

LAB REPORT

IMAGE PROCESSING (COMPRESSION)

December 1, 2018

Instructor

Christophe STOLZ

ASRAF ALI Abdul Salam Rasmi

NAZIFI Nahid

Masters in Computer Vision

Centre Universitaire Condorcet

Universite de Bourgogne

Chapter 1

Basic Compression Method

1.1 INTRODUCTION

Image compression is the process of encoding or converting an image file in such a way that it consumes less space than the original file. It minimizes the size of a graphics file in bytes without degrading the quality of the image to an unacceptable level. This process allows less use of storage memory space. It also reduces the time required for images to be transferred over the Internet or downloaded from Web pages.

1.2 ARITHMETIC CODING

Arithmetic coding is a common algorithm used in both loss-less and lossy data compression techniques. It is an entropy encoding technique in which the frequently seen symbols are encoded with fewer bits than rarely seen symbols. It converts the entire input data into a single floating point number n where $(0.0 \leq n \leq 1.0)$. The interval is divided into sub-intervals in the ratio of the probability of occurrence frequencies. For a start point and end point of an entire range the lower-limit of a character range is the upper limit of the previous character given by start point + cumulative frequency X (end - start). Therefore, each interval corresponds to one symbol. The first symbol restricts the tag position to be in one of the intervals. The reduced interval is partitioned recursively as more symbols are processed. Once the tag falls into an interval, it never gets out of it [1].

1.2.1 Test 1

In this test we tried to Encode and Decode the DNA sequence with symbols (alphabets) {A, C, T, G} with probabilities {0.5, 0.3, 0.15, 0.05}, and the sequence is ACTAGC.

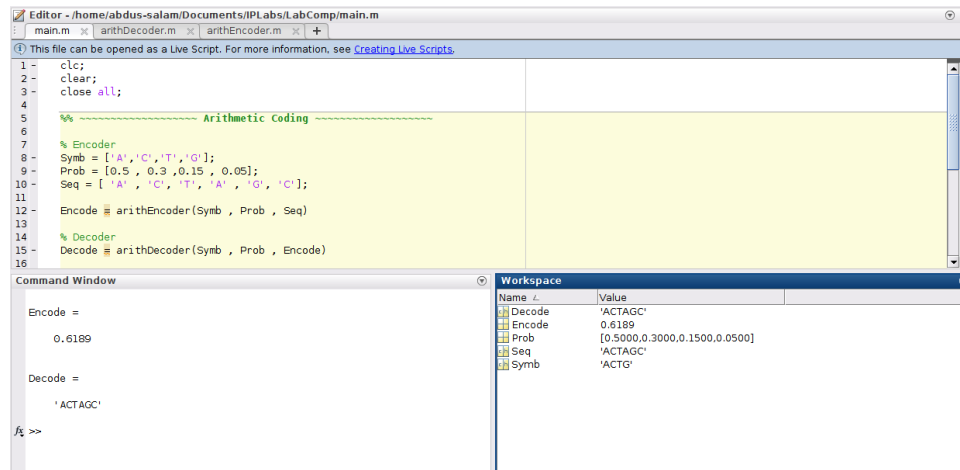


Figure 1.1: Output of DNA Sequence

1.2.2 Test 2

In this test we tried to Encode the sequence BE_A_BEE.

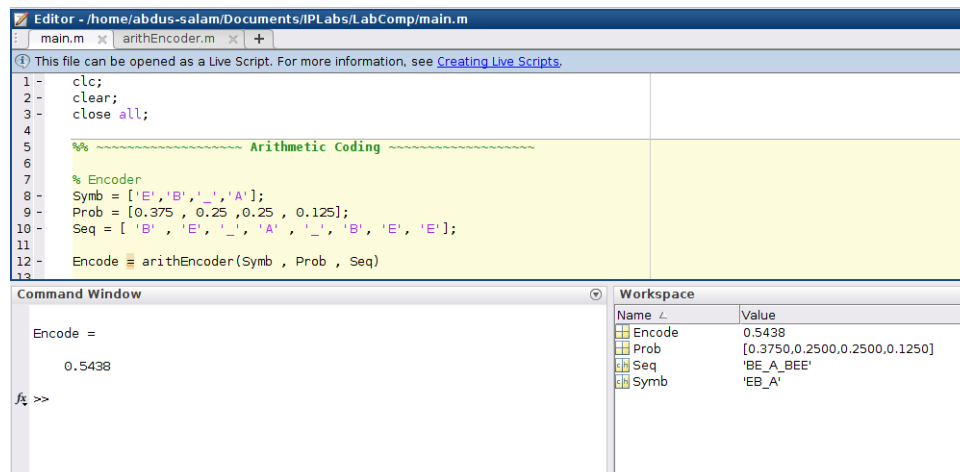


Figure 1.2: Output of BE_A_BEE Sequence

1.3 HUFFMAN CODING

Huffman coding is a loss-less data compression algorithm. The idea is to assign variable-length binary codes to input characters. The lengths of the assigned codes are based on the frequencies of corresponding characters. The most frequent character gets the smallest code and the least frequent character gets the largest code. The variable-length codes assigned to input characters are Prefix Codes, means the codes (bit sequences) are assigned in such a way that the code assigned to one character is not prefix of code assigned to any other character. This is how Huffman Coding makes sure that there is no ambiguity when decoding the generated bit stream [2].

In this exercise we tried to code an Image using Huffman coding considering the probabilities of pixel intensities. To perform this we created a function call *probVector.m* which takes gray-scale image as input and returns an array containing probabilities of pixel intensities.

1.4 SIMPLIFIED JPEG ENCODER

The term "JPEG" is an acronym for the Joint Photographic Experts Group, which created the standard. JPEG is a commonly used method of lossy compression for digital images. The degree of compression can be adjusted, allowing a selectable trade-off between storage size and image quality. JPEG typically achieves 10 : 1 compression with little perceptible loss in image quality. JPEG uses a lossy form of compression based on the discrete cosine transform (DCT). This mathematical operation converts each block of the image from the spatial (2D) domain into the frequency domain (so called transform domain) [3].

In this exercise we performed JPEG encoding for a grayscale image. The output of Encoding are shown below.



Figure 1.3: Input Image



Figure 1.4: Level Shifted Input Image



Figure 1.5: Compressed Image

Bibliography

- [1] M. Nelson and J. Gailly, *The Data Compression Book*. Wiley, 1995.
- [2] “Huffman coding.” <https://www.geeksforgeeks.org/huffman-coding-greedy-algo-3/>.
- [3] W. Pennebaker and J. Mitchell, *JPEG: Still Image Data Compression Standard*. Chapman & Hall digital multimedia standards series, Springer US, 1992.