Lab Report 10 Shuchen Zhang (szhan114)

Introduction:

In this lab, we examine the lossless compression coding method called run-length coding. Both run length coding of binary image, run length coding of image bit planes and predictive coding are explored. For each image, the entropy, total number of runs, total number of bits requires and the number of bits per pixel and the histogram of run-length are examined.

Methods:

Part 1:

For both lena and baboon image, the entropy H of the image is calculated using:

$$H = -\sum_{i=0}^{255} P(i)\log_2(P(i))$$

the total number of bits required can be calculated as the product of entropy and the total number of pixels in the image.

Part 2 :

In this part, we examine the run length coding which is a lossless compression technique. The image is scanned using “raster-snake” order, the consecutive scanned pixel with the same pixel intensity are consider in the same run. The run length for each run are stored and the probability of occurrence of each possible run length are displayed as histogram. The total number of bits needed can be calculated as (entropy * total number of runs). The average number of bits per pixel is (total number of bits)/(number of pixels in the image).

Part3:

8 bitplane images of lena image are generated and each image are measured and examined using the methods in part 2.

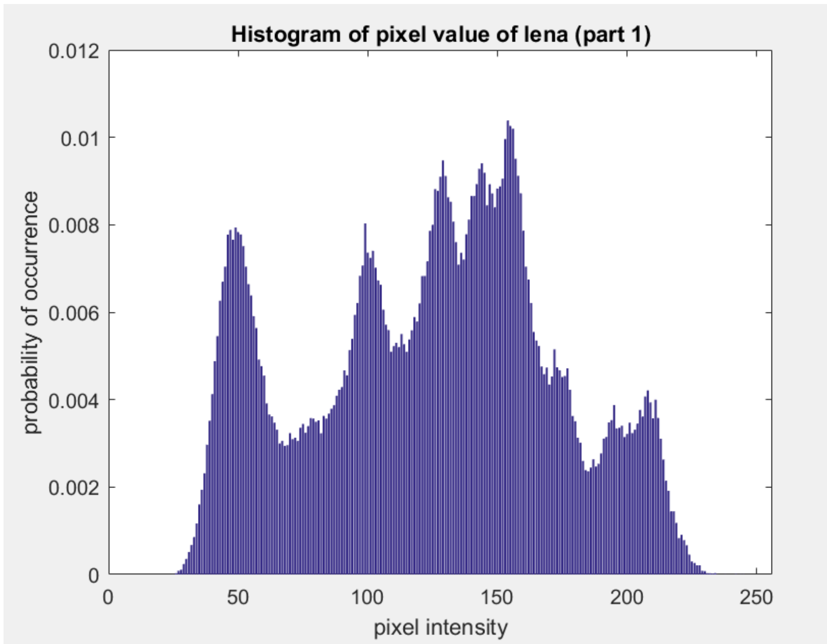
Part4:

In this part, we examined predictive coding. In predictive coding, we assume that a pixel’s intensity value will be very likely as the previous scanned neighbor’s intensity value. We predict the current pixel’s value from its causal neighbor and generate a predicted image. The error image would be the difference between the predicted image and the original image. If given the first pixel value and the error(difference) image pixel values, the original pixels can be reconstructed. In this way, we could compress the difference image losslessly in order to compress the original image lossly. The histogram of difference pixel value are generated. The bits required and the entropy are also calculated.

Part 1:

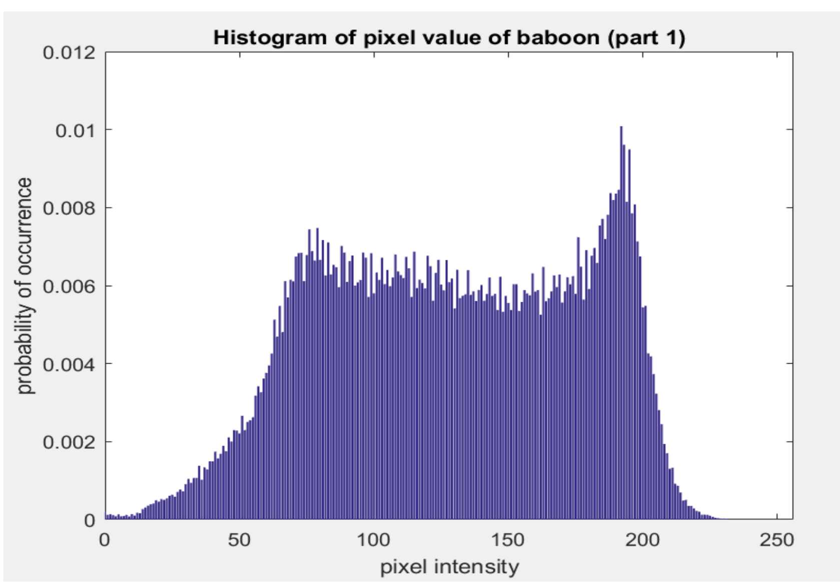
Lena:

a. Entropy values: 7.44551
b. Total number of bits: 1.95179e+06



Baboon:

a. Entropy values: 7.474432
b. Total number of bits: 195938e+06



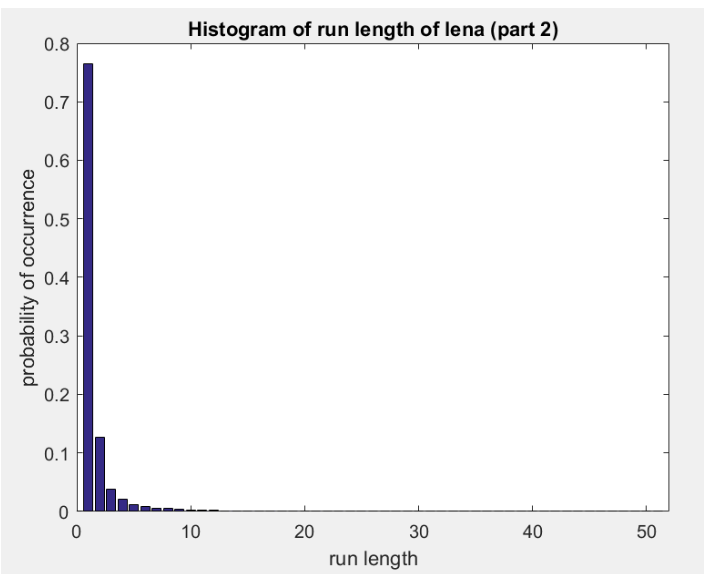
Q2: The entropies for both image are quite similar but the shape of resulting histogram is quite different. So I cannot deduce a noticeable correlation between entropy and shape of the histogram. I think this conclusion is generally true.

<u>Part 2:</u>

Q3: weaskedforit image compress better because it requires less bits than lenaerrdiff image after runlength coding. So we could see in general, the image with many sections of consecutive pixels with similar pixel intensity values such as character image like weaskedforit image entropy of lenaerrdiff image is lower than weaskedforit image. Weaskedforit is easier to compress since it has more repeated data than lenaerrdiff with large chunk of black and white constant pixel intensity regions. The histogram of lena image has fewer possible run length and they are mostly smaller than the weaskedforit.

Lenaerrdiff.raw:

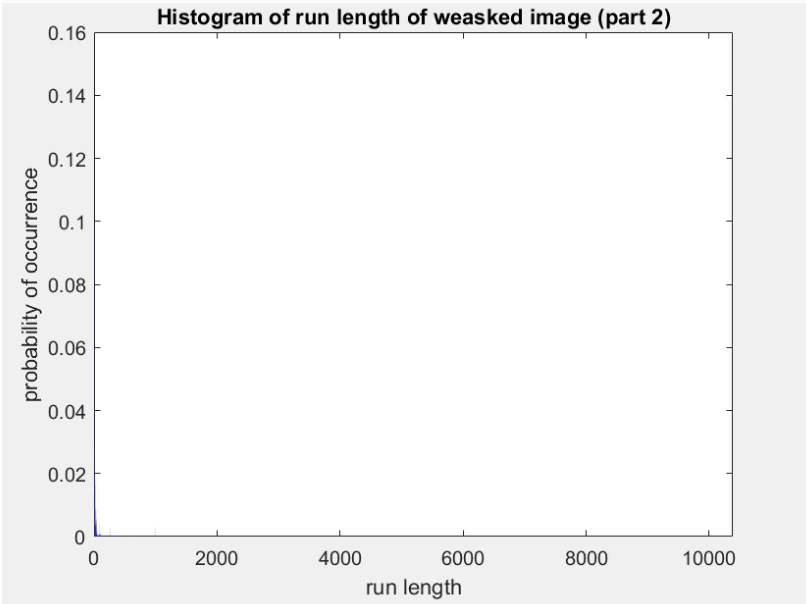
a. Run length entroy = 1.35382
b. Total number of runs = 159928
c. Total number of bits requires = 216514
d. Number of bits per pixel = 0.825936



Weaskedforit.raw:

a. Run length entropy = 5.17314
b. Total number of runs = 10224
c. Total number of bits requires = 52890.17

d. Number of bits per pixel = 0.201760

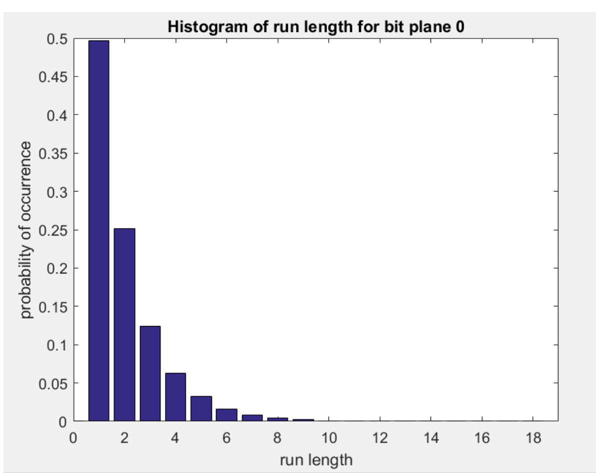


<u>Part 3:</u>

Q4. The most significant bit plane are most efficient to compress. The least significant bit plane is least efficient to compress. Because the least significant bit planes have much less long runs than the most significant bit plane.

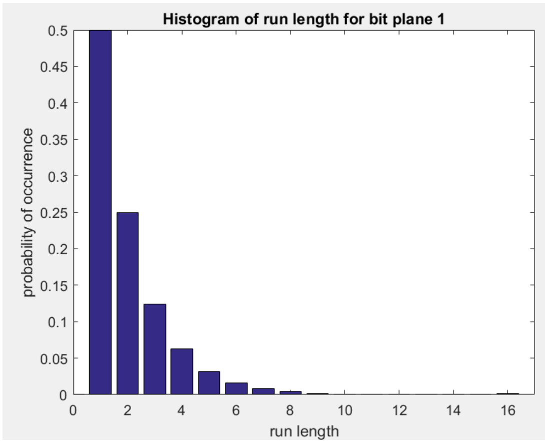
Lena_bit0:

a. Run length entropy = 2.01302
b. Total number of runs = 130213
c. Total number of bits requires = 262121
d. Number of bits per pixel = 0.999914



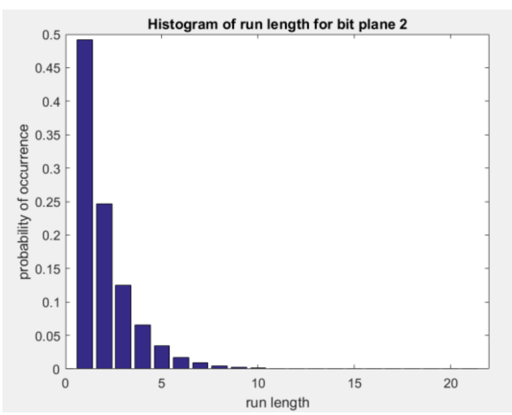
Lena_bit1:

a. Run length entropy = 2.00556
b. Total number of runs = 130700
c. Total number of bits requires = 262126
d. Number of bits per pixel = 0.999932



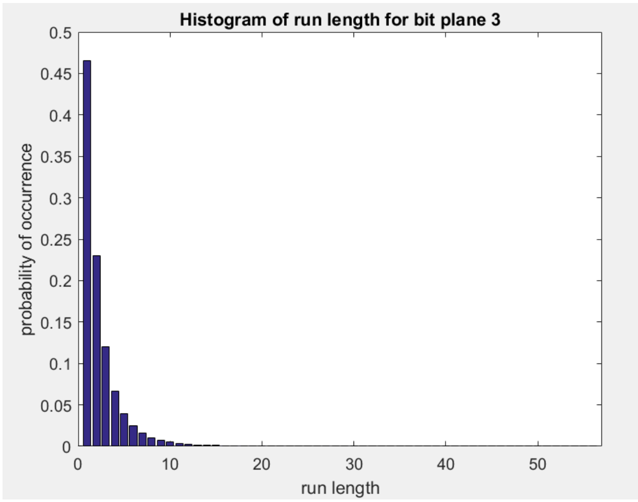
Lena_bit2:

a. Run length entropy = 2.05427
b. Total number of runs = 127512
c. Total number of bits requires = 261944
d. Number of bits per pixel = 0.999237

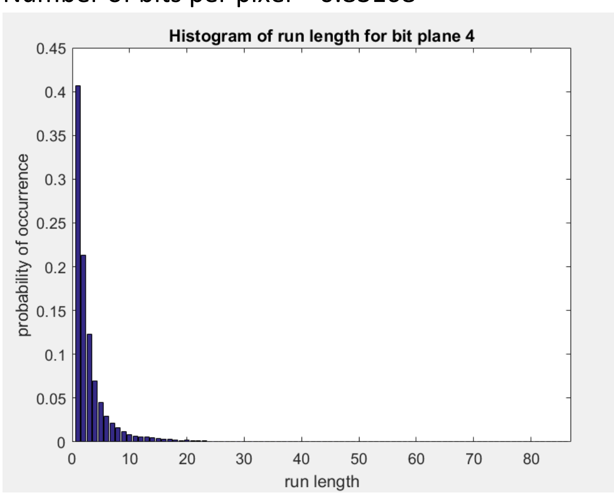


Lena_bit3:

a. Run length entropy = 2.34021
b. Total number of runs = 107856
c. Total number of bits requires = 252405
d. Number of bits per pixel = 0.96285

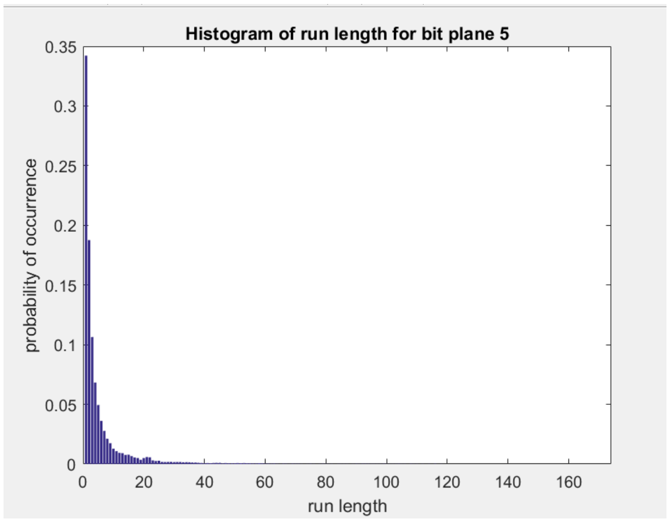
e. 

Lena_bit4:

a. Run length entropy = 2.8132
b. Total number of runs = 77499
c. Total number of bits requires = 218020
d. Number of bits per pixel = 0.83168
e. 

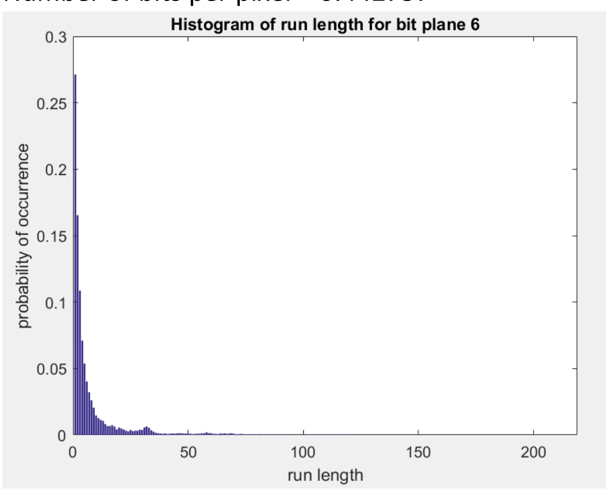
Lena_bit5:

a. Run length entropy = 3.49383
b. Total number of runs = 45667
c. Total number of bits requires = 159553
d. Number of bits per pixel = 0.608645

e. 

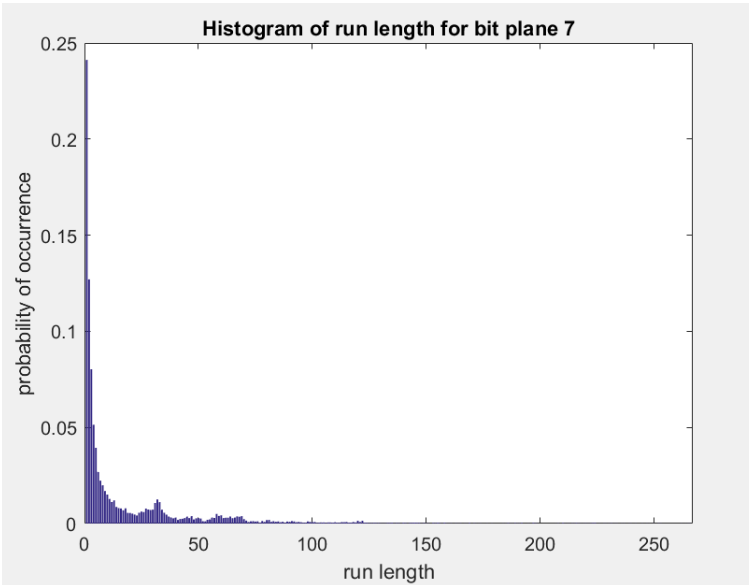
Lena_bit6:

a. Run length entropy = 4.08871
b. Total number of runs = 28387
c. Total number of bits requires = 116066
d. Number of bits per pixel = 0.442757



Lena_bit7:

a. Run length entropy = 5.00165
b. Total number of runs = 13926
c. Total number of bits requires = 69653
d. Number of bits per pixel = 0.265705

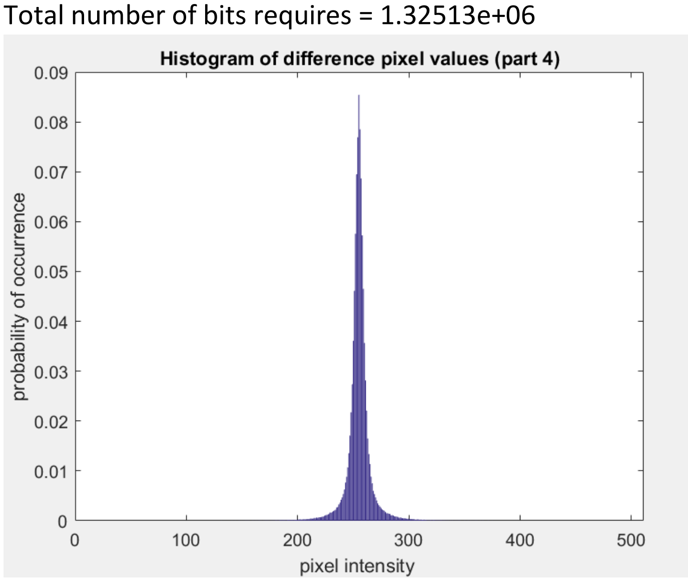


Total number of bits per pixel for all bit plane image of lena combined = 6.11072 bits

Part 4:

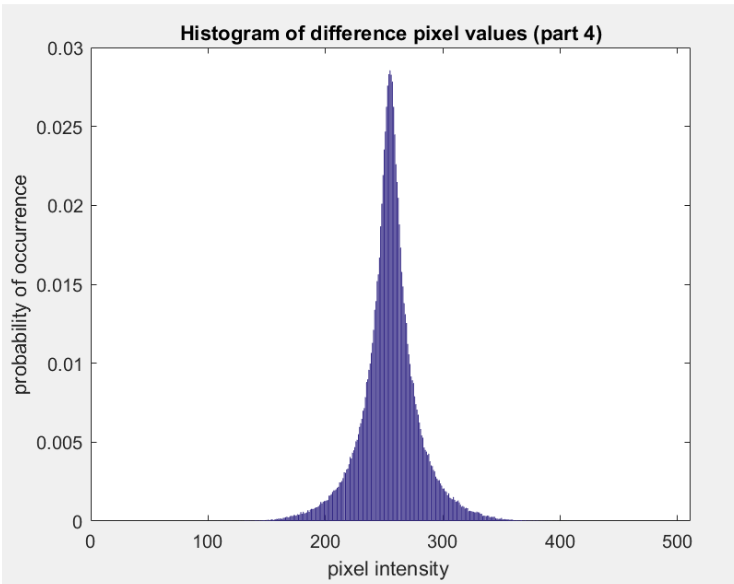
Q5. Compare the entropy result form part 4 and part 1: Both baboon and lena image has lower entropy in part 4 than part 1. Predictive coding performs better because each pixel's value can be vary similar with its causal neighboring pixel so that the resulting difference pixel (prediction error) value are low, thus the entropy value will be low.

Lena:

a. Entropy = 5.055
b. Total number of bits requires = 1.32513e+06
c. 

Baboon:

a. Entropy = 6.54798
b. Total number of bits requires = 1.71651e+06



Q6. Run length coding of a bi-level image can result in an expansion of the image in terms of bits per pixel if the image has mostly very short runs which means that the image has very few consecutive scanned pixels with same intensity value.

Q7. There are no 8-bit images for which run-length encoding each bit planes separately takes as many or more bits than Huffman coding done on 256-level image. The total number of bits requires is larger for lena image in part 1 which use Huffman coding than for 8 bit planes combined in part 3.

Conclusion:

In this lab, the lossless coding method run length coding method is examined. The run length coding is most efficient in compressing the image that has more long runs. The lossless predictive coding is more efficient because its assumption that each pixel's value is similar to its causal pixel so that the entropy is lowered.