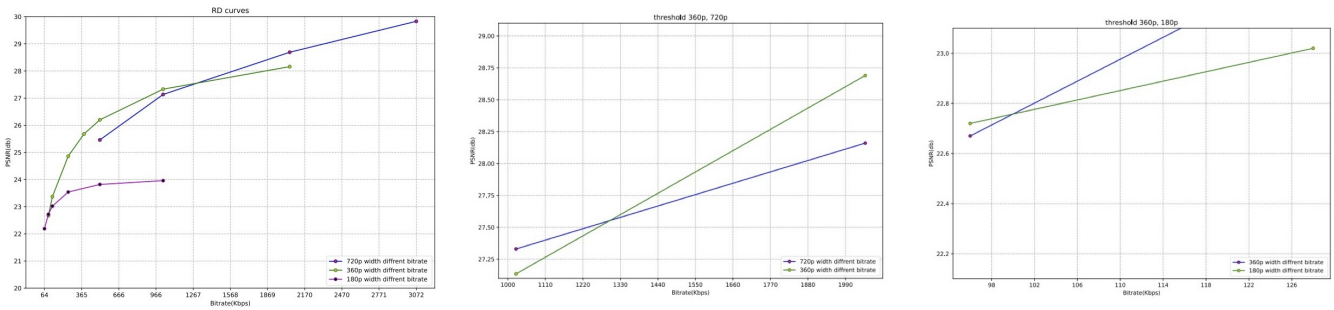

Figure 3: The RD Curves for three representations are shown as well as crossover bitrates.

2. I implement this part in Python according to the algorithm principles and processes given in the previous step. I used opencv, numpy, and matplotlib. The installation commands are: pip install opencv-python, pip install numpy, pip install matplotlib. Then, it can be calculated and implemented. First, I transcoded the corresponding 720P, 360P, and 180P videos, and then calculated the PSNR values of each .yuv video at different bit rates using the algorithm mentioned in the first question. And recorded the values (see Figure 2 dataset). Then enter the result into python and use matplotlib to draw the RD curve, and the result is as follows left image of Figure 3.

3. Translate the following sentences: In the video transcoding process, I set different bit rates, which directly affect the quality of the transcoded video. PSNR is the peak signal-to-noise ratio, used to measure video quality. The higher the PSNR, the better the video quality. The higher the bitrate of the video, the higher the PSNR value. This is because high bitrate videos can provide more information, thus making the video quality better. After downsampling the original video for transcoding, I reached the following conclusions, based on these three videos I analyzed one by one.

- Through the "dancing_org.1280x548.yuv" video, I found that the video without downsampling has a larger PSNR value affected by the bitrate. When using a lower bitrate for transcoding, the video quality is worse. When using a higher bitrate, the video quality will be better, and it can even approach the quality of the original video. Moreover, without downsampling, even at a low bitrate, the video quality can be acceptable.
- Through the "dancing.640x274.yuv" video, I found that videos without excessive downsampling are greatly affected by the bitrate during transcoding. Low bitrate videos have poor quality, and high bitrate videos have significantly improved quality. However, due to the loss of the image during downsampling, it is impossible to approach the quality of the original video.
- Through the "dancing.320x138.yuv" video, I found that the lower the resolution of the video, the lower the PSNR value compared to the original video. Also, increasing the transcoding bitrate later cannot improve the PSNR. The increase in PSNR value from 512Kbps to 1024Kbps is very small, indicating that if the resolution is too low after downsampling, it cannot improve image quality by increasing the bitrate and upsampling.

In summary, the video transcoding bitrate is an important factor affecting the quality of the transcoded video, and the resolution of downsampling has a huge impact on video quality.

4. From the RD curve chart, I can find that there is an intersection between the 720p video and 360p video curves, and the intersection point is located within the bitrate range of [1024, 2048]. From the RD curve chart, I can find that there is an intersection between the 180p video and 360p video curves, and the intersection point is located within the bitrate range of [96, 128].

3. Write here the quality and rate for your chosen representation at 138p. $\sim$ 90Kbps, 21.7dB

4. Write here the quality and rate for your chosen representation at 274p. $\sim$ 100Kbps, 22.76dB

5. Write here the quality and rate for your chosen representation at 720p. $\sim$ 1298Kbps, 27.5dB

6. The V_l representation at 274p has bitrate and PSNR = 599Kbps, 26.49 dB

7. Comment on any differences between the representation V_l at 274p and your chosen representation.

According to the slope formula:

$$k = \frac{(y_2 - y_1)}{(x_2 - x_1)} \tag{4}$$

Linear equation:

$$y - y_1 = k(x - x_1) \tag{5}$$

I can find the corresponding thresholds by setting up equations and calculating their intersection points.

- According to the calculations, the intersection points of the 720p and 360p lines are

$$y = 0.001519 * x + 25.58(720p) \tag{6}$$

$$y = 0.00081 * x + 26.5(360p) \tag{7}$$

  At [1298, 27.5], so the threshold between 720p and 360p is 1298 Kbps, as shown in the middle of Figure 3.

- According to the calculations, the intersection points of the 180p and 360p lines are

$$y = 0.021875 * x + 20.57(360p) \tag{8}$$

$$y = 0.0094 * x + 21.82(180p) \tag{9}$$

  At [100, 22.76], so the threshold between 180p and 360p is 100 Kbps, as shown on the right side of Figure 3.

However, the problem with this calculation method is: This is only an estimation scheme for the threshold, because the relationship between bitrate and PSNR is not strictly linear, so the estimated threshold is not accurate enough.

5. The bitrate of $V1$ is (1298-100)/2 = 599 Kbps. Use the following command for transcoding: "ffmpeg -s 640x274 -pix_fmt yuv420p -i dancing.640x274.yuv -b:v 599k -vcodec libx264 -pix_fmt yuv420p -r 24 -nal-hrd cbr test360_599k.mp4" The PSNR value calculated from the transcoded video in the code is: 26.49 dB.

6. The main purpose of video transcoding is to ensure quality while optimizing storage and transmission. I will make a balanced choice and select the most suitable standard, but this is not absolute and should be chosen according to the specific video situation. For this report, I will choose a bitrate of 512 Kbps for 360p videos.
Differences: After encoding at 512 Kbps, the video size is smaller, and the encoding speed is faster.
Advantages: Adopting this bitrate can save bandwidth as much as possible, and the video quality is not significantly different from that of V1.