Mohit Kumar
SR No. 04-01-03-10-51-21-1-19825
MTech Artificial Intelligence

Assignment 3

In this assignment, I explored the principal and independent components of natural images and the compression of video frames. In the first task, I extracted large number of patches from natural images from the Berkeley segmentation dataset and visualized the principal and independent components as images. In the second task I tried to understand the bit rate below which it is better to downsample the video spatially by 2 , compress the video and then upsample rather than directly compress at the desired bit rate.

- **Principal components of natural images** I used the train images in the Berkeley segmentation dataset as natural images. I extracted 8×8 patches from each of these images and performed Principal Component Analysis (PCA) and Independent Component Analysis(ICA) on the patches. I then visualized the principal and independent components as images. I use 1000 patches from every image ($\sim$ 200) in the dataset. Note that I considered random overlapping patches within an image.

  Figures 1 and 2 show the principal components and Figures 3 and 4 show the independent components.

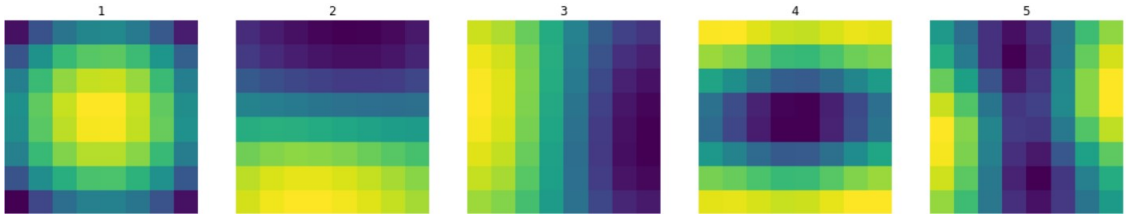

Figure 1: PCA components

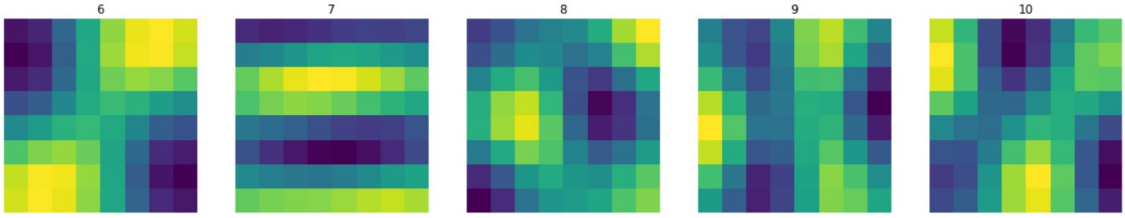

Figure 2: PCA components

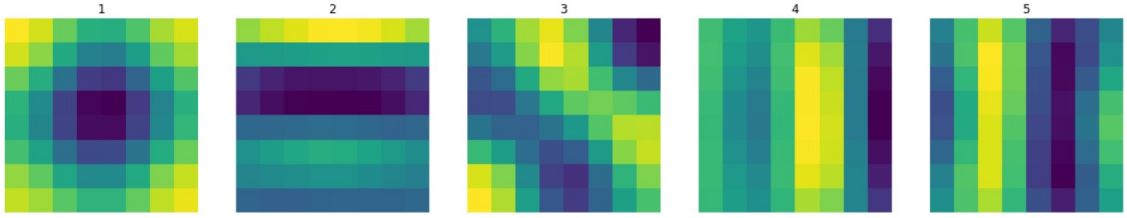

Figure 3: ICA components

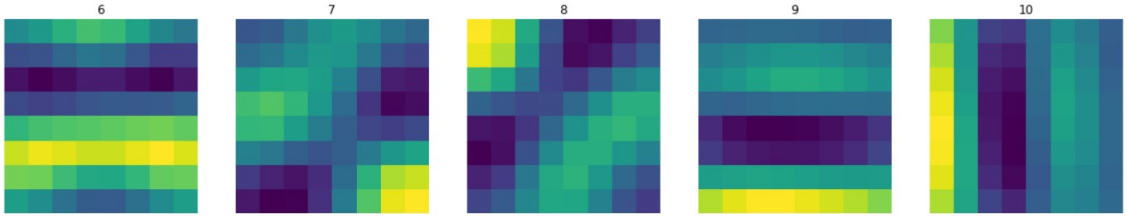

Figure 4: ICA components

For calculating the mutual information of a pair of nearby pixels, I chose the horizontal neighbors of a pixel in a patch. It was found to be **2.1**
The mutual information of the top two principal components in terms of their variance was found to be **0.024** for 1000 patches.
The mutual information of the top two independent components in terms of their variance was found to be **0.020** for 1000 patches.

- **Image compression and Resolution** I analysed the effects of bit rate used in video compression on the quality of videos. I tried to to understand the bit rate below which it was better to downsample the video spatially by 2 , compress the video and then upsample rather than directly compress the video at the desired bit rate. I used the bsl_25fps.yuv provided with this assignment and used H.264 as the compression scheme. I utilized the popular compression software ffmpeg to carry out the experiments. I compressed all the videos first with one I frame and rest being P frames and then with all I frames. Quality was measured in terms of PSNR in the luminance data, and VMAF for this experiment.

  The table 1 shows the settings I used for ffmepg for H.264 compression. I chose the values for keyint and min-keyint such that the compression is done first using one I frame and the rest being P frames and then using all I frames in the second case.

Table 1: Compression Settings (one I frame rest P frames)

Bit rate	$100 - 1000$ Kbps (step 25)
Video Codec	H.264
Audio Codec	Copy (default)
Key Int	300
Min Key Int	300
No Scenecut	1
B Frames	0
Log level	Panic
All other settings	ffmpeg default

Table 2: Compression Settings (all I frames)

Bit rate	$100 - 1000$ Kbps (step 25)
Video Codec	H.264
Audio Codec	Copy (default)
Key Int	1
Min Key Int	1
No Scenecut	1
B Frames	0
Log level	Panic
All other settings	ffmpeg default

Figures 5 and 6 show the plots between PSNR/VMAF and the bit rate for both the methods in case 1(One I frame rest P frames). Figures 7 and 8 correspond to case 2(all I frames). The figures include plots for both strategies (with and without dowmsampling and upsampling).

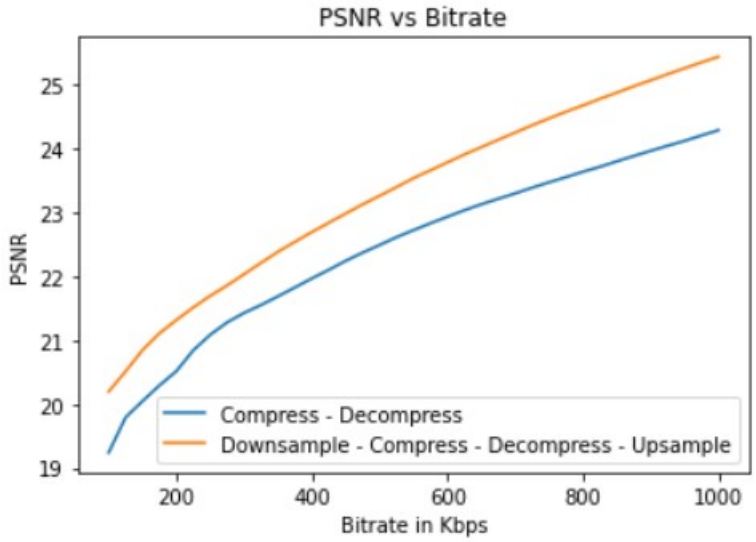

Figure 5: Case 1(one I frame rest P frames): PSNR

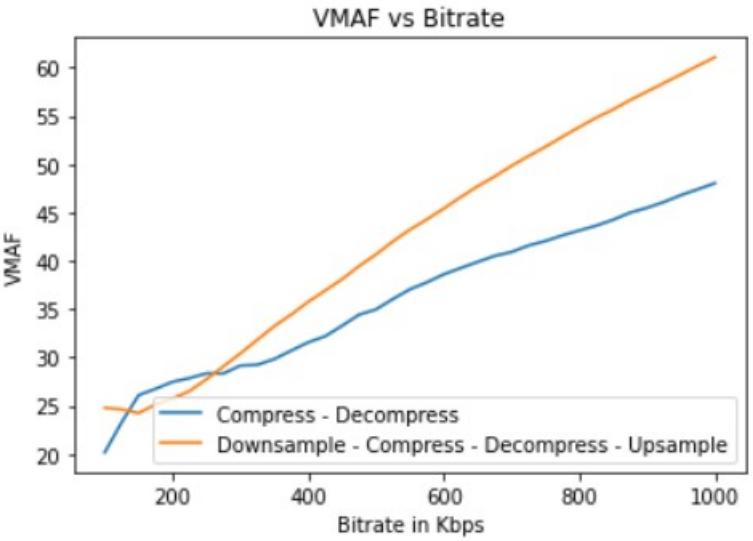

Figure 6: Case 1(one I frame rest P frames): VMAF

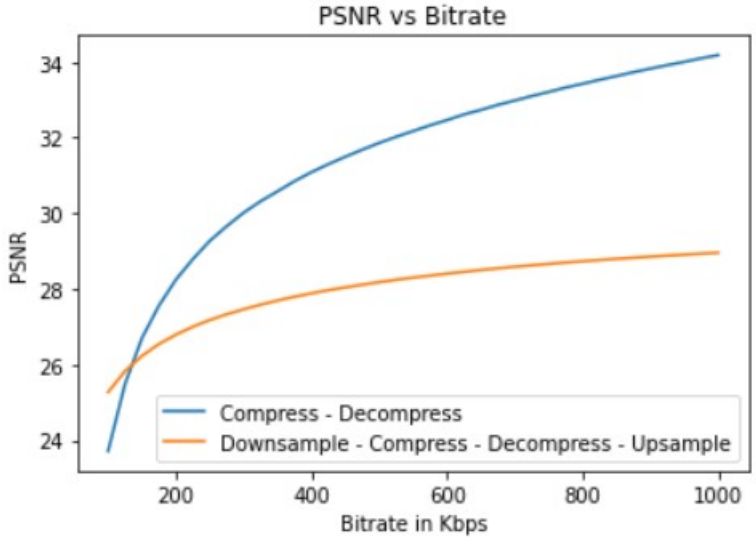

Figure 7: Case 2(all I frames): PSNR

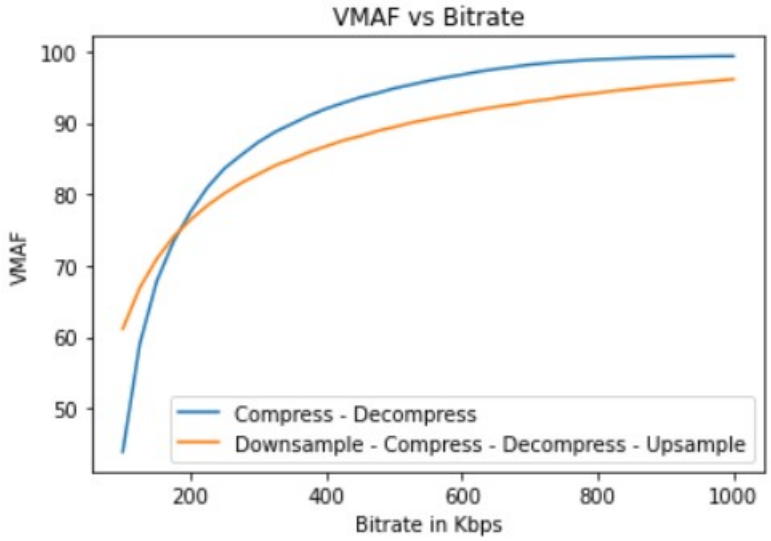

Figure 8: Case 2(all I frames): VMAF

Clearly, we can see that there is a switching between the two methods. In the first case(One I frame rest P frames) The switching bit rate is around **160 Kbps** above which compression-decompression in original dimensions is better and below which downsampling - compression decompression - upsampling is better.